

Nanoelectronics Era : Novel Device Technologies Enabling Systems on Chips

Navakanta Bhat

Abstract | The issues in scaling the Complementary Metal Oxide Semiconductor (CMOS) transistors in sub-100 nm regime are reviewed. The non-classical CMOS device technologies such as high-k gate dielectrics, strained silicon channel, Silicon On Insulator, multi-gate transistors, and metal gate electrodes, are discussed in detail. These techniques are expected to scale the CMOS devices to an ultimate limit of 5 nm physical gate length. The system level issues of growing power dissipation and increasing device to device variability in the chips should be overcome for the successful realization of complex systems in nanoelectronics technologies. A brief overview of the non-CMOS memory and logic device architecture is provided. The opportunities in building hybrid systems on chips by combining sensors and actuators along with the compute and storage functions on a single chip are described.

Associate Professor,
Department of Electrical
Communication
Engineering, Indian
Institute of Science,
Bangalore-560012
navakant@ece.iisc.ernet.in

1. Introduction

The growth in electronics industry since the invention of the first Integrated Circuit (IC) chip in 1958 [1], is probably unparalleled by any other industry sector. The cumulative aggregate growth rate (CAGR) is about 15% in the last 5 decades [2]. This has been possible due to the exponential increase in the component density integrated on a single chip, as predicted by Moore's law [3,4]. An implicit assumption in the Moore's law is that we can shrink the feature size of the transistor – the basic building block of an IC – at an exponential rate. The compound annual reduction rate of the transistor dimension is more than 10%, over the last 5 decades. In 1999, the transistor gate length crossed the 100nm barrier, leading to the volume production of the *nanoelectronics* chips. Today, nanoelectronics has been the most successful commercial manifestation of the nanotechnology, with the 65 nm CMOS technology in volume production.

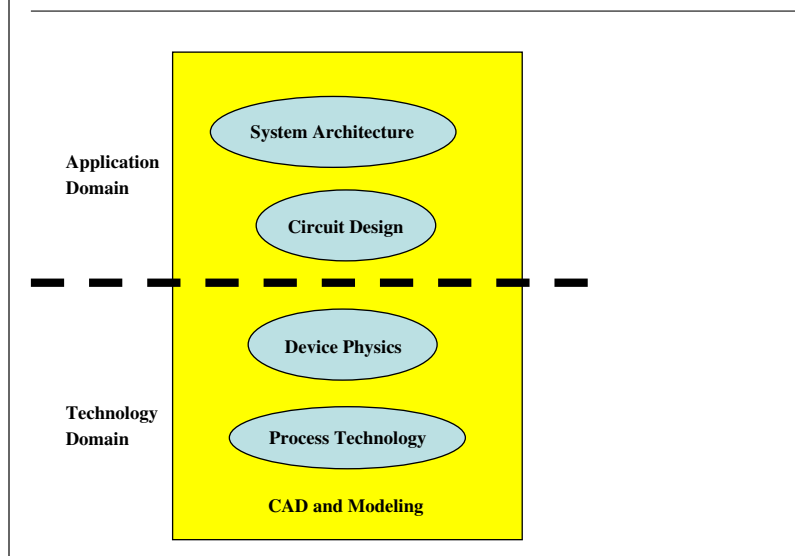
The history of silicon technology development underscores several innovations that enabled

continuous progress, overcoming supposedly insurmountable barriers. Some of these innovations include dual poly-silicon gates with n⁺ gates for *n-channel* (NMOS) and p⁺ gates for *p-channel* (PMOS) transistors, channel engineering with super steep retrograde and pocket halo implants, source/drain engineering through shallow extensions, shallow trench isolation to shrink the active device pitch, silicidation to overcome parasitic resistance, sub-wavelength optical phase shift lithography to pattern sub-100 nm features using 248 nm/193 nm deep UV source [5–7]. But to a large extent, the silicon technology has relied on the traditional scaling of classical CMOS device architecture. However in the nanoelectronics era, the technology scaling has become nontrivial due to fundamental limits imposed by device physics and materials technology. Whether and how we can go along the Moore's law curve will be dictated by the innovations in novel materials and our ability to architect the systems using non-classical device architectures. It is expected that non-classical CMOS transistors are essential in the short term and non

Keywords: nanoelectronics, non-classical CMOS, SOI, transport enhanced FETs, Multi-gate transistor, high-k gate dielectric, short channel effect, gate tunneling, non-CMOS memory and logic, MEMS, NEMS, SoC, variability, low power.

Moore's law: Moore's law is an empirical observation and extrapolation made by Gordon Moore in 1965 – *the component density on an integrated circuit doubles every year*. Although there has been an exponential growth in the component density in the last four decades, as predicted by Moore's law, the growth rate has decreased with the component density doubling every two to three years.

Figure 1: Hierarchical view of building an electronics Integrated Circuit (IC) chip.



CMOS transistors coupled with new information processing architectures will be required in the long term. We discuss some of the potential device structures in Section 2.

The success of electronics relies on two important aspects. On the one hand the building blocks of the chips, the transistors, have continued to shrink in their size. On the other hand the number of transistors reliably put together on a chip has grown exponentially. Several nomenclatures such as *VLSI-Very Large Scale Integration*, *ULSI-Ultra Large Scale Integration*, *GSI-Giga Scale Integration*, etc., are used to describe this. Today, we are able to integrate more than one billion transistors of sub-100nm size with very high yield on a silicon chip of 1cm² area. It is this capability of giga-scale integration using nano-scale building blocks that is very unique about the electronics systems. Figure 1 shows the different aspects involved in realizing the chip. It is important to recognize that the expertise required cuts across different domains, which are broadly divided into application and technology domains. The scaling of the device in nano regime requires a very good understanding of the physics of the device to engineer a high performance transistor. The process technology which can fabricate the device requires the expertise in material science, process chemistry and physics. The circuit design and system architecture requires the expertise in electronics design, computer science, communication and various other application domains. This complexity can be managed only through the Computer Aided Design (CAD) tools which require the expertise in computational science and programming. The

growth of nanoelectronics in the next two decades is hinged on innovations in all these hierarchical views of enabling an IC chip. This requires a highly interdisciplinary effort. In section 3 we briefly discuss on some of the issues in the application domain.

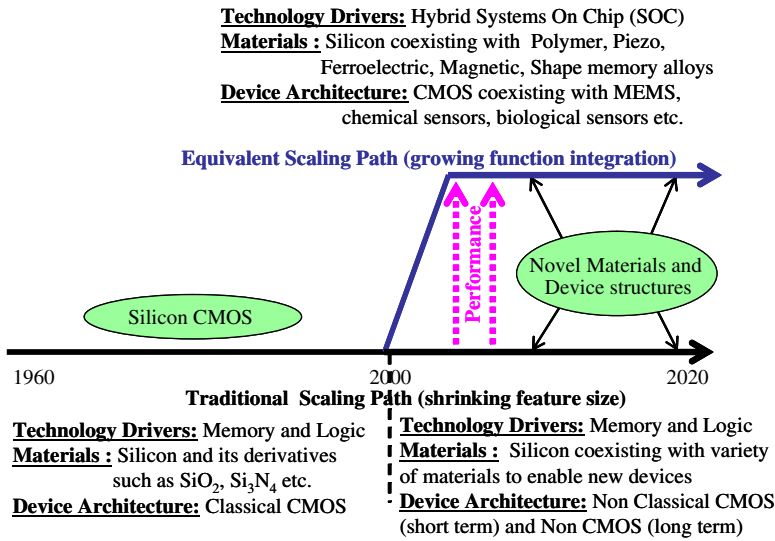
In the last few years, there has been another interesting development of building Systems on Chips (SoC), combining non-electronic energy domain components along with electronics subsystems of compute and storage elements on a single chip. The motivation for this phenomenon is illustrated in Fig. 2. The Silicon process technology lends itself to define non-electronics structures, such as micro/nano machined sensors and actuators. These are also referred to as Micro Electro Mechanical Systems (MEMS) and Nano Electro Mechanical Systems (NEMS). However, the SoCs include not only electronics and mechanical components but also optical, chemical and biological building blocks on the same chip. The nanoelectronics era is well poised for a greater confluence of multiple energy domain components on single chip, leading to penetration of such SoCs into newer application domains. We will discuss some of these developments and possibilities in Section 4.

2. Device Technology

Figure 3 shows a typical cross-section of classical NMOS transistor. The PMOS transistor is a complement of this structure with N type substrate doping and p⁺ source and drain regions. The CMOS technology utilizes both NMOS and PMOS transistors to build the circuits and systems. Both NMOS and PMOS transistors are created on bulk Silicon substrate with SiO₂ gate insulator and poly-crystalline silicon gate electrode. The MOS transistor is essentially a gate controlled switch, with the vertical electrical field controlling the channel conductivity (*field effect transistor – FET*). The gate length, L_g , of the transistor determines the switching speed of the transistor. The typical transistor dimensions on a 65 nm CMOS technology are illustrated in Fig. 3. The typical gate length is actually less than 65 nm, however the gate pitch (combination of gate length and the distance between neighboring gates) is of the order of 130 nm (2×65), and hence the label “65 nm CMOS technology”.

One of the important implications of scaling the MOS transistor in the sub-micron regime is the significant manifestation of the 2-dimensional nature of the MOS transistor. The conductivity of the MOS transistor channel, below the SiO₂ gate insulator, should be ideally controlled by the vertical

Figure 2: Traditional and Equivalent scaling paths in the Nanoelectronics Era.



Fermi level: The Fermi level is the top of the collection of electron energy levels at zero degree kelvin. In a semiconductor, the position of the Fermi level in the energy band-gap is governed by the relative concentration of electrons and holes.

electrical field set-up by the gate to source voltage (V_{gs}), and not by the lateral electric field set-up by the drain to source voltage (V_{ds}). In other words, V_{ds} should only influence the drift velocity of the carriers in the channel. However, with decreasing L_g , drain and source electrodes come into close proximity. Hence, the drain electric field starts influencing the height of the barrier at the source for the injection of free carriers into the channel. As a result, the drain electric field starts competing with the gate electrode to gain control over the channel conductivity. The transistor will no longer be an ideal gate controlled device. This phenomenon is referred to as Short Channel Effect (SCE). The immediate consequence of SCE is lowering of the threshold voltage (V_t) of the transistor with decreasing L_g as shown in Fig. 4. The short channel transistor needs to be designed properly so that V_t roll-off is controlled to yield the required value at the minimum L_g for a given technology. In a classical CMOS, this has been typically achieved by a combination of decrease in SiO₂ gate insulator thickness to increase the vertical electric field coupling to the channel, increase in substrate doping concentration to decrease the penetration of drain electric field towards source, and decrease in source/drain junction depths to reduce the coupling volume. However, these techniques will not be sufficient to scale the device in the nanoelectronics era. Several non-classical approaches are required in the immediate future to extend the scalability of CMOS architecture. Table 1 summarizes the overall projections of nanoelectronics device technology and the corresponding chip performance for the high end microprocessor.

2.1. Non classical CMOS

2.1.1. High-k gate dielectrics

Figure 5(a) shows the SiO₂ gate oxide thickness scaling trend in CMOS technology. For the 65 nm technology node, the SiO₂ thickness is about 1nm which is essentially two mono-layers of SiO₂film. Figure 5(b) shows the energy band diagram of the NMOS structure in the vertical direction for a positive applied bias. (This discussion holds good for the negative gate bias as well, except that the gate Fermi level, E_f , will be above the substrate and the electrons will be tunneling from the gate electrode). Typically, the conduction across the SiO₂ insulator layer is negligible due to very large band offset (~ 3.1 eV) between the conduction band of gate/substrate and the conduction band of the oxide. However, for an ultra-thin gate oxide with $T_{ox} < 5$ nm, carriers can quantum mechanically tunnel through the oxide and result in large gate leakage current. Furthermore, the tunneling probability increases exponentially with decrease in T_{ox} . Figure 6 shows the gate leakage current density for different gate oxide thickness. The gate leakage current density of 100 A/cm² indicates that, the transistor with 50 nm gate length results in a gate leakage of 50 nA per unit micron width (50 nA/ μ m). This is more than 20% of the total off state leakage current budget of the transistor. Hence it is not possible to scale T_{ox} any further. This calls for replacement of SiO₂ with an alternate gate insulator which has higher dielectric constant (k). The high-k dielectric of thickness of T_k is equivalent to effective oxide thickness of $EOT = 3.9/k \times T_k$, where SiO₂ dielectric constant is 3.9. Hence, it is possible to use thicker insulator and reduce the gate tunneling current (Fig. 7), while maintaining capacitive coupling of the gate electric field to the channel that is equivalent to thinner oxide. Table 2 lists the properties of some potential gate insulators that are being explored. In addition to high-k, the insulator should also have other important properties such as very good interface with Si resulting in minimal interface traps, thermodynamic stability, large band gap and band offset with Silicon. Among the various high-k dielectrics, HfO₂ and some of the rare earth metal oxides are very promising at this time [10–14]. The high-k gate dielectrics are expected to be introduced in the 45 nm CMOS technology node.

2.1.2. Transport enhanced FET

The increase in substrate doping concentration to suppress short channel effects has adversely affected the carrier mobility in the MOSFET channel. Fig. 8 shows the electron mobility in the channel of NMOS transistor for different technology

Table 1: Projected technology nodes from International Technology Roadmap for Semiconductors. The technology nodes are scaled by a factor of 0.7, corresponding to doubling of device density (except the year 2020). Note that the transistor size in any given *technology node* is significantly smaller than the *node nomenclature*.

Year of Production	2007	2010	2013	2016	2019	2020
Technology Node (nm)	65	45	32	22	16	14
Transistor Gate Length in Microprocessors circuits (nm)	25	18	13	9	6	5
Wafer diameter used in fab (inch)	12	12	18	18	18	18
Number of masks required for fabrication of Microprocessor	33	35	37	37	39	39
Number of Transistors in Microprocessor (billion)	1.1	2.2	4.4	8.8	17.7	17.7
Number of interconnect wiring levels in the Microprocessor chip	15	16	17	17	18	18
Number pins for packaged Microprocessor chip	1088	1450	1930	2568	3418	3760
Operating voltage (V)	1.1	1.0	0.9	0.8	0.7	0.7
Microprocessor frequency (GHz)	9.3	15.1	23	39.7	62.4	73.1
Chip power dissipation (Watts)	189	198	198	198	198	198

nodes [15]. The significant decrease in mobility of the scaled device is due to the increased impurity scattering in the channel, which gets worse with increased doping concentration. This necessitates a new class of non-classical CMOS transistors, namely the transport enhanced FETs. The technique relies on material engineering to enhance the channel mobility of the carriers. This class of devices includes strained-silicon MOSFETs and hetero epitaxially grown Germanium (Ge) or Gallium Arsenide (GaAs) channel MOSFETs on the silicon substrate.

The strained-silicon channel MOSFETs rely on a very well known phenomenon that the carrier mobility in silicon can be enhanced by introducing stress in the silicon lattice. The tensile strain enhances the electron mobility and the compressive strain enhances the hole mobility. Several techniques have been employed for the strain engineering in MOSFET. Figure 9 illustrates the strain engineering using Si-Ge. The inter-atomic distance in Ge (2.45 nm) is more than that of Si (2.36 nm) and the technique exploits this difference to create tensile or compressive strain. Si-Ge thin film is grown on a silicon substrate by grading the Ge concentration from 0 to about 30%. This graded layer typically contains the defects and ensures that the defect-free films can be grown on top. Then $\text{Si}_{0.7}\text{Ge}_{0.3}$ thin film is grown with an effective lattice constant higher than that of silicon. The thin silicon channel layer (about 10 nm) is epitaxially grown on $\text{Si}_{0.7}\text{Ge}_{0.3}$, thus leading to a strained silicon channel. A very important consideration in creating this structure is to ensure that the strained silicon film is defect-free, so that very high channel mobility can be obtained. A variation of this technique has been used for PMOS transistor to create compressive strain. The source/drain region is recessed into the substrate and the Si-Ge is selectively grown in the

source/drain regions. This induces a compressive stress in the channel due to the lateral confinement from the two directions of the channel, leading to enhanced hole mobility.

The other strain engineering technique relies on gate stack capping with appropriate material to induce stress in the underlying channel region. For example, capping the gate stack with Si_3N_4 can induce tensile stress and enhance the electron mobility. The isolation oxide in shallow trench region has also been utilized to induce strain in the channel region. One of the issues in utilizing the strain engineered MOSFETs in building the chips is to ensure that the NMOS and PMOS regions in the chip get different processing steps to create the respective channel regions. The strained silicon channel MOSFETs have been successfully integrated with channel mobility improvement of more than 30% [16,17]. The strained silicon MOSFETs are going through continuous improvement, and this technique will help scale the CMOS technology for the next 5 years or so. However, more radical techniques of transport enhancement will be required beyond that. Some of the potential candidates include Ge and GaAs MOSFETs on thin films grown hetero-epitaxially on the silicon substrate. The electron mobility in Ge is more than twice that of silicon and the hole mobility is about four times that of silicon. Similarly the electron mobility in GaAs is about six times that of Silicon. Historically, Ge and GaAs have never been amenable for building MOSFETs. While Si has an excellent insulator in terms of SiO_2 , which results in an ideal interface with minimal trap density, Ge and GaAs did not have this advantage. But now that SiO_2 is being replaced with alternate high-k gate dielectrics, similar technique can potentially be employed for Ge and GaAs as well. There are recent

Table 2: Comparison of potential high-k dielectrics with SiO₂

Material	Dielectric constant (k or ϵ_r)	Bandgap (eV)	Conduction band offset (eV)
SiO ₂	3.9	9	3.15
Si ₃ N ₄	7.9	5.3	2.4
Al ₂ O ₃	10	8.8	2.8
HfO ₂	16–30	4.5–6	1.5
ZrO ₂	12–16	5.8	1.5
TiO ₂	80–170	3.05	0
ReO ₂ *	12–30	2.5–5.5	1–3

*ReO₂ : Rare earth metal oxides

reports with high quality Germanium devices built using high-k gate dielectrics [18]. These techniques will potentially lead to “enhanced silicon CMOS” technologies, while continuing to exploit the silicon infrastructure.

2.1.3. Silicon On Insulator (SOI)

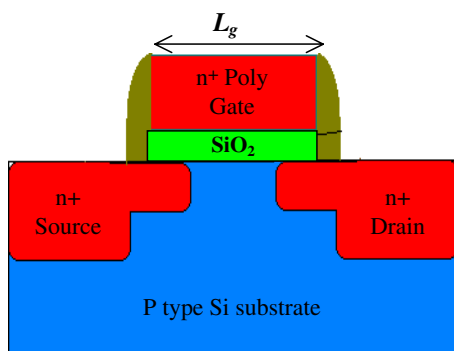
The SOI technology utilizes a very thin silicon film, on top of an insulator (typically SiO₂), to build devices. The difference between the construction of bulk Silicon devices and the SOI devices is schematically shown in Fig. 10. The insertion of SiO₂ underneath the active Silicon devices offers some unique advantages. In CMOS circuits, the propagation delay ($\tau = CV/I$) and the dynamic power ($P_d = CV^2f$) are directly proportional the load capacitance (C) which is charged and discharged to change the logic states. The load capacitance consists of three components: the gate capacitance of the driven gate (C_g), the source/drain capacitance of the driving gate (C_{ds}), and the interconnect capacitance (C_i) of the line connecting these two nodes. In the SOI technology C_{ds} is almost equal to zero, since the source/drain regions are placed directly on top of the

thick insulator. In addition, C_i is also significantly lower, since the device density in SOI is typically higher than the bulk technology. This is because the isolation between the transistors in the SOI technology is excellent because the transistors are in an island surrounded by an insulator. Hence the latch-up issue is non-existent in SOI, thus enabling very close placement of neighbouring transistors, leading to shorter interconnect lengths between any two nodes in a circuit. The SOI technology typically offers a performance improvement equivalent to one generation of bulk technology (i.e. n^{th} generation of SOI is equivalent to $(n + 1)^{th}$ generation of bulk) [19,20]. There are two variants of SOI device technologies: Partially Depleted SOI (PDSOI) and Fully Depleted SOI (FDSOI). In PDSOI, the thickness of the silicon film is typically 100 nm or more so that only part of the entire silicon body is depleted when the transistor is turned on. In contrast, the silicon film thickness for FDSOI is typically 50 nm or less so that the entire silicon body is depleted when the transistor is turned on. Each of these devices has their own advantages and disadvantages. The PDSOI devices are relatively easy to fabricate. However, they have a very unique problem of floating body induced threshold voltage (V_t) instability i.e. the V_t becomes a strong function of the charge stored in the floating body. This requires very challenging circuit design techniques because the circuit behaviour will have history dependence. On the other hand FDSOI devices have challenging fabrication process requirements due to thin Silicon body. However they do not suffer from V_t instability. In addition FDSOI devices offer a very good platform for more exotic non-classical CMOS transistors, namely the multi-gate MOSFETs.

propagation delay: The time required for a change in digital logic state at the input node of a logic gate to appear as the corresponding change in logic state at the output node.

dynamic power: The power dissipated in switching the output digital logic state of a logic gate.

Figure 3: The cross section of NMOS transistor and typical dimensions in 65 nm CMOS technology.



Poly-Si gate length, L_g	50 nm
Poly-Si gate thickness	100 nm
SiO ₂ thickness	1 nm
Source/Drain depth: extension and deep region	40 nm and 100 nm
Source/Drain doping	$\sim 10^{20}/\text{cm}^3$
Substrate doping	$\sim 10^{17}/\text{cm}^3$

2.1.4. Multi Gate FETs

The conventional technique of increasing the substrate doping concentration to suppress the short channel effect cannot be continued indefinitely. This is because the leakage at the drain/source to substrate junction due to tunneling will adversely affect the transistor's performance. An efficient technique to restore the gate control of the channel, without increasing the substrate doping, is essential. The scaling of MOSFETs has also been accompanied by the scaling of the operating voltages (V_{dd}). This is essential to ensure that the electric fields in the MOSFET are acceptable and the long term reliability

of the devices and circuits is not compromised. The scaling of the supply voltage also necessitates the scaling of the V_t , since the speed of the CMOS circuit depends on the drive current available from the transistor which is proportional to $(V_{dd} - V_t)$. This has resulted in an exponential increase in off state current of the transistor over the last few technology generations. This is because the sub-threshold conduction in a MOSFET is essentially governed by the sub-threshold slope (S) which is a non-scalable quantity [21], as shown in Fig. 11. For any given V_t the drive current can also be enhanced by preventing the mobility degradation. Hence, it is desirable to create a transistor with virtually intrinsic channel doping, while ensuring the adequate gate control of the channel. These requirements are the motivation for creating multi-gate FETs which include double-gate FET, tri-gate FET, Fin-FET and surround-gate FET [22–24]. By providing additional gates, the voltage on the gate can effectively couple to the channel thereby enhancing the FET performance. Figure 12 shows the typical structure of FinFET, wherein the gate wraps around the channel region. Fin-FETs with very good characteristics have been experimentally demonstrated. However, the manufacturable process integration has a long way to go. All these multi-gate devices require an ultra-thin silicon channel region which can be efficiently controlled.

Figure 4: V_t of NMOS decreases in the short channel regime.

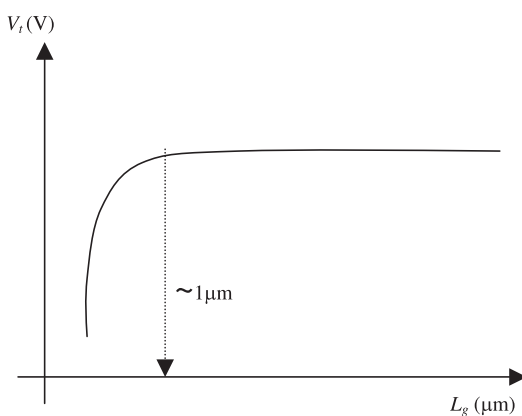


Figure 5: (a) Gate oxide thickness trend as a function of technology scaling; (b) Energy band diagram for NMOS with positive gate voltage.

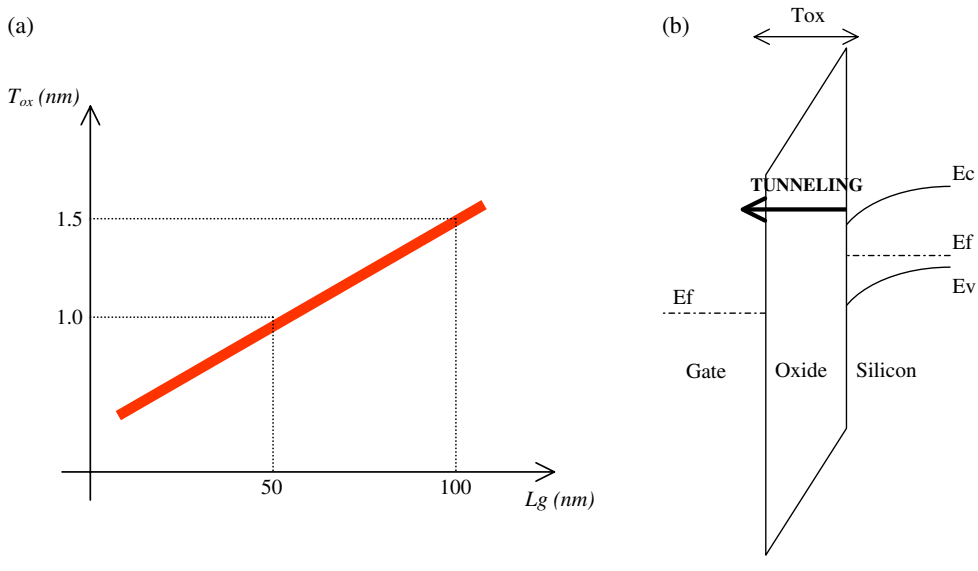
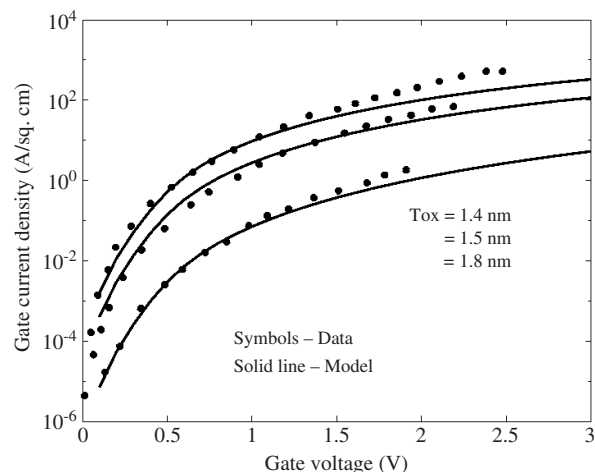


Figure 6: Gate leakage current due to direct tunneling [8].



2.1.5. Metal Gate Electrode

The early MOSFET technology was based on metal gate electrodes. However, the metal gate was replaced with highly doped poly-Si electrode, since the self aligned transistor was possible by doping the gate and source/drain during the single implantation step. The current CMOS technology is a dual gate technology with n⁺ poly-Si and p⁺ poly-Si gate electrodes for NMOS and PMOS respectively. However, the resistivity of the poly-Si will limit the performance in the sub-65 nm technology and hence the need for the metal gate electrode. However the choice of the material for metal gate electrode is not trivial because the work function of the gate plays an important role in deciding the V_t of the transistor. The NMOS and PMOS transistors may

Chalcogenides: The compounds consisting of chalcogen element – group VI element in the periodic table – can be switched between a high resistance amorphous phase and a low resistance crystalline phase. A very common chalcogenide used in phase change memory is Tellurium (Te) compound – GeSbTe.

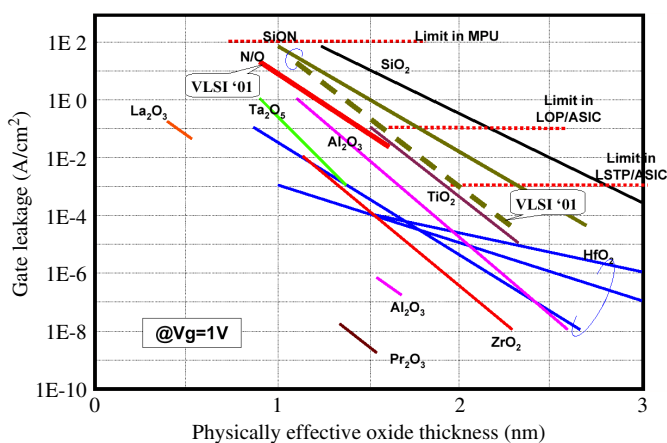
need different gate work functions, complicating the integration of the gate metal in the CMOS technology [25]. The work function engineering of the gate electrode will play an important role in deciding the V_t of the transistor and hence the choice of the appropriate material.

The combination of all these techniques may become essential to enable scaling of the CMOS device technology to an ultimate limit of about 5 nm physical gate length by the year 2020. However several innovations are required in semiconductor process technology to enable such devices. The detailed discussion of the process technology requirements is beyond the scope of this article. One of the key process technology requirements is the scalability of the photolithography process beyond 65 nm technology node. In the immediate future, 193 nm UV with immersion technique may be sufficient. However, alternative lithography techniques such as Extreme UV lithography, mask less lithography, imprint lithography, electron beam lithography may be necessary to reach the ultimate CMOS scaling limit.

2.2. Non CMOS architectures

The non-CMOS architectures will be required to continue the performance enhancement of nanoelectronics systems beyond the ultimately scaled CMOS. Some of these device structures will also require completely different information processing paradigms. Any electronics system fundamentally does two tasks: computation and storage. The computational element is referred to as “Logic” block and the storage element is referred to as “Memory” block. A particular system architecture combines these blocks to yield the required functionality. For example, the current CMOS technology is utilized to build systems based on digital – boolean architectures, operating on the binary variable.

Figure 7: Gate leakage current for high-k dielectrics [9].

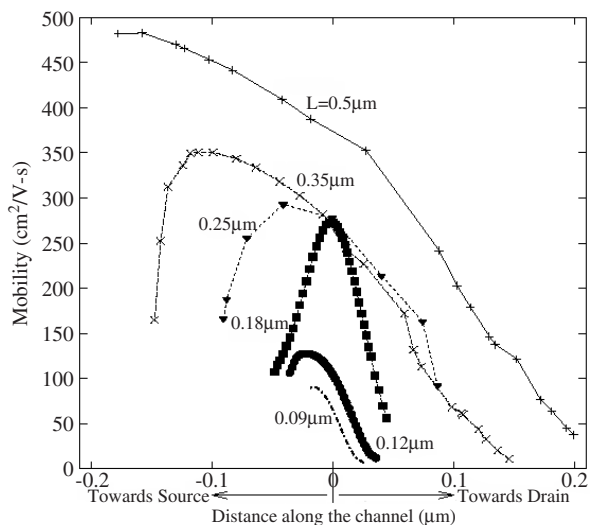


2.2.1. Non-CMOS Memory architectures

The conventional CMOS memories are charge-based, and hence they become very vulnerable to noise with scaling because the charge stored on the memory cell is decreasing with every technology generation. The alternate memory structures exploit different principles for storing the information.

Phase Change Memory (PCM) cell consists of a material (such as chalcogenides) whose phase can be changed reversibly between a high resistance state and low resistance state [26]. Ferroelectric RAM (FeRAM) exploits the effect of remnant polarization on a ferroelectric gate dielectric such as lead zirconate titanate (PZT) [27]. Magnetoresistive RAM (MRAM) utilizes the magnetic tunnel junction

Figure 8: Electron mobility in NMOS for different technology nodes [15].



Cellular Nonlinear Networks (CNN): Cellular neural networks (CNN) are systems based on a parallel computing architecture with an array of weak compute elements, wherein the communication is allowed only between the neighbouring compute units.

to modulate the resistance of memory cell [28]. Polymer memory relies on a memory cell with organic-metal-organic film stack, whose resistance can be programmed into high and low values to store information [29]. Molecular memory is a very broad class of memory cell, which stores information on a single molecule connecting two electrodes, by modulating the conductivity of the molecule [30]. It is expected that some of these alternate memory structures, especially the first three options, may be integrated along with the CMOS logic some time in the next decade.

2.2.2. Non-CMOS Logic architectures

This class of devices attempts to realize the fundamental element of a Logic block, *the transistor*,

in a completely radical fashion. In some cases such devices are not amenable to integrate into the conventional digital architectures and a completely novel system architecture paradigm will be required.

Nano-wire and nano-tube based 1-D transistor structures are potential candidates. The notable among them is the Carbon Nano Tube (CNT) transistor [31]. In order to integrate it with the conventional architecture, the issue of selectively defining the CNT based NMOS and PMOS transistors in giga-scale by growing the CNTs at precise locations on the chip should be resolved. The Resonant Tunneling Transistor (RTT) is another alternative with a negative differential resistance and very high switching speed [32]. Single Electron Transistors (SET) are structurally similar to the conventional MOSFET except that the source to substrate and drain to substrate junctions are tunnel junctions and the channel is ideally formed by a quantum dot [33]. A SET can be potentially used to represent a binary bit by a single electron, but alternate architectures such as Cellular Nonlinear Networks (CNN) will be required to realize such systems. Another class of transistors exploits the spin of the electrons, leading to a field that is popularly referred to as “Spintronics”. There are several concepts proposed to build such devices. The spin MOSFET is a hybrid device which relies on gate control (electro static channel control) and drain control (spin dependent scattering at the drain). However, this requires the efficient injection of spin polarized electrons into the channel, necessitating half-metallic-ferromagnetic contacts at the source and drain [34]. There are other concepts such as spin-torque transistor [35] and spin-gain transistor [36]. However, all these concepts are still far from surpassing the capability of the ultimately scaled CMOS.

The “Optoelectronics” or “Photonics” proposes to build digital computers by using photons for information processing and transmission at the speed of the light [37]. However, the component size will be limited by the diffraction phenomenon. Furthermore, the compatibility with the existing computer architecture is unclear. On the other hand, the integration of optical interconnects for chip level wiring in the CMOS technology is considered to be more promising. “Organic electronics” or “plastic electronics” builds the transistors on organic thin film substrate such as pentacene [38]. These devices are extremely slow due to low channel mobility, and hence they do not compete with the ultimate CMOS in building the conventional compute architectures. However, these devices are expected to complement the CMOS technology and extend the electronics applications into new domains

Figure 9: Strained silicon NMOSFET structure with tensile strain in the channel.

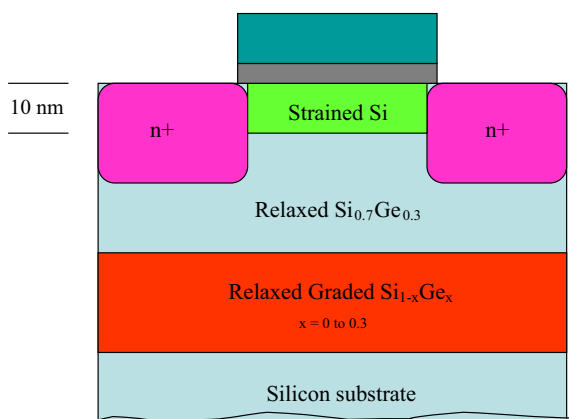
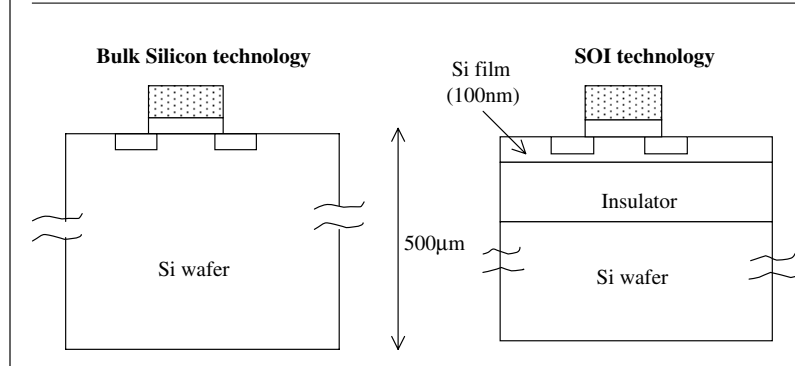


Figure 10: Schematic comparison of bulk MOSFET and SOI MOSFET.



such as flexible displays, wearable electronics, electronic paper, and low cost sensors. The organic electronics devices will utilize the conventional printing technology (*printable electronics*) and provide very low cost solutions. The Nano-Electro-Mechanical-Systems (NEMS) open up the possibility of nanomechanical computer where the digital bit is represented in mechanical domain, by displacement of cantilevers [39]. However, the speed of signal propagation is limited by the velocity of the sound. Quantum electronics and quantum computing exploit the phase information in the quantum mechanical wave function for information processing [40]. Unlike the binary state *bit* in the conventional computers, the *qubit* in quantum computing can hold the state of 1 or 0 or superposition of these at the same time. The underlying devices fall into the class of devices which are based on particles with 2 spin states. Due to inherent massive parallelism, such computers will be ultra-fast in applications such as cryptography. Several issues are yet to be overcome for such devices to become practical. Also, it is not clear whether all

Qubit: A unit of quantum information wherein the bit can take a superposition of 0 and 1 states due to quantum entanglement.

classes of problems addressed by CMOS computing architecture are amenable for quantum computing architecture.

The molecular transistor, or the possibility of creating a switching device on a single molecule, has raised the expectations for the fascinating field of “Molecular electronics” [41]. These devices can potentially result in very high density with the device size of the order of 1nm. However, some of the major issues include assembling these devices on giga-scale and interconnecting them appropriately to yield complex systems. In fact, a completely different system architecture may be required to utilize these devices. Another important issue relates to the fabrication of these devices. These devices will likely be fabricated using chemical synthesis, i.e. *atoms-up* approach, rather than the conventional *top-down* approach using photolithography. This could be a disruptive technology that can enhance the nanoelectronics system performance very significantly beyond the ultimately scaled CMOS.

It is expected that none of the non-CMOS architectures can really replace the ultimately scaled CMOS. Instead, CMOS architectures will co-exist with a variety of other device technologies and enable the nanoelectronics system performance enhancement beyond 2020.

3. Circuits and Systems Design

The two important issues in realizing the giga-scale integrated circuits is the exponential increase in the power dissipation and the device variability. It is absolutely essential to manage these issues for a successful scaling of the CMOS technology to reach the ultimate limit by 2020. Although some of the novel device technologies are attempting to address these issues to some extent, novel design techniques at the circuits and systems level are considered to be very important.

3.1. Power dissipation

As indicated in Table 1, the high performance chips are already consuming more than 150 W of power. The high power dissipation also impacts the performance and reliability of the chip and requires very elaborate heat extraction techniques from the chip. Figure 13 shows the power dissipation trends in Intel microprocessors [5]. The biggest concern is a dramatic increase in the standby power of the chip. This is especially troublesome in battery operated devices. This increase is attributed to the exponential rise in the sub-threshold leakage in nano transistors.

The circuit techniques utilizing the multi- V_t transistors are useful in this context [42]. This is based on the notion that the performance of a

Figure 11: Subthreshold conduction in MOSFET.

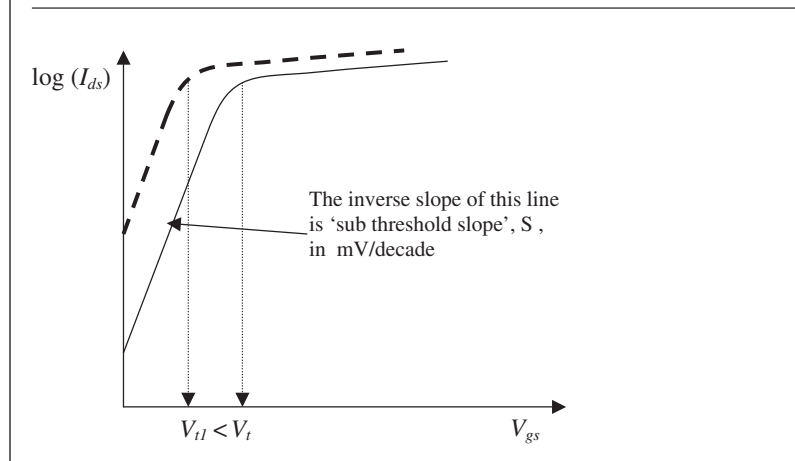


Figure 12: 3-Dimensional FinFET structure. The current flows from source (S) to drain (D) in lateral direction, through the channel wrapped around by the gate fin (G).

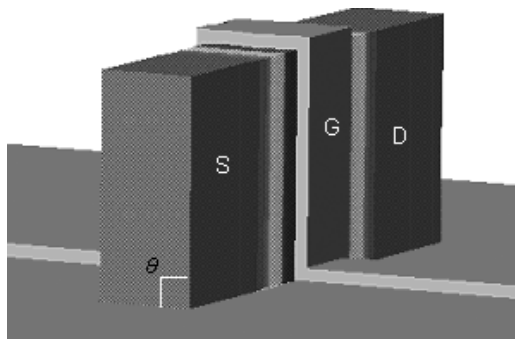
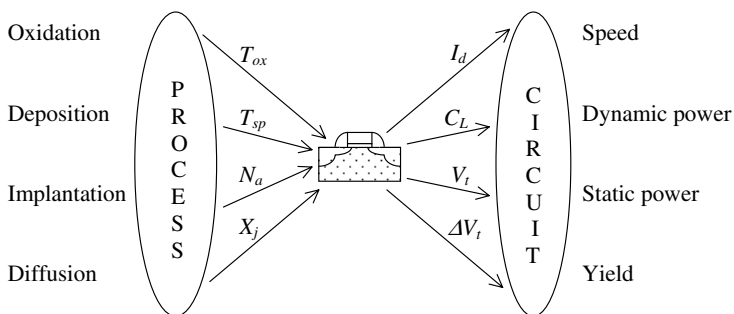


Figure 13: Power dissipation trends with technology scaling [5].



Figure 14: Impact of device variability on circuit performance.



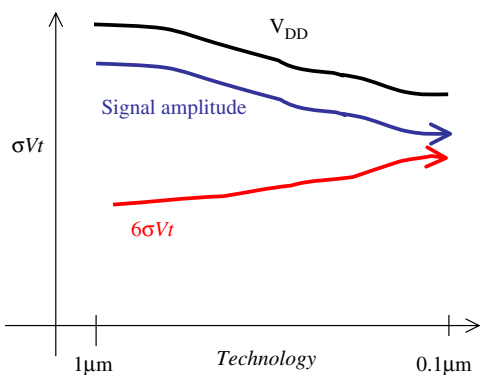
complex circuit with several million transistors is dictated by a few critical paths in the circuit. The number of transistors that occupy the critical paths is a very small fraction of the total number of transistors on the chip. It is sufficient to ensure only these transistors have a low- V_t . The rest of the transistors can have high- V_t and thereby reduce the static leakage power. The other techniques include the use of sleep transistor for power gating the unused circuit blocks on the chip [43]. Additionally, the system level innovations such as low power architectures exploiting parallelism for performance enhancements [44], software techniques to implement low power algorithms to implement energy efficient algorithm for a given task are beneficial to move forward [45].

3.2. Device variability

The device-to-device variability and its impact on the circuit design methodology has been regarded as a potential show-stopper for the scaling of CMOS technology. This refers to the statistical differences between pairs of identically designed and processed transistors. This arises due to statistical nature of underlying processes used to fabricate MOS transistor as shown in Fig. 14. The variation in the process parameters results in the variation in the structural parameters of the device such as junction depth (X_j), oxide thickness (T_{ox}), doping concentration (N_a). This in turn leads to variation in the electrical parameters of the device, eventually resulting in circuit parameter fluctuations. This results in a stringent trade-off between precision/speed versus yield for analog/digital VLSI application. Typically any given device parameter, such as V_t , has a normal distribution around its expected target value. The standard deviation, σ , of this distribution describes the extent of the variability. Figure 15 shows the variability trend with scaling. While the signal levels are going down with scaling, the noise due to variability is increasing with scaling [46]. This is because the intrinsic factors influencing the variability such as random dopant number fluctuation in nano transistors, get worse with scaling. Thus, it is becoming extremely difficult to create the nano transistors that are identical to each other.

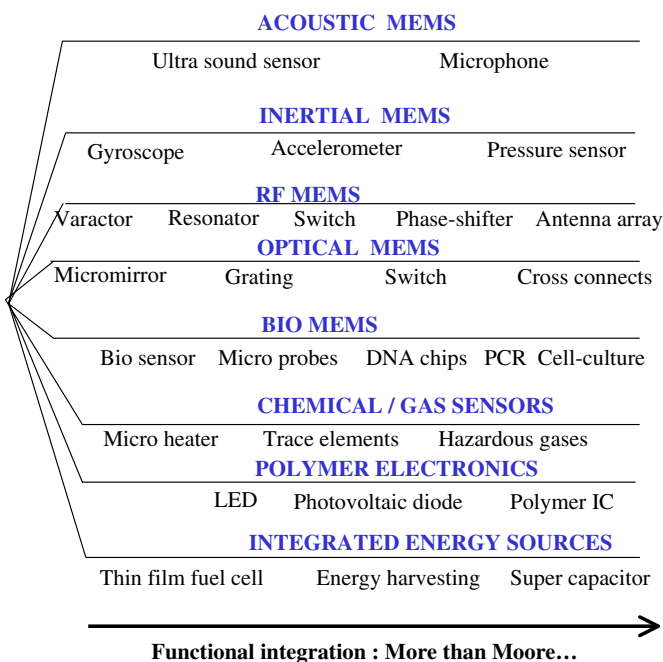
The conventional technique of over-designing for the worst case becomes very pessimistic. On the one hand, it has become very essential to accurately predict and model the variability and on the other hand it is essential to develop robust design techniques at the circuit and system levels. Recently statistical modeling techniques have been proposed to relate the circuit and device variability to the underlying semiconductor

Figure 15: Variability is increasing with technology scaling [46].



processes [47–49]. The circuit design techniques such as adaptive body biasing to tune the transistor parameter, and post fabrication corrections for non idealities are becoming essential [50–52]. The system architecture level techniques are also essential to handle variability in the long term. The adaptive neural network based architectures are attractive in this context [53,54]. The defect tolerant and bio-inspired architectures are capable of correcting the errors due to variability and realize very robust high performance systems [55,56].

Figure 16: Hybrid Systems on Chip enabled by the functional integration of non CMOS components



4. Hybrid Systems on a Chip

Traditionally, the electronics system performance has been enhanced in the context of Moore’s law by increasing the computing performance and storage capacity of the chips. However, for the past few years, it is becoming increasingly important to exploit the silicon technology infrastructure into other application domains. Specifically, the functional integration of nonelectronic domain components onto the chip is being driven by, what is popularly known as, the More than Moore philosophy. These chips are also referred to as Hybrid Systems on Chips because there is on chip interaction of multiple energy domains.

Figure 16 illustrates the different possibilities in terms of functional integration. In a majority of these cases the same set of semiconductor processes that are traditionally used to build transistors, are utilized to create micro/nano machined (MEMS/NEMS) structures on silicon. These structures can act as sensors and actuators and interface with the on chip electronics. Silicon has excellent mechanical properties, which enables the creation of robust mechanical building blocks such as beams, membranes, rigid masses, etc [57]. The success of such hybrid SOCs can be illustrated through some of these chips that have penetrated the consumer applications. Most of the airbag release systems deployed in automobiles employ a MEMS accelerometer chip [58]. It consists of a suspended poly-silicon proof mass created on top of Silicon substrate through surface micro-machining technique. The mass is configured as one of the plates of the capacitor, and hence the capacitance changes in response to acceleration. The movement of the proof mass is measured through capacitance sensing using the electronics interface circuit on the chip. This integration results in high reliability, low cost and high performance solution which can be deployed in a variety of applications. Another great success story has been the Digital Micromirror Display (DMD) device in the optical MEMS domain [59] which is routinely used to replace the LCD projectors. In this case, the pixel array of suspended micromirrors is created on top of electronics chip which provides the information to be displayed. Each of the micromirror is electrostatically actuated to control the reflection of the light to produce the appropriate grey level for the pixel. Similar opportunities exist in various other domains as illustrated in Fig. 16. These systems are also expected to be autonomous through harvesting the energy from the environment. Eventually the convergence of electronics and biology for *lab-on-a-chip* applications in diagnostics and therapy may be the biggest technological revolution in the 21st century.

5. Conclusions

The Silicon CMOS technology will continue to be scaled to an ultimate limit of about 5 nm physical gate length of a transistor, albeit requiring several innovations in process technology, materials engineering, and exploitation of new device physics. The nanoelectronics system performance is expected to scale according to the predictions of Moore's law along this path. However, the circuit and system level issues of power management and variability tolerance should be addressed through novel techniques. After the end of the CMOS roadmap, beyond 2020, radically new device technologies will be required for further enhancement of system performance. The performance enhancement of the SOCs will also evolve along a new trajectory of functional integration to include non electrical domain components. The sensors and actuators integrated along with the conventional compute and storage functions will open up new application domains.

A. Centre of Excellence in Nanoelectronics

The Ministry of Communication and Information Technology (MCIT), Government of India, has initiated an ambitious project to set-up two Centres of Excellence in Nanoelectronics at the Indian Institute of Science, Bangalore and the Indian Institute of Technology, Bombay. This project has an outlay of Rs. 100 crores over 5 years (2006–2010) to create comprehensive nanofabrication facilities (nanofabs) at these two institutes. These nanofabs will act as nodal centres in the country to facilitate the state-of-the-art experimental research. The nanofab infrastructure will enable the fundamental exploration in several areas of nano science and engineering and it is also expected to seed new technologies that will have societal and commercial impact.

At IISc, a large number of faculty members from various disciplines of science and engineering are involved in this project. The institute has taken a lead to foster this interdisciplinary effort by initiating the creation of a new building to support the nanofab infrastructure. The nanofab is expected to be fully functional by 2008. When completed, it would be on-par with some of the best university nanofabs elsewhere. The facility will also be available for other researchers in the country through the Indian Nanoelectronics Users Program (INUP).

Received 07 December 2006; revised 31 January 2007.

References

1. "The Chip that Jack Built Changed the World", <http://www.ti.com/corp/docs/kilbyctr/jackbuilt.shtml>

2. "International Technology Roadmap for Semiconductors", <http://www.itrs.net>
3. G. E. Moore, "Cramming more Components onto Integrated Circuits", *Electronics*, p.114,1965.
4. Bandyopadhyay, "Moore's law governs the Silicon revolution", p.78, *Proc. of IEEE*, 1998.
5. S. Thompson, P. Packan, M. Bohr, "MOS Scaling : Transistor Challenges for the 21st Century", *Intel Technology Journal*, 1998
6. Y. Taur, D.A.Buchanan, W. Chen, D.J. Frank, K.E.Ismail, S-H Lo, G. A. Sai-Halasz, R.G.Viswanathan, H-J C. Wann, S. J. Wind, H-S Wong, "CMOS Scaling into the Nanometer Regime", *Proc. of IEEE*, p. 486, 1997.
7. P. Gilbert et. al., "A high performance 1.5V, 0.1mm, Gate length CMOS technology with scaled copper metallization", p.1013, *International Electron Device Meeting*, 1998.
8. K. Maitra and N. Bhat, "Analytical approach to integrate the different components of direct tunneling current through ultrathin gate oxides in n-channel metal oxide semiconductor field-effect transistors", *Journal of Applied Physics*, Vol 93, No. 2, pp. 1064-1068, 15 January 2003
9. H. Iwai, *Personal communications*
10. T. Ando et. al., "Sub-1nm EOT HfSi_x/HfO₂ Gate Stack Using Novel Si Extrusion Process for High Performance Application", *Proceedings of VLSI Technology Symposium 2006*
11. K.L. Lee et. al., "Poly-Si/AlN/HfSiO stack for ideal threshold voltage and mobility in sub-100 nm MOSFETs", *Proceedings of VLSI Technology Symposium 2006*
12. J.A. Ng, Y. Kuroki, N. Sugii, K. Kakushima, S.-I. Ohmi, K. Tsutsui, T. Hattori, H. Iwai and H. Wong, "Effects of low temperature annealing on the ultrathin La₂O₃ gate dielectric; comparison of post deposition annealing and post metallization annealing", *Microelectronic Engineering*, p. 206, 2005.
13. M.P. Singh, C.S. Thakur, K. Shalini, S. Banerjee, N. Bhat, and S.A. Shivashankar, "Structural, Optical, and Electrical Characterization of Gadolinium Oxide Films Deposited by Low-pressure Metalorganic Chemical Vapour Deposition", *Journal of Applied Physic*, Vol 96, No. 10, pp. 5631-5637, 15 November 2004
14. M.P. Singh, C.S. Thakur, K. Shalini, N. Bhat, and S.A. Shivashankar, "Structural and Electrical Characterization of Erbium Oxide Films Grown on Si(100) by Low-pressure Metalorganic Chemical Vapour Deposition", *Applied Physics Letters*, October 2003
15. R. Srinivasan, "RF CMOS Device Engineering and Performance Analysis", *Ph.D. thesis, Indian Institute of Science*, 2006.
16. A.V-Y. Thean, "Strain-Enhanced CMOS Through Novel Process-Substrate Stress Hybridization of Super-Critically Thick Strained Silicon Directly on Insulator (SC-SSOI)", *Proceedings of VLSI Technology Symposium 2006*
17. K. Ota et.al., "Scalable eSiGe S/D technology with less layout dependence for 45-nm generation", *Proceedings of VLSI Technology Symposium 2006*
18. Krishnamohan, T. Krivokapic, Z. Uchida, K. Nishi, Y. Saraswat, K.C., "High-mobility ultrathin strained Ge MOSFETs on bulk and SOI with low band-to-band tunneling leakage: experiments", *IEEE Transactions on Electron Devices*, May 2006.
19. G. Shahidi, "SOI technology for the GHz era", *IBM Journal of Research and Development*, <http://www.research.ibm.com/journal/rd/462/shahidi.html> 2002.
20. James B. Kuo, Shih-Chia Lin, "Low-Voltage SOI CMOS VLSI Devices and Circuits", *Wiley-Interscience*
21. Taur and Ning, "Fundamentals of Modern VLSI Devices", *Cambridge University Press*
22. Jack Kavalieros, "Tri-Gate Transistor Architecture with High-k Gate Dielectrics, Metal Gates and Strain Engineering", *Proceedings of VLSI Technology Symposium 2006*

23. H. Shang, "Investigation of FinFET devices for 32nm technologies and beyond", *Proceedings of VLSI Technology Symposium 2006*
24. H. Lee, "Sub-5nm All-Around Gate FinFET for Ultimate Scaling", *Proceedings of VLSI Technology Symposium 2006*
25. P.F. Hsu, "Advanced Dual Metal Gate MOSFETs with High-k Dielectric for CMOS Application", *Proceedings of VLSI Technology Symposium 2006*
26. S. Lai, "Current status of the phase change memory and its future", *IEDM Technology Digest*, 2001
27. Y.M. Kang, "World Smallest 0.34 μ m² COB Cell 1T1C 64Mb FRAM with New Sensing Architecture", and Highly Reliable MOCVD PZT Integration Technology, *Proceedings of VLSI Technology Symposium 2006*
28. W.J. Gallagher, S.S.P. Parkin, "Development of the Magnetic Tunnel Junction MRAM at IBM : From first junctions to 16 MB MRAM Demonstrator Chip", *IBM Journal of Research and Development*, January 2006.
29. J. He, J., et.al. "Three-terminal organic memory devices," *Journal of Applied Physics*, 97, 2005.
30. W. Wu, et.al. "One-kilobit cross-bar molecular memory circuits at 30-nm half-pitch fabricated by nanoimprint lithography," *Applied Physics A*. 80.6, p. 1173, 2005.
31. Lin, Y., et.al. "Novel carbon nanotubes FET design with tunable polarity," *IEDM Technical Digest* p.687, 2004.
32. Matsuo, N., H. Kihara, Y. Takami. "Application of advanced metal-oxide-semiconductor transistor in next generation, silicon resonant tunneling MOS transistor, to new logic circuit," *Solid State Electronics*, p.1969, 2003.
33. Park, K-S., et al. "SOI Single-electron transistor with low RC delay for logic cells and SET/FET hybrid ICs," *IEEE Trans. Nanotechnol.* 4.2, p. 242, 2005.
34. Sugahara, S. and M. Tanaka, "A spin metal-oxide-semiconductor field-effect transistor using half-metallic-ferromagnet contacts for the source and drain", *Applied Physics Letters*, 84.13, pp. 2307-2309, 2004).
35. Bauer, G. E. W., A. Brataas, Y. Tserkovnyak, B. J. van Wees. "Spin Torque Transistor," *Applied Physics Letters*, p.3928 , 2003.
36. Nikonov, D. E., G. I. Bourianoff. "Spin-Gain Transistor in ferromagnetic semiconductors-the semiconductor Bloch-equations approach," *IEEE Transactions On Nanotechnology*, 4.2, pp. 206-214, 2005.
37. Caulfield, H. J., "Perspectives in Optical Computing," *Computer* pp. 22-25, 1998.
38. Dimitrakopoulos, C. D., S. Purushothaman, J. Kymissis, A. Callegari, J. M. Shaw. "Low-voltage Organic Transistors on Plastic Comprising High-dielectric Constant Gate Insulators," *Science*, pp. 822-824 (1999).
39. Despont, M., J. Brugger, U. Drechsler, U. Düring, W. Haberle, M. Lutwyche, H. Rothuizen, R. Stutz, R. Widmer, G. Binnig, H. Rohrer, and P. Vettiger. "VLSI-NEMS Chip for Parallel AFM Data Storage," *Sensors and Actuators*, pp. 100-107, 2000.
40. Quantum Information Science and Technology Roadmapping Project, <http://qjst.lanl.gov/> 2005.
41. Chen, Y., D. A. A. Ohlberg, X. Li, D. R. Stewart, R. S. Williams, J. O. Jeppesen, K. A. Nielsen, J. F. Stoddart, D. L. Olynick, E. Anderson. "Nanoscale molecular-switch devices fabricated by imprint lithography," *Applied Physics Letters*. p.1610, 2003.
42. J.T. Kao, A.P.Chandrakasan, "Dual-threshold voltage techniques for low-power digital circuits", *IEEE Journal of Solid-State Circuits*, p. 1009, 2000.
43. C. Long and L. He, Distributed sleep transistor network for power reduction, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, p. 937, 2004.
44. Anantha P. Chandrakasan, Samuel Sheng, and Robert W. Brodersen, Low-Power CMOS Digital Design", *IEEE Journal of Solid-State Circuits*, p. 473, 1992.
45. S. Hua, Gang Qu, and S. S. Bhattacharyya, "Energy-efficient embedded software implementation on multiprocessor system-on-chip with multiple voltages", *ACM Transactions on Embedded Computing Systems (TECS)*, p. 321, 2006.
46. Marcel J. M. Pelgrom, Hans P. Tuinhout and Maarten Vertgeer, "Transistor Matching in Analog CMOS Applications", *IEDM Tech. Dig.*, 1998, pp. 915-918.
47. H. C. Srinivasaiah and Navakanta Bhat, "Mixed Mode Simulation Approach to Characterize the Circuit Delay Sensitivity to Implant Dose Variations", *IEEE Transactions on CAD of Integrated Circuits and Systems*, Vol. 22. No.6, June 2003.
48. H. C. Srinivasaiah and Navakanta Bhat, "Monte Carlo Analysis of the Implant Dose Sensitivity in 0.1 μ m NMOSFET", *Solid State Electronics*, Volume 47, pp. 1379 1383, Aug. 2003.
49. B.P. Harish, Mahesh B. Patil and Navakanta Bhat, "On A Generalized Framework for Modeling the Effects of Process Variations on Circuit Delay Performance" To appear in *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*
50. R. Singh and N. Bhat, "An Offset Compensation Technique for Latch Type Sense Amplifier in High Speed Low Power SRAMs", *IEEE Transactions on VLSI Systems*, pp. 652-657, June 2004.
51. N. Bhat and S. Mukherjee, "Yield and Speed Enhancement of Semiconductor Integrated Circuits using Post Fabrication Transistor Mismatch Compensation Circuitry", *US patent #6934200*, 2005.
52. S. Narendra, A. Keshavarzi, B. A. Bloechel, Shekhar Borkar, and Vivek De, "Forward Body Bias for Microprocessors in 130-nm Technology Generation and Beyond", *IEEE J. of Solid-State Circuits*, p. 696, May 2003.
53. A. Gupta and N. Bhat, "Asymmetric Cross-Coupled Differential Pair Configuration to Realize Neuron Activation Function and its Derivative", *IEEE Transactions on Circuits and Systems Part II: Express Briefs*, Vol. 52, No. 1, pp. 10-13, January 2005
54. A. Gupta and N. Bhat, "Back-Gate Effect to Generate Derivative of Neuron Activation Function", *Analog Integrated Circuits and Signal Processing*, Kluwer 41 (1), pp. 89-92, October 2004
55. Heath, J. R., et al. "A Defect Tolerant Computer Architecture: Opportunities for Nanotechnology," *Science*, p. 1716, 1998.
56. Türel, Ö, J. H. Lee, X. Ma, and K. Likharev. "Neuromorphic architectures for nanoelectronic circuits," *Int. J. Circuit Theory and Applications*, p. 277 , 2004.
57. K. E. Peterson, "Silicon as a mechanical material," *Proc. IEEE*, vol. 70, pp. 420-457, May 1982.
58. Analog Devices Products, <http://www.analog.com/index.html>
59. M.R.Douglass, "DMD reliability: a MEMS success story" http://www.dlp.com/dlp_technology/images/dynamic/white_papers/153_Reliability_paper.pdf



Navakanta Bhat Navakanta Bhat received his B.E. in Electronics and Communication from University of Mysore in 1989, M.Tech. in Microelectronics from I.I.T. Bombay in 1992 and Ph.D. in Electrical Engineering from Stanford University, Stanford, CA in 1996. Then he worked at Motorola's Networking and Computing Systems Group in Austin, TX until 1999. At Motorola he worked on logic technology development and he was responsible for developing high performance transistor design and dual gate oxide technology. He joined the Indian Institute of Science, Bangalore in 1999 where he is currently an Associate Professor in the Electrical Communication Engineering department. His current research is focused Nano-CMOS technology and Integrated CMOS-MEMS sensors. The work spans the domains of process technology, device design, circuit design and modeling. He has several research publications in international journals and conferences and 3 US patents to his credit. He has received the Young Engineer Award (2003) from the Indian National Academy of Engineering. He is also the recipient of the Swarnajayanti fellowship (2005) from the Department of

Science and Technology, Govt. of India and Prof. Satish Dhavan award (2005) from the Govt. of Karnataka. He was the founding chair of the IEEE Electron Devices and Solid-State Circuits society, Bangalore chapter which was recognized as the Outstanding Chapter of the Year by the IEEE SSC society (2003) and IEEE EDS

society (2005). He has been on the program committees of several international conferences. He was the technical program chair for the international conference on VLSI design and Embedded systems (2007). He is a Distinguished Lecturer of the IEEE Electron Devices Society.