Standby Voltage Scaling for Reduced Power

B. Calhoun, A. Chandrakasan
Department of Electrical Engineering
Massachusetts Institute of Technology, Cambridge, MA

Abstract

Lowering V_{DD} during standby mode reduces power by decreasing both voltage and current. Measurements of a 0.13µm testchip show that reducing V_{DD} to near the point where state is lost gives the best power savings. We propose closed-loop voltage scaling that uses "canary" flip-flops for achieving these savings. This approach provides over 2X higher savings than optimal open-loop approaches without loss of state.

Introduction

The well-known increase in leakage currents with each new process generation has made leakage power a significant part of overall power consumption in both active and standby mode[1]. The components of leakage include subthreshold leakage, gate leakage, GIDL, and forward biased diode leakage[2]. At the 0.13µm technology node, subthreshold leakage dominates the other components. For energy-constrained systems or circuits with lifetime requirements, leakage reduction is especially important during standby mode. The BSIM2 model [3] incorporates a linearized body coefficient, γ, and linearized DIBL parameter, η, into its subthreshold current equation:

$$I_{subth} = I_o \cdot e^{\frac{V_{GS} - V_{T0} - \gamma V_S + \eta V_{DS}}{n \cdot V_{th}}} \cdot \left(1 - e^{-\frac{V_{DS}}{V_{th}}}\right) \quad (1)$$

Equation (1) indicates that lowering V_{DD} will produce a corresponding exponential reduction in subthreshold current resulting from DIBL. At extremely low V_{DS} values (~kT/q), the parenthetical term produces a more pronounced tail-off in leakage current. Lowering V_{DD} thus saves standby power by decreasing both I_{subth} and V_{DD}. Since gate leakage also decreases as V_{DD} falls, this technique still saves power when gate leakage becomes comparable to subthreshold leakage.

Previous work has suggested using voltage scaling for standby power reduction[4]. One approach uses diode stacks together with power gating MOSFETs to pinch in the rail voltages during standby[5]. The quantity and sizes of the devices used in the diode stack determine the reduced V_{DD} value during sleep mode. In order to preserve state, the authors of [5] do not reduce standby V_{DD} by more than about 40%. We examine what savings are possible from lowering the supply voltage and show that greater reduction of V_{DD} is preferable. At the same time, careful treatment of a circuit's storage elements will allow state preservation.

Potential Power Savings from Lowering V_{DD}

Equation (1) shows that lowering V_{DD} will provide power savings, but it is helpful to quantify those savings. Presumably, the standby power supply for a circuit can decrease to zero, but the circuit will lose all of its state. The optimal point for power savings using this technique is the lowest voltage for which the circuit retains state. In theory, the combinational logic in a circuit does not need to hold state. If two power supplies are readily available, the voltage supply to combinational logic can fall all the way to zero while sequential elements use a different power supply. The sequential supply may be decreased to save power, but it must remain above some minimum point to hold its state. Separating the power supplies increases power savings dramatically because all of the combinational logic draws zero power in standby while the sequential logic uses reduced power resulting from a lower V_{DD}. Separating the power supplies to the combinational and sequential logic is often impractical, so we will assume that the entire circuit under test shares the same power supply.

Fig. 1 shows simulated current reduction and power savings versus V_{DD} for a Leakage Feedback Flip-Flop (LFBFF)[6]. The LFBFF is a Multi-Threshold CMOS (MTCMOS) stor-

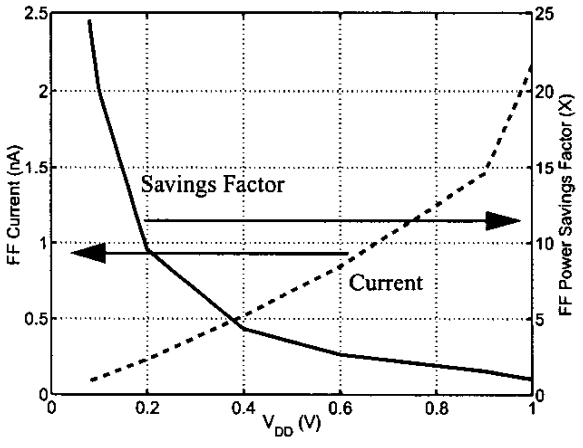

Fig. 1. Simulated Leakage Feedback Flip-Flop (LFBFF) current and power savings factor versus V_{DD}.

0-7803-7842-3/03/$17.00 © 2003 IEEE IEEE 2003 CUSTOM INTEGRATED CIRCUITS CONFERENCE

age device that retains its state even during sleep mode by selectively maintaining an active path power or ground. The LFBFF simulated in the figure is sized for critical path operation. The decreasing current and voltage create power savings relative to the nominal case at V_{DD}=1.0V. The curves stop at 80mV, which we later show to be the lower limit for a critical path LFBFF for holding both a '1' and a '0'. The plot suggests that a reduction in V_{DD} to a voltage above 400mV provides power savings less than 5X, while driving V_{DD} close to its minimum boosts the savings dramatically. Changing V_{DD} from 200mV to 100mV causes the savings to jump by over 2X. Exploiting the dramatic power savings that occur at low V_{DD} is viable only if the circuits continue to save state.

We fabricated a 0.13μm testchip that uses LFBFFs such as the one whose measurements appear in Fig. 1. The testchip implements a low power FPGA architecture for DSP applications. It consists of 12 Configurable Logic Blocks (CLBs) separated into 3 slices. Each CLB holds a 4-bit adder, a 4-bit register (LFBFF), and a 16-bit LookUp Table (LUT). The composition of the CLBs on the testchip makes it a fair representation of many generic circuits. The testchip uses MTCMOS logic in which high V_T devices act as power gating switches during standby mode. This logic style allows fast, low V_T devices on the critical path, but it enjoys the lower standby leakage of high V_T devices. While the circuits we tested are MTCMOS, the voltage scaling approach will work for single V_T designs as well.

Fig. 2 shows the measured dependence of current on V_{DD} in the testchip. It also shows the measured power savings for the testchip. These measurements show the current drawn by the testchip while the sleep signal is asserted and the high V_T devices limit the leakage current. Both curves follow the trend previously shown in the LFBFF simulation. The power savings increase rapidly for V_{DD} below about 200mV.

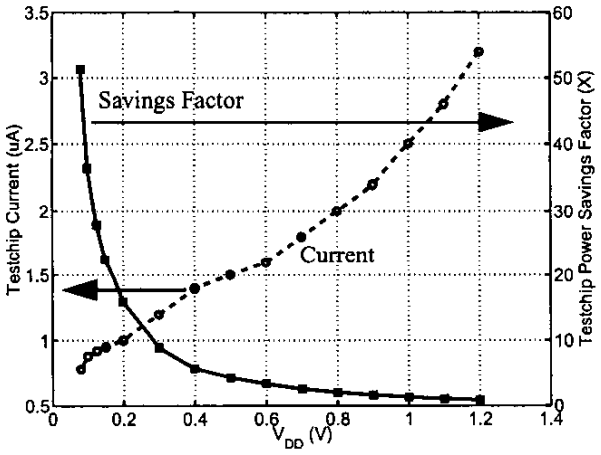

Fig. 2. Measured Testchip current and power savings versus V_{DD}.

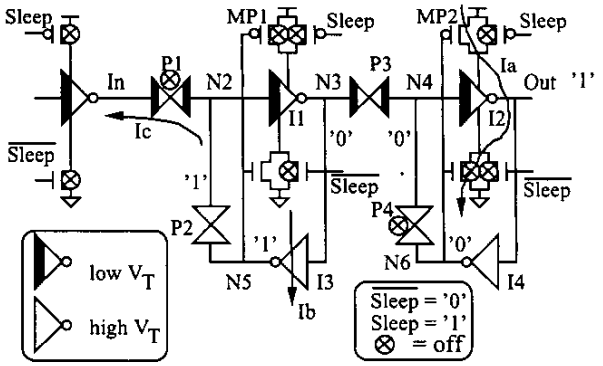

Fig. 3. Schematic of LFBFF in standby (sleep) mode holding a '1'.

Saving State at Low V_{DD}

A solid understanding of the mechanism of failure for the sequential elements in a circuit shows how sizing affects the minimum supply voltage for the circuit. The sequential elements will lose their state below that voltage value. Even though the analysis in this paper focuses on the LFBFF, the approach applies to any sequential circuit, including circuits with only one V_T. We will use the sizing analysis to implement a solution that can achieve the higher power savings by driving the supply voltage close to its minimum value.

A given implementation of LFBFFs has a lower bound on the supply voltage below which the device cannot hold state. Fig. 3 shows the LFBFF holding a logical '1' in sleep mode. The FF maintains the correct state as long as the nodes in the circuit all maintain the correct logical values. The logical '1' nodes are the initial point of failure for the FF as the power supply lowers. The two key nodes for storing a '1' are the N2/N5 node and the output node. For simplicity of analysis, we ignore the contribution of the sleep devices that are turned off by the sleep signal and consider only the conditionally off devices. This initial analysis assumes that MP1, MP2, and inverters I3 and I4 are all minimum sized.

First consider the output node holding a '1'. The output will retain its value as long as MP2 can supply all of the current, I_a, drawn by the two NMOS sleep devices with V_{SD}(MP2) near 0. V_{SG} of MP2 decreases with the power supply voltage so MP2 eventually transitions from the linear region to the subthreshold region. As MP2 loses current drive, its V_{SD} must increase to provide the current drawn by the NMOS FETs. Once this happens, the voltage at the output node begins to fall at a faster rate than the decreasing supply voltage, and the flip-flop fails to hold its state. This suggests that increasing the size of MP2 will extend the ability of the output node to hold a '1' at lower supply voltages. We must consider, however, that increasing the size of MP2 will at the same time reduce the FF's ability to hold a '0' at that node.

Fig. 4 shows a simulation of the failure of a minimum sized LFBFF to hold a '1'. This failure occurs with a supply voltage of 100mV. The figure shows that node N6 begins to increase its voltage value well before the FF actually fails. As the output node voltage (input to I4) begins to drop below the V_{DD} value, the PMOS device in inverter I4 begins to increase its subthreshold current as the NMOS in I4 loses some of its current drive. Node N6 must rise above zero to balance the current in the inverter. The increase of N6 aggravates the situation with I_a by reducing the gate drive for MP2 and increasing the gate drive of the NMOS sleep device. This feedback adds to the inability of MP2 to sustain I_a without the output voltage dropping even more. A similar situation causes the logical '1' at node N5 to transition to '0'. Fig. 3 shows two currents, I_b and I_c, that act to pull node N5/N2 toward ground. V_{SD} for the PMOS in inverter I3 must increase for that device to maintain I_b as the voltage supply drops. Similar to the process described above, N5 falls below the V_{DD} value.

The combination of these effects results in the failure of the minimum sized flip-flop to maintain a logical '1' below about 100mV. However, the mechanisms of failure suggest that the performance of the flip-flop can improve. For example, increasing the width of MP2 above minimum size allows that FET to sustain the large leakage current I_a without requiring a larger V_{SD}. Similarly, raising the size of the PMOS device in inverter I3 would hold N5 at a logical '1' for lower supplies.

Increasing the sizes of MP2 and the I3 PMOS in the critical path LFBFFs allows those devices to hold a logical '1' at a supply voltage of 80mV. Fig. 5 shows a simulation of the correct operation. Note that the output node and all the internal voltages settle to a steady state at the lower supply voltage and then return to the proper levels when the supply voltage rises again. This plot shows that careful sizing of the devices mentioned above can improve the ability to store a '1' at low supply voltages. However, the critical path LFBFF represents a compromise in sizing for saving a '1' and a '0'. The modifications that improve the FF's ability to hold a '1' simultaneously degrade the ability to hold a '0'. The critical path LFBFF simulated in Fig. 5 can hold a '1' or a '0' at and above 80mV in simulation. Measurements from the testchip show that the simulations give good approximations to the actual performance. The LFBFFs on the testchip are measured to hold both '1' and '0' consistently at and above 95mV.

Closed-Loop Standby Voltage Scaling

The minimum voltage scaling value in simulation shows some discrepancy from the value measured on the testchip. This indicates that an open-loop design for voltage scaling should be conservative. We propose that a closed-loop solution can achieve higher savings without losing state. By monitoring how close the critical path FFs are to failure, we can adjust the supply voltage for maximum savings under different environmental conditions such as varying temperatures.

We propose using "canary" LFBFFs sized to fail at higher supply voltages than the critical path flip-flops. The previous analysis shows that sizing within the LFBFF can adjust the failure point for holding a '0' or a '1'. Using this sizing analysis, we designed banks of LFBFFs designed to fail to hold a '1' or '0' at a range of higher voltages than the critical path flip-flops. A closed-loop control of the standby voltage supply based on feedback from the canary flip-flops can lower the supply voltage very close to the minimum value without causing the critical path flip-flops to fail.

Canary flip-flops can only provide a convincing guarantee that the core flip-flops are safe if they consistently fail at a higher supply voltage than the core flip-flops. To test the robustness of the canary FF approach, we designed two banks of canary

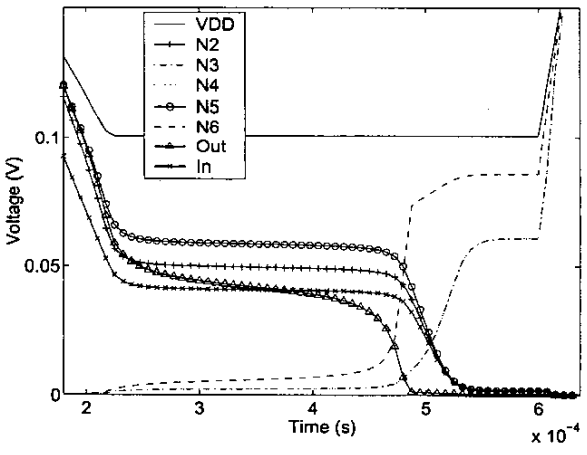

Fig. 4. Simulation of minimum sized LFBFF failing to hold a '1' at V_{DD}=100mV.

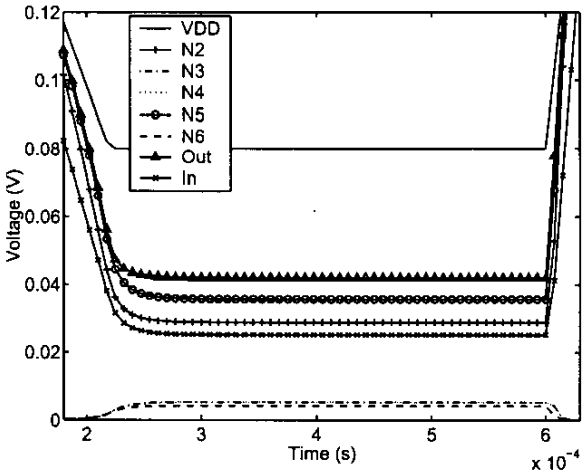

Fig. 5. Simulation of critical path LFBFF holding a '1' at V_{DD}=80mV.

FFs. The first bank is sized to fail at the typical process corner for nine different voltages above the critical path flip-flop while holding a '1'. The second bank serves the same task for storing a '0'.

Fig. 6 shows simulation results for the canary flip-flops at all process corners. The horizontal axis shows the sizing category for the FFs in each bank where number 1 is the critical path FF and number 10 is sized to fail at the highest voltage. The typical corner (TT) shows a smooth gradient of failure voltages across the sizing categories. The smoothness of the curve changes at other corners, but the critical path FFs always fail at the lowest supply voltage. This result indicates that the canary FFs will correctly alert a controller to "how close" the supply voltage comes to the ultimate failure point. Simulations of the canary FFs over different temperatures also maintain the gap between the canary and core failure points.

The process corner plot also tells exactly how much power savings the closed-loop approach can offer above the optimal open-loop design. The critical path FF fails at the worst case corner at just below 200mV. At the other corners, it operates down to below 100mV. Supposing that an aggressive open-loop design can safely scale V_{DD} to 200mV for all cases, the closed loop approach gives over 2X additional savings at any other corner by scaling down near 100mV (see Fig. 2). Thus, the closed loop approach offers significant power savings even over the optimal open-loop approach.

A controller using canary FFs has several other advantages as well. First, since the FF bank is designed to fail with a thermometer pattern, any notches in the failure sequence (i.e.- category 4 fails before category 5) would indicate large process variation within die. In response, the controller could scale

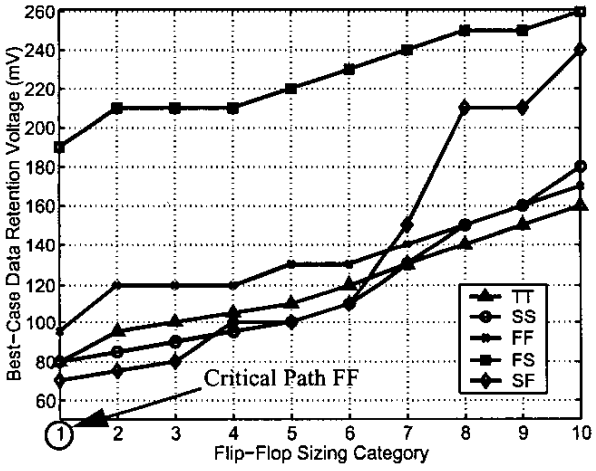

Fig. 6. Simulation of canary flip-flop banks at all process corners. The voltage plotted is the lowest supply voltage at which the given flip-flop continues to hold state correctly. Each line represents the worst case from the storing '0' and storing '1' banks for the given corner.

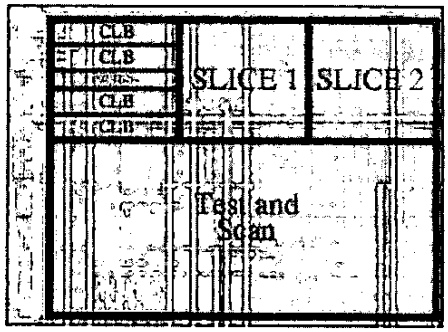

Fig. 7. Annotated Die Photo.

less aggressively. Secondly, the FF bank effectively estimates how close the core FFs are to failure. The controller can then decide how close to drive V_{DD} to the ultimate failure point. This range allows a trade-off between reliability and power savings. If the circuit state is critical, the controller can stop scaling after only 1 or 2 canary FFs fail. If the circuit state is not very important, the controller can scale more aggressively at the risk of generating some errors in the state. Finally, the closed-loop approach tracks changes in the environment.

Conclusion

We examine the potential savings from scaling V_{DD} in standby mode and conclude that lowering the supply close to the minimum value for holding state is worthwhile. Since open-loop approaches will need to be conservative, we propose a closed-loop approach using canary flip-flops. We show that the canary FFs consistently fail at higher supply voltages than the core FFs at all process corners. The success of the canary FFs makes a closed-loop approach to standby voltage scaling feasible. Measurements of the testchip in Fig. 7 show that this approach would give power savings of over 40X in our circuits. This is over 2X better than an optimal open-loop approach.

Acknowledgements

This work is funded by Texas Instruments; and by DARPA and the Air Force Research Lab, under agreement number F33615-02-2-4005. Ben is supported with an Infineon fellowship. We acknowledge Cypress Semiconductor for chip fabrication.

References

[1] S. Borkar, "Design Challenges of Technology Scaling", *IEEE Micro*, vol. 19, no. 4, pp. 23-29, July-Aug. 1999.

[2] A. Keshavarzi, K. Roy, and C. Hawkins, "Intrinsic Leakage in Low Power Deep Submicron CMOS ICs", *Int'l Test Conf.*, 1997.

[3] J. Sheu, D. Scharfetter, P. Ko, and M. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors", *JSSC*, vol.22, iss. 4, 1987.

[4] R. Krishnamurthy, et al, "High-performance and low-power challenges for sub-70nm microprocessor circuits", *CICC 2002*, 2002.

[5] T. Enamoto, et al, "A self-controllable-voltage-level (SVL) circuit for low-power, high-speed CMOS circuits", *ESSCIRC 2002*, 2002.

[6] J. Kao and A. Chandrakasan, "MTCMOS sequential circuits", *ESSCIRC 2001*, 2001.