[Pick the date]

TECHALONE.COM

FINFET

INTRODUCTION

Since the fabrication of MOSFET, the minimum channel length has been shrinking continuously. The motivation behind this decrease has been an increasing interest in high speed devices and in very large scale integrated circuits. The sustained scaling of conventional bulk device requires innovations to circumvent the barriers of fundamental physics constraining the conventional MOSFET device structure. The limits most often cited are control of the density and location of dopants providing high I_{on}/I_{off} ratio and finite subthreshold slope and quantum-mechanical tunneling of carriers through thin gate from drain to source and from drain to body. The channel depletion width must scale with the channel length to contain the off-state leakage I_{off}. This leads to high doping concentration, which degrade the carrier mobility and causes junction edge leakage due to tunneling. Furthermore, the dopant profile control, in terms of depth and steepness, becomes much more difficult. The gate oxide thickness t_{ox} must also scale with the channel length to maintain gate control, proper threshold voltage V_T and performance. The thinning of the gate dielectric results in gate tunneling leakage, degrading the circuit performance, power and noise margin.

Alternative device structures based on **silicon-on-insulator (SOI)** technology have emerged as an effective means of extending MOS scaling beyond bulk limits for mainstream high-performance or low-power applications .**Partially depleted (PD) SOI** was the first SOI technology introduced for high-performance microprocessor applications. The **ultra-thin-body fully depleted (FD) SOI** and the **non-planar FinFET** device structures promise to be the potential "future" technology/device choices.

In these device structures, the short-channel effect is controlled by geometry, and the off-state leakage is limited by the thin Si film. For effective suppression of the off-state leakage, the thickness of the Si film must be less than one quarter of the channel length.

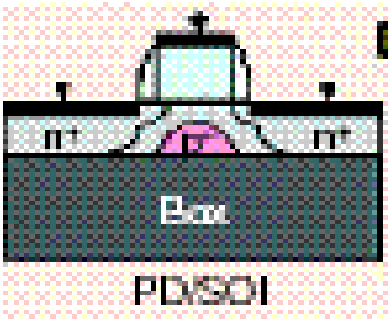
The desired V_T is achieved by manipulating the gate work function, such as the use of midgap material or poly-SiGe. Concurrently, material enhancements, such as the use of a) high-k gate material and b) strained Si channel for mobility and current drive improvement, have been actively pursued.

As scaling approaches multiple physical limits and as new device structures and materials are introduced, unique and new circuit design issues continue to be presented. In this article, we review the design challenges of these emerging technologies with particular emphasis on the implications and impacts of individual device scaling elements and unique device structures on the circuit design. We focus on the planar device structures, from continuous scaling of PD SOI to FD SOI, and new materials such as strained-Si channel and high-k gate dielectric.

PARTIALLY DEPLETED [PD] SOI

The PD floating-body MOSFET was the first SOI transistor generically adopted for high-performance applications, primarily due to device and processing similarities to bulk CMOS device.

The PD SOI device is largely identical to the bulk device, except for the addition of a buried oxide ("BOX") layer. The active Si film thickness is larger than the channel depletion width, thus leaving a quasi-neutral "floating" body region underneath the channel. The V_T of the device is completely decoupled from the Si film thickness, and the doping profiles can be tailored for any desired V_T .



The device offers several advantages for performance/ power improvement:

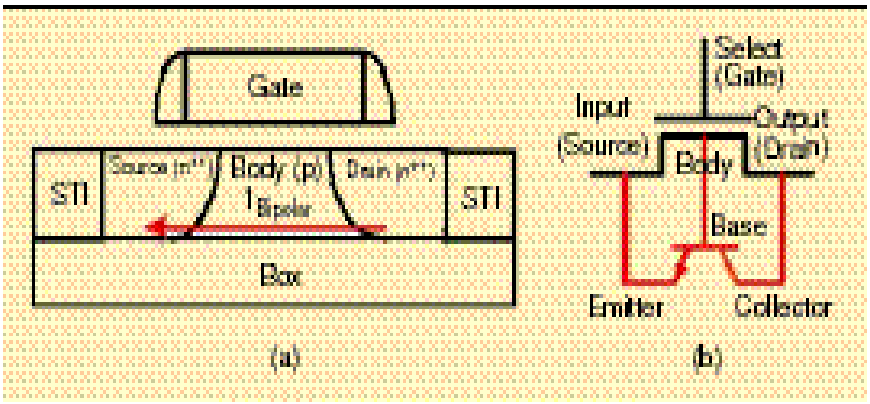
1) reduced junction capacitance,
2) lower average threshold due to positive V_{BS} during switching.
3) dynamic loading effects,in which the load device tends to be in high V_T state during switching

The performance comes at the cost of some design complexity resulting from the floating body of the device, such as

1) parasitic bipolar effect and
2) hysteretic V_T variation.

Parasitic Bipolar Effect

In PDSOI an n-p-n transistor is formed with source and drain as emitter & collector respectively and body as the base. The topology typically involves an "off" transistor with the source and drain voltage set up in the "high" state (hence body voltage at"high") When the source is subsequently pulled down, large overdrive is developed across the body-source junction, causing bipolar current to flow through the lateral parasitic bipolar transistor.This may result in circuit failure.



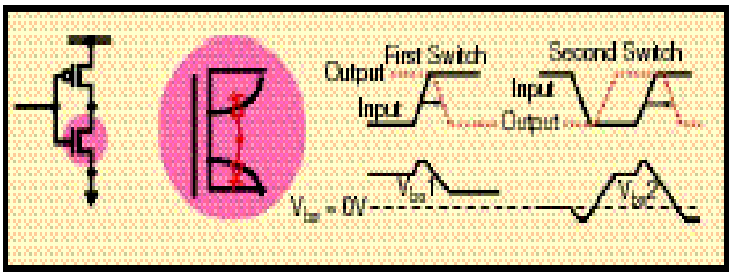
In SRAM bitline structures, the aggregate parasitic bipolar effect of the unselected cells on the selected bitline disturbs the read/write operations and limits the number of cells that can be attached to a bitline pair

Hysteretic V_T Variation

The hysteretic VT variation due to long time constants of various body charging/discharging mechanisms.

A commonly used gauge for hysteretic V_T variation (or "history effect" as it is known in the SOI community) is the disparity in the body voltages and delays between the so-called "first switch" and "second switch" . The "first switch" refers to the case where a circuit (e.g., inverter) starts in an initial quiescent state with input "low" and then undergoes an input-rising transition. In this case, the initial dc equilibrium body potential

of the switching nMOSFET is determined primarily by the balance of the back-to-back drain-to-body and body-to-source diodes. The “second switch” refers to the case where the circuit is initially in a quiescent state with input “high.” The input first falls and then rises (hence, the name “second switch”). For this case, the preswitch body voltage is determined by capacitive coupling between the drain and the body.



Input/output waveforms&nMOS body voltage for a PD SOI CMOS inverter under “first switch” & “second switch” condition is shown above

The duty cycle, slew rate, and output load also affect the hysteretic behavior of the circuits. A higher duty cycle increases hysteretic behavior due to higher switching activity causing a gain or loss of body charge and less time for the
device to return/settle to its initial equilibrium state

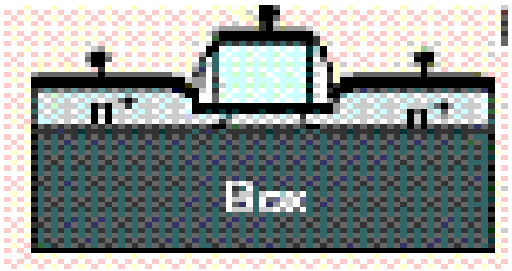
SCALING Si FILM: FROM PD SOI TO FD SOI

The major benefits of scaling/thinning of the Si film are: 1) reduction of junction capacitance for performance improvement, 2) better short channel roll-off, and 3) better soft error rate (SER) due to less charge generation and collection volume.

In addition, the history effect (disparity between first switch and second switch) is also reduced. The reduced junction capacitance improves delays of both the first and second switches. However, for the second switch the reduced junction capacitance reduces the capacitive coupling between the drain and the body .The resulting decrease in the pre-switch body voltage for the second switch partially offsets the performance improvement.

Unfortunately, the thinning of Si film degrades the body resistance, rendering body contacts less effective and eventually useless .Self-heating becomes more severe. Furthermore, as the film thickness is scaled below 50nm,the device may become dynamically fully depleted (or quasi-depleted);the body would become fully depleted under certain bias conditions or during certain circuit-switching transients. This necessitates a unified PD|FD device model with smooth and seamless transitions among different modes of operation. Typically, this is modeled by varying the built-in potential between the body and source junction, thus changing the amount of body charges the body-to-source junction diode can sink for a given change in the body potential .The presence of dynamic full depletion also complicates the static timing methodology. The various body voltage bounds, established based on the assumption of partial depletion need to be extended to cover this new phenomenon. Notice that dynamic depletion tends to occur first in long-channel, low V_T devices. For short channel devices, the proximity of the heavily doped "halo" regions to each other increases the effective body doping, and the device is less likely to be dynamically fully depleted. In a FD_SOI device, the channel depletion layer extends through the entire Si film. This significantly reduces the floating body effect (completely eliminating the floating-body effect with ultra-thin Si films). A raised source/drain structure is typically employed to overcome the large source/drain series resistance of the thin Si film. There are two approaches to achieve the desired VT. One can use the traditional dual P+/N+ poly-silicon with a highly doped channel. This approach has several drawbacks and limitations: a) V_T would be sensitive to Si film thickness variation, b) high doping degrades the carrier mobility and results in junction edge leakage due to tunneling, and c) in devices with ultra-thin body, the amount of dopant required for the desired V_T can not be realistically achieved. Excessively high body doping would turn the device into a "resistor" rather than a "transistor." Consequently, the preferred and more scalable approach is to build an "undoped" channel with the desired V_T set either by the source/drain halo or by the use of midgap gate materials. The use of undoped channel a) reduces the V_T sensitivity to Si film thickness variation, b) reduces dopant fluctuation

effect, c) reduces transverse electric field and impurity scattering, leading to higher mobility, and d) reduces band-to-band tunneling leakage at the junction edge**.**



The use of midgap gate materials may allow the use of a single electrode for both nMOSFETs and pMOSFETs.

As the Si film thickness is reduced, the gate has more control of the channel charges, and the subthreshold slope improves. The channel leakage in the device is limited by the Si film thickness and decreases as the film thickness is reduced. The fringing electric field from source/drain penetrates into the buried oxide underneath the channel, which causes back interface virtual biasing, resulting in increased I_{off} and degraded subthreshold slope. This can be suppressed by thinning the buried oxide at the expense of larger junction capacitance to the substrate. The FD SOI technology is , in general, quite "transparent" for design migration and the challenge is primarily in technology development and manufacturing.

MAJOR DESIGN ISSUES

Gate Oxide Tunneling Leakage

As the gate oxide thickness is scaled to maintain gate control $V_{T\ and}$ performance, gate insulator direct tunneling leakage increases. Nitrided oxide, which reduces the leakage by any order of magnitude , has been widely used in the industry to contain this leakage. Nevertheless, the oxide tunneling leakage increases by 2 5* for every 0.1 nm decrease in oxide thickness. This amounts to over a 30* increase per technology generation. On the

contrary, the channel leakage increases by about 3-5 per technology generation. As such , the oxide tunneling leakage has quickly approached I_{off} and will surpass I_{off} at room temperature for oxide thickness around 1.0nm or below , thus becoming a serious concern for overall chip leakage.

Furthermore, at 1.0nm, the tunneling leaking for nitrided oxide reaches 100A/cm^2, while the traditional reliability limitation for silicon dioxide gate insulator leakage is 1.0A/cm^2 .A recent study showed that, at 100 A/cm^2 , static CMOS and domino circuits in bulk CMOS still exhibit “acceptable functionality and noise margin”.

The oxide tunneling current consists of several components. The electron tunneling from the valence band (EVB) generates the substrate current in both nMOS and pMOS .This substrate current component is significantly less than the tunneling current between the gate and the channel, and its effect can usually be neglected in bulk CMOS .In PD SOI devices, however, this substrate current charges or discharges the body, thus changing V_T and affecting circuit operation. As this gate-to-body tunneling current has a weaker temperature dependence than the channel current, and other leakage and body charging/discharging current components, its effect is more pronounced at lower temperature.

The detailed study on a 34-kb L1 directory SRAM showed that the presence of in the gate-to-body tunneling current resulted in much more significant degradation “write” operation compared with the “read” operation. On the other hand , the initial cycle parasitic bipolar disturb resulting from the aggregate effect of unselected cells in the same bitline was reduced.

The gate-to-body tunneling current increases the disparity between the first switch and the second switch.

In FD SOI and FinFET with ultra-thin body, the gate tunneling leakage is significantly reduced. This is because 1) in these, an undoped or very lightly doped body (channel) is used and the depletion charges essentially equal to zero, thus reducing the vertical electric field in the channel, and 2) the quantum confinement effect in ultra-thin Si film results in a broader inversion charge distribution and lower vertical electric field at the bottom of the inversion layer. Consequently, the gate tunneling leakage is reduced by about

3-4X.If a high-k gate dielectric such as HfO_2 ,the reduction in gate current can exceed an order of magnitude as the increased physical thickness of the gate dielectric barrier makes the tunneling current more dependent upon the shape
of the potential well in Si. In general, scaling of the body thickness reduces the gate leakage current because the potential well becomes shallower. However, excessive scaling\thinning of body thickness below~5nm increases confinement of carriers toward the gate dielectric interface, and the gate current increases to approach that of the bulk devices. This may not be of concern since 5nm is close to the practical limit of body thickness in actual device technologies.

Self heating

The heat transfer is dominated by phonon transport in semiconductors and by electron transport in metals. The thermal conductivity of the buried oxide (1.4 W/m-.C) is about two orders of magnitude lower than that of Si (120 W/m-.C), giving rise to local self-heating in SOI devices. This is particularly a concern for devices that are "on" most or all the time (e.g., biasing elements, current source, current mirror, bleeder, etc.) and for circuits with high duty cycle and slow slew rate (such as clock distribution, I/O driver).

Scaling of the Si film degrades the thermal conductivity and increases the thermal resistance. In scaled SOI devices, both the channel length and Si film thickness are much smaller than the phonon mean free path for Si (~300 nm at room temperature), and the thermal conductivity is severely degraded due to phonon boundary scattering.

The thermal resistance increase is particularly significant for thinner Si film with thick buried oxide. As the Si film thickness is scaled further to approach the Phonon wavelength (~ tens of nm), the phonon confinement effect becomes significant. This is the mechanical/thermal analogy of the quantum confinement effect in electronic devices with an ultra-thin Si film. The boundary conditions change from the usual periodic boundary conditions for bulk materials to essentially zero displacements on the boundaries in SOI.

Soft Error Rate

The α-generated charges in SOI devices are substantially less than in bulk devices due to the presence of the buried oxide, and appreciable charge generation can only

occur when an α-particle hits the channel region. While scaling of the device reduces the charge generation volume, the Q_{crit} also decreases due to a lower capacitance at the cell's storage node and scaled V_{DD}.

In a PD SOI device, the total charges accumulated at the cell storage node can be significantly higher than the α-generated charges due to the parasitic bipolar effect. For properly scaled PD SOI devices, the parasitic bipolar gain is reduced, and the resulting overall single-event-upset-induced failure rate is less than that of bulk silicon. Furthermore, scaling/thinning of the Si film reduces the charge generation volume and the base-emitter (body-source) junction area of the parasitic bipolar transistor, thus improving SER as well.

Strained-Si channel And High-k Gate

Strained-Si surface channel CMOS has recently emerged as an effective means of extending scaling for future high-performance applications due to higher mobility and improved I_{on}. The lattice mismatch between the Si channel and the underlying relaxed SiGe layer results in biaxial tensile strain, which reduces the intervalley scattering by increasing sub-band splitting and enhances carrier transport by reducing conductivity effective mass.

Combining strained si-channel and SOI complements the improved I_{on} of strained Si channel device with the benefit of SOI. However, there are numerous design implications. The narrower bandgap of the SiGe layer causes a heterostructural band offset, which reduces V_T and increases I_{off}. The mobility enhancement for nMOS and pMOS may be quite different due to device design and process integration constraints, which may upset the established β(p/n strength) ratio of existing designs. The tensile strain is "biaxial", so mobility enhancements (therefore I_{on} improvement) are the same along X- and Y-axis. However, in some high-density design (eg: SRAM cell), "bent gates" at a 45° angle are sometimes used, which would result in disparity in mobility enhancement and I_{on} improvement. The SiGe layer with 20% Ge has a 70% higher dielectric constant and a 10% lower built-in potential due the narrower band gap, resulting in higher junction capacitance. Furthermore, higher body doping density could be needed to compensate for the V_T

reduction which further increases the junction capacitance. The thermal conductivity of the SiGe layer is about 15X lower than that for Si thus aggravating the self-heating effect.

The presence of the SiGe layer also significantly affects the floating-body effect. For 20% Ge content, the band gap is about 90% of that of Si. This narrower bandgap results in higher (~10X) intrinsic carrier density n_i , and thus proportionally higher recombination current at the body-to-source junction. The narrower bandgap and higher dielectric constant of the SiGe layer, and the higher body doping to compensate for the lowered V_T caused by the band offset, give rise to larger band-to-band tunneling current and trapped-assisted tunneling current at the drain-to-body junction. The latter effect may overpower the increase in recombination current at the body-to-source-junction, resulting in more significant floating-body effect.

High-k gate dielectric has recently been pursued to contain the gate leakage and extend device scaling. Most of the potential high-k gate insulators have lower bandgap than SiO_2 and therefore must be thicker to keep the tunneling leakage down. These materials also have charge-trapping related V_T instability and mobile degradation. The integration of high-k gate dielectric with strained-Si channel significantly enhances the mobility. Notice that, as the high-k gate material offers higher gate capacitance per unit area, some circuit resizing/retuning may be necessary, especially in the critical paths where device capacitances tend to dominate.

INTRODUCTION TO DOUBLE_GATE CMOS

Innovative device architectures will be necessary to continue the benefits that previously acquired through rote scaling. Double-gate CMOS (DGCMOS) offers distinct advantages for scaling to very short gate lengths. Furthermore, adoption of gate dielectrics with permittivity substantially greater than that of SiO_2 (so-called “high-k materials”) may be deferred if a DGCMOS architecture is employed. Previously, serious structural challenges have made adoption of DGCMOS architecture untenable. Recently, through use of the delta device, now commonly referred to as the FinFET, significant advances in DGCMOS device technology and performance have been demonstrated. Fabrication in

FinFET-DGCMOS is very close to that of conventional CMOS process, with only minor disruptions, offering the potential for a rapid deployment to manufacturing. Planar product designs have been converted to FinFET–DGCMOS without disruption to the physical area, thereby demonstrating its compatibility with today's planar CMOS design methodology and automation techniques.

Overcoming Obstacles By Doubling Up

CMOS technology scaling has traversed many anticipated barriers over the past 20 years to rapidly progress from 2μm to 90 nm rules. Currently, two obstacles, namely subthreshold and gate-dielectric leakages, have become the dominant barrier for further CMOS scaling, even for highly leakage-tolerant application such as microprocessors.

DOUBLE GATE FET

Double-gate CMOS (DGCMOS) offers distinct advantages for scaling to very short gate lengths. Fabrication of FinFET-DGCMOS is very close to that of conventional CMOS process, with only minor disruptions, offering the potential for a rapid deployment to manufacturing. Planar product designs have been converted to FinFET-DGCMOS without disruption to the physical area, thereby demonstrating its compatibility with today's planar CMOS design methodology and automation techniques.

Overcoming Obstacles By Doubling Up

CMOS technology scaling has traversed many anticipated barriers over the past 20 years to rapidly progress from 2 μm to 90nm rules. Currently, two obstacles, namely subthreshold and gate-dielectric leakages, have become the dominant barrier for further CMOS scaling, even for highly leakage-tolerant applications such as microprocessors.

Double-gate (DG) FETs, in which a second gate is added opposite the traditional (first) gate, have better control over short-channel effects [SCEs]. SCE limits the minimum channel length at which an FET is electrically well behaved.

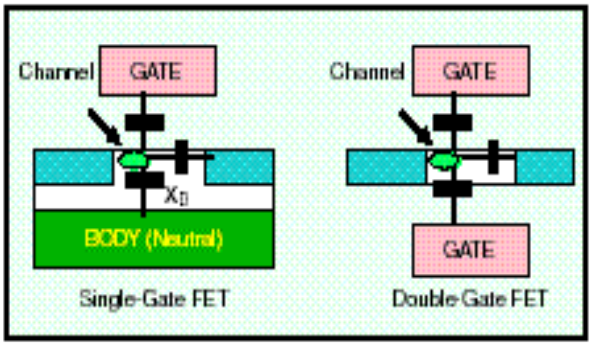


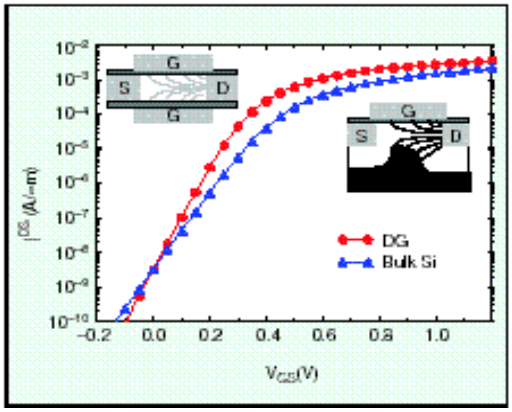
Figure schematically illustrates the advantage of DG-FETs.

As the channel length of an FET is reduced, the drain potential begins to strongly influence the channel potential, leading to an inability to shut off the channel current with the gate. This short-channel effect is mitigated by use of thin gate oxide (to increase the influence of the gate on the channel) and thin depletion depth below the channel to the substrate, to shield the channel from the drain. Gate oxide thickness has been reduced to the point where, at 90 nm CMOS, the power drain from gate leakage is comparable to the power used for switching of circuits. Thus, further reduction of the thickness would lead to unreasonable power increases.

Alternatively, further decrease of the depletion region X_D degrades gate influence on the channel and leads to a slower turn on of the channel region.

In DG-FETs, the longitudinal electric field generated by the drain is better screened from the source end of the channel due to proximity to the channel of the second gate, resulting in reduced short-channel effects, in particular, reduced drain induced- barrier lowering (DIBL) and improved subthreshold swing (S). Therefore, as CMOS scaling becomes limited by leakage currents, DGCMOS offers the opportunity to proceed beyond the performance of single-gate (SG) bulk-silicon or PDSOI CMOS. Both the DIBL and subthreshold swing for the DG device are dramatically improved relative to those of the bulk-silicon counterpart. From a bulk-silicon device design perspective, increased body doping concentration could be employed to reduce DIBL; however, at some point it would also increase the subthreshold swing, thereby requiring higher threshold voltage V_T to keep the subthreshold current adequately low. Similarly, decreasing the body doping

concentration could improve the subthreshold swing but degrade DIBL. Hence a compromise is necessary for the bulk-silicon device design. Note that, for a scaled bulk-silicon (or PD SOI) device, a highly doped channel/halo must be used to control severe SCEs, and lower S for extremely short L_{eff} could not be achieved by use of low channel/halo doping.



In Figure, simulations of the IDS–VGS characteristics of DG and SG FETs shows the steeper turn on of the DG-FET, which results from the gate coupling advantage. This property enables the use of lower threshold voltage for the DG-FET for a given off-current. As a direct result, higher drive currents at lower power-supply voltages V_{DD} are attainable.

Double_Gate Threshold Voltage

The very thin silicon body associated with fully depleted DG-FETs suggests that the centering of V_T could be a challenging proposition. Three basic techniques have been explored both theoretically and experimentally, namely, use of body doping, use of asymmetric gate work function, and use of symmetric mid-gap work-function gate-electrodes.

Adequate body doping can be achieved by directly doping the silicon body or by use of "halo" (also known as "pocket") ion implants introduced laterally from the gate edges, or a combination of these two techniques. One technique uniquely available to DG-FETs is the use of asymmetric gates, wherein the two gate electrodes are of materials of differing work functions.

Body doping has been the technique of choice for V_T centering in both bulk and PD-SOI planar CMOS technologies. Adequate body doping can be achieved by directly doping the silicon body or by use of "halo" ion implants introduced laterally from the gate edges, or a combination of these two techniques.

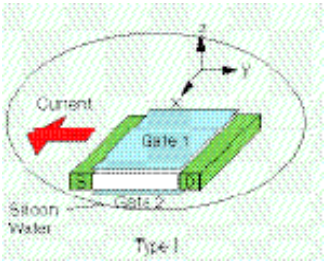
Metal gates offer the possibility of centering threshold voltage with asingle work function for both gate electrodes without relying on body doping. Many metals with workfunctions near the middle of the silicon bandgap exist. Use of these metals in SG-FETs is problematic since the V_T of such devices is typically above 0.5 V , which is too high for most CMOS applications. Lower V_T may be achieved by counter-doping the body, which results in buried conduction channels in the off-state, thus degrading short channel effects. Metal gates on DG-FETs, on the other hand, naturally achieve the V_Ts in the vicinity of 0.2 V(the exact value depending on the details) and good short channel characteristics.

DOUBLE-GATE TAXONOMY

Numerous structures for DG-FETs have been proposed and demonstrated. These structures may be classified into one of the three basic categories.

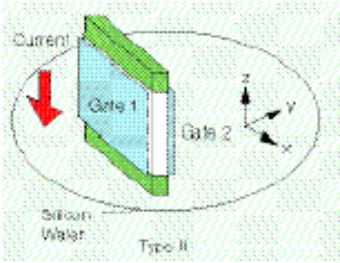
Type I, The Planar DG-FET:

This is a direct extension of a planar CMOS process with a second, buried gate



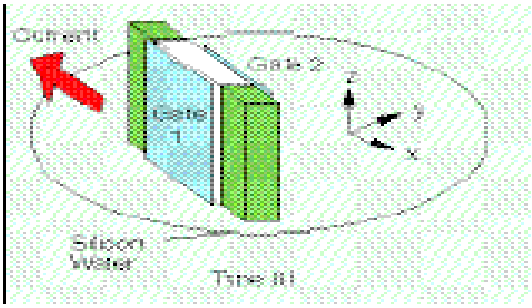
Type II, The Vertical DGFET:

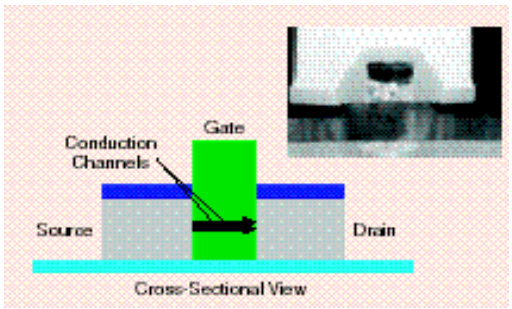
Here bthe silicon body has been rotated to a vertical orientation on the silicon wafer with the source and drain on the top and bottom boundaries of the body, and the gates on either side.

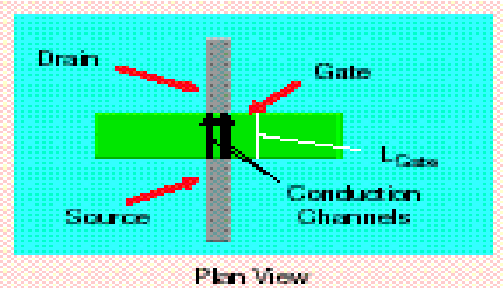


Type III Non Planar FinFET:

In FinFET the silicon body has been rotated on its edge into a vertical orientation so only the source and drain regions are placed horizontally about the body, as in a conventional planar FET. Referred to as FinFETs as the silicon resembles the dorsal fin of a fish.



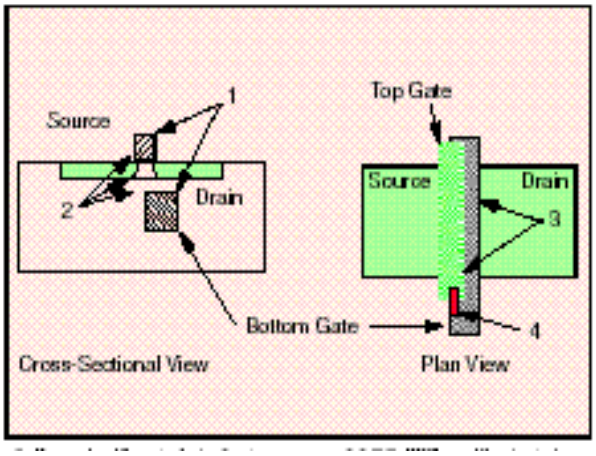




THE DOUBLE-GATE CHALLENGE

DG-FETs have been the subject of much research for over 20 years; hence, if DGCMOS offers significant advantage over SG devices, one must question why DG devices have not played a significant role on the CMOS technology scene to date.

The four major obstacles to DGCMOS are represented schematically.



The first three issues are closely related to one another and consist of

1. definition of both gates to the same image size accurately
2. self-alignment of the source/drain regions to both top and bottom gates

3. alignment of the two gates to one another.

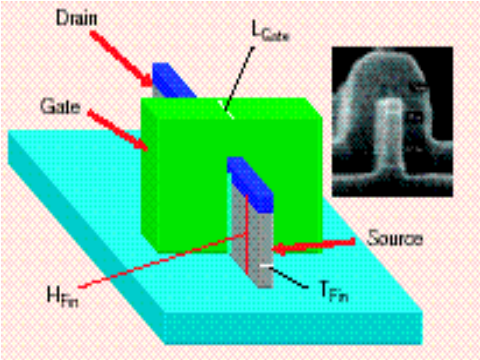
   These three goals are critical for short devices to provide high drive current and low gate capacitance simultaneously.

4. The fourth obstacle is that of providing an area-efficient means of connecting the two gates with a low-resistance path

Type I planar DG-FETs are severely challenged to deliver all of the first three requirements since the "second" gate is buried below a layer of active silicon. The fourth hurdle also challenges the planar DG-FET; a process module is required to define the additional contact to the buried gate if space is not to be lost to it, and a low-resistance gate material must be introduced in the buried oxide.

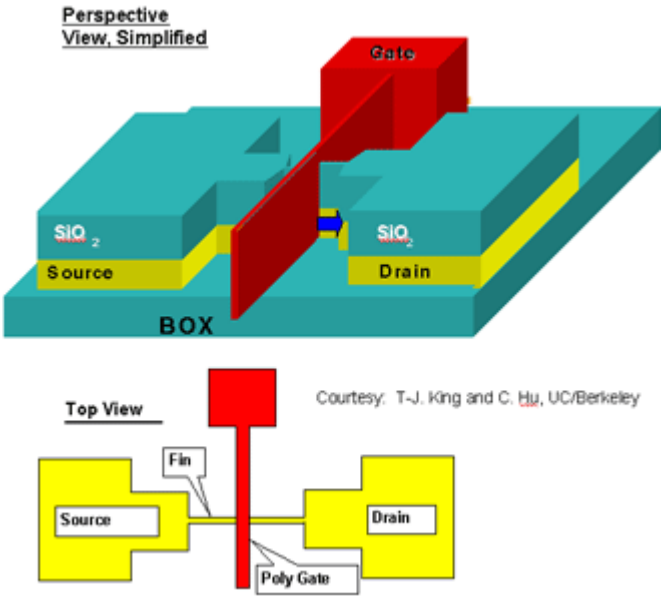
Type II vertical DG-FETs typically address problems 1 and 4 quite successfully. In this case the gate length is usually defined by the thickness of a deposited gate-electrode material, which automatically makes both gates the same length and self-aligned to each other. Similarly, the source and drain junctions can be symmetrically defined to have the same alignment to both gates; however, unique challenges are presented to defining both self-alignment of the bottom junction to the gates and to keeping the parasitic series resistances associated with the bottom junction low. Furthermore, a space-efficient low-capacitance contact scheme to the lower junction requires a high-wire act in process integration. While high drive currents have been achieved with Type II structures, high performance (e.g., low capacitance) and CMOS integration have met with limited progress.

Type III vertical fin-type DG-FETs have the advantages access to both gates, and both sides of source and drain, from the front of the wafer. Gate length is conventionally defined since the direction of the current is in the wafer plane. Gate width, however, is no longer controlled by lithography; rather, the width is given by twice the height of the silicon fin H_{Fin} .



Finfet

FEATURES OF FINFET

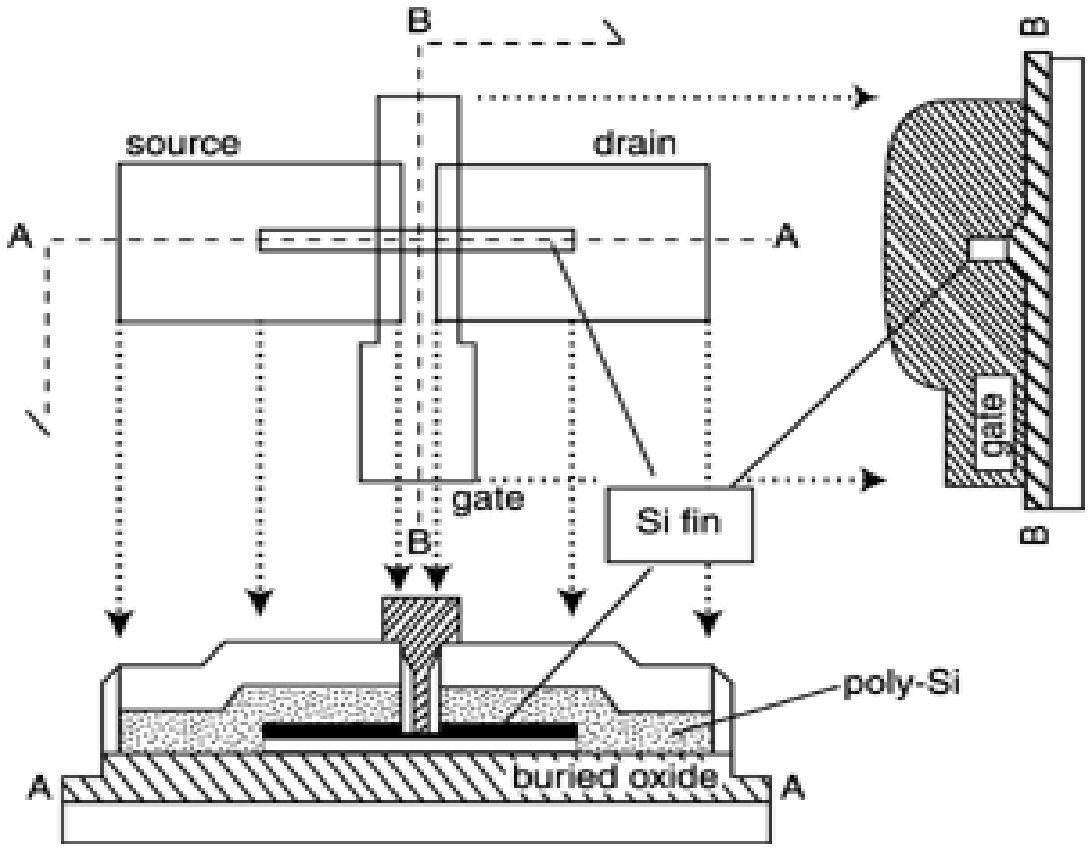


Finfet consists of a vertical Si fin controlled by self_aligned double gate.

Main Features of Finfet are

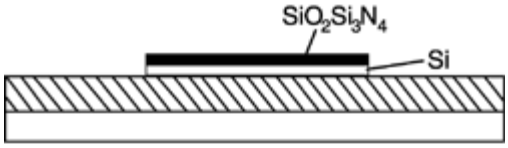
1) Ultra thin Si fin for suppression of short channel effects
2) Raised source/drain to reduce parasitic resistance and improve currrent drive
3) Gate_last process with low_T,high_k gate dielectrics
4) Symmetric gates yield great performance,but can built asymmetric gates that target V_T

Finfets are designed to use multiple fins to achieve larger channel widths.Source/Drain pads connect the fins in parallel. As the number of fins is increased ,the current through the device increases.For eg: A 5 fin device 5 times more current than single fin device.

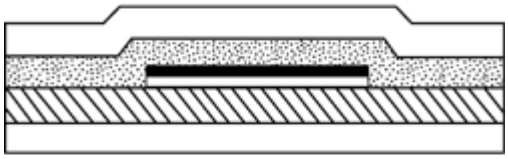


Schematic explaining the parts of a FinFET

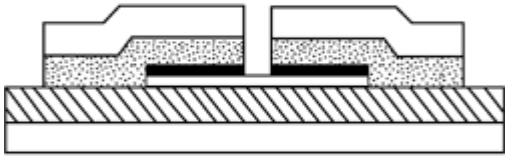
PROCESS FLOW OF FINFET



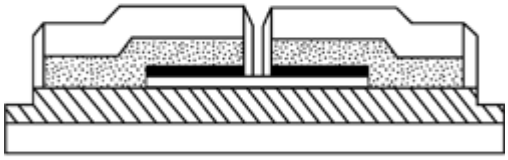
(1) After depositing Si_3N_4 and SiO_2 stacked layer, Si fin was formed.



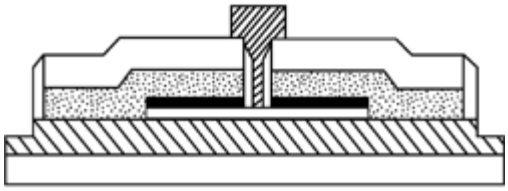
(2) Phosphorus-doped-poly Si and SiO_2 stacked layer deposited.



(3) Source and drain were etched while Si fin was covered by the mask layer.



(4) Spacer SiO_2 layer was etched down into buried oxide layer.



(5) After depositing B-doped SiGe, gate pattern was delineated.

FinFET-DGCMOS Process Flow in Detail

A conventional SOI wafer can be used as starting material, except that the alignment notch of the wafer is preferably rotated 45° about the axis of symmetry of the wafer. The reason for this deviation is to provide{100} planes on silicon fins that are oriented along the conventional "x" and "y" directions on the wafer.

The process of defining fins and source/drain silicon is very similar to that used to define trench isolation in today's CMOS. Patterns are defined and etched into the active top silicon layer in both processes. The conventional process requires additional processing to fill and planarize the isolation trenches; the FinFET process, on the other hand, proceeds directly to channel processing, such as sacrificial oxidations, masked ion implantations for channels, or specialized passive elements, followed by the gate dielectric module.

Gate deposition and etch are very similar, with less-severe demands on the selectivity of the gate-electrode etch to gate oxide, since the oxide surface is orthogonal to the etch direction. Ion implantation of source/drain species and halos(or pockets) must differ for obvious geometrical reason but otherwise are largely similar to conventional planar implantation steps. Conventional $CoSi_2$ or $NiSi_2$ processes are used to silicide the tops of the mesas and the gate, for contacts to source/drain and gate, respectively.

How To Convert Planar To FinFET Technology

As described above, FinFET processing on SOI wafers uses standard manufacturing process modules. To etch the ultra thin (T_{SI}=15nm) fins, spacer lithography [side wall image transfer] is used. Since the SIT process always generates an even number of fins, an extra process step is needed for removal of fins to allow odd number of fins or otherwise break fin" "loops" where needed. This means, that for conversion of an existing design, two additional levels have to be introduced, namely the " fin" and the "Trim" level. All other design levels remain the same.

Consider now a planar design to be converted for processing in the 90 nm FinFET technology node. The FinFET height H_{Fin} together with the fin pitch (determined by photolithography) defines the FinFET device width W_{Fin} within the given silicon width of the planar device, to get the same or better device strength . For automatic Fin and Trim generation, Fin-GEN, a software tool, has been developed, which takes the active area and

poly gate levels, and, based on special FinFET ground rules, generates the additional levels.the circuit (as well as other β-ratio sensitive circuitry) may additionally require manual adjustment on the number of fins in the N- and P-devices after automatic addition of fins in the N- and P- devices after automatic addition of the FinFET levels.

Besides device width quantization, other factors like width variation, threshold variation, and self-heating must be taken into account when designing with FinFETs. A process with multiple threshold voltages and multiple gate oxide thickness is required to take full advantage of this new device.

As already discussed, the width quantization imposes some restrictions on the device strength flexibility, but most of them can be absorbed easily when converting an existing design or starting a new design, respectively. Of course, as stated earlier, latches, dynamic circuit styles in general, and SRAM cells need careful optimization when designing with FinFETs.

Discrete devices and circuits for analog applications require special attention. As an example, consider a driver/receiver circuit with an ESD protection diode. In a planar process the protection voltage is proportional junction length of the diode. In FinFET technology the same junction length per fin pitch may be only about one-eigth of that of the planar device.

Another example is the total output driver impedance matching, which is usually implemented with a planar resistor requiring a silicon block resistor, on a silicon island to adjust output impedance (including the wire to the pad) to 50 Ω. For such applications, and analog circuits in general, special devices may be necessary for optimized designs using FinFETs.

APPLICATIONS OF FinFETs

DG devices like Fin FETs offer unique opportunities for microprocessor design.compared to a planar process in the same technology node, FinFETs have reduced channel and gate leakage currents. This can lead to considerable power reductions when converting a planar design to fin FET technology. Utilizing fin FETs would lead to a reduction in total power by a factor of two, without compromising performance.

Another possibility to save power arises when both gates can be controlled separately. The second gate can be used to control the threshold voltage of the device, thereby allowing fast switching on one side and reduced leakage currents when circuits are idle.

Finally, separate access to both gates could also be used to design simplified logic gates. This would also reduce power, and save chip area, leading to smaller, more cost-efficient designs. However chip designs using finFETs must cope with quantization of device width, since every single transistor consists of an integral number of fins,each fin having the same height.

SIMULATION OF VERTICAL DOUBLE-GATE SOI MOSFETS USING DEVICE3D

Introduction

This article will present the simulation methodology of a self-aligned double-gate MOSFET structure (FinFET) using SILVACO 3-D simulation suite. The double-gate MOSFET is one of the most attractive alternative to classical MOSFET structure for gate length down to 20nm. The main advantage of the FinFET is the ability to drastically reduce the short channel effect. In spite of his double-gate structure, the FinFET is closed to its root, the conventional MOSFET in layout and fabrication. 3-D numerical simulations of the FinFET are performed in this article, in order to validate the basic principles and to uncover several important aspects: evaluation of the length , width and quantum effects.

Device Features

The features of the structure are shown in Figure 1 are: (1) a transistor is formed in a vertical ultra –thin Si fin and is controlled by a double-gate, which considerably reduced short channel effects; (2) the two gates are self aligned and are aligned to S/D; (3) S/D is raised to reduce the access resistance; (4) Up to date gate process: low temperature, high -k dielectrics can be used and (5) the structure is quasi-planar because Si Fin is relatively short [1,2].

Device Simulation

The 3-D SILVACO simulation suite including ***Device3D, DevEdit3D*** and ***TonyPlot3D,*** allows device engineers to study deep sub-micron devices which are 3-D by nature like the FinFET presented above. Furthermore, 3-D simulations give access to data impossible to measure like charge distribution, potential, electric field and current lines.

A 3-D FinFET structure was designed by using ***DevEdit3D***. This is an advanced tool for structure editing and mesh generation. The device structure was realized by drawing first the FinFET, from the bottom view (Figure 2),in a (x,y) plane before extending it in the z-direction.

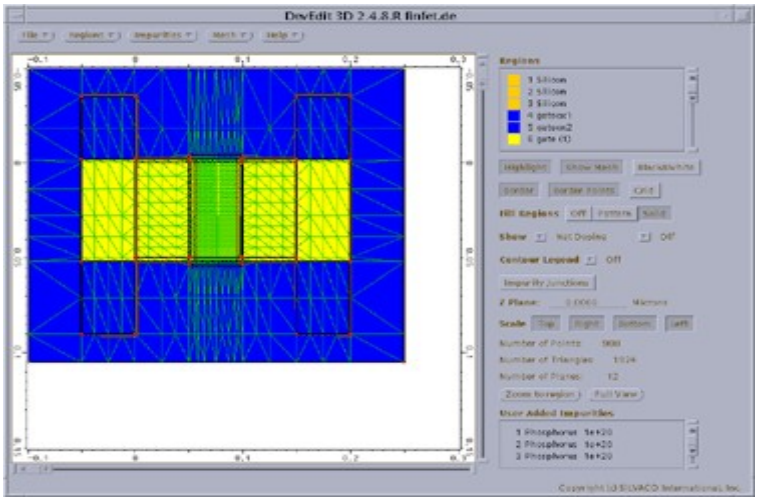


Figure 1: Illustration of DEVEDIT3D used to build the FinFET structure

The z-direction in this case corresponds to the vertical to the substrate. The final 3-D structure is shown in ***TonyPlot3D*** (Figure 3).

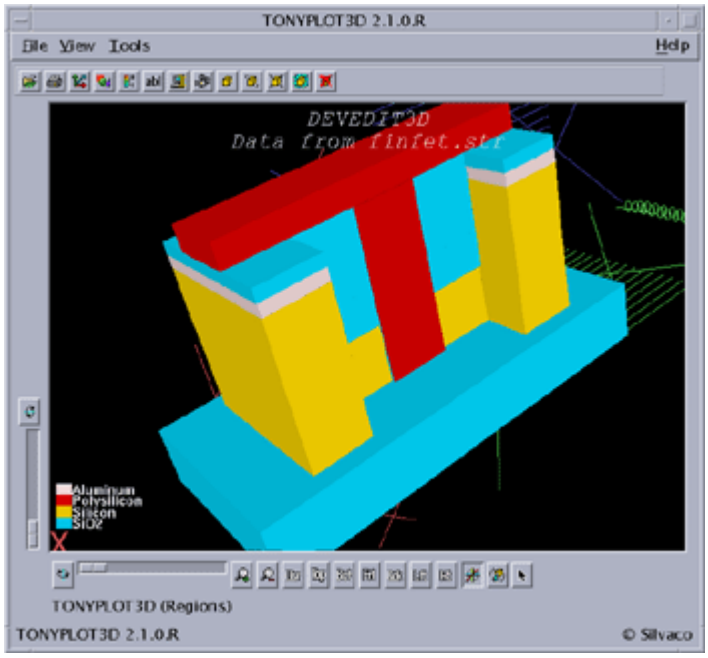


Figure 2: Plot of a 50-nm FinFET 3-D structure for a width of 50nm

The basic characteristics of this FinFET was Tox=2nm length=50nm width=50nm and Fin height=50nm. Note that we have defined a parameterized structure for subsequent use in our automation tool, which make much more easier any kind of variation (length, width) to perform large scale simulation.

The main physical effects (mobility, carrier statistics, recombination) were expressed by a set of models universally used for simulating the MOS technology: mobility dependence of the electric field and doping level, Boltzmann statistics and Schokley-Read-Hall generation recombination mechanisms [3].

Simulation Results

Typical I-V characteristics of a 50-nm gate length are shown in Figure 4. The leakage current caused by DIBL was well suppressed.

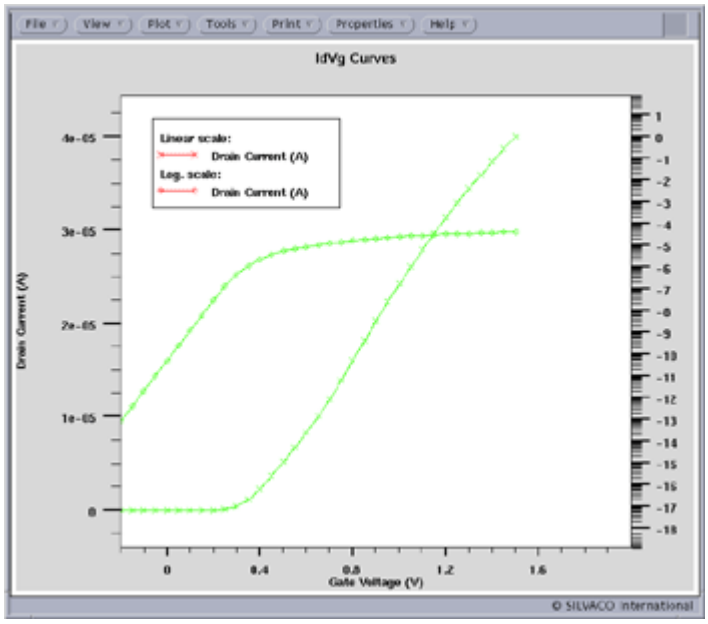


Figure 3: 50-nm FinFET IdVg curves for a width of 50nm.

The roll-off of a FinFET with a width of 50nm is well controlled as can be seen in Figure 5. This result can be correlated to the good control of the channel potential due to the double gate.

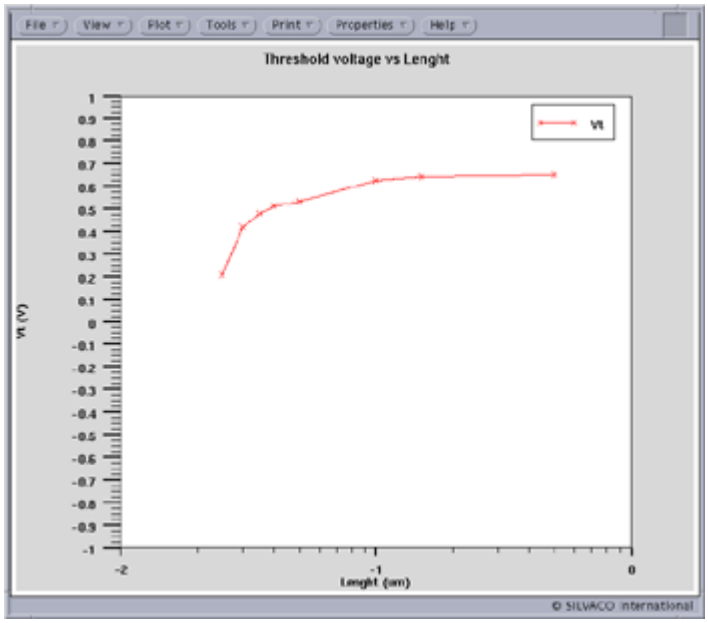


Figure 4: Threshold voltage as a function of gate length for a width of 50-nm.

The width of this FinFET is adjusted by the number of Si fins. Let say you want to double the width of your device then you have to put 2 Fins between source and drain (Figure 6).

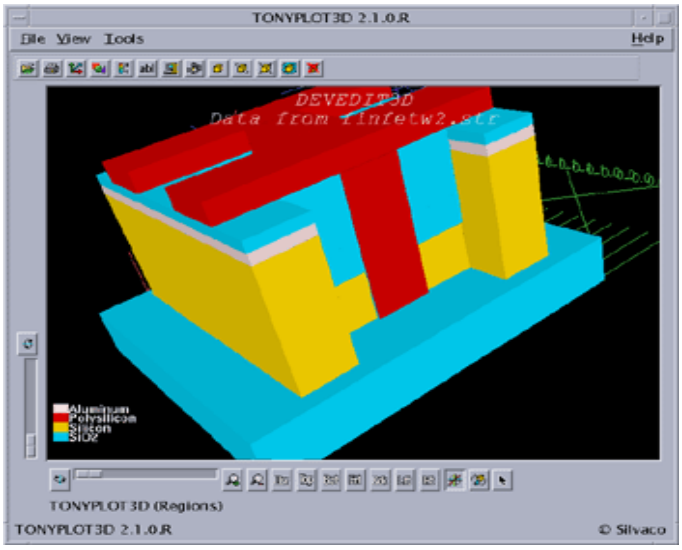


Figure 5: Structure of a 2-parallel channel device. Gate length 50-nm.

Note that this can be achieved very simply using the "mirror" feature in ***DevEdit3D***. The resulting I-V curve can be seen

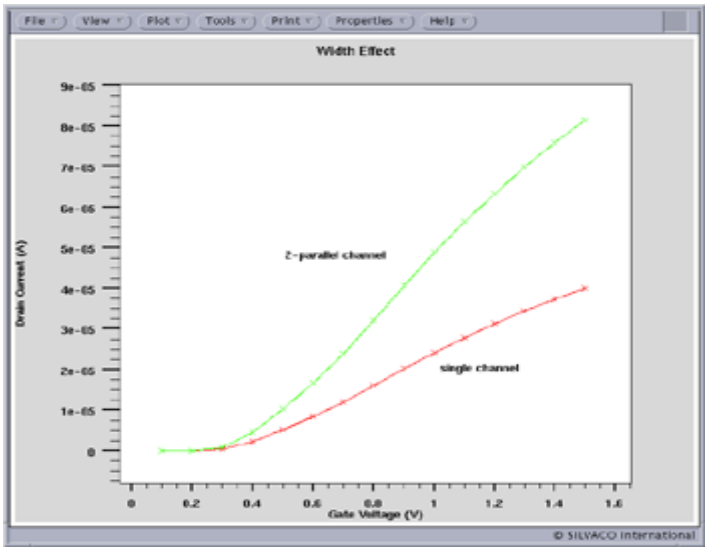


Figure 6: Drain current comparison between single and 2-parallel channel device. Gate length 50-nm.

Finally we have made simulations using our quantum module named ***Quantum3D***. The result is plotted in Figure 8. One can see a shift in the threshold voltage indicating some quantum effect. This correction is quite small as indicated in [2].

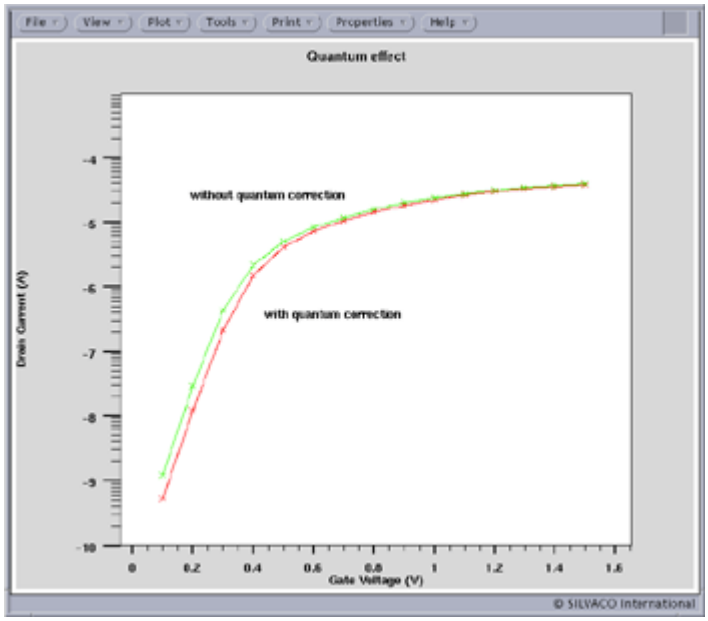


Figure 7: Quantum effect in a 50-nm with a width of 50nm.

Sub 50-nm FinFETs were successfully simulated using 3-D SILVACO simulation tools. It is very easy to study the impact of the geometry and doping of this 3-D device using ***Device3D***. Indeed more and more people take a look at this novel structure since it is an attractive successor to the single-gate MOSFET.

CONCLUSIONS

Simulations show that this structure should be scalable down to 10 nm. Formation of ultra thin fin (0.7 L_g, for a lightly doped body) is critical for suppressing short channel effects. This structure was fabricated by forming the S\D before the gate, a technique that may be needed for future high-k dielectric and metal-gate technologies that cannot tolerate the high temperatures required for S\D formation. Further performance improvement is possible by using a thinner gate dielectric and thinner spacers. Despite its double gate structure, the FinFET is similar to the conventional MOSFET with regard to layout and fabrication. It is an attractive successor to the single gate MOSFET by virtue of its superior electrostatic properties and comparative ease of manufacturability. Industrial research groups such as Intel, IBM and AMD have shown interests in developing similar devices, as well as mechanisms to migrate mask layouts from Bulk-MOS to FinFETs. Issues such as gate work function engineering, high quality ultra thin fin lithography and source\drain resistance need to be resolved and a high-yield process flow needs to be established by process researchers before FinFETs can be used in commercial ICs. Device researchers need to understand and model quantum effects, and circuit design researchers need to exploit the packing density afforded by the quasi-planar device to design efficient architectures.

REFERENCES:

1. S. Thompson, P. Packan and M. Bohr, "MOS scaling , Transistor challenges for the 21st century," Intel Tech.J.,vol.Q3, pp1-19,1998
2. C.H.Wann, H. Noda, T. Tanaka, M.Yoshida and C. Hu, "A comparative study of advanced MOSFET concepts ," IEEE Trans. Electron Devices, vol. 43, no. 10, pp 1742-1753, Oct. 1996
3. D.Hisamoto, W.C. Lee, J.Keidzerski, H.Takeuchi, K.Asano, C.Kuo, T.J.King, J.Bokor and C.Hu, "A folded channel MOSFET for deep-sub-tenth micron era,"in IEDM Tech. Dig. 1998, pp 1032-1034
4. D.Hisamoto, W.C. Lee, J.Keidzerski, H.Takeuchi, K.Asano, C.Kuo. T.J.King, J.Bokor and C.Hu, "FinFET-a self-aligned double-gate MOSFET scalable beyond 20 nm," IEEE Trans.Electron Devices, vol.47, pp. 2320-2325, Dec. 2000.
5. X. Huang, W.C. Lee, C.Kuo, D.Hisamoto, L. Chang, J. Keidzerski, E. Anderson, H.Takeuchi, Y.K. Choi, K.Asano, V.Subramanian, T.J.King, J.Bokor, C.Hu, " Sub-50 nm FinFET: PMOS," in IEDM Tech. Dig.1999.,pp 67-70
6. H.S. Wong, K. Chan and Y.Taur, "Self-aligned (top and bottom) double-gate MOSFET with a 25 nm thick silicon channel," in IEDM Tech.,Dig.1997,pp. 427-430
7. J. Hergenrother et al, " The vertical replacement-gate (VRG) MOSFET: A 50 nm vertical MOSFET with lithography-independent gate length," in IEDM Tech. Dig. 1999 , pp. 75-78

CONTENTS

ACKNOWLEDGEMENT

First, and foremost I thank **God Almighty** for making this venture a success.

I extend my sincere gratitude to **Prof. N. Premachandran**, Principal, Govt. Engineering College, Thrissur and **Prof. K. P. Indiradevi,** Head of Electronics & Communication Department, Govt. Engineering College, Thrissur for providing me with necessary infrastructure.

I would like to convey a deep sense of gratitude to the seminar coordinator **Mrs. C. R. Muneera** for the timely advices.

I also extend my sincere thanks to my friends and seniors for their help.

ABSTRACT

The introduction of FinFET Technology has opened new chapters in Nano-technology. Simulations show that FinFET structure should be scalable down to 10 nm. Formation of ultra thin fin enables suppressed short channel effects. It is an attractive successor to the single gate MOSFET by virtue of its superior electrostatic properties and comparative ease of manufacturability.