Leakage Power Reduction Techniques in CMOS VLSI Circuits – A Survey

[1]D.vijayalakshmi, [2]Dr P.C Kishore Raja

[1]Assistant Professor, BIT, Bangalore
[2]HOD, Department of E&C, Saveetha University

Abstract: **This paper covers the various techniques used to reduce Leakage power in CMOS circuits. The CMOS has been the leading technology in today's world of mobile communication due to its low power consumption. Reduction of leakage power in CMOS has been the research interest for the last couple of years. In CMOS integrated circuit design there is an important trade-off between technology scaling and static power consumption. In today's CMOS technology the leakage power consumption plays a significant role. As we approach Nano-scale design the total chip power consumption becomes dependent on leakage power. Increasing the battery life in mobile wireless communication and mobile computing and similar other applications is the topic of research nowadays. Further, since the leakage of battery exists even when devices are in idle state makes leakage power loss most critical in CMOS VLSI circuits. Many techniques have been evolved to tackle the problem and it is still in progress. This paper also focuses on a new technique called scan chain technique.**

Keywords: **CMOS, Leakage power, VLSI circuits, multimedia applications, Static power, Nano Scale, LSSR, DUT**

1. Introduction

The rapid growth in semiconductor technology through the use of deep-sub micron processes has led the feature sizes to be shrinking; thereby integrating extremely complex functionality on a single chip. In the ever increasing market of mobile hand-held devices used all over the world today, the battery-powered electronic system forms the backbone. To maximize the battery life, the tremendous computational capacity of portable devices such as notebook computers, personal communication devices (mobile phones, pocket PCs, PDAs), hearing aids and implantable pacemakers has to be realized with very low power requirements.

The power dissipation has become a very critical design metric due to device miniaturization and rapid growth towards wireless communication. The longer the battery lasts; the better is the device . The power dissipation has not diminished even with the scaling down of the supply voltage. The problem of heat removal and power dissipation is getting worse as the magnitude of power per unit area has kept growing. There is a little help from advanced cooling and packaging strategies with the rapid increase in power consumption of present day chips. Also, the cost associated with the packaging and the cooling of such devices is becoming prohibitive. In addition to cost, the issue of reliability is a major concern. It is already reported that Component failure rate roughly doubles for every 10°C increase in operating temperature. Following Moore's law, with the on-chip devices doubling every two years, minimizing the power consumption has become currently an extremely challenging area of research. Leakage power of a CMOS transistor depends on gate length and oxide layer thickness. To decrease the dynamic power, the supply voltage is decreased which leads to the performance degradation. To increase the operating speed the device, the threshold voltage should also be scaled down along with the supply voltage, which results in exponential increase in the sub-threshold leakage current, thereby increase in the static power dissipation. The main components of leakage current in a MOS transistor are shown in the figure given below.(Figure1)

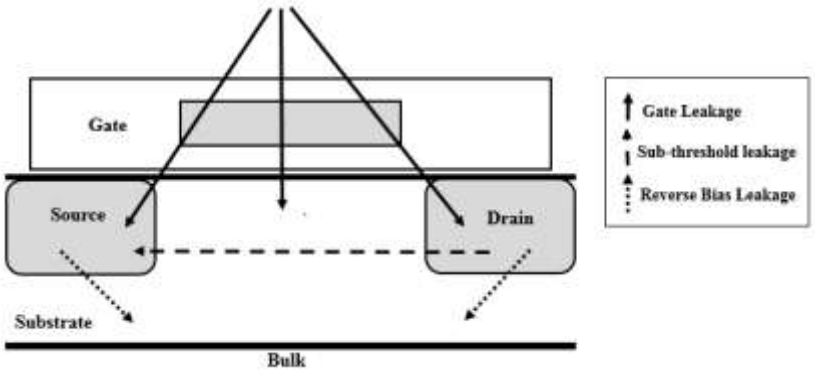

Figure 1

With the increase in the leakage current more and more, as will be seen that it becomes proportional to the total power dissipation as given by following equation.

$$Pleak = I\ leak * VDD \qquad (1)$$

Many techniques have been come into existence to overcome the leakage power problem in the nano-scale technology, but those techniques have tradeoff between area, delay and also active power. Some of those techniques are as described in this section.

2. Various Techniques to reduce Leakage power

2.1. Dual Vt and MTCMOS

This was the earliest suggested technique to reduce the leakage power. Dual VT technique is a variation in MTCMOS, in which the gates in the critical path use low-threshold transistors and high-threshold transistors for gates in non-critical path. Both the methods requires additional mask layers for each value of Vt in fabrication, which is a complicated task depositing two different oxides thickness, hence making the fabrication process complex. Moreover the techniques also suffer from turning-on latency i.e., the idle circuit cannot be used immediately after reactivated since some time is needed to return to normal operating condition. The latency is typically a few cycles for former method, and for Dual technology, is much higher. When the circuit is active, these techniques are not effective in controlling the leakage power.

2.2. Sleep Mode Approach

This method was developed to overcome the disadvantages of the dual Vt and MTCMOS technique. It

is one of the most commonly known traditional approaches for sub threshold leakage power reduction is the sleep approach. In this sleep approach, additional transistors (sleep transistors) are inserted in between the power supply and ground.

As explained in in this technique an additional "sleep" PMOS transistor is placed between VDD and the pull-up network of a circuit and an additional "sleep" NMOS transistor is placed between the pull-down network of the circuits and GND . These sleep transistors turn off the circuit by cutting off the power rails. The sleep transistors are turned on when the circuit is active and provide very low resistance in the conduction path so that circuit's performance will not get affected due to these additional transistors. During the standby mode the leakage power is reduced in the circuit by making transistors turned off which introduces large resistance in the conduction path. Thus leakage power can be reduced effectively by switching off the power source. These types of techniques are also called gated-VDD and gated- GND.

2.3. Stack Approach

The sleep technique though proved to be better than dual Vt and MTCMOS technique but however could not give a satisfying result in reducing the leakage power. This led to design a new better circuit and in this race they suggested a new technique called the stack technique which forces a stack effect by breaking down an existing transistor into two half size transistors. The circuit is as shown below (figure2).

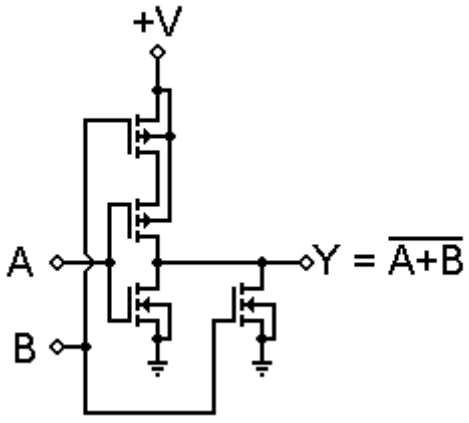

Figure 2. Stack mode approach using 2 input NOR gate

It is shown that induced reverse bias between the two transistors results when the two transistors are turned off together resulting in sub-threshold leakage current reduction. But the disadvantage is increase delay significantly between divided transistors which could limit the usefulness of the approach.

2.4. Sleepy Keeper Approach

In this Approach we can Analyze various problems faced by conventional CMOS circuit, that the basic problem with traditional CMOS is that the transistors are used only in their most efficient, and naturally inverting, i.e., PMOS transistors are connect to VDD and the NMOS transistors are connect to GND. It is a well known fact that the PMOS transistors are not efficient at passing GND and that the NMOS transistors are not efficient at passing VDD. However, to maintain a value of '1' in sleep mode, given that the '1' value has already been calculated, the sleepy keeper approach uses this output value of '1' and an NMOS transistor

connected to VDD to maintain output value equal to '1' when in sleep mode. To tackle this problem the authors suggested a new design using sleepy keeper approach in which an additional single NMOS transistor placed in parallel to the pull-up sleep transistor connects VDD to the pull-up network. When in sleep mode, this NMOS transistor is the only source of VDD to the pull-up network since the sleep transistor is off. An additional single PMOS transistor placed in parallel to the pull-down sleep transistor is the only source of GND to the pull-down network. The suggested circuit can be seen below (Figure3)

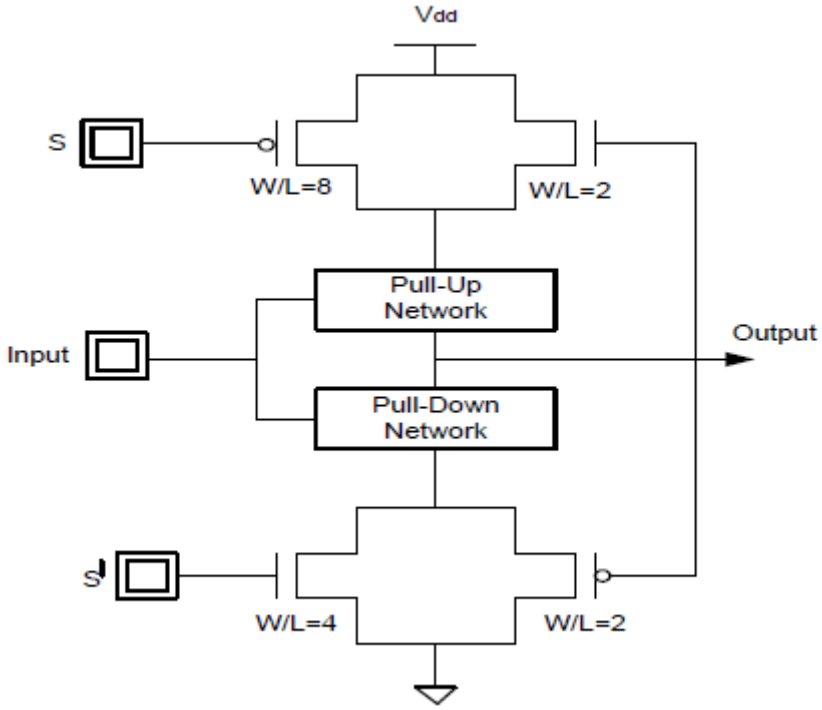

Figure 3. Sleepy Keeper circuit

However it was reported that major disadvantage faced by this technique was the reduction of power by very less percentage which was not able to fulfill the current demands of present requirement of the VLSI designed circuits.

2.5. LECTOR Technique

This is one of the low power retention techniques. This technique suggested a CMOS circuit in which two extra Leakage Control Transistors (a P-type and an N-type) is inserted within the gate, in which the gate terminal of each Leakage Control Transistor is controlled by the source of the other.

The circuit can be viewed in the figure given below.(figure 4)

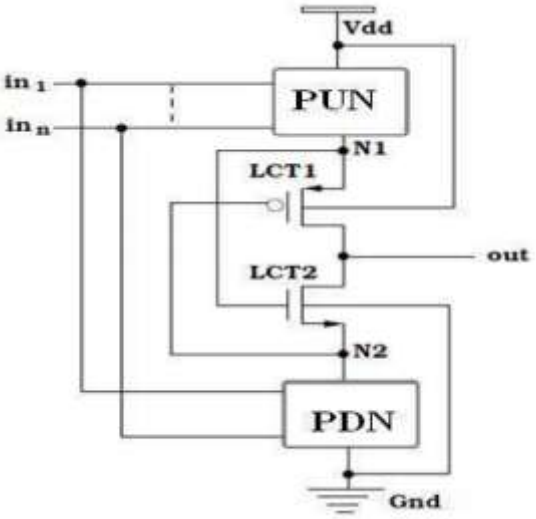

Figure 4. LECTOR circuit design

The basic idea behind this approach was for reduction of leakage power is the effective stacking of transistors in the path from supply voltage to ground. It is Observed that "a state with more than one transistor OFF in a path from supply voltage to ground

is far less leaky than a state with only one transistor OFF in any supply to ground path." In their method they introduced two leakage control transistors (LCTs) in each CMOS gate such that one of the LCTs is near its cutoff region of operation. They illustrated that their Leakage Control Transistor technique (LECTOR) with the case of a NAND gate .A CMOS NAND gate with the addition of two leakage control transistors.

3. over view

It is seen that overall leakage power is mainly due to sub-threshold leakage . In addition, gate-oxide leakage is another possible contributor to leakage power. A possible solution widely studied is the potential use of high-k (high dielectric constant) gate insulators. Again the aim of this paper is to review the reduction of the sub threshold leakage component of static power consumption by various methods implemented till date.

It was seen that there was a reduction in magnitude sub threshold leakage power with application of dual threshold voltage (Vth) techniques, along with the sleep and sleepy stack approaches. As per the reviews the major advantage of the sleepy stack approach over the sleep approaches is that the sleepy stack approach saves exact logic state. But the sleepy stack approach comes with a demerit that each transistor in the original, base case, traditional CMOS design results in three transistors in the sleepy stack equivalent. The goal of this paper is to give the review of the latest approaches to achieve large reduction of leakage power in CMOS circuits. The new approaches are called LSSR (Lector Stack State Retention Technique) and Scan chain Technique.

4. LSSR (Lector Stack State Retention Technique)

In order to achieve low power LSSR is formed by combining two previously done approaches namely LECTOR approach and Forced stack approach. Since it combines the two above mentioned techniques it has the features of both the approaches and thus is much beneficial than the previous works done.

The circuit is proposed by introducing two gated leakage transistors between pull up and pull down networks with high threshold voltage, and then stack effect is added to pull up and pull down networks by dividing each transistor in to half size transistors.

As per the works being in progress it is believed that this new technique LSSR can be proved to be much better than the earlier works done. LSSR which can achieve better leakage reduction by maintaining exact logic state(state retention) than the other techniques.

5. SCAN CHAIN TECHNIQUE

Scan chain is a technique used in design for testing. The objective is to make testing easier by providing a simple way to set and observe every flip-flop in an IC. The basic structure of scan includes the following set of signals in order to control and observe the scan mechanism.

1.Scan_in and scan_out define the input and output of a scan chain. In a full scan mode usually each input drives only one chain and scan out observe one as well.

2. A scan enable pin is a special signal that is added to a design. When this signal is asserted, every flip-flop in the design is connected into a long shift register.

3. Clock signal which is used for controlling all the FFs in the chain during shift phase and the capture phase. An arbitrary pattern can be entered into the chain of flip-flops, and the state of every flip-flop can be read out

The main of this paper was to give a review of the various steps taken towards the reduction of the leakage power for VLSI designs. A major thrust towards the low power design of CMOS is actually due to recent technological advances in wireless communication because the usable time of a mobile device is heavily restricted by its battery life. With the growing complexity of mobile devices, such as with a digital camera, multimedia services, Video Conferencing, global positioning system (GPS) etc are the features which make the battery power problem more challenging. The leakage current of a logic gate is a strong function of its input values. The reason is that the input values affect the number of OFF transistors in the NMOS and PMOS networks of a logic gate. For example, the minimum leakage current of a two-input NAND gate corresponds to the case when both its inputs are zero. In this case, both NMOS transistors in the NMOS network are off, while both PMOS transistors are on. The effective resistance between the supply and the ground is the resistance of two OFF NMOS transistors in series. This is the maximum possible resistance. If one of the inputs is zero and the other is one, the effective resistance will be the same as the resistance of one OFF NMOS transistor. This is clearly smaller than the previous case. If both inputs are one, both NMOS transistors will be on. On the other hand, the PMOS transistors will be off. The effective resistance in this case is the resistance of two OFF PMOS transistors in parallel. Clearly, this resistance is smaller than the other cases. There is also the "stack effect" i.e., the phenomenon whereby the leakage current through a stack of two OFF transistors of W/L ratios each is lower than that of a single OFF transistor with a W/2L ratio. This is mainly because of the body effect, which causes an increase in the effective resistance of the two transistor chain compared to that of a single transistor. They formulate the problem of finding the MLV using a series of Boolean Satisfiable problems. Using this vector to drive the circuit while in the STANDBY state, they reduce the circuit leakage by as much as 35%. Having found the minimum leakage pattern, one can use this vector to drive the circuit while in the sleep mode.

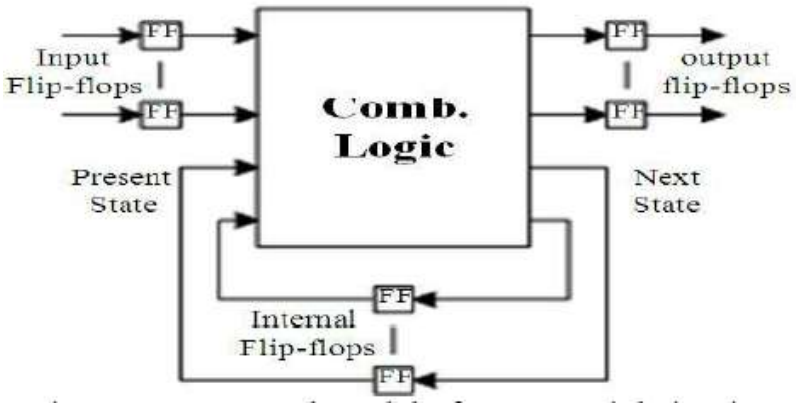

Fig. 4. A general model of a sequential circuit

In Figure 4, we consider a sequential circuit comprised of a combinational circuit and a set of flip-flops. In the scan-based designs, the flip-flops are connected in such a way that they enable two modes of operation: Normal mode and test mode. In the normal mode, the flip-flops are connected as shown in Figure 2. At each clock cycle, the next state is stored in the flip-flops. In the test mode, the flip-flops are reconfigured and form one or more shift registers, called scan registers or scan chains. At each clock cycle the values of the flip-flops are shifted. The values can be observed through the output of the last flip-flop of the scan chain. Furthermore, the values can be shifted into the scan-chain through the input of the first flip-flop in the chain. In this paper, we assume that all internal and external (input and output) flip-flops are included in the scan chain. This type of circuit is called full-scan. Full scan chains convert the problem of testing a sequential circuit to that of a combinational one. In other words, the input and internal flip- flops can be treated as primary inputs of the circuit, whereas the output and internal flip-flops are considered as the primary outputs. In order to test a circuit, the circuit is first switched to the test mode and the present state value is shifted into the flip- flops. After that the circuit is switched to the normal mode and operates for one or more cycles under the externally provided input values. In the next step, the circuit is switched back to the test mode and the next state value is shifted out. Different patterns are widely used to efficiently test the logic of DUT's. While additional functional tests might be necessary to fill some test gaps, a well prepared scan test allows detecting of avery high percentage of manufacturing failures, requiring a drastically smaller amount of testdata and test time compared to functional tests.

To enable a scan test for a chip design, additional test logic must be inserted; this is called "scan insertion". Scan insertion consists of two steps:

1. Replace plain memory cells like flip flops or latches by scan cells.

2. Connect these together forming one or more chains.
Scan cells can be operated in two modes, the functional/mission mode used during normal operation and the scan mode that allows shifting through the scan chains.

In VLSI circuit design, scan chains are introduced to improve the testability of integrated circuits [14]. After logic synthesis, all flip-flops in the circuits are replaced with scan flip flops. These scan flip-flops are connected sequentially to form a scan chain(or multiple scan chains) in a single chip. Each scan flip-flop in the scan chain has two input sources: the output of the previous flip-flop in the scan chain and the output of the combinational circuits. During normal operation, the response at the state outputs is captured in the flip-flop. In testing mode, test vectors are shifted into the registers through the primary input pads and the test output values are shifted out through the primary output pads.

To solve the problem faced various works have been implemented and still technicians are working on this field. However through this paper we get to know the advantages and disadvantages of various works done .We conclude that all the above mentioned circuit may lead to much large reduction of leakage power than the general stack and sleep approaches. Finally it is concluded that the optimized layout will also play an important role in reducing the leakages.

Acknowledgment
I kindly acknowledge all the authors for their Reference work without whom this paper is not possible.

References

[1] Moore, G. E. No exponential is forever: but "Forever" can be delayed! In: IEEE int. Conf. Solid state circuits, 2003. Proceedings… IEEE, 2003, p. 20-23.
[2] Sheu, B. J. et al. BSIM: Berkeley Short-Channel IGFET Model for MOS Transistors. IEEE Journal of Solid State Circuits, New York, v.SC-22, n.4, p. 558-566, Aug. 1987.
[3] Guindi, R. S.; Najm, F. N. Design techniques for gate-leakage reduction in

CMOS circuits. In: Int. Symp. Quality electronic design, 2003. Proceedings…, IEEE, 2003, p.61-65.
[4] Mukhopadhyay, S. et. al. Gate Leakage Reduction for Scaled Device Using
Transistor Stacking. IEEE Trans. on VLSI Systems, New York, v.11, n.4, p. 716-730, Aug 2003.
[5] Narendra, S. G.; Chandrasekhar, A. Leakage in Nanometer CMOS Technologies. New York: Springer, 2006. Pp. 307-310.
[6] Mukhopadhyay, S. et al. Accurate Estimation of Total Leakage in Nanometer-Scale Bulk CMOS Circuits Based on Device Geometry and Doping Profile. IEEE Trans. On CAD of IC and Systems, New York, v.24, n.3, p. 363-381, Mar. 2005.
[7] KIM, C. H.; ROY, K. Dynamic Vth Scaling Scheme for Active Leakage Power Reduction. In: Design Automation And Test In Europe Conference, 2002. Proceedings… IEEE, 2002, p. 163-167.
[8] Ferre, A. and Figueras, J., "Characterization of Leakage Power in CMOS Technologies", IEEE International Conference on Electronics, Circuits and Systems, Vol. 2, 1998, pp. 85 -188.
[9] Abdollahi, A.; Fallah, F.; Pedram, M., "Runtime mechanisms for leakage current reduction in CMOS VLSI circuits" Low Power Electronics and Design, 2002. ISLPED '02. Proceedings of the 2002 International Symposium on, 2002, Page(s): 213 -218.
[10] K. Roy et al., Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits, Proceedings of the IEEE, 2003, pp. 305-327.
[11] M.C. Johnson et al., Models and algorithms for bounds on leakage in CMOS circuits, IEEE TCAD, 1999, pp. 714-725.
[12] Swarup Bhunia, Hamid Mahmoodi, Debjyoti Ghosh, and Kaushik Roy, "Power Reduction in Test-Per-Scan BIST with Supply Gating and Efficient Scan Partitioning," Proceedings of the Sixth International Symposium on Quality Electronic Design (ISQED'05), pp. 3636-3642.
[13] H. Mizuno, K. Ishibashi, T. Shimura, T. Hattori, S. Narita, K. Shiozawa, S. Ikeda, K. Uchiyama, "A 18uAStandby-Current 1.8V 200MHz Microprocessor with Self Substrate-Biased Data-Retention Mode," Proc. IEEE International Solid-State Circuits Conference, pp.280-281, 1999.
[14] AGARWAL, A. et al. Leakage Power Analysis and Reduction for Nanoscale Circuits, IEEE Micro, Los Alamitos, v.26, n.2, p 68-80, Mar. 2006
[15] S. Narendra, S. Borkar, V. De, D. Antoniadis, and A. Chandrakasan, "Scaling of stack effect and its application for leakage reduction," Intl. Symp. Low Power Electronic and Design, 2001, pp. 195-200.
[16] Wei Wang et al. "Leakage current optimization techniques during test based on Don't Care Bit Assignment," Journal of computer science and technology, Sept 2007, pp. 673-680.
[17] K. Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage ReductionTechniques in Deep-Submicrometer CMOS Circuits", In Proc. IEEE, vol. 91, pp. 305-327, Feb., 2003.
[18] M. Johnson, D. Somasekhar, L.-Y.Chiou, and K. Roy, "Leakage control with efficient use of transistor stacks in single threshold CMOS," IEEE Trans. VLSI Systems., vol. 10, no. 1, pp. 1-5, Feb. 2002.
[19] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, "1-V power supply high-speed digital circuit technology with multi-threshold voltage CMOS," IEEE J. Solid-State Circuits, vol. 30, pp. 847-854, Aug. 1995.
[20] Steven Keeping, "Design techniques for extending Li-ion battery life", November 19, 2013.
[21] Ken Bigelow, "Inside Computer Logic Gates".
[22] Se Hun Kim and Vincent J. Mooney III , "Sleepy Keeper : a New Approach to Low-Leakage Power VLSI Design", in VLSI SOC conference 2006, PP. 367-372.
[23] N. Hanchate and N. Ranganathan," Lector: A technique for leakage reduction in CMOS circuits", IEEE Transactions on Very Large Scale Integration (VLSI) Systems , vol. 12, no. 2, pp. 196-205, February 2004.