Invited Paper

Challenges for Nanoscale MOSFETs and Emerging Nanoelectronics

Yong-Bin Kim[†]
Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA

Complementary metal-oxide-semiconductor (CMOS) technology scaling has been a main key for continuous progress in silicon-based semiconductor industry over the past three decades. However, as the technology scaling enters nanometer regime, CMOS devices are facing many serious problems such as increased leakage currents, difficulty on increase of on-current, large parameter variations, low reliability and yield, increase in manufacturing cost, and etc. To sustain the historical improvements, various innovations in CMOS materials and device structures have been researched and introduced. In parallel with those researches, various new nanoelectronic devices, so called "Beyond CMOS Devices," are actively being investigated and researched to supplement or possibly replace ultimately scaled conventional CMOS devices. While those nanoelectronic devices offer ultra-high density system integration, they are still in a premature stage having many critical issues such as high variations and deteriorated reliability. The practical realization of those promising technologies requires extensive researches from device to system architecture level. In this paper, the current researches and challenges on nanoelectronics are reviewed and critical tasks are summarized from device level to circuit design/CAD domain to better prepare for the forthcoming technologies.

Keywords: Nanoscale complementary metal-oxide-semiconductor, Scaling, Small-geometry effects, Short channel effects, Silicon-on-insulator technology, Nanoelectronics

1. INTRODUCTION

The relentless evolution of electronics, information technology (IT), and communications has been mainly enabled by continuous progress in silicon-based complementary metal-oxide-semiconductor (CMOS) technology. This continuous progress has been maintained mostly by its dimensional scaling, which results in exponential growth in both device density and performance. The reduction in cost-per-function has steadily been increasing the economic productivity with every new technology generation. In addition to its scalability, the unique device properties such as high input resistance, self isolation, zero static power dissipation, simple layout and process steps have made CMOS transistors as the main components of the current integrated circuits (ICs). Today CMOS ICs are everywhere and indispensible in our life, ranging from portable electronics to telecommunications and transportation [1].

However, as device scaling continues for the 21st century, it turns out that the historical growth, doubled circuit density and increased performance by about 40% every technology generation, followed by "Moore's Law," [2, 3] cannot be maintained only by the conventional scaling theory. Increasing leakage current does not allow further reduction of threshold voltage, which in turn impedes further supply voltage scaling for the historical speed improvement. Accordingly, generated higher electric fields inside of the transistor worsen device reliability and increase leakage currents. Moreover, the required high channel doping causes significant challenges such as mobility degradation, increased band-to-band tunneling (BTBT), gate-induced drain leakage (GIDL) and random dopants induced threshold voltage fluctuations [4].

Due to those ever increasing short channel effects (SCEs), there have been several strategies introduced for CMOS device to extend Moore's Law. A few examples of those are; increasing electrostatic control over the channel by means of the continuous equivalent oxide thickness (EOT) scaling with high-k/metal gate stack, multi-gate structures for higher drive current at the constant over-drive voltage, improving carrier mobility by adopt-

[†] Author to whom all correspondence should be addressed:
E-mail: ybk@ece.neu.edu

ing high mobility channel materials, and strain engineering and reducing parasitics.

In addition, various researches have been actively carried out in device domain to find an alternative device to continue to sustain Moore's Law. Among these efforts, various kinds of alternative memory and logic devices (beyond CMOS devices) have been proposed. These nanodevices take advantages of the quantum mechanical phenomena and ballistic transport characteristics under low supply voltage and consume low power. Furthermore, due to their extremely small sizes, those devices are expected to be used for ultra-high density integrated electronic components having billions of devices in a single chip. However, it also increases defects and variations both during manufacture and chip operations.

In the introductory part of this paper, after the CMOS technology scaling trend and "Small-Geometry Effects" are reviewed to understand what impedes further down-scaling of planar bulk metal-oxide-semiconductor field-effect transistors (MOSFETs), the current and tentative efforts to cope with the limiting factors are discussed. After the CMOS scaling issues are reviewed, present and promising nanotechnology devices such as silicon-on-insulator (SOI) devices, Multi-gate FETs, Si nanowire/cabon nanotube FETs are investigates in terms of their advantages/drawbacks, basic operation principles, current state of development, and challenges for commercialization along with other possible options. Finally, this paper will conclude with a discussion of the various challenges to be resolved for nanoelectronics to be a full-brown technology from device to system architecture level.

2. CHALLENGES FOR NANOSCALE MOSFETs

It is important to have a solid understanding of the phenomena that occurs in nanoscale MOSFETs since it not only provides useful interpretation of the anomalies that we might encounter in their SPICE simulations but also provides us the future possibilities and limitations of the device scaling. In this section, we will review these phenomena starting from a brief overview of scaling trend of silicon-based planar bulk MOSFETs and highlight the challenges to be solved. Current approaches to control these small-geometry effects such as increased leakage currents, threshold voltage variations, weakened gate controllability over the channel, and increased in S/D resistances will be discussed.

2.1 Scaling trend of MOSFETs

Since the first MOSFET was successfully demonstrated in 1960 [5], in order to keep up with the increasing demand for faster, smaller, and cheaper products, lots of efforts have been made to shrink the size of MOSFET switching devices in semiconductor industry. Without any significant fundamental changes, the basic structure of MOSFETs has been evolved to continue its successful geometric scaling. The fundamental principles to scale down devices have been guided by the ideal scaling principles formulated by Robert Dennard in 1974 [6], and it is based on a simple electrostatics; if the device dimensions (W, L, and t_{ox}) and voltages (V_{DD} and V_{th}) are scaled down by the factor of α with the increased doping concentration by α (>1), all electric fields in the scaled transistor will remain the same as it was in the original device, hence named "constant-field scaling." Historically, the geometric scaling factor of each new generation has been $\alpha = \sqrt{2}$ so the number of transistors per unit area doubles every technology generation [2, 3]. To achieve higher circuit density and

performance while maintaining device reliability and constant power dissipation per unit area, there have been considerable reductions in the device dimensions and supply and threshold voltages. Recently, however, as the supply voltage (V_{DD}) scaling approaches to 1 V, the conventional scaling has been deviated from the ideal constant-field scaling due to the difficulty of further lowering the threshold voltage (V_{th}). This fundamental problem stems from the nonscalabe characteristic of the thermal voltage ($V_T = kT/q$) which makes the relatively fixed sub-threshold swing (S) at the constant temperature [7-9]. This, in turn, makes the sub-threshold leakage current increase exponentially as the V_{th} reduces. Therefore, there exists a lowest possible value of V_{th}, which is determined by the application constraints related to power consumption and circuit functionality. In addition, with the knowledge of increased V_{th} variation in nanoscale MOSFETs, it is necessary to make an enough margin for V_{th} variation to ensure that V_{th} stays well away from the lowest possible value of V_{th} [10]. Furthermore, especially for high-performance (HP) logic technology, it is required to keep a certain level of V_{DD}-V_{th}, which determines the drive current and hence performance in a chip. It makes further down-scaling of the supply voltage difficult. Fig. 1 shows the difficulty on supply voltage scaling. Under fixed V_{th}, reducing V_{DD} trades off performance (speed) and leakage power. The power density trend between the active and standby power is shown in Fig. 2. These trends in technology scaling have made us enter a new era in achieving high performance under constrained power [8].

2.2 Issues for nanoscale MOSFETs: *Categorized based on where the problems occur*

2.2.1 Channel

1) Sub-threshold leakage current: The Sub-threshold leakage current is the weak inversion conduction current, which is dominated by the diffusion current flowing between the drain and source when $|V_{GS}| < |V_{th}|$. It is considered as one of non-ideal characteristics of MOSFET as a switching device and contributes major portions of the standby leakage power dissipation. This weak inversion conduction current can be expressed based on the Eq. (1) [11],

$$I_{subth} = \mu C_{dep} \left(\frac{W}{L} \right) V_T^2 \left(exp \left(\frac{V_{GS} \cdot V_{th}}{n V_T} \right) \right) \left(1 \cdot exp \left(- \frac{V_{DS}}{V_T} \right) \right) \qquad (1)$$

where $C_{dep} = \sqrt{\varepsilon_{si} q N_{sub} / 4 \varphi_B}$ denotes the capacitance of the depletion region under the gate area, V_T is the thermal voltage which equals to kT/q, and n is the sub-threshold parameter and expressed as $1 + C_{dep}/C_{ox}$.

Since, from Eq. (1), I_{subth} increases exponentially with both increasing V_{GS} and decreasing V_{th}, the partial derivative of $log_{10} I_{subth}$ with respect to V_{GS} yields a constant slope called "sub-threshold slope (SS)" and equals to

$$ss = \frac{\partial (log_{10} I_{subth})}{\partial V_{GS}} = \frac{1}{ln 10 \bullet I_{subth}} \frac{\partial I_{subth}}{\partial V_{GS}} \qquad (2)$$

This parameter shows how abruptly the transistor turns off with decreasing gate voltage. In order to turn off the transistor effectively, *S* must be designed to be as small as possible. Fig. 3 shows that *S* is always greater than 2.3 V_T (~60 m*V/dec*) at room temperature and shows how well the channel surface potential can be controlled by the gate contact. *S* can be made smaller (close to ~60 m*V/dec*) by using a thinner gate oxide thickness

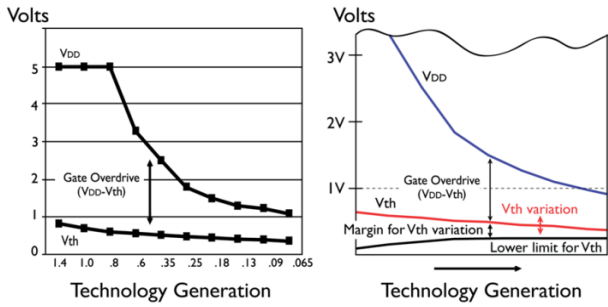Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

95

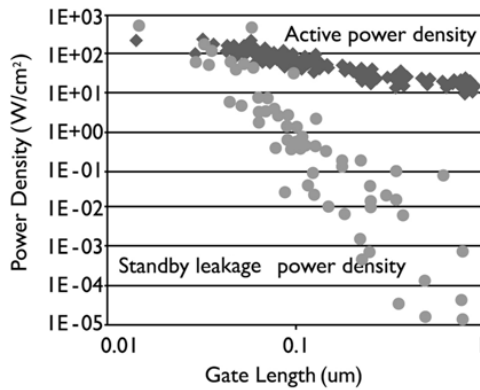Fig. 1. Trend of supply voltage and threshold voltage scaling.

Fig. 2. Power density trends between the active power and standby leakage power with CMOS Technology scaling: supply voltage scaling is essential to decrease overall power dissipation.
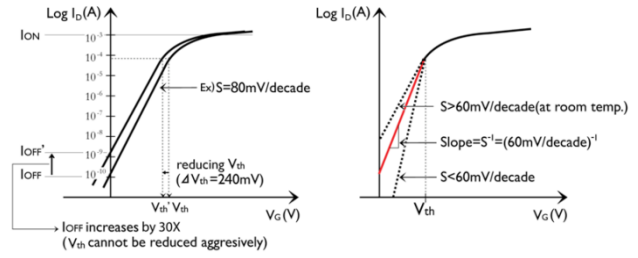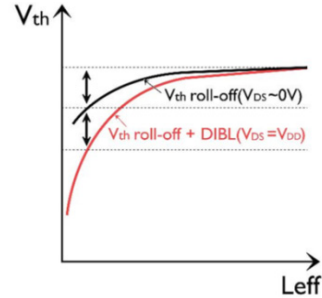

Fig. 3. Characteristics of sub-threshold conduction.

Fig. 4. Threshold voltage roll-off and drain induced barrier lowering (DIBL).

Fig. 5. Cross-section of an advanced MOSFET.

(resulting in larger C_{ox}) or a lower substrate doping concentration (resulting in the larger depletion width beneath the channel, hence reduced C_{dep}). In addition, under lower temperature operation, S can be reduced since S is a function of T. For MOS a transistor built in SOI technology, the sub-threshold swing is usually better than in bulk technology. In fact the sub-threshold swing of SOI devices can even reach the optimum value (2.3 V_T) depending on whether their bulk is fully depleted or partially depleted. This makes SOI a promising candidate for ultra low-power CMOS applications though a lot of improvements to the current process technologies still have to be made until this new technique can be applied to commercial products on a large scale [11, 12].

2) Threshold voltage variation: Since the threshold voltage (V_{th}) variation is directly related to the device speed and sub-threshold leakage current, it has to be minimized. It is generally explained in terms of 1) V_{th} *roll-off (or falloff)* and 2) *Drain induced barrier lowering (DIBL).*

The transistors with a different channel length (L) in the same wafer, even in the same die, yield difference V_{th}. The threshold voltage reduction due to the reduced channel length represents V_{th} *roll-off.* Further V_{th} reduction caused by increasing drain voltage describes *DIBL* as shown in Fig. 4. Both phenomena stem from the lowered potential barrier between the drain and source due to the relatively increased charge-sharing effect between the channel depletion region and source/drain depletion regions comparing to long-channel device case. This charge-sharing effect makes a transistor require less gate voltage to deplete the substrate beneath the gate dielectric and makes V_{th} decrease [11-13].

If the depletion region around the drain continues to extend to the source depletion region with the further increasing drain
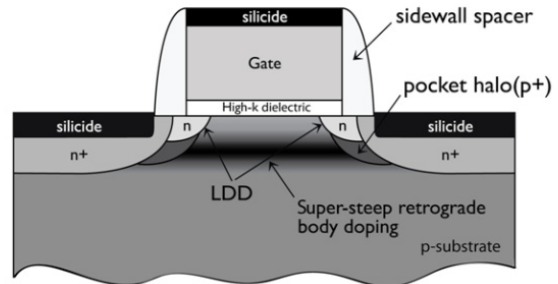
voltage and finally merges together before junction breakdown occurs, the drain current increases due to the existence of a parasitic current path located well below the gate. This phenomenon is known as *punch-through* and added up to the sub-threshold leakage current. In order to prevent the punch-through effect, one might think that an increase of the overall substrate doping level may be helpful to reduce D/S depletion regions. However, the higher doping also increases the sub-threshold swing (S) (as explained in section 2.2.1), BTBT current and GIDL current, and also it causes carrier mobility degradation. Therefore, it is not the best option to reduce the leakage current [4, 12, 14].

All these phenomena are collectively referred to as "short-channel effects (SCEs)" and tend to increase standby leakage power dissipation. To mitigate SCEs, an advanced MOSFET, depicted in Fig. 5, uses the silicidation of S/D junctions and lightly doped drain/source (LDD) structure with pocket halo implant and super-steep retrograde body doping [12, 15-17].

Additionally, increasing random dopant variations (RDF) and line-edge roughness (LER), particularly non-rectangular gate can cause significant statistical V_{th} variation and in turn severely affect circuit stability as well as performance distribution [18-20].

3) Carrier mobility degradation: Since the rate of the supply voltage scaling has been reduced while the geometric scal-

96

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

ing keeps the same historical rate, the electric fields inside the MOSFETs keep increasing. The drift velocity of the carriers is proportional to the longitudinal electric field across the channel at low field ($<10^3$ V/cm). After that point, however, the increasing rate of the carriers' velocity decreases with the increasing longitudinal field in Si at room temperature. Finally, the carriers reach their maximum velocity of $v_{sat} \sim 10^7$ cm/sec when the electric field exceeds $\sim 3 \times 10^4$ V/cm for electrons and $\sim 10^5$ V/cm for holes (here the channel length (L) is assumed to be much greater than the average distance (l) between scattering events such as $L >> 10$ nm.) This carrier mobility degradation is called "*Velocity Saturation*," and originates from various scattering mechanisms such as optical phonon scattering, phonon dispersion, phonon absorption as well as emission, and the energy band non-parabolicity [21, 22].

Another high electric field is developed between the gate and the channel due to the aggressive gate oxide thickness scaling with relatively constant supply voltage, which limits the charge carriers to a narrower region below the oxide-silicon interface, leading to more carrier scattering and hence lower mobility. Furthermore, the increased body doping by means of suppressing SCEs degrades the carrier mobility [21].

4) Hot carrier effects (HCEs): The high electric fields in a device also cause reliability problems such as threshold voltage shifts and trans-conductance degradation due to "Hot Carrier Effects (HCEs)." As the average velocity of carriers in the channel saturates by the increased scattering rate at the high electric fields, the carriers can attain high kinetic energy. Once those hot carriers obtain sufficient energy to overcome barriers, they might migrate into the unwanted area such as the gate dielectric, gate, or substrate of a transistor. Especially, the highly accelerated hot carriers near the drain region can generate new electron-hole pairs by collision with the silicon atoms, which is called "impact ionization." Impact ionization can cause significant increase in substrate current or carrier injection into the gate dielectric, which causes charges to get trapped in the gate oxide. This causes threshold voltage shifts and therefore the device becomes unstable and even can fail [19, 21].

The lightly LDD MOSFETs, shown in Fig. 5, have been widely used to reduce the maximum longitudinal electric field in the channel by means of the additional voltage drop across the higher resistive LDD region. As a result, the LDD technique increases hot carrier immunity in a device, but also degrades the device performance [15, 21].

5) Direct source to drain tunneling: The ultimate physical scaling limit of MOSFETs is direct source-to-drain tunneling. If the barrier width (transistor channel length) between source and drain becomes small enough for electrons to tunnel through the barrier without any additional gate bias, MOSFETs no longer can be used as a switch. When we consider only over-barrier transition of electrons but the direct tunneling, from the well-known Shannon-von Neumann-Landauer (SNL) expression, the minimum energy barrier to separate two different states of electrons is $E_{SNL} = k_B T(\ln 2) = 0.017$ eV at room temperature (T = 300 K), where k_B is the Boltzmann constant and T is the temperature. Therefore, the smallest energy required to process a bit is $E_{bit} > E_{SNL} = k_B T(\ln 2) = 0.017$ eV. Also, according to quantum mechanics the minimum channel length (barrier width) able to resist E_{bit}, not to allow tunneling, is about 5 nm. In order to achieve a smaller channel length, it is necessary to increase the Ebit. However this, in turn, tradeoffs with the growth of total power dissipation in a chip. Therefore, enhanced cooling techniques become the critical issues to achieve this goal [23, 24].

2.2.2 Gate

One of the key innovation enablers for 45 nm process technology and beyond is the high-k/metal gate transistor, which is regarded as one of the biggest developments in transistor design in 40 years [25]. High-k/metal gate have enabled the continuous EOT scaling. In this section, the motivation for using high-k/metal gate stack in advanced CMOS devices is discussed.

1) Direct tunneling gate leakage current: With the continuous device scaling, the gate oxide thickness has been accordingly reduced to maintain the gate controllability over the channel. However, as the gate oxide thickness scales below 2 nm, the direct tunneling (DT) gate leakage increases exponentially due to quantum mechanical tunneling. The DT gate leakage current can not only increase standby power dissipation but also limit the proper device operation [26]. These serious problems can be solved by replacing SiO_2 with higher permittivity (high-k) gate dielectrics, which allows a physically thicker dielectric layer to have an EOT [27]. Recently, high-k dielectric based on Hafnium and dual metal gate has been introduced to increase transistor performance while also reducing gate leakage as gate dielectric thickness actually increased while also the gate capacitance increased [25].

2) Gate depletion: Polycrystalline silicon (poly-Si) has been widely used as the gate electrode material of MOSFETs since it replaced aluminum on account of its superior thermal stability to higher processing temperature. In addition, the use of poly-Si as gate electrode has reduced the overall number of processing steps by means of self-aligned processing, where the gate itself is used as a hard mask during ion implantation for the source and drain junction formation. This technique allows tighter overlap between gate and source/drain regions and results in lower parasitic capacitance. Lastly, by controlling its doping concentration, the work function of the poly-Si can be easily tuned [28].

However, even heavily doped poly-Si gate electrode has a certain resistance which contributes to a considerable RC time delay. Furthermore, when the doping concentration of the poly-Si is not sufficient, at high gate bias during inversion, a region in the poly-Si gate electrode adjacent to the poly-Si/SiO_2 interface becomes depleted with carriers, which is called "Poly-depletion Effect." This results in increased EOT and hence significant decrease in the drive capability of transistor. To improve gate depletion, the number of active dopants in the gate electrode can be increased, but this solution is limited as current doping concentrations are already nearing the equilibrium solid solubility in silicon [27, 28].

For these reasons poly-Si is not suitable for the future gate electrode and new materials are required. Fortunately, metal gate electrode with a suitable work function (WF) shows many advantages such as lower gate sheet resistance and no boron penetration and poly depletion effect. However, the selection of metal gate substitutes for poly-Si, which are compatible with the new high-k gate dielectrics, is not as advanced [29].

2.2.3 Drain/source

1) Parasitic resistance (suppress on-current): As the device gets smaller, the influence of parasitic resistance on on-current increases significantly. Therefore, proper control of parasitic resistance in a MOSFET becomes more important to achieve further performance improvement. The total parasitic resistance, which accounts for the voltage drop between channel and S/D contacts, can be divided into the four components such as overlap resistance, extension resistance, deep resistance, and silicide-

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

97

diffusion contact resistance [30]. Shallower S/D junction depths are desired to suppress SCEs effectively; however, since the shallower junction also increases the sheet resistance, the S/D doping must be increased accordingly to keep the sheet resistance constant. However, as the maximum solid solubility of dopants is limited to ~10^{20} cm^{-3} on the doping density, the series resistance increases with further reduction in junction depth. In addition, the ultra-shallow junctions are difficult to form and also cause significant increase in BTBT leakage current [31]. Furthermore, since the silicide-to-junction contact resistance accounts for nearly half of the total parasitic resistance in state-of-the-art MOSFETs, it is necessary to reduce the contact resistivity to meet the scaling requirement. Contact resistivity can be lowered by increasing the active dopant concentration at the silicide-to-junction interface or reducing the Schottky barrier height in contact region [32].

2) Parasitic capacitance: In digital applications, lower capacitance, along with higher drive current, is the main factor for high performance CMOS circuit design, because the conventional CMOS inverter delay model is given by $C_G V_{DD}/I_D$, where C_G is gate capacitance, V_{DD} is supply voltage, and I_D is drain on-current. Thus, with geometry scaling, the performance improvement has been achieved because the gate capacitance consisting of both intrinsic gate capacitance ($C_{int} = C_{GC} + 2C_{OV}$) and parasitic capacitance ($C_{para} \sim C_{fringe} + C_{gco}$) decreases in proportion to the gate length reduction in micro MOSFET, where C_{GC} is gate-to-channel capacitance, C_{OV} is drain/source-to-gate electrode overlap capacitance, C_{fringe} is inner/outer fringing capacitance, and C_{gco} is gate-to-contact capacitance. However, as MOSFET enters into nano regime, gate capacitance does not decrease in proportion to the gate length reduction due to relatively increased parasitic capacitance. Therefore, in order to sustain performance improvement from scaling, parasitic capacitance reduction technique is required [33].

2.2.4 Substrate (bulk)

Substrate leakage current consists of impact ionization current, reverse-biased junction leakage current (mostly BTBT leakage current), and GIDL current.

1) Reverse-biased junction leakage current (reverse-biased band-to-band tunneling current): Reverse-biased junction leakage current (I_{REV}) is the current flowing between the source/drain (S/D) and the substrate through the parasitic reverse-biased *pn*-junction diode in the off-state MOSFET. I_{REV} mainly consists of the diffusion and drift of minority carriers near the depletion region edge and the generation of electron-hole pairs in the depletion region of the reverse-biased *pn*-junction. The amount of I_{REV} depends on the junction area and doping concentration. If both S/D and substrate regions are heavily doped, BTBT dominates the I_{REV} since the electric field across the junction depletion region increases. If the high electric field (>10^6 V/cm), so that the voltage drop across the junction is bigger than the band gap of silicon, is formed across the reverse-biased junctions of the source/drain (S/D) regions, especially with increasing S/D voltage or reverse body bias, significant amount of BTBT current flows through the S/D to substrate junctions. In nanometer devices, higher channel and S/D doping with shallow junction depths are required to minimize SCEs, there is significant increase in BTBT current [34].

Recently high-mobility channel materials, such as Ge, strained-Ge (s-Ge), s-SiGe, and III-V materials (GaAs, InAs, InSb, and InGaAs), are actively being researched as candidates for future channel materials of highly scaled MOSFETs. However,
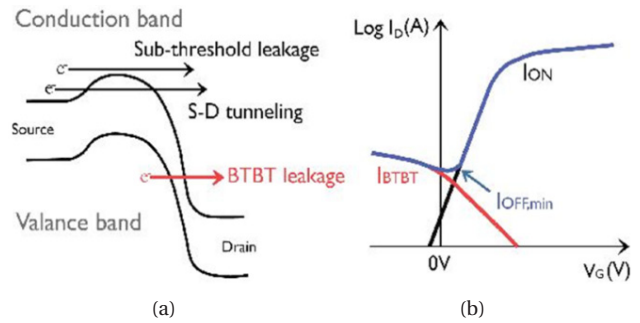

Fig. 6. (a) Various leakage mechanisms in a MOSFET (left), (b) $I_{OFF, min}$ is the minimum achievable leakage current in a MOSFET. In low E_G materials it is generally limited by I_{BTBT}.

since most of high-mobility materials have lower band gaps than Si, BTBT leakage currents increase significantly. Therefore, various device structures to minimize BTBT currents are being developed [35]. The various leakage mechanisms in a MOSFET are illustrated in Fig. 6.

2) GIDL current (Surface BTBT current): GIDL current, also called surface BTBT current, has become one of the major off-state leakage current components in state-of-the-art MOSFETs. When the drain of n-MOSFET is biased at the supply voltage (V_{DD}) and the gate is biased at either zero or negative voltage, a depletion region is formed under the gate and drain overlap region. In the same way as the BTBT current, if the high electric field is formed in the narrower depletion region as a result of the reverse-bias between channel and drain, a significant amount of surface BTBT current flows through drain to substrate junctions due to twisting of bandgaps [36, 37]. With higher supply voltage, thinner oxide thickness, LDD, reverse body bias (RBB) technique, and high mobility channel materials having smaller bandgaps enhance the GIDL current.

2.3 Summary

As the device scaling is approaching its physical size limitations, the technology cycle is getting slow down due to increasing power consumption, process variation, and fabrication cost. To continue the historical growth for the next decade, various studies have been conducted to overcome ever increasing challenges. As explained in section 2.1, nowadays device scaling tradeoffs between performance and power consumption, therefore technological innovations which can achieve high performance through very low power are required. These efforts have made conventional MOSFETs evolve in various ways with new materials, new structures, and so on. In the next section, the evolution of MOSFET structures is reviewed.

3. EVOLUTION OF MOSFET

The conventional device scaling trend, which is mostly based on reducing its physical size, is no longer valid for the future technology generations. Therefore, several strategies have been introduced to extend Moore's Law, such as increasing electrostatic control over the channel by means of the continued EOT scaling with high-k/metal gate stack and multi gate structures for higher drive current at constant over-drive voltage (V_{DD}-V_{th}), improving mobility of carriers by adopting high mobility channel materials and strain engineering, reducing parasitics, and so on [38].
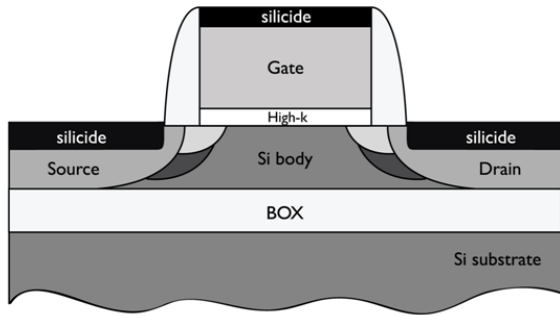
98

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

Fig. 7. Cross-section of a SOI MOSFET.

3.1 Silicon on insulator

The advantages of SOI technology come from its buried oxide (BOX) layer (a cross-section of a SOI device is shown in Fig. 7). With the reduction of the parasitic capacitances, mostly as a result of the reduced drain/source junction capacitances, SOI devices yield improved switching speed and reduced power consumption. The operating speed is also improved since the isolated channel from substrate bias prevents the increase in a threshold voltage of stacked SOI transistors. In addition, the perfect lateral and vertical isolation from substrate provides latch-up and inter-device leakage free CMOS technology, reduction in various interferences, and better soft error immunity. Moreover, SOI technology offers tighter transistor packing density and simplified processing [39].

SOI transistors are classified into two types; "partially depleted (PD) SOI," if the silicon film (typically 100 nm or more) on the BOX layer is thicker than the depletion region depth beneath the gate oxide, and "fully depleted (FD) SOI," if the body (silicon film) thickness is thin enough (typically 50 nm or less) or the doping concentration of the body is low enough to be fully depleted. FD SOI transistors have superior advantages over PD SOI transistors in terms of extremely low sub-threshold swing (<65 mV/decade), no floating-body effects, and low threshold voltage variation with temperature (2-3 times less). However, since FD SOI transistors are even more sensitive to process variation such as the silicon film layer variation resulting in threshold voltage fluctuation, PD SOI devices were commercially introduced first. With careful device design and advanced process techniques, fully depleted ultrathin-body SOI (FD UTB SOI) devices are considered as one of the best scaling options. Undoped or very lightly doped UTB devices minimize impurity scattering and reduce threshold voltage variation resulting from random dopant fluctuation [40].

Another important merit of SOI technology is that it provides the cornerstone for new device structures such as multi gate field-effect transistors (MuGFETs), which includes more than one gate into a single device. The double-gate (DG) FET shown in Fig. 8(b) is the first step for those multi gate devices. Since the transverse electric field induced by V_{DS} is shared by both top and bottom channels, it mitigates SCEs [38, 41]. However, since the gate-channel-gate stacked structure is not suitable for Si-planar technology to fabricate, the complex fabrication processes are required to form DG FETs. These problems are considerably eased with a self-aligned FinFET (Fig. 8(c)) [42]. In addition to DG FET and FinFET, tri-gate, gate-all-around (GAA) (or surround-gate), and nanowire FET structures (Figs. 8(b)-(f)) show the better drive current at fixed V_{DD} in order. Especially, a thin nanowire channel FET, shown in Fig. 8(f), provides the most effective electrostatic characteristics via the gate bias.

On the other hand there is a significant drawback in SOI technology. Since the BOX, which has approximately 100 times lower thermal conductivity than that of silicon, prevents thermal con-
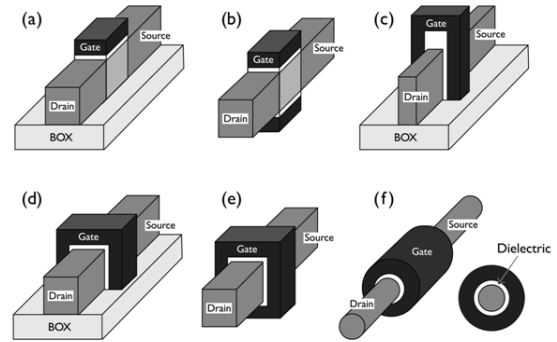

Fig. 8. Various SOI device: (a) Single gate SOI transistor, (b) double gate planar SOI transistor, (c) double gatenon-planar FinFET, (d) tri-gate FET, (e) quadruple-gate (or gate-all-around) FET, and gate-all-around (or surrounding gate) FET (nanowire FET).
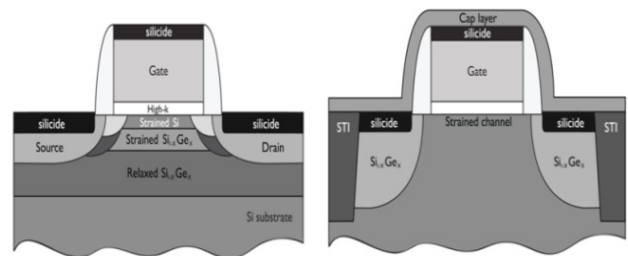

Fig. 9. MOSFET with a strained-silicon channel.

duction path from SOI transistors to the substrate, SOI transistors are easily affected by the thermal heating generated in the channel, which is called "Self-Heating Effects." Consequently, the self-heating degrades the mobility of carriers and causes the threshold voltage shift. These effects get worse with FD structures because they use thinner silicon films [43].

3.2 Stress engineering

Mobility loss resulted from higher channel doping and scaled gate dielectrics should be compensated to meet the performance targets of the future technologies. In the parallel with SOI technology, there has been more straight forward and cost effective way to improve device performance and scalability. That is mobility-enhancement technology. Both the high mobility channel materials (such as Ge and GaAs) and the strained channel by means of stress offer mobility enhancement. Various CMOS fabrication processes can be used to induce appropriate strain to the channel region of the MOSFETs. Strain engineering can be largely divided into two: local strain and global strain. Local strain exploits the process steps such as shallow trench isolation (STI), silicidation at the source/drain (S/D) region, and nitride CESL (Contact Etch Stop Liners). The local strain techniques were first used because they were more promising for manufacturing. However, since they have strong device geometry dependence, it makes their scaling behavior less predictable. The most innovative way on strained-Si was focused on biaxial global strain generated by depositing a thin layer of silicon on a relaxed SiGe virtual substrate. This results in enhanced carrier transport in the strained Si layer, and mobility enhancements of 110% for electrons and 45% for holes. A cross-section of transistors with strained Si-channel fabricated by means of both the global and local strains is shown in Fig. 9 [44-46].

Appropriately applied strain to the MOSFET channel can significantly improve the device performance by means of modify-

ing energy bands and increasing mobility of carriers in silicon. Strain is being widely used in HP logic technologies. SiGe in the PMOS S/D regions, stress memorization technique (SMT) for the NMOS, and a nitride stress-capping layer for both devices have been commonly employed in state-of-the-art MOSFETs since its introduction at the 90nm node. Strain is expected to continue CMOS technology scaling for the time being [47].

3.3 Possible solutions after CMOS scaling limits

As dimensional scaling of CMOS transistors is reaching their fundamental physical limits, various researches have been actively carried out to find an alternative way to continue to follow Moore's law. Among these efforts, various kinds of alternative memory and logic devices, so called "Beyond CMOS Devices," have been proposed. These nanodevices take advantage of the quantum mechanical phenomena and ballistic transport characteristics under lower supply voltage and hence low power consumption. Those devices are expected to be used for ultra high density integrated electronic computers due to their extremely small size. In the following sections, the general overview will be discussed for the promising emerging logic nanodevices such as nanowire (NW) transistors, carbon nanotube field-effect transistors (CNTFETs), graphene nanoribbon (GNR) transistors, single electron transistors (SETs), and quantum-dot cellular automata (QCA) in terms of their basic operation principles, current state of development, and challenges for commercialization.

4. PROMISING NANODEVICES BEYOND CMOS

There are a number of nanoelectronic devices that can possibly supplement or replace current CMOS devices. These include nanowire or carbon nanotube transistors, graphene FETs, single electron transistors, and spin transistors. These devices offer sizes of a few nanometers and can be self-assembled.

4.1 Nanowire (NW) transistors

Nanowire field-effect transistors (NWFETs), gate-all-around or surround gate FETs with a thin nanowire channel, have drawn much attention and have been considered as promising candidates for continuous CMOS scaling since their nonplanar geometry provides superior electrostatic control of the channel than the conventional planar structures. The increasing attention in nanowire research stems from several key factors: their high-yield reproducible electronic properties [48-50], cost-effective "bottom-up" fabrication which circumvents some fabrication challenges, higher carrier mobility by means of the reduction of scattering resulting from the crystalline structure, smooth surfaces and the ability to produce radial and axial nanowire heterostructures [51, 52], better scalability resulting from the fact that diameter of nanowires can be controlled down to well below 10 nm [53, 54].

However, due to their smaller nanowire diameters, the inversion charge changes from surface inversion to bulk inversion due to quantum confinement. Thus, variations in nanowire dimensions due to fabrication imperfections can lead to perturbations in the carrier potential and scattering that degrade the charge transport characteristics. Also, variations in nanowire diameters may lead to a variation in FET threshold voltage. Reducing variability is therefore a key challenge in making nanowire FETs a viable technology. Furthermore, quantum confinement effects make modeling of nanowire transistors a complex problem. While there are many studies in the literature on the modeling of

nanowire transistors based on nonequilibrium Green function or Monte Carlo approaches, the physics related to the operation of nanowire transistors needs to be well articulated so that simple compact models, including ballistic transport and realistic subband parameters, can be developed for circuit design using SPICE-like simulators [55].

4.2 Carbon nanotube field-effect transistors (CNTFETs)

Since they were discovered in 1991 [56], due to their unique material properties [57], carbon nanotubes (CNTs) have received worldwide attention from many different fields of scientists and engineers.

CNTs are graphene, which is a two-dimensional honeycomb lattice of carbon atoms, sheets rolled up into cylinders. They show either metallic or semiconducting properties depending on the direction how CNT are rolled up (chirality). Since the band-gap of semiconducting CNTs is inversely proportional to their diameters, threshold voltage can be easily controlled [58-60].

With their superior material properties, such as large current carrying capacity, the excellent mechanical and thermal stability, and high thermal conductivity, the metallic nanotubes are attractive as future interconnects [61]. Along with these properties, the semiconducting nanotubes also show great advantages as a channel material of high-performance (HP) FETs. It is easier for CNTFETs to incorporate high-k dielectrics due to the absence of dangling bonds. Also, since both NMOS and PMOS transistors show almost identical I-V characteristics, it becomes a significant advantage for CMOS circuit design. Furthermore, they are very attractive to Si-based semiconductor industry for the following reasons: 1) CNTFETs show considerable improvement in device performance metrics such as low power and high speed 2) their operating principles and devices structure are similar to Si-based CMOS transistors; therefore, the CMOS design infrastructure could be reused [62].

Even though CNT-based electronics presents great promise, there are several difficult challenges to be solved. No existing processes for synthesis or growth of nanotubes can produce tubes with identical diameters and chiralities. The starting material will have to be purified if far more selective growth processes cannot be developed. Many progresses are being reported in purification. However, there is still a big gap between these promising early results and the extremely tight control of placement that would be needed. Furthermore, an even tighter gate control through ultrathin high-k gate dielectrics and extremely abrupt doping profiles are required for the device fabrication [63].

4.3 Graphene nanoribbon (GNR) transistor

Graphene, a single atomic layer of carbon atoms arranged into a two-dimensional (2D) hexagonal lattice, or simply thought as an unrolled CNT, shows great potential for nanoelectronic devices since its discovery in 2004 [64]. It shares many of the advantages of metallic CNTs such as the high carrier mobility for ballistic transport, high carrier velocity for fast switching, excellent thermal conductivity, and etc. The potential to produce wafer-scale graphene films with full planar processing for devices promises high integration potential with conventional CMOS fabrication processes, presenting a significant advantage over CNTs [65].

Two-dimensional graphene has a zero band-gap (E_G) semi-metal characteristic. However, by patterning graphene into a few nanometer wide GNR, a band-gap can be induced [66]. Similar to the CNT cases, GNRs have energy band-gap which is inversely proportional to their width, and width confinement down to the
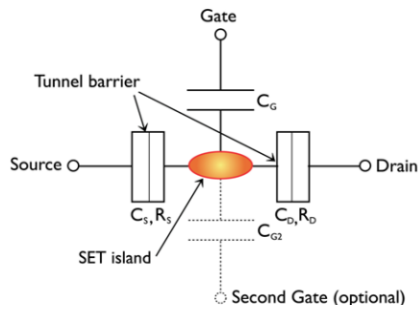
100

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

Fig. 10. Schematic of a basic Single electron transistor.

Fig. 11. Four dot QCA cells.

sub-10 nm scale is essential to open a band-gap that is sufficient for room temperature transistor operation. Unlike CNTs, which are mixtures of metallic and semiconducting materials, recent samples of chemically derived sub-10 nm GNRs have shown all-semiconducting behavior [67]. GNRs, under the assumption that the edges of ribbons are passivated by hydrogen atoms, have two main types of armchair-edge and zigzag-edge GNRs (AGNRs and ZGNRs). ZGNRs are predicted to be metallic by a simple tight-binding model, but a band-gap exists in more advanced and spin-unrestricted simulations [68]. For digital circuit applications, the focus has been on using AGNRs as the channel material. AGNRs have an electronic structure that is closely related to that of zigzag CNTs. The band-gap in AGNRs originates from quantum confinement, and edge effects play a critical role [68, 69].

Due to the atomically thin and nanometer-wide geometries of GNRs, variability and defects are projected to have a larger impact on GNRFET circuit performance and reliability in comparison to the conventional silicon devices. Variability, for example, can arise from the difficulty of control of the GNR width, edge roughness, or oxide thickness during fabrication. Defects may occur during fabrication due to a charge impurity in the gate insulator or a lattice vacancy, and result in a large performance variation or a nonfunctional device. Not only must each of the variability and defect mechanisms be identified and studied, the models developed for GNRs and GNRFETs must be capable of predicting their effects in isolation as well as together in a systematic manner [69].

Although the GNR material promises ultra-small, fast, and low-energy FETs, two key effects of variability and defects - leakage and low noise margins - are significant. For example, the variation of the channel width by a couple of angstrom changes the leakage current by orders of magnitude, and a single Coulomb charge impurity can lower the FET on-current by about 30%. Dense memories, which are the biggest prospect for graphene-based devices, are particularly susceptible to variations and defects with near-zero noise margins and an increase in leakage power of over 5 X. This assessment of the effects of variability, defects, and parasitics indicate their important role on circuit performance. These effects must be carefully considered in the performance assessment and design optimization for future graphene-based electronics technology [70].

4.4 Single-electron transistors (SETs)

Due to their small size and low-power dissipation at good speed, SETs are very attractive devices for future large-scale integration. The basic structure of SET consists of three-terminals such as drain, gate, source, and the second gate is an optional. A schematic of SET, as shown in Fig. 10, is analogous to that of conventional MOSFETs. However, SET has a tiny conductive island coupled to a gate electrode with gate capacitance C_g. Source and
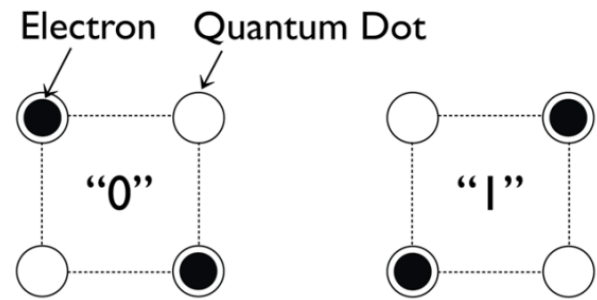
drain electrodes are connected to the island through a tunnel barrier (junction). The tunnel barrier, which controls the motion of every single electron, consists of two conductors separated by thin layer and it is modeled as tunneling resistances $R_{D, S}$ and junction capacitances $C_{D, S}$. The increased gate bias attracts electrons to the island only through either drain or source tunnel barrier, and the number of electrons in the island only has a fixed integer. Therefore, the increased gate bias makes electrons flow one by one when a small voltage is applied between the source and drain electrodes by means of the "Coulomb blockade" phenomenon [71-73].

New applications and architectures that exploit the unique functionality of room temperature operating SET circuits have been developed, especially by monolithic integration of SETs with FET circuits to complement the conventional Si CMOS performance. Representative examples include SET/CMOS hybrid multi-value logic circuits [74], multiband filtering circuits [75], analog pattern matching circuits (discussed more fully in the next section) [76], associative recognition tasks [77], and others [78], in which characteristic Coulomb blockade oscillations of SETs are typically utilized to reduce the number of devices. Note that certain aspects of the circuit performance, especially the room temperature operation [76, 79], already exceed the theoretical evaluation of the logic gate parameters for 2 nm SETs. These devices have a theoretically estimated maximum operation temperature T~20 K, integration density n~10^{11} cm^{-2}, and speed of the order of 1 GHz [80]. However, large threshold voltage variation continues to impede the realization of large scale SET circuits, making it difficult for SETs to compete directly with CMOS devices used to implement Boolean logic operations. Engineering breakthroughs are needed to eliminate the size and background charge fluctuations in order to suppress the threshold voltage variations.

The majority of the SET circuits demonstrated to date employ so called "voltage state logic," where a bit is represented by the voltage of capacitor charged by many electrons. The problem of the low fan-out for this scheme can be overcome by reducing the capacitance and/or by combining with conventional FET circuits. Truly single-electron approaches, representing a bit by a single electron ("bit state logic") [81] and the use of a single electron as a source of random number generations [82], have been limited to laboratory demonstrations. The problem of the limited fan-out, which is caused by using only a single electron in the truly single-electron devices, may be solved by innovative circuit designs such as the binary-decision-diagram [83].

4.5 Quantum-dot cellular automata (QCA)

As the number of transistors placed onto a single chip increases, the heat generated during a switching cycle can no longer be removed and this may result in considerable limitation of the speed of operation. There is a new paradigm which takes

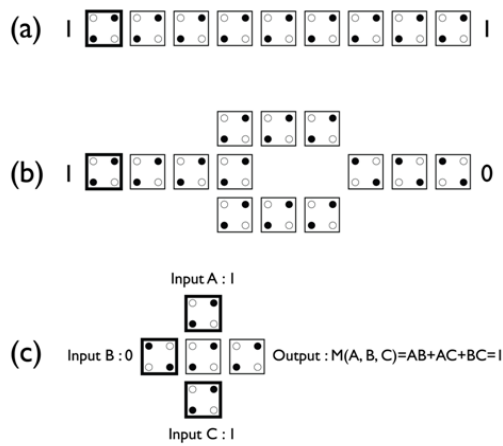Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

101

Fig. 12. Fundamental QCA devices (a) binary wire, (b) QCA inverter, (c) majority logic gate.

advantage of the reduction of the feature size. One nanostructure paradigm, proposed by Lent et al. [84-86], is QCA, which exploits arrays of coupled quantum dots to implement Boolean logic functions. The advantage of QCA lies in the extremely high packing densities possible due to the small size of the dots, the simplified interconnection, and the extremely low power-delay product [87].

A schematic diagram of a four-dot QCA cell is shown in Fig. 11. This is the simplest non-clocked QCA cell. The cell consists of four quantum dots positioned at the comers of a square. The cell contains two extra mobile electrons, which are allowed to tunnel between neighboring sites of the cell, but not out of the cell. If the tunnel barriers between cells are sufficiently high, the electrons will be well localized on individual dots. The Coulomb repulsion between the electrons makes them occupy antipodal sites in the square as shown. For an isolated cell, there are two energetically equivalent arrangements, polarizations, of the extra electrons that we can denote as binary 1 and binary 0. The two polarization states of the cell will not be energetically equivalent if other cells are nearby, because the Coulomb interaction with other cells breaks the degeneracy [87].

A QCA wire is shown in Fig. 12(a). The left-most cell is fixed with polarization representing the input. The difference between input and outputs cells in QCA arrays is simply that inputs are fixed and outputs are free to change. The ground state configuration of the remaining free cells is the one with each cell polarized in the same way as the input cell. We can consider this transmission of the input signal from one end to the other. Cells that are positioned diagonally from each other tend to anti-align. This behavior is employed to construct an inverter as shown in Fig. 12(b). Fig. 12(c) shows the fundamental QCA logical device, a three-input majority gate, from which more complex circuits can be built. The central cell, labeled the device cell, has three fixed inputs labeled A, B, and C. The device cell has its lowest energy state if it assumes the polarization of the majority of the three input cells. It is possible to "reduce" a majority logic gate by fixing one of its three inputs in the 1 or 0 state. In this way, a reduced majority logic gate can also serve as a programmable AND/OR gate. Combined with the inverter shown above, this AND/OR functionality ensures that QCA devices provide logical completeness [87].

For practical QCA circuits, clocked control is necessary. In clocked circuits, the system is switched smoothly between the ground state corresponding to the old inputs and the state corresponding to the new inputs. Clocking allows cells to be locked in a particular state, performing as a latch. Thus general purpose computing is possible, and in particular, latches make it possible to implement pipelined architectures as required in all high-performance systems [88].

A severe problem of QCA arises from the fact that QCA are single-electron devices and they are very sensitive to the background charge. Today, no viable solutions to the background charge immune single-electron systems are known. Another serious drawback of QCA devices is that room-temperature operation is not achievable with solid-state QCA systems [79].

5. TECHNICAL CHALLENGES FOR ANOELECTRONICS

While nanoelectronics presents the opportunity to incorporate billions of devices into a single chip, it also increases defects and variations both during manufacture and chip operations. In this chapter, technical challenges are summarized for nanoelectronics technology ranging from device to system architecture level.

5.1 Devices

Due to their extremely small size (tens of nanometers and below), quantum-mechanical effects such as ballistic transport, tunneling effects and quantum interference start to predominate in nanodevice operation. Therefore, good understanding of quantum-mechanical phenomena is required to exploit the properties of nanodevices.

Emerging Research Device (ERD) Working Group evaluated nanodevices in terms of nine relevance criteria; scalability, performance, energy efficiency, on/off ratio for memory devices, gain for logic devices, operational reliability, operation temperature, CMOS technology compatibility, and CMOS architectural compatibility. For logic applications, since the number of transistors in a chip is limited not by their size but by the maximum allowable power dissipation (~100 W/cm^2), 1D structure devices such as NW transistors and CNTFETs are promising because they show higher I_{ON}/I_{OFF} ratio rate at lower supply voltage. In spite of their higher risk, channel replacement materials also have potential. While both of those charge-based channel replacement logic technologies show better device characteristics than other devices using computational state variables, state variable devices are expected to give us a solution for limitations of charge-based technology in long-term future. Due to their lacks of reliability, it is not likely for revolutionary nanoelectronic devices to replace CMOS technology entirely without compulsory needs for those devices. However, if used to complement ultimately scaled CMOS, they might present more possible ways to extend the functionality or reduce the power dissipation with their unique properties [89].

5.2 Interconnect

Suitable interconnect technology must be developed based on new device concepts. In order to connect extremely small devices, interconnects are required to be less than 10 nm in diameter. As the diameters of metallic interconnect wires scale down to the mean free path of electrons, both the surface scattering from the boundaries of ultra-narrow conductors and the grain boundary scattering would impede electronic conduction in the wires [90]. As a result, the circuit will be slower and it is not possible to achieve performance improvement from scaling. Furthermore, electromigration resulting from movement of lots of electrons and electron scattering makes another serious problem. Therefore, various potential interconnect technologies ranging from nanowires to carbon nanotubes and quantum wire have been

102

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

extensively researched in order to satisfy the requirements such as low resistivity, large current carrying capacity, easy of fabrication, and isolation with low-k dielectric materials for ultra-high density applications [91].

5.3 Architectures

Nanoarchitecture can be defined as the organization of basic computational structures composed of nanodevices assembled into a system that computes something useful. The design trends of nanoelectronics also can be summarized as increased emphasis on parallelism, strict constraints on localized interconnection, regularity in nanofabrics, and repair of faulty circuits' offline and possibly online. Particularly, the high unreliability of nanodevices is the most critical problem in constructing nanoelectronic systems. The defect prone fabrication process, especially from bottom-up self-assembly process, increases manufacturing defects. At the same time, the high occurrence of transient faults and variations results in higher dynamic fault occurrences as well. Therefore, following issues are needed to be investigated and solved:

(1) How to use huge numbers of nanodevices in an effective way, (2) How to deal with the offline and online errors, (3) How to ensure system reliability without performance degradation, and (4) How to maintain power dissipation in tolerable level [92]?

Even after the four fundamental architectural questions are answered, a few fundamental efforts should be made in order to exploit benefits of nanotechnology and minimize its drawbacks in architecture level. The efforts include mapping more algorithms into parallel forms, finding alternative computation and information representation models, and designing fault tolerant, reconfigurable, and power efficient systems [93].

5.4 Circuit design

To design circuit, first of all, we must be able to implement all necessary primitive logic functions and latches out of nanoelectronic devices. Since most of the bottom-up fabrication techniques are only feasible for non-restorable two-terminal devices as of today, one of the key design challenges is to provide signal restoration.

Besides it is necessary to bridge the dimension gap between the nanometer-scale pitch devices and the lithographic-scale devices without nullifying the benefits from nanoelectronic systems in order to provide an interface between the nanodevice/nanowire and CMOS sub-systems.

Moreover, variability and reliability problem are expected to be the major concern in nanoelectronic circuit design during manufacture and chip operations in the same way as the conventional top-down CMOS technology has encountered. Furthermore, small signal levels cause lower signal-to-noise margins due to reduced supply voltage. Because these problems are even worsened with increasing thermal fluctuations and crosstalks, nanoelectronic devices will be much more unreliable. Due their molecular sizes, there will be always a non-trivial probability that individual contacts may be inadequate or devices are missing. Along with the doping and topology variations, this may show up as either variations in device parameters or even a complete lack of connection at some crosspoints. In addition to manufacturing defects, nanodevices may be susceptible to persistent changes over their operational lifetime due to ionizing-particle strikes, aging, movement of individual atoms, and temperature-induced variations. Furthermore, with the state stored in small numbers of electrons, ionizing particle strikes and statistical switching effects can induce a high rate of soft errors.

Therefore, new novel fault and defect tolerance circuit design

technique and methodologies have to be developed for nano-electronic low-power, high-speed and high-density systems to make the circuit reliable out of the unreliable components (devices and wires). These can be achieved with designing circuit with very regular structures rather than arbitrary geometries and [93, 94].

5.5 CAD

The primary purpose of computer-aid design (CAD) tools is to offer an effective means for designing and fabricating electronic circuits in various ways. It comprises modeling of process steps, behavior of the electronic devices and physical systems on computers, providing the analysis of design variants, changing the expression of designs in a form suitable for manufacturing, and so on. They also can execute many functions, including many forms of verification (geometrical and electrical design rule checking, several types of simulation) and various aspects of design synthesis (generation of masks from layout, automatic placement and routing, logic synthesis and optimization) [95-97].

As the designs increase in size and complexity, the importance of CAD tools are being enhanced. Since nanoelectronics uses a higher level of device integration than their CMOS counterparts, it will be required to revise or reinvent many existing synthesis, physical design, and verification methodologies and techniques to handle complex nanoscale designs. Due to the huge number of nanodevices in integrated circuits, it is necessary to raise the level of design entry and abstraction from register-transfer level to the architecture and system level in order to manage design complexity and increase design productivity with increasing defects and fault rates. Therefore, new design flows and tools are needed to optimize nanoelectronic circuits for better performance and yield. The most urgent step is to develop suitable tool sets that can address the nanofabrication processes and unique features of nanotechnology devices and fabrics. The next required step for circuit integration is to develop design tool sets that can optimize the circuits having error resilience with the aid of statistical design techniques. Those tools should be able to analyze and quantify the key sources of nanodevice and nanowire parameter variations and defects [93, 94].

5.6 Fabrication

One of the grand challenges for nanoelectronics is to establish a new manufacturing paradigm which enables the continuation of the cost-per-function reduction well down to several nanometers. Conventional lithography technique has been a main enabler for large-scale CMOS integration. However, this increases process and mask complexity as the technology generation advances below 100 nm node, significantly increasing manufacturing cost. Therefore, a nanofabrication technique which can control and produce billions of nanodevices in a quick, reliable and cost-effective way is a prerequisite to nanaoelectronics manufacturing.

Largely, there are two different fabrication methods for nanoelectronic circuits. The top-down fabrication refers to the technique employing many steps of sophisticated lithographic and etching techniques in order to pattern a substrate. On the other hand, in the bottom-up approach, individual devices and wires are manufactured first from nanomaterials, and then assembled into a circuit. This bottom-up technique includes gas-phase methods such as chemical vapor deposition (CVD), metal-organic chemical vapor deposition (MOCVD), atomic layer deposition (ALD), liquid phase methods, and etc. Although the top-down approach is dominant at present time, it is expected that bottom-up approach will play an important role in near

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

103

future since many of those methods are both cost-effective and offer high throughput.

While bottom-up self-assembly presents some advantages, the drawback of this is the difficulty on generating arbitrary patterns which can be formed using photolithography. Furthermore, due to increased defect densities, self-assembled circuits and architectures will have to be designed for defect tolerance [98].

6. CONCLUSIONS

As dimensional scaling of CMOS transistors is reaching its fundamental limits, various researches have been actively carried out to find an alternative way to continue to follow Moore's law. Among these efforts, various kinds of alternative memory and logic devices, so called "beyond CMOS" devices, have been proposed. These nanodevices take advantage of the quantum mechanical phenomena and ballistic transport characteristics under lower supply voltage and hence low power consumption. Moreover, due to their extremely small size, those devices are expected to be used for ultra-density integrated electronic computers. While nanoelectronics presents the opportunity to incorporate billions of devices into a single system, it also increases defects and variations both during manufacturing and chip operations. Therefore, an additional constraint, 'reliability,' has to be added to the conventional low-power, high-speed and high-density design consideration. Due to this additional constraint, it is necessary to develop a new knowledge and reliability paradigm for nanoelectronics in order to enable industries to predict, optimize and design upfront reliability and performance of nanoelectronics. The new paradigm should be in a multi dimensional space covering devices, interconnects, architecture, circuit design, CAD, and fabrication issues.

REFERENCES

[1] H. Iwai, Extended Abstracts 2008 8th International Workshop on Junction Technology (IWJT '08) (Shanghai, China 2008 May 15-16, IEEE Press) p. 1. [DOI: 10.1109/IWJT.2008.4540004].

[2] G. E. Moore, Electronics **38**, (1965).

[3] G. E. Moore, International Electron Devices Meeting. Technical Digest (Washington, DC 1975 Dec. 1-3, IEEE Group on Electron Devices) p. 11.

[4] International Technology Roadmap for Semiconductors (ITRS) 2007 Edition. Available from: http://www.itrs.net/links/2007ITRS/Home2007.htm.

[5] D. Kahng and M. M. Atalla, the IRE Solid-State Device Research Conference (Pittsburgh, PA 1960 Jun., Carnegie Institute of Technology).

[6] R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, IEEE J. Solid-State Circuits **SC-9**, 256 (1974).

[7] D. A. Antoniadis, I. Aberg, C. Ní Chleirigh, O. M. Nayfeh, A. Khakifirooz, and J. L. Hoyt, IBM J. Res. Dev. **50**, 363 (2006) [DOI: 10.1147/rd.504.0363].

[8] M. Horowitz, E. Alon, D. Patil, S. Naffziger, R. Kumar, and K. Bernstein, IEEE International Electron Devices Meeting. IEDM Technical Digest (Washington, DC 2005 Dec. 5-7, IEEE Group on Electron Devices) p. 7. [DOI: 10.1109/IEDM.2005.1609253].

[9] H. S. P. Wong, D. J. Frank, P. M. Solomon, C. H. J. Wann, and J. J. Welser, Proc. IEEE **87**, 537 (1999) [DOI: 10.1109/5.752515].

[10] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H. S. P. Wong, Proc. IEEE **89**, 259 (2001) [DOI: 10.1109/5.915374].

[11] B. Razavi, *Design of Analog CMOS Integrated Circuits* (McGraw-Hill, Boston, MA, 2001).

[12] M. Stockinger, *Optimization of Ultra-Low-Power CMOS Transistors*, Ph.D. dissertation (Vienna, Austria 2000, Institute for Microelectronics).

[13] R. R. Troutman, IEEE J. Solid-State Circuits **14**, 383 (1979).

[14] I. M. Bateman, G. A. Armstrong, and J. A. Magowan, 19th International Electron Devices Meeting. Technical Digest (Washington, DC 1973 Dec. 3-5, IEEE Group on Electron Devices) p. 147.

[15] S. Ogura, P. J. Tsang, W. W. Walker, D. L. Critchlow, and J. F. Shepard, IEEE J. Solid-State Circuits **15**, 424 (1980).

[16] S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era, Vol. 3: The Submicron MOSFET* (Lattice Press, Sunset Beach, CA, 1986).

[17] I. De and C. M. Osburn, IEEE Trans Electron Devices 46, 1711 (1999) [DOI: 10.1109/16.777161].

[18] T. Mizuno, J.-I. Okamura, and A. Toriumi, IEEE Trans. Electron Devices **41**, 2216 (1994) [DOI: 10.1109/16.333844].

[19] A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies, and S. Saini, IEEE Trans. Electron Devices **48**, 722 (2001) [DOI: 10.1109/16.915703].

[20] Y. Ye, F. Liu, S. Nassif, and Y. Cao, Proceedings of the 45th Annual Design Automation Conference (Anaheim, CA 2008 Jun. 8-13, ACM/IEEE) p. 900. [DOI: 10.1145/1391469.1391698].

[21] R. F. Pierret, *Semiconductor Device Fundamentals* (Addison-Wesley, Reading, MA, 1996) p. 691.

[22] B. L. Anderson and R. L. Anderson, *Fundamentals of Semiconductor Devices* (McGraw-Hill Higher Education, Boston, 2005) p. 124, p. 425.

[23] P. A. Gargini, International Symposium on VLSI Technology Systems and Applications (VLSI-TSA) (Hsinchu 2008 Apr. 21-23, IEEE) p. 10. [DOI: 10.1109/VTSA.2008.4530775].

[24] V. V. Zhirnov, R. K. Cavin Iii, J. A. Hutchby, and G. I. Bourianoff, Proc. IEEE **91**, 1934 (2003) [DOI: 10.1109/JPROC.2003.818324].

[25] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers, V. Naydenov, T. Khondker, S. Sarkar, and P. Singh, 2007 IEEE Asian Solid-State Circuits Conference (A-SSCC) (Jeju, Korea 2007 Nov. 12-14, IEEE) p. 14. [DOI: 10.1109/AS-SCC.2007.4425784].

[26] P. J. Wright and K. C. Saraswat, IEEE Trans. Electron Devices **37**, 1884 (1990) [DOI: 10.1109/16.57140].

[27] G. D. Wilk, R. M. Wallace, and J. M. Anthony, J. Appl. Phys. **89**, 5243 (2001) [DOI: 10.1063/1.1361065].

[28] G. M. T. Wong, *An Investigation of the Work Function of Metal Gate Electrodes for Advanced CMOS Applications, Ph.D. dissertation* (Palo Alto, CA 2008, Stanford University).

[29] K. S. Chang, M. L. Green, J. R. Hattrick-Simpers, I. Takeuchi, J. S. Suehle, O. Celik, and S. De Gendt, IEEE Trans. Electron Devices **55**, 2641 (2008) [DOI: 10.1109/TED.2008.2003091].

[30] S. D. Kim, C. M. Park, and J. C. S. Woo, IEEE Trans. Electron Devices **49**, 457 (2002) [DOI: 10.1109/16.987117].

[31] T. Krishnamohan, *Physics and Technology of High Mobility, Strained Germanium Channel, Heterostructure MOSFETs*, Ph.D. dissertation (Palo Alto, CA 2006, Stanford University).

[32] M. Ieong, IEEE Nanotechnology Materials and Devices Conference (NMDC) (Gyeongju, Korea 2006 Oct. 22-25, IEEE) p. 88. [DOI: 10.1109/NMDC.2006.4388702].

[33] M. C. Chang, C. S. Chang, C. P. Chao, K. I. Goto, M. Ieong, L. C. Lu, and C. H. Diaz, IEEE Trans. Electron Devices **55**, 84 (2008) [DOI: 10.1109/TED.2007.911348].

[34] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, Proc. IEEE **91**, 305 (2003) [DOI: 10.1109/JPROC.2002.808156].

[35] K. C. Saraswat, International Symposium on VLSI Technology Systems and Applications (VLSI-TSA) (Hsinchu 2007 Apr. 23-25, IEEE) p. 1. [DOI: 10.1109/VTSA.2007.378944].

[36] S. A. Parke, J. E. Moon, H.-j. C. Wann, P. K. Ko, and C. Hu, IEEE Trans. Electron Devices **39**, 1694 (1992) [DOI:

10.1109/16.141236].

[37]　Y. Xiaobin, P. Jae-Eun, W. Jing, Z. Enhai, D. Ahlgren, T. Hook, Y. Jun, V. Chan, S. Huiling, L. Chu-Hsin, R. Lindsay, P. Sungjoon, and C. Hyotae, IEEE International Integrated Reliability Workshop Final Report (IRW 2007) (South Lake Tahoe, CA 2007 Oct. 15-18, IEEE) p. 70. [DOI: 10.1109/IRWS.2007.4469224].

[38]　L. Chang, Y. K. Choi, D. Ha, P. Ranade, S. Xiong, J. Bokor, C. Hu, and T. J. King, Proc. IEEE **91**, 1860 (2003) [DOI: 10.1109/JPROC.2003.818336].

[39]　T. Sakurai, A. Matsuzawa, and T. Douseki, *Fully-Depleted SOI CMOS Circuits and Technology for Ultra-Low Power Applications* (Springer, Dordrecht, The Netherlands, 2006).

[40]　R. Simonton, *Special Report SOI Wafer Technology for CMOS ICs [Electronic Document]* (Simonton Associates, 2002) Available from: http://www.icknowledge.com/threshold_simonton/soi-technology.pdf.

[41]　A. Jakubowski and L. LUkasiak, Mater. Sci. -Poland **26**, 5 (2008).

[42]　D. Hisamoto, W. C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T. J. King, F. Jeffrey Bokor, and C. Hu, IEEE Trans. Electron Devices **47**, 2320 (2000).

[43]　A. K. Sharma and A. Teverovsky, *Reliability Evaluation of Fully Depleted SOI (FDSOI) Technology for Space Applications [Electronic Document]* (NASA Electronic Parts and Packaging (NEPP) Program, 2001). Available from: http://nepp.nasa.gov/docuploads/f8b88988-9a2d-462b-986eb801f50978a9/eval_fdsoiparti_neppfinalreport.pdf.

[44]　N. Mohta and S. E. Thompson, IEEE Circuits Devices Mag. **21**, 18 (2005) [DOI: 10.1109/MCD.2005.1517386].

[45]　W. Chee, S. Maikop, and C. Y. Yu, IEEE Circuits Devices Mag. **21**, 21 (2005) [DOI: 10.1109/MCD.2005.1438752].

[46]　C. K. Maiti, International Workshop on Physics of Semiconductor Devices (IWPSD 2007) (Mumbai, India 2007 Dec. 16-20, IEEE) p. 52. [DOI: 10.1109/IWPSD.2007.4472453].

[47]　C. Auth, M. Buehler, A. Cappellani, C.-h. Choi, G. Ding, W. Han, S. Joshi, B. McIntyre, M. Prince, P. Ranade, J. Sandford, and C. Thomas, Intel Tech. J. **12**, 77 (2008) [DOI: 10.1535/itj.1202.01].

[48]　C. M. Lieber and Z. L. Wang, MRS Bull. **32**, 99 (2007).

[49]　W. Lu and C. M. Lieber, J. Phys. D: Appl. Phys. **39**, R387 (2006) [DOI: 10.1088/0022-3727/39/21/R01].

[50]　A. M. Morales and C. M. Lieber, Science **279**, 208 (1998) [DOI: 10.1126/science.279.5348.208].

[51]　W. Lu, J. Xiang, B. P. Timko, Y. Wu, and C. M. Lieber, Proc. Natl. Acad. Sci. U.S.A. **102**, 10046 (2005) [DOI: 10.1073/pnas.0504581102].

[52]　J. Xiang, W. Lu, Y. Hu, Y. Wu, H. Yan, and C. M. Lieber, Nature **441**, 489 (2006) [DOI: 10.1038/nature04796].

[53]　Y. Wu, Y. Cui, L. Huynh, C. J. Barrelet, D. C. Bell, and C. M. Lieber, Nano Lett. **4**, 433 (2004) [DOI: 10.1021/nl035162i].

[54]　W. Lu, P. Xie, and C. M. Lieber, IEEE Trans. Electron Devices **55**, 2859 (2008) [DOI: 10.1109/TED.2008.2005158].

[55]　M. J. Kumar, M. A. Reed, G. A. J. Amaratunga, G. M. Cohen, D. B. Janes, C. M. Lieber, M. Meyyappan, L. E. Wernersson, K. L. Wang, R. S. Chau, T. I. Kamins, M. Lundstrom, B. Yu, and C. Zhou, IEEE Trans. Electron Devices **55**, 2813 (2008) [DOI: 10.1109/TED.2008.2006781].

[56]　S. Iijima, Nature **354**, 56 (1991) [DOI: 10.1038/354056a0].

[57]　H. Dai, Surf. Sci. **500**, 218 (2002) [DOI: 10.1016/S0039-6028(01)01558-8].

[58]　M. S. Dresselhaus, G. Dresselhaus, and P. Avouris, *Carbon nanotubes: Synthesis, Structure, Properties, and Applications* (Springer, Berlin; New York, 2001).

[59]　P. Avouris, J. Appenzeller, R. Martel, and S. J. Wind, Proc. IEEE **91**, 1772 (2003) [DOI: 10.1109/JPROC.2003.818338].

[60]　N. Srivastava and K. Banerjee, IEEE/ACM International Conference on Computer-Aided Design (ICCAD-2005) (San Jose, CA 2005, IEEE/ACM) p. 383. [DOI: 10.1109/ICCAD.2005.1560098].

[61]　V. V. Zhirnov, J. A. Hutchby, G. I. Bourianoff, and J. E. Brewer, IEEE Circuits Devices Mag. **21**, 37 (2005) [DOI: 10.1109/MCD.2005.1438811].

[62]　M. Lundstrom, Proceedings of the 2002 International Symposium on Low Power Electronics and Design (ISLPED '02) (Monterey, CA 2002, IEEE) p. 172.

[63]　J. Appenzeller, Proc. IEEE **96**, 201 (2008) [DOI: 10.1109/JPROC.2007.911051].

[64]　K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004) [DOI: 10.1126/science.1102896].

[65]　W. A. De Heer, C. Berger, E. Conrad, P. First, R. Murali, and J. Meindl, IEEE International Electron Devices Meeting (IEDM) (Washington, DC 2007 Dec. 10-12, IEEE) p. 199. [DOI: 10.1109/IEDM.2007.4418901].

[66]　M. Y. Han, B. Ozyilmaz, Y. Zhang, and P. Kim, Phys. Rev. Lett. **98**(2007) [DOI: 10.1103/PhysRevLett.98.206805].

[67]　X. Li, X. Wang, L. Zhang, S. Lee, and H. Dai, Science **319**, 1229 (2008) [DOI: 10.1126/science.1150878].

[68]　Y. W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **97**(2006) [DOI: 10.1103/PhysRevLett.97.216803].

[69]　K. Mohanram and J. Guo, International Conference on Computer-Aided Design (ICCAD) (San Jose, CA 2008, IEEE) p. 412. [DOI: 10.1109/ICCAD.2008.4681607].

[70]　M. Choudhury, Y. Yoon, J. Guo, and K. Mohanram, 45th Design Automation Conference (DAC) (Anaheim, CA 2008, ACM) p. 272. [DOI: 10.1109/DAC.2008.4555822].

[71]　Y. Takahashi, A. Fujiwara, Y. Ono, and K. Murase, 30th IEEE International Symposium on Multiple-Valued Logic (ISMVL'2000) (Portland, OR 2000, IEEE) p. 411.

[72]　K. K. Likharev, Proc. IEEE **87**, 606 (1999) [DOI: 10.1109/5.752518].

[73]　C. Wasshuber, *Computational Single-Electronics* (Springer, Wien; New York, 2001).

[74]　H. Inokawa, A. Fujiwara, and Y. Takahashi, IEEE Trans. Electron Devices **50**, 462 (2003) [DOI: 10.1109/TED.2002.808421].

[75]　K. W. Song, Y. K. Lee, J. S. Sim, H. Jeoung, J. D. Lee, B. G. Park, Y. S. Jin, and Y. W. Kim, IEEE Trans. Electron Devices **52**, 1845 (2005) [DOI: 10.1109/TED.2005.852730].

[76]　M. Saitoh, H. Harata, and T. Hiramoto, IEEE International Electron Devices Meeting (IEDM) (San Francisco, CA 2004 Dec 13-15, IEEE) p. 187. [DOI: 10.1109/IEDM.2004.1419104].

[77]　S. Bandyopadhyay and V. Roychowdhury, Jpn. J. Appl. Phys. **35**, 3350 (1996) [DOI: 10.1143/JJAP.35.335].

[78]　K. S. Park, S. J. Kim, I. B. Back, W. H. Lee, J. S. Kang, Y. B. Jo, S. D. Lee, C. K. Lee, J. B. Choi, J. H. Kim, K. H. Park, W. J. Cho, M. G. Jang, and S. J. Lee, IEEE Trans Nanotechnol. **4**, 242 (2005) [DOI: 10.1109/TNANO.2004.837857].

[79]　V. V. Zhirnov, J. A. Hutchby, G. I. Bourianoffls, and J. E. Brewer, IEEE Circuits Devices Mag. 21, 37 (2005) [DOI: 10.1109/MCD.2005.1438811].

[80]　R. H. Chen, A. N. Korotkov, and K. K. Likharev, Appl. Phys. Lett. **68**, 1954 (1996) [DOI: 10.1063/1.115637].

[81]　K. Nishiguchi, A. Fujiwara, Y. Ono, H. Inokawa, and Y. Takahashi, Appl. Phys. Lett. **88**, 183101 (2006) [DOI: 1063/1.2200475].

[82]　K. Uchida, T. Tanamoto, R. Ohba, S. I. Yasuda, and S. Fujita, IEEE International Devices Meeting (IEDM) (San Francisco, CA 2002 Dec. 8-11, IEEE) p. 177.

[83]　N. Asahi, M. Akazawa, and Y. Amemiya, IEEE Trans. Electron Devices **44**, 1109 (1997) [DOI: 10.1109/16.595938].

[84]　C. S. Lent and P. D. Tougaw, J. Appl. Phys. **74**, 6227 (1993) [DOI: 10.1063/1.355196].

[85]　P. D. Tougaw and C. S. Lent, J. Appl. Phys. **75**, 1818 (1994) [DOI: 10.1063/1.356375].

[86]　C. S. Lent, P. D. Tougaw, W. Porod, and G. H. Bernstein, Nanotechnology **4**, 49 (1993) [DOI: 10.1088/0957-4484/4/1/004].

[87]　G. L. Snider, A. O. Orlov, V. Joshi, R. A. Joyce, Q. Hua, K. K. Yada-

Trans. Electr. Electron. Mater. 11(3) 93 (2010): Y.-B. Kim

105

valli, G. H. Bernstein, T. P. Fehlner, and C. S. Lent, 9th International Conference on Solid-State and Integrated-Circuit Technology (ICSICT 2008) (Bajing, China 2008 Oct. 20-23) p. 549. [DOI: 10.1109/ICSICT.2008.4734600].

[88] G. L. Snider, A. O. Orlov, R. K. Kummamuru, R. Ramasubramaniam, I. Amlani, G. H. Bernstein, C. S. Lent, J. L. Merz, and P. Wolfgang, Proceedings of the 2001 1st IEEE Conference on Nanotechnology (IEEE-NANO 2001) (Maui, HI 2008 Oct. 28-30, IEEE) p. 465. [DOI: 10.1109/NANO.2001.966468].

[89] J. A. Hutchby, R. Cavin, V. Zhirnov, J. E. Brewer, and G. Bourianoff, Computer **41**, 28 (2008) [DOI: 10.1109/MC.2008.154].

[90] A. K. Goel, *High-Speed VLSI Interconnections*, 2nd ed. (Wiley-Interscience; IEEE Press, Hoboken, NJ, 2007).

[91] A. K. Goel, IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2008) (Niagara Falls, ON 2008 May 4-7, IEEE) p. 189. [DOI: 10.1109/CCECE.2008.4564521].

[92] A. Orailoglu, 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Workshop on Dependable and Secure Nanocomputing (Edinburgh, UK 2007 Jun. 25-28, IEEE/IFIP). Available from: http://www.laas.fr/WDSN07/WDSN07_files/Texts/WDSN07-8D-04-Orailoglu.pdf.

[93] R. I. Bahar, D. Hammerstrom, J. Harlow, W. H. Joyner Jr, C. Lau, D. Marculescu, A. Orailoglu, and M. Pedram, IEEE Computer, **40**, 25 (2007)[DOI: 10.1109/MC.2007.7].

[94] A. DeHon and K. K. Likharev, IEEE/ACM International Conference on Computer-Aided Design (ICCAD-2005) (San Jose, CA 2005 Nov. 6-10, IEEE) p. 375. [DOI: 10.1109/ICCAD.2005.1560097].

[95] G. D. Hachtel and F. Somenzi, *Logic Synthesis and Verification Algorithms*, 2nd print., with corrections ed. (Kluwer Academic Publishers, Boston, 1998), pp. 25-41.

[96] G. De Micheli, *Synthesis and Optimization of Digital Circuits* (McGraw-Hill, New York, 1994), pp. 75-97.

[97] R. Zhang, *Computer-Aided Design Algorithms and Tools for Nanotechnologies*, Ph.D. dissertation (Princeton, NJ 2008, Princeton University).

[98] S. C. Goldstein, Proceedings of Government Microcircuit Applications and Critical Technology Conference (GOMAC Tech 04) (Monterey, CA 2004).