Analysis of Deeply Scaled Multi-Gate Devices with Design Centering across Multiple Voltage Regimes

Shuang Chen, Xue Lin, Alireza Shafaei, Yanzhi Wang, Massoud Pedram

Department of Electrical Engineering, University of Southern California, Los Angeles, U.S.A.

***Abstract*—This work aims at finding a design-centered FinFET model with small geometric for circuit and system level simulations and performance prediction of next-generation systems on chip. A number of devices including the ITRS 7nm multi-gate device are used as examples. While adjusting design parameters for the transistors, a design centering step is included in which the gate workfunction is carefully adjusted to account for the increased power dissipation due to gate length variations. Using a cross-layer framework, compact device models and standard cell libraries are built up for circuit-level and system-level simulations. Simulation results of SRAM cells as well as some combinational/sequential benchmark circuits are shown to compare the device performance in different technologies.**

Keywords—FinFET, Near-threshold computing, Design centering, SRAM.

I. Introduction

In order to keep up with the pace of Moore's law [1], efforts are required on a number of topics, one of which is finding a proper structure for deeply-scaled semiconductor devices. Compared to conventional planar CMOS devices, FinFET devices provide stronger control over the channel through gate coupling, thus offering a solution to some major issues in nanoscale transistors, e.g. high leakage power, short channel effects, etc. [2] This being the case, it is necessary to properly determine design parameters of future generations of FinFET devices, which can not only help determining whether FinFETs will be the ultimate choice towards the end of ITRS roadmap, but also identify and overcome potential issues and challenges facing circuit, architecture, and system designs. In this context, one crucial aspect that should be accounted for is the process variations in design parameters such as the physical gate length, gate work-function, etc. This calls for a design centering step to be done before final decisions can be made on the device design parameters. In this work, a 7nm-gate-length FinFET model (referred to as 7nm-L_g in this paper) and the ITRS 7nm multi-gate FinFET device model [3] (referred to as 7nm-ITRS in this paper) are used as examples to show the design centering process and the performance of the design centered version of these devices when used in SRAM cells as well as combinational and sequential circuits. Sentaurus Device, HSPICE, and Synopsys Design Compiler are used for device simulations, circuit simulations, and post-synthesis simulations, respectively.

II. Device Modeling

In this work, we use the 2D structure suggested by ITRS [3] as shown in Fig. 1. The nominal values of major design parameters are shown in Table I. Please note that for 7nm-L_g FinFET transistors, the effective channel length is larger than the physical gate length, which means that there is an underlap instead of an overlap between the gate and the source/drain diffusion. For the purpose of design centering, we assume a Gaussian distribution for the physical gate length with a standard deviation of 0.8nm as suggested in [4]. Since the leakage current increases super-linearly when the gate length decreases, the average leakage current is larger than the leakage current specified with the nominal gate length. To limit the average leakage current below a desired level (e.g. 100nA/μm as in ITRS reports), the gate workfunction is adjusted. Then, lookup-table-based Verilog-A models are extracted and standard cell libraries are generated using methods proposed in [5]. Based on considerations of delay, power consumption, and reliability, we generate three libraries for 7nm-L_g FinFETs under supply voltages of 0.3V, 0.4V and 0.5V, respectively, and four libraries for 7nm-ITRS FinFETs under supply voltages of 0.4V, 0.5V, 0.6V, and 0.7V, respectively.

III. Results and Discussion

As shown in Fig. 2, the leakage current is generally a convex decreasing function of the physical gate length. After design centering, the nominal leakage current for both 7nm-L_g FinFETs and 7nm-ITRS FinFETs are set to 70nA/μm (measured with $V_{dd} = 0.45$V and $V_{dd} = 0.8$V, respectively). The I_d-V_g curves with $V_{ds} = 0.5$V and the I_d-V_d curve with $V_{gs} = 0.5$V are shown in Fig. 3. Since we adjust the leakage current only through the gate workfunction, the sub-threshold slope (SS) does not degrade. Using the second-order derivative logarithmic (SDL) method, the threshold voltage is 0.25V for nfets and 0.22V for pfets of 7nm-L_g devices, and 0.29V for nfets and 0.28V for pfets of 7nm-ITRS multi-gate devices. Therefore, the characterized libraries for both technologies cover both the near-threshold operation regime and the super-threshold operation regime. The performance of design centered FinFET-based circuits on a set of benchmark circuits are shown in Table II. As can be seen, compared to the 45nm technology, the 7nm-ITRS technology achieves up to 16.7x improvement in timing (delay or minimum clock cycle), 19.0x reduction in leakage power, and 70.6x reduction in dynamic power consumption. Significant improvement can also be found for the 7nm-L_g devices over 45nm devices. The performance of SRAM cells are shown in Table III and Fig. 4. One can see that SRAM cells can function correctly with a supply voltage as low as 0.4V for the 7nm-ITRS technology and 0.3V for the 7nm-L_g technology. And the read performance can be improved by using 8T-SRAM cells in both technologies.

IV. Conclusion

In this paper, we present a design centering approach for nanoscale FinFET devices that can be incorporated into a cross-layer simulation framework. Compact models and standard cell libraries are generated for 7nm-ITRS and 7nm-L_g devices. And the performance of these devices are analyzed based on simulation results.

References

[1] G. E. Moore, *Electronics Magazine*, 1965. [2] L. Chang, *et al*, *Proc. of the IEEE*, 2003. [3] [Online] http://www.itrs.net/reports.html. [4] T. Matsukawa *et al*, *VLSIT*, 2009. [5] S. Chen, *et al*, *S3S Conference*, 2014.

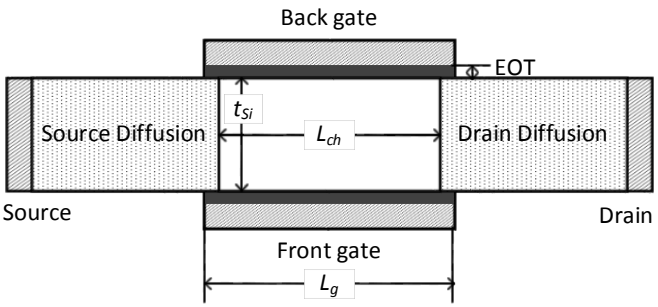

Fig. 1. FinFET structure model used in Sentaurus Device simulations

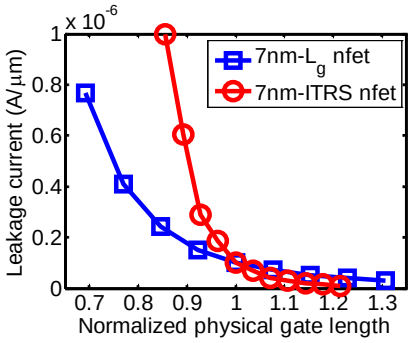

Fig. 2. Leakage current with different gate lengths. The leakage current is a convex decreasing function of the physical gate length, thus making the average leakage current subject to process variations larger than the leakage current measured at nominal gate length.

TABLE I. LIST OF DESIGN PARAMETERS

Parameter Name	7nm-L_g	7nm-ITRS
Physical gate length (L_g)	7.06nm	13.9nm
Effective channel length[a] (L_{ch})	10.32nm	11.0nm
Channel width (t_{Si})	3.80nm	4.4nm
Equivalent oxide thickness[b] (EOT)	0.33nm	0.69nm
Fin Height[c]	15.20nm	17.6nm

[a] Effective channel length larger than physical gate length means that an underlap exists between the gate and the source/drain diffusion area.

[b] The physical gate oxide thicknesses are 1.30nm and 2.49nm for 7nm-L_g and 7nm-ITRS FinFET devices, respectively.

[c] The fin height is set to four times the channel width.

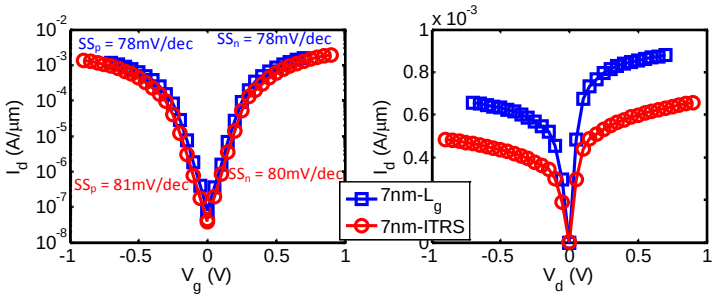

Fig. 3. (a) I_d-V_g curves with V_{ds} = 0.5V. The subthreshold slopes (SS) are also listed. (b) I_d-V_d curve with V_{gs} = 0.5V.

TABLE II. PERFORMANCE COMPARISON BETWEEN DIFFERENT TECHNOLOGIES IN LOGIC CIRCUITS

Library	Delay (ns)		Clock cycle (ns)		Energy per operation (J)		Leakage power (W)		Dynamic power (W)	
	C3540[a]	Multiplier	Arbiter	FIFO	C3540	Multiplier	Arbiter	FIFO	Arbiter	FIFO
7nm-ITRS 0.4V	0.215	0.280	0.12	0.12	2.98e-15	2.75e-14	5.13e-07	4.30e-06	1.39e-05	9.77e-05
7nm-ITRS 0.5V	0.135	0.170	0.08	0.10	5.01e-15	4.79e-14	7.71e-07	6.13e-06	2.77e-05	2.00e-04
7nm-ITRS 0.6V	0.095	0.126	0.06	0.08	8.65e-15	6.98e-14	1.12e-06	8.93e-06	7.61e-05	3.82e-04
7nm-ITRS 0.7V	0.080	0.105	0.04	0.06	1.12e-14	9.81e-14	1.79e-06	1.20e-05	1.55e-04	4.71e-04
7nm-L_g0.3V	0.140	0.180	0.08	0.09	8.83e-16	7.13e-15	4.58e-07	4.25e-06	8.50e-06	7.03e-05
7nm-L_g0.4V	0.087	0.110	0.06	0.08	1.92e-15	1.79e-14	8.40e-07	6.67e-06	1.27e-05	3.01e-04
7nm-L_g0.5V	0.060	0.085	0.05	0.04	3.81e-15	3.18e-14	1.32e-06	1.08e-05	4.51e-05	2.01e-04
Nangate 45nm 1.1V	1.050	2.900	0.60	1.00	4.23e-13	3.41e-12	9.75e-06	7.03e-05	9.82e-04	3.12e-03

[a] "C3540" is a combinational benchmark circuit from the ISCAS benchmark set, "Multiplier" is a 16-bit multiplier, "Arbiter" is a 4-to-1 4-bit arbiter, and "FIFO" is a FIFO with a width of 8 bits and a depth of 64.

Four libraries are generated for 7nm-ITRS devices with the supply voltage varies from 0.4V to 0.7V, and three libraries are generated for 7nm-L_g devices with the supply voltage varying from 0.3V to 0.5V. The Nangate 45nm library is provided as a baseline. Since the threshold voltage is around 0.24V for7nm-L_g devices and 0.28V for 7nm-ITRS devices. The generated libraries covers both the near-threshold and the super-threshold operation regime.

TABLE III. PERFORMANCE OF FINFET-BASED 6T-SRAM CELLS UNDER DIFFERENT SUPPLY VOLTAGES[a]

Library	Read latency (ps)	Write latency (ps)	Leakage power (nW)	Read energy (aJ)	Write energy (aJ)	SNM[b]
7nm-ITRS 0.4V	8.73	1.80	1.09	17.39	6.25	20%
7nm-ITRS 0.5V	5.01	1.46	1.79	22.00	11.21	18%
7nm-ITRS 0.6V	3.49	1.29	2.67	26.62	17.87	17%
7nm-ITRS 0.7V	2.69	1.24	3.76	31.18	28.62	15%
7nm-L_g0.3V	42.02	7.38	0.40	13.11	4.16	18%
7nm-L_g0.4V	15.84	4.28	0.71	17.76	9.19	15%
7nm-L_g0.5V	7.30	2.62	1.12	23.10	15.11	17%

a. Minimum sized FinFETs (i.e. FinFETs with one fin) are used as access transistors, pull-up transistors, and pull-down transistors to implement the 6T SRAM cells.

b. The static noise margin (SNM) is normalized by the supply voltage and specified as the minimum SNM in hold, read, and write operations.

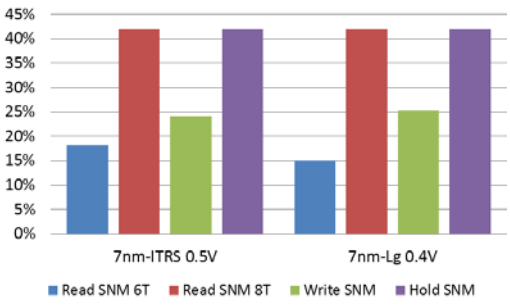

Fig. 4. Comparison of SNM in read, write, and hold operations. The SNM bottleneck appears in read operations for 6T-SRAM cells. By using 8T-SRAM cells, the SNM in read operations can increase from 15%~20% to around 40%.