INFORMATION AND COMMUNICATION TECHNOLOGIES

COORDINATION AND SUPPORT ACTION

EUROSOI+

European Platform for Low-Power Applications on Silicon-On-Insulator Technology

Grant Agreement nº 216373

D4.15 EUROSOI+ Focused Reports: Benchmark of UTB vs. FinFET vs. DG SOI transistors. Which SOI device is favourite and for which application?

Due date of deliverable: 30-06-2008
Actual submission date: 24-07-2008

Start date of project: 01-01-2008 Duration: 39 months

Project coordinator: Prof. Francisco Gámiz, UGR
Project coordinator organisation: University of Granada, Spain Rev.1

Project co-funded by the European Commission within the Seventh Framework Programme (FP7)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of contents

1.- INTRODUCTION

The silicon MOSFET has been scaled down for more than thirty years to obtain higher performance and higher levels of integration. Recently, the miniaturization rate has accelerated, and the gate length is now less than 40 nm. Further miniaturization of silicon metal-oxide semiconductor field-effect transistors (MOSFETs) into nanoscale complementary MOS (CMOS) will significantly affect advances in future information technology. The International Technology Roadmap for Semiconductors (ITRS) [1] predicts that gate lengths in mass-produced CMOS transistors will be less than 10 nm in the year 2016. However, it is well recognized that it will be extremely difficult for bulk-Si MOSFET technology to meet industry-specified performance targets for both drive current and leakage current in a near future. New techniques and architectures to overcome the scaling and performance limits of conventional CMOS devices are urgently needed. One of the most promising avenues is the utilization of SOI-like device structures. Indeed the use of very thin silicon films enables excellent scalability due to superior short channel-effect (SCE) immunity.

This report describes three specific devices, namely, ultra-thin body MOSFET, Double-gate SOI MOSFET and FinFETs as potential technological boosters for future device scaling.

2.- FROM BULK TO THIN FILM

Bulk devices require a heavy channel doping to control short-channel effects. With a heavily doped channel, carrier mobility is severely degraded due to impurity scattering and an increased transverse electric field. Above a concentration of $10^{18}cm^{-3}$, mobility is expected to be noticeably affected by channel dopants [2]. In addition, significant depletion charge in the channel will form, thus increasing the average vertical field experienced by carriers in the inversion layer and increasing the effects of phonon and interface scattering [3]. Furthermore, this increased depletion charge will result in a larger depletion capacitance and subthreshold slope. As a result, for a given off-state leakage current specification, the threshold voltage must be raised, thus reducing the on-state drive current. A large channel doping will also inevitably enhance band-to-band tunneling leakage between the body and drain [4]. This will be especially critical because abrupt halo doping profiles in the channel are currently used to localize the heavy channel doping whereas abrupt drain doping profiles are desirable for the reduction of series resistance. Together, these effects will greatly increase band-to-band tunneling, which could eventually become the dominant off-state leakage mechanism in the transistor.

Advanced MOSFET structures (Fig. 1) such as the ultrathin-body (UTB) silicon-on-insulator (SOI) single-gate transistor and the double-gate (DG) transistor can be scaled more aggressively than the classic bulk-Si structure [5, 6] and, hence, may be adapted for IC massive production [7]. These structures do not rely on heavy doping; however, they make use of different materials, process technology requirements and associated challenges.

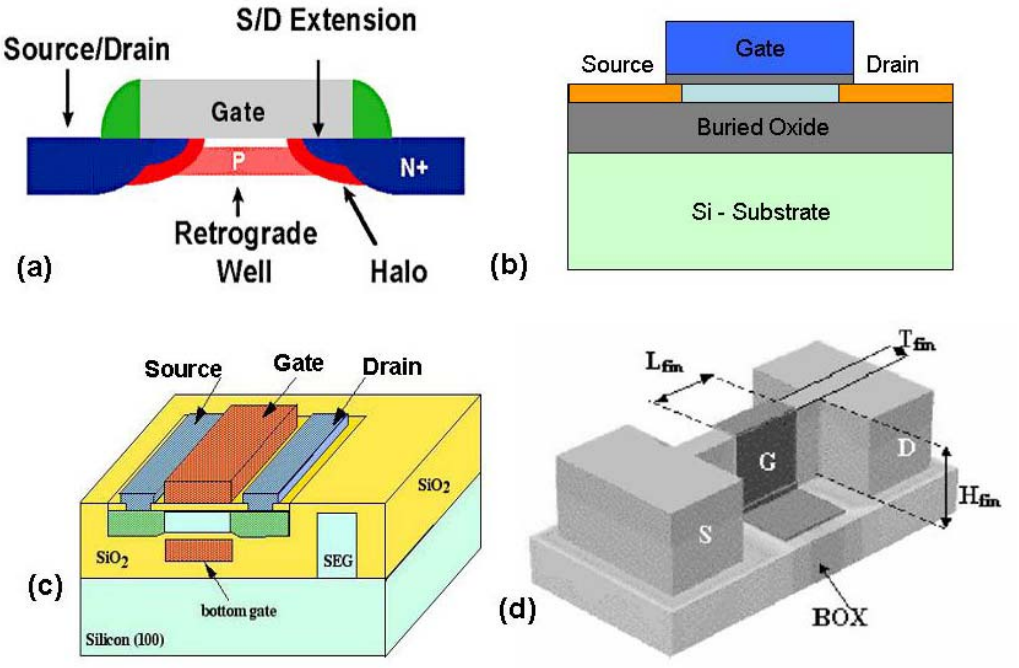

Figure1: Scaled silicon devices: (a) advanced bulk MOS, (b) Ultra thin body SOI, (c) Planar double gate SOI, and (d) FinFET (Portions of this figure were taken from references [22] and [23])

In UTB and DG devices, short-channel effects are controlled by a thin silicon film, thus allowing for gate-length scaling down to the 10-nm regime [8] without the use of channel dopants. With a lightly doped channel, DG and UTB devices have negligible depletion charge and capacitance, which yields a steep subthreshold slope. DG devices show even better subthreshold slope than their UTB counterparts due to better short-channel effect control by the DG structure. Lower transverse electric field and negligible impurity scattering contribute to increased mobility that further improves the drive current in both devices. With respect to circuit performance, DG and UTB devices provide an additional advantage in that the capacitive load is decreased by the elimination of both depletion and junction capacitances.

For DG and UTB devices, the enhancement in drive current at a given off-state current specification leads directly to an improvement in inverter delay. An additional speedup (5%–10%) results from the elimination of depletion and junction capacitances.

3.- ULTRA-THIN BODY MOSFETs

For UTB MOSFETs, short-channel effects are strongly dependent on the body thickness (T_{Si}), and can be evaluated by the ratio of the gate length (L_G) to the body thickness. SCEs are sufficiently suppressed when the ratio between the two device dimensions (L_G /T_{Si}) is larger than 4 and this criterion has been tested for different experimental devices. This value may be further reduced by scaling the gate-oxide thickness or increasing the channel doping concentration. Nevertheless, metal gate work-function engineering is necessary to adjust and control the threshold voltage of UTB MOSFETs.

In UTB devices, the body thickness should be as thin as possible for ultimate scalability; however, the series resistance of the source and drain regions could limit transistor drive currents [9, 10]. By using selective deposition, raised source/drain structures can be created to reduce the parasitic resistance. One drawback of this structure, however, is that it inevitably results in increased overlap capacitance. Thus, spacer width between the gate and raised source/drain should be optimized.

A recent trend is to use ultrathin BOX (10 nm) which contributes to reducing SCEs.

4.- DOUBLE-GATE FINFET

The DG device is electrostatically more robust than a single-gate UTB MOSFET because two gates are used to control the channel from both sides, thus allowing for additional gate length scaling by at least a factor of two. The addition of a second gate electrode not only halves the effective body thickness of the device and enables twice as thick films to be used (L_G /T_{Si} ≈ 2), but also eliminates the penetration of the drain electric field through the buried oxide,

which improves gate control of the channel. Numerous methods have been proposed and demonstrated in the past to fabricate DG devices [11, 12]; however, many suffer from process complexity.

The FinFET was, thus, proposed as a more practical DG structure with a simpler and more manufacturable process similar to conventional SOI CMOS processes. The FinFET uses a single-gate material deposited over a silicon fin to form perfectly aligned gates along the fin sidewalls. The fin width is the most important process variable because it determines the body thickness, which governs short-channel effects. Another distinct advantage of ultra-narrow fins is that the lateral gates can control the potential at the bottom interface (fin-BOX), making the device to be immune to radiation effects and to the field penetration from drain into the BOX.

The off-state leakage current density increases dramatically as the fin width increased because gate control of the channel is worsened [8]. Channel mobility [13] and threshold voltage [14, 15] can also be sensitive to this dimension. It is, thus, important to achieve small and controllable dimensions for the fin width. This can be accomplished using optical or electron beam lithography, but both inevitably lead to critical dimension (CD) variation and line edge roughness (LER). A key issue for the FinFET is that adequate suppression of short-channel effects requires that the fin width be approximately half of the gate length [8, 16] such that a sublithographic patterning technology is needed for fin formation. This is a clear departure from historic device scaling where gate length was at the limit of lithographic capabilities.

In order to obtain higher drive current multiple fins must be placed parallel all straddled by a single-gate line [17]. The achievable fin pitch, thus, determines the amount of layout area required for a device.

Due to their vertical nature, the Fin-channels are formed on the side surfaces of the fin, therefore, crystal orientation and surface quality become critical issues. The mobility dependence on the crystal orientations in FinFETs have attracted much attention. For example, sidewall channels in the (110) plane when oriented parallel and perpendicular to the wafer flat of a standard (100) wafer can achieve enhanced hole mobility as compared with (100) silicon surfaces while electron mobility is degraded in (110) surfaces.

5.- SCALING LIMITS

For digital circuit applications, an ultimate limit will eventually be determined by the ability to maintain proper transistor operation, namely, an ultimately scaled device must have the capability to be turned both on and off. With scaling of transistor gate lengths, the primary difficulty lies in the control of the off state leakage current.

The leakage current of a MOSFET in the off state will be comprised of three different components: Thermionic emission above the channel potential barrier, band-to-band tunneling between the body and drain, and quantum mechanical tunneling between the source and drain.

Thermionic emission is primarily controlled by the channel potential barrier height, which, in turn, is affected by SCEs. Band-to-band tunneling is controlled by the power supply voltage and the gradient of the S/D doping profile. Direct tunnelling of carriers from the source to drain may also occur at extremely small gate lengths because the channel potential barrier width is very small.

By calculating these three leakage current components, the off-state leakage can be estimated. Scaling of the gate length results in increased leakage current because gate control of the channel is reduced, allowing for increased DIBL which reduces the barrier height of the channel potential barrier, and thus enhancing the thermionic emission, the dominant leakage component. Therefore, the scaling limit of UTB and DG MOSFETs is a strong function of the body thickness.

The minimum acceptable body thickness for a thin-body MOSFET is currently considered to be around 5 nm [18]. This is because the series resistance in the source and drain regions introduced by such a thin film may become unacceptable even with the incorporation of a raised source/drain technology. In addition, quantum confinement in the thin body may cause an intolerable shift in the threshold voltage.

Scaling limit for DG MOSFETs can be obtained under the assumption that body thicknesses cannot be scaled below 5 nm. With the additional assumption that equivalent oxide thickness scaling will be limited to 1 nm (due to gate leakage current even with alternative gate dielectric materials), off state leakage current targets can be met at gate lengths down to 10 nm. Triple-gate FINFETs with Ω configuration are rather similar to gate-all-around (GAA) transistors which offer an additional window for scaling used ($L_G \approx 5$ nm).

As a conclusion, the DG MOSFET structure (in its FinFET version) has been shown to be the most scalable transistor design due to its dual-gate nature, in which the two gate electrodes provide significant control of the channel.

6.- CIRCUITS

In order to assess the potential circuit performance of the aforementioned devices, L. Chang et al have carried out 2-D device simulations to perform mixed-mode simulation of various benchmark circuits [6, 7]. Three benchmark ring oscillator circuits were simulated for FinFETs, UTB, and bulk devices: a fan-out of 4 (FO4) inverter, a NAND pull-down stack, and a pass-gate multiplexer. The FO4 inverter delay is a standard technology benchmark used to predict the delay of more complex circuits. The NAND and pass-gate structures consider the impact of the body effect on performance.

As a result of these simulations the DGSOI-FinFET outperforms the UTB and Bulk-MOSFETs for all supply voltages and circuits due to a combination of improved subthreshold swing, higher carrier mobility, and reduced gate capacitance from simultaneous switching of the two gates. The UTB device

shows slightly degraded subthreshold swing and does not benefit from the reduction in gate-switching capacitance. The bulk MOSFET, with its doping substrate and nonideal subthreshold swing, is significantly slower than FinFETs and UTB MOSFETs.

Moreover, double gate devices with independent gate control option (separate contacts for back and front gates) have been developed [19, 20]. Such DG devices are referred to as Independent or Isolated Gate (IG) devices. For example, back gate bias can be used to dynamically adjust the threshold voltage of the front gate to tune the power and performance requirement of a circuit. Independent control of the front and the back gate can be very attractive for circuit design.

As an additional advantage, it has been shown that a circuit designed with bulk-CMOS transistors can be directly translated to DG technology by replacing each transistor with a DGMOS, where the front and the back gates are tied together. However, the directly translated circuit style does not allow the possibility of independent control of front and back gates.
Finally, it has been shown that FinFETs can also be effective for low-power SRAM circuits [21]. Optimization of the gate sidewall spacer thickness is suggested to simultaneously minimize leakage current. Optimization of spacer thickness results in around 70% reduction in SRAM cell leakage and reduced cell read failure probability (by 200X) compared to a conventional FinFET SRAMs [22].

7.- CONCLUSIONS

Significant advantages in double-gate MOSFET device technology and performance have been demonstrated through the use of the fin channel type of device architecture. This is because the fin-channel structure circumvents the obstacles that are common to the double gate MOSFET.
The major obstacles are: self-alignment of the source/drain regions with both top and bottom gates, alignment of the two gates to one another and connection of the two gates without area penalty. The obstacles are closely related to the reduction of parasitic resistance and capacitance that is required to obtain high performance for practical applications. The planar channel devices have serious problems surmounting these three obstacles and the vertical-channel devices face challenges in self-alignment of the source/drain regions with both top and bottom gates. Because of geometrical configuration, fin-channel devices can circumvent the obstacles the same as conventional bulk-CMOS.

Moreover, the DG device shows a larger enhancement over traditional bulk-Si MOSFETs than the UTB devices because of improved short-channel effects. All above conclusions apply to any SOI-like technology including the recent GeOI approach.

8.- REFERENCES

[1] The International Technology Roadmap for Semiconductors (ITRS), Semiconductor Industry Association, San Jose, CA.

[2] T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr, “Scaling challenges and device design requirements for high performance sub-50 nm gate length planar CMOS transistors,” in *Symp. VLSI Technology Dig. Tech. Papers*, 2000, pp. 174–175.

[3] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, “On the universality of inversion layer mobility in Si MOSFET’s: part I—effects of substrate impurity concentration,” *IEEE Trans. Electron Devices*, vol. 41, pp. 2357–2362, Dec. 1994.

[4] Y. Taur, C. H. Wann, and D. J. Frank, “25 nm CMOS design considerations,” in *Int. Electron Devices Meeting Tech. Dig.*, 1998, pp. 789–792.

[5] C. H. Wann, K. Noda, T. Tanaka, M. Yoshida, and C. Hu, “A comparative study of advanced MOSFET concepts,” *IEEE Trans. Electron Devices*, vol. 43, pp. 1742–1753, Oct. 1996.

[6] L. Chang, Y-K Choi, D. Ha, P. Ranade, S. Xiong, J. Bokor, C. Hu, T-J King, “Extremely Scaled Silicon Nano-CMOS Devices” Proceedings of the IEEE, 91(11), Nov 2003, pp. 1860-1873.

[7] L. Chang, Y-K Choi, J. Kedzierski, N. Lindert, P. Xuan, J. Bokor, C. Hu, and T-J King, “Moore´s Law lives on” IEEE Circuits & Devices Magazine, January pp. 35-42, 2003.

[8] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu, “Gate-length scaling and threshold voltage control of double-gate MOSFETs,” in Int. Electron Devices Meeting Tech. Dig., 2000, pp. 719–722.

[9] S.-D. Kim, C.-M. Park, and J. C. S. Woo, “Advanced model and analysis of series resistance for CMOS scaling into nanometer regime—part II quantitative analysis,” *IEEE Trans. Electron Devices*, vol. 49, pp. 467–472, Mar 2002.

[10] B. Doyle, R. Arghavani, D. Barlage, S. Datta, M. Doczy, J. Kavalieros, A. Murthy, and R. Chau, “Transistor elements for 30 nm physical gate lengths and beyond,” *Intel Tech. J.*, vol. 6, pp. 42–54, May 2002.

[11] J.-H. Lee, G. Taraschi, A.Wei, T. A. Langdo, E. A. Fitzgerald, and D.A. Antoniadis, “Super self-aligned double-gate (SSDG) MOSFET’s utilizing oxidation rate difference and selective epitaxy,” in *Int. Electron Devices Meeting Tech. Dig.*, 1999, pp. 71–74.

[12] H.-S. P. Wong, K. K. Chan, and Y. Taur, “Self-aligned (top and bottom) double-gate MOSFET with a 25 nm thick silicon channel,” in *Int. Electron Devices Meeting Tech. Dig.*, 1997, pp. 427–430

[13] M. Shoji and S. Horiguchi, “Electronic structures and phonon limited electron mobility double-gate silicon-on-insulator Si inversion layers,” *J. Appl. Phys.*, vol. 85, pp. 2722–2731, 1999

[14] Y.-K. Choi, D. Ha, T.-J. King, and C. Hu, “Nanoscale ultrathin body PMOSFET’s with raised selective germanium source/drain,” *IEEE Electron Device Lett.*, vol. 22, pp. 447–448, Sept. 2001.

[15] H. Majima, H. Ishikuro, and T. Hiramoto, “Threshold voltage increase by quantum mechanical narrow channel effect in ultra-narrow MOSFETs,” in *Int. Electron Devices Meeting Tech. Dig.*, 2001, pp. 379–382.

[16] D. Hisamoto, T. Kaga, and E. Takeda, “Impact of the vertical SOI ‘DELTA’ structure on planar device technology,” *IEEE Trans. Electron Devices*, vol. 38, pp. 1419–1424, June 1991.

[17] S. Tang, L. Chang, N. Lindert, Y.-K. Choi, W.-C. Lee, X. Huang, V. Subramanian, J. Bokor, T.-J. King, and C. Hu, “FinFET: a quasiplanar double-gate MOSFET,” in *2001 IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2000, pp. 118–119.

[18] D. J. Frank, S. E. Laux, and M. V. Fischetti, “Monte Carlo simulation of a 30 nm dual-gate MOSFET: how short can Si go?,” in *Int. Electron Devices Meeting Tech. Dig.*, 1992, pp. 553–556.

[19] D. Fried, et. al, “A Fin-type independent-double-gate NFET”, Device Research Conference, 2003, pp. 45-46.

[20] L. Mathew, et. al, “CMOS vertical multiple independent gate field effect transistors (MIGFET)”, Int. SOI Conf., 2004, pp. 187-188.

[21] H. Ananthan and K. Roy, “Technology and circuit design considerations for width-quantized quasi-planar double-gate SRAM,” *IEEE Trans. Electron Devices*, vol. 53(2), pp. 242–250, 2006.

[22] K. Roy, H. Mahmoodi, S. Mukhopadhyay, H. Ananthan, A. Bansal, and T. Cakici, “Double-Gate SOI Devices for Low-Power and High-Performance Applications”, Proc. Int. Conference on Computer Aided Design, pp. 217-224, 2005.

[23] F. Dauge, J. Pretet, S. Cristoloveanu, A. Vandooren, L. Mathew, J. Jomaah, B.Y. Nguyen; Solid-State Electronics 48 (2004) 535–542.