INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS

2009 Edition

EMERGING RESEARCH DEVICES

THE ITRS IS DEVISED AND INTENDED FOR TECHNOLOGY ASSESSMENT ONLY AND IS WITHOUT REGARD TO ANY COMMERCIAL CONSIDERATIONS PERTAINING TO INDIVIDUAL PRODUCTS OR EQUIPMENT.

TABLE OF CONTENTS

Scope 1	
Difficult Challenges 2	
Introduction	2
Device Technologies	2
Materials Technologies	3
Nano-information Processing Taxonomy	3
Emerging Research Devices	4
Memory Taxonomy and Devices	4
Memory Taxonomy	5
Memory Devices	5
Logic and Alternative Information Processing Devices	8
Logic Devices	10
Non-FET, Non Charge-based "Beyond CMOS" Devices	
Emerging Research Architectures	20
Introduction	20
Logic Architecture Benchmarking	20
Architectural requirements for a competitive logic device	21
Architecture Assessment Methodology	
Observations	
Quantitative Results	
Accelerators in the Future Logic Architectures	
	∠⊃ 22
Challenges in Memory Systems	23
Opportunities for Architectural Exploitation of Emerging Memory Devices	
Research Needs	
Inference Computing	
Introduction	
Architecture for Inference	
Architecture	
Algorithm Characteristics for Bayesian Inference	26
Performance Limits of Information Processing Architectures	26
Emerging Memory and Logic Devices—A Critical Assessment	27
Introduction	27
Technologies beyond CMOS	27
Overall Technology Requirements and Relevance Criteria	27
Electronic Charge-based Nanoscale Devices	
Alternate Computational-State-Variable Nanoscale Devices	29
Potential Performance Assessment for Memory and Logic Devices	29
Methodology	
Results	

Fundamental Guiding Principles— "Beyond CMOS" Information Processing	
Introduction	41
Grand Challenges	41
Computational State Variable(s) other than Solely Electron Charge	41
Non-thermal Equilibrium Systems	41
Novel Energy Transfer Interactions	41
Nanoscale Thermal Management	41
Sub-lithographic Manufacturing Process	41
Alternative Architectures	41
Endnotes/References	

LIST OF FIGURES

Figure ERD1	A Taxonomy for Emerging Research Information Processing Devices (The technology entries are representative but not comprehensive.)	4
Figure ERD2	Exemplary Design Space for a NAND2 Circuit with Fan-out of 1 in Various New Emerging Research Switches. Red points are charge based; other po are examples of switches using other state variables	ints .22
Figure ERD3a-d	Technology Performance Evaluation for a) STT MRAM, b) Nanothermal Memory, c) Nanoionic Memory, and d) Ferroelectric Memory	.34
Figure ERD3e-h	Technology Performance Evaluation for e) Electronic Effects Memory, f) Nanomechanical Memory, g) Macromolecular Memory, and h) Molecular Memory	.35
Figure ERD4a-d	Technology Performance Evaluation for a) Unconventional Geometry MOSFETs, b) CNT MOSFETs, c) Nanowire, and d) Ge MOSFETs	.36
Figure ERD4e-h	Technology Performance Evaluation for e) Tunnel MOSFETs, f) III-V Compound Semiconductor MOSFETs, g) GNR MOSFETs, and h) I MOSFETs	.37
Figure ERD4i-I	Technology Performance Evaluation for i) Negative Cg FETs, j) Atomic Switches, k) Spin Transistors, and I) NEMS Devices.	.38
Figure ERD4m-p	Technology Performance Evaluation for m) Pseudospintronic Devices, n) SETs, o) Molecular Switches, and p) Spin Wave Devices	.39
Figure ERD4q-r	Technology Performance Evaluation for q) Nanomagnetic Devices, and r) Moving Domain Wall Devices	.40

LIST OF TABLES

Table ERD1	Emerging Research Devices Difficult Challenges	.2
Table ERD2	Memory Taxonomy	.4
Table ERD3	Current Baseline and Prototypical Memory Technologies	.5
Table ERD4	Transition Table for Emerging Research Memory Devices	.5
Table ERD5	Emerging Research Resistance-based Memory Devices— Demonstrated and Projected Parameters	.5
Table ERD6	Transition Table for Emerging Research Logic Devices	.9
Table ERD7a	MOSFETS: Extending the channel to the End of the Roadmap	.9
Table ERD7b	Charge based Beyond CMOS: Non-Conventional FETs and other Charge- based information carrier devices	.9
Table ERD7c	Alternative Information Processing Devices	.9
Table ERD8	Research and Technology Development Schedule proposed for Carbon- based Nanoelectronics to impact the Industry's Timetable for Scaling Information Processing Technologies	10
Table ERD9	Prototypical Criteria for Emerging Device Architectures	21
Table ERD10	Summary of the Attributes of Several Emerging Research Devices and Projections for their Application Spaces	22
Table ERD11	Challenges in Existing Memory Systems Illustrated using Today's Memory Devices	23
Table ERD12	Sample of Potential Unique Opportunities for Architectural Exploitation of Emerging Memory Devices	24
Table ERD13	Summary of Hardware Requirements for the Three Basic Approaches to Probabilistic Inference	26
Table ERD14	Potential Evaluation for Emerging Research Memory Devices	31
Table ERD15	Potential Evaluation for Emerging Research Logic and Alternate Information Processing Devices	ı 31

EMERGING RESEARCH DEVICES

SCOPE

Continued dimensional and functional scaling of CMOS is driving information processing¹ technology into a broadening spectrum of new applications. Many of these new applications are enabled by performance gains realized by CMOS scaling. As one goal, the ERD chapter assesses technologies being explored to replace the silicon channel and source/drain regions with new, high mobility and high carrier velocity materials to sustain CMOS performance gains to and beyond the 16 nm generation. These channel and source/drain replacement materials include Ge, III-V compound semiconductors, graphene nanoribbons, carbon nanotubes, or nanowires. (Material challenges related to emerging research devices are addressed in the complementary chapter entitled *Emerging Research Materials*). Eventually, because dimensional scaling of CMOS will approach fundamental limits, several new information processing devices are being explored to sustain the historical integrated circuit scaling cadence and reduction of cost/function into future decades. Therefore, another goal of this chapter is to survey, assess, and catalog viable new information processing devices for their long-range potential, technological maturity, and to identify the scientific/technological challenges gating their being accepted by the semiconductor industry as having acceptable risk for further development. An ancillary goal of this chapter is to stimulate invention and development of viable device concepts that extend dimensional 1, performance, and/or functional scaling of information processing substantially beyond "ultimately scaled" CMOS (i.e., scaled to the end of the roadmap).

These goals are accomplished by addressing two technology-defining domains: 1) extending the functionality of the CMOS platform via heterogeneous integration of these new technologies onto this platform, and 2) stimulating invention of a new information processing paradigm. A related goal is to provide an objective, informative resource for the constituent nanoelectronics communities pursuing: 1) research, 2) tool development, 3) funding support, and 4) investment each directed to developing a new information processing technology. These communities include university, research institute, and industrial research laboratories; tool vendors; research funding agencies; and the semiconductor industry. The potential and maturity of each emerging research device (ERD) technology are reviewed and assessed in this chapter to identify the most important scientific and technological challenges that must be overcome for a candidate device to become a viable technology.

The scope of the ERD Chapter is restricted to information manipulation, transmission, and storage. Within this scope, the chapter is intended to accomplish two objectives. First is to gather in one place substantive, alternative concepts for memory, information processing devices, and information processing nanoarchitectures that, if successful, will substantially extend the Roadmap beyond CMOS. This discussion provides a window into these candidate approaches.

The second objective is to provide a balanced, critical assessment of these emerging new device technologies. A brief section also is included to propose a set of fundamental principles that will likely govern successful extension of information processing technology substantially beyond that attainable solely with ultimately scaled CMOS. This chapter, therefore, provides an ITRS perspective on emerging new device technologies and serves as a bridge between CMOS and the realm of microelectronics beyond the end of CMOS dimensional and equivalent performance scaling.

The ERD chapter introduces a new feature in this year's edition—highlighting an emerging research information processing technology that is exhibiting high potential and is rapidly maturing. Among several emerging information processing technologies assessed and considered, carbon-based nanoelectronics is highlighted this year with an enhanced review. Carbon-based nanoelectronics, including carbon nanotubes, graphene, and graphene nanoribbons, was chosen for emphasis because of its relative performance and maturity when applied to MOSFETs and its high potential for realizing a new information processing paradigm.

The ERD chapter is divided into four categories: 1) memory devices, 2) information processing or logic devices, 3) information processing nanoarchitectures, and 4) a critical assessment of each technology entry. Some detail is provided for each entry regarding operation principles, advantages, technical challenges, maturity, level of research activity, and current and projected performance. Also included is a device and architectural focus combining emerging research devices offering specialized, unique functions as heterogeneous core processors integrated with a CMOS platform technology. This represents the nearer term focus of the chapter, with the longer term focus remaining on discovery of an alternate information processing technology to eventually replace digital CMOS.

As in previous editions, the chapter includes "transition tables." The purpose of these transition tables is twofold. The first is to track technologies that have appeared in or have been removed from the 2007 tables and so provide a very

short explanation of the reason for this change. The second purpose is to identify and provide a placeholder for technologies that are considered important but do not meet the criteria for full inclusion into the more detailed tables. These technologies may be expected to become more or less visible in future editions of the roadmap and hence the name.

DIFFICULT CHALLENGES

INTRODUCTION

The semiconductor industry is facing two classes of difficult challenges related to extending integrated circuit technology to and beyond the end of CMOS dimensional scaling. One set relates to extending CMOS beyond its ultimately scaled density and functionality by integrating, for example, a new high speed, dense, and low power memory technology on the CMOS platform. Another class of challenges is to extend information processing substantially beyond that attainable by CMOS alone using an innovative combination of new devices and architectural approaches for extending CMOS and, eventually, inventing a new information processing platform technology. The difficult challenges are presented in Table ERD1.

Table ERD1Emerging Research Devices Difficult Challenges

Difficult Challenges ≥ 16 nm and < 16 nm	Summary of Issues and opportunities
Scale high-speed, dense, embeddable, volatile, and non-volatile memory technologies to and beyond the 16 nm technology generation.	SRAM and FLASH scaling will reach definite limits within the next several years (see PIDS Difficult Challenges). These are driving the need for new memory technologies to replace SRAM and FLASH memories.
	Identify the most promising technical approach(es) to obtain electrically accessible, high- speed, high-density, low-power, (preferably) embeddable volatile and non-volatile RAM
	The desired material/device properties must be maintained through and after high temperature and corrosive chemical processing Reliability issues should be identified & addressed early in the technology development
Scale CMOS to and beyond the 16 nm technology generation.	Develop new materials to replace silicon as an alternate channel and source/drain to increase the saturation velocity and maximum drain current in MOSFETs while minimizing leakage currents and power dissipation for technology scaled to 16 nm and beyond.
	Develop means to control the variability of critical dimensions and statistical distributions (e.g., gate length, channel thickness, S/D doping concentrations, etc.)
	Accommodate the heterogeneous integration of dissimilar materials The desired material/device properties must be maintained through and after high temperature and corrosive chemical processing Reliability issues should be identified & addressed early in t
Extend ultimately scaled CMOS as a platform technology into new domains of application.	Discover and reduce to practice new device technologies and a primitive-level architecture to provide special purpose optimized functional cores heterogeneously integrable with silicon CMOS.
	Invent and develop a new information processing technology eventually to replace CMOS
Continue functional scaling of information processing technology substantially beyond that attainable by ultimately scaled CMOS.	Ensure that a new information processing technology is compatible with the new memory technology discussed above; i.e., the logic technology must also provide the access function in a new memory technology.
	Bridge a knowledge gap that exists between materials behaviors and device functions.
	Accommodate the heterogeneous integration of dissimilar materials
	The desired material/device properties must be maintained through and after high temperature and corrosive chemical processing
	Reliability issues should be identified & addressed early in the technology development

DEVICE TECHNOLOGIES

Difficult challenges related to emerging research devices are divided into those related to memory technologies and those related to information processing or logic devices. (Refer to Table ERD1.) One challenge is the need of a new memory technology that combines the best features of current memories in a fabrication technology compatible with CMOS process flow scaled beyond the present limits of SRAM and FLASH. This would provide a memory device fabrication technology required for both stand-alone and embedded memory applications. The ability of an MPU to execute programs is limited by interaction between the processor and the memory, and scaling does not automatically solve this problem. The current evolutionary solution is to increase MPU cache memory, thereby increasing the floor space that SRAM occupies on an MPU chip. This trend eventually leads to a decrease of the net information throughput. In addition to auxiliary circuitry to maintain stored data, volatility of semiconductor memory requires

external storage media with slow access (e.g., magnetic hard drives, optical CD, etc.). Therefore, development of *electrically accessible non-volatile* memory with *high speed* and *high density* would initiate a revolution in computer architecture. This development would provide a significant increase in information throughput beyond the traditional benefits of scaling when fully realized for nanoscale CMOS devices.

A related challenge is to sustain scaling of CMOS logic technology to and beyond 16 nm. One approach to sustaining performance gains as CMOS scaling matures in the next decade is to replace the strained silicon MOSFET channel (and the source/drain) with an alternate material offering a higher potential quasi-ballistic-carrier velocity and higher mobility than strained silicon. Candidate materials include strained Ge, SiGe, a variety of III-V compound semiconductors, and graphene. Introduction of non-silicon materials into the channel and source/drain regions of an otherwise silicon MOSFET (i.e., onto a silicon substrate) is fraught with several very difficult challenges. These challenges include heterogeneous fabrication of high-quality (i.e., defect free) channel and source/drain materials on non-lattice matched silicon, minimization of band-to-band tunneling in narrow bandgap channel materials, elimination of Fermi level pinning on III-V and Ge surfaces, and fabrication of high- κ gate dielectrics on the passivated channel materials. Additional challenges are to sustain the required reduction in leakage currents and power dissipation in these ultimately scaled CMOS gates and to introduce these new materials into the MOSFET while simultaneously minimizing the increasing variations in critical dimensions and statistical fluctuations in the channel (source/drain) doping concentrations.

A longer-term challenge is invention and reduction to practice of a new manufacturable information processing technology addressing "beyond CMOS" applications. For example, emerging research devices might be used to realize special purpose processor cores that could be integrated with multiple CMOS CPU cores to obtain performance advantages. These new special purpose cores may provide a particular system function much more efficiently than a digital CMOS block, or they may offer a uniquely new function not available in a CMOS-based solution. Solutions to this challenge beyond the end of CMOS scaling also may lead to new opportunities for such an emerging research device technology to eventually replace the CMOS gate as a new information processing primitive element.

MATERIALS TECHNOLOGIES

The most difficult challenge for Emerging Research Materials is to deliver materials with controlled properties that will enable operation of emerging research devices in high density at the nanometer scale. To improve control of material properties for high density devices, research on materials synthesis must be integrated with work on new and improved metrology and modeling. These important objectives are addressed in the chapter entitled *Emerging Research Materials*.

NANO-INFORMATION PROCESSING TAXONOMY

Information processing to accomplish a specific system function, in general, requires several different interactive layers of technology. The objective off this section is to carefully delineate a taxonomy of these layers to further distinguish the scope of this chapter from that of the Emerging Research Materials Chapter and the Design Chapter.

One comprehensive top-down list of these layers begins with the required application or system function, leading to system architecture, micro- or nanoarchitecture, circuits, devices, and materials. As shown in Figure ERD1 below, a different bottom-up representation of this hierarchy begins with the lowest physical layer represented by a computational state variable and ends with the highest layer represented by the architecture. In this more schematic representation, focused on generic information processing at the device/circuit level, a fundamental unit of information (e.g., a bit) is represented by a computational state variable, for example, the position of a bead in the ancient Abacus calculator or the charge or voltage state of a node capacitance in CMOS logic. A device provides the physical means of representing and manipulating a computational state variable among its two or more allowed discrete states. Eventually, device concepts may transition from simple binary switches to devices with more complex information processing functionality perhaps with multiple fan-in and fan-out. The device is a physical structure resulting from the assemblage of a variety of materials possessing certain desired properties obtained through exercising a set of fabrication processes. An important layer, therefore, encompasses the various materials and processes necessary to fabricate the required device structure, which is the domain of the *ERM chapter*. The data representation is how the computational state variable is encoded by the assemblage of devices to process the bits or data. Two of the most common examples of data representation are binary digital and continuous or analog signaling. This layer is within the scope of the ERD chapter. The architecture plane encompasses three subclasses of this Taxonomy: 1) nanoarchitecture or the physical arrangement or assemblage of devices to form higher level functional primitives to represent and enable execution of a computational model, 2) the computational model that describes the algorithm by which information is processed using the primitives, e.g., logic, arithmetic, memory, cellular nonlinear network (CNN); and 3) the system-level architecture that describes the conceptual structure and functional behavior of the system exercising the computational model. Subclass 1) is within the scope of the ERD chapter, and subclasses 2) and 3) above are within the scope of the Design chapter.

The elements shown in the red-lined yellow boxes represent the current CMOS platform technology that is based on electronic charge as a binary computational state variable. This state variable serves as the foundation for the von Neumann computational system architecture. Analog data representation also is included in the current CMOS platform technology. The other entries grouped in these five categories summarize individual approaches that, combined in some yet to be determined highly innovative fashion, may provide a new highly scalable information processing paradigm.



A Taxonomy for Nano Information Processing Technologies



EMERGING RESEARCH DEVICES

MEMORY TAXONOMY AND DEVICES

The memory technologies tabulated in this section are a representative sample of published research efforts (circa 2007 – 2009) selected to describe some attractive alternative approaches. Historically, very few memory research options yield practical memory devices, and including a particular approach here does not in any way constitute advocacy or endorsement. Conversely, not including a particular concept in this section does not in any way constitute rejection of that approach. This listing does point out that existing research efforts are exploring a variety of basic memory mechanisms. These mechanisms include electronic charge isolated by surrounding dielectrics; remnant polarization on a ferroelectric gate dielectric and resistance change caused by a variety of phenomena. Table ERD2 is an organization or taxonomy of the existing and emerging memory technologies into four categories. A strong theme is the need to monolithically integrate each of these memory options onto a CMOS technology platform in a seamless manner. Fabrication technologies are sought that are modifications of or additions to a CMOS platform technology. A goal is to present the end user with a device that behaves similar to the familiar silicon memory chip.

Table ERD2Memory Taxonomy

Because each of these new approaches attempts to mimic and improve on the capabilities of a present day memory technology, key performance parameters are provided in Table ERD3 for existing baseline and prototypical memory

technologies. These parameters provide relevant benchmarks against which the current and projected performance of each new research memory technology may be compared.

Table ERD3 Current Baseline and Prototypical Memory Technologies

The Emerging Research Memory technology entries in the current version of the roadmap differ in several respects from the 2007 edition. These changes in technology entries dropped and added to this section are captured in the Transition Table for Emerging Research Memory Devices (Table ERD4). The changes are: 1) Drop Nanofloating Gate Memory; 2) Replace Insulator Resistance Change Memory with 3) Fuse/Anti-fuse Memory, 4) Ionic Memory, and 5) Electronic Effects Memory, and lastly 6) Add an entry for Nano-mechanical memory. The reasons and motivations for these changes are given in the Table ERD4.

Table ERD4Transition Table for Emerging Research Memory Devices

This section is organized around a set of eight technology entries shown in the column headers of Table ERD5. These entries were selected using a systematic survey of the literature to determine the areas of greatest worldwide research activity. Each technology entry listed has several sub-categories of devices that are grouped together to simplify the discussion. Key parameters associated with the technologies are listed in the table. For each parameter, three numbers for performance are given which indicate: 1) minimum performance, satisfactory for practical application, 2) theoretically predicted performance values based on calculations and early experimental demonstrations, 3) up-to-date experimental values of these performance parameters reported in the cited technical references.

Table ERD5Emerging Research Resistance-based Memory Devices—
Demonstrated and Projected Parameters

The last row in Table ERD5 contains the number of papers on the particular device technology published in the last two years. It is meant to be a gauge of the amount of research activity currently taking place in the research community and it is a primary metric that determines which of the candidate devices are included in this table. The tables have been extensively footnoted and details may be found in the indicated references. The text associated with the table gives a brief summary of the operating principles of each device and as well as significant issues that are not captured in the table.

MEMORY TAXONOMY

Table ERD2 provides a simple way to categorize memory technologies. In this scheme, equivalent functional elements that make up a cell are identified. For example, the familiar DRAM cell that consists of an access transistor and a capacitor storage node is labeled as a 1T1C technology. Other technologies such as MRAM where data is stored as the spin state in a magnetic material can be represented as a 1T1R technology. Here the resistance "R" indicates that the cell readout is accomplished by sensing the current through the cell. The utility of this form of classification reflects the trend to simplify cells (i.e., reduce cell area) by reducing the number of equivalent elements to a minimum. Thus, early in the development of a given technology it is common to see multi-transistor multi-x (x equals capacitor or resistor) cells. As learning progresses, the structures are scaled down to a producible 1T1x form. The near ideal arrangement is to incorporate the data storage element directly into the transistor structure such that a 1T cell is achieved. In ultra-dense nanoelectronic memory arrays, instead of the transistor "T," a two terminal non-linear diode-like element may be used with a resistive memory element. Such structure is represented as 1D1R technology.

An important property that differentiates emerging memory technologies is whether data can be retained when power is not present. Nonvolatile memory offers essential use advantages, and the degree to which non-volatility exists is measured in terms of the length of time that data can be expected to be retained. Volatile memories also have a characteristic retention time that can vary from milliseconds to (for practical purposes) the length of time that power remains on.

MEMORY DEVICES

Ferroelectric FET Memory—The Ferroelectric FET (FeFET)² is a 1T memory device where a ferroelectric capacitor is integrated into the gate stack of a FET. The ferroelectric polarization directly affects charges in the channel and leads

to a defined shift of the output characteristics of the FET. At the channel interface, a high quality insulator is required to guarantee a low interface state density. For this reason it is very difficult to fabricate a FeFET with excellent electrical properties.³ When a ferroelectric film is deposited on directly a Si substrate, a diffusion of atoms from and into the ferroelectric film may result in the degradation of its electrical properties. In order to avoid the diffusion problem, an insulating buffer layer is inserted between a ferroelectric film and Si substrate.⁴ Hence, the resulting gate structure consists of a metal-ferroelectric-insulator-semiconductor (MFIS) gate stack. Sometimes, another metal layer is introduced between the ferroelectric and the insulator (MFMIS). The FeFET device scales as a MOSFET. However, scaling is projected to end approximately with the 22 nm generation, because the insulation layer becomes too thin and the properties of the ferroelectric with respect to thickness dependence of the coercive field will not allow further reduction.⁵ In the last decade, many attempts have been made to fabricate FeFET-based nonvolatile memories. The major challenge is the long-term reliability related to the ferroelectric-semiconductor interface. Recently, new materials for the FeFET stack were reported, such as organic ferroelectrics, nanotubes⁶, nanowires⁷, and graphene⁸, Using an organic ferroelectric film as a gate dielectric allows for elimination of the buffer layer, due to lower crystallization temperature of organic materials, and therefore suppression of the diffusion.⁹ Another important challenge is the rather short retention time, approximately 30 days, for the FeFET.

Nanoelectromechanical memory (NEMM)—the NEMM is based on a bi-stable nano-electromechanical switch. In this concept, mechanical digital signals are represented by displacements of solid nanoelements (e.g., nanowires, nanorods or nanoparticles), which result in closing or opening an electrical circuit. The original concept of NEMM was a carbon nanotube cross-bar memory.¹⁰ Each memory element is based on a suspended crossed CNT. A cross-bar array of CNTs forms mechanically bi-stable, electrostatically-switchable device elements at each cross point, and the memory state is read out as the junction resistance. Several different modifications of suspended-beam NEMMs are currently being explored using different materials including carbon nanotubes¹¹, Si¹², Ge¹³, and TiN.¹⁴ A difficult challenge of the suspended-beam NEMM is scalabily: according to a recent study, it might be difficult to achieve low-voltage (~1 V) operation for the beam length less than 50 nm.¹⁵

There are also emerging NEMM concepts other than suspended beam memory. Recently, a nanoelectromechanical device was proposed in which an iron nanoparticle shuttle is controllably moved by an electrical signal within a hollow nanotube channel.¹⁶ Different positions of the nanoparticle within the channel results in different channel resistances. Another recent work reported a bi-stable current-voltage behavior in discontinuous 5-10 nm thin films of graphitic sheets; a nanoelectromechanical mechanism was proposed for this effect.¹⁷

Spin Transfer Torque Memory (STTRAM)—STTRAM is an advanced version of the MRAM with a different WRITE mechanism. The memory cell consists of a semiconductor isolation device and a magnetic tunnel junction (MTJ) with two ferromagnetic layers separated by a MgO-based tunneling barrier layer in which thickness is controlled to approximately 1 nm.¹⁸ Switching MTJ states occurs as a result of the flow of spin-polarized conduction electrons through the MTJ. The spin-polarized current transfers angular momentum to the spins of the core electrons in the magnetic free layer causing it to switch. The READ operation of the STTRAM is the same as that of the MRAM. Utilizing spin-polarized electrons for WRITE operation allows for a considerable reduction both in the WRITE energy and the cell size¹⁹, and greatly improves the writing current scalability into future technology generations. Key factors in STTRAM development include reducing STT writing current and voltage (energy) while maintaining adequate thermal stability. Use of perpendicular magnetic anisotropic films shows particular promise for this requirement.²⁰ Additional key factors include establishing adequate margins between read/write voltages and write/ breakdown voltages for reliable high speed operation.

Nanothermal Memory—In Nanothermal Memory elements, consisting of a nano-scale metal-insulator-metal (MIM) structure, typical resistive switching phenomena are based on thermal effects which result in unipolar switching characteristics. Two different thermal processes are operative. One is initiated by a voltage-induced partial dielectric breakdown in which the material in a discharge filament is significantly modified due to Joule heating. Because of the current compliance, only a weak conductive filament with a controlled resistance is formed. This filament may be composed of the electrode metal transported into the insulator, carbon from residual organics²¹, or decomposed insulator material such as sub-oxides.²² During the reset transition, this conductive filament is disrupted thermally because of high power density in the order of 10¹² W/cm³ generated locally. This mechanism is referred to as the thermochemical type (or the fuse – antifuse type). One candidate out of many for the insulator is NiO, first reported in the 1960s.²³ Recently, the filamentary nature of the conductive path in the ON-state has been confirmed (for NiO²⁴ and TiO²⁵). Pt/NiO/Pt thin film cells have been successfully integrated into CMOS technology to demonstrate nonvolatile memory operation.²⁶ A critical parameter for this unipolar switching effect seems to be the value of the current compliance. It should be also noted that while thermal effects seem to be dominant in the unipolar resistive switching,

there are indications, that additional processes also play a role^{27, 28, 29}, For example, electrochemical effects, such as a thermally-assisted redox process, appear to play a role in the nanothermal memory operating mechanism.^{30, 31, 32}

The other type of Nanothermal memory is nanowire-structured phase change cell^{33, 34}, for which the underlying phase transformation between amorphous and crystalline phases is similar to the conventional phase-change memory (PCM) described in *Table ERD3*. Compared to prototypical PCM, the switching current and therefore the write energy could be considerably reduced in nanowire-PCM cells.³⁵ The principle challenge is fabrication of the nanowire crossbar memory cells containing the required select diodes, perhaps using self-assembly or directed-assembly technology.

Nanoionic Memory-The Nanoionic memory operation is based on a change in resistance of a MIM structure caused by ion (cation or anion) migration combined with redox processes involving the electrode material or the insulator material, or both.^{36, 37} The insulator is required to exhibit ionic conductivity. The material class is comprised of oxides, higher chalcogenide (including glasses), and semiconductors, as well as organic compounds including polymers. One variant is based on cation transport, cathodic reduction and growth of metallic filaments. Primarily Ag and Cu based systems have been successfully realized in demonstrator cells.^{38, 39} A filament thus formed connecting opposite electrodes leads to a low-resistance state, while reversal of the polarity of the applied voltage causes oxidation which dissolves the filament and restores the high-resistance state. Other variants originate from anion (e.g., oxygen ion) transport and redox reactions that introduce an electronic conduction within the insulator material itself.^{40, 41} In some cases, a formation process is required before the bi-stable switching can be started. Often, the conduction is of filamentary nature. If this effect can be controlled, memories based on this bi-stable switching process can be scaled to very small feature sizes. The switching speed is limited by the ion transport. If the active distance, which is relevant for the redox controlled bi-stable switching, is small (in the < 10 nm regime) the switching time can be as low as a few nanoseconds. Precise predictions are not yet possible, because many details of the mechanism of the reported phenomena are still unknown. Developing an understanding of the physical mechanisms governing switching of the Nanoionic Memory is a key challenge for this technology,

Electronic Effects Memory—The Electronics Effects memory includes three different mechanisms operative in a MIM structure: 1) Charge injection and trapping, 2) Mott transition, and 3) Ferroelectric polarization effects.

1) Charge injection and trapping can be a cause for changes in resistance.^{42, 43, 44} In the charge-trap model (Simmons-Verderber theory⁴⁵), charges are injected by Fowler-Nordheim tunneling at high electric fields and subsequently trapped at sites such as defects within the (semi) insulator or semiconductor. This modifies the interface barrier structure and, hence, the resistance of the device. A materials issue that requires further investigation is cycling fatigue – defect formation during switch cycles.⁴⁶ The formation of defects may limit both the lifetime and the dimensional scaling due statistical uncertainty of defect density distribution.⁴⁷ Also, there is a fundamental concern on scalability of the charge-trapping memory below 100 nm.

2) In the *Mott Transition Memory*, charge injection induces a transition from strongly correlated to weakly correlated electrons, resulting in an insulator-metal transition. The Mott transition mechanism was reported for several oxide systems such as PCMO - (Pr,Ca)MnO₃,^{48, 49}, STO - SrTiO₃:Cr⁵⁰, CeO₂/LCMO⁵¹, TiO₂/TiN⁵², and VO₂.⁵³ A generic model has been presented.⁵⁴ A critical issue for this type of device is the sensitivity of the behavior of correlated electrons to small changes in parameters, including charge density, strain, disorder, and local chemical composition.⁵⁵ Thus, precise control of the physical and chemical structure of the material and interfaces is crucial. The transition between the two electrical states can, in principle, be induced by different external stimuli, such as electric field, photo- or thermal excitation.⁵⁶ While the electric field-induced transitions are used in the memory device, a strong thermal component may be present and the device can undergo a thermal phase transition between metallic and insulating state. In some cases, e.g., for the VO₂ Mott memory, the metal-insulator phase transition temperature is ~ 68° C, and it might be difficult to control the thermal environment to prevent such a transition.

More recently, a new metal-insulator transition device has been proposed which is based on formation of a quasi-twodimensional electron gas (2DEG) at the ferroelectric oxide interface, such as KNbO₃/BaTiO₃.⁵⁷ By controlling the polarization, the interface may be highly conductive (2DEG) or insulating. Ferroelectrics are expected to have long retention times and exhibit a resistance difference of many orders of magnitude, which would be very suitable for memory operation.

3) *Ferroelectric polarization* can modify the tunneling properties of ultrathin films or modify the Schottky-type spacecharge layer in adjacent semiconducting layers^{58, 59} resulting in ferroelectric resistive switching. Experimentally, the realization of ferroelectric tunnel devices is challenging as they require ultrathin (~nm) high-quality ferroelectric barrier layers.⁶⁰ Recently, robust ferroelectricity has been obtained in highly strained BaTiO₃ films down to 1 nm thickness, and resistive readout of the polarization state by measuring the tunnel current by conductive-tip AFM has been demonstrated.⁶¹ In this work, scalability down to 70 nm has been achieved.

Macromolecular Memory—Macromolecular memory, sometimes referred to as polymer or organic memory, consists of a memory element, which is a thin-film of organic material.⁶² It is, in some cases, similar to molecular memory, but extreme scaling is not important, while reduced fabrication cost is emphasized.

The active organic insulator layer in the macromolecular MIM memory often contains embedded metal components, which could be a thin metal layer, or metal nanoclusters^{63, 64}, or metal ions in organometallic materials such CuTCNQ and AgTCNQ (TCNQ=7,7,8,8-tetracyano-p-quinodimethane).⁶⁵ A memory device based on C₆₀ molecules embedded in polymer film was also reported⁶⁶; each of these structures can exhibit two states of different conductivities, at the same applied voltage. The WRITE operation is performed by applying a voltage pulse to the structure, which results in reversible switching between a low-resistance and a high-resistance state. After transition occurs, the device remains in one of two states after turning off the power. The ERASE operation is performed by application of a reverse voltage pulse.

Experimental results suggest that the embedded metal layer plays a critical role in bistable I-V characteristics of many macromolecular memory elements.⁶⁷ The memory operation mechanisms are still unclear. Some research⁶⁸ suggests that the changes in resistance could be due to trapping the charge in the discrete metal nanocrystals, indicating that this type of memory falls into the class of electronic effect memories. In some cases, there are also indications of ionic mechanisms.⁶⁹

Molecular Memory-Molecular memory is a broad term encompassing different proposals for using individual molecules or small clusters of molecules as building blocks of memory cells. In the molecular memory, data are stored by applying an external voltage that causes a transition of the molecule into one of two possible conduction states. Data is read by measuring resistance changes in the molecular cell. The concept emphasizes extreme scaling; in principle, one bit of information can be stored in the space of a single molecule. Computing with molecules as circuit building blocks is an exciting concept with several desirable advantages over conventional circuit elements. Because of their small size, very dense circuits could be built, and bottom-up self-assembly of molecules in complex structures could be applied to augment top-down lithography fabrication techniques. As all molecules of one type are identical, molecular switches should have identical characteristics, thus reducing the problem of variability of components. However, the success of molecular electronics depends on our understanding of the phenomena accompanying molecular switching, where currently many questions remain.⁷⁰ Early experiments on the reversible change in electrical conductance generated considerable interest.^{71, 72} However, further studies revealed several serious challenges for single/few molecule devices due to extreme sensitivity of the device characteristics to the exterior parameters such as contacts, reproducible nanogap, environment etc. Also, there are multiple mechanisms contributing to the electrical characteristics of the molecular devices, e.g., the conductivity switching as an intrinsic behavior of molecular switches may often be masked by other effects, such as e.g., in some cases, formation of metal filaments along the molecule attached between two metal electrodes.⁷³ In other cases, intrinsic molecular switching has been reported, and a 160-kbit molecular memory has been fabricated.⁷⁴ Molecular memory is viewed as a long term research goal. The knowledge base for molecular electronics needs further fundamental work, which is currently under way.⁷⁵

LOGIC AND ALTERNATIVE INFORMATION PROCESSING DEVICES

The first three editions of the ERD Logic section have evaluated alternative logic technology entries in terms of their potential to displace scaled CMOS devices in high performance general purpose computing. The conclusion reached in those editions was that none of the alternative technologies surveyed had a high potential for displacing scaled CMOS devices on the ITRS roadmap scheduled for the 2020's. This conclusion and similar findings elsewhere have driven a tremendous interest in "More than Moore" applications and underlying technologies. Therefore, in light of the huge amount of research activity taking place in alternative technologies, a natural question is whether there are some useful information processing functions other than general purpose Boolean logic where the particular physical characteristics of an emerging technology could offer an advantage relative to and combined with scaled CMOS. This edition of the Emerging Research Device chapter addresses this and related questions.

As in previous sections of the ERD chapter, this section includes a transition table, *Table ERD6*. The transition table provides the disposition of technologies transitioned into and out of *Table ERD7a* from 2007 to 2009. It also previews new technologies that may be included in the 2011 edition. The Transition Table (Table ERD6) contains four new Technology Entries included for the first time in the Logic and Information processing section.

Table ERD6Transition Table for Emerging Research Logic Devices

Any analysis of projected performance of alternative devices for non-Boolean applications is intimately related to the associated architectural configuration. The current industry trend towards heterogeneous asymmetric multicore processors is consistent with the idea that future systems could support dedicated coprocessors utilizing novel devices for specialized applications. These dedicated coprocessors and accelerators would be integrated as one or more cores dedicated to specific operations in an otherwise conventional general purpose CMOS-based system. Specific examples might include an image recognition or speech recognition coprocessor, a Bayesian inference engine for data mining, or an associative memory unit for synthesis applications.

The Emerging Research Device Logic section was expanded in 2007 to include consideration of novel devices for general purpose, Boolean logic, as well as special purpose applications described above. That edition was therefore organized around two tables rather than one. The first table in 2007 was very similar to the logic table in previous editions of the ERD chapter. It contained parameters for the alternative technology entries relevant to evaluating performance of that technology for general purpose Boolean logic. The second table evaluated technologies that might be useful for alternative information processing but could not be considered viable technologies for general purpose computing. It was and is well understood that the novel applications and novel devices will require novel architectures and cannot be thought of as drop in replacements for CMOS.

In this edition, the Emerging Research Logic Device section is expanded further and organized around three tables. These tables are labeled: ERD7a. "MOSFET: Extending the Channel of MOSFETs to the End of the Roadmap"; ERD7b "Charge based Beyond CMOS: Non-Conventional FETs and other Charge-based information carrier devices," and ERD7c; "Non-FET, Non Charge-based 'Beyond CMOS Devices," respectively. The titles are very indicative of the content of the tables. The first table, ERD7a, contains extensions and enhancements to current MOSFETs. They are charge based and utilize basic field effect functionality. Table ERD7b entries all involve electron transport but the switching function is inherently different from the field effect transistor and include such effects as quantum mechanical tunneling and Coulomb blockade. Table ERD7c entries involve information carriers other than electronic charge and effect such as spin wave interference and magnetic exchange coupling. It is likely that these technology entries will not be suitable for general purpose computing but might be suitable for special purpose computing such as cryptography, image processing, and inference engines.

Table ERD7a MOSFETS: Extending the channel to the End of the Roadmap

Table ERD7bCharge-based Beyond CMOS: Non-Conventional FETs and other Charge-based
Information Carrier Devices

Table ERD7cAlternative Information Processing Devices

Finally, the ITRS' International Roadmap Committee (IRC), recognizing that it may be timely to accelerate development of one or two of the most promising proposals for well-defined new information processing devices, requested the Emerging Research Devices and Emerging Research Materials working groups to recommend one or two of the most promising emerging research device technologies for detailed roadmapping and accelerated development.

In response, the ERD/ERM working groups conducted a study encompassing a Workshop entitled "Maturity Evaluation for Selected beyond CMOS Emerging Technologies" and a subsequent ERD/ERM working group meeting to develop a recommendation to the IRC.⁷⁶ The objective of this study was to evaluate several "Beyond CMOS" candidate information processing technologies, followed by near-consensus selection of one or two Technology Entries having high potential for enabling a paradigm shifting information processing technology. The single candidate selected was "Carbon-based Nanoelectronics," which was recommended to the ITRS' IRC for *their* decision regarding its potential. The IRC supported the ERD/ERM Working Group's choice of Carbon-based Nanoelectronics for additional focused Roadmapping and accelerated development.

Table ERD8 presents a roadmap for "Carbon-based Nanoelectronics," employed as a candidate *possible* solution for a "channel replacement material," for extending CMOS to the end of the Roadmap. Focused research necessary to develop either carbon nanotubes or graphene for this application may provide a technology platform and requisite scientific knowledge base to enable discovery of a new paradigm for processing information for "beyond CMOS."

Table ERD8Research and Technology Development Schedule proposed for Carbon-basedNanoelectronics to impact the Industry's Timetable for Scaling Information Processing Technologies.

LOGIC DEVICES

As mentioned previously, this edition of the ERD includes two tables devoted to logic devices and one table devoted to alternative information processing devices and technologies. The first table contains MOSFET structures with alternative channel materials or different geometry and which use field effect functionality in the channel to modulate the current flow. It uses a parameter set relevant to high performance, general purpose logic. All quantitative parameters have separate entries for best projected value and best demonstrated value and the values for each are referenced. It should be especially noted that all the entries in a selected column do not necessarily relate to the same device as is implied by separate references. In general, the entries in a given column refer to the best reported values for that particular metric and the different entries will in general come from separate devices

MOSFET: Extending the Channel of MOSFETs to the End of the Roadmap

Carbon Nanotube FETs—The primary potential advantages of Carbon Nanotube FETs are the high mobility of charge carriers and the potential to minimize the subthreshold slope (i.e., minimize the short channel effects) by a surround gate geometry. On the other hand, there are multiple challenges to achieving this, including: 1) the ability to control bandgap, 2) growth of the nanotubes in required locations and directions, 3) control of charge carrier type and concentration, 4) deposition of a gate dielectric, and 5) formation of a low resistance electrical contact.

In the past two years, a recently developed CNT FET compact model⁷⁷ projects an ideal 25-50X increase in switching speed over 32nm CMOS, but only a 2-10X improvement when parasitic resistances and capacitances are included in practical circuits. A FET has been fabricated with a measured f_T of 4GHz⁷⁸ and a projected f_T of 30GHz. A FET has been fabricated with a measured f_T of 4GHz⁷⁸ and a projected f_T of 30GHz. A FET has been fabricated with a measured f_T of 4GHz⁷⁸ and a projected f_T of 30GHz. A FET has been fabricated with a semiconducting CNT channel and a metallic CNT gate electrode and modeling analysis projects a 5ps gate delay.⁷⁹ A CNT FET memory device⁸⁰ with HfO₂ gate demonstrated read and write switching times of 100 ns which was hypothesized to be the result of fast traps at the CNT-HfO₂ interface.

Of the remaining challenges, most progress has been achieved in sorting carbon nanotubes with chemical techniques to achieve a distribution of CNTs with controlled band gap, but the purity is still many orders of magnitude less than would be required for fabrication of integrated circuits. While researchers have used field controlled growth to fabricate a CNT gate aligned over a CNT channel and others have used dielectrophoresis to assemble CNTs in higher density, additional increases in density are needed. Others have deposited highly aligned CNTs on quartz substrates, transferred the CNTs to 4-inch silicon wafers and fabricated a variety of operating CNT circuits.^{80A} However, the ultimate goal of directly depositing only semiconducting CNTs as a high mobility channel replacement material on a silicon wafer and fabricating CNT circuits in an otherwise conventional CMOS process flow still remains elusive. Little progress if any has been made in controlling carrier type and concentration and this is still done by attaching molecules to the surface of the CNT. The ability to deposit a 20nm HfO₂ surround gate dielectric has been demonstrated, but the role of fast interface traps in operation as a FET needs to be understood. Contact formation on small diameter nanotubes has produced wide variability and recent characterization indicates that the Schottky barrier height of Pd-CNTs is close to the bandgap of the nanotubes

Graphene Nanoribbon FETs—Graphene materials offer the potential of the extremely high carrier mobilities available to CNTs (without the need for controlling CNT chirality), combined with the promise of patterning graphene nanoribbons using conventional processes. Work on graphene field effect transistors (FETs), while still at an early stage, is proceeding at a rapid pace. Beginning with the first description of the electric field effect in graphene in 2004 ⁸¹, graphene FETs using bottom gating⁸², top-gating^{83, 84, 85}, dual-gating⁸⁶, and side-gating.⁸⁷ have now been demonstrated using various combinations of exfoliated⁸⁸, epitaxial^{89, 90, 91}, stamped exfoliated⁹², organically-grown⁹³, and chemically-derived graphene.⁹⁴

A majority of graphene FET research uses exfoliated graphene to form the channel material. Back-gated graphene FETs with SiO₂ dielectric were shown⁹⁵ to have field-effect mobilities up to 10,000cm²/V s (note that although back-gating is not desirable for FET-based circuitry, this mobility value does act as a useful comparison to top-gate mobility values). A SiO₂ dielectric top-gated transistor (W=265 nm/L=7.3 μ m)⁹⁶ has been demonstrated with hole mobilities of 710 cm²/V s and electron mobilities of 530 cm²/V s. An HfO₂ dielectric top-gate graphene FET⁹⁷ was reported to yield Ion/Ioff~7 (T=1.7K) and a transconductance of 2.9 mS/ μ m. More recently, a dual-gate Al₂O₃/SiO₂ dielectric graphene FET (L=2.4 μ m)⁹⁸ was demonstrated with electron mobility of 8600 cm²/Vs.

FET devices using epitaxial graphene comprise a smaller proportion of the research but give similar results. For example,⁹⁹ mobilities of 535 cm²/V s on polystyrene dielectric top-gated graphene FETs have been reported , and SiO₂ dielectric top-gate transistors (W=50 μ m/L=400 μ m)¹⁰⁰ have been demonstrated with electron mobilities ~1500 cm²/V s, hole mobilities ~ 3400 cm²/V s, and Ion/Ioff~2.5. Additionally, HfO₂ dielectric top-gated transistors (W=5 μ m/L=10 μ m) have been reported¹⁰¹ with C-face mobility of 5000 cm²/V s and Ion/Ioff~2.5-3 and Si-face mobility of 1200 cm²/V s and Ion/Ioff ~5.

The predictions of high current densities, extraordinary mobilities, and superior FET performance,^{102, 103} all with the goal of compatibility with CMOS process and temperature range¹⁰⁴ continue to drive the rapid pace of innovation in graphene FETs. This innovation, accompanied by evidence of tunable Ion/Ioff via bandgaps^{105, 106} and increasing intrinsic carrier mobilities in the range of 7×10^4 cm²/V-s at room temperature^{106A} and 2-3 x 10^5 cm²/V s at 5K^{107, 108} is likely to bring rapid advance in this area over the next few years, hence ERD's identification of graphene as a possible solution for extending CMOS to the end of the Roadmap (*Table ERD8*).

An important problem with graphene for digital applications is its zero bandgap which in turn will result in a very small Ion/Ioff ratio. To open up the band gap one either has to build devices with graphene nano ribbons (d < 5nm) or find other methods to locally open up the bandgap. Very little is known about the transport properties of these narrow ribbons, although passivation of the ribbon edges is a major challenge. An important application space for graphene may be RF with discrete elements and high linearity requirements.

Nanowire Field-Effect Transistors (NWFETs)—Nanowire field-effect transistors are FET structures in which the conventional planar MOSFET channel is replaced with a semiconducting nanowire. Such nanowires have been demonstrated with diameters as small as 0.5 nm.¹⁰⁹ They may be composed of a wide variety of materials, including silicon, germanium, various III-V compound semiconductors (GaN, AlN, InN, GaP, InP, GaAs, InAs), II-VI materials (CdSe, ZnSe, CdS, ZnS), as well as semiconducting oxides (In₂O₃, ZnO, TiO₂), etc.¹¹⁰ Importantly, at low diameters, these nanowires exhibit quantum confinement behavior, i.e., 1-D conduction, that may permit the reduction of short channel effects and other limitations to the scaling of planar MOSFETs. To first order the 1D effect observed in transport is related to a 1D density of states and leads to somewhat modified charge carrier scattering. Electrostatic behavior is to first order classically described.

Important progress has been made in the manufacture of semiconducting nanowires for use as FET channels, for which there are two principal methods. The first method is nanoimprint lithography, by which semiconducting channels are formed through a printing or stamping process.¹¹¹ The second is catalyzed chemical vapor deposition.^{112, 113} In particular, the vapor-liquid-solid (VLS) growth mechanism has been used to demonstrate a variety of nanowires, including core-shell and core-multishell heterostructures.^{114, 115} Heterogeneous composite nanowire structures have been configured in both core-shell and longitudinally segmented configurations using group IV and compound semiconductor materials. The longitudinally segmented configurations are grown epitaxially so that the material interfaces are perpendicular to the axis of the nanowire. This allows significant lattice mismatches without significant defects. Vertical transistors have been fabricated in this manner using Si¹¹⁶, InAs^{117, 118}, and ZnO¹¹⁹, with quite good characteristics. Core-shell gate-all-around configurations¹²⁰ display excellent gate control and few short channel effects.

Circuit functionality of nanowire devices has been demonstrated, including individual CMOS logic gates¹²¹, as well as a PMOS ring oscillator that was fabricated successfully and exhibited \sim 12 MHz operation.¹²² The performance of the test circuit was limited by interconnect capacitance and thus did not achieve the THz operation that is predicted to be the intrinsic capability of such devices.¹²³ Issues such as nanowire placement, contacting, and integration remain to be solved before widespread adoption can occur.

To take advantage of the better scaling behavior of NWs, new technologies are needed to build NW based circuits at the required integration densities. This problem will require innovation in NW contacts and integration flows. At this point it is not clear if vertical wire configurations are a useful solution for this problem

III-V channel replacement devices—High mobility III-V compound semiconductor materials are attractive candidates as channel replacement materials for nMOSFETs. However, in general, there is a trade-off between mobility and bandgap; high mobility materials such as InAs have narrower bandgap. There is also a tradeoff between low bandgap enabling lower voltage operation (to address the power/performance balance) versus excess leakage current. In addition to the narrow bandgap, the energy difference between the lowest and the second lowest conduction bands tends to be small, which increases the population of electrons in the second lowest conduction band making the mobility worse.¹²⁴ Therefore, ternary compound semiconductors such as InGaAs have attracted much attention because of their moderate bandgap and the acceptable energy difference between the lowest and the second lowest

conduction band minima. However, challenges for III-V channels are not only channel material selection but also gate stacks^{125, 126}, stress engineering¹²⁷, growth on Si¹²⁸, surface passivation¹²⁹, and low-resistance S/D formation. A number of high-k gate dielectrics such as ZrO_2^{130} , $Al_2O_3^{131}$, HfO_2^{132} , and $HfAlO_{133}$ for use with III-V structures have been investigated but no clear winner has emerged. Although III-V materials have small piezoreistance coefficient, stress engineering does shows a performance enhancement¹³⁴ with applied stress. Attachment of III-V structure on Si by direct bonding¹³⁵ has been demonstrated. Although surface passivation is another major concern, it is reported that Inrich InGaAs has no Fermi-level pinning.^{136, 137} Ga-rich Ga₂O₃(Gd₂O₃)/InGaAs has a free-moving Fermi-level near the conduction and valence-band edges¹³⁸, and Silane-Ammonia surface passivation technology makes it possible to realize Dit of IE11eV-1cm⁻².¹³⁹

Buried, short-channel III-V HEMT structures have been fabricated which show^{140, 141} clear performance advantages over conventional Si MOSFETs. However, similar III-V MOSFETs incorporating a surface-channel do not show similar improvements relative to Si MOSFETs and will require significant improvements to be commercially feasible. The conventional HEMT design may not be suitable for digital applications since it may not meet the density requirements for a competitive gate conductor pitch. Therfore, a self-aligned HEMT solution is required. Furthermore the HEMT operating voltage is restricted due to excessive gate leakage

Ge channel replacement devices—Germanium as a channel replacement material has attracted great attention because of its excellent electron and hole mobilities. In particular, the hole mobility of strained germanium is much better than that of silicon.^{142, 143} On the other hand, because of the small bandgap, band-to-band tunneling leakage current, or GIDL (gate-induced drain leakage) can be large. Therefore, any potential use of Ge as a channel replacement material requires an ultrathin Ge film in order to control the leakage.¹⁴⁴ The gate stack is another difficult challenge for Ge MOSFETs. A Si cap layer is frequently used to make a good interface as well as to reduce the electric field inside the germanium to control BTBT leakage reduction.¹⁴⁵ High quality GeO₂/Ge interfaces have been studied by several research groups.^{146, 147, 148} High-κ gate dielectrics such as ZrO₂, HfO₂¹⁴⁹, and SrGex¹⁵⁰, have also been investigated. Although high hole mobility in pFETs has been demonstrated by many research groups, electron mobility in nFETs is not very good in spite of high electron mobility in bulk Ge. Because currently available stressed Si PMOS technology out performs unstressed Ge based PFETs, only stressed Ge channels may be competitive. Also Ge may not have a scaling advantage over Si because the lower bandgap will require more graded junctions and therefore an increase of Rext. Low resistance S/D formation is relatively easy for Ge pFETs, because of Fermi-level pinning at the valence band edge. On the other hand, low resistance S/D formation is very difficult for Ge nFETs. Recently, operation of short-channel Ge MOSFETs with gate lengths less than 80nm have been reported.^{151, 152} Although the progress of EOT and gate-length scaling will be required, Ge or Ge-rich pMOSFETs are good candidates for future generation MOSFETs. The performance of the n-channel Ge MOSFET also needs substantial improvement.

Unconventional Geometries for FET devices—"Unconventional geometries for FET devices" are defined as FET structure other than the conventional planar MISFET structure. The basic operation principles for these devices are, however, very similar to the more conventional planar MISFET case. In this section the focus is on three-dimensional multi-gate FETs in different configurations including vertical channel devices (Surround Gate Transistor (SGT)¹⁵³ Vertical Replacement Gate (VRG)¹⁵⁴ and horizontal channel devices.^{155, 156, 157, 158}). Most of these devices are fabricated as Fully-Depleted (FD) channel FETs.

Threshold voltage Vt in multi-gate FETs can be adjusted by work function engineering of the gate material, not by the channel doping. The work function engineering in the gate stack is mandatory for CMOS applications to achieve proper Vt for both n/p type channels. Due to its excellent gate controlled electrostatics, FD channel multi-gate FETs have smaller subthreshold swing and higher punch through immunity in the short gate length region, compared to the conventional single gate MOSFET case. One restriction for FD channel multi-gate FETs is the channel thickness. For example, the channel thickness should be less than $1/2\sim2/3$ of the minimum gate length for the double gate FET. Such fully depleted devices with thin channels, however, will challenge junction design for Rext engineering.

Even though multi-gate FETs were first conceived and developed several years ago, device integration and several applications are currently being discussed in the technical literature. Recent papers discuss logic, analog / RF ¹⁵⁹ and memory cells including SRAM¹⁶⁰, DRAM¹⁶¹, and NAND flash memories.^{162, 163, 164} It is also expected that the device characteristic variability caused by the random dopant fluctuation (RDF) is reduced in this structure, due to smaller dopant concentration in the channel^{165, 166, 167} relative to conventional planar MOSFETs.

Unconventional geometry devices can potentially deliver high device integration density and power efficiency at good speed if the concerns are solved for threshold voltage adjustment and large parasitic resistance and capacitance. Use of merged epitaxial-Si S/D¹⁶⁸ or dopant segregated Schottky S/D have been explored for FinFETs¹⁶⁹ to reduce the parasitic resistances. Device reliability issues and ESD characteristics for these devices are now under investigation

and they should be improved.¹⁷⁰ Once such concerns are resolved, FD channel multi-gate FETs should become plausible candidates for future CMOS alternatives for below the 22 nm generation, where the suppression of both short channel effect and characteristic variability are important for VLSI applications.

Charge based Beyond CMOS: Non-Conventional FETs and other Charge-based information carrier devices

As mentioned previously, the 2009 edition of the ERD chapter is built around three tables. The first table contains extensions to CMOS derived from new, high mobility channel replacement materials and unconventional geometries. The parameters chosen are those relevant to high performance general purpose computing. The second table contains charge based devices potentially suitable for logic applications but having functionality quite different from the Field Effect Transistor. Since these devices can potentially apply to logic applications, they are characterized by the same set of parameters as the first table.

Tunnel FETs—Tunnel FETs are gated reverse-biased p-i-n junctions that are expected to have OFF–ON transitions much more abrupt than conventional MOSFETs, whose 60-mV/dec subthreshold swing limit is set by the thermal injection of carriers from the source to the channel.^{171, 172, 173} Without a gate voltage, the width of the energy barrier between the intrinsic region and the p+ region is wider than the minimum for significant band-to-band tunneling probability, and the device is in the OFF-state. As the positive gate voltage increases, the bands in the intrinsic region are pushed down in energy, narrowing the tunneling barrier and allowing tunneling current to flow. Band-to-band tunneling (BTBT) is a quantum-mechanical phenomenon, expected to provide a much more abrupt transition between ON and OFF states of a three terminal switch compared to the 60mV/decade MOSFET limit. Tunnel FETs are actively investigated due to their potential for low standby leakage current and as enablers of future logic circuits operating with a supply voltage smaller than 0.5V and saving many decades of Ioff. This is essentially due to the fact that the subthreshold swing of Tunnel FET can be significantly lower than the 60mV/decade MOSFET limit at room temperature. Recent reports suggest that Tunnel FETs could be also considered as promising candidates for the high performance switch, by using appropriate heterostructure architectures¹⁷⁴ and/or exploiting low band-gap materials such as III-V compound semiconductors, Ge, SiGe, or graphene. Tunnel FETs are expected to match or even outperform the speed performance of CMOS (in terms of equivalent CV/I metrics) at the same supply voltage.

Many detailed device simulations have predicted that BTBT FETs could produce subthreshold swings below the thermal limit in conventional semiconductor materials such as silicon¹⁷⁵ or SiGe¹⁷⁶, carbon-nanotube (CNT)¹⁷⁷ or graphene based transistors.¹⁷⁸ Since the tunneling current is determined by the bandgap and effective mass of the material, the silicon TFET seems limited by its low on-state current density and, probably, only high amounts of strain at the tunneling junction (>3GPa) can improve their behavior.¹⁷⁹ The first CNT based TFET was¹⁸⁰ demonstrated with a subthrehold slope of 40mV/decade and a Si TFET¹⁸¹ was demonstrated with S=52.8mV/decade and Ion=12.1 μ A/ μ m at a 1V supply voltage (which is about two orders of magnitude lower than a high-performance n-channel MOSFET). However, such small subthreshold slopes have been reported rather as point values and not as average values on more than three decades of current, which is essential for their success. To improve the drive current, the most effective way is to change the channel material to a material with a narrower band gap and smaller effective mass. Tunnel FETs were¹⁸² demonstrated on Si_xGe_(1-x)OI substrates, and tunnel FETs on strained Ge substrates¹⁸³ were fabricated which simultaneously showed a subthreshold swing below 60mV/dec and high Ion (300 μ A/ μ m).

Key challenges for tunnel FETs include optimization of the device architecture for high Ion combined with an average subthreshold swing lower than 60mV/decade over at least four decades of current. Particularly, engineering of the source tunneling region (junction abruptness, band-gap, carrier effective mass) and enhancement of gate control on internal electric field are of major importance for achieving an experimental tunnel FET device, matching the predictions made by numerical simulations. With this respect, tunnel FETs can benefit from heterostructures needing low bandgap materials on silicon advanced platforms, which define a clear technological challenge. There is a stringent need of more experimental demonstration of tunnel FET architectures enabling the sub-0.5V low power switch. Moreover, for future design of integrated circuits based on tunnel FETs, development of device compact modeling is needed.

Impact Ionization MOS (IMOS)—The fundamental limit of the MOSFET subthreshold slope at (kt/e) Log10 (60mV/decade) is linked to the diffusion type of transport in the subthreshold regime and has significant implications on the lowest supply voltage allowable for the design of ultra scaled devices. At the sub 30nm scale, devices will have to face short channel and DIBL effects. The strongest design lever that can limit power dissipation (either static or dynamic) is the decrease of supply voltage, while overdrive is needed to get high performance. That is why a steep subthreshold voltage is needed while maintaining a low switching threshold. The microelectronics research community has thus envisaged the use of limiting phenomena other than electrostatic voltage barriers and diffusive

transport for switching. In addition to Tunnel FETs, impact ionization based FETs, called IMOS^{184, 185}, have been proposed as candidates to meet these requirements. Generally, low bandgap materials are used to build IMOS devices. The I-MOS is composed of a PIN diode, whose intrinsic area is partially covered by a gate.

The attractiveness of IMOS comes from its potential co-integration with CMOS. The use of IMOS devices has been evaluated to potentially reduce by 75% CMOS inverter switching current and to improve the static noise margin of 6T SRAM cells by 22%.^{186, 187}

IMOS is used in an avalanche mode following the impact ionization regime in order to obtain a very steep increase of current via carrier multiplication by electron hole pair generation.¹⁸⁸ It should be pointed out that avalanche charge multiplication is an intrinsically slow and statistical process with a stochastic nature that will inject additional variability into IMOS device design.

A high electric field is needed to produce impact ionization¹⁸⁹ and potentially limits the supply voltage to values higher than the bandgap of some materials that would otherwise be good channel replacement candidates.¹⁹⁷ In general¹⁹⁰ a drift zone is necessary to allow generation of sufficient electrons and this will induce an offset in the electrical output characteristics.

Impact ionization generates high energy carriers which induce increased trapping in the gate dielectric¹⁹¹, but, using the IMOS in aTFET mode will limit the degradation issues.¹⁹²

The literature generally reports numerous subthreshold slopes lower than $60 \text{mV/dec}^{193, 194, 195, 196}$ with the record lowest value at 2mV /dec.¹⁹⁷ However, geometries are generally relaxed relative to the most advanced CMOS generations. For bulk silicon, the smallest functional device demonstrated 5.3V avalanche breakdowndown in a 40nm device but also demonstrated a large leakage current.¹⁹⁸ Thin SOI will help to reduce the DIBL that can appear in small geometries, as well as in devices operating under the avalanche regime.

Increased drivability due to higher impact ionization generation rates and reduced breakdown voltages would make the device more scalable allowing use of germanium in which impact ionization is higher than in silicon.^{199, 200} The use of IMOS in a TFET mode increases its drivability as well as its reliability.²⁰¹ An additional challenge is the integration density possible with the IMOS device.

Spin Transistor—Spin transistors can be classified into 'Non-Conventional Charge-based Extended CMOS Devices'. They exhibit the transistor behavior with functions of magnetoresistive devices. The most important feature is the control of transistor output via spin or magnetization. Spin transistors can be divided into two categories, i.e., spin-FET and spin-MOSFET. Although the source and drain of both the devices are composed of a ferromagnetic material, their operating principles are quite different. In the spin-FET, the switching operation can be achieved by spin precession or dephasing of spin-polarized carriers injected in the channel. On the other hand, relative magnetization configurations of the source and drain are used to modify the output current for the spin-MOSFET.

The Datta-Das spin-FET²⁰² employs the Rashba spin-orbit interaction to induce spin precession of spin-polarized carriers in the channel with the magnetization configuration of the source/drain fixed. Therefore, materials with a strong spin-orbit interaction are required for the channel effectively to induce the Rashba effect. The unique output characteristics, including oscillating output current or negative differential resistance, are attractive. However, the Datta-Das spin-FET channel length is not scalable, and is not likely to achieve a high on-current to off-current (on/off) ratio.²⁰³ Several modified versions of original Datta-Das spin-FET have been proposed.^{204, 205, 206}

The Hall-Flatte spin-FET²⁰⁷ is the FET type of spin transistor with an insulated gate structure. Gate-controlled spinorbit interaction is used for its switching. Coherent spin transport without the spin-orbit interaction causes the off-state, and dephasing induced by the spin-orbit interaction is applied to the on-state. A high spin injection efficiency of 100% is necessary in order to achieve high performance. The problems are similar for the Datta-Das spin-FET and the Hall-Flatte spin-FET.²⁰⁸ Neither the Datta-Das spin-FET nor the Hall-Flatte spin-FET has made major progress or a breakthrough for their realization since the previous 2007 ITRS edition.

In contrast the channel of a spin-MOSFET²⁰⁹ is comprised of silicon with a very weak spin-orbit interaction. Therefore, the output current is modulated by magnetization configurations of the source and drain. Since the spin-MOSFET requires no spin precession of spin-polarized carriers in the channel, it has an excellent scalability similar to conventional MOSFETs. In contrast to the spin-FET, the cutoff state of the spin-MOSFET is simply achieved by a gate bias condition in the same manner as for ordinary MOSFETs (i.e., magnetization-configuration-independent cutoff condition is used as its off-state). In addition, the spin-MOSFET exhibits two different current drive capabilities that are controlled by magnetization configurations of the source and drain. These features play an essential role for reconfigurable logic and nonvolatile logic applications.

Although the spin-MOSFET has not been realized yet, there has been important progress and verification for its elemental technology. In unrelated but relevant experiments, spin injection and spin transport in Si were experimentally demonstrated by using a spin-LED²¹⁰, a non-local multi-terminal device²¹¹, and a hot-electron-based spin-injector/detector device.²¹² It is particularly worth noting that coherent spin transport in Si over 350 microns was observed.²¹³ An important remaining requirement is the injection of a high percentage of spin-polarized electrons from a half-metal source into the channel. Another major challenge is how these devices will be interconnected. One approach is to use spin-wires and another is for all wiring to be done with conventional spin independent charge carrier flow.

Single-electron Transistors (SETs)—SETs²¹⁴ are three-terminal devices that switch on/off tunnel currents conveying electrons that are being transported one by one from source to drain through a small island. Potentially, SETs can be applied to general purpose Boolean logic, but significant circuit and architecture changes will be required. SETs can potentially deliver high device density and power efficiency at good speed if the issues of the large threshold voltage variation and the low current drivability can be solved.

Most of the new applications and architectures have focused on how to utilize the unique characteristics of SETs, i.e., Coulomb blockade oscillations, to provide a new function or to perform a function with a reduced number of devices compared to CMOS. They include digital applications such as SET/CMOS hybrid multi-value logic circuits²¹⁵ and reconfigurable logic circuits²¹⁶ as well as analogue applications such as multiband filtering circuits²¹⁷, analog pattern matching circuits²¹⁸, and associative recognition tasks.²¹⁹ Due to the above noted issues which make it difficult for SETs to compete directly with CMOS²²⁰, developing a hybrid circuit with CMOS would help utilize the novel functionalities of SETs in functional circuits. While early SET devices required cryogenic temperatures, room-temperature SET operation has been reported, more recently.^{221, 222} However, the key issue to broader application remains the large threshold voltage variation caused by stray charges. The large variability in device performance continues to impede large-scale integration. Engineering breakthroughs, possibly with the bottom-up approach, are needed to eliminate the size and background charge fluctuations.²²³

SET-based logic is usually a "voltage state logic" where a bit is represented by the voltage of capacitor charged with many electrons. There is another approach called "charge state logic" in which a bit is represented by a single electron, typically implemented with single-electron manipulation/transfer devices. This includes the binary-decision-diagram circuits²²⁴ and the quantum-dot cellular automaton circuits.²²⁵ The merit of charge state logic is that there is no static power dissipation. However, the issue of the high bit error rate requires unconventional circuit designs with fault-tolerant schemes, or novel applications such as stochastic information processing which use single electrons as physical random numbers.^{226, 227} At this point, the feasibility at the circuit level is not clear.

Research has progressed since 2007 in the fabrication of room-temperature operating devices by means of ultra narrow silicon channel²²⁸ and silicidation of silicon nanowires²²⁹, as well as the new circuit application such as reconfigurable logic circuits²³⁰ and random-number-based circuits.²³¹

NEMS Switch—This discussion builds on the maturity evaluation report of the 2008 ERD/ERM Working Group²³², which summarizes the motivations and performance attributes of nanoelectromechanical (NEM) switches. Here, recent developments in NEM switches for logic are reviewed especially to obtain benchmarking data for *Table ERD7b*. The primary advantages of the NEM switch over CMOS are low static power dissipation and radiation tolerance. Recent simulation results²³³ explore NEM relay design strategies for digital logic and I/O that can significantly improve the energy efficiency of the whole VLSI system. By exploiting the low effective threshold voltage and zero leakage achievable with these relays, it is shown that NEM relay-based adders can theoretically achieve an order of magnitude or more improvement in energy efficiency over CMOS adders with ns-range delays and with no area penalty. In addition, design approaches have matured^{234, 235, 236} and the first submicron demonstrations have appeared using metal cantilever²³⁷ and carbon nanotube approaches.^{238, 239} While switches have been demonstrated, the efforts have not yet demonstrated the construction of logic functions or the direct measurement of switching frequency or energy. A key challenge is to achieve volt-level operation at GHz switching speed with reliable wear and repeatability. Furthermore, NEM devices and circuits must be able to handle switching current transients without burn-out. Approaches to minimize switching energy in which mechanical energy is stored and released to assist the transition between logic states merit further development.²⁴⁰

Negative gate capacitance FET—Based on the energy landscapes of ferroelectric capacitors²⁴¹, it has been suggested that by replacing the standard insulator of a MOSFET gate stack with a ferroelectric insulator of appropriate thickness it should be possible to implement a step-up voltage transformer that will amplify the gate voltage, thus leading to values of S lower than 60 mV/decade and enabling low voltage/low power operation; this device is called a negative capacitance FET. The main advantage of such a device²⁴² is that it involves no change in the basic physics of the FET

and thus does not affect its current drive or impose other restrictions; thus, high Ion levels, similar to advanced CMOS would be achievable with lower voltages.

To date, the concept of a negative capacitance Fe-FET has been reported based on simulations and analytical derivations.^{243, 244} An experimental attempt to demonstrate small swing Fe-FET, based on a PVDF organic ferroelectric gate stack was reported at IEDM 2008²⁴⁵; however, the very small current region (~1pA) in which this small swing has been observed make the data susceptible to leakage errors. Therefore, further experimental proof is needed. Other mechanisms such as avalanche breakdown and polaronic effects within the oxide have been cited as possibly providing the positive feedback needed for negative capacitance.²⁴⁶ An analog mechanical equivalent of negative capacitance has been reported²⁴⁷, showing that BaTiO₃ can have a negative stiffness which can be stabilized by combining it with positive stiffness materials and a large amplification of total stiffness can be experimentally observed. This property is the mechanical analog of the negative capacitance; thus one can speculate that, in theory, there is no fundamental reason why an inherently unstable system cannot be stabilized and a practical benefit in reducing the subthreshold swing obtained.

The major challenge concerns identification of adapted gate materials (ferroelectrics and/or oxides) that can experimentally support the demonstration of a negative capacitance in a FET gate stack. One should note that the negative swing region is unstable and its stabilization can be achieved only by adding in series a positive capacitance. It was mentioned that such a device would not show any hysteresis. Due to the difficulty of capacitance matching (negative with positive) over a large range of voltage, obtaining simultaneously a zero hysteresis and a sharp turn on characteristic may prove to be challenging. In principle, the scalability of the device should be similar to the one of a MOSFET. However, it is not yet clear what materials should compose the gate stack in order to achieve the positive feedback and if this negative capacitance property could be maintained when scaling down the dielectric thickness of the operation voltages. Moreover, the negative capacitance itself, even if achievable, is not enough for a high performance switch; other challenges include the response time of the polarization in ferroelectric materials. Organic ferroelectrics, such as PVDF, have response time of the order of only hundreds of ns and a limited range of temperature operation. The inorganic ferroelectrics have much faster response times but their high processing temperature makes their integration into CMOS difficult.

Alternative Information Processing Devices- The previous two Emerging Research Logic Device tables considered devices which utilized electronic charge as the information carrier and could be applied to Boolean logic operations. This section will involve information cariers other than electronic charge and employ physical phenomena much different than FETs, such as spin wave interference and magnetic exchange coupling, to provide a new information processing functionality. It is likely that these technology entries will not be suitable for general purpose computing but might be suitable for special purpose computing such as cryptography, image processing, and inference engines.

This section is motivated by the observation that some of the emerging research devices may have unique physical response characteristics better suited for applications involving recognition, mining, and synthesis than general purpose computation. The current industry trend toward heterogeneous multicore systems will, in principle, allow inclusion of such devices into a hybrid system architecture combining special purpose processors containing novel devices with more conventional, general purpose processors all integrated onto a silicon CMOS platform.

The characteristics of present day CMOS devices have been optimized to give a strongly non-linear bi-stable response function that maximizes the I_{on}/I_{off} ratio and minimizes the sub-threshold slope. These characteristics are very well suited to binary logic operations and in some sense, the CMOS device may be thought of as the natural device for binary Boolean logic. However, the physics of some of the alternative devices being investigated are quite different from the physics of CMOS devices. Examples include response characteristics quite different from the strongly bistable response of CMOS devices. Examples include response functions with a peaked structure (either in voltage or frequency), response functions with more than two stable states, periodic response functions, and sigmoidal response functions with tunable slopes and curvatures. Each of these response functions may be associated with some "natural" application or algorithmic application. Some of those applications will be discussed here.

The other factor relevant to consideration of alternative information processing devices is that some, if not most, future information processing will be done on information which is something other than binary data sets represented by electronic charge. This information includes optical images, image sequences, speech, and data sets derived from physical sensors. This data set (or data set in the form of a time series) is usually analog in nature. In order to process analog signals in general purpose digital computers, A/D and D/A conversions must be done and these can add significantly to the thermal budget, energy budget, and overall cost. In some of these cases, it may be more efficient to process the data in its original analog representation rather than convert everything to a digital (i.e., binary, electrical) representation and use a general purpose processor. In other cases, especially those in which total power or total energy is strongly constrained, a hybrid analog/digital system may be optimal.

NON-FET, NON CHARGE-BASED "BEYOND CMOS" DEVICES

The entries in the Non-FET, Non Charge-Based 'Beyond CMOS Devices table include novel devices that may prove useful for various information processing tasks other than high performance, general purpose computing. Some of the more specialized tasks include associative processing, communication, multivalued logic and ferromagnetic elements for non-volatility and radiation hardness and error tolerance. In general, these may require a functional organization other than von Neumann architecture. The task or application and architectural configuration are noted in the table and linked to the entries in the Emerging Research Architecture section.

Collective Spin Device—Ferromagnetic logic devices are a class of alternative logic devices that use the local magnetization orientation of a domain of a ferromagnetic material to store the computational state. In the nomenclature adopted here, FM devices are distinct from spin devices, which are based on the individual dynamics of spin of one or a few charge carriers. FM devices have the potential of being non-volatile and radiation hard, which is derived from the properties of the ferromagnetic materials themselves. While many ferromagnetic metals have Currie temperatures well above room temperature, the Currie temperatures of most ferromagnetic semiconductors are still limited to well below room temperature.

The scaling of FM devices was analyzed and compared to CMOS.^{248, 249, 250} A majority of these scaling studies found that the FM devices are uniformly slower than CMOS, but their switching energy is smaller and in some cases denser. Theoretically they can provide similar computation throughput at lower power dissipation although none of the FM devices have been reduced to practice.

Magnetic amplifiers designed to amplify or restore a magnetic signal have been proposed in Mn:GaAs²⁵¹ and Mn:Ge.²⁵². Both concepts operate by triggering a spontaneous ferromagnetic transition in a bipolar-like heterostructure by modulating the charge density in a dilute magnetic semiconductor (DMS) channel. The goal is to develop an element for future magnetic circuits that could restore magnetic state to overcome losses inherent in any physical system. Fabrication efforts are underway to experimentally demonstrate the operation of this device.²⁵³ A major challenge to such devices is the lack of a demonstrated room temperature DMS material.

Spin communication structures that rely on propagation of coherent spin waves without electron current flow have been designed, simulated, fabricated, and measured.²⁵⁴ Signals are excited and propagated in a Fe ferromagnetic film and the propagation length is measured as a function of frequency and applied field. Ferromagnetic spin waves with a frequency of a few GHz were clearly detected.

Magnetic switches based on asymmetric nanorings have been designed and tested.²⁵⁵ These are based on competition between the exchange energy and the magnetostatic energy in nanomagnets. The relative probability of exciting different magnetic modes and hence the interaction energy in symmetric nanorings is dictated by the ring geometry and cannot be altered after fabrication. However, progress is being made²⁵⁶ in engineering the internal magnetic state of the rings through introduction of geometric features and defects. These features can control the evolution of the magnetic states of individual rings to enable specialized computational response characteristics. Another approach is to use asymmetric nanorings, which allow tuning the asymmetry electrically, acting as a gate to modulate the interaction.

Self check-pointing architectures²⁵⁷ have recently been demonstrated that use spin torque memory devices integrated with traditional CMOS logic to provide high-performance nonvolatile storage that can dramatically reduce off state leakage power relative to a similar all CMOS device. The proposed microprocessor uses the magnetoelectronic devices to "snapshot" the state of the currently executing program at regular intervals, providing protection against power failures and enabling rapid context switching. Research results for novel ferromagnetic devices, examples of which are discussed in this section, represent important steps towards a viable magnetic technology based on collective magnetic effects in ferromagnetic materials.

Moving Domain Wall Devices—Domain Wall (DW) logic devices²⁵⁸ are formed by ferromagnetic wires in which domain walls propagate in separate regions with different directions of magnetization. DW devices can have the architecture of AND, OR, and NOT logic gates as well as majority gates. An external "clocking" magnetic field is usually required to move the domain walls. The method for generating such pulsed magnetic fields remains the major challenge of this type of device. Recently a NOT gate functionality formed by moving a domain wall between two nanocontacts was demonstrated²⁵⁹ at a relatively high frequency of 7.6MHz. A domain wall shift register has been run with an even higher speed of 25MHz.²⁶⁰ Recent results show that the current-driven domain wall propagation in room temperature metal could reach as fast as 100 m/s.²⁶¹ A concept called "magnetic racetrack memory"²⁶²was proposed to achieve a new data storage system, based on the current-induced domain wall propagation, with performance comparable to DRAM and cost comparable to Hard Disk Drive (HDD). It has been demonstrated by writing and reading a train of 100 pulses in an array of magnetic nanowires.²⁶³ Basic studies probing such details can lead to new

device structures, e.g., use of spin polarized currents to move domain walls in the specially shaped metal for the memory and amplifier applications.²⁶⁴ The challenges of current-driven domain wall devices are the high power dissipation caused by high currents required and the relatively slow switching speeds

Atomic Switch—The atomic switch is an MIM electrochemical switch that uses a local oxidation/reduction process to form metallic nanofilaments connecting two dissimilar metallic electrodes thereby establishing a low-resistance state. Reversal of the polarity of the applied voltage enables the redox process to dissolve the nanofilaments thereby resetting the high resistance state. The atomic switch is one member of a group referred to as the Resistance Random Access Memories (ReRAM).²⁶⁵ One distinguishing difference between the ReRAM and the atomic switch is the atomic switch has a reversible electrode for introducing metal atoms (cations) into the ionic conductive materials to form a conductive path between the reversible electrode and the counter electrode²⁶⁶, This mechanism has been experimentally²⁶⁷ and theoretically²⁶⁸ confirmed. On the other hand, both electrodes of ReRAMs are inert, and the diffusion of oxygen vacancies is controlled to form a conductive path.

The atomic switch was initially developed as a two-terminal device using a sulfide^{269, 270, 271}, which has demonstrated the logical operation based on the crossbar architecture²⁷² with potential scalability down to 20 nm.²⁷³ Memristive operation²⁷⁴, which has been demonstrated using a TiO₂ switch²⁷⁵, also shows the potential of electrochemical switches for logical application. More recently, nonvolatile three-terminal operation has been demonstrated²⁷⁶, where formation and annihilation of one or more metal filaments between the source and drain electrodes are controlled using metal cations supplied from the gate electrode.

The advantages of the three-terminal atomic switch are high ON/OFF ratio, low ON-resistance, nonvolatility, and low power consumption. An adequate retention time of the metal filament has been confirmed.^{277, 278} Switching speed and cyclic endurance are challenges for the three-terminal atomic switch, although a switching speed of 10 MHz and a switching repetition of 10¹¹ cycles have been confirmed by the two-terminal atomic switch.²⁷⁹

In the last few years, an atomic switch fabrication process using fully CMOS compatible materials has been developed to enable the formation of atomic switches in the metal interconnect layers of CMOS devices.^{280, 281, 282, 283, 284, 285} This has enabled development of a new type of programmable logic device²⁸⁶, which can facilitate logic circuits based on the CMOL architecture to realize new type of nonvolatile logic circuits with CMOS devices.

Switching speed, cyclic endurance, uniformities of the switching bias voltage and resistances both for the on-state and the off-state need to be improved for general usage as a logic device. There have been many reports on resistive switching using metal oxides, in which some devices worked as unipolar devices and others as bipolar devices. For this reason, understanding the device physics is a most important and urgent issue. In addition, development of the architecture for nonvolatile devices is desired to be compatible with other nonvolatile logic devices.

Molecular Devices—Molecular devices have remained a very active research area with significant activity in the three principle areas of Contacts, Density functional theory and Molecular Switching.²⁸⁷ Research in Molecular Contacts has focused on the length-dependent changes in molecular orbital alignment and coupling with contact states. Experimental measurements of thermopower on a series of phenylenediamines, phenylenedithiols, and alkanedithiols were made and found to agree well with corresponding calculations.²⁸⁸ In a related study, the electronic transport properties were found to depend strongly on the nature of the contact itself (chemical bond or physical contact).²⁸⁹

Electronic transport properties strongly depend on the specific materials with special interest in the use of Si as a substrate. It was shown that the Si (111) and Si (100) surfaces form bonds with C either via a Si-C or Si-O-C bond. These materials are useful as template for molecular electronic junctions²⁹⁰, although possibly the layers are not completely ordered.

Different research groups have tried to assemble monolayers of molecules on Si (111) 7x7 surfaces but the resulting monolayers were disordered. However on reconstructed surfaces of Si (111) 3x3, Ag hexagonal-ordered arrays of C70 were achieved. The integration and electrical characterization of molecular monolayers were investigated by contacting them with conducting polymer PEDOT/PSS291].

In parallel with experimental investigations, theory groups are increasingly active in the area of electronic properties of single molecules and Self Assembled Monolayers (SAMs) and are able to address progressively more complex issues. For example, conductance histograms extracted from STM measurements have been compared to quantum chemistry calculations. They propose that certain combinations of different S-Au couplings and trans/gauche conformations determine the conductance through the Au/alkanedithiol/Au junctions

A detailed general picture comprising charge trapping coupled with electronic and vibrational polarization effects has been applied to weak and strong coupling of the molecules. This picture can explain a surprisingly large range of phenomena including NDR, dynamical switching, switching noise, current hysteresis, etc.²⁹²

The ultimate goal of molecular hysteretic switching however remains difficult to achieve. Early work showed the clearest direct evidence so far, but several more recent attempts to demonstrated switching (some which did not use electrical control) showed various degrees of success:

Azobenzene derivatives in the trans form, adsorbed in a homogeneous two-dimensional layer, have been shown to be collectively switched with spatial selectivity, thus forming a periodic pattern of *cis* isomers. The probability of a molecule switching is not equally distributed, but is strongly dependent on both the surrounding molecules and the supporting surface, which precisely determine the switching capability of each individual molecule. Consequently, exactly the same lattices of cis isomers are created in repeated erasing and reswitching cycles. These results demonstrate a conceptually new approach to spatially addressing single functional molecules.²⁹³

A second approach, light controlled conductance switching of ordered metal-molecule-metal devices, has also been demonstrated.²⁹⁴ Specifically, the group demonstrated reversible, light-controlled conductance switching of molecular devices based on photochromic diarylethene molecules. These devices consist of ordered, two-dimensional lattices of gold nanoparticles, in which neighboring particles are bridged by switchable molecules. They independently confirm that reversible isomerization of the diarylethenes employed is at the heart of the room-temperature conductance switching.

Another approach uses organic functional molecules with optoelectronic properties that undergo reversible structural transformation in response to external stimuli such as light, protons, and metal ions. These functional molecules are potentially suitable for molecular states, logic gates and combinational logic circuits.^{295, 296, 297}

Bilayer, "Pseudospintronic devices" and the BiSFET in particular—"Pseudo-spins" are discrete degrees of freedom other than spin which can still be treated much like spin. There are in principle many possible forms of pseudospin and many possible device applications. As one example, collective pseudospin effects, much like those for spin in a ferromagnet, may be particularly interesting.

In the case of the proposed "Bilayer pseudoSpin Field-Effect Transistor" (BiSFET), "top" and "bottom" graphene layers take the place of spin up and spin down. BiSFET operation would employ gate-modified interlayer collective many-body exciton tunneling/recombination in graphene bilayer systems to achieve ultra-low power (~0.008aJ) yet still high speed (~100GHz) above room temperature switching.²⁹⁸ Furthermore, while pseudospin based, the BiSFET is also still charge based, alleviating any need to convert between spin-state variables and charge-state variables.

In bilayer systems, it is possible for electrons in one layer to pair with holes (both Fermions) in the other layer resulting in electron-hole-pairs/excitons (Bosons). The excitons can then condense, altering the quantum wave-functions in the bilayer qualitatively. This qualitative change effectively shorts the two layers, reducing the tunnel resistance from a large value to a value essentially limited only by the contacts. The dramatic reduction in tunnel resistance applies only for small interlayer bias, however, because too much current destroys the condensate and produces a negative differential resistance (NDR), and does so for interlayer voltages potentially small compared to k_BT .²⁹⁹

So far such condensates have been observed only at very low temperatures and under high magnetic fields in GaAs/AlGaAs double quantum well systems.^{300, 301, 302} However, as a consequence of a synergy of multiple properties of graphene— single atomic layer thickness, nearly perfect electron-hole symmetry in the band structure, low density of states, and zero bandgap—it has recently been predicted that this condensate could occur above room temperature in otherwise weakly coupled and oppositely charged graphene double/bi-layer systems.^{303, 304}

The proposed BiSFET would use gates to control the relative electron and hole carrier densities in the opposing layers and, thus, the quality of the exciton condensate and the peak interlayer current before the onset of NDR.

The BiSFET is not a drop in replacement for CMOS. However, with gate-controllable NDR and a clocked power supply voltage, SPICE simulations confirm that it is possible to arrange two or a few BiSFETs into gate-switchable bistable systems to form at least all of the basic logic functions—Invert, NAND, NOR, AND, OR. More complex and novel functionality including memory may be possible as well, leading to sequential logic. With high peak current densities, 100 GHz clock frequencies are projected, and with sub- k_BT onset NDR and supply voltages only on the scale of k_BT in these simulations, very low power operation on the order of 0.008aJ per switching event per BiSFET is shown to be possible.³⁰⁵ At this stage, however, the BiSFET is only a concept based on novel predicted physics in a novel material system. Furthermore, the proposed BiSFET requires radically different ways of implementing digital logic now being explored in SPICE-based simulations. And even if theory holds, fabrication of devices with the necessary degree of control of surface quality, work functions, lithography, etc. will impose numerous challenges as well.

Nanomagnetic logic devices (NML formerly known as MQCA)—Nanomagnetic devices exploit magnetic phenomena for logic based on physically-coupled single-domain nanomagnets.³⁰⁶ This scheme is based on the shape-dependent switching of magnetic elements³⁰⁷, and the use of an applied magnetic-field clock for switching.³⁰⁸ A three-input majority-logic gate, functioning as a universal nanomagnetic logic element has been experimentally demonstrated.³⁰⁹ A recent review of nanomagnetic logic is presented in "Magnetic Quantum-dot Cellular Automata: Recent Developments and Prospects"³¹⁰.

Nanomagnetic logic has potential advantages relative to CMOS of being non-volatile, dense, low-power, and radiation-hard. Such magnetic elements are compatible with MRAM technology, which can provide input-output interfaces. Compatibility with MRAM also promises a natural integration of memory and logic. Nanomagnetic logic also appears to be scalable to the ultimate limit of using individual atomic spins.³¹¹

Key challenges for nanomagnetic logic include the need for a local magnetic field for switching. Recent work has focused on local clocking fields generated by current-carrying wires.³¹² Simulations show that such a scheme is capable of producing the required magnetic-field strengths. A reasonably complete analysis of the energy required by the circuitry and the wires for creating the clocking fields plus the energy consumed in magnet switching, shows that the total energy required for a complete nanomagnetic logic system (a 32 bit ripple-carry adder (RCA)) is much less than that projected for a comparable CMOS circuits.^{313, 314} The energy required for the wires to create the clocking field is distributed among a large number of nanomagnets and is thus difficult to apportion among individual nanomagnets or even individual nanomagnetic logic gates. The results of "Clocking Structures and Power Analysis for Nanomagnet-based Logic Devices,"³¹⁵ project a total energy dissipation for a complete 32-bit RCA (including the wires for the clocking field) of 4E-15 J. Potential avenues to further reduce the switching power include the engineering of materials properties for switching, such as biaxial anisotropy³¹⁶, or the use of other physical phenomena, such as the spin-torque transfer effect.

RSFQ—RSFQ devices, systems and circuits have been developed, prototyped, and fabricated. Design and fabrication lines for RSFQ systems exist. They could become an important technology if the correct market driver emerges. For these devices cryogenic operation, cost and material integration issues limit their application space, and, therefore, they are not included in *Table ERD7c*.

EMERGING RESEARCH ARCHITECTURES

INTRODUCTION

The objective of the Emerging Research Architecture (ERA) section is to identify possible applications for emerging logic and memory devices. This is a difficult challenge because in many cases, circuit-level models and/or architecture-level models for these devices and their interconnect systems either do not exist or they are very primitive. Moreover, the envisioned applications for these new devices can take many forms; i.e., i) as a drop-in replacement for CMOS, ii) as supplemental devices that complement and coexist with CMOS devices, or iii) as devices whose unusual properties can provide unique functionality for selected information processing applications. With this in mind, this ERA section is organized to reflect the potential application space for emerging research devices from an architecture perspective. The first sub-section is focused on Benchmarking methodologies for various emerging logic devices versus end-of-ITRS-CMOS and contains proposed benchmarking metrics as well as examples of a benchmarking exercise. The second sub-section is focused on architectures for memory hierarchies of emerging memory devices and proposes methodologies to assess the performance of these systems. The third sub-section provides an example of a possible novel application of emerging research devices in inference computing that may be able to leverage these devices to perform computation in a highly efficient non-digital fashion. Finally, the last subsection describes an approach to estimating bounds on the best possible performance achievable for configurations of interconnected devices based on a bottom-up formulation that relies on limiting device and interconnect models for physical layout and for energy usage.

LOGIC ARCHITECTURE BENCHMARKING

Proposed post-CMOS replacement devices are very different from their CMOS counterparts, and often pass computational state variables (or tokens) other than charge. Alternative state variables include collective or single spin, excitonic, plasmonic, photonic, magnetic, qubit, and even material domains (e.g., ferromagnetic). With the multiplicity

of programs characterizing the physics of proposed new structures, it is critical to examine circuits and architectures in which these devices might complete transactions effectively. Data suggest that as devices approach quantummechanical limits, fundamental relationships between practical minimum delay and energy consumption across these disparate technologies may become evident.

ARCHITECTURAL REQUIREMENTS FOR A COMPETITIVE LOGIC DEVICE

The circuit designer and architect depend on the logic switch to exhibit specific desired characteristics in order to insure successful realization of a wide range of applications. These characteristics³¹⁷ which have since been supplemented in the literature, include:

- Inversion and Flexibility (can form an infinite number of logic functions)
- Isolation (output does not affect input)
- Logic gain (output may drive more than one following gate and a High Ion/Ioff Ratio)
- Logical completeness the device is capable of realizing any arbitrary logic function.
- Self Restoring / Stable (Signal quality restored in each gate)
- Low cost Manufacturability (Robust operating physics, acceptable process tolerance)
- Reliability (Aging, wear-out, radiation immunity)
- Performance (transaction throughput improvement)
- "Span of Control" is an important means of connecting device performance and area to communication performance, relating time to space. The metric measures how other devices may be contacted within a characteristic delay of the switch, and is dependent not only on switch delay, but switch area as well as communication speed.³¹⁸ Successful architectures need effective fan-out.

Devices with intrinsic properties supporting the above features will be adopted more readily by the industry.

ARCHITECTURE ASSESSMENT METHODOLOGY

Several approaches can be taken to assess the impact of a particular device/interconnect technology on information processing architectures. These include the use of a set of quantitative metrics to compare various High Performance Computing (HPC) alternatives that capture the impact of the devices and their interconnect system on the acceleration of a particular function in a computing system. Table ERD9 contains an exemplary set of metrics which have been used by multiple institutions to characterize various logic devices and their interconnect systems in the context of architecture applications.

Table ERD9 Prototypical Criteria for Emerging Device Architectures

OBSERVATIONS

A number of common themes have emerged in the observations made during recent studies of post-CMOS replacement switches.³¹⁹ A few noteworthy concepts are shared below.

- Most of the architectures that have been considered to date in the context of new devices utilize binary logic to implement von Neumann computing structures. In this area, CMOS implementations are difficult to supplant because they are very competitive across the spectrum of energy, delay, and area. Novel electronbased devices appear to be the best candidates as a drop-in replacement for CMOS for binary logic applications.
- 2) As the behavior of these emerging research devices becomes better understood, novel architectures that leverage these features will begin to appear.
- 3) The low voltage energy-delay tradeoff conundrum which plagues CMOS appears to extend into many of the post-CMOS proposed structures as well.
- 4) Even in the Boolean regime, the trend from multi-core to many-core is supported by the energy-delay exhibited by these new switches. Extending device "fan-out," the number of logic gates which can receive a preceding gate's output, remains a profound challenge for the devices examined. This may promote alternative architectures that do not require high fan-out to improve through-put.
- 5) Increasing functional integration and on-chip switch count will continue to grow. To that end, in any logic architectural alternative, both flexible rich logic circuit libraries and reconfigurability will be required for new switch implementations.

- 6) Patterning, precision layer deposition, material purity, dopant placement, and alignment precision critical to CMOS will continue to be important in the realization of architectures using these new switches.
- Assessment of novel architectures using new switches must also include the transport mechanism for the information tokens. Fundamental relationships connecting information generation with information communication spatially and temporally will dictate CMOS' successor.

Table ERD10 provides a brief overview of the attributes and possible applications for several emerging research devices in the binary logic domain.

Table ERD10Summary of the Attributes of Several Emerging Research Devices and Projections for
their Application Spaces

QUANTITATIVE RESULTS

Preliminary analyses sponsored by the SRC/NRI (to be published in the Proceedings of the IEEE in 2010) surveyed the potential logic opportunities afforded by emerging research switches using a variety of information tokens and communication transport mechanisms. Specifically, the projected effectiveness of these devices used in a number of logic gate configurations was evaluated, and normalized to CMOS at the 15nm generation as captured by ITRS.

The data from this study corroborates anecdotal insights from earlier works, suggesting that envisioned early new switch structures are superior to CMOS in logic settings in energy and area, but inferior to CMOS in delay. Median delays for various complex logic gates expressed in the new switches modeled ranged from 11X to 24X the delay of CMOS in the same technology era. Median modeled energy consumption in these complex logic gates, on the other hand, was superior to CMOS, ranging from 0.004X to 0.29X the energy of 15nm CMOS. Similarly, the median area for complex logic circuits expressed in new switches ranged from 0.23X to 0.62X that of CMOS. Figure ERD2 below illustrates energy, delay, and area design space for the NAND2 circuit with a fan-out of 1 expressed in some of the emerging research switches. These data are of course only a first glimpse at the advantages and challenges associated with new switches, and will change dramatically as these alternatives are more thoroughly developed.



Figure ERD2 Exemplary Design Space for a NAND2 Circuit with Fan-out of 1 in Various New Emerging Research Switches. Red points are charge based; other points are examples of switches using other state variables

ACCELERATORS IN THE FUTURE LOGIC ARCHITECTURE LANDSCAPE

It is clear that CMOS will remain the primary basis for High Performance Computing (HPC) for many years. New logic and memory devices will likely gain entry into HPC via the performance of special functions supporting the core processors that are implemented primarily in CMOS. One potential area for entry is that of accelerators that could off-load specific computations from the core processors and provide overall improvement in system performance. Examples of widely-used accelerators include: Encryption/Decryption, Compression/Decompression, Floating Point Units, Digital Signal Processors, etc. As a general rule, an accelerator is considered as an adjunct to the core processors if replacing its software implementation improves overall core processor throughput by approximately ten percent. An accelerator that is offered as a CMOS replacement should offer a performance improvement relative to its CMOS implementation of an order of magnitude. There may be instances where the unique characteristics of emerging devices can be used to advantage in implementing an accelerator. Emerging research switches with properties which more effectively perform these specific transactions most likely will require processing which differentiate them from conventional CMOS; it is unlikely than a common substrate can host both simultaneously. Enablements, such as 3D integration, which allow disparate processes to play together, will become increasing valuable tools.

EMERGING RESEARCH MEMORY ARCHITECTURES

INTRODUCTION

The objective of the ERA section is to identify possible applications for emerging logic and memory devices. This section focuses specifically on memory devices. The open question is to what extent will emerging research memory devices create opportunities to solve circuit and system problems in new and useful ways? Is there potential for an all-purpose memory device? Or will emerging research memory devices simply be drop in replacements for SRAM, DRAM and flash? This section has been organized to reflect the problems in existing memory sub-systems, and the opportunities that emerging memory devices might create. The second section summarizes the current challenges for memory systems. The third section lists some of the opportunities that might be uniquely addressed by emerging memory devices. The last section briefly addresses a research agenda to address these challenges and opportunities.

CHALLENGES IN MEMORY SYSTEMS

The general challenges that arise from deployment of today's available memory devices are summarized in Table ERD11. Scaled area factor is the most important challenge. Core density scales poorly today and must scale better in future devices. Also today, peripheral circuit area is often larger than core memory area. Scaled area factor is the prime determiner of cost and is likely to get worse in emerging devices. Power consumption is an enormous challenge The power overhead of the memory hierarchy must be reduced by more than an order of magnitude in order to sustain scaling of computational performance at constant power levels. What makes this challenge particularly difficult is the need to scale up bandwidth dramatically at the same time. In addition, it is important to remember that latency minimization has been the traditional goal of the memory hierarchy designer, as it has a first order impact on performance. While latency might be traded for reduced power consumption in future systems, this balancing cannot be one sided. There are strong tradeoffs between factors above that in the past have lead to radically different design points. For example, fill factor can be balanced against power and/or bandwidth. Approaches are needed that are globally Pareto optimal.

Table ERD11 Challenges in Existing Memory Systems Illustrated using Today's Memory Devices

The Soft Error Rate (SER) of logic and SRAM do not scale well in future technology generations. (DRAM is less likely to suffer an SER problem.) This problem must be resolved at the device level for SRAM successors, if at all possible. In order to cope with logic and memory failures, in high reliability systems, the memory hierarchy is regularly check-pointed (copied to disk) in order to provide a degree of fault tolerance. The energy and bandwidth costs of check-pointing gets worse in future systems and it would be highly desirable to have solid state storage that can negate the need of check-pointing to disk. Today Flash cannot provide this function due to the write cycle limit.

Of course the ultimate goal of an emerging memory device is to provide a universal device. The ultimate device will have the speed and rewritability SRAM, the density of DRAM and the persistent storage of flash. Such a device of course could lead to a flat memory hierarchy, or at least a flatter one, especially with respect to latency management. If the device also has a low process complexity, so it can be easily embedded in logic, then a whole range of interesting possibilities becomes available. For example, today, the only mainstream employment of memory in logic is the use of SRAM to implement the look-up-tables and manage the switchbox routing in FPGAs. Since the area overhead of SRAM is quite high, this limits the potential range of application of memory devices in logic systems.

The final challenge is cost. The rate of increase in the device count in DRAM with succeeding new generations has slowed due in part to device issues, but largely because device scaling has slowed and the fill factor (the % of the chip devoted to the DRAM cells) goes down with each generation (the "scaled area factor" challenge). In the DDR3 memory architecture, only roughly half the chip area is allocated to the memory cores. This is in part due to the overhead of the peripheral circuits, but is largely due to the use of DRAM area to provide faster bandwidth interfaces. Similarly, the main challenge preventing the wide spread adoption of solid state storage, in place of rotating media, is its cost premium.

OPPORTUNITIES FOR ARCHITECTURAL EXPLOITATION OF EMERGING MEMORY DEVICES

The highest priority is most probably development of energy optimized, high bandwidth memory hierarchies. Even with today's memory devices, there is a need for development of theory and models to minimize the average energy per memory access across the entire memory hierarchy from the register file to disk. Energy optimized memory hierarchies will have different organizations than latency optimized ones. Emerging research memory devices have potential to permit further reductions in energy per access, through architectural and circuit exploitation. Refer to Table ERD12. For example, the energy per access of a memory is roughly proportional to the number of rows times the number of columns. If area-efficient small sub-arrays can be built, there is potential for considerable power savings. The use of 3DIC with Through Silicon Vias (TSV) is very likely going to lead to dramatically different ways to build memories, including the low-power provisioning of high bandwidth and low latency.

The most interesting opportunity is the potential to exploit emerging research memory devices in logic. This opportunity becomes possible for a number of reasons. Several of the emerging devices are simple to fabricate, and thus could be easily included in an SOC integrated process flow. Alternatively the likely widespread adoption of 3DICs with TSVs allows the intimate mating of different technologies. Furthermore, most emerging research memory devices are non-volatile. This permits their use as programmable elements within a logic device. Numerous opportunities exist. Revisiting the architectural concepts of "processor in memory" and "logic in memory" is likely to be worthwhile. There are tremendous opportunities for incorporating more programmable logic and interconnect in devices, e.g., using such devices to create smaller, more energy efficient logic look-up tables and programmable switchboxes. Mixing these elements into Application Specific Processors could broaden the range of the latter, and blur the dividing line with FPGAs further. Embedded non-volatile memory could be used for "background" checkpointing – checkpointing without writing everything to disk, reducing the performance and energy overhead of providing reliability. Note that many of these applications require only medium-term storage. The device does not have to be optimized for ten year retention. Many investigators have looked into using Programmable memory devices used in crossbar configurations as dense configurable two-level logic, e.g., concepts in nano-cross-bars and CMOL, particularly using $4F^2$ memory cells, have been considered Inference computing (see next section) could benefit from a persistent analog memory.

Table ERD12Sample of Potential Unique Opportunities for Architectural Exploitation of Emerging
Memory Devices

As mentioned many of these possibilities are enhanced by the potential of devices with a $4F^2$ footprint, particularly resistive memories. It is important to note that to achieve a small footprint one must avoid having a transistor in the cell. However, as resistive-only arrays are not scalable, a 1D1R (resistor with a diode) cell is needed to build anything beyond tiny arrays.³²⁰ Since the resultant geometry is very regular, there is potential for faster lithography scaling than for logic. Sub-22 nm lithography is a lot easier for very regular structures, and this could be exploited with such a device. Thus these arrays might be particularly dense. They could either be built in large (slow) arrays, providing potential for a cost-effective solid-state disk, or in small (fast) arrays for energy efficient memory. Since the length of the bit and word lines are much reduced by using a $4F^2$ cell, their energy overhead, and RC delay can be can be minimized in small memory arrays. However, the overhead of the peripheral circuits might in turn be quite substantial. It will be valuable to investigate low-area overhead peripheral circuits, for example using inverters as sense amps, or using emerging research logic devices in the periphery.

RESEARCH NEEDS

Many research needs arise from the challenges and opportunities described above. Most of the needs are hopefully self-evident from that discussion, and don't warrant repetition here. However, it is important to point out that much of the research needs to be addressed in a cross-disciplinary fashion. Many of the research needs require simultaneous investigation at the device, circuit, and architectural levels.

INFERENCE COMPUTING

INTRODUCTION

Inference computation seeks to utilize prior knowledge, usually expressed in terms of conditional probabilities, to infer event outcome probabilities based on observed system events. Suppose there is a system with n states, denoted by \underline{x} , and suppose that a subset of those states, say \underline{u} , is observed. Let the remaining states in \underline{x} be denoted by \underline{y} . Then the task of the inference engine is to compute the conditional probabilities, $p(\underline{y}|\underline{u})$ over all possible values of \underline{y} so as to infer the most probable value of \underline{y} . Even if each component of the state vector can only take on a finite number of values, the computational problem of identifying the most probable values of the unobserved states, \underline{y} , is very large and explodes as the state dimension increases, or if any of the variables are continuous. Moreover, these computations often require substantial accuracy. Fortunately, not all state variables are interdependent and Bayesian networks can be encoded with conditional probabilities at each state variable node and the inference calculations can be somewhat simplified. Bayesian networks express in a graph format the structure of probabilistic relationships between several random variables where conditional dependence is encoded by the network edges and this reduces the inference computation problem somewhat but it is still exponential in the number of variables. The basic computation in Bayesian networks is probabilistic inference which is fundamental to almost all problems involved with intelligent computing.

ARCHITECTURE FOR INFERENCE

In mapping inference computations to hardware there are a number of issues to be considered, including, the type and degree of parallelism (multiple, independent threads versus data parallelism), the arithmetic precision, inter-thread communication, and local storage requirements, etc.

There are a large number of variations on basic Bayesian techniques, over a number of different fields, from communication theory and pattern recognition, to computer vision, robotics, and speech recognition. However, for this review of inference as a computational model, three general families of algorithms are considered:

- 1. Inference by Analysis;
- 2. Inference by Random Sampling, and
- 3. Inference Using Distributed Representations
- 1. Analytic Inference

While a number of techniques are used to perform inference in Bayesian Networks including Variable Elimination and Dynamic Programming, most Bayesian Networks are evaluated using variations of Bayesian Belief Propagation (BBP).³²¹ Data are input to the network by setting certain variables to known or observed values (the "evidence"). Bayesian Belief Propagation is then performed to find the probability distributions of the free variables. Analytic techniques generally require significant precision and dynamic range, generally in the form of floating point representations. Dynamic range requirements and limited parallelism make them good candidates for multi-core architectures, but not necessarily for more advanced nano-scale computation. And they typically do not scale particularly well to very large networks.

2. Random Sampling

Another approach to performing inference is by the use of random sampling techniques, most of which fall under the general category of Monte Carlo Simulations. As with analytic techniques, evidence is input by setting some nodes to known values. Then random samples of the free variables are generated. Two commonly used techniques are Adaptive Importance Sampling and Markov Chain Monte Carlo simulation. These techniques basically use adaptive sampling techniques to do a guided, adaptive search of the model's state space. For large complex Bayesian Structures, such random sampling is often the only way to evaluate the network. However, random sampling suffers from the fact that as the size of the Network increases, increasingly larger sample sets are required to obtain sufficiently accurate statistics. So it too has scaling limitations, though they tend to be less restrictive than many analytic techniques.

Monte Carlo techniques are computationally intensive and are massively parallel. In addition, arithmetic precision requirements are significantly relaxed relative to analytical techniques. Consequently sampling techniques map cleanly to simpler, massively parallel, low precision computing structures. These techniques may also benefit from morphic cores with hardware accelerated random number generation, such as by pCMOS (probabilistic CMOS)³²².

3. Distributed Data Representation Networks

Probabilistic inference is basically a computationally intensive (NP-Hard) constraint satisfaction problem that is compounded when the state space is scaled. Consequently, scaling is a key motivation for searching for more radical approaches to performing inference efficiently and in a more massively parallel form. One promising approach is to use distributed data representations (DDR). Such networks are actually a different way to structure Bayesian Networks, and although analytic and sampling techniques are possible with these structures, they also allow different kinds of massively parallel execution. Although DDR Bayesian Networks are very promising, they are also the most limited in successful demonstrations of real applications.

Computing with DDRs can be thought of as the computational equivalent of spread spectrum communication. In a distributed representation, individual nodes do not represent an entire variable or concept, but are the result of the interaction of a group of units typically configured in a network structure, and often each unit can participate in several representations. Representing data in this manner more easily allows incremental, integrative, decentralized adaptation, and the computational and communication loads are spread more evenly across the system.

DDRs also appear to be an important computational principle in neural systems. Biological neural circuits perform inference over huge knowledge structures in fractions of a second, using extremely slow, and unreliable devices. Understanding this computation is an active area of research which is developing concurrently with many new ideas in device technology.

One example of this approach is the work in Numenta.³²³ This model starts with an approximation to a general Bayesian module, which is then combined into a hierarchy to form what they call a Hierarchical Temporal Memory (HTM). Issues related to hardware architectures for Bayesian Inference over DDR structures and how they may be implemented with emerging devices are now being studied by several groups.

ARCHITECTURE

Mapping Bayesian Networks to a multi-core implementation is straightforward: just implement each node as a task, and connect them via basic inter-task message passing. A traditional symmetric multi-processing (SMP) based, multi-core machine would most certainly provide good performance. However, this approach breaks down as the system is scaled to very large networks. Among other things, Bayesian Networks tend to be storage intensive. Therefore, implementation issues such as data structure organization; memory management and cache utilization become important considerations. In fact, a potentially serious performance constraint may be access to primary memory, and unusual memory referencing patterns may significantly compromise cache memory.

One promising massively parallel approach is that of associative processing³²⁴, which has been shown to approximate Bayesian inference. This has the potential for significant levels of parallelism, and for exciting new implementations using new device technologies. Using morphic cores for heterogeneous multi-core structures, such massively parallel implementations of Bayesian networks become relevant. More study is needed to explore radical new implementation technologies, such as analog-based soft constraint satisfaction, and how they may be used to do inference.

ALGORITHM CHARACTERISTICS FOR BAYESIAN INFERENCE

Table ERD13 summarizes the hardware requirements for the three basic approaches to probabilistic inference.

Table ERD13Summary of Hardware Requirements for the Three Basic Approaches to ProbabilisticInference

PERFORMANCE LIMITS OF INFORMATION PROCESSING ARCHITECTURES

For more than three decades, improvement in integrated circuit technology, primarily due to the benefits of feature size scaling, has been a critical factor in supporting the corresponding advances in information processing system performance. As CMOS approaches scaling limits dictated by fundamental physical considerations, the question arises whether significant improvements in overall computational performance could be achieved without device scaling, e.g., by innovative architectures? This question is especially relevant in light of the worldwide effort to find an alternative technology for information processing that will continue to yield the exponential gains in performance.

A recent study explored the connection of the device physics in the *Boltzmann-Heisenberg limits* and the parameters of the digital circuits implemented from these devices.³²⁵ An abstraction for a *Minimal Turing Machine* has been offered based on limiting devices and circuits, thus the terminology *Turing-Heisenberg Rapprochement*. Recently there have been efforts to extend physical estimates for device performance limits to architectures using a bottom-up approach

based on device models via the application of statistical models and fundamental physical considerations. The elements of the approach are: (1) the determination of minimal area layouts for devices (Microsystems) and circuits (Macrosystems), (2) the development of free energy estimates, and (3) the enforcement of energy conservation constraints.

This work is based on the premise that an information processing architecture is the integration of functionally different and active components to form a system capable of computing. After the layout of the system has been determined, the techniques of thermodynamics can be used to estimate the energy needed to operate an ensemble of devices in some pre-determined manner.³²⁶ Although real architectures are vastly more complex than the simplified ensembles used to date in these studies, the approach should provide a baseline for comparison between different architectures. The information processing engine is similar to Carnot's ideal heat engine in many respects: (1) the engine efficiency sets the upper efficiency limit for any practical engine, and (2) the efficiency is determined from equilibrium thermodynamics and hence is ideal in its operation. The characterization and development of the ideal engine has been one of the critical enablers for increased application of scientific and engineering principles in the industrial revolution. The discipline of thermodynamics itself resulted from the need to increase the efficiency of utility heat engines. In addition, innovations in steam and internal combustion engines have been driven by the need to extract maximum efficiency, guided by Carnot's Law. Although it is difficult to determine such a limit for computational engines, it is believed that a similar formulation for computation would be an initial significant step forward in this direction. In addition, a global optimization in terms of energy efficiency may dictate ways of optimizing architecture and in evaluating architectural efficiency. It seems that since any core technology must be described in time and space, the proposed methodology should be extensible to estimate the possible achievable performance for information processors implemented in that technology. An example of a preliminary result is the finding that a higher number of switching levels by more Microsystems leads to fewer Macrosystem units for given system planar dimensions. This projection seems to be consistent with the trend toward multi-core systems being employed by industry today.

EMERGING MEMORY AND LOGIC DEVICES—A CRITICAL ASSESSMENT

INTRODUCTION

The purpose of this section is first to introduce a set of overall technology requirements and evaluation or relevance criteria and second, based on these criteria, to offer an assessment of the potential of each emerging research technology entry considered in this chapter to perform one of two complementary functions—1) eventually replace CMOS with a highly scalable, high performance, low power information processing device technology, or 2) provide a memory or storage technology capable of scaling either volatile and/or nonvolatile memory technology beyond the 16 nm generation. This critical review only assesses the long term potential of each emerging research device (memory and logic) technology to replace ultimately scaled CMOS logic or to provide a memory device technology of increased functional density and performance. This review does not assess the nearer-term potential of each emerging research device technology used to provide special functions to augment CMOS in, for example, an accelerator application.

Assessing the long-range potential of emerging new device and information processing technologies at best is quite difficult, yet providing such an objective and balanced judgment is important. Effective allocation of limited resources requires the semiconductor industry and the research community to consider the long-term potential and advantages offered by a new device technology compared to the projected performance of fully scaled CMOS or of related memory technologies. The intent is to provide such a technically grounded, objective benchmarking for current emerging research device technologies.

Additionally, electronic charge-based approaches will be discussed separately from those approaches proposing use of a new means for "computational state variable" and data representation. This separate discussion addresses an important question related to new charge-based information processing approaches concerning the fundamental limits of an elemental switch (size, energy, speed, etc.).

TECHNOLOGIES BEYOND CMOS

OVERALL TECHNOLOGY REQUIREMENTS AND RELEVANCE CRITERIA

[1] Scalability—First and foremost the major incentive for developing and investing in a new information processing technology is to discover and exploit a new domain for scaling information processing functional density and throughput per Joule substantially beyond that attainable by ultimately-scaled CMOS. Silicon-based CMOS has

provided several decades of scaling of MOSFET densities. The goal of a new information processing technology is to replicate this success by providing additional decades of functional and information throughput rate scaling using a new technology. In other words, it should be possible to articulate a Moore's law for the proposed technology over additional decades.

[2] Performance—Future performance metrics will be very similar to current performance metrics. They are cost, size, and speed. A future information processing technology must continue to provide (at least) incremental improvements in these parameters beyond those attainable by ultimately scaled CMOS technology. In addition, nanodevices that implement both logic and memory in the same device would revolutionize circuit and nanoarchitecture implementations.

[3] Energy Efficiency—Energy efficiency appears likely to be the limiting factor of any beyond CMOS device using electronic charge or electric current as a computational state variable. It also appears likely that it will be a dominant criterion in determining the ultimate applicability of alternate state variable devices. Clock speed versus density trade-offs for electron transport devices will dictate that for future technology generations, clock speed will need to be decreased for very high densities or conversely, density will need to be decreased for very high clock speeds. Nanoscale electron transport devices will best suit implementations that rely on the efficient use of parallel processing to minimize energy dissipation more than on fast switching.

[4A] OFF/ON or "1/0" Ratio (Memory Devices)—The OFF/ON ratio of a memory device is the ratio of the access resistance of a memory storage element in the OFF state to its access resistance in the ON state. For non-volatile memories, the OFF/ON ratio represents the ratio between leakage current of an unselected memory cell to the read current of a selected cell. In cross-point memories, a very large OFF/ON ratio is required to minimize power dissipation and maintain adequate read signal margin.

[4B] Gain (Logic Devices)—The gain of nanodevices is an important limitation for presently used combinatorial logic where gate fan-outs require significant drive current and low voltages make gates more noise sensitive. New logic and low-fan-out memory circuit approaches will be needed to use most of these nanodevices for computing applications. Signal regeneration for large circuits of nanodevices may need to be accomplished by integration with CMOS. In the near-term integratability of nanodevices with silicon CMOS is a requirement due to the need for signal restoration for many logic implementations and to be compatible with the established technology and market base. This integration will be necessary at all levels from design tools and circuits to process technology.

[5] Operational Reliability—Operational reliability is the ability of the memory and logic devices to operate reliably within their operational error tolerance given in their performance specifications. The error rate of all nanoscale devices and circuits is a major concern. These errors arise from the difficulty of providing highly precise dimensional control needed to fabricate the devices and also from interference from the local environment, such as spurious background charges in SETs. Large-scale and powerful error detection and correction schemes will need to be a central theme of any architecture and implementations that use nanoscale devices.

[6] Operational Temperature—Nanodevices must be able to operate close to a room temperature environment for most practical applications with sufficient tolerance for higher temperature (e.g., 100°C) operation internal to the device structure.

[7] CMOS Technological Compatibility—The semiconductor industry has been based for the last 40 years on incremental scaling of device dimensions to achieve performance gains. The principle economic benefit of such an approach is it allows the industry to fully apply previous technology investments to future products. Any alternative technology as a goal should utilize the tremendous investment in infrastructure to the highest degree possible.

[8] CMOS Architectural Compatibility—This criterion is motivated by the same set of concerns that motivate the CMOS technological compatibility, namely the ability to utilize the existing CMOS infrastructure. Architectural compatibly is defined in terms of the logic system and the data representation used by the alternative technology. CMOS utilizes Boolean logic and a binary data representation and ideally, an alternative technology would need to do so as well.

ELECTRONIC CHARGE-BASED NANOSCALE DEVICES

An important issue regarding emerging charge-based nanoelectronic switch elements is related to the fundamental limits to the scaling of these new devices, and how they compare with CMOS technology at its projected end of scaling. The 2009 ITRS projects the scaling of CMOS first level metal $\frac{1}{2}$ pitch to 8 nm by 2024. This generation represents a physical gate length for a MPU/ASIC device of ~7 nm with an average power dissipation of approximately 100W/cm² in 2022. A recent analysis³²⁷ concluded that the fundamental limit of scaling an electronic charge-based switch is only a factor of 3× smaller than the physical gate length of a CMOS MOSFET in 2024. Furthermore the density of these switches is limited by maximum allowable power dissipation of approximately

100W/cm², and not by their size. The conclusion of this work is that MOSFET technology scaled to its practical limit in terms of size and power density will asymptotically reach the theoretical limits of scaling for charge-based FET-like devices. Consequently, application of emerging charge-based logic technologies, such as 1D structures (nanowires and nanotubes), may be best suited for use as a replacement of the silicon channel in an otherwise silicon-based MOSFET technology infrastructure. Further, new charge-based devices, such as the Tunnel FET, may provide a path for substantially lowering power dissipation in a CMOS-compatible technology requiring minimum impact on the current design or fabrication infrastructure. However, use of novel charge-based switches to develop a completely new information processing technology, may not be justified to obtain a relatively modest maximum of 3× scaling in size or speed. This conclusion is particularly true since the device density is limited by power dissipation and not by the size of the binary switch.

ALTERNATE COMPUTATIONAL-STATE-VARIABLE NANOSCALE DEVICES

In this context, the term "*computational state variable*" refers to the notion of the finite state machine introduced by Turing in 1930s. The idea is that there are numerous ways to represent, manipulate, and store computational information or logic state. The earliest example of a finite state storage device was the abacus, which represents numerical data by the position of beads on a string. In this example, the computational state variable is simply a physical position, and the operator accomplishes readout by looking at the abacus. The operator's fingers physically move the beads to perform the data manipulations. Early core memories used the orientation of magnetic dipoles to store state. Similarly, paper tapes and punch cards used the presence or absence of holes to store the state of the computational variable. Several possible new computational state variables include: magnetic dipole (e.g., electron or nuclear spin state), molecular state, phase state, strongly correlated electron state, quantum qubit, photon polarization, etc. The question is: can a new computational state variable together with its physical representation be realized that will scale information processing technology additional decades in terms of functional density, speed, and power similar to that provided by CMOS over the past 40 years? This is the question addressed in this Critical Assessment of the technology entries proposed for memory and information processing applications.

POTENTIAL PERFORMANCE ASSESSMENT FOR MEMORY AND LOGIC DEVICES

The long-term potential performance is critically reviewed for each new memory and logic research device technology discussed in this chapter. Some of these technologies have been proposed to scale CMOS to the end of the current roadmap, while others have been proposed to eventually replace ultimately-scaled CMOS logic or memory device technology to sustain functional scaling beyond the current roadmap. In both cases, it is assumed in the following that all technological challenges and barriers impeding each research device from reaching its full potential have been resolved. Consequently, the objective of this analysis is to evaluate the ultimate performance projected for each research device technology compared to the performance projected for CMOS, or a related memory technology, scaled to the end of the roadmap.

The question addressed by this ERD Critical Review is different than that addressed by the Emerging Research Materials (ERM) Critical Review of materials and fabrication process challenges in the ERM chapter. The question addressed in the ERM chapter's Critical Review is: to what extent will research materials and fabrication processes eventually be developed to attain the projected properties of a research device? Consequently, the outcome of the ERM and ERD Critical Reviews for related technologies in some instances may be somewhat different.

For example, consider application of III-V compound semiconductors used as a channel replacement material in an otherwise silicon MOSFET structure. The ERD Critical Review of this technology, conducted by several members of the ERD working group, concludes that a III-V (e.g., InGaAs) alternate n-channel MOSFET has a significant performance advantage compared to the ultimately-scaled, end-of-the-roadmap silicon MOSFET. This assessment assumes the required material properties (e.g., electron mobility, electron velocity, dielectric/semiconductor interface state density, etc.) of the III-V channel are attained. Conversely, the ERM Critical Review considers the likelihood of realizing these required III-V compound semiconductor material properties needed to attain a high performance, III-V MOSFET. This ERM Review, performed by several ERM working group members, concludes that III-V materials will not likely yield material properties sufficient to obtain III-V MOSFET performance superior to that projected for the ultimately scaled silicon MOSFET.

METHODOLOGY

Some of the nanoscale devices discussed in this chapter are charge-based structures proposed to extend CMOS to the end of the current roadmap. Other nanoscale devices proposed as candidates to provide new means of information processing in the "Beyond CMOS-scaling" domain offer new computational state variables and will likely require new fabrication technologies. A set of relevance or evaluation criteria, defined above in the section entitled "Overall Technology Requirements and Relevance Criteria," are used to parameterize the extent to which proposed "CMOS Extension" and "Beyond CMOS" technologies are applicable to information processing applications. The Relevance

Criteria are: 1) Scalability, 2) Performance, 3) Energy Efficiency, 4) Gain (Logic) or OFF/ON Ratio (Memory), 5) Operational Reliability, 6) Operational Temperature, 7) CMOS Technological Compatibility, and 8) CMOS Architectural Compatibility.

Each CMOS extension and beyond-CMOS-scaling emerging research nanoscale memory and logic device technology is evaluated against each Relevance Criterion according to a single factor. For logic, this factor relates to the *projected potential performance* of a nanoscale device technology, assuming its successful development to maturity, *compared to that for silicon CMOS scaled to the end of the Roadmap*. For memory, this factor relates the *projected potential performance* of each nanoscale memory device technology, assuming its successful development to maturity, *compared to that for ultimately scaled current silicon memory technology which the new memory would displace*. Performance potential for each criterion is assigned a value from 1–3, with "3" substantially exceeding ultimately-scaled CMOS, and "1" substantially inferior to CMOS or, again, a comparable existing memory technology. This evaluation is determined by a survey of the ERD Working Group members composed of individuals representing a broad range of technical backgrounds and expertise.

Logic—Individual Potential for Emerging Research Logic Devices Related to each Technology Relevance Criterion

	Substantially exceeds ultimately scaled CMOS digital information processing technology (Relevance Criteria $1-5$)
	6) or is compatible with CMOS operating temperature
3	7) or is monolithically integrable with CMOS wafer technology
	8) or is compatible with CMOS wafer technology
	(i.e., Substantially Better than Silicon CMOS Digital Information Processing Technology)
	Comparable to ultimately scaled CMOS digital information processing technology
	(Relevance Criteria 1 – 5)
	6) or requires a very aggressive forced air cooling technology
2	7) or is functionally integrable (easily) with CMOS wafer technology
	8) or can be integrated with CMOS architecture with some difficulty
	(i.e., Comparable to Silicon CMOS Digital Information Processing Technology)
	Substantially (2×) inferior to ultimately scaled CMOS digital information processing technology.(Relevance Criteria 1 – 5)
	6) or requires very aggressive liquid cooling technology
1	7) or is not integrable with CMOS wafer technology
	8) or can not be integrated with CMOS architecture
	(i.e., Substantially Worse than Silicon CMOS Digital Information Processing Technology)

Memory—Individual Potential for Emerging Research Memory Devic	ces
Related to each Technology Relevance Criterion	

3	Substantially exceeds the appropriate ultimately scaled Baseline Memory Technology (Relevance Criteria $1-5$)	
	6) or is compatible with CMOS operating temperature	
	7) or is monolithically integrable with CMOS wafer technology	
	8) or is compatible with CMOS wafer technology	
	(i.e., Substantially Better than ultimately scaled Silicon Baseline Memory Technology)	
	Comparable to the appropriate ultimately scaled Baseline Memory Technology (Relevance Criteria $1 - 5$)	
2	6) or requires a very aggressive forced air cooling technology	
	7) or is functionally integrable (easily) with CMOS wafer technology	
	8) or can be integrated with CMOS architecture with some difficulty	
	(i.e., Comparable to Silicon ultimately scaled n Baseline Memory Technology)	
	Substantially (2×) inferior to the appropriate ultimately scaled Baseline Memory Technology (Relevance Criteria $1-5$)	
1	6) or requires very aggressive liquid cooling technology	
	7) or is not integrable with CMOS wafer technology	
	8) or can not be integrated with CMOS architecture	
	(i.e., Substantially Worse than ultimately scaled Silicon Baseline Memory Technology)	
Overall Potential Assessment (OPA) = Potential Summed over the Eight		

Relevance Criteria for each Technology Entry

Maximum Overall Potential Assessment (OPA) = 24

Minimum Overall Potential Assessment (OPA) = 8

Overall Potential Assessment for Technology Entries

Potential for the Technology Entry is projected to be significantly better than silicon CMOS or baseline memory (compared using the Technology Relevance Criteria) $(OPA \ge 20)$	Potential
Potential for the Technology Entry is projected to be slightly better than silicon CMOS or baseline memory (compared using the Technology Relevance Criteria) (OPA >16–20)	Potential
Potential for the Technology Entry is projected to be significantly less than silicon CMOS or baseline memory (compared using the Technology Relevance Criteria) (OPA < 16)	Potential

RESULTS

Tables ERD14 and ERD15 summarize the results of the ERD critical review. The color scale is defined in the table above entitled "Overall Potential Assessment for Technology Entries." The color represents the overall assessment for each emerging research memory and logic technology. White indicates the ERD Working Group's judgment of a relatively high potential for a fully matured research device technology to excel compared to ultimately-scaled CMOS for logic or compared to the current memory technology to be replaced. Conversely, magenta indicates a relatively low potential. Green provides additional granularity from a moderately high potential to a lower assessment for potential. On a scale of 1–3, the numbers given in each box are the average of the responses for that technology/relevance criterion received from members of the ERD Working Group. The error bars indicate the average response \pm the standard deviation.

Table ERD14 Potential Evaluation for Emerging Research Memory Devices

Table ERD15Potential Evaluation for Emerging Research Logic and Alternate Information Processing
Devices

Assignment of the relative ratings for each Technology Entry for memory and for logic is the collective judgment of the ERD Working Group and is intended to be a guideline, but not proscriptive. These ratings taken together with the numerical tables and descriptive text are intended to provide the reader with ERD Working Group's perspective on each Technology Entry following two years of conducting several workshops, reviewing the literature, and engaging in lively discussions within the Working Group. This evaluation is illustrated in further detail for each Memory Technology in Figures ERD3a through ERD3h and for each Logic Technology in Figures ERD4a through ERD4r.

The results displayed in the Memory and Logic Critical Review *Tables ERD14* and *ERD15* are interpreted briefly below beginning with the highest rated technology entry and ending with the lowest rated entry. The discussions are centered on the Technology Entries having a higher potential.

Emerging Research Memory Technologies

Each of the emerging research memory technology entries, assessed in Table ERD14, has some potential for filling the role of becoming a universal memory, but each also has one or more gating challenges to fulfilling their potential. (Detailed discussion of these gating challenges is found in the section on Emerging Research Memory technologies.)

One example of this, the spin torque transfer magnetic RAM (SSTT-MRAM), is evaluated to have a relatively good potential as a next-generation memory device combining the advantages of the high speed of SRAM, the scalability of DRAM, and the non volatility of FLASH. However, the OFF/ON ratio is of some concern, since an increasingly large OFF/ON ratio is required as the bit line voltage becomes lower. Also of concern is the energy dissipation per operation.

The Nanothermal Memory technology category encompasses two different technologies: the Nanowire Phase Change Memory (NW/PCM) and the Fuse/Antifuse Memory technology. This category also is seen to have relatively good potential to offer a memory technology for scaling to and beyond the 16 nm generation. Two areas of concern are the energy efficiency and the operational reliability. The energy efficiency challenge, concern for which increased this year as shown in Figure ERD3b, is related to the switching current and therefore the write energy needed to obtain the ON (set) and OFF (reset) states. The concern for operational reliability is related to the state-change mechanism of this category depending principally upon thermally activated processes. A related issue is to clarify and understand the possible role of other physical processes, e.g., electrochemical effects, in changing the resistive state of the fuse/antifuse technology. The Nanoionic Memory is another broad category of memory technology, in which the resistance switching mechanism is thought to depend principally on ionic (cation or anion) transport in either an insulator or an electrolytic ion conductor. While this category is thought to be promising, two areas of concern are raised. The more serious challenge is the Operational Reliability followed by a slight concern regarding a potentially low OFF/ON ratio. The concern with Operational Reliability is related to an inability to obtain complete reversibility of ionic drift processes exacerbated by possible thermal processes as well. As shown in Figure ERD3c, the critical evaluation of this category did not change appreciably from 2007.

Evaluated to be somewhat lower in potential, the Ferroelectric FET Memory is encumbered by concerns in four areas. The two more serious challenges are Operational Reliability and OFF/ON Ratio and two of lesser concerns are Performance and Energy Efficiency. Operational Reliability of the FeFET RAM is limited by the time dependant remnant polarization of the ferroelectric gate dielectric, particularly governed by the dielectric-semiconductor interface, which changes the threshold voltage in time. Also, the FeFET memory element may not be scalable beyond the 22nm generation. Figure ERD3d shows that evaluation for OFF/ON ratio declined in 2009, but for the other criteria was similar to that of 2007; both 2007 and 2009 evaluations were significantly lower than the initial evaluation in 2005 for several criteria.

The Electronic Effects Memory is the last emerging research memory evaluated as having some potential for becoming a viable memory technology in the long term. While this is seen to have significant scaling potential, three important challenges are identified: Performance, Operational Reliability, and ON/OFF Ratio. Evaluation of this memory category declined considerably (> 1.5 in Overall Potential Assessment or OPA) from the 2007 evaluation, as shown in Figure ERD3e.

A difficult challenge of the suspended-beam Nanomechanical Memory (NEMM) is scalabily. According to a recent study, it might be difficult to achieve low-voltage (\sim 1 V) operation for the beam length less than 50 nm. Consequently, the NEMM may not be competitive in the long term.

Neither of the last two memory technologies evaluated, Macromolecular Memory and Molecular Memory, is considered to have long range potential for high-performance computing due to low expectations for Performance, Operational Reliability, and OFF/ON Ratio.

Emerging Research Logic Technologies

The results for Emerging Research Logic and Alternative Information Processing technologies are displayed in *Table ERD 15* and Figures ERD4a–4r. While the color coding highlights eleven logic technologies as being most promising, the top seven are clustered within one Overall Potential Assessment (OPA) point (OPA = 17.6 to 18.7). These top seven, (Unconventional Geometry FETs, CNT-FETs, NW-FETs, Ge MOSFETs, Tunnel FETs, III-V MOSFETs, and GNR FETs) are all electron charge-based FET structures targeting extension of CMOS to the end of the current Roadmap. This assessment is consistent with the preliminary results of a study discussed in the Architecture Section in which electron charge-based and non-charge based devices are benchmarked using a number of CMOS logic gate configurations. The remaining four of this green-colored sub-group (IMOS, Negative Cg FET, Atomic Switch, and Spin Transistors) are clustered at a lower point (OPA =16.5 to 16.1). The last seven of the Logic group, colored magenta, are not rated to have a long range potential performance exceeding that for ultimately scaled CMOS.

The Unconventional MOSFET structures offer high device integration density and power efficiency at good speed if concerns are solved for threshold voltage adjustment and large parasitic resistances and capacitances. A detailed assessment of this category is given in Figure ERD4a.

Carbon Nanotube FETs offer high carrier mobility, high quasi-ballistic charge carrier velocity, and a tubular structure ideally suited for minimized short channel effects (i.e., abrupt turnoff of channel current), if a gate-all-around process is developed. Other challenges include the ability to obtain single wall semiconducting nanotubes, control the bandgap energy, control growth of nanotube position/direction, and control the carrier type and density. As shown in Figure ERD4b, the critical assessment of CNTs in 2009 is similar to that of 2007, with the exception of CMOS Architectural Compatibility, which increased

Nanowire FETs offer an appealing approach to scaling CMOS with attributes similar to CNT FETs and challenges related to growth and fabrication as well as parasitic resistances and capacitances. The 2009 and 2007 critical assessments for Nanowire FETs are similar, except the higher assessments this year for Architectural and Technological Compatibility, as shown in Figure ERD4c.

Semiconductor materials proposed to replace the silicon channel and source/drain in an otherwise silicon MOSFET technology have gained attention for their potential to enable performance scaling of CMOS to the end of the current Roadmap. These materials include carbon nanotubes, germanium, III-V compound semiconductors, and graphene nanoribbons. Carbon nanotubes are discussed above and graphene nanoribbons are discussed below. Germanium and III-V compound semiconductors offer similar advantages and face similar challenges. In particular, a III-V n-channel MOSFET can be integrated with a germanium p-channel MOSFET to maximize high n-channel III-V and p-channel germanium carrier mobilities and carrier velocities. Principal challenges include several material and fabrication process issues discussed in the ERM chapter. Figures ERD4d (germanium) and ERD4f (III-V compound semiconductors) indicate very similar (rather positive) assessments for these technologies from 2007 to 2009 with one exception: germanium is now viewed as being relatively compatible with silicon CMOS technology.

The Tunnel MOSFET offers an appealing concept for substantially lowering the energy dissipated in a switching device by substituting a tunneling process for a thermionic process for injecting charge carriers into the channel of a MOSFET. The major challenge is to simultaneously obtain a sharp subthreshold slope (much less than 60mV/decade) with a high on current, Ion. This is discussed in detail above in the Logic Section and in the ERM Chapter. Further the Tunnel FET may have a problem with Operational Reliability due to high sensitivity of device operation to slight variations of the tunnel structure and the resulting tunnel barrier. This assessment is illustrated in Figure ERD4e.

While graphene nanoribbon (GNR), used as a channel replacement material, offers an attractive alternative, it faces several important challenges discussed above in the Logic Section and in the ERM Chapter. Assuming solution to these several materials and process challenges (including development of a viable epitaxial growth technology), GNR may not offer sufficient device gain to be competitive. This assessment, compared to 2005 and 2007 assessments for a related 1D structure, is illustrated in Figure ERD4g.

The IMOS device offers an advantage similar to that of the Tunnel FET, namely to reduce the subthreshold slope, but faces several additional serious concerns related to scalability, performance, and operational reliability. Assessments illustrating these concerns are given in Figure ERD4h.

The negative Cg FET offers another approach to lowering the energy dissipated in switching a FET, but faces a significant challenge of identifying gate dielectrics (ferroelectrics and oxides) that can support demonstration of a negative capacitance in a FET gate stack. Another major concern is the Operational Reliability as shown in Figure ERD4i.

The Spin Transistor category represents two different device structures. One is the Spin FET and the other is the Spin MOSFET. In both instances the device complements the usual field effect behavior of a MOSFET with additional functionality of magnetoresistive devices. Consequently, the Spin Transistors may enable more complex transfer functions with fewer devices than CMOS is able to provide. Notwithstanding considerable focused research on developing these devices, none have been realized experimentally. Further, there are concerns with the potential Scalability, Gain, Operational Reliability, and CMOS technological compatibility of these devices, as illustrated in Figure ERD4k. However, the 2009 assessment of the Spin Transistor is significantly better in several criteria than the 2005 and 2007 assessments, as seen in Figure ERD4k.

The long-term potential of the last seven of the evaluated devices [i.e., NEMs Switches, Pseudospintronic (e.g., BISFET), SETs, Molecular Switches, Spin Wave Devices, Nanomagnetic (e.g., M:QCA), and Moving Domain Wall (i.e., ferromagnetic) Devices] are all viewed as being limited by important challenges related to their projected performance, gain, operational reliability, and CMOS technological compatibility, as shown in Figures ERD41 through ERD4r. First, however, many of these proposed information processing device technologies need to be realized and demonstrated experimentally to provide a more concrete and realistic basis for their critical assessment.



Figure ERD3a-d Technology Performance Evaluation for a) STT MRAM, b) Nanothermal Memory, c) Nanoionic Memory, and d) Ferroelectric Memory



Figure ERD3e-h Technology Performance Evaluation for e) Electronic Effects Memory, f) Nanomechanical Memory, g) Macromolecular Memory, and h) Molecular Memory



Figure ERD4a-d Technology Performance Evaluation for a) Unconventional Geometry MOSFETs, b) CNT MOSFETs, c) Nanowire, and d) Ge MOSFETs



Figure ERD4e-h Technology Performance Evaluation for e) Tunnel MOSFETs, f) III-V Compound Semiconductor MOSFETs, g) GNR MOSFETs, and h) I MOSFETs



Figure ERD4i-l Technology Performance Evaluation for i) Negative Cg FETs, j) Atomic Switches, k) Spin Transistors, and l) NEMS Devices.



Figure ERD4m-p Technology Performance Evaluation for m) Pseudospintronic Devices, n) SETs, o) Molecular Switches, and p) Spin Wave Devices.



Figure ERD4q-r Technology Performance Evaluation for q) Nanomagnetic Devices, and r) Moving Domain Wall Devices

FUNDAMENTAL GUIDING PRINCIPLES— "BEYOND CMOS" INFORMATION PROCESSING

INTRODUCTION

In considering the many disparate new approaches proposed to provide order of magnitude scaling of information processing beyond that attainable with ultimately scaled CMOS, the Emerging Research Devices Working Group proposes the following comprehensive set of guiding principles. We believe these "Guiding Principles" provide a useful structure for directing research on any "Beyond CMOS" information processing technology to dramatically enhance scaling of functional density and performance while simultaneously reducing the energy dissipated per functional operation. Further this new technology would need to be realizable using a highly manufacturable fabrication process.

GRAND CHALLENGES

COMPUTATIONAL STATE VARIABLE(S) OTHER THAN SOLELY ELECTRON CHARGE

These include spin, phase, multipole orientation, mechanical position, polarity, orbital symmetry, magnetic flux quanta, molecular configuration, and other quantum states. The estimated performance comparison of alternative state variable devices to ultimately scaled CMOS should be made as early in a program as possible to down-select and identify key trade-offs.

NON-THERMAL EQUILIBRIUM SYSTEMS

These are systems that are out of equilibrium with the ambient thermal environment for some period of their operation, thereby reducing the perturbations of stored information energy in the system caused by thermal interactions with the environment. The purpose is to allow lower energy computational processing while maintaining information integrity.

NOVEL ENERGY TRANSFER INTERACTIONS

These interactions would provide the interconnect function between communicating information processing elements. Energy transfer mechanisms for device interconnection could be based on short range interactions, including, for example, quantum exchange and double exchange interactions, electron hopping, Förster coupling (dipole–dipole coupling), tunneling and coherent phonons.

NANOSCALE THERMAL MANAGEMENT

This could be accomplished by manipulating lattice phonons for constructive energy transport and heat removal.

SUB-LITHOGRAPHIC MANUFACTURING PROCESS

One example of this principle is directed self-assembly of complex structures composed of nanoscale building blocks. These self-assembly approaches should address non-regular, hierarchically organized structures, be tied to specific device ideas, and be consistent with high volume manufacturing processes.

ALTERNATIVE ARCHITECTURES

In this case, architecture is the functional arrangement on a single chip of interconnected devices that includes embedded computational components. These architectures could utilize, for special purposes, novel devices other than CMOS to perform unique functions.

ENDNOTES/REFERENCES

- [1] Information processing refers to the input, transmission, storage, manipulation or processing, and output of data. The scope of the ERD Chapter is restricted to data or information manipulation, transmission, and storage.
- [2] H. Ishiwara, "Current status of ferroelectric-gate Si transistors and challenge to ferroelectric-gate CNT transistors," Curr. Appl. Phys. 9 (2009) S2-S6.
- [3] H. Ishiwara, "Current status of ferroelectric-gate Si transistors and challenge to ferroelectric-gate CNT transistors," Curr. Appl. Phys. 9 (2009) S2-S6.
- [4] H. Ishiwara, "Current status of ferroelectric-gate Si transistors and challenge to ferroelectric-gate CNT transistors," Curr. Appl. Phys. 9 (2009) S2-S6.
- [5] M. Fitsilis, Y. Mustafa, R. Waser, "Scaling the ferroelectric field effect transistor," Integrated Ferroelectrics 70: 29-44 2005.
- [6] W. Y. Fu et al. "Intrinsic memory function of carbon nanotube-based ferroelectric field-effect transistor," Nano Letters 9 (2009) 921.
- [7] L. Liao et al. "Ferroelectric transistors with nanowire channel: Toward nonvolatile memory applications," ACS Nano 3 (2009) 700.
- [8] Y. Zheng et al., "Gate-controlled nonvolatile graphene-ferroelectric memory," Appl. Phys. Lett. 94 (2009) 163505
- [9] Ishiwara, "Current status of ferroelectric-gate Si transistors and challenge to ferroelectric-gate CNT transistors," Curr. Appl. Phys. 9 (2009) S2-S6.
- [10] T. Rueckes, et al., "Carbon nanotube-Based Nonvolatile Random Access Memory for Molecular Computing," Science 289 (2000): 94-97.
- [11] J. E. Jang et al., "Nanoelectromechanical switch with low voltage drive," Appl. Phys. Lett. 93 (2008) 113105.
- [12] R.L. Badzey, et al., "A controllable nanomechanical memory element, Appl. Phys. Lett. 85 (2004): 3587-3589".
- [13] J. Andazane et al., "Two-terminal nanoelectromechanical devices based on germanium nanowires," Nano Lett. 9 (2009) 1824.
- [14] W. W. Jang et al. "NEMS switch with 30 nm-thick beam and 20 nm-thick air-gap for high density non-volatile memory applications," Solid-State Electronics 52 (2008) 1578-1583.
- [15] W. Y. Choi, T. Osabe and T-S. King Liu, "Nano-electro-mechanical nonvolatile memory (NEMory) cell design and scaling," IEEE Trans. Electron Dev. 55 (2008) 3482-3488.
- [16] G. E. Begtrup, W. Gannett, T. D. Yuzvinsky, V. H. Crespi, and A. Zettl, "Nanoscale reversible mass transport for archival memory," Nano Lett. 9 (2009) 1835-1838.
- [17] Y. Li, A. Sinitskii, and J. M. Tour, "Electronic two-terminal bistable graphitic memories," Nature Materials 7 (2008) 966-971.
- [18] M. H. Kryder and C. S. Kim, "After hard drives what comes next?" 2009 Intern. Conf.on Magnetics, Sacramento, CA, 6-7 May, 2009.
- [19] J-G. Zhu, "Magnetoresistive random access memory: The path to competiveness and scalability," Proc. IEEE 96 (2008) 1786-1797.
- [20] T. Kishi et al. "Lower-current and fast switching of a perpendicular TMR for high speed and high density spin-transfertorque MRAM," IEDM 2008, pp. 1-4.
- [21] H. Pagnia and N. Sotnik, "Bistable switching in electroformed metal-insulator-metal devices," Physica Status Sol. A, 108, 11-65 (1988).
- [22] F. A. Chudnovskii, L. L. Odynets, A. L. Pergament, and G. B. Stefanovich. "Electroforming and switching in oxides of transition metals: the role of metal-insulator transition in the switching mechanism," J. Solid State Chem., 122, 95-9 (1996).
- [23] J. C. Bruyere and B. K. Chakraverty, "Switching and Negative Resistance in Thin Films of Nickel Oxide" Appl. Phys. Lett. 16, 40 (1970).

- [24] D. C. Kim, S. Seo, S. E. Ahn, D. -S. Suh, M. J. Lee, B. -H. Park, I. K. Yoo, I. G. Baek, H. -J. Kim, E. K. Yim, J. E. Lee, S. O. Park, H. S. Kim, U-In Chung, J. T. Moon, and B. I. Ryu. Electrical observations of filamentary conductions for the resistive memory switching in NiO films. Applied Physics Letters, 88, 202102-1-3 (2006).
- [25] B. J. Choi, D. S. Jeong, S. K. Kim, C. Rohde, S. Choi, J. H. Oh, H. J. Kim, C. S. Hwang, K. Szot, R. Waser, B. Reichenberg, and S. Tiedke. Resistive switching mechanism of TiO₂ thin films grown by atomic-layer deposition. J. Appl. Phys. 98, 33715-1-10 (2005).
- [26] I. G. Baek, M. S. Lee, S. Seo, M. J. Lee, D. H. Seo, D. S. Suh, J. C. Park, S. O. Park, H. S. Kim, I. K. Yoo, I. G, and J. T. Moon. Highly scalable nonvolatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses. 2004 Intern. Electron Dev. Meeting, San Francisco, CA, USA, 13/12/2004-15/12/2004, Piscataway, NJ, USA: IEEE, 587-90 (2005).
- [27] T-N. Fang, S. Kaza, S. Haddad, A. Chen, Y.-C. Wu, Z. Lan, S. Avanzino, D. Liao, C. Gopalan, S. Choi, S. Mahdavi, M. Buynoski, Y. Lin, C. Marrian, C. Bill, M. VanBuskirk, and M. Taguchi, "Erase Mechanism for Copper Oxide Resistive Switching Memory Cells with Nickel Electrode," Tech. Dig. Int. Electron Devices Meet. 2006, pp.789-792.
- [28] K. Kinoshita et al, "Lowering the Switching Current of Resistance Random Access Memory Using a Hetero Junction Structure Consisting of Transition Metal Oxide," Jap. J. Appl. Phys. 45, L991 (2006).
- [29] D. Lee, D. J. Seong, I. Jo, F. Xiang, R. Dong, S. Oh, H. Hwang, "Resistance Switching of Copper doped MoOx Films for Nonvolatile Memory Application," Appl. Phys. Lett. 90, 122104 (2007).
- [30] K. Tsunoda, K. Kinoshita, H. Noshiro, Y. Yamazaki, T. Iizuka, Y. Ito, A. Takahashi, A. Okano, Y. Sato, T. Fukano, M. Aoki, and Y. Sugiyama, "Low Power and High Speed Switching of Ti-doped NiO ReRAM under the Unipolar Voltage Source of less than 3 V," Tech. Dig. Int. Electron Devices Meet., 2007, pp.767-770.
- [31] H. Shima, F. Takano, H. Muramatsu, H. Akinaga, Y. Tamai, I. H. Inoue, and H. Takagi, "Voltage polarity dependent low-power and high-speed resistance switching in CoO resistance random access memory with Ta electrode," Appl. Phys. Lett. 93 (2008) 113504.
- [32] Z. Wei et al., "Highly Reliable TaOx ReRAM and Direct Evidence of Redox Reaction Mechanism," Tech. Dig. Int. Electron Devices Meet., 2008, pp.293-296.
- [33] B. Yu, S. H. Ju, X. H. Sun, G. Ng, T. D. Nguyen, M. Meyyappan, and D. B. Janes," Indium selenide nanowire phasechange memory," Appl. Phys. Lett. 91 (2007) 1333119.
- [34] Y. Jung, C. Y. Yang, S. H. Lee, and R. Agarwal, "Phase-change Ge-Sb nanowires: Synthesis, memory switching, and phase-instability," Nano Letters 9 (2009) 2103-2108.
- [35] B. Yu, X. H. Sun, S. H. Ju, D. B. Janes, and M. Meyyappan, "Chalcogenide-nanowire-based phase change memory," IEEE Trans. Nanotechn. 7 (2008) 496-502.
- [36] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," Nature Materials 6 (2007) 833-840
- [37] R. Waser, R. Dittman, G. Staikov, and K. Szot, "Redox-based resistive switching memories nanoionic mechanisms, prospects, and challenges," Adv. Mat. 21 (2009) 2632-2663.
- [38] M.N. Kozicki, C. Gopalan, M. Balakrishnan, M. Park, and M. Mitkova, "Non-Volatile Memory Based on Solid Electrolytes," Proceedings of the 2004 Non-Volatile Memory Technology Symposium, 10-17 (2004).
- [39] T. Sakamoto, H. Sunamura, H. Kawuara, T. Hasegawa, T. Nakayama, and M. Aono, "Nanometer-scale switches using copper sulfide," Appl. Phys. Lett. 82 (2003) 3032.
- [40] H. Pagnia and N. Sotnik, "Bistable Switching in Electroformed Metal-Insulator-Metal Devices," Phys. Stat. Sol. (a) 108 (1988) 11.
- [41] A. Beck, J.G. Bednorz, Ch. Gerber, C. Rossel, and D. Widmer, "Reproducible Switching Effect in Thin Oxide Films for Memory Applications," Appl. Phys. Lett. 77 (2000) 139.
- [43] A. Sawa, T. Fujii, M. Kawasaki, and Y. Tokura. Interface resistance switching at a few nanometer thick perovskite manganite active layers. Applied Physics Letters, USA, 88, 232112-1-3 (2006).
- [44] T. Fujii, M. Kawasaki, A. Sawa, H. Akoh, Y. Kawazoe, and Y. Tokura. Hysteretic current-voltage characteristics and resistance switching at an epitaxial oxide Schottky junction SrRuO₃/SrTi_{0.99}Nb_{0.01}O₃. Applied Physics Letters, USA, 86, 12107-1-3 (2005).
- [45] J. G. Simmons and R. R. Verderber, "New conduction and reversible memory phenomena in thin insulating films," Proc. R. Soc. London, Ser. A 301, no 1464 (1967) 77.

- [46] C. Papagianni, Y. B. Nian, Y. Q. Wang, N. J. Wu, A. Ignatiev, Impedance study of reproducible switching memory effect, 2004 Nov-volatile Memory Symposium, 2004, Nov. 15-17, pp.125-128.
- [47] S. Tsui, Y. Q. Wang, Y. Y. Xue, C. W. Chu, Mechanism and scalability in resistive switching of metal Pr_{0.7}Ca_{0.3}MnO₃ interface, Appl. Phys. Lett 89 (2006) 123502.
- [48] A. Asamitsu, Y. Tomioka, H. Kuwahara, and Y. Tokura. Current switching of resistive states in magnetoresistive manganites. Nature, 388, 50-2 (1997).
- [49] D. S. Kim, Y. H. Kim, C. E. Lee, and Y. T. Kim. Colossal electroresistance mechanism in a Au/Pr_{0.7}Ca_{0.3}MnO₃/Pt sandwich structure: evidence for a Mott transition. Physical Review B 74, 174430-1-6 (2006).
- [50] G. I. Meijer, U. Staub, M. Janousch, S. L. Johnson, B. Delley, T. Neisius, and. Valence states of Cr and the insulatorto-metal transition in Cr-doped SrTiO3. Phys. Rev. B, 72, 155102 (2005).
- [51] R. Fors, S. I. Khartsev, and A. M. Grishin, "Giant resistance switching in metal-insulator-manganite junctions: Evidence for Mott transition," Phys. Rev. B 71 (2005) 045305.
- [53] C. Ko and S. Ramanathan, "Stability of electrical switching properties in vanadium dioxide thin films under multiple thermal cycles across the phase transition boundary," J. Appl. Phys. 104 (2008) 086105.
- [54] M. J. Rozenberg, I. H. Inoue, and M. J. Sanchez. Nonvolatile memory with multilevel switching: a basic model. Physical Review Letters, USA, 92, 178302/1-4 (2004).
- [55] C. H. Ahn et al., Electrostatic modification of novel materials, Rev. Modern Physics 78 (2006) 1185-1212.
- [56] C. Ko and S. Ramanathan, "Stability of electrical switching properties in vanadium dioxide thin films under multiple thermal cycles across the phase transition boundary," J. Appl. Phys. 104 (2008) 086105.
- [57] M. K. Niranjan, Y. Wang, S. S. Jaswal, and E. Y. Tsymbal, "Prediction of a switchable two-dimensional electron gas at ferroelectric oxide interfaces," Phys. Rev. Lett. 103 (2009) 016804.
- [58] P. W. M. Blom, R. M. Wolf, J. F. M. Cillessen, and M. P. C. M. Krijn, "Ferroelectric Schottky Diode," Phys. Rev. Lett. 73 (1994) 2107.
- [59] H. Kohlstedt, N. A. Pertsev, J. Rodriguez Contreras, and R. Waser "Theoretical current-voltage characteristics of ferroelectric tunnel junctions," Phys. Rev. B, 72 (2005) 125341.
- [60] J. Son, J. Cagnon and S. Stemmer, "Electrical properties of epitaxial SrTO3 tunnel barriers on (001) Pt/SrTiO3 substrates," Appl. Phys. Lett. 94 (2009) 062903.
- [61] V. Garcia, et al., "Giant tunnel electroresistance for non-destructive readout of ferroelectric states," Nature 460 (2009) 81-84.
- [62] J. C. Scott and L. D. Bozano, "Nonvolatile memory elements based on organic materials," Adv. Mat. 19 (2007) 1452-1463.
- [63] J. Ouyang, C. W. Chu, C. R. Szmanda, L. P. Ma, and Y. Yang, "Programmable polymer thin film and non-volatile memory device," Nature Materials v. 3, no 12 (2004) 918-922.
- [64] J. He, L. P. Ma, J. Wu and Y. Yang, "Three-terminal organic memory devices," J. Appl. Phys. v. 97 (2005) 064507.
- [65] Muller R, Naulaerts R, Billen J, et al., CuTCNQ resistive nonvolatile memories with a noble metal bottom electrode, Appl. Phys. Lett. 90 (2007) 063503.
- [66] A. Kanwal and M. Chowalla, "Stable, three layered organic memory devices from C₆₀ molecules and insulating polymers," Appl. Phys. Lett. 89 (2006) 203103.
- [67] L. P. Ma, J. Liu, and Y. Yang, "Organic electrical bistable devices and rewritable memory cells," Appl. Phys. Lett. v. 80, no. 16 (2002) 2997-2999.
- [68] L. D. Bozano, B. W. Kean, V. R. Deline, J. R. Salem, and J. C. Scott, "Mechanism for bistability in organic memory elements," Appl. Phys. Lett. v. 84, no. 4 (2004) 607-609.
- [69] N. B. Zhitenev, A. Sidorenko, D. M. Tennant and R. A. Cirelli, "Chemical modification of the electronic conducting states in polymer nanodevices," Nature Nanotechnology 2 (2007) 237-242.
- [70] V. V. Zhirnov and R. K. Cavin, "Molecular Electronics: Chemistry of molecules or physics of contacts?" Nature Materials 5 (2006) 11-12.

- [71] M. A. Reed, J. Chen, A. M. Rawlett, et al., Molecular random access memory cell, Appl. Phys. Lett. v. 78, no. 23 (2001) 3735-3737.
- [72] J. M. Tour, L. Cheng, D. P. Nackashi, Y. X. Yao, A. K. Flatt, S. K. St Angelo, T. E. Mallouk, P. D. Franzon, "NanoCell electronic memories," J. Amer. Chem. Soc., 125 (2003): 13279-13283.
- [73] C. N. Lau, D. R. Stewart, R. S. Williams, M. Bockrath, Direct observation of nanoscale switching centers in metal/molecule/metal structures, NANO LETTERS 4 (2004): 569-572.
- [74] J. E. Green, J. W. Choi, A. Boukai, Y. Bunimovich, E. Johnston-Halperin, E. Delonno, Y. Luo, B. A. Sheriff, K. Xu, Y. Shik Shin, H-R. Tseng, J. F. Stoddart, and J. R. Heath, "A 160-kilobit molecular memory electronic memory patterned at 10¹¹ bits per square centimeter," Nature 445 (2007) 414-417.
- [75] T. Pro, J. Buckley, K. Huang, A. Calborean, M, Gely, G. Delapierre, G. Ghibaudo, F. Duclairoir, J-C. Marchon, E. Jalaguier, P. Maldivi, B. De Salvo, and S. Deleonibus, IEEE Trans. Nanotechnol. 8 (2009) 204-212.
- [76] In no sense does this mean that research addressing emerging research device technologies not selected for accelerated development should be abandoned. Conversely, the ERD Working Group believes that several promising candidate technologies considered in this exercise, but not recommended for accelerated development at this time, should remain the subject of fundamental research to clarify their potential.
- [77] J. DENG, A. LIN, G. C. WAN, and H.-S. P. WONG, "Carbon Nanotube Transistor Compact Model for Circuit Design and Performance Optimization," ACM Journal on Emerging Technologies in Computing Systems, Vol. 4, Article 7, 2008.
- [78] A. Le Louarn, F. Kapche, J.-M. Bethoux, H. Happy, G. Dambrine V. Derycke, P. Chenevier, N. Izard, M. F. Goffman, and J.-P. Bourgoin, "Intrinsic current gain cutoff frequency of 30 GHz with carbon nanotube transistors," Appl. Phys. Lett. 90, 233108, (2007).
- [79] J. Svensson, Y. Tarakanov, D.S. Lee, J. M. Kinaret, Y. W. Park, and E. E. B. Campbell, "A carbon nanotube gated carbon nanotube transistor with 5 ps gate delay," Nanotechnology 19, 325201, 2008.
- [80] M. Rinkio, A. Johansson, G. S. Paraoanu, and P. Torma, "High-Speed Memory from Carbon Nanotube Field-Effect Transistors with High-K Gate Dielectric," Nano Lettters 9 643-647, 2009.
- [80a] N. Patil, A. Lin, E.R. Myer s, K. Ryu, A. Badmaev, C. Zhou, H.-S. Philip Wong, and S. Mitra, "Wafer-Scale Growth and Transfer of Aligned Single-Walled Carbon Nanotubes," IEEE Trans on Nanotechnology 4 498-504, 2009.
- [81] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A A, "Electric field effect in atomically thin carbon films," Science, Vol. 306, pp. 666–669 (2004).
- [82] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A A, "Electric field effect in atomically thin carbon films," Science, Vol. 306, pp. 666–669 (2004).
- [83] Gong Gu, Shu Nie, R. M. Feenstra, R. P. Devaty, W. J. Choyke, Winston K. Chan and Michael G. Kane, "Field effect in epitaxial graphene on a silicon carbide substrate," Applied Physics Letters, Vol. 90, 253507, (2007).
- [84] Kedzierski, J., Pei-Lan Hsu, Healey, P., Wyatt, P.W., Keast, C.L., Sprinkle, M., Berger, C., de Heer, W.A., "Epitaxial Graphene Transistors on SiC Substrates," IEEE Transactions on Electron Devices, Vol. 55 (8), pp. 2078 - 2085 (Aug. 2008).
- [85] Lemme, M.C., Echtermeyer, T.J., Baus, M., Kurz, H., "A Graphene Field-Effect Device," IEEE Electron Device Letters, Vol. 28 (4), pp. 282 - 284 (April 2007).
- [86] Seyoung Kim, Junghyo Nah, Insun Jo, Davood Shahrjerdi, Luigi Colombo, Zhen Yao, Emanuel Tutuc, and Sanjay K. Banerjee, "Realization of a high mobility dual-gated graphene field-effect transistor with Al2O3 dielectric," Applied Physics Letters, Vol. 94, 062107, (2009).
- [87] Xuebin Li, Xiaosong Wu, Michael Sprinkle, Fan Ming, Claire Berger, Raghunath Murali, Farhana Zaman, James Meindl, Walter de Heer, "Electric Field Effect in Epitaxial Graphene Devices," 2008 APS March Meeting, March 10– 14, 2008, New Orleans, Louisiana, Abstract Q29.00006.
- [88] Inanc Meric, Melinda Y. Han, Andrea F. Young, Barbaros Ozyilmaz, Philip Kim, Kenneth L. Shepard, "Current saturation in zero-bandgap, top-gated graphene field-effect transistors," Nature Nanotechnology, Vol. 3, 654 - 659 (2008).
- [89] Gong Gu, Shu Nie, R. M. Feenstra, R. P. Devaty, W. J. Choyke, Winston K. Chan and Michael G. Kane, "Field effect in epitaxial graphene on a silicon carbide substrate," Applied Physics Letters, Vol. 90, 253507, (2007).

- [90] Kedzierski, J., Pei-Lan Hsu, Healey, P., Wyatt, P.W., Keast, C.L., Sprinkle, M., Berger, C., de Heer, W.A., "Epitaxial Graphene Transistors on SiC Substrates," IEEE Transactions on Electron Devices, Vol. 55 (8), pp. 2078 - 2085 (Aug. 2008).
- [91] Yanqing Wu, Peide D. Ye, Michael A. Capano, Tian Shen, Yi Xuan, Yang Sui, Minghao Qi, James A. Cooper Jr., "Epitaxially grown graphene field-effect transistors with electron mobility exceeding 1500cm2/Vs and hole mobility exceeding 3400 cm2/Vs," ISDRS 2007, December 12-14, 2007, College Park, MD, USA.
- [92] Xiaogan Liang, Zengli Fu, and Stephen Y. Chou, "Graphene Transistors Fabricated via Transfer-Printing in Device Active-Areas on Large Wafer," Nano Letters, Vol. 7 (12), pp. 3840-3844 (2007).
- [93] Tomohiko Mori, Yoshihiro Kikuzawa, Hisato Takeuchi "N-type field-effect transistor based on a fluorinatedgraphene," Organic Electronics, Vol. 9, pp. 328–332 (2008).
- [94] Xaolin Li, Xinran Wang, Li Zhang, Sangwon Lee, Hongjie Dai, "Chemically Derived, Ultrasmooth Graphene Nanoribbon Semiconductors," Science, Vol. 319. no. 5867, pp. 1229 - 1232 (29 February 2008).
- [95] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A A, "Electric field effect in atomically thin carbon films," Science, Vol. 306, pp. 666–669 (2004).
- [96]Lemme, M.C., Echtermeyer, T.J., Baus, M., Kurz, H., "A Graphene Field-Effect Device," IEEE Electron Device Letters, Vol. 28 (4), pp. 282 284 (April 2007).
- [97] Inanc Meric, Melinda Y. Han, Andrea F. Young, Barbaros Ozyilmaz, Philip Kim, Kenneth L. Shepard, "Current saturation in zero-bandgap, top-gated graphene field-effect transistors," Nature Nanotechnology, Vol. 3, 654 - 659 (2008).
- [98]Seyoung Kim,, Junghyo Nah, Insun Jo, Davood Shahrjerdi, Luigi Colombo, Zhen Yao, Emanuel Tutuc, and Sanjay K. Banerjee, "Realization of a high mobility dual-gated graphene field-effect transistor with Al2O3 dielectric," Applied Physics Letters, Vol. 94, 062107, (2009).
- [99]Gong Gu, Shu Nie, R. M. Feenstra, R. P. Devaty, W. J. Choyke, Winston K. Chan and Michael G. Kane, "Field effect in epitaxial graphene on a silicon carbide substrate," Applied Physics Letters, Vol. 90, 253507, (2007).
- [100]Yanqing Wu, Peide D. Ye, Michael A. Capano, Tian Shen, Yi Xuan, Yang Sui, Minghao Qi, James A. Cooper Jr., "Epitaxially grown graphene field-effect transistors with electron mobility exceeding 1500cm2/Vs and hole mobility exceeding 3400 cm2/Vs," ISDRS 2007, December 12-14, 2007, College Park, MD, USA.
- [101]Kedzierski, J., Pei-Lan Hsu, Healey, P., Wyatt, P.W., Keast, C.L., Sprinkle, M., Berger, C., de Heer, W.A., "Epitaxial Graphene Transistors on SiC Substrates," IEEE Transactions on Electron Devices, Vol. 55 (8), pp. 2078 - 2085 (Aug. 2008).
- [102]Gengchiau Liang, Neophytos Neophytou, Dmitri E. Nikonov and Mark S. Lundstrom, "Performance Projections for Ballistic Graphene Nanoribbon Field-Effect Transistors," IEEE Transactions on Electron Devices, Vol. 54 (4), p. 677 (APRIL 2007).
- [103]Grassi, R., Gnudi, A., Gnani, E., Reggiani, S., Baccarani, G., "Graphene nanoribbons FETs for high-performance logic applications: Perspectives and challenges," Proceedings of the 2008 ICSICT Oct 20-23, 2008, pp. 365-368.
- [104]Claire Berger, Zhimin Song, Tianbo Li, Xuebin Li, Asmerom Y. Ogbazghi, Rui Feng, Zhenting Dai, Alexei N. Marchenkov, Edward H. Conrad, Phillip N. First, and Walt A. de Heer, "Ultrathin Epitaxial Graphite: 2D Electron Gas Properties and a Route toward Graphene-based Nanoelectronics," J. Phys. Chem. B, Vol. 108, 19912-19916 (2004).
- [105] Melinda Y. Han, Barbaros Ozyilmaz, Yuanbo Zhang, and Philip Kim, "Energy Band-Gap Engineering of Graphene Nanoribbons," Physical Review Letters, Vol. 98, 206805 (2007).
- [106]Xiaolin Li, Xinran Wang, Li Zhang, Sangwon Lee, Hongjie Dai, "Chemically Derived, Ultrasmooth Graphene Nanoribbon Semiconductors," Science, Vol. 319 (5867), pp. 1229 – 1232 (29 February 2008).
- [106a] F. Chen, J. Xia, D.K. Ferry, and N. Tao. "Dielectric screening enhanced performance in graphene FET." Nano Letters, vol. 9, pp. 2571-2574, 2009.
- [107] K.I. Bolotin, K.J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, H.L. Stormer, "Ultrahigh electron mobility in suspended graphene," Solid State Communications, Vol. 146, pp. 351–355 (2008).
- [108] M. Orlita, C. Faugeras, P. Plochocka, P. Neugebauer, G. Martinez, D. K. Maude, A.-L. Barra, M. Sprinkle, C. Berger, W. A. de Heer, and M. Potemski, "Approaching the Dirac Point in High-Mobility Multilayer Epitaxial Graphene," Physical Review Letters, Vol. 101, 267601 (2008).

- [109] D. D. D. Ma, C. S. Lee, F. C. K. Au, S. Y. Tong, and S. T. Lee, "Small-diameter silicon nanowire surfaces," Science, vol. 299, pp. 1874-1877, 21 Mar. 2003.
- [110] H. Yan and P. Yang, "Semiconductor nanowires: functional building blocks for nanotechnology," The Chemistry of Nanostructured Materials, P. Yang, ed. River Edge, NJ: World Scientific, 2004.
- [111] R. Beckman, E. Johnston-Halperin, Y. Luo, J. E. Green, and J. R. Heath, "Bridging dimensions: demultiplexing ultrahigh-density nanowire circuits," Science, vol. 310, no. 5747, pp. 465-468, 21 Oct. 2005.
- [112] Cui, L. J. Lauhon, M. S. Gudiksen, J. Wang, and C. M. Lieber, "Diameter-controlled synthesis of single-crystal silicon nanowires," Appl. Phys. Lett., vol. 78, no. 15, pp. 2214-2216, 9 April 2001.
- [113] Y. Wu and P. Yang, "Germanium nanowire growth via simple vapor transport," Chem. Mater., vol. 12, pp. 605-607, 2000.
- [114] J. Xiang, W. Lu, Y. Hu, Y. Wu, H. Yan, and C. M. Lieber, "Ge/Si nanowire heterostructures as high-performance field-effect transistors," Nature, vol. 441, pp. 489-493, 2006.
- [115] W. Lu and C. M. Lieber, "Nanoelectronics from the bottom up," Nature Materials, vol. 6, pp. 841-850, Nov. 2007.
- [116]B. Yang, K. D. Buddharaju, S. H. G. Teo, N. Singh, G. Q. Lo, and D. L. Kwong, "Vertical silicon-nanowire formation and gate-all-around MOSFET," IEEE Elect. Dev. Lett., vol. 29, no. 7, Jul. 2008.
- [117] C. Thelander, L. E. Froberg, C. Rehstedt, L. Samuelson, and L.-E. Wenersson, "Vertical enhancement-mode InAs nanowire field-effect transistor with 50-nm wrap gate," IEEE Elect. Dev. Lett., vol. 29, no. 3, p. 206, Mar. 2008.
- [118] J.-L. Autran and D. Munteanu, "Beyond silicon bulk MOS transistor: new materials, emerging structures and ultimate devices," Revue de l'Electricite et de l'Electronique, no.4, pp. 25-37, April 2007.
- [119] H. T. Ng, J. Han, T. Yamada, P. Nguyen, Y. P. Chen, and M. Meyyappan, "Single crystal nanowire vertical surroundgate field-effect transistor," Nano Letters, vol. 4, no. 7, pp. 1247-1252, 2004.
- [120] K. H. Yeo et al., "Gate-All-Around (GAA) Twin Silicon Nanowire MOSFET (TSNWFET) with 15 nm Length Gate and 4 nm Radius Nanowires," Proc. Int'l Electron Devices Meeting (IEDM '06), 11-13 Dec. 2006.
- [121] B. A. Sheriff, D. Wang, J. R. Heath, and J. N. Kurtin, "Complementary symmetry nanowire logic circuits: experimental demonstrations and in silico optimizations," ACS Nano, vol. 2, no. 9, pp. 1789-1798, 2008.
- [122] R. S. Friedman, M. C. McAlpine, D. S. Ricketts, D. Ham, and C. M. Lieber, "High-speed integrated nanowire circuits," Nature, vol. 434, no. 7037, p. 1085, Apr. 2005.
- [123] W. Lu, P. Xie, and C. M. Lieber, "Nanowire transistor performance limits and applications," IEEE Trans. Elect. Dev., vol. 55, no. 11, pp. 2859-2876, Nov. 2008.
- [124] K. C. Saraswat, C. O. Chui, D. Kim, T. Krishnamohan and A. Pethe, "High Mobility Materials and Novel Device Structures for High Performance Nanoscale MOSFETs," Tech. Dig. of IEDM, 26.2, 2006.
- [125] N. Goel, D. Heh, S. Koveshnikov, I. Ok, S. Oktyabrsky, V. Tokranov, R. Kambhampati, M. Yakimov, Y. Sun, P. Pianetta, C. K. Gaspe, M. B. Santos, J. Lee, S. Datta, P. Majhi, and W. Tsai, "Addressing The Gate Stack Challenge For High Mobility InGaAs Channels For NFETs," Tech. Dig. of IEDM, p363, 2008.
- [126] D. Varghese, Y. Xuan, Y. Q. Wu, T. Shen, P.D. Ye, and M. A. Alam, "Multi-probe Interface Characterization of InGaAs/Al2O3 MOSFET," Tech. Dig. of IEDM, p379, 2008.
- [127]H-C. Chin, X. Gong, X. Liu, Z. Lin, Y-C. Yeo, "Strained In0.53Ga0.47As n-MOSFETs: Performance Boost with insitu Doped Lattice-Mismatched Source/Drain Stressors and Interface Engineering," Symp. VLSI Tech. Dig. Tech. Papers, p244, 2009.
- [128] M. Yokoyama, T. Yasuda, H. Takagi, H. Yamada, N. Fukuhara, M. Hata, M. Sugiyama, Y. Nakano, M. Takenaka, and S. Takagi, "High Mobility metal S/D III-V-On-Insulator MOSFETs on a Si substrate using direct wafer bonding," Symp. VLSI Tech. Dig. Tech. Papers, p242, 2009.
- [129] Y. Xuan, Y. Q. "Wu, T. Shen, T. Yang, P. D. Ye, "High performance submicron inversion-type enhancement-mode InGaAs MOSFETs," Tech. Dig. of IEDM, p637, 2007.
- [130] N. Goel, D. Heh, S. Koveshnikov, I. Ok, S. Oktyabrsky, V. Tokranov, R. Kambhampati, M. Yakimov, Y. Sun, P. Pianetta, C. K. Gaspe, M. B. Santos, J. Lee, S. Datta, P. Majhi, and W. Tsai, "Addressing The Gate Stack Challenge For High Mobility InGaAs Channels For NFETs," Tech. Dig. of IEDM, p363, 2008.

- [131] Y. Xuan, Y. Q. "Wu, T. Shen, T. Yang, P. D. Ye, "High performance submicron inversion-type enhancement-mode InGaAs MOSFETs," Tech. Dig. of IEDM, p637, 2007.
- [132] (118) Y. Xuan, Y. Q. "Wu, T. Shen, T. Yang, P. D. Ye, "High performance submicron inversion-type enhancementmode InGaAs MOSFETs," Tech. Dig. of IEDM, p637, 2007.
- [133] (118) Y. Xuan, Y. Q. "Wu, T. Shen, T. Yang, P. D. Ye, "High performance submicron inversion-type enhancementmode InGaAs MOSFETs," Tech. Dig. of IEDM, p637, 2007.
- [134] H-C. Chin, X. Gong, X. Liu, Z. Lin, Y-C. Yeo, "Strained In0.53Ga0.47As n-MOSFETs: Performance Boost with insitu Doped Lattice-Mismatched Source/Drain Stressors and Interface Engineering," Symp. VLSI Tech. Dig. Tech. Papers, p244, 2009.
- [135] M. Yokoyama, T. Yasuda, H. Takagi, H. Yamada, N. Fukuhara, M. Hata, M. Sugiyama, Y. Nakano, M. Takenaka, and S. Takagi, "High Mobility metal S/D III-V-On-Insulator MOSFETs on a Si substrate using direct wafer bonding," Symp. VLSI Tech. Dig. Tech. Papers, p242, 2009.
- [136] D. Varghese, Y. Xuan, Y. Q. Wu, T. Shen, P.D. Ye, and M. A. Alam, "Multi-probe Interface Characterization of InGaAs/Al2O3 MOSFET," Tech. Dig. of IEDM, p379, 2008.
- [137] Y. Xuan, T. Shen, M. Xu, Y. Q. Wu, and P. D. Ye, "High-performance Surface Channel In-rich InGaAs MOSFETs with ALD High-k as Gate Dielectric," Tech. Dig. of IEDM, p371, 2008.
- [138] T. H. Chiang, W. C. Lee, T. D. Lin, D. Lin, K. H. Shiu, J. Kwo, W. E. Wang, W. Tsai, "Approaching Fermi Level Unpinning in Oxide-InGaAs," Tech. Dig. of IEDM, p375, 2008.
- [139] H-C. Chin, X. Gong, X. Liu, Z. Lin, Y-C. Yeo, "Strained In0.53Ga0.47As n-MOSFETs: Performance Boost with insitu Doped Lattice-Mismatched Source/Drain Stressors and Interface Engineering," Symp. VLSI Tech. Dig. Tech. Papers, p244, 2009.
- [140]D.-H. Kim and J. A. del Alamo, "Logic-Performance of 40nm InAs HEMTs," Tech. Dig. of IEDM, p629, 2007
- [141] N. Waldron, D.-H. Kim, and J. A. del Alamo, "90nm Self-aligned Enhancement-mode InGaAs HEMT for Logic Applications," Tech. Dig. of IEDM, p633, 2007.
- [142] T. Irisawa, S. Tokumitsu, T. Hattori, K. Nakagawa, S. Koh, and Y. Shiraki, "Ultrahigh room-temperature hole Hall and effective mobility in Si0.3Ge0.7/Ge/Si0.3Ge0.7 heterostructure," Applied Physics Letters, 81, p847, 2002.
- [143] T. Krishnamohan, Z. Krivokapic, K. Uchida, and Y. Nishi, "High-Mobility Ultrathin Strained Ge MOSFETs on Bulk and SOI With Low Band-to-Band Tunneling Leakage: Experiments," IEEE Trans. Electron Devices, 53, p990, 2006.
- [144] J. Mitard, C. Shea, B. DeJaeger, A. Pristera, G. Wang, M. Houssa, G. Eneman, G. Hellings, W-E. Wang, J. C. Lin, F. E. Leys, R. Loo, G. Winderickx, E. Vrancken, A. Stesmans, K. DeMeyer, M. Caymax, L. Pantisano, M. Meuris, M. Heyns, "Impact of EOT scaling down to 0.85nm on 70nm Ge-pFETs technology with STI," Symp. VLSI Tech. Dig. Tech Papers, p82, 2009.
- [145] J. Mitard, C. Shea, B. DeJaeger, A. Pristera, G. Wang, M. Houssa, G. Eneman, G. Hellings, W-E. Wang, J. C. Lin, F. E. Leys, R. Loo, G. Winderickx, E. Vrancken, A. Stesmans, K. DeMeyer, M. Caymax, L. Pantisano, M. Meuris, M. Heyns, "Impact of EOT scaling down to 0.85nm on 70nm Ge-pFETs technology with STI," Symp. VLSI Tech. Dig. Tech Papers, p82, 2009.
- [146] T. Takahashi, T. Nishimura, L. Chen, S. Sakata, K. Kita and A. Toriumi, "Proof of Ge-interfacing Concepts for Metal/High-k/Ge CMOS - Ge-intimate Material Selection and Interface Conscious Process Flow -," Tech. Dig. of IEDM, p697, 2007.
- [147] Y. Nakakita, R. Nakane, T. Sasada, H. Matsubara, M. Takenaka and S. Takagi, "Interface-controlled Self-Align Source/Drain Ge pMOSFETs Using Thermally-Oxidized GeO2 Interfacial Layers," Tech. Dig. of IEDM, p877, 2008
- [148] M. Kobayashi, T. Irisawa, B. M. Kope, Y. Sun, K. Saraswat, H.-S. P. Wong, P. Panetta, Y. Nishi, "High quality GeO2/Ge Interface Formed by SPA Radical Oxidation and Uniaxial Stress Engineering for High Performance Ge NMOSFETs," VLSI Tech. Dig. Tech Papers, p76, 2009.
- [149] T. Takahashi, T. Nishimura, L. Chen, S. Sakata, K. Kita and A. Toriumi, "Proof of Ge-interfacing Concepts for Metal/High-k/Ge CMOS - Ge-intimate Material Selection and Interface Conscious Process Flow -," Tech. Dig. of IEDM, p697, 2007.
- [150] Y. Kamata, A. Takashima, Y. Kamimuta, and T. Tezuka, "New Approach to Form EOT-Scalable Gate Stach with Strontium Germanide Interlayer for High-k/Ge MISFETs," VLSI Tech. Dig. Tech Papers, p78, 2009.

- [151] T. Takahashi, T. Nishimura, L. Chen, S. Sakata, K. Kita and A. Toriumi, "Proof of Ge-interfacing Concepts for Metal/High-k/Ge CMOS - Ge-intimate Material Selection and Interface Conscious Process Flow -," Tech. Dig. of IEDM, p697, 2007.
- [152] T. Yamamoto, Y. Yamashita, M. Harada, N. Taoka, K. Ikeda, K. Suzuki1, O. Kiso, N. Sugiyama and S. Takagi, "High Performance 60 nm Gate Length Germanium p-MOSFETs with Ni Germanide Metal Source/Drain," Tech. Dig. of IEDM, p1041, 2007.
- [153] H. Takato, K. Sunouchi, N. Okabe, A. Nitayama, K. Hieda, F. Horiguchi and F. Masuoka: IEDM Tech. Dig., pp. 222-225, 1988.
- [154] J. M. Hergenrother, D. Monroe, F. P. Klemens, A. Kornblit, G. R. Weber, W. M. Mansfield, M. R. Baker, F. H. Baumann, K. J. Bolan, J. E. Bower, N. A. Ciampa, R. A. Cirelli, J. I. Colonell, D. J. Eaglesham, J. Frackoviak, H. J. Gossmann, M. L. Green, S. J. Hillenius, C. A. King, R. N. Kleiman, W. Y.-C. Lai, J. T.-C. Lee, R. C. Liu, H. L. Maynard, M. D. Morris, S.-H. Oh, C. –S. Pai, C. S. Rafferty, J. M. Rosamilia, T. W. Sorsch and H.-H. Vuong : IEDM Tech. Dig., pp. 75-78, 1999.
- [155] D. Hisamoto, W.-C. Lee, J. Kedzierski, E. Anderson, H. Takeuchi, K. Asano, T.-J. King, J. Bokor and C. Hu: IEDM Tech. Dig., pp. 1032-1034, 1998.
- [156] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios and R. Chau: Symp. on VLSI Tech., pp. 133-134, 2003.
- [157] S. Harrison, P. Coronel, F. Leverd, R. Cerutti, R. Palla, D. Delille, S. Borel, S. Jullian, R. Pantel, S. Descombes, D. Dutartre, Y. Morand, MP. Samson, D. Lenoble, A. Talbot, A. Villaret, S. Monfray, P. Mazoyer, J. Bustos, H. Brut, A. Cros, D. Munteanu, J-L.Autran and T. Skotnicki: IEDM Tech. Dig., pp. 449-452, 2003.
- [158] S.-Y. Lee, E-J. Yoon, S.-M. Kim, C. W. Oh, M. Li, J.-D. Choi, K.-H. Yeo, M.-S. Kim, H.-J. Cho, S.-H. Kim, D.-W. Kim, D. Park and K. Kim: Symp. on VLSI Tech., pp.200-201, 2004.
- [159] G. Knoblinger, C. Pacha, F. Kuttner, A. Marshall, C. Russ, P. Haibach, P. Patruno, T. Schulz, K.v. Arnim, J. P. Engelstaedter, L. Bertolissi, W. Xiong, C. R. Cleavelin and K. Schruefer: ESSDERC Tech. Dig., pp.65-68, 2006.
- [160] [8] E. J. Nowak, B. A. Rainey, D. M. Fried, J. Kedzierski, M. Ieong, W. Leipold, J. Wright and M. Breitwisch: IEDM Tech. Dig., pp. 411-414, 2002.
- [161] [10] M. Yoshida, J. Kahng, C. Lee, S.-M. Jang, H. Sung, K. Kim, H.-J. Kim, K.-H. Jung, W. Yang, D. Park and B.-I. Ryu: Symp. on VLSI Tech., pp.34-35, 2006.
- [162] S.-K. Sung, S.-H. Lee, B. Y. Choi, J. J. Lee, J.-D. Choe, E. S. Cho, Y. J. Ahn, D. Choi, C.-H, Lee, D. H. Kim, Y.-S. Lee, S. B. Kim, D. Park and B.-I. Ryu: Symp. on VLSI Tech., pp. 106-107, 2006.
- [163] Y. J. Ahn, J.-D. Choe, J. J. Lee, D. Choi, E. S. Cho, B. Y. Choi, S.-H. Lee, S.-K. Sung, C.-H. Lee, S. H. Cheong, D. K. Lee, S. B. Kim, D. Park and B.-I. Ryu: Symp. on VLSI Tech., pp. 108-109, 2006.
- [164]M. Yoshida, J. Kahng, C. Lee, S.-M. Jang, H. Sung, K. Kim, H.-J. Kim, K.-H. Jung, W. Yang, D. Park and B.-I. Ryu: Symp. on VLSI Tech., pp.34-35, 2006.
- [165] A. Dixit, K. G. Anil, E. Baravelli, P. Roussel, A. Mercha, C. Gustin, M. Bamal, E. Grossar, R. Rooyackers, E. Augendre, M. Jurczak, S. Biesemans, K. DeMeyer: IEDM Tech. Dig., pp. 709-712, 2006.
- [166] A. V-Y Thean, Z-H. Shi, L. Mathew, T. Stephens, H. Desjardin, C. Parker, T. White, M. Stoker, L. Prabhu, R. Garcia, B-Y. Neugen, S. Murphy, R. Rai, J. Conner, B. E. White, S. Venkatesan: IEDM Tech. Dig. pp. 881-884, 2006.
- [167]S. Inaba, H, Kawasaki, K, Okano, T. Izumida, A. Yagishita, A. Kaneko, K. Ishimaru, N. Aoki and Y. Toyoshima: IEDM Tech. Dig., pp.487-490, 2007.
- [168] M. Guillorn, J. Chang, A. Bryant. M. Fuller, O. Dokumaci, X. Wang, J. Newbury, K. Babich, J. Ott, B. Haran, R. Yu, C. lavoie, D. Klaus, Y. Zhang, E. Sikorski, W. Graham, B. To, M. Lofaro, J. Tornello, D. Koli, B. Yang, A. Pyzyna, D. Neumeyer, M. Khater, A. Yagishita, H. Kawasaki and W. Haensch: Symp. on VLSI Tech., pp. 12-13, 2008.
- [169] A. Kaneko, A. Yagishita, K. Yahashi, T. Kubota, K. Matsuo, I. Mizushima, K. Okano, H. Kawasaki, T. Izumida, T. Kanemura, N. Aoki, A. Kinoshita, J. Koga, S. Inaba, K. Ishimaru, Y. Toyoshima, H. Ishiuchi, K. Suguro, K. Eguchi, Y. Tsunashima: IEDM Tech. Dig. pp. 893-896, 2006.
- [170] H. Gossner, C. Russ, F. Siegelin, J. Schneider, K. Schruefer, T. Schulz, C. Duvvury, C. R. Cleavelin and W. Xiong: IEDM Tech. Dig., pp. 101-104, 2006.

- [171] 1.J. Quinn, G. Kawamoto, B. McCombe, "Subband Spectroscopy by Surface Channel Tunneling," Surface Sci, vol. 73, 1978, pp. 190-196.
- [172]2.T. Baba, "Proposal for Surface Tunnel Transistors," Jpn. J. Appl. Phys., vol. 31, 1992, pp. L455-L457.
- [173] Q. Zhang, W. Zhao, and A. Seabaugh, "Low-subthreshold-swing tunnel transistors," IEEE Electron Device Lett., vol. 27, no. 4, Apr. 2006, pp. 297–300.
- [174] O.M. Nayfeh, C.N. Chleirigh, J. Hennessy, L. Gomez, J.L. Hoyt, D.A. Antoniadis, IEEE Electron Device Letters, Volume 29, Issue 9, Sept. 2008, pp. 1074 – 1077.
- [175] K. Boucart, A.M. Ionescu, "Double-Gae Tunnel FET With High-κ Gate Dielectric," IEEE Transactions on Electron Devices, Volume 54, Issue 7, July 2007, pp. 1725 – 1733.
- [176] K. K. Bhuwalka, J. Schulze, and I. Eisele, "Performance enhancement of vertical tunnel field-effect transistor with SiGe in the δp + layer," Jpn. J. Appl. Phys., vol. 43, no. 7A, Jul. 2004, pp. 4073–4078.
- [177] S.O. Koswatta, M.S. Lundstrom, D.E. Nikonov, "Performance Comparison Between p-i-n Tunneling Transistors and Conventional MOSFETs," IEEE Transactions on Electron Devices, Volume 56, Issue 3, March 2009, pp. 456 – 465.
- [178] C. Aydin, A. Zaslavsky, S. Luryi, S. Cristoloveanu, D. Mariolle, D. Fraboulet, and S. Deleonibus, "Lateral interband tunneling transistor in siliconon-insulator," Appl. Phys. Letters, vol. 84, pp. 1780-1782, 2004.
- [179] 8.S.O. Koswatta, M.S. Lundstrom, D.E. Nikonov, "Performance Comparison Between p-i-n Tunneling Transistors and Conventional MOSFETs," IEEE Transactions on Electron Devices, Volume 56, Issue 3, March 2009, pp. 456– 465.
- [180] J. Appenzeller, Y.-M. Lin, J. Knoch, and Ph. Avouris, "Band-to-Band Tunneling in Carbon Nanotube Field-Effect Transistors," Phys. Rev. Lett., vol. 93, no. 19, 2004, pp. 196805-1-4.
- [181] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. King Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," IEEE Electron Device Lett., vol. 28, no. 8, Aug. 2007, pp. 743–745.
- [182] 12.F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, S. Deleonibus, "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance," Tech. Dig. IEEE IEDM, 2008, pp. 163-166.
- [183] T. Krishnamohan, D. Kim, S. Raghunathan, and K. Saraswat, "Double-Gate Strained-Ge Heterostructure Tunneling FET (TFET) With Record High Drive Currents and 60mV/dec Subthreshold Slope," Tech. Dig. IEEE IEDM, 2008, pp. 947-949.
- [184] K. Gopalakrishnan, P.B. Griffin and J. Plummer,"I-MOS: A Novel Semiconductor Device with a Subthreshold Slope lower than kT/q," IEEE International Electron Devices Meeting, p. 289-292, December 2002.
- [185] W.Y. Choi, J.Y. Song, J.D. Lee, Y.J. Park and B.J. Park, "70-nm Impact-Ionization Metal-Oxide-Semiconductor (I-MOS) Devices Integrated with Tunneling Field-Effect Transistors (TFETs)," IEEE International Electron Devices Meeting, p. 975-978, December 2005.
- [186] W.Y. Choi, J.Y. Song, J.D. Lee, Y.J. Park and B.J. Park, "70-nm Impact-Ionization Metal-Oxide-Semiconductor (I-MOS) Devices Integrated with Tunneling Field-Effect Transistors (TFETs)," IEEE International Electron Devices Meeting, p. 975-978, December 2005.
- [187] F. Mayer, C. Le Royer, G. Le Carval, L. Clavelier and S.Deleonibus, "Static and Dynamic TCAD Analysis of IMOS Performance: From the Single Device to the Circuit," IEEE Transaction On Electron Devices, Vol.53, Issue 8, p. 1852-1857, August 2006.
- [188] W. Shockley, "Problems Related to p-n Junctions in Silicon," Solid-States Electronics, Vol. 2, N°1, p. 35-60, January 1961.
- [189] A.G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," Physical Review, Vol. 109, N°5, March 1958.
- [190] P.A. Wolff, "Theory of Electron Multiplication in Silicon and Germanium," Physical Review, Vol. 95, N°6, p. 1415-1420, September 1954.
- [191] F. Mayer, C. Le Royer, D. Blachier, L. Clavelier, and S. Deleonibus, « Avalanche Breakdown Due to 3-D Effects in the Impact-Ionization MOS (I-MOS) on SOI: Reliability Issues », IEEE Transaction on Electron Devices, Vol.55, N°6, p. 1373-1378, June 2008.

- [192] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B.Previtali, and S. Deleonibus, "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance," IEDM 2008, San Francisco, Dec 15-17, 2008. Digital Object Identifier 10.1109/IEDM.2008.4796641.
- [193] K. Gopalakrishnan, P.B. Griffin and J. Plummer,"I-MOS: A Novel Semiconductor Device with a Subthreshold Slope lower than kT/q," IEEE International Electron Devices Meeting, p. 289-292, December 2002.
- [194] W.Y. Choi, J.Y. Song, J.D. Lee, Y.J. Park and B.J. Park, "70-nm Impact-Ionization Metal-Oxide-Semiconductor (I-MOS) Devices Integrated with Tunneling Field-Effect Transistors (TFETs)," IEEE International Electron Devices Meeting, p. 975-978, December 2005.
- [195] F. Mayer, C. Le Royer, G. Le Carval, L. Clavelier C. Tabone and S. Deleonibus, "Co-integration of 2mV/dec Subthreshold Slope Impact Ionization MOS (I-MOS) with CMOS," ESSDERC, September 2006.
- [196] Charbuillet, C.; Monfray, S.; Dubois, E.; Bouillon, P.; Judong, F.; Skotnicki, T.; "High Current Drive in Ultra-Short Impact Ionization MOS (I-MOS) Devices," International Electron Device Meeting, 11-13 Dec. 2006, Digital Object Identifier 10.1109/IEDM.2006.346983.
- [197] F. Mayer, C. Le Royer, G. Le Carval, L. Clavelier C. Tabone and S. Deleonibus, "Co-integration of 2mV/dec Subthreshold Slope Impact Ionization MOS (I-MOS) with CMOS," ESSDERC, September 2006.
- [198] Charbuillet, C.; Monfray, S.; Dubois, E.; Bouillon, P.; Judong, F.; Skotnicki, T.; "High Current Drive in Ultra-Short Impact Ionization MOS (I-MOS) Devices," International Electron Device Meeting, 11-13 Dec. 2006, Digital Object Identifier 10.1109/IEDM.2006.346983.
- [199] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B.Previtali, and S. Deleonibus, "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance," IEDM 2008, San Francisco, Dec 15-17, 2008. Digital Object Identifier 10.1109/IEDM.2008.4796641.
- [200] F. Mayer, C. Le Royer, D. Blachier, L. Clavelier, and S. Deleonibus, « Avalanche Breakdown Due to 3-D Effects in the Impact-Ionization MOS (I-MOS) on SOI: Reliability Issues », IEEE Transaction on Electron Devices, Vol.55, N°6, p. 1373-1378, June 2008.
- [201] F. Mayer, C. Le Royer, J.-F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B.Previtali, and S. Deleonibus, "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance," IEDM 2008, San Francisco, Dec 15-17, 2008. Digital Object Identifier 10.1109/IEDM.2008.4796641.
- [202] S. Datta and B. Das, Appl. Phys. Lett. 56, 665 (1990).
- [203] S. Bandyopadhyay and M. Cahay, Appl. Phys. Lett. 85, 1433 (2004).
- [210] J.C. Egues, G. Burkard and D. Loss, Appl. Phys. Lett. 82, 2658 (2003).
- [205] J. Schliemann, J.C. Egues, D. Loss, Phys. Rev. Lett. 90, 146801 (2003).
- [206] S. Bandyopadhyaya, M. Cahay, Physica E 25, 399 (2005).
- [207] K. C. Hall and M. E. Flatté, Appl. Phys. Lett., 88, 162503 (2006).
- [2108] S. Bandyopadhyay and M. Cahay, http://arxiv.org, cond-mat/0604532, and cond-mat/0607659.
- [209] S. Sugahara and M. Tanaka, Appl. Phys. Lett. 84, 2307 (2004), S. Sugahara, IEE Proc. Circuits, Devices & Systems 152, 35 (2005).
- [210] B. T. Jonker, G. Kioseoglou, A. T. Hanbicki, C. H. Li, and P. E. Thompson, Nature Physics 3, 542 (2007).
- [211] O. M. J. van 't Erve, A. T. Hanbicki, M. Holub, C. H. Li, C. Awo-Affouda, P. E. Thompson, and B. T. Jonker, Appl. Phys. Lett. 91, 212109 (2007).
- [212] I. Appelbaum, B. Huang and D. J. Monsma, Nature 447, 295 (2007).
- [213] B. Huang, D. J. Monsma, and I. Appelbaum, Phys. Rev. Lett. 99, 177209 (2007).
- [214] Nanoelectronics and Information Technology. Ed. Rainer Wasser. Wiley-VCH, 2003. 425-444.
- [215] H. Inokawa, A. Fujiwara, and Y. Takahashi. "A multiple-valued logic and memory with combined single-electron and metal-oxide-semiconductor transistors." IEEE Transactions on Electron Devices 50, 462-470 (2003).
- [216] C. K. Lee, S. J. Kim, S. J. Shin, J. B. Choi, and Y. Takahashi. "Single-electron-based flexible multivalued logic gates." Appl. Phys. Lett. 92, 093101 (2008).

- [217] K. W. Song, Y. K. Lee, J. S. Sim, H. Jeoung, J. D. Lee, B. G. Park, Y. S. Jin, and Y. W. Kim. "SET/CMOS hybrid process and multiband filtering circuits." IEEE Transactions on Electron Devices 52, 1845-1850 (2005).
- [218] M. Saitoh, H. Harata, and T. Hiramoto. "Room-temperature demonstration of integrated silicon single-electron transistor circuits for current switching and analog pattern matching." IEDM Tech. Dig. 187-190 (2004).
- [219] Bandyopadhyay S. and Roychowdhury V. "Computational paradigms in nanoelectronics: Quantum coupled single electron logic and neuromorphic networks." Japan Journal of Applied Physics, Pt 1. 35.6A (1996) 3350-3362.
- [220] V. V. Zhirnov, J. A. Hutchby, G. I. Bourianoff, and J. E. Brewer. "Emerging research logic devices." IEEE Circuits and Devices 21 (3): 37-46 (2005).
- [221 M. Saitoh, H. Harata, and T. Hiramoto. "Room-temperature demonstration of integrated silicon single-electron transistor circuits for current switching and analog pattern matching." IEDM Tech. Dig. 187-190 (2004).
- [222] M. Kobayashi and T. Hiramoto, "Experimental study on quantum confinement effects in silicon nanowire metaloxide-semiconductor field-effect transistors and single-electron transistors.". J. Appl. Phys. 103, 053709 (2008).
- [223] A. Fujiwara, M. Nagase, S. Horiguchi and Y. Takahashi. "Threshold voltage of Si single-electron transistor." Jpn. J. Appl. Phys. 42, 2429-2433 (2003).
- [224] N. Asahi, M. Akazawa, Y. Amemiya. "Single-electron logic device based on the binary decision diagram." IEEE Transactions on Electron Devices 44, 1109-1116 (1997).
- [225] A. I. Csurgay, W. Porod, and C. S. Lent, "Signal processing with near-neighbor-coupled time-varying quantum-dot arrays.," IEEE Transactions on Circuits and Systems I, 47, 1212 1223 (2000).
- [226] T. Yamanaka, T. Morie, M. Nagata, and A. Iwata, "A single-electron stochastic associative processing circuit robust to random background-charge effects and its structure using nanocrystal floating-gate transistors.," Nanotechnology 11, 154-160 (2000).
- [227] K. Nishiguchi, Y. Ono, A. Fujiwara, H. Inokawa, and Y. Takahashi. "Stochastic data processing circuit based on single electrons using nanoscale field-effect transistors." Applied Physics Letters 92 062105 (2008).
- [228] M. Kobayashi and T. Hiramoto, "Experimental study on quantum confinement effects in silicon nanowire metaloxide-semiconductor field-effect transistors and single-electron transistors.". J. Appl. Phys. 103, 053709 (2008).
- [229] A. Zwanenburg, A. A. van Loon, G. A. Steele, C. E. W. M. van Rijmenam, T. Balder, Y. Fang, C. M. Lieber, and L. P. Kouwenhoven, "Ultrasmall silicon quantum dots" Journal of Applied Physics 105, 124314 (2009).
- [230] C. K. Lee, S. J. Kim, S. J. Shin, J. B. Choi, and Y. Takahashi. "Single-electron-based flexible multivalued logic gates." Appl. Phys. Lett. 92, 093101 (2008).
- [231] K. Nishiguchi, Y. Ono, A. Fujiwara, H. Inokawa, and Y. Takahashi, "Stochastic data processing circuit based on single electrons using nanoscale field-effect transistors." Applied Physics Letters 92 062105 (2008).
- [232] J. Hutchby ed., "Maturity evaluation for selected beyond CMOS emerging technologies," Workshop and ERD/ERM Working Group Meeting (July 12-13, 2008).
- [233] Fred Chen2, Hei Kam1, Dejan Markovic3, Tsu-Jae King Liu1, Vladimir Stojanovic2, Elad Alon1, "Integrated Circuit Design with NEM Relays," International Conference on Computer-Aided Design, 2008. ICCAD 2008. IEEE/ACM 10-13 Nov. 2008 Page(s):750 - 757.
- [234] Kaul et al., "Electromechanical carbon nanotube switches for high- frequency applications," Nano Lett. 6, 942-947 (2006).
- [235] Pruvost et al., "Design optimization of NEMS switches for suspended- gate single-electron transistor applications," IEEE Trans. Nanotechnology, 8, 174-184 (2009).
- [236] Yousif et al., "CMOS considerations in nanoelectromechanical carbon nanotube-based switches," Nanotechnology 19, 285204 (2008).
- [237] Jang et al., "Fabrication and characterization of a nanoelectromechanical switch with 15-nm-thick suspension air gap," Appl. Phys. Lett. 92, 103110 (2008).
- [248] Kinaret et al., "A carbon-nanotube-based nanorelay," Appl. Phys. Lett. 82, 1287-1289 (2003).
- [239] Kaul et al., "Electromechanical carbon nanotube switches for high- frequency applications," Nano Lett. 6, 942-947 (2006).

- [240] Akarvardar et al., "Energy-reversible complementary NEM logic gates," Device Research Conference, 69-70 (2008).
- [241] S. Salahuddin and S. Datta, "Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices," Nanoletters, Vol. 8, No. 2, 2008, pp. 405-410.
- [242] V. V. Zhirnov and R. K. Cavin, "Nanoelectronics: Negative capacitance to the rescue?," Nature Nanotechnology 3, 2008, pp. 77 – 78.
- [243] S. Salahuddin and S. Datta, "Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices," Nanoletters, Vol. 8, No. 2, 2008, pp. 405-410.
- [244] S. Salahuddin and S. Datta, "Can the subthreshold swing in a classical FET be lowered below 60 mV/decade?," Techn. Digest of IEDM 2008.
- [245] G.A. Salvatore, D. Bouvet, A.M. Ionescu, "Demonstration of subthrehold swing smaller than 60mV/decade in Fe-FET with P(VDF-TrFE)/SiO2 gate stack," Techn. Digest of IEDM 2008.
- [246] S. Salahuddin and S. Datta, "Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices," Nanoletters, Vol. 8, No. 2, 2008, pp. 405-410.
- [247] T. Jaglinski, D. Kochmann, D. Stone, R.S. Lakes, "Composite Materials with Viscoelastic Stiffness Greater Than Diamond," Science, 2007, 315, 620.
- [248] [F1] D. E. Nikonov, G. I. Bourianoff, and P. A. Gargini, "Simulation of Highly Idealized, Atomic Scale Magnetic Quantum Cellular Automata Logic Circuits," Journal of Nanoelectronics and Optoelectronics, v.3, no. 1, pp. 3-11 (2008).
- [249] Shoun Matsunaga, Jun Hayakawa1, Shoji Ikeda, Katsuya Miura, Haruhiro Hasegawa, Tetsuo Endoh Hideo Ohno, and Takahiro Hanyu, "Fabrication of a Nonvolatile Full Adder Based onLogic-in-Memory Architecture Using Magnetic Tunnel Junctions", Applied Physics Express 1 (2008) 091301.
- [250] Mingqiang Bao,1,a Alexander Khitun,1 Yina Wu,1 Joo-Young Lee,1 Kang L. Wang,1 and Ajey P. Jacob2, "Coplanar waveguide radio frequency ferromagnetic parametric amplifier," APPLIED PHYSICS LETTERS 93, 072509, 2008.
- [251] Nikonov DE, Bourianoff GI. "Spin gain transistor in ferromagnetic semiconductors The semiconductor Blochequations approach." IEEE Transactions On Nanotechnology 4 (2): 206-214 Mar 2005.
- [252] Jingjing Chen, Kang L. Wang, and Kosmas Galatsis. "Electrical field control magnetic phase transition in nanostructured MnxGe1-x." Applied Physics Letters 90, 012501, 2007.
- [253] Nikonov DE, Bourianoff GI. "Spin gain transistor in ferromagnetic semiconductors The semiconductor Blochequations approach." IEEE Transactions On Nanotechnology 4 (2): 206-214 Mar 2005.
- [254] Mingqiang Bao, Jooyoung Lee, Alexander Khitun, and Kang L Wang. "3-D mapping of spin wave propagation in CoFe thin film." Applied Physics Letters.
- [255] F. Q. Zhu, G.W. Chern, O. Tchernyshyov, X. C. Zhu, J. G. Zhu, and C. L. Chien. "Magnetic Bistability and Controllable Reversal of Asymmetric Ferromagnetic Nanorings." PRL 96, 027205 (2006).
- [256] X. S. Gao,1 A. O. Adeyeye,2,a and C. A. Ross3, "Magnetization reversal process in elongated Co rings with engineered defects" JOURNAL OF APPLIED PHYSICS 103, 063906 2008
- [257] Shoun Matsunaga, et. al. "Fabrication of a Nonvolatile Full Adder Based on Logic-in-Memory Architecture Using Magnetic Tunnel Junctions" Applied Physics Express, 1 (2008) 091301
- [258] R. P. Cowburn, "Digital nanomagnetic logic," Device Research Conference, 2003, pp. 111-114, 23-25 June 2003.
- [259] P. Xu, K. Xia, C. Gu, L. Tang, H. Yang, J. Li. "An all-metallic logic gate based on current-driven domain wall motion". Nature Nano, v. 3, no. 2, pp. 97-100 (2008).
- [260] M. Hayashi, L. Thomas, R. Moriya, C. Rettner, and S. S. P. Parkin. "Current-Controlled Magnetic Domain-Wall Nanowire Shift Register," Science, v. 320, no. 5873, 209-211 (2008).
- [261] M. Hayashi, L. Thomas, C. Rettner, R. Moriya, Y. B. Bazaliy, and S.P.Parkin. "Current driven domain wall velocities exceeding the spin angular momentum transfer rate in permalloy nanowires." Phys. Rev. Lett., 98, 037204 (2007).
- [262] S. S. P. Parkin, U.S. Patent No. 6,834,005 (2004).

- [263] S. S. P. Parkin, M. Hayashi, and L. Thomas. "Magnetic Domain-Wall Racetrack Memory" Science v. 320, no. 5873, p. 190-4 (2008).
- [264] S. E. Barnes, J. Ieda, and S. Maekawa. "Magnetic memory and current amplification devices using moving domain walls." Appl. Phys. Lett. 89, 122507 (2006).
- [265] R. Waser and M. Aono, 'Nanoionics-based resistive switching memories' Nature materials, 6 (2007) 833.
- [266] T. Sakamoto, H. Sunamura, H. Kawaura, T. Hasegawa, T. Nakayama and M. Aono, Appl. Phys. Lett., 'Nanometerscale switches using copper sulfide', 82 (2003) 3032.
- [267] T. Sakamoto, K. Lister, N. Banno, T. Hasegawa, K. Terabe, and M. Aono, 'Electronic transport in Ta2O5 resistive switch', Appl. Phys. Lett., 91 092110 (2007).
- [268] Z. Wang, T. Gu, T. Tada and S. Watanabe, 'Excess-silver-induced bridge formation in a silver sulfide atomic switch', Appl. Phys. Lett. 93 (2008) 152106.
- [269] M. Kundu, T. Hasegawa, K. Terabe and M. Aono, 'Effect of sulfurization condition on structural and electrical properties of copper sulfide films', J. Appl. Phys., 103 (2008) 073523.
- [270] M. Kundu, T. Hasegawa, K. Terabe and M. Aono, 'Effect of sulfurization conditions and post deposition annealing treatment on the structural and electrical properties of silver sulfide films', J. Appl. Phys., 99 (2006) 103501.
- [271] M. M. Masis, S. J. van der Molen, W. T. Fu, M. B. Hesselberth and J. M. van Ruitenbeek, 'Conductance switching in Ag2S devices fabricated by in situ sulfurization', Nanotechnol., 20 (2009) 095710.
- [272] K. Terabe, T. Hasegawa, T. Nakayama and M. Aono, 'Quantized conductance atomic switch', Nature 433 (2005) 47.
- [273] Ch. Liang, K. Terabe, T. Hasegawa, R. Negishi, T. Tamura and M. Aono, 'Ionic-Electronic Conductor Nanostructures: Template-Confined Growth and Nonlinear Electrical Transport', Small, 10 (2005) 971.
- [274] D. B. Strukov, G. S. Snider, D. R. Stewart and R. S. Williams, 'Missing Memristor found', Nature 453 (2008) 80.
- [275] J. J. Yang, M. D. Pickett, X. Li, D. A. A. Ohlberg, D. R. Stewart, and R. S. Williams, 'Memristive switching mechanism for metal/oxide/metal nanodevices', Nature Nanotech. 3 429 (2008).
- [276] N. Banno, T. Sakamoto, N. Iguchi, H. Kawaura, S. Kaeriyama, M. Mizuno, K. Terabe, T. Hasegawa, and M. Aono, 'Solid-Electrolyte Nanometer Switch', IEICE Trans. Electron., E89-C (2006) 1492.
- [277] N. Banno, T. Sakamoto, S. Fujieda and M. Aono, 'On-state reliability of solid-electrolyte switch', Proc. Int. Reliabil. Phys. Symp., Phoenix, 2008, p.707.
- [278] R. Soni, P. Meuffels, H. Kohlstedt, C. Kugeler and R. Waser, 'Reliability analysis of the low resistance state stability of Ge0.3Se0.7 based solid electrolyte nonvolatile memory cell', Appl. Phys. Lett., 94 (2009) 123503.
- [279] M. N. Kozicki M. Park and M. Mitkova, 'Nanoscale memory elements based on solid-state electrolytes', IEEE Trans. Nanotechnol., 4 (2005) 331.
- [280] T. Sakamoto, K. Lister, N. Banno, T. Hasegawa, K. Terabe, and M. Aono, 'Electronic transport in Ta2O5 resistive switch', Appl. Phys. Lett., 91 092110 (2007).
- [281] M. N. Kozicki M. Park and M. Mitkova, 'Nanoscale memory elements based on solid-state electrolytes', IEEE Trans. Nanotechnol., 4 (2005) 331.
- [282] K. Aratani, K. Ohba, T. Mizuguchi, S. Yasuda, T. Shiimoto, T. Tsushima, T. Sone, K. Endo, A. Kouchiyama, S. Saaki, A. Maesaka, N. Yamada and H. Narisawa, 'A Novel Resistance Memory with High Scalability and Nanosecond', in IEDM Tech. Digest, p. 783-786 (2007).
- [283] C. Schindler, G. Staikov and R. Waser, 'Electrode kinetics of Cu-SiO2-based resistive switching cells: Overcoming the voltage-time dilemma of electrochemical metallization memories', Appl. Phys. Lett., 94 (2009) 072109.
- [284] M. Haemori, T. Nagata and T. Chikyow, 'Impact of Cu electrode on switching behavior in a Cu/HfO2/Pt structure and resultant Cu ion diffusion', Appl. Phys. Express, 2 (2009) 061401.
- [285] Y. Naitoh, Y. Morita, M. Horikawa, H. Suga and T. Shimizu, 'Non-volatile resistive switching using silicon nanogap junction', Appl. Phys. Express, 1 (2008) 103001.
- [286] S. Kaeriyama, T. Sakamoto, H. Sunamura, M. Mizuno, H. Kawaura, T. Hasegawa, K. Terabe, T. Nakayama, and M. Aono, 'A nonvolatile programmable solid-electrolyte nanometer switch', IEEE J. Solid-State Circuits, 40 168 (2005).
- [287] A. Lozanne, Nature materials 7, 10 (2008)

- [288] J. A. Malen et al., Nano Letters 9, 1164 (2009).
- [289] H. Akkerman and B. de Boer, J. Phys.: Condens. Matter 20, 013001 (2008).
- [290] F. Thieblemont et al., Adv. Mater. 20, 3931 (2008).
- [291] P. A. van Hal et al., Nature Nanotechnology 3, 749 (2008).
- [292] M. Galperin, et al., Science 319, 1056 (2008).
- [293] C. Dri et al., Nature Nanotechnology 3, 649 (2008).
- [294] S. J. van der Molen et al., Nano Letters 9, 76 (2009).
- [295] Y. S. Zhao, et al., Adv. Mater. 20, 2859 (2008)
- [296] G. Jiang, et al., Adv. Mater. 20, 2888 (2008).
- [297] R. Ballardini, et al., Adv. Funct. Mater. 17, 740 (2007).
- [298] S. K. Banerjee, L. F. Register, E. Tutuc, D. Reddy, and A. H. MacDonald, "Bilayer PseudoSpin Field-Effect Transistor (BiSFET): A Proposed New Logic Device," in IEEE-EDL 30, 158-160 (2009).
- [299] Jung-Jung Su and A.H. MacDonald, "How to make a bilayer exciton condensate flow," Nature Physics 4, 799 802 (2008)].
- [300] I. B. Spielman, J. P. Eisenstein, L. N. Pfeiffer, K. W. West, "Resonantly enhanced tunneling in a double layer quantum Hall ferromagnet," Phys. Rev. Lett. 84, 5808 (2000).
- [301] E. Tutuc, M. Shayegan, D. A. Huse, "Counterflow measurements in strongly correlated GaAs hole bilayers: Evidence for electron-hole pairing," Phys. Rev. Lett. 93, 036802 (2004).
- [302] L. Tiemann, W. Dietsche, M. Hauser, K. von Klitzing, "Critical tunneling currents in the regime of bilayer excitons," New J. Phys. 10, 045018 (2008).
- [303] H. Min, R. Bistritzer, Jung-Jung Su, A. H. MacDonald, "Room temperature superfluidity in graphene bilayers," in Phys. Rev. B 78, 121401 (2008).
- [304] C.-H. Zhang, Y. N. Joglekar, "Excitonic condensation of massless fermions in graphene bilayers," in Phys. Rev. B 77, 233405 (2008).
- [305] S. K. Banerjee, L. F. Register, E. Tutuc, D. Reddy, and A. H. MacDonald, "Bilayer PseudoSpin Field-Effect Transistor (BiSFET): A Proposed New Logic Device," in IEEE-EDL 30, 158-160 (2009).
- [306] G. Csaba, A. Imre, G. H. Bernstein, W. Porod, and V. Metlushko, "Nanocompting by Field-Coupled Nanomagnets," IEEE Transactions on Nanotechnology 1 (4), 209-213 (2002).
- [307] A. Imre, G. Csaba, A. Orlov, G. H. Bernstein, W. Porod, and V. Metlushko, "Investigation of Shape-Dependent Switching of Coupled Nanomagnets," Superlattices and Microstructures 34, 513-518 (2003).
- [308] G. Csaba, W. Porod, and A. I. Csurgay, "A computing architecture composed of field-coupled single-domain nanomagnets clocked by magnetic fields," International Journal of Circuit Theory and Applications 31, 67-82 (2003).
- [309] A. Imre, G. Csaba, L. Ji, A. Orlov, G. H. Bernstein, and W. Porod, "Majority Logic Gate for Magnetic Quantum-Dot Cellular Automata," Science, Vol. 311, 205-208, 13 January 2006.
- [310] A. Orlov, A. Imre, L. Ji, G. Csaba, W. Porod, and G. H. Bernstein, "Magnetic Quantum-dot Cellular Automata: Recent Developments and Prospects," J. Nanoelec. Optoelec. 3 (1), 55-68 (2008).
- [311] Dmitri E. Nikonov, George I. Bourianoff, and Paolo A. Gargini, "Suitability for Digital Logic and Scaling of Atomistic Magnetic QCA," DRC, 2008.
- [312] M.T. Niemier, M.T. Alam, X.S. Hu, G.H. Bernstein, W. Porod, M. Putney, and J. DeAngelis, "Clocking Structures and Power Analysis for Nanomagnet-based Logic Devices," Proceedings of International Symposium on Low Power Electronics and Design, p. 26-31 (2007).
- [313] M.T. Niemier, M.T. Alam, X.S. Hu, G.H. Bernstein, W. Porod, M. Putney, and J. DeAngelis, "Clocking Structures and Power Analysis for Nanomagnet-based Logic Devices," Proceedings of International Symposium on Low Power Electronics and Design, p. 26-31 (2007).
- [314] G. Csaba, P. Lugli, A. Csurgay, and W. Porod, "Simulation of Power Gain and Dissipation in Field-Coupled Nanomagnets," Journal of Computational Electronics 4 (1/2), 105-110 (2005).

- [315] M.T. Niemier, M.T. Alam, X.S. Hu, G.H. Bernstein, W. Porod, M. Putney, and J. DeAngelis, "Clocking Structures and Power Analysis for Nanomagnet-based Logic Devices," Proceedings of International Symposium on Low Power Electronics and Design, p. 26-31 (2007).
- [316] D. B. Carlton, N. C. Emley NC (Emley, E. Tuchfeld, and J. Bokor, "Simulation Studies of Nanomagnet-Based Logic Architecture," Nano Letters 8, 4173-4178 (2008).
- [317] "The evolution of digital electronics towards VLSI," Keyes, R.W.; IEEE Transactions on Electron Devices, Volume 26, Issue 4, Apr 1979 Page(s):271 279.
- [318] "Will Physical Scalability Sabotage Performance Gains"? Doug Matzke, IEEE Computer, Volume 30, Issue 9, September, 1997, Page: 37 39.
- [319] "Boolean Logic and Alternative Information-Processing Devices". George Bourianoff, et.al., Computer, May 2008, pp. 38-46.
- [320] C. Amsinck, N. DiSpigna, D. Nackashi, P. Franzon, "Scaling constraints in nanoelectronic random-access memories," Nsnotechnology 16(10), Oct. 2005, pp. 2251-2260.
- [321] Pearl, J., Probabilistic Reasoning in Intelligent Systems Networks of Plausible Inference, Morgan Kaufman, 1988 & 1997.
- [322] Probabilistic Design: A Survey of Probabilistic CMOS Technology and Future Directions for Terascale IC Design, Lakshmi N. B. Chakrapani, Jason George, Bo Marr, Bilge E. S. Akgul, and Krishna V. Palem VLSI-SoC: Research Trends in VLSI and Systems on Chip, De Micheli, Giovanni; Mir, Salvador; Reis, Ricardo (Eds.), ISBN: 978-0-387-74908-2, December 2007.
- [323] Jeff Hawkins and Dileep George at Numenta, Inc., http://www.numenta.com/.
- [324] Associative Processors offer a label as output in response to an input vector that usually represents some pattern. The Associative Processor computes that stored labeled pattern that most closely approximates, in a Bayesian sense, the input vector to the machine and displays the associated label as an output. The 'knowledge base' for the machine is that set of labeled vectors that serve as a basis for its labeling decision.
- [325] V.V. Zhirnov and R.K. Cavin, "Scaling Beyond CMOS: Turing-Heisenberg Rapprochement," ESSDERC 2009, Athens, Greece.
- [326] S. Shankar, V. Zhirnov, and R. Cavin, "Computation from Devices to System-Level Thermodynamics," ESC 2009, Grenoble, France.
- [327] V. V. Zhirnov, R. K. Cavin, J. A. Hutchby, G. I. Bourianoff, "Limits to Binary Logic Scaling A Gedankin Model, Proc. IEEE, November 2003.