

ANALYSIS OF ULTRA-LOW-POWER CMOS WITH PROCESS AND DEVICE SIMULATION

G. Schrom†, D. Liu‡, Ch. Pichler†, Ch. Svensson‡, and S. Selberherr†

†Institute for Microelectronics, Technical University Vienna,
Gusshausstrasse 27–29, A–1040 Wien, Austria

‡Department of Physics and Measurement Technology, Linköping University,
S–58183 Linköping, Sweden

Abstract

The feasibility and the limitations of ultra-low-power CMOS technologies are investigated using process and device simulation, followed by post-processing of the simulated IV data. On the basis of simplified modern state-of-the-art processes and special scaling a set of possible ultra-low-power CMOS processes was developed and analyzed for their performance on the gate level.

1. Introduction

By drastically decreasing the supply voltage and the threshold voltages, a great reduction of the power consumption can be achieved at the expense of an increase in gate delay. This can be compensated to a certain extent by employing parallelism in the systems design so that, for a given overall performance, the total power consumption is drastically reduced compared to conventional CMOS techniques [1]. We demonstrate the feasibility of ultra-low-power CMOS and determine a lower limit for the supply voltage depending on the type of digital circuit technique.

2. Process Technology

The processes under consideration are recessed-well dual-gate processes with a very thin gate oxide (5nm and below) to obtain controllably low threshold voltages. The source/drain dopings are formed by single shallow implants and a conventional furnace anneal. The G/S and G/D overlap capacitances can be controlled with a spacer formed prior to the S/D implants. As a consequence of the low voltages, ultra-low-power processes differ from conventional CMOS processes in several points: Because of the low V_{DD} , the hot-carrier problem virtually does not exist and therefore an LDD process is not necessary. Also, no GIDL can occur. As for very low V_{DD} the devices must operate in the weak-inversion regime, the difference of the carrier mobilities μ_n, μ_p can be roughly compensated by adjusting the threshold voltages to achieve symmetric inverter transfer characteristics. This compensation does not work, however, in the transient case because the speed is mainly determined by the strong-inversion part of the input characteristics. The sub-threshold behavior is crucial because it determines the achievable ratio of I_{on}/I_{off} which is limited by $e^{V_{DDq}/kT}$ and decreases as $V_{Tn,p}$ are made smaller. Therefore, 'zero- V_T ' transistors are not desirable. On the other hand,

if $V_{Tn,p}$ are too high the speed becomes unacceptably low. A major challenge is to achieve controllably low threshold voltages. Although the adjustment of $V_{Tn,p}$ with a bulk bias seems very attractive, this method is unlikely to be accepted for digital circuit design because of the significant overhead.

Another problem can arise from the very thin gate insulator. If one uses very thin thermally grown gate oxides (below 5nm), boron diffusion can considerably degrade the device behavior. Also, boron segregation causes deleterious effect, especially, in the sub-threshold regime. On the other hand, there are several options for the gate insulator. Nitrides or oxinitrides may be good alternatives to conventional (pure SiO_2) gate oxides. Silicon nitride can be used as an effective diffusion barrier and ultra thin Si_3N_4 layers with low defect densities are easier to fabricate [2]. The lower tunneling barrier compared to that of SiO_2 is still acceptable for ultra-low-power CMOS because of the low voltages and, also, because a controllably small gate current is allowed.

3. Process and Device Simulation

Both process and device simulation were done using VISTA with the SFC (Simulation Flow Controller) to allow for quick process design and evaluation [3]. For the electrical characterization of the devices MINIMOS 6.0 was used to calculate a matrix of drain currents $I_D(V_G, V_D)$ over a range of V_G and V_D for the PMOS and NMOS transistor. Based on these data, a fast and accurate table-driven DC analysis of simple gates and inverters is possible. The bulk effect could also be included but for the given devices and voltages it was found not to be significant. The dynamic behavior was estimated from capacitance data obtained by AC analysis with MINIMOS. Fig. 1 and Fig. 2 show the doping profiles of static-logic CMOS devices.

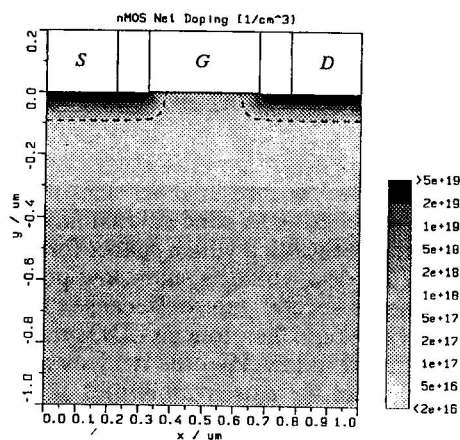


Figure 1: NMOS doping profile, process A ($V_{DD} = 200\text{mV}$)

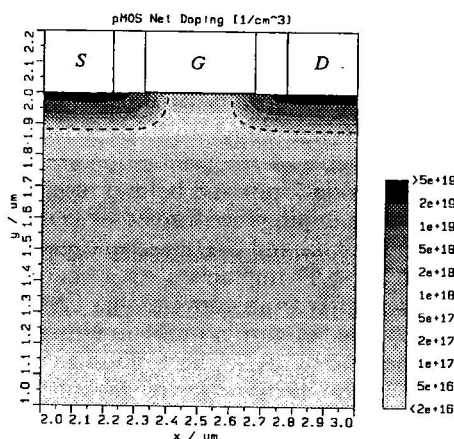


Figure 2: PMOS doping profile, process A ($V_{DD} = 200\text{mV}$)

4. Results and Discussion

The simulated processes were a $0.35\mu\text{m}$ process (A) for static logic and a $0.5\mu\text{m}$ process (B) for dynamic logic. The processes were designed for proper DC characteristics but were not optimized for speed. The device characteristics for process A are shown in Fig. 3 and 4. Fig. 5 and 6 show the inverter transfer curves. Fig. 7 and Fig. 8 show the noise margins and the inverter delay as a function of the supply voltage. From

Fig. 7 it can be seen that a ring oscillator built with process A would work even at $V_{DD} = 80\text{mV}$ and by using additional inverters at the gate inputs and outputs one could also design digital circuits for $V_{DD} < 100\text{mV}$ but the overhead would be considerable. For Process B, the ratios of I_{on}/I_{off} in Table 1 are in the order of 10^4 and the ratio of τ_l/τ_d is about 2300 which is rather low for dynamic logic. For process A, it can be seen from Table 2 that for 3-input NANDs with minimal transistors the high-noise-margin NM_H is already very low. From these data we conclude that the limits for the supply voltage will be at 200mV for static logic and 500mV for dynamic logic with a fan-in of 3 at $T = 300\text{K}$. The ultimate limit for the CMOS supply voltage is given by the thermal voltage as $V_{DD} > X \cdot kT/q$ where X is a factor depending on the type of digital circuit technique and on the process technology. We found that $X_{stat} < 8$ and $X_{dyn} < 20$ is sufficient for a fan-in of 3.

References

- [1] D. Liu and Ch. Svensson. Trading Speed for Low Power by Choice of Supply and Threshold Voltages. *IEEE J.Solid-State Circuits*, 28(1):10–17, 1993.
- [2] T. Morimoto, H.S. Momose, S. Takagi, K. Yamabe, and H. Iwai. Ultrathin Nitride Gate MISFET Operating with Tunneling Gate Current. In *Proc.22ND Int.Conf. on Solid-State Devices and Materials*, pages 361–364, Sendai, Japan, 1990.
- [3] Ch. Pichler and S. Selberherr. Process Flow Representation within the VISTA Framework. In S. Selberherr, H. Stippel, and E. Strasser, editors, *Simulation of Semiconductor Devices and Processes*, volume 5, pages 25–28. Springer, 1993.

Table 1: Simulated device characteristics. The threshold voltage was defined as $|I_D(V_T)| = 1\mu\text{A}/\mu\text{m}$. All voltages are in V, all currents are in $\text{A}/\mu\text{m}$

process	V_{DD}	$V_{T,n}$	$V_{T,p}$	$I_{off,n}$	$I_{off,p}$	$I_{on,n}$	$I_{on,p}$
A	0.2	0.067	-0.059	$0.14 \cdot 10^{-6}$	$0.27 \cdot 10^{-6}$	$16.7 \cdot 10^{-6}$	$9.4 \cdot 10^{-6}$
B	0.5	0.26	-0.24	$0.7 \cdot 10^{-9}$	$2.8 \cdot 10^{-9}$	$25.6 \cdot 10^{-6}$	$16.7 \cdot 10^{-6}$

Table 2: Noise margins (in $\%V_{DD}$) for a simple inverter and a 3-input NAND gate, and inverter delay, leakage time, switching energy, and static power consumption

process	$NM_{H,inv}$	$NM_{L,inv}$	$NM_{H,gate}$	$NM_{L,gate}$	t_d	t_l	E_s	P_{stat}
A	28	23	13	39	0.29ns	7.2ns	0.65fJ	41nW
B	38	44	31	49	0.55ns	1.3 μs	4.3fJ	0.88nW

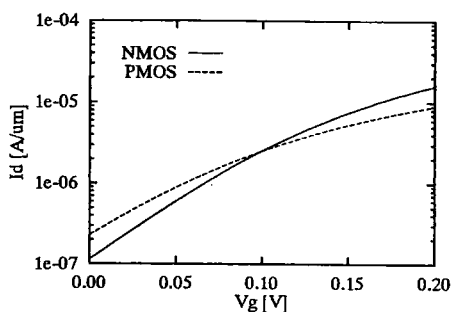


Figure 3: Input characteristics, process A ($V_{DD} = 200\text{mV}$)

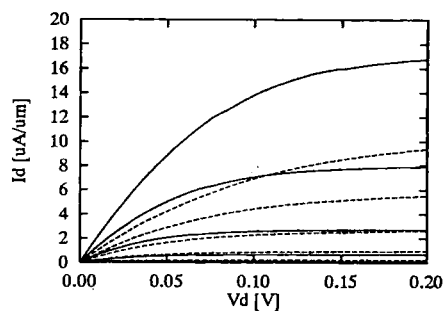


Figure 4: Output characteristics, process A ($V_{DD} = 200\text{mV}$)

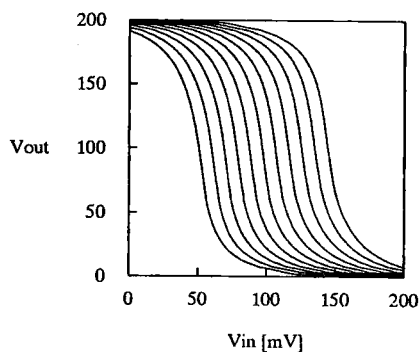


Figure 5: Inverter transfer characteristics for $W_n/W_p = 0.1...10$, process A at $V_{DD} = 200\text{mV}$

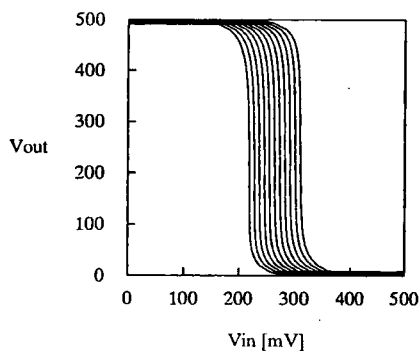


Figure 6: Inverter transfer characteristics for $W_n/W_p = 0.1...10$, process B at $V_{DD} = 500\text{mV}$

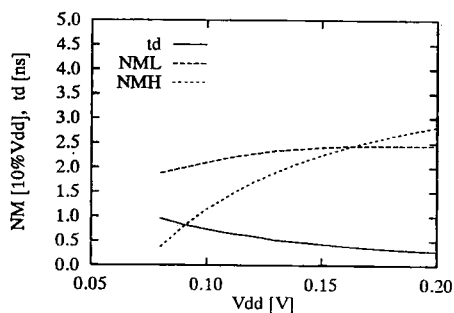


Figure 7: Noise margins and delaytime vs. V_{DD} , process A

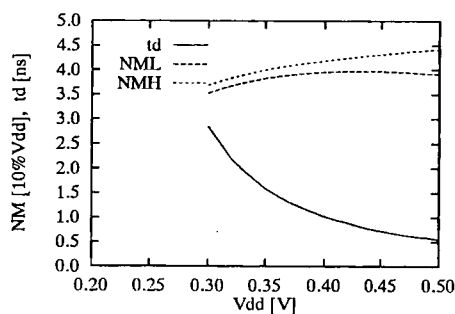


Figure 8: Noise margins and delaytime vs. V_{DD} , process B