Novel Approaches to Low Leakage and Area Efficient VLSI Design

Tajrian Izma[1], Parag Barua[1], Md. Rejaur Rahman[1], Prianka Sengupta[1] and M.S. Islam[2*]
[1]*Department of Electrical and Electronic Engineering (EEE), BRAC University, Dhaka-1212, Bangladesh*
[2]*Department of EEE, Uttara University, Dhanmondi, Dhaka-1209, Bangladesh*
[*]Corresponding author: islams@eee.buet.ac.bd, parag_ndc@yahoo.com

Abstract: **The development of digital integrated circuits is challenged by higher power consumption. The combination of higher clock speeds, greater functional integration, and smaller process geometries has contributed to significant growth in power density. Scaling improves transistor density and functionality on a chip. Scaling helps to increase speed and frequency of operation and hence higher performance. As voltages scale downward with the geometries threshold voltages must also decrease to gain the performance advantages of the new technology but leakage current increases exponentially. Thinner gate oxides have led to an increase in gate leakage current. Today leakage power has become an increasingly important issue in processor hardware and software design. With the main component of leakage, the sub-threshold current, exponentially increasing with decreasing device dimensions, leakage commands an ever increasing share in the processor power consumption. In 65 nm and below technologies, leakage accounts for 30-40% of processor power. According to the International Technology Roadmap for Semiconductors (ITRS) [1], leakage power dissipation may eventually dominate total power consumption as technology feature sizes shrink. While there are several process technology and circuit-level solutions to reduce leakage in processors, we propose novel approaches for reducing both leakage and dynamic power with minimum possible area and delay trade off.**

I. INTRODUCTION

For the most recent CMOS feature sizes (e.g., 90nm and 65nm), leakage power dissipation has become an overriding concern for VLSI circuit designers. International technology roadmap for semiconductors (ITRS) [1] reports that leakage power dissipation may concern for low-power chip designers since dynamic power accounted for 90% or more of the total chip power. Therefore, many previously proposed techniques, such as voltage and frequency scaling, focused on dynamic power reduction. However, as the technology feature size shrinks, e.g., to 90 nm and 65 nm, static power dissipation has become a great challenge for current and future technologies. In this paper, we proposed two new approaches for leakage reduction. Hence we compared our approaches with prior techniques in this paper.

II. NOVEL STACKED SLEEP APPROACH

In this Novel technique there are two sleep transistors is forcedly stacked in both pull-up and pull-down networks. So that we called it "Stacked Sleep" approach (Fig. 1). For maintaining the exact logic state, two extra transistors are used with the pull-up and pull-down networks. Basically the sleep transistor is used in this technique to reduce the leakage power in sleep mode. In this novel technique leakage is reduced in two ways, firstly stacked effect [2] of sleep transistor and secondly, sleep transistor effect.

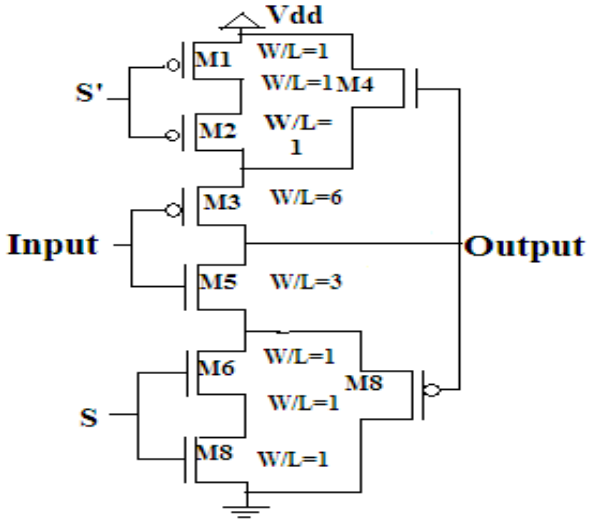

Fig. 1: Novel stacked sleep approach

Sleep transistors induce extra internal resistance and according to the Narendra *et al.* study stack effect increases the channel length of the circuit. Sleep transistor is turned off when the exact generic logic circuit are not in use. Moreover sleep transistor isolated the logic network and that can help to reduce the leakage power dramatically during sleep mode. As like as the prior approaches dual V_{th} can be applied in this novel technique. Now we explain how the stacked sleep works during active mode and during sleep mode. Also, we explain how the leakage power is saved by using the stacked sleep approach. Despite all the sleep transistors used in this technique are turned on during active mode and turned off during sleep mode. During active mode (Fig.1), s=1 and s'=0 are asserted and thus all sleep transistors are turned ON. Since, the entire sleep transistors in pull-up network (M1 and M2) and pull-down network (M6 and M8) are turned ON during active mode, so that circuit will perform faster. Furthermore, high V_{th} is used in this technique which can incur some delay penalty but reduces the leakage substantially. During sleep mode s=0 and s'=1 are asserted

and so all sleep transistors in both pull-up network (M1 and M2) and pull-down network (M6 and M8) are turned OFF. Although all sleep transistors are turned OFF but the circuit still maintains exact logic level by the help of two parallel transistors (M4 and M8). When sleep transistors are turned OFF, induce stack effect suppresses leakage power consumption. By combining this effect stacked sleep approaches can achieve low leakage. However its increase some delay penalty.

III. NOVEL VARIABLE BODY BIASING TECHNIQUE

Variable body biasing technique is combined structure of forced sleep and sleep transistor techniques. But the basic difference of this circuit is that source of PMOS is connected with body of another PMOS in pull-up network. Meanwhile, in pull-down network, source of NMOS is connected with body of another NMOS.

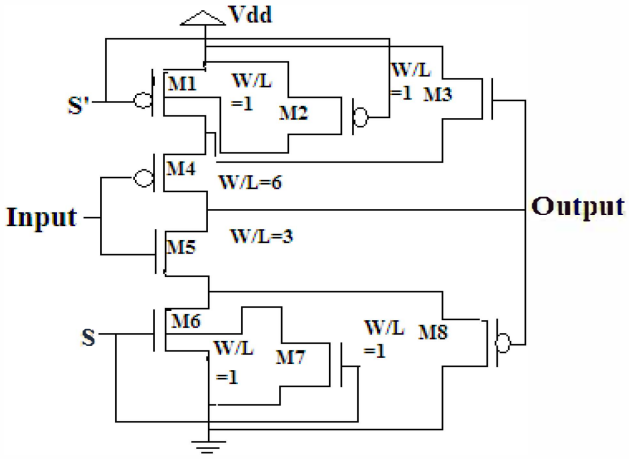

Fig. 2: Novel variable body biasing technique

Hence we named it "Variable body biasing" technique (Fig. 2). This Novel technique helps to reduce the leakage power by changing the body bias of transistors dramatically. Moreover, the source voltages of PMOS transistor (M2) and NMOS transistor (M7) depend on the gate voltages of M1 and M6 transistor, respectively. Thus this effect changes threshold voltage and reduces the leakage power substantially. During active mode, s=1 and s'=0 are asserted, thus turned ON the entire sleep transistors. When all sleep transistors are turned ON then the internal resistance in between increases. During sleep mode, s=0 and s'=1 are asserted; all sleep transistors in pull-up network (M1 and M2) and pull-down network (M6 and M7) are turned OFF. Meanwhile, for retaining the exact logic state in sleep mode two extra transistors (M3 and M8) are paralleled with the sleep transistors in both pull-up and pull-down networks. So it is a state saving technique. However, high and low V_{th} can be applied in this novel technique.

IV. SIMULATION RESULTS

We compared the proposed stacked sleep approach and variable body biasing technique with previous approaches namely dual sleep, sleepy keeper and sleepy stack. Here we compare our design approaches in terms of power consumption (dynamic and static), propagation delay and area with the prior approaches [2-7]. Moreover, in the novel techniques we used low V_{th} almost for all transistors. High V_{th} can be applied also. Our two novel techniques are suitable for generic logic design and memory cell. For calculating power consumption and delay we used Avant! Star-HSPICE software [8]. All the simulation processes are done for a chain of four inverters and a SRAM cell. The inverter chains have aspect ratio W/L= 6 for PMOS and W/L= 3 for NMOS and rest of transistors aspect ratios are W/L=1. We choose five different technologies from BSIM4 PTM [9] to observe the changes of power and delay as the technology feature size shrinks. The chosen technologies, i.e., 22nm, 32nm, 45nm, 90nm and 130nm, use supply voltages of 0.8V, 0.9V, 1V, 1.2V and 1.3V, respectively. We do consider single V_{th} for proposed methods but for sleepy stack, sleepy keeper and dual sleep both single and dual V_{th} are used. We measure the dynamic and static power of the chain of 4 inverters by asserting '1' and '0' repeatedly. For the chain of 4 inverters, we measure two different propagation delay values: one when an output goes high and another when an output goes low. We take the larger value as the worst case propagation delay of the chain of 4 inverters. We directly measure the area of two novel techniques from an actual full layout we did for each. We perform two different experiments. We first compare the stacked sleep and variable body biasing with three well-known techniques, i.e., sleepy stack, sleepy keeper and dual sleep, while scaling transistors technologies for chain of four inverters and secondly for SRAM cell. We use the "Delvto" option of HSPICE to change V_{th}.

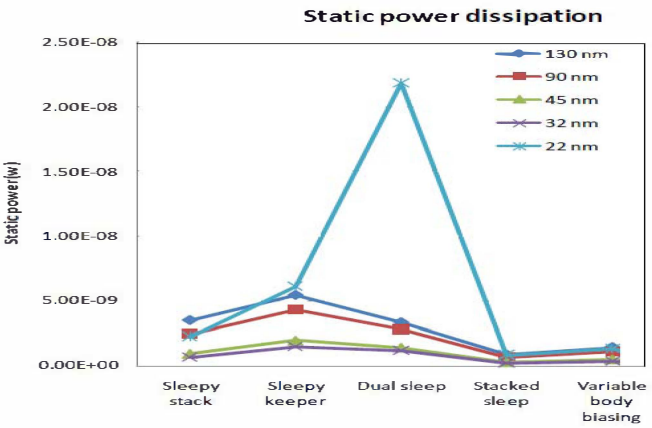

Fig. 3: Static power dissipation (chain of four inverters)

The static power dissipation, dynamic power dissipation, propagation delay and area for a chain of four inverters are shown in Fig.3, Fig.4, Fig.5 and Fig.6, respectively.

The static power dissipation, dynamic power dissipation, propagation delay and area for a SRAM cell are shown in Fig.7, Fig.8, Fig.9 and Fig.10, respectively.

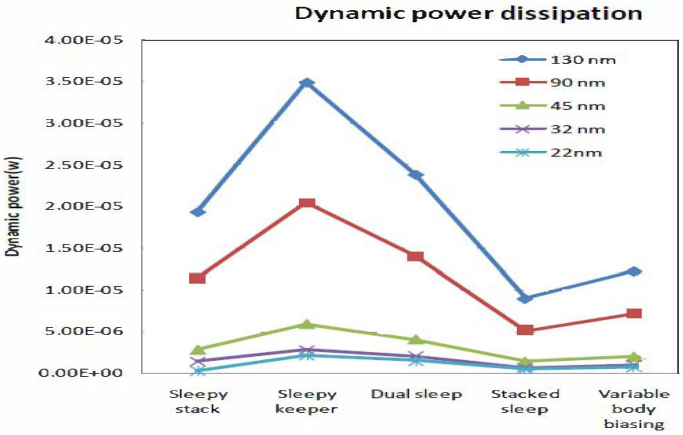

Fig. 4: Dynamic power dissipation (chain of four inverters)

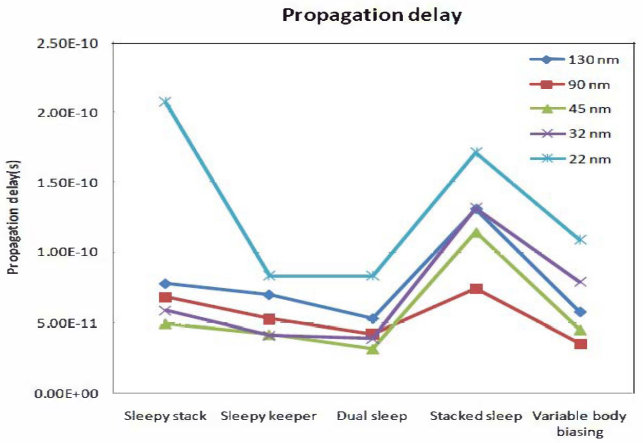

Fig. 5: Propagation delay (chain of four inverters)

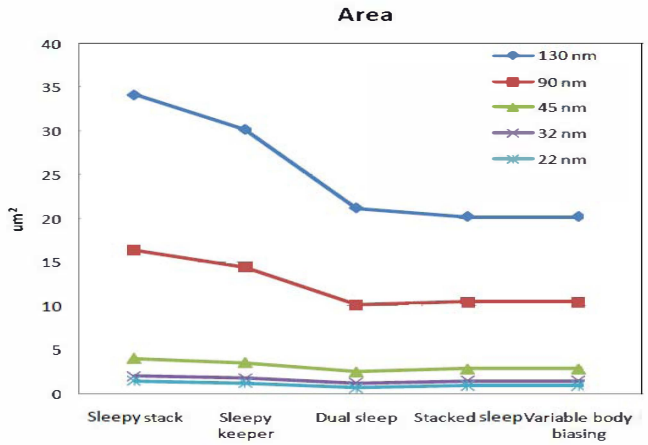

Fig. 6: Area (chain of four inverters)

The comparison of stacked sleep approach (32 nm technology) and variable body biasing technique (32 nm technology) with dual sleep method for a chain of four inverters and a SRAM cell are shown in Table-1, Table-2 Table-3, Table-4 respectively. Here "+" and "–" indicate better and worst performance while comparing stacked sleep and variable body biasing with the dual sleep.

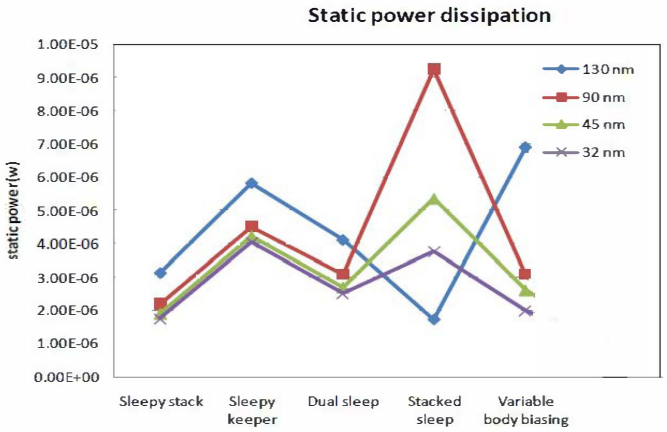

Fig. 7: Static power dissipation (SRAM cell)

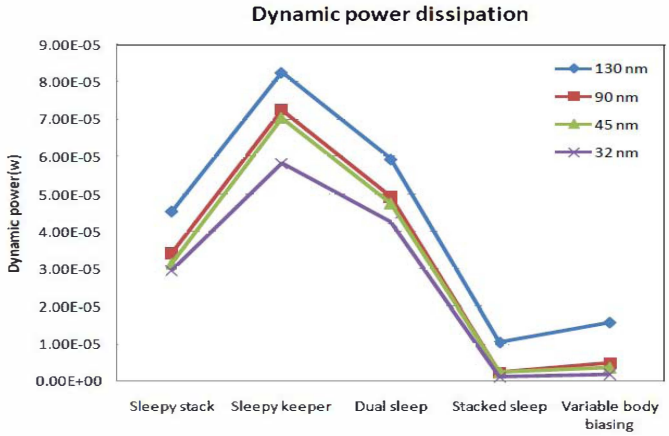

Fig. 8: Dynamic power dissipation (SRAM cell)

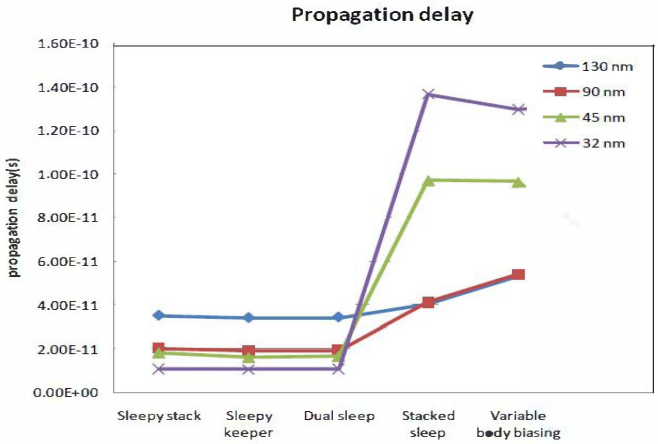

Fig. 9: Propagation delay (SRAM cell)

In case of a chain of four inverters stacked sleep approach shows 83.41%, 65.49% better and 238.08%, 20.31% worse than dual sleep method in static power, dynamic power, propagation delay and area, respectively.

In case of a chain of four inverters variable body biasing technique shows 69.29%, 50.44% better performance and 102.49%, 20.31% worse performance than the Dual sleep method in terms of static power, dynamic power, propagation delay and area, respectively.

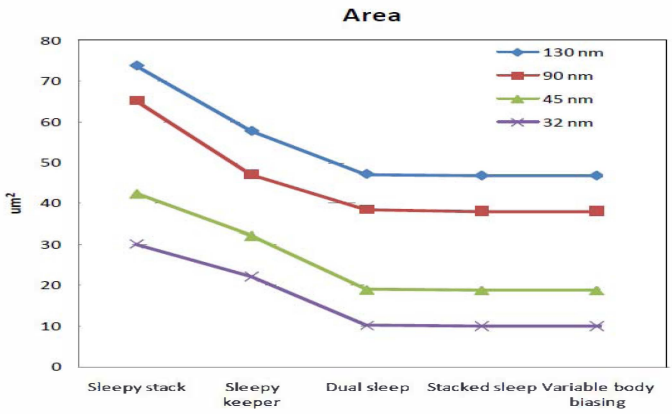

Fig. 10: Area (SRAM cell)

In case of a SRAM cell stacked sleep shows 97.11%, 1% better performance and 49.59%, 112.6% worss performance than the Dual sleep method in terms of dynamic power, area, static power and propagation delay, respectively.

Table-1: Comparison of stacked sleep approach in case of a chain of four inverters.

Method	Static power	Dynamic power	Propagation delay	Area
Dual sleep	+83.41%	+65.49%	-238.08%	-20.31%

Table-2: Comparison of stacked sleep approach in case of a SRAM cell.

Method	Static power	Dynamic power	Propagation delay	Area
Dual sleep	-49.59%	+97.11%	-112.6%	+1%

Table-3: Comparison of variable body biasing in case of chain of four inverters

Method	Static power	Dynamic power	Propagation delay	Area
Dual sleep	+69.29%	+50.44%	-102.49%	-20.31%

Table-4: Comparison of variable body biasing in case of a SRAM cell.

Method	Static power	Dynamic power	Propagation delay	Area
Dual sleep	+20.69%	+95.33%	-102.9%	+1%

Meanwhile, in case of SRAM cell variable body biasing shows 20.69%, 95.33%, 1% better performance and 102.49% degraded performance than the dual sleep method in terms of static power, dynamic power, area and propagation delay respectively. For this, stacked sleep and variable body shows best performance in power reduction and area while some delay penalty.

V. CONCLUSIONS

CMOS technology in nanometer scale faces great challenge due to sub-threshold leakage power consumption. The earlier approaches and our proposed methods can be effective in some ways but no previous approaches could find the exact solution of power dissipation. So depending on the designer, they choose different techniques based on the chosen technology. In this paper we actually proposed two different novel approaches for leakage reduction as well as area and delay concern. We compared our proposed methods with the previous approaches. Our two proposed methods can be applied in both generic logic and memory circuit. Multi-threshold voltage can be applied in this technique. Our proposed method had excellent tradeoffs between power, delay and area than the prior approaches. Therefore, novel stacked sleep and variable body biasing techniques represent a new way in the VLSI designer's working area.

REFERENCES

[1] International Technology Roadmap for Semiconductor Industry Association, 2002.[Online] Available http://public.itrs.net

[2] J.C. Park, V.J. Mooney III and P. Pfeiffenberger, "Sleepy Stack Reduction of Leakage Power," *Proceeding of the International Workshop on Power and Timing Modeling, Optimization and Simulation*, pp. 148-158, September 2004.

[3] Sleepy Keeper: a New Approach to Low- leakage Power VLSI Design Vincent J. Mooney III, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA USA (mooney@ece.gatech.edu)

[4] N. Karmakar, M. Z. Sadi, M. K. Alam and M. S. Islam, "A novel dual sleep approach to low leakage and area efficient VLSI design" *Proc. 2009 IEEE Regional Symposium on Micro and Nano Electronics (RSM2009)*, Kota Bharu, Malaysia, August 10-12, 2009, pp. 409-414.

[5] A. P. Chandrakasan, Sheng, S., and Brodersen, R. W., "Low- Power CMOS Digital Design," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 4, pp. 473–484, April 1992.

[6] M. Johnson, D. Somasekhar, L. Y. Chiou, and K. Roy, "Leakage Control with Efficient Use of Transistor Stacks in Single Threshold CMOS," *IEEE Transactions on VLSI Systems*, vol. 10, no. 1, pp. 1–5, February 2002.

[7] Linda E. M. Brackenbury, "Design of VLSI Systems- A Practical Introduction",macmilan education limited.

[8] Avant! Corporation, http://www.avanticorp.com

[9] PTM Model [http://ptm.asu.edu/latest.html]