FinFETs for Nanoscale CMOS Digital Integrated Circuits

Tsu-Jae King
Advanced Technology Group, Synopsys, Inc.
Mountain View, California, USA
tsujae@synopsys.com

Abstract

Suppression of leakage current and reduction in device-to-device variability will be key challenges for sub-45nm CMOS technologies. Non-classical transistor structures such as the FinFET will likely be necessary to meet transistor performance requirements in the sub-20nm gate length regime. This paper presents an overview of FinFET technology and describes how it can be used to improve the performance, standby power consumption, and variability in nanoscale-CMOS digital ICs.

Keywords

MOSFET, thin-body transistor, CMOS, SRAM, digital IC

INTRODUCTION

The steady miniaturization of the metal-oxide- semiconductor field-effect transistor (MOSFET) with each new generation of complementary-MOS (CMOS) technology has yielded continual improvements in integrated-circuit (IC) performance (speed) and cost per function over the past several decades, to usher in the Information Age. Continued transistor scaling will not be as straightforward in the future as it has been in the past, however, because of fundamental materials and process technology limits [1]. Suppression of leakage current to minimize static power consumption, and reduction in device-to-device variability to increase yield (and thereby lower cost), will be key challenges for transistor scaling in the sub-20nm gate length (L_G) regime. This paper describes how these challenges can be mitigated by adopting the "FinFET" transistor structure, in order to sustain the rapid growth of the industry and usher in the age of ambient intelligence and ubiquitous computing.

BULK-SI MOSFET SCALING CHALLENGES

In order to scale the classical bulk-Si MOSFET structure (Fig. 1a) down to ~10nm L_G, heavy halo and channel doping (greater than 1×10^{18} cm^{-3}) will be required to suppress leakage current and short-channel effects [2]. As a result, field-effect carrier mobilities will be degraded, resulting in incommensurate improvements in transistor drive current with L_G scaling [3]. Thin-body transistor structures (Figs. 1b and 1c) rely not on heavy channel doping but on a sufficiently thin channel/body region ($T_{Si} < L_G$) to limit leakage current. The use of a lightly doped or undoped channel provides immunity to variations in threshold voltage (V_T)

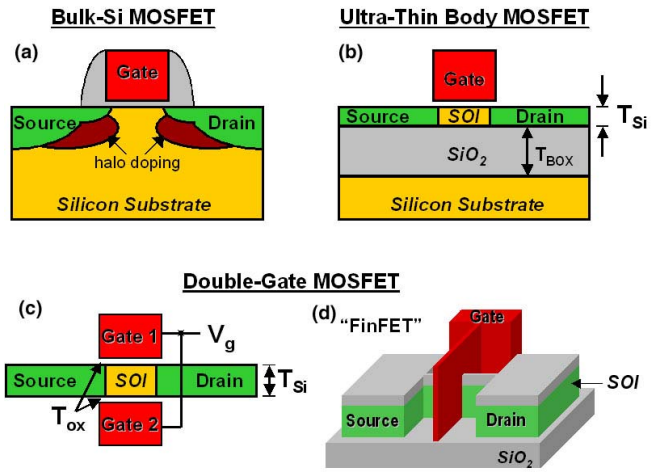

Figure 1: Schematic diagrams of MOSFET structures (a) classical bulk-Si, (b) ultrathin-body (UTB), (c) double-gate (DG), and (d) FinFET

resulting from statistical dopant fluctuations in the channel, as well as enhanced carrier mobility for higher transistor drive current because of the lower transverse electric field in the inversion layer. Therefore, thin-body MOSFETs offer improved circuit performance as compared to the bulk-Si transistor structure [4]. V_T adjustment can be achieved without channel doping by tuning the gate work function (Φ_M) and/or by engineering the source/drain lateral doping profiles to adjust the electrical channel length (L_{eff}) during the manufacturing process.

FINFET TECHNOLOGY

The quasi-planar FinFET (Fig. 1d) offers the superior scalability [5] and lower gate leakage current [6] of a double-gate (DG) MOSFET structure, together with a process flow and layout similar to that of the conventional MOSFET [7]. It is therefore being investigated by many companies [8-9] for sub-65nm CMOS technologies. FinFETs with gate lengths down to 10 nm have already been demonstrated with excellent control of short-channel effects and <0.5 ps intrinsic delay [10,11]. One advantage of this vertical transistor structure is that it is relatively immune to gate line-edge roughness, a major source of variability in planar nanoscale FETs [12]. In order to suppress short-channel effects, the thickness of the lightly doped or undoped FinFET channel/body (*i.e.* the fin width) must be ~1.5× smaller than L_G [13].

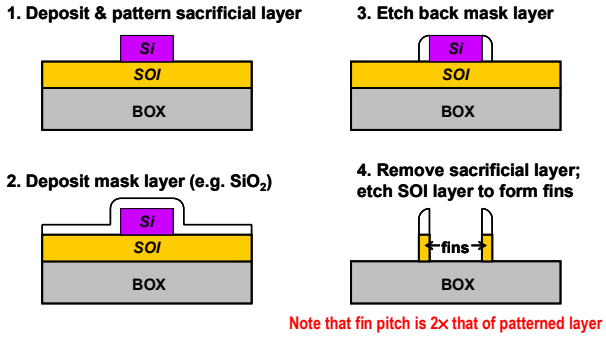

Figure 2: Sequence of schematic cross-sections illustrating the process for forming fins with sub-lithographic width using sidewall spacers.

Narrow Fin Formation

Sub-lithographic fins can be formed by using "spacers", formed along the sidewalls of a sacrificial patterned layer, as a hard mask (Fig. 2) [14]. The width of the spacers is determined by the thickness of the deposited spacer layer, and can be very uniform across a wafer so that FinFET performance variability due to variations in fin width are minimized. Another advantage of the spacer lithography process is that it provides for a doubling of fin density.

Optimization of Fin Orientation

The transconductance of a FinFET is dependent on its layout orientation, due to carrier mobility anisotropy in crystalline Si [8]. The channel surfaces of a FinFET lie in the (110) crystallographic plane when the fin is oriented parallel or perpendicular to the wafer flat or notch of a standard (100) wafer. Hole mobility is enhanced, while electron mobility is degraded, for a (110) Si surface as compared with a (100) Si surface [15]. To simultaneously achieve maximum NMOS and PMOS drive currents, a (100) sidewall surface for NMOS and (110) sidewall surface for PMOS is desirable. This can be achieved by orienting the PMOS fins to be either perpendicular or parallel to the flat or notch of a (100) wafer and orienting the NMOS fins to be rotated at a 45° angle (Fig. 3). Non-Manhattan layout geometry may pose a yield issue for sub-wavelength lithography, however. An alternative approach is to use a (110) wafer to allow the optimal CMOS FinFET sidewall surface orientations to be achieved with Manhattan layouts.

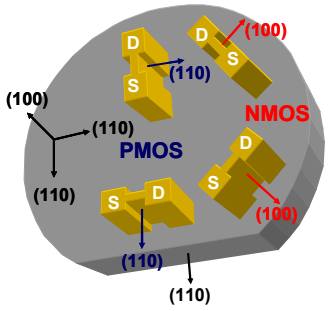

Figure 3: Si fin orientation for optimal CMOS FinFET performance. PMOS devices have (110) fin sidewall surfaces, while NMOS devices have (100) surfaces.

FinFET Design Considerations

As shown in Fig. 4a, the effective channel width W of a single-fin FinFET is twice the Si fin height H_{FIN}. Large W can be practically achieved by using multiple fins in parallel [16], but with variations limited to increments of $2\times H_{FIN}$. Taller fins allow for higher layout area efficiency, but with a trade-off in design flexibility. If the fin aspect ratio H_{FIN}/T_{Si} were to be limited to 3 (dictated by Si dry-etch process capability), $H_{FIN} = 3\times T_{Si} = 2\times L_{Gmin}$, where L_{Gmin} is the minimum gate length. Thus, W would be constrained to be an integer number of $4\times L_{Gmin}$.

Performance Enhancement Approaches

Techniques for increasing the average velocity of carriers in the channel – without significantly impacting cost and device reliability – will ultimately be necessary in order for the industry to maintain its historic 17%-per-year performance improvement rate [17]. Approaches to enhancing carrier mobility include the use of a strained capping layer [18], a strained gate electrode [19], or strained S/D regions (using epitaxial $Si_{1-x}Ge_x$ [20] or silicide [21]), in addition to optimization of the channel surface crystal orientation and current flow direction [22]. Some of these methods have already been shown to be effective for enhancing FinFET performance [23,24].

Parasitic source/drain series resistance and contact resistance will ultimately limit FinFET performance in the nanoscale regime [25]. The use of thick source/drain (S/D) regions, *e.g.* formed by selective growth of Si [8], $Si_{1-x}Ge_x$, or Ge [26], can help to alleviate this issue, particularly if low specific contact resistivity ($\rho_c<10^{-8}$ Ω-cm^2) contacts can be formed by silicidation/germanidation of the S/D fin surfaces.

Power consumption will be a primary design constraint for sub-65nm CMOS technologies, so that active leakage (V_T) control will be necessary for optimization of energy *vs.* delay trade-offs in future ULSI systems. The two gate electrodes of a FinFET can be electrically isolated, by using a masked etch to remove the gate material directly over the fin [27], to allow for independent operation. The resultant "multiple-independent gate" (MIG) FET can be operated as a back-gated UTB FET with the capability for dynamic V_T control (Fig. 4b).

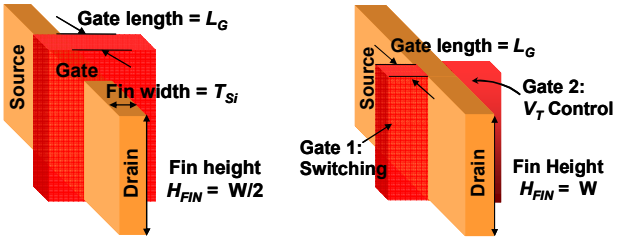

Figure 4: Close-up view of FinFETs with key parameters labeled. (a) double-gate (DG) (b) multiple independent gates (MIG) for dynamic V_T control.

FINFET-BASED SRAM DESIGN

Static memory (SRAM) arrays occupy a large fraction of the chip area in many of today's ICs, and will account for even more chip area in future designs. The issues of transistor leakage current (static power dissipation) and device-to-device variability (reduced static noise margin, *i.e.* cell stability) pose serious challenges for scaling conventional SRAM technology beyond the 65nm node. These issues can be alleviated by using FinFETs rather than classical bulk-Si MOSFETs in the memory cells.

Mixed-mode device-circuit simulation has been employed to compare the DC behavior of various SRAM cell designs for the 45nm node [28]. The circuit schematic and layout for a FinFET-based six-transistor (6-T) SRAM cell is shown in Fig. 5. (The fin sidewalls are assumed to lie along (100) crystallographic planes.) As compared against a bulk-Si 6-T cell, the FinFET-based 6-T cell achieves larger read margin (175mV *vs.* 135mV, for 1V supply voltage) due to better suppression of leakage current. The read margin can be improved further by increasing the "beta ratio" of the cell, *i.e.* the relative strength of the pull-down n-channel transistors to the access transistors. This can be done by using 2-fin pull-down devices or by rotating the access transistors so that the fin sidewalls lie along (110) crystallographic planes, with slight area penalty (~15% increase in cell layout area).

Back-Gated FETs for Dynamic Feedback

By connecting the storage node to the back-gate of the access transistor, as shown in Fig. 6, the strength of the access transistor can be selectively decreased. For example, if the stored bit is a "0", the back-gate of the corresponding access transistor is biased at 0V, decreasing its strength. This effectively increases the beta ratio during the read cycle and thus improves the read margin (by 71%). Although the BG access transistor is weaker than a DG access transistor, the "0" storage node in the 6-T design with feedback stays closer to V_{SS} than in the 6-T design without feedback, so that the BG access transistor has more gate overdrive. The net result is that the cell read performance is negligibly impacted.

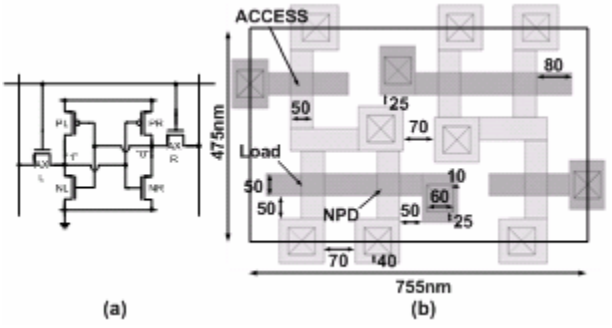

Figure 5: (a) Circuit schematic and (b) layout for a 6-T SRAM cell utilizing double-gate FinFETs [28]. (The large rectangle delineates one memory cell.) Dimensions are indicated in nanometers.

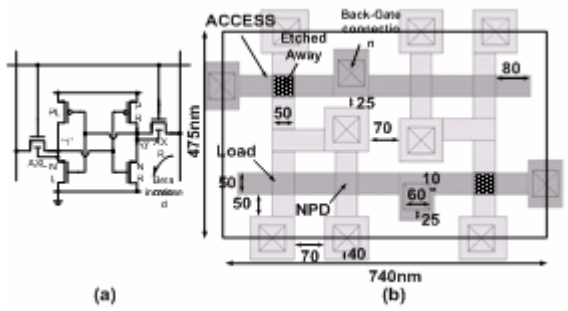

Figure 6: (a) Circuit schematic and (b) layout for a 6-T SRAM cell using back-gated access transistors to provide dynamic feedback [28]. The dark squares indicate where the gate material is removed.

Note that the back-gate connections in Fig. 6 do not have any associated layout area penalty. Neither is there a significant penalty in static power consumption [28].

The impact of L_G and T_{Si} variations ($3\sigma_{LG} = 3\sigma_{TSi} = 10\%$ of L_G) for FinFETs and the impact of statistical dopant fluctuations and L_G variations in bulk-Si MOSFETs ($3\sigma_{LG} = 10\%$ of L_G) were studied via Monte Carlo and mixed-mode simulations. The simulated cell read margin distributions are shown in Fig. 7. These results show that FinFET-based cell designs provide larger SNM with tighter distribution, which facilitates supply-voltage scaling to reduce dynamic power consumption.

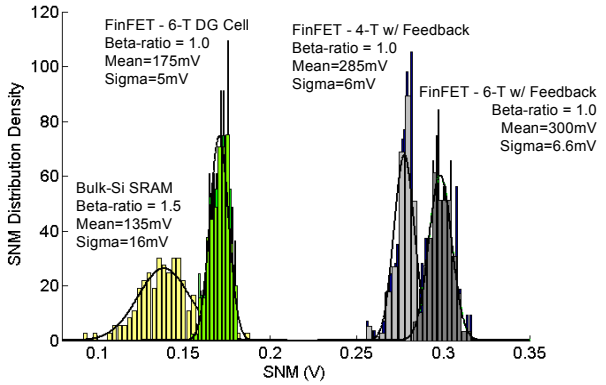

Figure 7: Simulated impact of process-induced variations on SRAM static noise margin (SNM) [28].

SUMMARY

The FinFET is a manufacturable double-gate MOSFET structure which achieves superior control of short-channel effects and higher drive current as compared to the classical bulk-Si MOSFET structure. By eliminating the need for heavy channel/body doping, it also can provide immunity to random variations associated with the discreteness of dopant atoms. Therefore, it is a promising solution to surmount the challenges of increasing leakage current and device-to-device variability for future high-density, low-power digital ICs. Judicious application of BG FinFETs can yield dramatic improvements in SRAM cell stability, to facilitate memory technology scaling.

ACKNOWLEDGEMENTS

This work resulted from fruitful collaborations with my present and former colleagues at UC-Berkeley, in particular Borivoje Nikolić, Hideki Takeuchi, Yang-Kyu Choi, Leland Chang, Sriram Balasubramanian, Kyoungsub Shin, and Hei Kam. Research funding from the SRC (Advanced Devices and Technology Program), the MARCO Focus Center for Advanced Materials, Structures and Devices, and the MARCO Focus Center for Circuits, Systems, and Software is gratefully acknowledged.

REFERENCES

[1] D. J. Frank *et al*., *Proc. IEEE*, Vol. 89, p.259 (2001).

[2] A. Hokazono *et al*., *IEDM Tech. Dig.,* p.639 (2002).

[3] D. Antoniadis, *Symp. VLSI Tech. Dig.*, p.2 (2002).

[4] L. Chang *et al*., *Proc. IEEE.*, Vol. 91, No. 11, p.1860 (2003).

[5] C. Wann *et al*., *IEEE Trans. Electron Devices*, Vol. 43, p.1742 (1996).

[6] L. Chang *et al*., *IEDM Tech. Dig.*, p.99 (2001).

[7] N. Lindert *et al*., *IEEE Electron Device Letters*, Vol. 22, p.487 (2001).

[8] J. Kedzierski *et al*., *IEDM Tech. Dig.*, p.437 (2001).

[9] F.-L. Yang *et al*., *Symp. VLSI Tech. Dig.*, p.104 (2002).

[10] B. Yu *et al*., *IEDM Tech. Dig.*, p.251 (2002).

[11] F.-L. Yang *et al*., *Symp. VLSI Tech. Dig.*, p.196 (2004).

[12] A. R. Brown *et al*., *IEEE Trans. Nanotechnology*, Vol. 1, p.195 (2002).

[13] N. Lindert *et al*., *59th Device Research Conf. Digest*, p. 26 (2001).

[14] Y.-K. Choi *et al*., *IEEE Trans. Electron Devices*, p.436 (2002).

[15] T. Sato *et al*., *Physical Review B*, Vol. 4, p.1950 (1971).

[16] D. Hisamoto *et al*., *IEDM Tech. Dig.*, p. 1032 (1998).

[17] Semiconductor Industry Association *Int'l Technology Roadmap for Semiconductors*, 2003 Edition, http://public.itrs.net/Files/2003ITRS/Home2003.htm

[18] F. Ootsuka *et al*., *IEDM Tech. Dig*., p.575 (2000).

[19] K. Ota *et al*., *IEDM Tech. Dig*., p.27 (2002).

[20] T. Ghani *et al*., *IEDM Tech. Dig*., p.978 (2003).

[21] C.-H. Ge *et al*., *IEDM Tech. Dig*., p.73 (2003).

[22] M. Yang *et al*., *IEDM Tech. Dig*., p.453 (2003).

[23] P. Verheyen *et al*., *Symp. VLSI Tech. Dig*. (2005).

[24] K. Shin *et al*., *Device Research Conf. Dig*., p.201 (2005).

[25] H. Kam *et al*., *Proc. Silicon Nanoelectronics Workshop*, p.9 (2004).

[26] N. Lindert *et al*., *IEEE Int'l SOI Conference*, p.111 (2001).

[27] L. Mathew *et al*., *IEEE Int'l Conf. on Integrated Circuit Design and Technology*, p.97 (2004).

[28] Z. Guo *et al*., *Proc. ISLPED* (2005).