4.2

A New Technique for Standby Leakage Reduction in High-Performance Circuits

Yibin Ye, Shekhar Borkar and Vivek De
Microcomputer Research Labs
Intel Corporation, Hillsboro, OR 97124, USA

Abstract

A new standby leakage control technique, which exploits the leakage reduction offered by transistor stacks with "more than one 'off' device", demonstrates 2X reduction in standby leakage power for a 32-bit static CMOS adder in a low-Vt, sub-1V, 0.1 µm technology. Leakage reduction is achieved with minimal overheads in area, power and process technology. The dynamics of leakage reduction due to transistor stacks, and its influence on the overall leakage power of large circuits are elucidated for the first time.

I. Introduction

A number of techniques such as multiple-threshold (MTCMOS) [1], dual-Vt [2] and reverse body bias [2,3] have been proposed in the past for reduction of processor leakage power during standby mode. In this paper, we propose a new standby leakage control scheme which exploits the large reduction in leakage current achievable by simultaneously turning off *more than one* transistor in nMOS or pMOS "stacks" (*i.e.* series-connected devices) between supply & ground. Typically, a large circuit block contains a significant number of logic gates where transistor "stacks" are already present (*e.g.* pMOS stack in NOR or nMOS stack in NAND gates). The technique described here enables effective leakage reduction during standby mode by installing a vector at the inputs of the circuit block so as to maximize the number of nMOS and pMOS stacks with "more than one 'off' device". In contrast to techniques reported in the past [1-3], the proposed scheme offers leakage reduction with minimal overheads in area, power and process technology. In particular, this technique can eliminate the need for a high-Vt device for standby leakage reduction in a sub-1V, 0.1 µm technology.

We use extensive circuit simulations of individual logic gates and a 32-bit static CMOS adder, designed in a sub-1V, 0.1 µm technology, to 1) elucidate the dynamics of leakage reduction due to transistor stacks, 2) examine its influence on the overall leakage power of the adder during both active and standby modes of operation, and 3) determine the standby leakage reductions yielded by application of the new leakage control technique. Two different Vt values are considered throughout the analysis. The low-Vt is 100mV smaller than the high-Vt.

II. Leakage Reduction due to Transistor Stacks

A 2-input NAND gate is used to illustrate the dynamics of leakage reduction in 2-transistor stacks with both devices 'off' (Fig. 1). From the dc solution of nMOS subthreshold current characteristics (Fig. 1), it is clear that the leakage current through a 2-transistor stack is approximately an order of magnitude smaller than the leakage of a single transistor. This reduction in leakage is due primarily to 1) negative gate-to-source biasing and 2) body-effect induced Vt increase in M1, or 3) reduced drain-to-source voltage in M2 which causes its Vt to increase, as the voltage Vm at the intermediate node converges to ~100mV (Fig. 1). Thus, as shown in Fig. 2, smaller amounts of leakage reduction are obtained at higher temperatures due to larger subthreshold swing. For 3- or 4-transistor stacks, the leakage reduction is found to be 2-3X larger (Fig. 3) in both nMOS and pMOS.

The time required for the leakage current in transistor stacks to converge to its final value is dictated by the rate of charging or discharging of the capacitance at the intermediate node by the subthreshold drain current of M1 or M2. This time constant (Fig. 4) is, therefore, determined by 1) drain-body junction and gate-overlap capacitances per unit width, 2) the input conditions immediately before the stack transistors are turned 'off', and 3) transistor subthreshold leakage current, which depends strongly on temperature and Vt. Therefore, the convergence rate of leakage current in transistor stacks increases rapidly with Vt reduction and temperature increase (Figs. 4 & 5). For minimum-Vt devices in a sub-1V, 0.1 µm technology, this time constant in 2-nMOS stacks at 110°C ranges from 5-50ns depending on input conditions before both devices are turned 'off'.

III. Dependence of Adder Leakage on Input Vectors

Increase in the active and standby leakage of the 32-bit static CMOS Kogg-Stone adder with Vt-reduction is shown to be smaller than that in individual transistors (Fig. 6) due to the presence of a significant number of transistor stacks in the design. The standby leakage power varies by 30%-40% (Fig. 7) depending on the input vector, which determines the number of transistor stacks in the design with more than one 'off' device. The adder leakage during active operation is dictated by the sequence of input vectors as well as the operating clock frequency (Fig. 8). Magnitude of the stack leakage time constant at elevated temperatures relative to the time interval between consecutive switching events determines the extent of convergence of the leakage to steady-state value. As a result, the active leakage corresponding to each input vector becomes higher as the clock frequency increases from 100 to 1000 MHz (Fig. 8), resulting in larger average leakage power at higher frequencies.

IV. Standby Leakage Control by Input Vector Activation

Fig. 9 shows an implementation of the new leakage reduction technique where a "standby" control signal, derived from the "clock gating" signal, is used to generate and store a predetermined vector in the static input latches of the adder during "standby" mode so as to maximize the number of nMOS and pMOS stacks with "more than one 'off' device". Since the desired input vector for leakage minimization is encoded by using a NAND or NOR gate in the feedback loop of the static latch (Fig. 9), minimal penalty is incurred in adder performance. As shown in Fig. 10, up to 2X reduction in standby leakage can be achieved by this technique. In order that the additional switching energy dissipated by the adder and latches, during entry into and exit from "standby mode", be less than 10% of the total leakage energy saved by this technique during standby, the adder must remain in standby mode for at least 5 µs (Fig. 11).

V. Conclusions

We demonstrate a new standby leakage control technique, which exploits the leakage reduction offered by transistor stacks with "more than one 'off' device". Up to 2X reduction in standby leakage power can be achieved by this technique with minimal overheads in area, power and process technology. We also elucidate the dynamics of leakage reduction due to transistor stacks, and its influence on overall leakage power of large circuits.

References

[1] S. Thompson *et. al.*, *1997 Symp. VLSI Tech.*, pp. 69-70
[2] S. Mutoh *et. al.*, *IEEE JSSC*, Aug. 1995, pp. 847-854
[3] T. Kuroda *et. al.*, *IEEE JSSC*, Nov. 1996, pp. 1770-1779

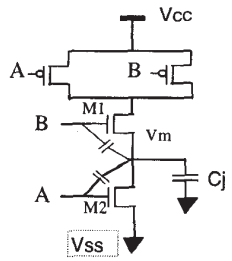


Fig. 1(a): 2-nMOS stack in a NAND gate

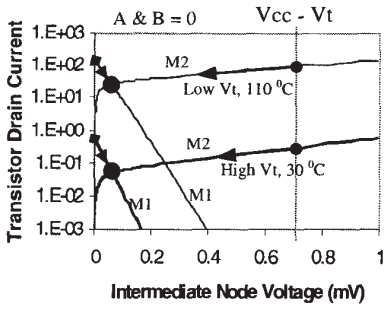


Fig 1(b): DC solution of 2 NMOS stacks

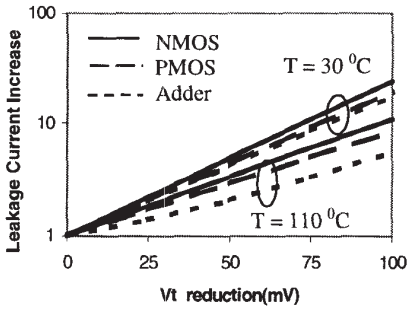


Fig. 6: Leakage current increase with Vt reduction for single transistors and an adder

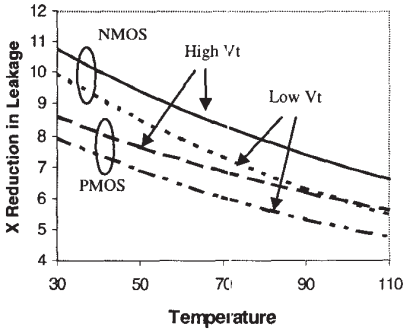


Fig.2: Leakage reduction by 2-nMOS & 2-pMOS stacks at different Vt and T

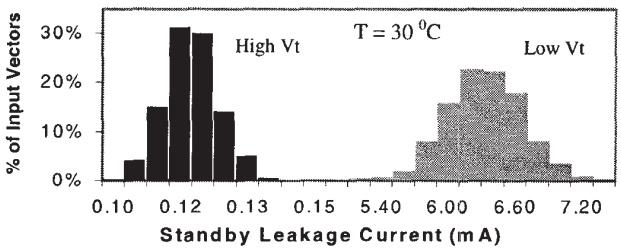


Fig. 7: Distribution of standby leakage current in the adder for a large number of random input vectors

	High Vt	Low Vt
2 NMOS	10.7X	9.96X
3 NMOS	21.1X	18.8X
4 NMOS	31.5X	26.7X
2 PMOS	8.6X	7.9X
3 PMOS	16.1X	13.7X
4 PMOS	23.1X	18.7X

Fig. 3: Leakage current reduction by 2-, 3-, & 4-transistor stacks at T=30°C

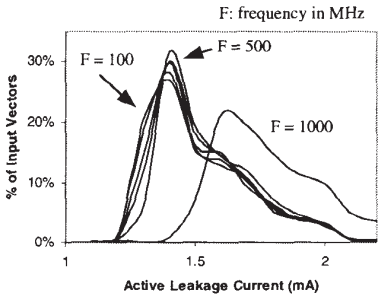


Fig. 8(a): Distribution of active leakage current of the adder with low-Vt devices at different clock frequencies

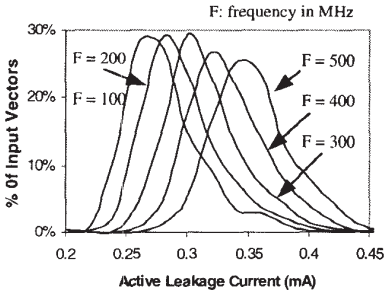


Fig. 8(b): Distribution of active leakage current of the adder with high-Vt devices at different clock frequencies

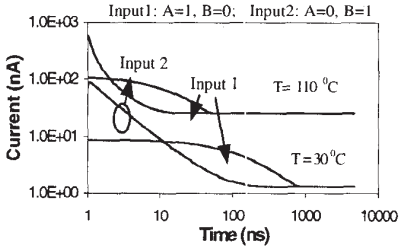
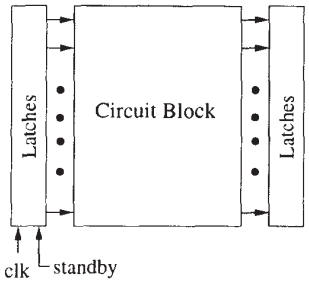
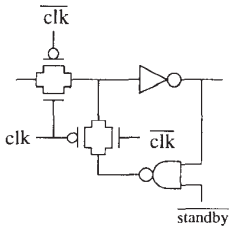


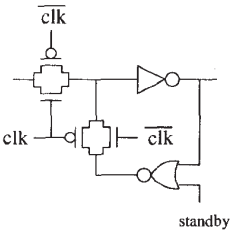
Fig. 4: Temporal behavior of leakage current in transistor stacks for different temperatures and initial input conditions



(a) Block diagram



(b) A latch to store "0"



(c) A latch to store "1"

Fig. 9: An implementation of the standby leakage control scheme through input vector activation

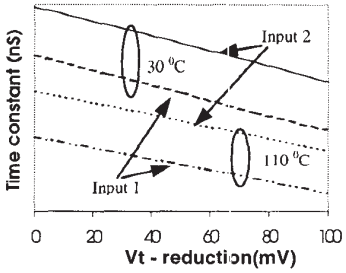


Fig. 5: Dependence of the time constant for stack leakage on Vt, temperature and initial input conditions

		% Reduction
High Vt	Avg.	35.4%
	Worst	60.7%
Low Vt	Avg.	33.3%
	Worst	56.5%

Fig. 10: Adder leakage current reduction by "best" input vector activation compared t average and worst standby leakage

	High Vt	Low Vt
Savings	2.2uA	0.0384mA
Overhead	1.64 nJ	1.84 nJ
Min. time in standby	84 uS	5.4 uS

Fig. 11: Standby leakage power savings and the minimum time required in standby mode