Beyond the conventional transistor

by H.-S. P. Wong

This paper focuses on approaches to continuing CMOS scaling by introducing new device structures and new materials. Starting from an analysis of the sources of improvements in device performance, we present technology options for achieving these performance enhancements. These options include high-dielectric-constant (high-*k*) gate dielectric, metal gate electrode, double-gate FET, and strained-silicon FET. Nanotechnology is examined in the context of continuing the progress in electronic systems enabled by silicon microelectronics technology. The carbon nanotube field-effect transistor is examined as an example of the evaluation process required to identify suitable nanotechnologies for such purposes.

1. Introduction

The semiconductor industry has been so successful in providing continued system performance improvement year after year that the Semiconductor Industry Association (SIA) has been publishing roadmaps for semiconductor technology since 1992. These roadmaps represent a consensus outlook of industry trends, taking history as a guide. The recent roadmaps [1] incorporate participation from the global semiconductor industry, including the United States, Europe, Japan, Korea, and Taiwan. They basically affirm the desire of the industry to continue with Moore's law [2], which is often stated as the doubling of transistor performance and quadrupling of the number of devices on a chip every three years. The phenomenal progress signified by Moore's law has been achieved through scaling of the metal-oxide–semiconductor field-effect transistor (MOSFET) [3, 4] from larger physical dimensions to smaller physical dimensions, thereby gaining speed and density.

Shrinking the conventional MOSFET beyond the 50-nm-technology node requires innovations to circumvent barriers due to the fundamental physics that constrains the conventional MOSFET. The limits most often cited [4–12] are 1) quantum-mechanical tunneling of carriers through the thin gate oxide; 2) quantum-mechanical tunneling of carriers from source to drain, and from drain to the body of the MOSFET; 3) control of the density and location of dopant atoms in the MOSFET channel and source/drain region to provide a high on-off current ratio; and 4) the finite subthreshold slope. These fundamental limits have led to pessimistic predictions of the imminent end of technological progress for the semiconductor industry [4]. On the other hand, the push to scale the conventional MOSFET continues to show remarkable progress [13, 14].

Instead of reiterating the considerations of device scaling limits here, we refer the reader to our previous analyses [8–10] as well as analyses by others in the literature [4–7, 11, 12]. We focus this paper instead on approaches to circumvent or surmount the barriers to device scaling. The organization of this paper is as follows. We first address opportunities for the silicon MOSFET, focusing primarily on approaches that depart from conventional scaling techniques (for example, doping profile control, thin silicon dioxide gate dielectrics, SOI). Topics covered include high-dielectric-constant (high-*k*)

©**Copyright** 2002 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/02/$5.00 © 2002 IBM

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

Table 1 Device performance improvement opportunities.

Source of improvement	*Parameters affected*	*Method*
Charge density	1. S (inverse subthreshold slope) 2. Q_{inv} at a fixed off-current	1. Double-gate FET. 2. Lowered operating temperature.
Carrier transport	1. Mobility (μ_{eff}) 2. Carrier velocity 3. Ballistic transport	1. Strained silicon. 2. High-mobility and -saturation-velocity materials (e.g., Ge, InGaAs, InP). 3. Reduced mobility degradation factors (e.g., reduced transverse electric field, reduced Coulomb scattering due to dopants, reduced phonon scattering). 4. Shorter channel length. 5. Lowered operating temperature.
Ensuring device scalability to a shorter channel length	1. Generalized scale length (λ). 2. Channel length (L_g)	1. Maintaining good electrostatic control of channel potential (e.g., double-gate FET, ground-plane FET, and ultrathin-body SOI) by controlling the device physical geometry and providing means to terminate drain electric fields. 2. Sharp doping profiles, halo/pocket implants. 3. High gate capacitance (thin gate dielectrics, metal gate electrode) to provide strong gate control of channel potential.
Parasitic resistance	1. R_{ext}	1. Extended/raised source/drain. 2. Low-barrier Schottky contact.
Parasitic capacitance	1. C_{jn} 2. C_{GD}, C_{GS}, C_{GB}	1. SOI. 2. Double-gate FET.

gate dielectric, metal gate electrode, double-gate FET, and strained-silicon FET. The second part of this paper examines the space between conventional microelectronics technology and the more exploratory nanotechnology. Such a wide spectrum of nanotechnologies are being explored today that it is impossible to make even a modest attempt to cover the field. The approach adopted in this paper is to select an example, the carbon nanotube field-effect transistor, to illustrate both the opportunities offered by nanotechnologies and the most important questions that must be answered before such technologies can find practical use. The example is therefore chosen for illustrative purposes rather than an implied suggestion of eventual technological utility.

2. Silicon MOSFET

For digital circuits, a figure of merit for MOSFETs for unloaded circuits is CV/I, where C is the gate capacitance, V is the voltage swing, and I is the current drive of the MOSFET. For loaded circuits, the current drive of the MOSFET is of paramount importance. Historical data indicate that scaling the MOSFET channel length improves circuit speed, as suggested by scaling theory [3]. Reference [15] illustrates data on the CV/I metric from recent literature. The off-current specification for CMOS has been rising rapidly to keep the speed performance high. While 1 nA/μm was the maximum off-current allowed in the late 1990s [8], off-currents in excess of 100 nA/μm are proposed today [13]. This trend obviously cannot continue, since the on-current increases only linearly as off-current increases exponentially in a typical device design tradeoff. Means to mitigate the standby power increase must be found.

Keeping in mind both the CV/I metric and the benefits of a large current drive, we note that device performance may be improved by 1) inducing a larger charge density for a given gate voltage drive; 2) enhancing the carrier transport by improving the mobility, saturation velocity, or ballistic transport; 3) ensuring device scalability to achieve a shorter channel length; and 4) reducing parasitic capacitances and parasitic resistances. **Table 1** summarizes these opportunities and proposed technology options for capitalizing on them. These options generally fall into two categories: new materials and new device structures. In many cases, the introduction of a new material requires the use of a new device structure, or vice versa. Throughout the discussion, we direct attention to areas of device physics and materials science that must be better understood in order to advance the technology.

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

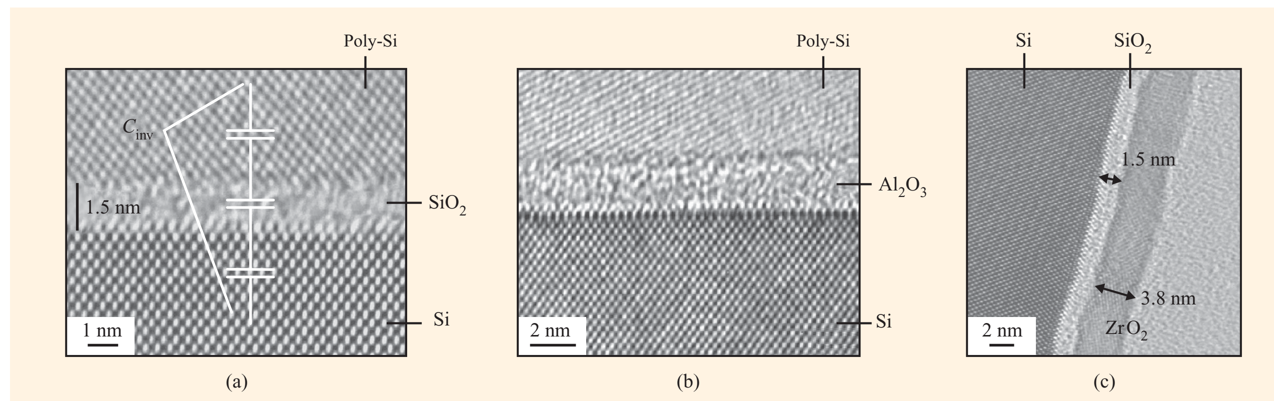

Figure 1

(a) Transmission electron micrograph (TEM) of a conventional silicon dioxide (oxynitride) with a physical thickness of 1.5 nm. (b) TEM of a 2.2-nm Al_2O_3 with an equivalent electrical thickness of 1 nm. (c) TEM of a 3.8-nm ZrO_2 on a 1.5-nm interfacial silicon dioxide. Adapted with permission from Gusev et al. [20]; © 2001 IEEE.

MOSFET gate stack

Continued device scaling requires the continued reduction of the gate dielectric thickness. This requirement arises from two different considerations: controlling the short-channel effect and achieving a high current drive by keeping the amount of charge induced in the channel large as the power-supply voltage decreases. In both cases, to a first approximation, it is the electrical thickness that is important. The electrical thickness at inversion is determined by the series combination of three capacitances in the gate stack: the depletion capacitance of the gate electrode, the capacitance of the gate dielectric, and the capacitance of the inversion layer in the silicon [**Figure 1**, part (a)].

On the other hand, the direct tunneling current through the gate dielectric grows exponentially with decreasing physical thickness of the gate dielectric [16]. This tunneling current has a direct impact on the standby power of the chip and puts a lower limit on unabated reduction of the physical thickness of the gate dielectric. It is likely that tunneling currents arising from silicon dioxides (SiO_2) thinner than 0.8 nm cannot be tolerated, even for high-performance systems [10].

Solutions that reduce the gate tunneling current and gate capacitance degradation due to polysilicon depletion are explored through introduction of new materials: high-dielectric-constant gate dielectrics and metal gate electrodes.

High-k gate dielectric

A gate dielectric with a dielectric constant (k) substantially higher than that of SiO_2 (k_{ox}) will achieve a smaller equivalent electrical thickness (t_{eq}) than the SiO_2, even with a physical thickness (t_{phys}) larger than that of the SiO_2 (t_{ox}):

$$t_{eq} = \left(\frac{k_{ox}}{k}\right) t_{phys} .$$

Replacing the SiO_2 with a material having a different dielectric constant is not as simple as it may seem. The material bulk and interface properties must be comparable to those of SiO_2, which are remarkably good. Basic material properties such as thermodynamic stability with respect to silicon, stability under thermal conditions relevant to microelectronic fabrication, low diffusion coefficients, and thermal expansion match are some critical examples. In addition, interface traps of the order of a few 10^{10} $cm^{-2}eV^{-1}$ and bulk traps of the order of a few 10^{10} cm^{-2} are common among SiO_2 and the closely related oxynitrides [17, 18]. Charge trapping and reliability for the gate dielectrics are particularly important considerations.

Thermal stability with respect to silicon is an important consideration, since high-temperature anneals are generally employed to activate dopants in the source/drain as well as the polysilicon gate. Although many binary and ternary oxides are predicted to be thermally stable with respect to silicon [19], recent research on high-dielectric-constant gate insulators have focused primarily on binary metal oxides such as Ta_2O_5, TiO_2, ZrO_2, HfO_2, Y_2O_3, La_2O_3, Al_2O_3, and Gd_2O_3 and their silicates [20]. **Table 2** compares the properties of the common high-k gate dielectrics reported in the literature. The dielectric constant of these materials generally ranges from 10 to 40, which is about a factor of 3 to 10 higher than SiO_2. Leakage current reduction from $10^3\times$ to $10^6\times$, in comparison with SiO_2 of the same electrical thickness, is generally achieved experimentally for high-k gate dielectrics [21]. The benefits of using a very-high-dielectric-constant material to simply replace SiO_2 for

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

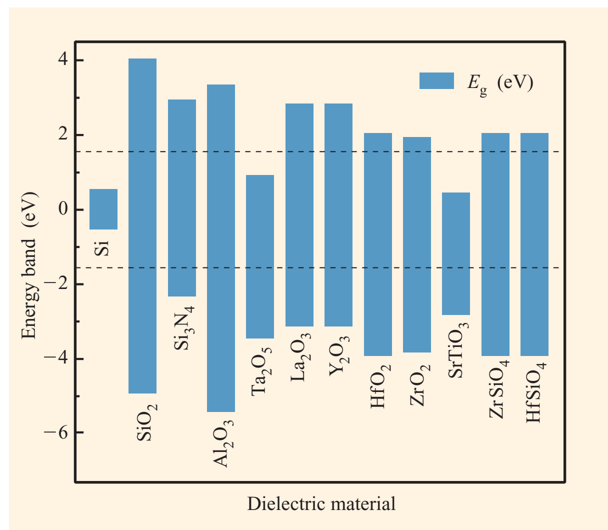

Figure 2

Bandgap and band alignment of high-k gate dielectrics with respect to silicon. Data from Robertson [25], with permission. The dashed line represents 1 eV above/below the conduction/valence bands.

the same electrical thickness are limited because of the presence of two-dimensional electric fringing fields from the drain through the physically thicker gate dielectric [10, 22]. The drain fringing field lowers the source-to-channel potential barrier and lowers the threshold voltage in a way similar to the well-known drain-induced barrier lowering (DIBL), in which the drain field modulates the source-to-channel potential barrier via coupling through the silicon substrate. The use of higher-k materials must therefore be combined with a concurrent reduction of the electrical thickness.

A large silicon-to-insulator energy barrier height is desirable because the gate direct-tunneling current is exponentially dependent on the (square root of the) barrier height [23]. In addition, hot-carrier emission into the gate insulator is also related to the same barrier height [24]. The high-k material should therefore not only have a large bandgap, but also have a band alignment which results in a large barrier height. **Figure 2** illustrates the bandgap and band alignment for several high-k gate dielectrics calculated by Robertson [25]. Most high-k materials that have other desirable properties do have relatively low band offsets and small bandgaps. Aluminum oxide (Al_2O_3) is probably the only material that has a bandgap and band alignment similar to those of SiO_2.

Figure 1 illustrates examples of thin gate dielectrics: SiO_2, Al_2O_3, and ZrO_2 with an interfacial SiO_2 layer. These dielectrics are only a few atoms thick. The thin dielectric films can be deposited by sputtering, sol–gel, physical vapor deposition (PVD), metallo-organic chemical vapor deposition (MOCVD), or atomic-layer deposition (ALD). Deposition uniformity does not appear to be a significant issue. However, integration of the deposited

Table 2 Selected material and electrical properties of high-k gate dielectrics. Data compiled from Robertson [25], Gusev et al. [20], Hubbard and Schlom [19], and other sources.

Dielectric	*Dielectric constant (bulk)*	*Bandgap (eV)*	*Conduction band offset (eV)*	*Leakage current reduction w.r.t. SiO_2*	*Thermal stability w.r.t. silicon (MEIS data)*
Silicon dioxide (SiO_2)	3.9	9	3.5	N/A	>1050°C
Silicon nitride (Si_3N_4)	7	5.3	2.4		>1050°C
Aluminum oxide (Al_2O_3)	~10	8.8	2.8	10^2–10^3×	~1000°C, RTA
Tantulum pentoxide (Ta_2O_5)	25	4.4	0.36		Not thermodynamically stable with silicon
Lanthanum oxide (La_2O_3)	~21	6*	2.3		
Gadolinium oxide (Gd_2O_3)	~12				
Yttrium oxide (Y_2O_3)	~15	6	2.3	10^4–10^5×	Silicate formation
Hafnium oxide (HfO_2)	~20	6	1.5	10^4–10^5×	~950°C
Zirconium oxide (ZrO_2)	~23	5.8	1.4	10^4–10^5×	~900°C
Strontium titanate ($SrTiO_3$)		3.3	−0.1		
Zirconium silicate ($ZrSiO_4$)		6*	1.5		
Hafnium silicate ($HfSiO_4$)		6*	1.5		

*Estimated value.

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

dielectric with the rest of the device fabrication process requires further research and development in several areas. If a conventional self-aligned polysilicon gate is used, the dielectric film must be able to withstand rapid thermal anneals (RTAs) up to at least 950°C for dopant activation in the polysilicon gate. The typical thermal treatments during a polysilicon gate CMOS process pose potential problems such as formation of silicates and interfacial SiO_2. In addition, diffusion (for example, boron, oxygen) through the gate dielectric is a serious concern. If a metal gate electrode is employed (using a low-temperature process), many of the thermal stability concerns can be relieved.

Figure 3(a) shows the electrical characteristics of an 80-nm polysilicon gate n-FET using Al_2O_3 as the gate dielectric, as reported by Buchanan et al. [26]. This work and that of others (for example, [21]) illustrates some of the obstacles for high-k gate dielectrics: 1) There are a significant number of traps and fixed charges in the film (or at the interfaces), leading to flat-band voltage shifts (up to 450 mV) and voltage bias instability; 2) the traps raise questions of reliability as channel hot carriers and carriers from gate tunneling traverse the gate dielectric, resulting in trap generation; and 3) the mobility of carriers in the FET channel is severely degraded (up to a factor of 2) for high-k gate dielectrics [**Figure 3(b)**].

The cause of the mobility degradation is not clear at present. Presumably, some of the differences can be attributed to the difficulty of obtaining accurate estimates of the effective electric field due to the charge trapping. Coulomb scattering due to the trapped charge alone cannot explain entirely the mobility degradation observed. Another source of mobility lowering may be found in remote phonon scattering [27]. The static dielectric constant of a high-bandgap high-k material derives its high dielectric constant primarily from ionic polarizability, since the large bandgap results in a small electronic polarizability. The ionic polarizability is associated with the "soft" metal–oxygen bonds with low-energy phonons. Fischetti et al. [27] studied the scattering of electrons in the inversion layer by surface optical phonons and suggested that there is generally an inverse relation between surface-optical-phonon-limited mobility and the static dielectric constant: the higher the dielectric constant, the lower the surface-optical-phonon-limited mobility.

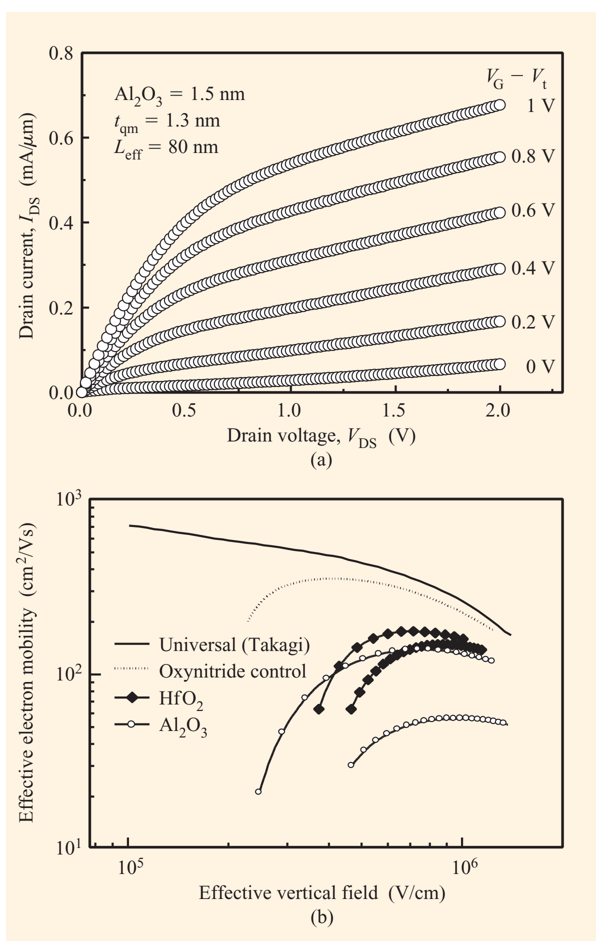

Figure 3

Electrical characteristics of a polysilicon-gated Al_2O_3 n-FET. (a) Drain current vs. drain voltage characteristics of an 80-nm-channel-length n-FET. Reproduced with permission from Buchanan et al. [26]; © 2000 IEEE. (b) Effective electron mobility of long-channel FET compared with the universal mobility curve [60]. Two HfO_2 curves show the effect of surface preparation. The Al_2O_3 curves illustrate the range of mobility for Al_2O_3 gate stacks. Mobility approximately twice as high as that of [26] is achieved due to improved processing. Reproduced with permission from Gusev et al. [21]; © 2001 IEEE.

Metal gate electrode

A metal gate electrode has several advantages compared to the doped polysilicon gate used almost exclusively today. Gate capacitance degradation due to the depletion of the doped polysilicon gate typically accounts for 0.4–0.5 nm of the equivalent-oxide thickness of the total gate capacitance at inversion. This is a substantial amount, considering that a gate equivalent oxide of less than 1.5 nm (at inversion) is required for sub-50-nm CMOS. The thermal instability of most high-k gate dielectrics may require the use of a low thermal budget process after the gate dielectric deposition. While junction activation may be performed prior to gate dielectric deposition, the high-temperature gate polysilicon activation step necessarily occurs after the gate dielectric formation. A further potential benefit of metal gate electrodes is the elimination of carrier mobility degradation due to plasmon scattering from the gate electrode. The plasmon frequency

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.