ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 9, March 2013

Leakage Power Reduction in Clustered Sleep Transistors

P.Sreenivasulu, Dr. K.Srinivasa Rao, Pavan Kumar Kaja, Dr. A.Vinaya babu

Abstract ***Low-power design has changed the design and verification paradigm forever. The industry is in dire need of a systematic approach to low-power verification. Power consumption of Very Large Scale Integrated (VLSI) circuits has been growing at an alarmingly rapid rate. This increase in power consumption, coupled with the increasing demand for portable/hand-held electronics, has made power consumption a dominant concern in the design of VLSI circuits today. Both leakage and dynamic power are increasing with every switch to smaller process geometries. However, due to process scaling trends, leakage power has become a major component of the total power consumption in VLSI circuits. Here in this paper we aimed to discuss some leakage power reduction techniques. One of the technique to reduce the leakage of a circuit is MTCMOS (Multi-threshold CMOS) approach is discussed.***

Key Words: **Low power, Static dissipation, dynamic dissipation, leakage current, MTCMOS, power gating.**

I. INTRODUCTION

To achieve higher density and performance and lower power consumption, CMOS devices have been scaled for more than 30 years. Transistor delay times decrease by more than 30% per technology generation, resulting in doubling of microprocessor performance every two years. Supply voltage has been scaled down in order to keep the power consumption under control. Hence, the transistor threshold voltage has to be commensurately scaled to maintain a high drive current and achieve performance improvement. With the advent of technology, the reduction of the supply voltage has become vital to reduce dynamic power and avoid reliability problems in deep submicron (DSM) regimes. However, reducing alone causes serious degradation to the circuit's performance. One way to maintain performance is to scale down both and the threshold voltage. However, reducing increases the sub threshold leakage current exponentially. This problem escalates in DSM technologies. The sub threshold leakage current can be approximately formulated as

$$I_{\text{leakage}} = I_0 e^{(V_{\text{gs}} - V_{\text{th}})/nV_T}$$

where , is the gate oxide capacitance, is the width to length ratio of the leaking MOS device, is the zero bias mobility, is the gate to source voltage, is the thermal voltage which is about 26 mV at , and is the sub threshold swing coefficient given by with being the depletion layer capacitance of the source/drain junctions. From (1), it is evident that the leakage current is exponentially proportional to. Therefore, leakage could be reduced by increasing or reducing. Over the past decade, several techniques have been proposed to reduce leakage power during the standby mode by increasing. In the Variable Threshold CMOS (VTCMOS) approach [1], the threshold voltage is controlled dynamically through varying the substrate bias voltage. In this scheme, all transistors have low threshold voltage (LVT) and the substrate bias is altered so as to: 1) compensate for fluctuations in the active mode and accordingly minimize delay variations and 2) reduce leakage current in the standby mode. Two drawbacks to the VTCMOS approach are: 1) since is proportional to the square root of the substrate voltage, a large change in the later is thus required to change by effective values and 2) VTCMOS requires a triple-well structure as well as a charge-pump circuit to produce the substrate voltage.

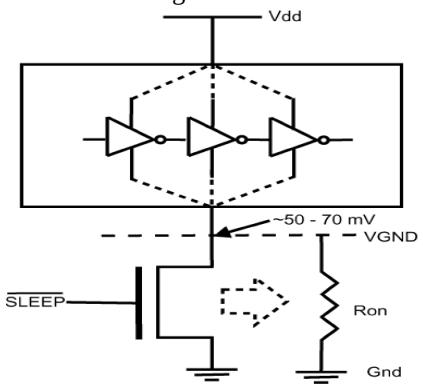

Fig. 1. Logic block with an nMOS sleep transistor.

Another technique is the Multi Voltage CMOS (MVCMOS) scheme [2]. The MVCMOS technique employs LVT transistors whose gate voltages are driven in the sleep mode to larger than and smaller than for the PMOS and NMOS, respectively.

II. LEAKAGE REDUCTION TECHNIQUES

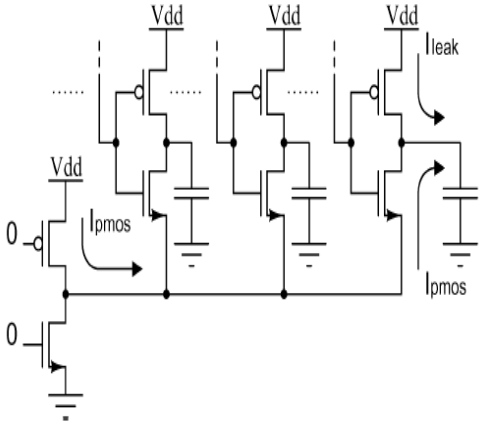

Fig. 2. Charging Of The Zero Output Nodes With A Pmos Pull-Up.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 9, March 2013

For a CMOS circuit, the total power dissipation includes dynamic and static components during the active mode of operation. In the standby mode, the power dissipation is due to the standby leakage current. Dynamic power dissipation consists of two components. One is the switching power due to charging and discharging of load capacitance. The other is short circuit power due to the nonzero rise and fall time of input waveforms. The static power of a CMOS circuit is determined by the leakage current through each transistor.
The dynamic (switching) power and leakage power are expressed as

$$P_D = \alpha f C V_{dd}^2$$
$$P_{LEAK} = I_{LEAK} \cdot V_{dd}$$

III. POWER GATING AND MULTI-THRESHOLD CMOS

The most natural way of lowering the leakage power dissipation of a VLSI circuit in the STANDBY state is to turn off its supply voltage. This can be done by using one PMOS transistor and one NMOS transistor in series with the transistors of each logic block to create a virtual ground and a virtual power supply as depicted in Figure 4. Notice that in practice only one transistor is necessary. Because of their lower on-resistance, NMOS transistors are usually used. The technique we used for calculating the size of the sleep transistor. Techniques used for gate clustering and assignment. Our results are summarized in Section IV-D. In Section V, four hybrid clustering techniques are proposed and compared. Section VI introduces the noise on the virtual ground rail as a design criterion and results for the techniques are shown in Section VII when noise is taken into account. If a current-time graph is constructed of the discharged currents, , and overlap in time in Case I. On the other hand, no overlap in time occurs for Case II. An intermediate case occurs when the discharged currents "partially" overlap, if the LVT logic blocks have slightly different discharge times.

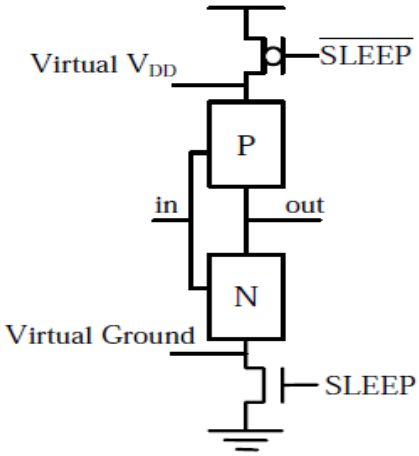

Fig 3: Power Gating Circuit.

In the ACTIVE state, the sleep transistor is on. Therefore, the circuit functions as usual. In the STANDBY state, the transistor is turned off, which disconnects the gate from the ground. Note that to lower the leakage, the threshold voltage of the sleep transistor must be large. Otherwise, the sleep transistor will have a high leakage current, which will make the power gating less effective. Additional savings may be achieved if the width of the sleep transistor is smaller than the combined width of the transistors in the pull-down network. In practice, Dual VT CMOS or Multi-Threshold CMOS (MTCMOS) is used for power gating [14][15]. In these technologies there are several types of transistors with different VT values. Transistors with a low VT are used to implement the logic, while high-VT devices are used as sleep transistors.

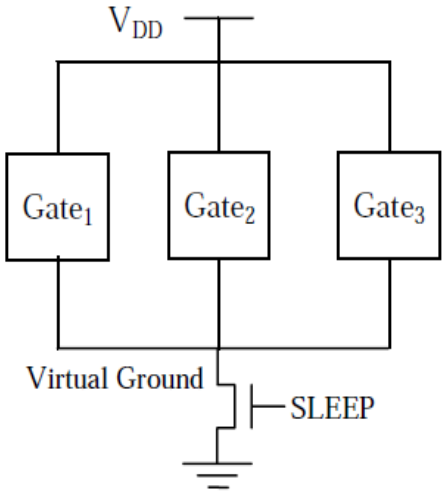

Fig 4: Using One Sleep Transistor for Several Gates.

This gives an *expected* discharge current value. The switching activity of a gate is computed by multiplying the probability that the output of the gate will be at zero by the probability it will be at one [13]. If the switching activity is not accounted for, the design problem would be very pessimistic and the sleep transistor will be oversized, causing substantial increase in leakage and dynamic power dissipation as well as in the die size. It is very unlikely that the clustered gates would have their worst-case current discharge at the same time. This has been deduced by exhaustively applying all input vectors to the CLA adder benchmark.

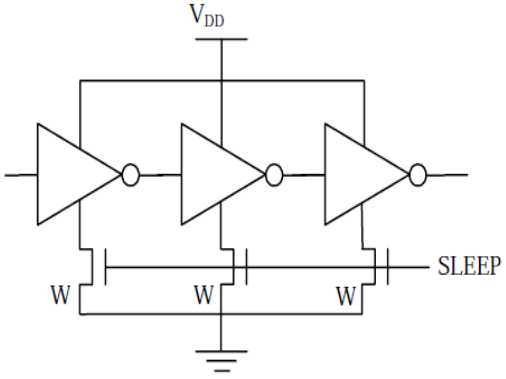

Fig 5: Sleep Transistor Sharing.

With the advent of portable electronic devices and high density VLSI circuits, power dissipation has emerged as a major design concern. Simultaneously, technological advances result in shrinking of the minimum feature size as well as in lowering of the supply and threshold voltages which cause a dramatic increase of the leakage current.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 9, March 2013

IV. LEAKAGE CONTROL IN ACTIVE MODE

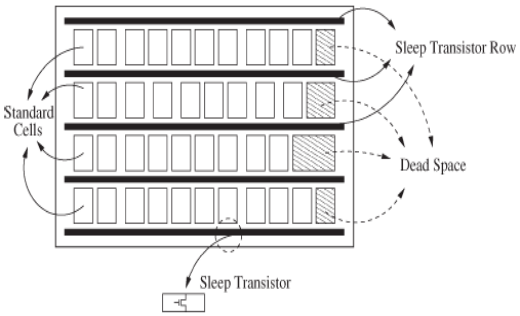

Fig. 6. Proposed Design Methodology.

For circuits whose power consumption is dominated by the leakage power in STANDBY mode, the aforementioned techniques can be used to reduce the power consumption. On the other hand, if a circuit's power consumption is not dominated by the leakage in STANDBY mode, then it is necessary to consider the total power consumption (including switching power dissipation and active mode leakage) and optimize the circuit to reduce it as described next.

- Multiple Threshold Cells
- Long Channel Devices
- Minimum Leakage Vector Method
- Stack Effect-based Method
- Sizing with Simultaneous Threshold and Supply Voltage Assignment

A. *CLUSTERING PHASE*

Our clustering algorithm uses a constrained minimization approach with a single cost function under given timing constraints. The clustering algorithm is based on an iterative row selection scheme, in which, starting from an initial solution where all rows are selected (i.e., gated), rows are progressively eliminated from the solution until we reach a maximal set of rows which minimizes the leakage cost function, under given timing constraints. To facilitate the vector comparisons and to offer an automated design environment, every discharge current at the output of a gate is represented by a vector.

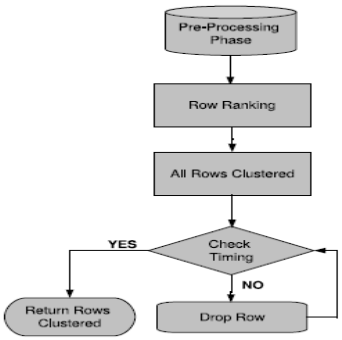

Fig 7: Clustering Algorithm.

The proposed scheme can be easily modified to work with only a single set of sleep transistors. The load-dependent delay *di* of a gate *i* in the absence of a sleep transistor can be expressed as

$$d^i \propto \frac{C_L V_{dd}}{(V_{dd} - V_{\mathrm{tL}})^{\alpha}}$$

Where *CL* is the load capacitance at the gate output, *V*tL is the low-voltage threshold = 350 mV, *Vdd* = 1.8 V, and α is the velocity saturation index (≈1.3 in 0.18-μm CMOS technology).

In the presence of a sleep transistor, the propagation delay of a gate can be expressed as

$$d^i_{\mathrm{sleep}} = \frac{KC_L V_{dd}}{(V_{dd} - 2V_x - V_{\mathrm{tL}})^{\alpha}}$$

where *Vx* is the potential of the virtual rails, as shown in Fig. 1, and *K* is the proportionality constant. can see now that in the active mode, the logic gate operates not at the true supply rails (*Vdd*) but at the virtual supply rails that are offset by a magnitude *Vx* from the supply rails on either side. Hence, the effective supply voltage seen by the gate is *Vdd* − 2*Vx*. Let us suppose that *I*sleepON is the current flowing in the gate during the active mode of operation. During this mode, the sleep transistor is in the linear region of operation. Using the basic device equations for a transistor in the linear region, the drainto-source current in the sleep transistor (which is the same as *I*sleepON) is given by

$$I_{\mathrm{sleep_{ON}}} = \mu_n C_{\mathrm{ox}} (W/L)_{\mathrm{sleep}} \left((V_{dd} - V_{\mathrm{tH}}) V_x - \frac{V_x^2}{2} \right)$$

$$I_{\mathrm{sleep_{ON}}} \simeq \mu_n C_{\mathrm{ox}} (W/L)_{\mathrm{sleep}} (V_{dd} - V_{\mathrm{tH}}) V_x.$$

The sub threshold leakage current *I*leak in sleep mode will be determined by the sleep transistor and is expressed as given in [13] as

$$I_{\mathrm{leak}} = \mu_n C_{\mathrm{ox}} (W/L)_{\mathrm{sleep}} e^{1.8} V_T^2 e^{\frac{V_{\mathrm{gs}} - V_{\mathrm{tH}}}{nV_T}} \left(1 - e^{\frac{-V_{\mathrm{ds}}}{V_T}} \right)$$

Where μn is the *N*-mobility, *C*ox is the oxide capacitance, *V*tH is the high-threshold voltage (= 500 mV), *VT* is the thermal voltage = 26 mV, and *n* is the sub threshold swing parameter. Equation (2) establishes a relation between delay of a gate *di* sleep and *Vx*. By replacing *Vx* in (4) in terms of *di* sleep [using (2)], we get the dependence between (*W/L*) sleep and *di* sleep (assuming the ON current is constant for each gate). Thus, a range of (*W/L*)sleep for the sleep transistor would correspond to a range of gate delays. Finally, (*W/L*) sleep in (5) can be replaced in terms of *di* sleep, hence establishing a relationship between gate delay and gate leakage. The final relation between leakage and delay can be expressed as

$$I_{\mathrm{leak}} = \mu_n C_{\mathrm{ox}} e^{1.8} V_T^2 e^{\frac{V_{\mathrm{gs}} - V_{\mathrm{tH}}}{nV_T}} \left(1 - e^{\frac{-V_{\mathrm{ds}}}{V_T}} \right) \frac{I_{\mathrm{sleep_{ON}}}}{\mu_n C_{\mathrm{ox}} (V_{dd} - V_{\mathrm{tH}})} \times \frac{d_{\mathrm{sleep}}^{1/\alpha}}{(V_{dd} - V_{\mathrm{tL}}) d_{\mathrm{sleep}}^{1/\alpha} - (KC_L V_{dd})^{1/\alpha}}.$$

To illustrate our techniques, six benchmarks are used as test vehicles: a 4-bit carry look ahead (CLA) adder, a 32-bit priority checker, a six-bit array multiplier design, a four-bit

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 9, March 2013

ALU/Function Generator (74181 ISCAS'85 benchmark), a 32-single error correcting circuit (C499 ISCAS'85 benchmark), and a 27-bit channel interrupt controller (CIC) (C432 ISCAS'85 benchmark). These benchmarks have been chosen to offer a variety of circuits with different structures that employ various gates with different fanouts. The 4-bit CLA adder will be first used to demonstrate the proposed techniques. Results pertaining to all other benchmarks will be provided in Section IV-D. Fig. 2 shows a schematic diagram of the CLA adder which consists of 28 gates – . All gates are implemented in 0.18- CMOS technology. To illustrate our proposed technique, a preprocessing stage of gate currents is described in the next section. This stage will be utilized to solve the BP problem, and later on to solve the SP problem.

Algorithm: *Time_Frame_Partitioning*($MIC(C_i,T_j)$, n)

1: **Output:** An efficient partitioning
2: /* step 1: mark the candidate time units */
3: **for** $j \leftarrow 1$ to NUM_TF **do**
4: **if** (the n+1 largest $MIC(C_i)$ occurs in T_j) **then**
5: mark(T_j);
6: **end if**
7: **end for**
8: /* step 2: *n-way* partitioning */
9: use n cuts to separate the marked T_j;
10: **return**

Fig 8. Variable Length *N-Way* Partitioning Algorithm.

V. SIMULATION RESULTS

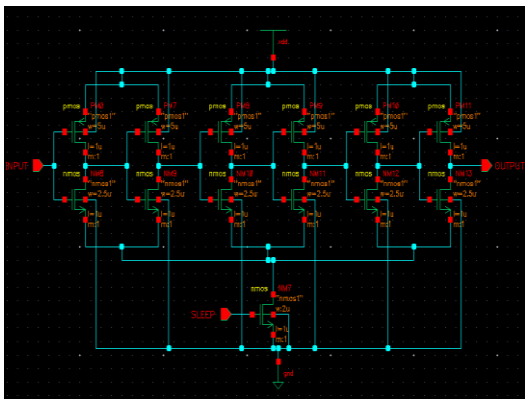

Fig 9. Sleep Transistor

The wire and junction capacitance associated with the virtual ground line should actually help reduce the ground bounce by serving as a local charge sink or reservoir for current [9] as shown in Fig. 23. However, this capacitance would have to be extremely large in order to offset the effects of a poorly sized sleep transistor. The *RC* network serves as a low pass filter, where the *RC* time constant must be large enough such that the virtual ground voltage can only rise to a fraction of its peak dc value.

In order to include ground bounce as a design criterion, dynamic and leakage power are reduced under two constraints that must be achieved simultaneously. Firstly, the speed degradation is set to never exceed 5%, and secondly, ground bounce is also set to never exceed 50 mV. Based on these constraints, the circuit is guaranteed to achieve sufficient speed and noise margins.

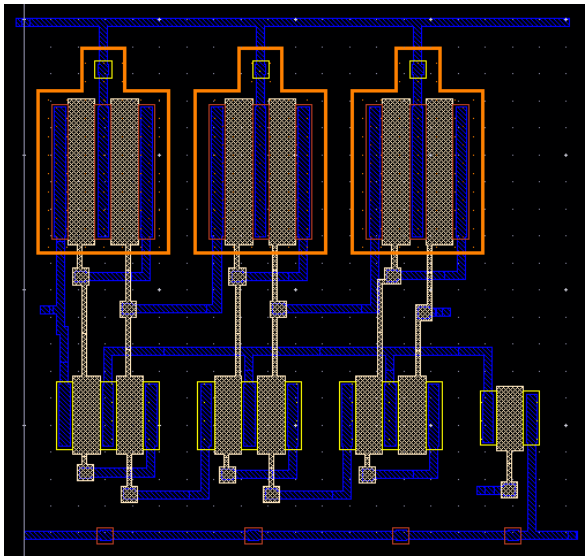

Fig 10. Final Parasitic Extraction Layout

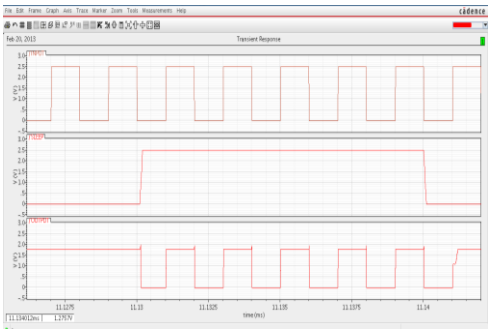

Fig 11. Output Wave Forms Of Sleep Transistor

There are two major leakage mechanisms in scaled devices: sub threshold leakage and gate leakage[10]. Both of these leakage components are increasing very rapidly with technology scaling. Furthermore, to get sufficiently large capacitance, decaps are usually made from thin gate-oxide MOS transistors with long and wide channels.

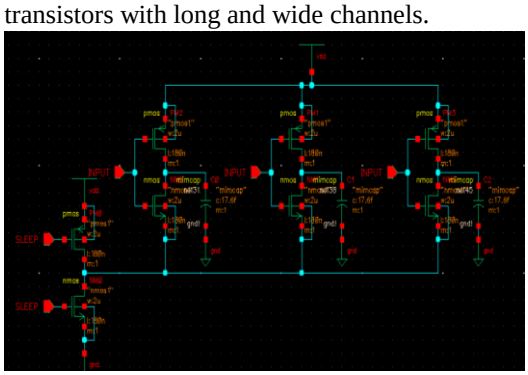

Fig 12.Sleep Transistor 2

Power grid noise is caused by simultaneous switching of the core logic and I/O buffers. Sleep transistors introduce power-gating noise when they are switching on and off.

This gives an *expected* discharge current value. The switching activity of a gate is computed by multiplying the probability that the output of the gate will be at zero by the probability it will be at one [13]. If the switching activity is not accounted for, the design problem would be very pessimistic and the sleep transistor will be oversized, causing substantial increase in leakage and dynamic power dissipation

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 9, March 2013

as well as in the die size. It is very unlikely that the clustered gates would have their worst-case current discharge at the same time. This has been deduced by exhaustively applying all input vectors to the CLA adder benchmark.

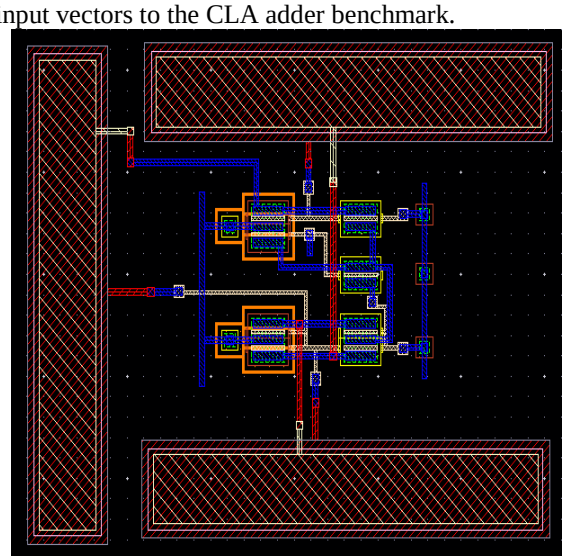

Fig 13.Final Layout of Sleep Transistor 2

VI. CONCLUSION

In this paper, we have presented a new, layout-aware power gating methodology for leakage power reduction in nanometer CMOS circuits. Our methodology allows row-based, clustered power-gating and it features minimal perturbation of the original layout compared to existing sleep transistor insertion techniques. This favors fast design closure and makes the methodology suitable for the implementation as a CAD tool. Sleep transistors are effective to reduce both dynamic and leakage power. Sleep transistor can be viewed as an essential part of the power/ground network. Leakage power consumption is becoming dominant in deep sub-micron CMOS technologies, and different approaches for limiting it are now appearing in the scientific literature. Experimental data show leakage power reductions around 80% (total power savings, accounting for cell dynamic and internal power, are around 19%), with a circuit delay increase of 5% caused by active-mode slow-down due to the insertion of the sleep transistors and an average area overhead around 2.5%.

REFERENCES

[1] F. Fallah and M. Pedram, "Standby and active leakage current control and minimization in CMOS VLSI circuits," IEICE Trans. Electron., vol. E88-C, no. 4, pp. 509–519, Apr. 2005.

[2] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-sub micrometer CMOS circuits," Proc. IEEE, vol. 91, no. 2, pp. 305–327, Feb. 2003.

[3] V. Khandelwal and A. Srivastava, "Leakage control through fine-grained placement and sizing of sleep transistors," in Proc. IEEE/ACM Int. Conf. CAD (ICCAD), San Jose, CA, Nov. 2004, pp. 533–536.

[4] M. Keating, D. Flynn, R. Aitken, A. Gibbons, and K. Shi, Low Power Methodology Manual: For System-on-Chip Design. New York: Springer, 2007.

[5] H. Jiang, M. Marek-Sadowska, and S. R. Nassif, "Benefits and costs of power-gating technique," in Proc. IEEE Int. Conf. Comput. Des.(ICCD), San Jose, CA, Oct. 2005, pp. 559–566.

[6] S. Mutoh, S. Shigematsu, Y. Matsuya, H. Fukuda, and T. Kaneko, "1-V power supply high-speed digital circuit technology with multithreshold voltage CMOS," IEEE J. Solid-State Circuits, vol. 30, no. 8, pp. 847–854, Aug. 1995.

[7] J. Kao, A. Chandrakasan, and D. Antoniadis, "Transistor sizing issues and tool for multi-threshold CMOS technology," in Proc. ACM/IEEE Des. Autom. Conf. (DAC), Anaheim, CA, Jun. 1997, pp. 409–414.

[8] J. Kao, A. Chandrakasan, and S. Narendra, "MTCMOS hierarchical sizing based on mutual exclusive discharge patterns," in Proc. ACM/ IEEE Des. Autom. Conf. (DAC), San Francisco, CA, Jun. 1998, pp. 495–500.

[9] M. Anis, S. Areibi, and M. Elmasry, "Design and optimization of multithreshold CMOS (MTCMOS) circuits," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 22, no. 10, pp. 1324–1342, Oct. 2003.

[10] M. Anis, S. Areibi, S. Mahmoud, and M. Elmasry, "Dynamic and leakage power reduction in MTCMOS circuits using an automated efficient gate clustering technique," in Proc. ACM/IEEE Des. Autom. Conf. (DAC), New Orleans, LA, Jun. 2002, pp. 480–485.

[11] W.Wang, M. Anis, and S. Areibi, "Fast techniques for standby leakage reduction in MTCMOS circuits," in Proc. IEEE Syst.-on-Chip Conf. (SOCC), Santa Clara, CA, Sep. 2004, pp. 21–24.