LEAKAGE CURRENT REDUCTION USING INPUT VECTOR CONTROL APPROACH

A DISSERTATION

Submitted in partial fulfillment of the requirements for the award of the degree

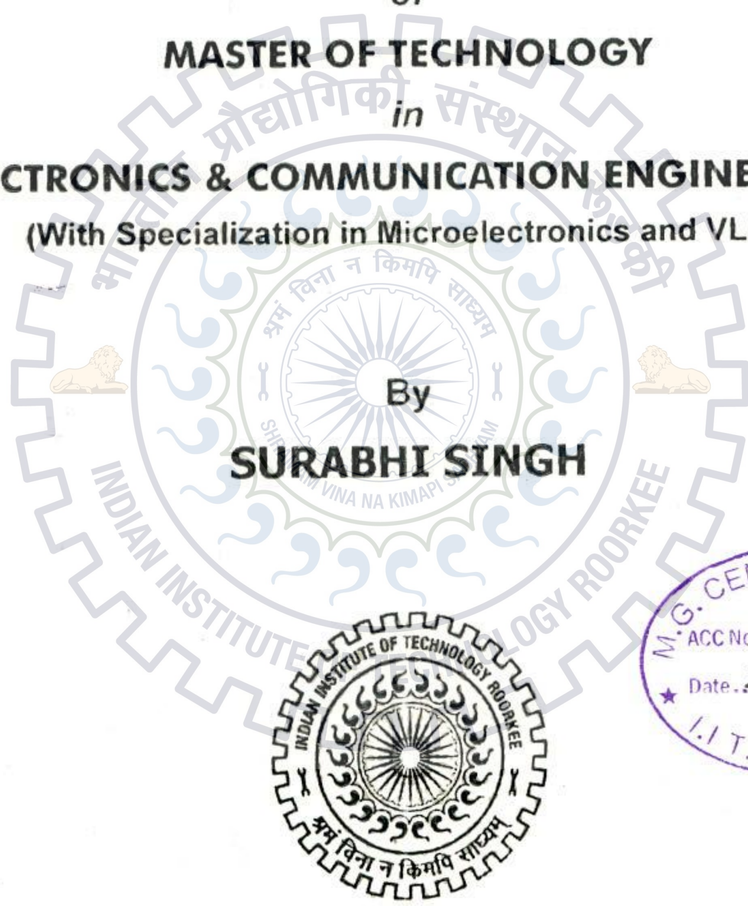
of

MASTER OF TECHNOLOGY

in

ELECTRONICS & COMMUNICATION ENGINEERING

(With Specialization in Microelectronics and VLSI)



By

SURABHI SINGH

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE - 247 667 (INDIA)

JUNE, 2013

CANDIDATE'S DECLARATION

I hereby declare that the work, which is presented in this dissertation, titled **"Leakage Current Reduction using Input Vector Control Approach"**, being submitted in partial fulfillment of the requirements of the award of the degree of **Master of Technology** with specialization in **Microelectronics and VLSI**, in the Department of Electronics and Communication Engineering, Indian Institute of Technology, Roorkee is an authentic record of my own work carried out from July 2012 to June 2013, under the guidance and supervision of **Dr. S. Dasgupta and Dr. B. K. Kaushik**, Department of Electronics and Communication Engineering, Indian Institute of Technology, Roorkee.

The results embodied in this report have not been submitted for the award of any other degree or diploma.

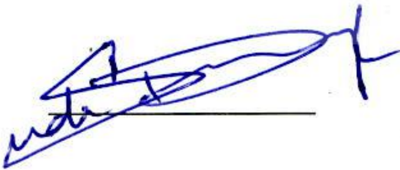
Date: 17-06-13

Place: Roorkee

Surabhi Singh

CERTIFICATE

This is certified that the statement made by the candidate is correct to the best of our knowledge and belief.



(Dr. S. Dasgupta)

E&C Department

Indian Institute of Technology, Roorkee

Roorkee-247667, Uttarakhand (India)

(Dr. B. K. Kaushik)

E&C Department

Indian Institute of Technology, Roorkee

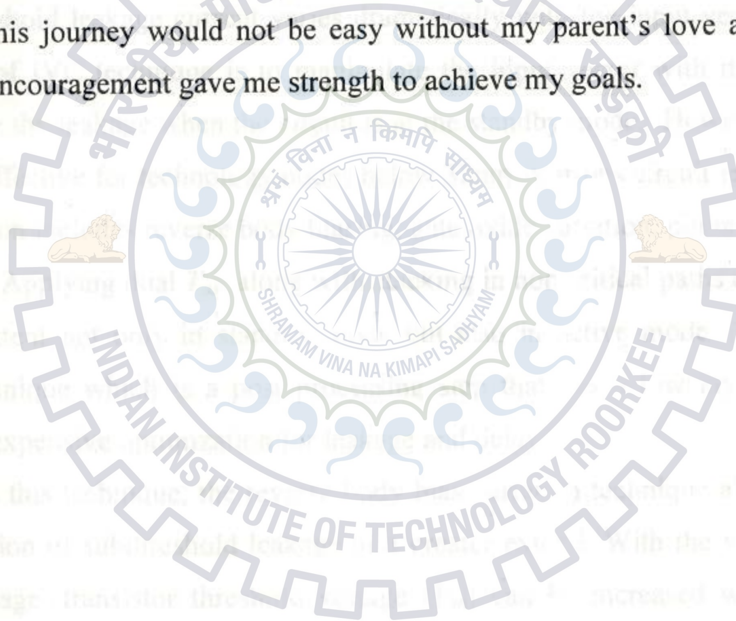
Roorkee-247667, Uttarakhand (India)

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my supervisors **Dr. S. Dasgupta** and **Dr. B. K. Kaushik**, Department of Electronics and Communication Engineering, IIT Roorkee, for allowing me to carry out this project work and providing me with valuable suggestions and support throughout my project. With consistent discussions and encouragement, right from early stage of my work helped me to understand the different aspects of research study. I would also like to thank **Dr. Anand Bulusu** sir for his valuable suggestions which helped me to improve my work even more.

I want to thank all my Ph.D. seniors, especially Baljit Kaur and Manoj Kr. Majumder for always helping me and encouraging me throughout my work. I also like to thank all my classmates and friends for making my M. Tech journey memorable. Besides them, I want to mention that this journey would not be easy without my parent's love and support. Their blessings and encouragement gave me strength to achieve my goals.



Surabhi Singh
June 2013
Roorkee, india

ABSTRACT

To achieve higher packing density and improved performance, CMOS devices have been continuously scaled down from last few decades. The reduction in feature size makes power consumption to be a critical issue in nanoscale CMOS VLSI circuits. In order to maintain the power consumption under control, supply voltage (V_{dd}) and threshold voltage (V_{th}) has to be commensurately scaled to achieve an improved performance. However, the reduction in V_{th} results in substantial increment in leakage currents which degrades device performance in nanoscale regime.

Input vector control (IVC) is a popular technique for leakage power reduction. It utilizes the transistor stack effect in CMOS gates by applying a minimum leakage vector (MLV) to the primary inputs of combinational circuits during the standby mode. CMOS gate's sub threshold leakage current varies dramatically with the input vector applied to the gate, the idea of IVC technique is to manipulate the input vector with the help of a sleep signal to reduce the leakage when the circuit is at the standby mode. However, this technique becomes less effective for technology nodes below 32nm. Various circuit level techniques for leakage reduction includes reverse body biasing, gate oxide variation, pin reordering and logic reconstruction. Applying dual T_{ox} along with stacking in non critical paths can reduce leakage to a greater extent not only in standby mode but also in active mode. We can apply pin reordering technique which is a post processing step that has a low layout impact and is therefore an inexpensive optimization for leakage and delay.

Besides this technique, the reverse body bias variation technique along with stacking helps in reduction of subthreshold leakage to a greater extend. With the variation in reverse body bias voltage, transistor threshold voltage (V_{th}) can be increased which substantially reduces subthreshold leakage current. On the other hand, gate oxide thickness (T_{ox}) can be increased to modify threshold voltage and reduce subthreshold and gate tunneling leakage in a CMOS VLSI circuit. These approaches should be carefully used as the increase in T_{ox} and V_{th}, the device suffers from adverse short channel effects. Along with these techniques, pin reordering and logic reconstruction can be used which are low cost and powerful techniques for leakage and delay reduction.

TABLE OF CONTENTS

CANDIDATE'S DECLARATION i

ACKNOWLEDGEMENT ii

ABSTRACT iii

TABLE OF CONTENTS iv

LIST OF FIGURES vi

LIST OF TABLES vii

1. INTRODUCTION 1

1.1 Leakage current mechanisms 2

1.1.1 Gate tunneling current 3

1.1.2 Subthreshold current 4

1.1.3 BTBT current 5

1.2 Problem statement 6

1.3 Organization of thesis 8

2. LEAKAGE CURRENT REDUCTION TECHNIQUES 9

2.1 Leakage reduction using stacking effect of transistor 9

2.2 Leakage reduction using pin reordering and logic reconstruction 10

2.3 Leakage reduction using Dual-T_{ox} technique 12

2.4 Leakage reduction using RBB voltage 13

3. BASIC GATE REPLACEMENT TECHNIQUE 14

3.1 Basic gate replacement algorithm 15

4. LEAKAGE REDUCTION MECHANISMS BASED ON STACKING EFFECT 18

4.1 Leakage Reduction using Modified Gate Replacement Technique for CMOS VLSI circuit 18

4.1.1 Modified gate replacement technique 18

4.1.2 Results and discussion 20

4.2 Modified Gate Replacement Algorithm for Leakage Reduction using Dual-T_{ox} in CMOS VLSI circuits 22

4.2.1 Results and discussion 24

4.2.1.1 Analysis of leakage current for different oxide thickness 24

4.2.1.2 Analysis of leakage current for gate replacement using Dual-T_{ox} transistors 26

4.2.1.3 Analysis of leakage current for benchmark circuits 27

4.3 Modified Gate Replacement Algorithm for Leakage Reduction using RBB variation in CMOS VLSI circuits 27

4.3.1 Modified gate replacement technique using RBB variation 28

4.3.2 Results and discussions 30

4.3.2.1 Effect of RBB on leakage currents 30

4.3.2.2 Effect of stacking on leakage current 31

4.3.2.3 Effect of leakage on benchmark circuits 32

CONCLUSION AND FUTURE SCOPE 34

References 36

Publication list 39

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
SHRAMAM VINA NA KIMAPI SADHYAM

LIST OF FIGURES

Figure 1.1 Different leakage current mechanisms 3

Figure 1.2 Mechanisms of gate tunneling leakage 4

Figure 1.3 BTBT leakage current in reverse biased junction 6

Figure 2.1 Stacking effect of at NAND gate having two inputs 9

Figure 2.2 Effect of stacking on leakage current 10

Figure 2.3 (a) Biasing conditions for a NMOS transistor and (b) Significant contribution states 12

Figure 2.4 Influence of ordering of transistors on delay 12

Figure 3.1 Gate replacement and its effect on fanout of gates 15

Figure 3.2 Example for gate replacement. (a) C17 with total leakage 831.08 nA (b) C17 after gate replacement with total leakage 476.88 nA 16

Figure 3.3 Pseudo code for basic gate replacement technique 17

Figure 4.1 Pseudo code for the modified gate replacement algorithm 20

Figure 4.2 (a) Minimum leakage current ($I_{leakage}$) and (b) number of replaced gates for different benchmark circuits using conventional and modified gate replacement technique 21

Figure 4.3 Pseudo code of the modified gate replacement algorithm 23

Figure 4.4 Variation of propagation delay and leakage current with (a) oxide thickness and (b) width of PMOS transistor 25

Figure 4.5 Leakage current reduction for different Benchmark Circuits 27

Figure 4.6 Pseudo code of the modified gate replacement technique using reverse body bias voltage variation 30

Figure 4.7 Variation of (a) subthreshold, gate and BTBT leakage current and (b) total leakage current for different reverse body bias voltages using a CMOS transistor at 45nm technology node 31

Figure 4.8 Leakage current for different benchmark circuits by RBB variation and stacking effect using RBB variation 33

LIST OF TABLES

Table 4.1 Percentage reduction in leakage using modified gate replacement technique *w.r.t.* Conventional one 22

Table 4.2 Variation of leakage current and propagation delay for different T_{ox} of a PMOS transistor 25

Table 4.3 Variation in leakage current and propagation delay for different widths of PMOS transistor at T_{ox} = 2nm 26

Table 4.4 Leakage current for different input states for a two input NAND gate using T_{ox} variation and stacking effect along with dual-T_{ox} transistors 26

Table 4.5 Leakage reduction using modified replacement technique for benchmark circuits 27

Table 4.6 Variation of different leakage currents with reverse body bias voltage at 45nm technology node for an NMOS transistor 31

Table 4.7 Leakage current for 2-input NAND Gate at 45nm technology node 32

Table 4.8 Leakage current for 3-input NAND Gate in Standby and Active mode at 45nm technology node 32

Table 4.9 Leakage Current for Different Benchmark Circuits by Reverse Body Bias (RBB) and Stacking Effect using RBB variation 33

Chapter 1
INTRODUCTION

From the last few decades, CMOS transistors have been continuously scaled for achieving higher packing density along with improvement in the device performance. The shrinking of feature size results in increased power consumption which should be considered as a critical aspect in a nanoscale CMOS VLSI circuits. In case to maintain the power consumption within limits, supply voltage (V_{dd}) of CMOS transistor must also be scaled. Moreover, for maintaining gate drive current threshold voltage (V_{th}) has to be scaled to achieve an improvement in performance [1]. However, reduction in V_{th} results in increase of leakage currents of CMOS VLSI circuits. Furthermore, to avoid the adverse effect of short channel in MOS transistor, it is desirable to reduce oxide thickness (T_{ox}) in same proportion as that of the channel length. The reduction in V_{th} and T_{ox} further results in higher subthreshold and gate tunneling current. Apart from this, the use of the higher value of the substrate doping density in scaled devices causes flow of leakage currents from drain region and source region to body terminals of a MOS transistor. This leakage current usually referred as band to band tunneling leakage current (BTBT). Along with these three, the other leakage mechanisms which mainly occur in small geometries are gate induced drain (GIDL) and punch through leakage currents. The GIDL occurs due to higher electric fields in drain and substrate of MOS transistor which is biased in accumulation region. On the other hand, punch through occurs due to extending of drain region to substrate and source region to substrate in the channel [1]. These two components can be neglected when the device is operating in normal modes.

The leakage currents occurring in nanometer CMOS VLSI circuits is basically due to gate, subthreshold and reverse biased *pn* junction leakage which degrades device performance. The reduction of these leakage currents at various technology nodes is the primary concern in current technology nodes. There are different dominating leakages at different technology nodes. For the technology nodes up to 65nm, the subthreshold current dominates the gate and BTBT leakage .This leakage can be effectively reduced by the stacking effect of transistors [2-6]. As we move further to lower technology nodes up to 32nm, the gate leakage becomes the major contributor after the subthreshold leakage. Although BTBT leakage do not contributes significantly to the overall leakage current but it

degrades device performance. This work presents various algorithms to reduce these leakages at different technology nodes which are primarily based on the gate replacement algorithm [7]. The gate replacement algorithm uses the stacking effect in the transistors to reduce subthreshold leakage by increasing number the transistors which are in 'off' state in a transistor stack. To achieve this, the gate which produces higher leakage is replaced by another gate incorporating an extra *SLEEP* signal. This replaced gate is the gate with same functionality as that of previous gate in active mode while possessing a lower leakage in standby mode of operation [7-11]. Along with stacking effect, various other circuit level techniques can be employed for further leakage reduction. Various circuit level techniques for leakage reduction includes reverse body biasing, gate oxide variation, pin reordering and logic reconstruction [12-20]. With the variation in gate oxide thickness (T_{ox}), threshold voltage (V_{th}) can be modified which results in reduction of the subthreshold leakage and gate tunneling leakage in a CMOS VLSI circuit [12, 16]. On the other hand, reverse body bias voltage variation substantially reduces subthreshold leakage current while increasing BTBT leakage in nanometer regime [17, 18]. These approaches should be carefully used as the increase in T_{ox} and V_{th}, the device suffers from adverse short channel effects. Along with these techniques, pin reordering and logic reconstruction [19-22] can be used which are low cost and powerful techniques for leakage and delay reduction. These approaches are applied to various benchmark circuits using HSPICE simulator for analyzing leakage current and delay in nanometer regime [23-25].

1.1 Leakage current mechanisms

In recent nanoscale VLSI circuits, one can observe three main sources of leakage currents referred as gate, subthreshold, and reverse biased *pn* junction leakage [1] as presented in Fig. 1.1. As we know, the current flow in the channel region of the MOS transistor is due to the formation of inversion layer. In case, the gate bias voltage is not sufficient to invert the channel, there are still carriers present in channel which suffers potential barrier. Due to this potential barrier, no current flows in the channel. This state is called weak inversion state. The subthreshold current streams from the drain region to source region when the transistor is operating in weak inversion. Conversely, gate leakage arises due to the continuous scaling of thickness of oxide. It results in a high field region between gate and channel. Using this phenomenon, electrons/holes tunnel from substrate to gate through the potential barrier across oxide which is known to be gate oxide tunneling current. Apart

from this, reverse biased *pn* junction leakage is the adverse effect of reverse biased voltage applied in MOS devices. For nanoscale VLSI circuits, the gate and subthreshold leakages are primarily dominated because of the reduced channel length and gate oxide thickness [1, 3].

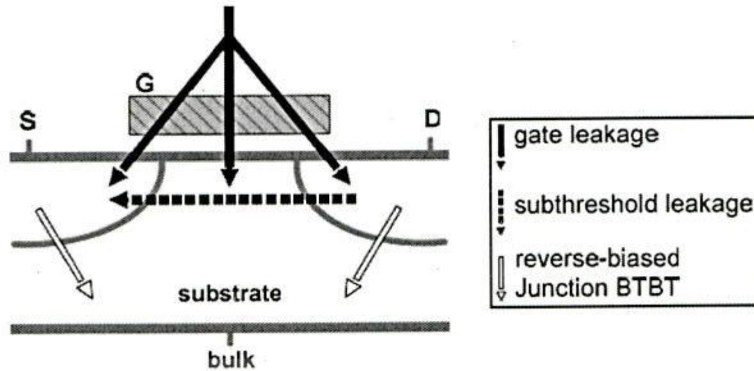


Fig. 1.1 Different leakage current mechanisms

1.1.1 Gate Tunneling Current

The aggressive scaling of device in nanometer regime results in increasing short channel effects that degrades the device performance in nanoscale technology nodes. To control short channel effect, scaling of oxide thickness becomes necessary. Continuous scaling in oxide thickness results in higher electric fields which in turn lead to current due to direct tunneling through transistor gate oxide. Gate oxide tunneling arises because of the tunneling of electrons from source, drain and channel region into the gate of the transistor through gate oxide potential barrier. This leakage phenomenon primarily based on three mechanisms which are referred as electron conduction-band tunneling (ECB), electron valence-band tunneling (EVB) and holes valence-band tunneling (HBV) as shown in Fig. 1.2. In ECB, tunneling of electrons takes place from conduction band of substrate to the gate conduction band or vice versa. In case of second mechanism, electrons tunnels from the valence band of substrate to the gate conduction band. In hole valence band tunneling, holes tunnels from valence band of substrate to the valence band of gate [1-3].

With the increase in thickness of gate oxide, the exponential drop of tunneling current can be formulated as [1],

$$I_g=(A.B).(W.L)e^{-C\frac{T_{ox}}{V_{gs}}.\alpha} \tag{1.1}$$

where $A = q_3/8\pi h\varphi_b$, $B=(V_{gs}/T_{ox})^2$, $C=8\pi\frac{\sqrt{2m_{ox}}\varphi_b^{3/2}}{3hq}$, α is considered as constant which lies between 1 to 0.1 and mainly depends on the oxide potential drop. The Planck's constant and barrier height of electrons/holes in the conduction/valence band are$h$ and φ_b respectively.

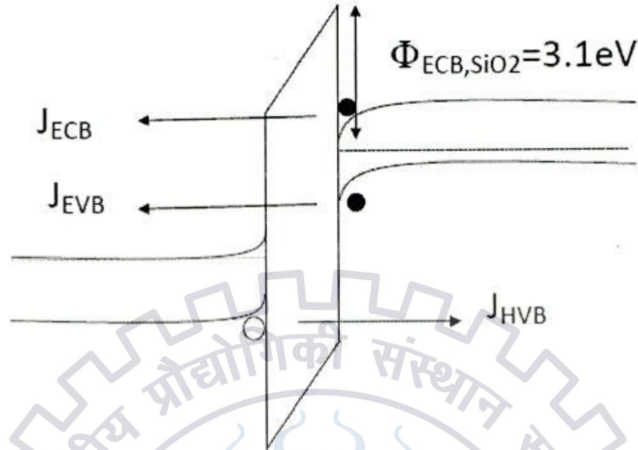


Fig. 1.2 Mechanisms of gate tunneling leakage

1.1.2 Subthreshold Current

With the continuous technology scaling, supply voltage also scaled to lower the dynamic power consumption and preserve the device reliability. Along with supply voltage, threshold voltage scaling becomes necessary for reasonable gate over-drive. Due to the aggressive scaling of V_{th}, an exponential increment in subthreshold leakage is observed in nanometer CMOS VLSI circuits. As we know, the current flow in the channel region is basically due to formation of inversion layer. If gate bias voltage is not sufficient enough to invert the channel, there are still few carriers present. These carriers do not contribute in current flow due to potential barrier present in the channel region. This region is called weak inversion state. In this state, current flow is basically dominated by diffusion of carriers and results in subthreshold current [1-3].

The current which streams from the source to drain region is due to the drift and diffusion of carriers. In strong inversion state, the current flow is dominated by drift current whereas in the region of weak inversion, current flow occurs due to the diffusion of carriers. In weak inversion, the minority carrier concentration is quiet low. Also, channel suffers from small horizontal field ($V_{gs}<V_{th}$) but a small longitudinal field still appears due to the drain to source voltage (V_{ds}). Even if the value of the transistor's gate terminal voltage reduces below

V_{th}, small leakage current still flows due to process of diffusion of the minority carriers between the source region and drain region of the MOS transistor as presented in Fig. 1.1. Thus, the subthreshold leakage current can be expressed as [1],

$$I_{sb} = I_0 . e^{\frac{V_{gs}-V_{th}}{\frac{\eta kT}{q}}} \left(1-e^{-\frac{V_{ds}}{\frac{kT}{q}}}\right) \tag{1.2}$$

where $I_0 = \mu_0 C_{ox} (\frac{W}{L}) (\frac{kT}{q})^2 (1 - e^{1.8})$, L and W are consider to be length and width of a MOS transistor, μ_0 *is* the mobility in case of low electric field, C_{ox} is the capacitance between the gate and oxide region, k and q are termed as Boltzmann's constant and charge of electron. Beside this, swing factor in the sub threshold region is denoted as N.

1.1.3 BTBT leakage

The MOS transistor consists of two *pn* junctions, *i.e.*, drain/source region to substrate region. The reverse voltage applied across these junctions result in flow of *pn* junction leakage current. This leakage is mainly a strongly depends on area across junction and the concentration of doping in *pn* regions. If the *p* and *n* regions of a MOS transistor are heavily doped, the *pn* junction leakage is primarily due to tunneling from band to band region (BTBT).

The BTBT resulting as effect of high electric which occurs between reverse biased drain/source to substrate regions in a MOS transistor. This high electric field causes current streams through junction due to electron tunneling from the valence band to conduction band as shown in Fig. 1.3. Tunneling occurs when the sum of applied reverse biasing voltage (V_{app}), build-in voltage (V_{bi}), and total potential drop in junctions is larger than the band gap as presented in Fig. 1.3. The density of tunneling current across the *pn* junction is given by [1-3],

$$J_{BTBT} = A \frac{E V_{app}}{E_g^{1/2}} exp\left(-B \frac{E_g^{3/2}}{E}\right) \tag{1.3}$$

where $A = \sqrt{2m^* q^3}/4\pi^3 h^2$, $B = 4\sqrt{2m^*}/3hq$. Here, effective mass of electron and the field across junction are m^* *and* E_g respectively. Charge of electron and $1/2\pi$ of the Plank's constant are denoted as q and h.

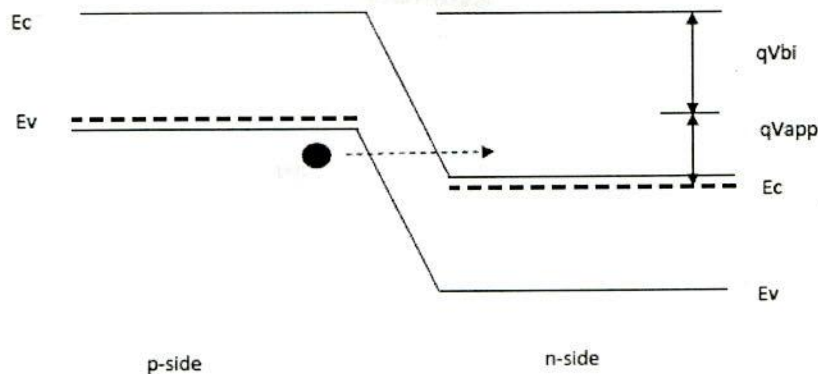


Fig. 1.3 BTBT leakage current in reverse biased junction

1.2 Problem statement

From the past few decades, CMOS devices continuously scaled to attain high density and lower consumption in power. Consequently, leakage current that becomes major contributor to the total dissipated power in CMOS VLSI circuits. The major contributor of leakage currents in nanoscale devices are basically subthreshold, gate tunneling and BTBT leakage current. Hence, reduction of these leakages becomes an important issue, especially, in low power VLSI circuits. Among the existing circuit level approach, stacking effect of transistor becomes most efficient way of reducing the subthreshold leakage. Gate replacement approach is primarily based on stacking effect which replaces gates at worst leakage state (WLS) in a topological order with another gates incorporating an extra sleep signal [7]. The replacement by another gate maintains the circuit functionality in active mode while having lower leakage in standby mode. The main issue with this approach is that if total gates which are at WLS are higher, then number of replaced gates will increase which increase circuit area and delay. Also this approach becomes less effective as we move down to lower technology nodes where gate and BTBT leakages become considerable.

The conventional gate replacement approach replaces each gate at WLS in topological order with another gate having same functionality in active mode while reduced leakage in standby mode [7]. The negative aspect of this approach is that on replacement, output of that gate may change. In this case if fanout of replaced gate is higher, the leakages of all fanout gates will be affected. This can lead to increase in overall circuit leakage. On further replacement of fan out gates, number of replacements increases. To reduce the number of replacements, the modified gate replacement technique is proposed which skip the gates with

higher fanout. It increases the overall circuit leakage and area. These skipped gates are further considered for replacement followed by an AND/OR gate. The use of AND/OR gate is to ensure that on replacement output of gate do not affect fanout gates. This replacement is made permanent only if the overall leakage of the circuit reduces. In case, if circuit delay also considered as a critical factor, the replacement should be done on non critical paths only. This approach proves to be effective up to 65nm technology nodes. However, at lower technology nodes, the gate leakage also comes into consideration below 65nm. The gate replacement which becomes less effective for lower technology nodes should be modified. For this purpose, gate oxide variation along with the stacking effect can be utilized for gate replacements. The gate leakage which occurs due to tunneling of electrons shows a strong dependence on the oxide thickness (T_{ox}) [12-16]. Therefore, even a little change in value of T_{ox} the gate leakage increases exponentially. Utilizing this concept, transistors having higher value of gate leakage are replaced by the transistors consist of high T_{ox}. The drawback of this technique is that the total delay of the CMOS circuits increases with increase in T_{ox}. Hence, to reduce the additional delay, the width of MOS transistor is also increased. Using this method, a modified gate replacement technique is proposed which can be applied after choosing optimized value of T_{ox} on the basis of trade off in leakage and delay. Once this optimum T_{ox} chosen, replacement of gates can be done to achieve reduction in both gate and subthreshold leakage.

Along with subthreshold and gate leakage, the reverse biased leakage across *pn* junction also contributes to the total leakage current in lower technology nodes and degrades device performance. The subthreshold leakage effectively reduces using stacking effect of transistors. This approach results in reduced leakage current for increase in circuit area and delay. To overcome this problem, it is preferable to use the RBB technique that increases the V_{th} of transistors in standby mode and reduces the subthreshold leakage [17,18]. However, this technique aggravates the short channel effects. Beside this, the gate leakage primarily depend on different input vectors applied to a CMOS circuit. With suitable choice of different input vectors, gate leakage can be effectively reduced. In comparison to the subthreshold and gate leakage, the BTBT leakage current can be ignored for the 45nm technology nodes. However, the reverse bias voltage variation increase the amount of BTBT leakage. Consequently, the value of reverse body bias is cautiously chosen that results in reduction of leakage without any significant increase in BTBT leakage current. Using the gate replacement along with the reverse body bias technique, a new approach for leakage

reduction is applied by considering the increase in area and short channel effects in CMOS VLSI circuits.

1.3 Organization of thesis

The organization of dissertation report is presented as:

Chapter 1 introduces different leakage mechanisms dominating in nanometer CMOS VLSI circuits. Beside this, different circuit level techniques for leakage reduction are also discussed along with the issues related to present techniques.

A detailed description about various circuit level techniques is presented in chapter 2. The briefs about the different techniques such as stacking, reverse body bias (RBB) variation, pin reordering and logic reconstruction, T_{ox} variation, etc. are illustrated in this chapter.

Chapter 3 demonstrates the basic gate replacement technique that primarily depends on the stacking effect of transistors. Using this technique, the gates at worst leakage state are replaced by another gate incorporating an extra *SLEEP* signal. This technique effectively increases overall delay and circuit area.

The above mentioned problem is removed in chapter 4 by using a modified gate replacement technique that utilizes the variation of T_{ox} and RBB. This modified technique is applied to different benchmark circuits to analyze the overall leakage current and number of replaced gates.

Finally, chapter 5 draws a brief summary. Some interesting points regarding the future work on basis of leakage reduction mechanism are also discussed.

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
SHRAMAM VINA NA KIMAPI SADHYAM

Chapter 2

LEAKAGE CURRENT REDUCTION TECHNIQUES

In nanometer regime, the performance of CMOS VLSI circuits is mainly limited by three major sources of leakages such as subthreshold, gate and BTBT leakage. Among these, subthreshold leakage current dominates the three leakages in the nanoscale technology nodes. As we move down to lower technology nodes up to 32nm, the gate and BTBT leakage also becomes significant which degrades device performance. This chapter presents various circuit based techniques for leakage reduction along with their effect on different leakage mechanisms.

2.1 Leakage reduction using stacking effect of transistors

Subthreshold leakage that flows through a larger number of stacks of 'off' transistors reduces substantially. This is referred as the stacking effect in transistors [2-3]. Let us consider a NAND gate having two inputs as presented in Fig. 2.1. When both the transistors M_1 and M_2 are switched off, the positive potential drop across the node (V_M) results in negligible drain current flowing through it.

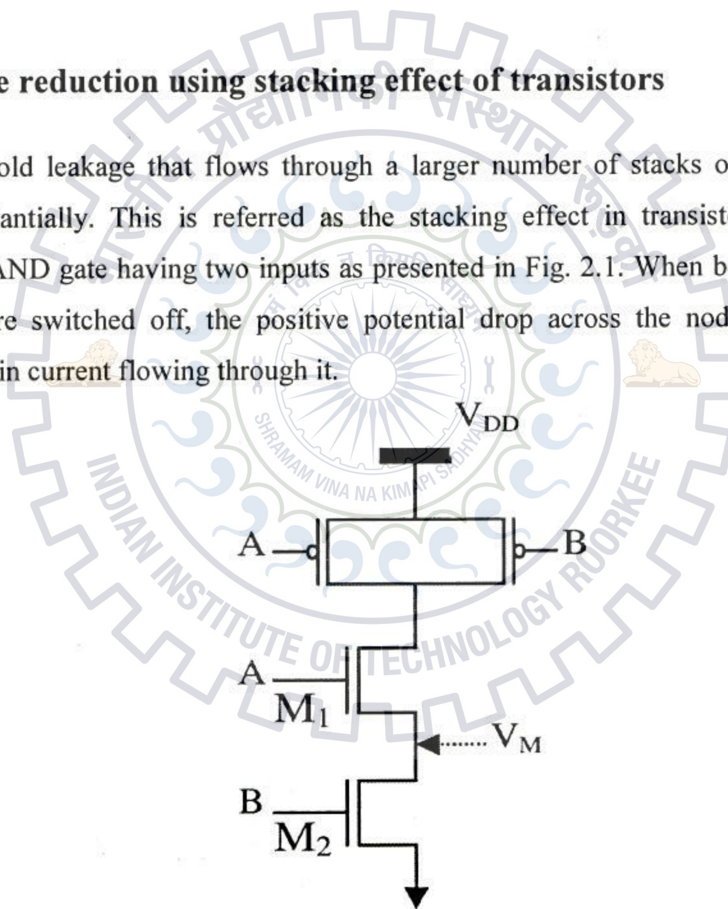


Fig. 2.1 Stacking effect ofa NAND gate having two inputs

Due to positive value of potential V_M at the transitional node following effects can be observed [1]:

1) Gate to the source region voltage (V_{GS1}) of transistor M_1 appears as negative that results in reduced subthreshold current.
2) Potential between body and source (V_{BS1}) of transistor M_1 becomes more negative, leads to increase in threshold voltage ofM_1. Consequently, the subthreshold leakage reduces due to larger body effect.
3) Potential drop between drain to the source (V_{DS1}) of transistor M_1 is reduced, which results in increasing threshold voltage of transistor M_1 which also reduces the sub threshold leakage.

The leakage of a two transistor stack is lesser than the leakage of a single transistor in order of magnitude. Thus, as number of 'OFF' transistors in stack increases, the overall leakage reduces as presented in Fig. 2.2.

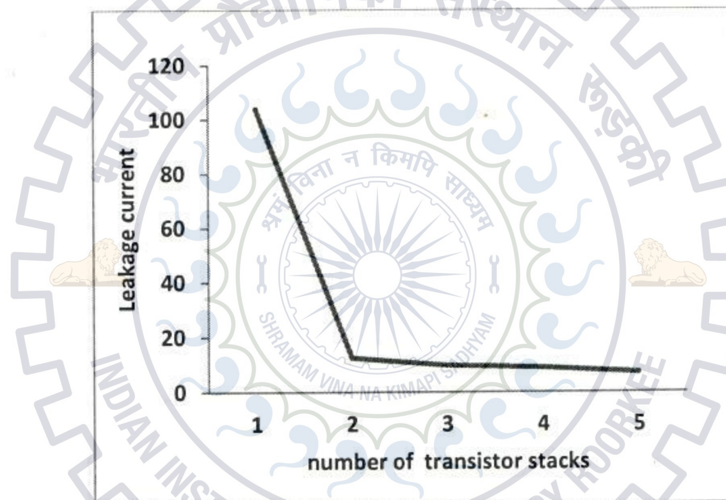


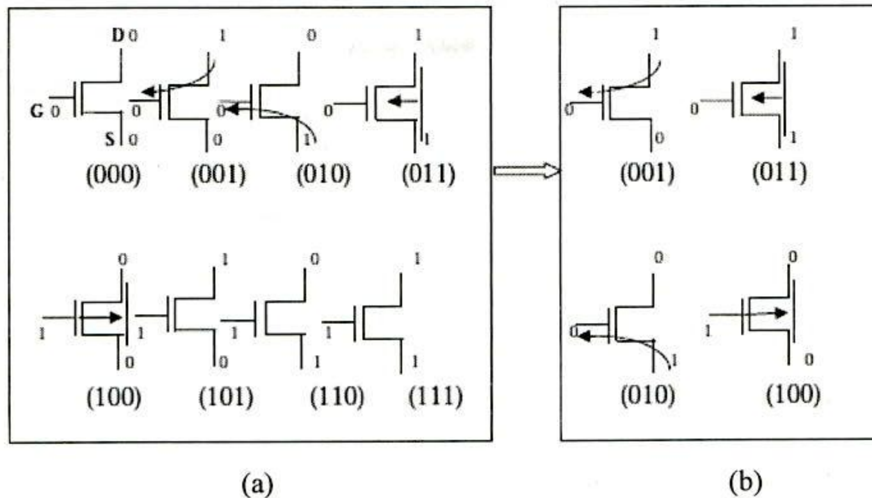
Fig. 2.2 Effect of stacking on leakage current

2.2 Leakage reduction using pin reordering and logic reconstruction

The dependence of subthreshold leakage on different input states is primarily based on stacking effect of transistors. According to stacking effect, the subthreshold leakage reduces when more number of transistors in stack is in 'off' condition. For example, the state (000) exhibits lower subthreshold leakage than state (001) [3-7]. As we move towards lower technology nodes, the gate leakage also increases considerably. The total leakage is therefore the sum of these two leakages. Although a state can produce lowest subthreshold leakage but not necessarily exhibits lower gate leakage. This section determines the gate leakage current

dependence of single transistor on different biasing conditions. Based on these conditions, one can estimate the total leakage of the circuit. Let us consider an NMOS transistor constituting of eight possible bias conditions (states) which can be grouped into four different categories. The first one comprises of the states (000) and (111). Among these two, no tunneling in gate occurs. The (100) and (011) becomes the second kind. For (100), electrons tunnelling from the channel region, source region and the drain region occurs directly into the gate region. For (011), the approach is same as that of (100)besidethe direction of current which is reversed. The next type consists of (001) and (010). As in the CMOS transistor source/drain terminals can be interchangeably used without having any change in operation, thus these states have no difference. In case of this state, electron tunnels directly from gate side to source/drain region. The last category includes (101) and (110) are basically termed as transient states which cannot occur in steady state of operation and becomes eventually equal to (111) bias state. Finally , we observe that there remains only four types of different states, i.e., (001), (010), (011) and (100) as shown in Fig. 2.3, which are significantly contributes to the total gate leakage current in comparison with the remaining four types of biasing states. Depending on these states gate leakage for all transistors is computed. On the basis of these states and stacking effect, the best state is choosen that gives lowest total leakage including gate and subthreshold leakage.

If delay is also considered for the leakage reduction than states must be choosen which gives lower leakage along with lower propagation delay. In a complex CMOS circuits, some input signals might be more critical than others. This is due to different propagation delays of logic gates along different paths which makes some signal arrive later than others. The path through the logic producing maximum delay, determines the ultimate speed of the structure is called critical path. Moving the transistor in critical path more closed to the output results in reduction of propagation delay. As demonstrated in Fig. 2.4, signal IN_1 is assumed to be critical signal. Furthermore, when IN_2 and IN_3 are at high logic, then IN_1 undergoes a transition from 0 to 1. Assume that the load capacitance C_L is initially charged at high potential. Incase, no path to *GND* exists untill transistor M_1 is turned ON, the delay between arrival of IN_1 and the output is therefore determined by the time taken by C_L, C_1 and C_2 to discharge. In the second case, capacitors C_1 and C_2 are already discharged when IN_1 changes states. In this case only C_L has to be discharged which results in a smaller delay [21].



(a) (b)

Fig. 2.3 (a) Biasing conditions for a NMOS transistor and (b) Significant contribution states.

Logic reconstruction is a low cost technique for reducing delay. In this technique, we can manipulate the logic equations to reduce fanin of the gate that reduces the total propagation delay of the CMOS VLSI circuit.The gate with high fanin can be partitioned into a logic recombination consisting of gates with lower fanin maintaining the same functionality [21].

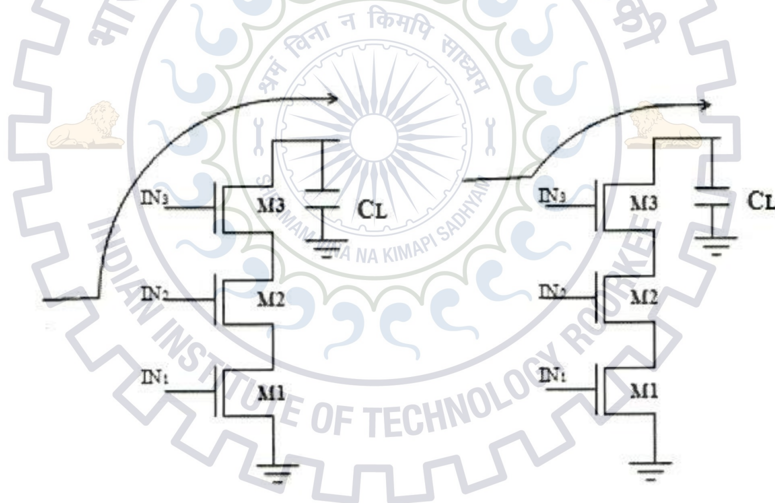


Fig. 2.4 Influence of ordering of transistors on delay

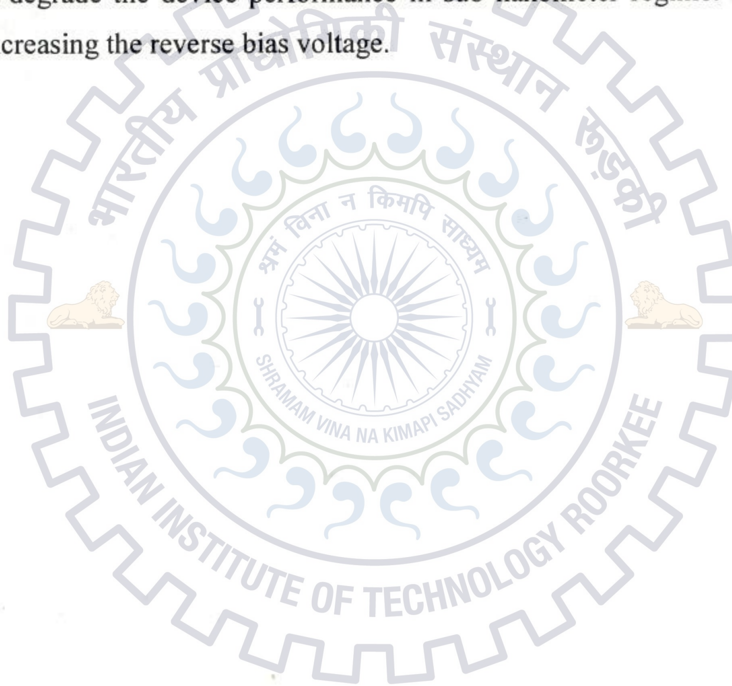
2.3 Leakage reduction using Dual-T_{ox} technique

Lower value of threshold voltage can result in larger subthreshold leakage which posses an exponential dependence on V_{th}. A major cause of the gate leakage (I_{gate}) is a result of the tunneling due to electrons from gate region through the gate oxide. The electron tunneling probability shows a strong dependence on the electric field which is applied and

thickness of the barrier (T_{ox}). As a result, even a little change in value of T_{ox} has remarkable effect on gate leakage current (I_{gate}) [12-16]. Apart from this, gate oxide thickness helps in modification of the threshold voltage (V_{th}). Thus, dual V_{th} can easily be applied in any circuit by depositing the oxide with varying thickness in noncritical paths.

2.4 Leakage reduction using RBB voltage

Subthreshold leakage that dominates other leakages in nanometer regime primarily depends exponentially on V_{th}[17-18]. With the increase in reverse bias body voltage, the V_{th} also increases. This higher V_{th} results in reduction of subthreshold current efficiently. However, with the further increment in the value of reverse biasing, the BTBT leakage also increases that degrade the device performance in sub nanometer regime. Hence, limitation imposed on increasing the reverse bias voltage.



Chapter 3

BASIC GATE REPLACEMENT TECHNIQUE

Input vector control (IVC) is one of the most efficient approach for the reduction in leakage power. It basically employs the transistor stacking effect in the CMOS gates by application of the minimum leakage states (MLS) [6-8] to the primary inputs of a VLSI circuits during standby mode of operation. This technique helps in reduction of both leakage i.e. subthreshold and gate leakage. The conventional method of gate replacement replace gate $G(\vec{x})$ by that of another $\tilde{G}(\vec{x}, \text{SLEEP})$, where $\vec{x}$ is the primary input vector at G, such that[7]:

1) $\tilde{G}(\vec{x}, 0) = G(\vec{x})$ when circuit operates in active mode ($SLEEP = 0$);
2) $\tilde{G}(\vec{x}, 1)$ has smaller value of leakage than $G(\vec{x})$ when the circuit operates in standby mode ($SLEEP = 1$).

The above mentioned conditions guarantees for maintaining the circuit's correct functioning in active mode, while reducing the leakage of the gate G when operating at the standby mode. This condition can result in changing the output of gate replaced. Therefore, the output of this gate affects the leakage current flowing in other gates too and thus they should be replaced carefully. Fig. 3.1 shows the basic gate replacement mechanism for different fanout gates. In Fig. 3.1(a), the gate G is considered to be replaced by $\tilde{G}$ that alters the output 0 to 1. If the output of G goes to the input of H (fanout gate), then the following conditions can be occurred:

1) In the case of Figs. 3.1(b) and (d), replacement has no effect on output of other gates.
2) For 3.1(c) and (f), replacement of G not only alters the output of H, but also makes H possessing WLS. Therefore, replacement of gate H by $\tilde{H}$can solve the problem. This process continues until all the logic gates are visited.
3) Finally, replacement in Fig. 3.1(e), place away both the gates G and the gate H far from states possessing worst leakage. It change the output H which is a NOR gate that can be conducted for similar methodology.

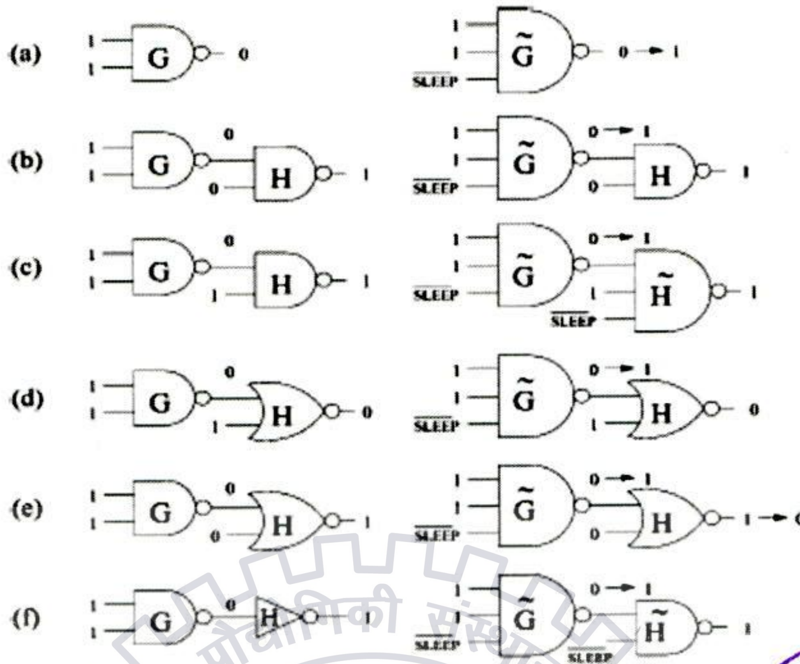


Fig. 3.1 Gate replacement and its effect on fanout of gates

3.1 Basic gate replacement algorithm

Considering a C17 benchmark circuit from the 'MCNC91' suite [24-25] as presented in Fig. 3.2(a). An in-depth search evaluates the minimum leakage state (MLS) as (00010) which produce lower leakage. With this minimum leakage vector, corresponding minimum leakage current obtained which equals to 831.08 nA. Here, we find that gate $G_3$is inworst leakage state which result in current of 454.5 nA with input state as [1, 1].This accounts for more than fifty percent of the total leakage. We observe here that a major contributor of the overall leakage are the gates that posses WLS. Instead of manipulating the primary input states, we can use approach of replacing these gates contributing more to the total leakage. In order to replace the NAND G_3 consisting of two inputs, by that of a NAND gate $\widetilde{G_3}$consisting of three inputs includingthe *SLEEP* signal which act as control signal is added to original gate. In active mode of operation, $\overline{SLEEP} = 1$ and possessing the same output as that of gate G_3. However, in the standby mode $\overline{SLEEP} = 0$ and $\widetilde{G_3}$ draws a current of 94.87 nA. This replacement of gate may result in changing of the output of this replaced gate in the sleep mode which in turn affects further the leakage which flows in gates G_5 and G_6. In such cases, we replace these fanout gates in the same manner. This now result into the new observed circuit's leakage to be as 476.88 nA with a 43% further decrement in comparison to the

original leakage current which is approximately equal to 831.08 nA. The proposed technique is different in concept from the IVC techniques already existing. Exclusively, the approach of IVC considers complete circuit and then find most suitable input vector which provides smaller leakage. The technique of gate replacement focuses on the gates that posses worst leakage contributing states under a particular vector and replacing them to obtain overall reduction in the leakage [7]. The pseudo code for basic gate replacement technique is presented in Fig 3.3.

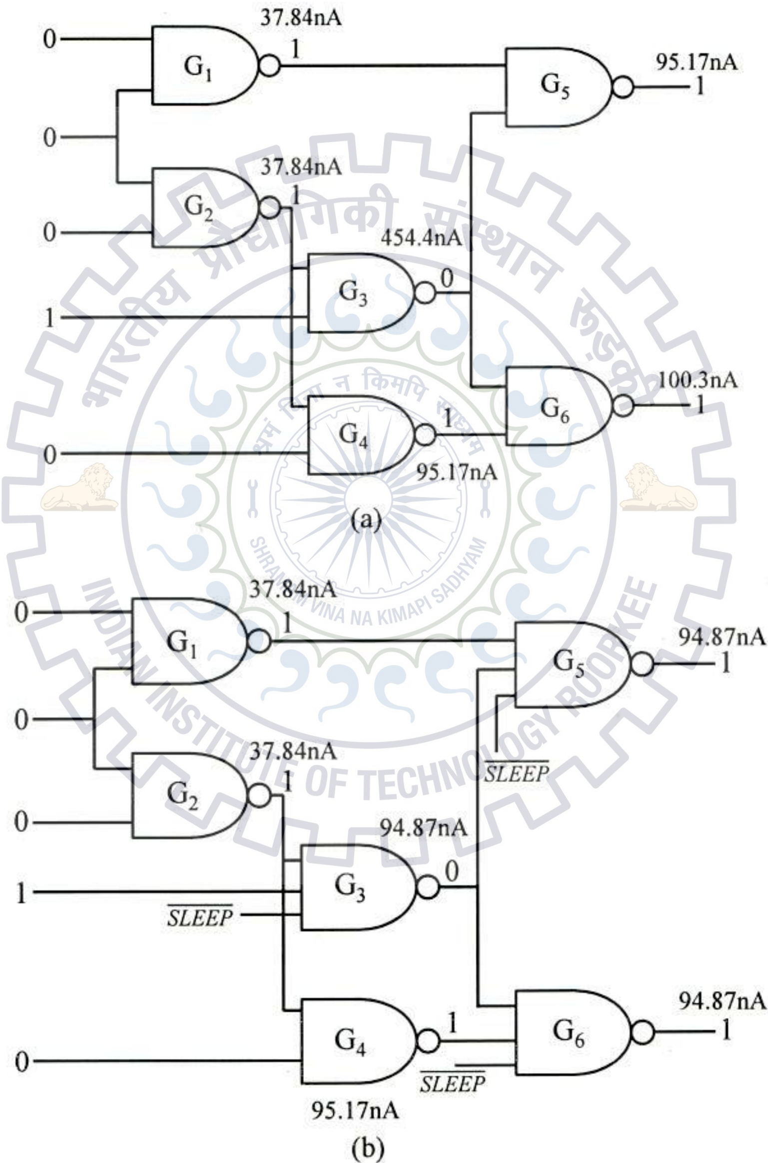


Fig. 3.2 Example for gate replacement (a) C17 with leakage of 831.08 nA (b) C17 after gate replacement with total leakage 476.88 nA

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

Inputs: $\{G_1, G_2, ...\}$: these gates are sorted in topological order,

$\{x_1, x_2, ...\}$: it is input state vector,

SLEEP: control sleep signal.

Output: circuit which maintains same operation as that of original when *SLEEP* = 0 and reduced value of leakage in case of *SLEEP* =1.

Gate replacement algorithm:

```
1. for each gate Gi ∈ { G1, G2,.........}
2. if (Gi is at WLS and not marked )
3.   include Gi in the selection S;
4.   while (there is new addition to S)
5.     for each newly selected gate G in S
6.       if (there exists library gate G̃ meets the conditions
7.         temporarily replace G by G̃;
8.         if (output of G is changed due to this replacement)
9.           include G's unmarked fanout gate Gj in S;
10.          compute the total leakage change of gates in S;
11.          if ( there is leakage reduction )
12.            mark all the gates Gi in the selection S;
13.            make the replacements in lines 7.9. or 10 permanent;
14.          else mark gate Gj only;
15.    empty the selection S;
16. else mark Gi if it has not been marked yet;
```

Fig. 3.3 Pseudo code for basic gate replacement technique

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

Chapter 4

Leakage reduction mechanisms based on stacking effect

In recent years, leakage power dominates the dynamic power in nanoscale CMOS VLSI circuits. This chapter provides a detailed description for reduction of the total leakage in nanoscale VLSI circuit. As the stacking effect becomes less effective at lower technology nodes, various other techniques like reverse body bias variation, gate oxide variation can be used along with stacking for leakage reduction. The different mechanisms for leakage reduction have been proposed in this section which primarily based on gate replacement technique.

4.1 Leakage Reduction using Modified Gate Replacement Technique for CMOS VLSI Circuit

The drawback of conventional gate replacement technique is that it primarily focuses on the higher number of replacements which affects the circuit area and delay. A novel approach of reduction in leakage current is proposed which is primarily based on the conventional gate replacement technique. This approach is more effective in circuits with higher logic depth. Using this technique, one has to replace all the gates marked as WLS. That encouragingly affects the circuit performance such as delay and area. In modified technique, we can neglect the gates with high fanout that leads to increase in overall leakage current. A comparative analysis is performed between the conventional and modified gate replacement mechanisms. Using the modified technique, the overall leakage current and number of replacements are reduced by 13.5% and 33.5% respectively as compared to the conventional one.

4.1.1 Modified gate replacement technique

Using this technique, all the logic gates in a CMOS circuit is visited by topological order. The gates with lower leakage are skipped and gates at WLS are replaced. After each replacement, we will check the total leakage of the circuit that has to be reduced. Using the replacement, if the overall leakage of the logic circuit increases, it can be assumed that the output of replaced gate increases the leakage of other gates too. Therefore, we will skip the

gate without any replacement. Next, we will visit other gates at WLS and replace them if overall leakage reduces, otherwise skip those gates too. Once all the gates are visited, we will again check all the skipped gates which were at WLS. We can replace these gates again using gate replacement with an AND gate following that logic gate. The advantage of using AND gate with *SLEEP* signal is that the output of the gate is always low in standby mode. Using this technique, the replacement has been done only if the overall leakage reduces. By using this approach, we achieve reduced leakage current for lower number of replacements. This approach is more effective when logic depth of the circuit is high. Using this approach of replacement, a new algorithm is proposed that can efficiently follow the modified technique as shown in Fig. 4.1. It is observed that the leakage current and number of replaced gates are significantly reduced for modified gate replacement technique. The reason behind is that the modified technique replaces the gates that reduces the overall leakage, whereas the conventional technique [7] replaces all the gates marked as WLS. Therefore, the modified technique reduces the number of replacements in a logic circuit that can encouragingly reduce the circuit delay and area.

Inputs: $\{G_1, G_2, ...\}$: these gates in a circuit sorted in a topological order.

$\{S_1, S_2, ...\}$: it is the subset of gates that includes all those gates which are at WLS but not replaced.

Output: circuit which maintains same operation as that of original when *SLEEP* = 0 and reduced value of leakage in case of *SLEEP* =1.

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

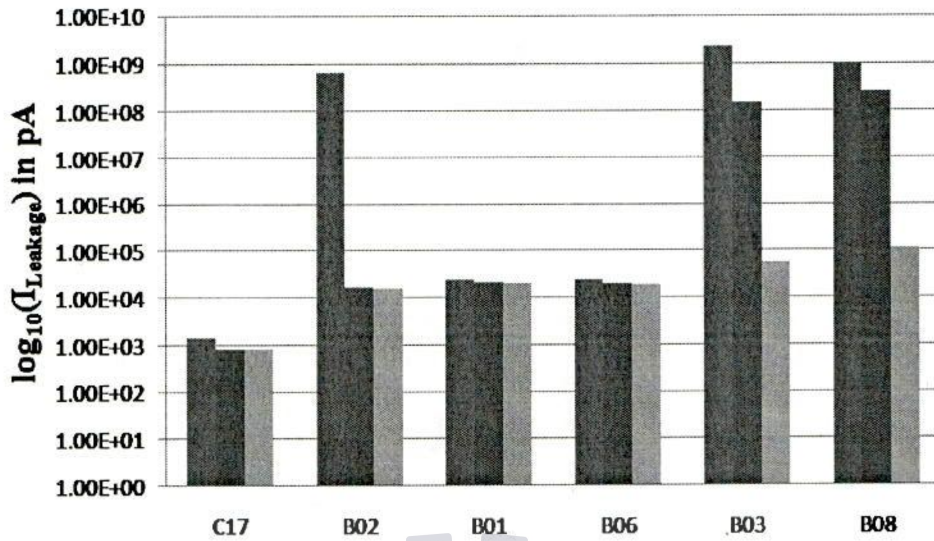
Modified gate replacement algorithm:

```
1. foreach gate G_i ∈ { G_1, G_2, G_3, ......... G_n}
2.    if (G_i is at WLS and not marked )
3.      replaceG_itemporarily
4.        if (Total leakage reduces)
5.          make the replacement permanent and mark the gate;
6.        else move gate to S;
7.    else mark and move to next gate;
8. go to step 1, till all gates in G_i are marked.
9. for each gate S_i ∈ { S_1, S_2, S_3, ......... S_n}
10.   if (S_i is not marked )
11.     replace S_i temporarily followed by 2-input AND gate;
12.       if (Total leakage reduces)
13.         make the replacement permanent and mark thegate;
14.       else mark and move to next gate;
15.   else move to next gate;
16.go to step 8, till all gates in S_i are marked.
```

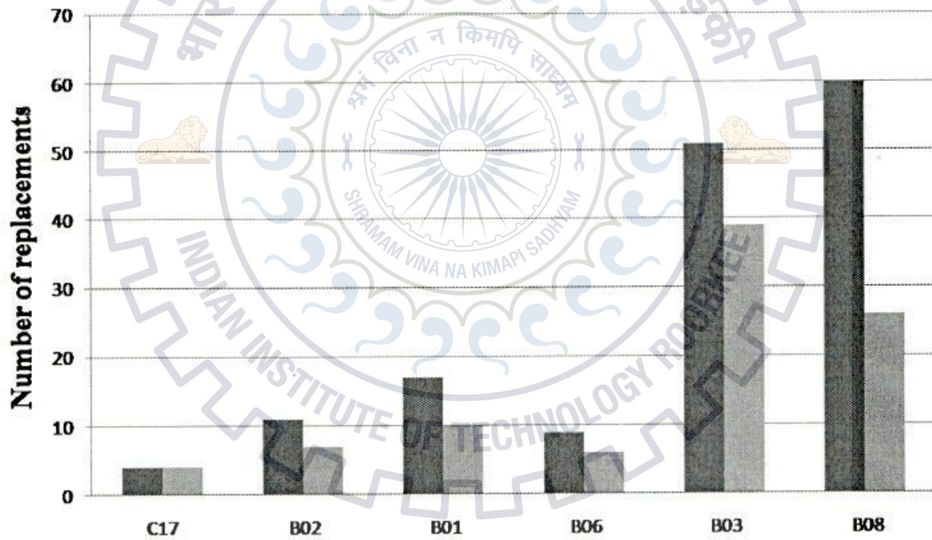
Fig. 4.1 Pseudo code for the modified gate replacement algorithm

4.1.2 Results and discussion

This section analyzes the minimum leakage finding technique by using the above mentioned concept of modified gate replacement. Different benchmark circuits of ITC'99 [24] and ISCAS [25] series are used to find the minimum leakage current. In this section, a comparative analysis is performed in different benchmark circuits in terms of leakage current and number of replaced gates. Figs. 4.2(a) and 4.2(b) demonstrates the comparative analysis for leakage current and number of replaced gates respectively. It is observed that the leakage current and number of replaced gates are significantly reduced for modified gate replacement technique. The reason behind is that the modified technique replaces the gates that reduces the overall circuit leakage, whereas the conventional technique [7] replaces all the gates marked as WLS. Therefore, the modified technique reduces the number of replacements in a logic circuit that can encouragingly reduce the circuit delay and area. The percentage reduction in minimum leakage current and number of replaced gates are summarized in Table 4.1. It is observed that the overall reduction in leakage and replacements for modified technique are 13.5% and 33.5% respectively as compared to the conventional gate replacement mechanism [7].



(a)



(b)

Fig. 4.2 (a) Minimum leakage current ($I_{leakage}$) and (b) number of replaced gates for different benchmark circuits using conventional [7] and modified gate replacement technique

Table 4.1 Percentage Reduction in Leakage using Modified Gate Replacement Technique *w.r.t.* Conventional one for different Benchmark Circuits

Benchmark circuits		Conventional gate replacement mechanism [4]		Modified gate replacement mechanism		% reduction in leakage by modified technique	% reduction in logic gates replaced by modified technique
Circuit name	Number of logic gates	Number of gates replaced	Leakage current (nA)	Number of gates replaced	Leakage current (nA)		
C17	6	4	0.81	4	0.81	0	0
B02	26	11	16.63	7	15.94	4.12	36.36
B01	45	17	21.18	10	19.78	6.58	41.17
B06	57	9	19.64	6	19.18	2.34	33.33
B08	170	60	245147.00	26	111529.00	54.50	56.67

4.2 Modified Gate Replacement Algorithm for Leakage Reduction using Dual-*Tox* in CMOS VLSI circuits

Different leakage mechanisms occurs in nanometer regime which includes the subthreshold and gate leakage current that degrades device performance. Based on the conventional gate replacement technique of leakage reduction, this section presents a novel algorithmas presented in Fig. 4.3. This technique employs the stacking effect using dual-T_{ox} transistors. This approach is more effective for lower technology nodes wherein the gate leakage dominates the subthreshold leakage. The stacking effect, used with dual-T_{ox} transistors, efficiently reduces the gate and subthreshold leakage in both the standby and active mode. Apart from this, leakage current can be further reduced using the pin reordering technique. The conventional gate replacement technique [7] efficiently reduces the subthreshold leakage without considerable reduction in gate leakage. Theoretically, the probability of electron tunneling can be referred as a strong function of the oxide thickness (T_{ox}). Therefore, a small change in T_{ox} can have a tremendous impact on gate leakage. Using this concept, the transistors with higher gate leakage can be replaced by the transistors with higher T_{ox} values [12-16]. The different quantitative values of T_{ox} limits the application of this approach, thus, it is required to find a most appropriate value of T_{ox} that will effectively reduce the gate leakage with lesser delay penalty.

Input: $\{G_1, G_2, \ldots. G_{n-1}, G_n\}$: gates in a circuit arranged in topological order

Output: a circuit that posses same operation in active mode as that of original along with lower leakage in standby

Modified gate replacement algorithm using T_{ox} variation

```
1. for each gate Gi ∈ {Gn, Gn-1, Gn-2, ......... G2, G1}
2.   if (Gi is at WLS and not marked )
3.     replace Gi temporarily;
4.       if (Total leakage reduces)
5.         Apply pin reordering;
6.           if (Total leakage reduces)
7.             Make the changes permanent and move to next gate Gi-1
               at WLS;
8.           else go to next step;
9.       else for each gate Gj ∈ { Gi-1, Gi-2, Gi-3, ......... Gn-1, Gn}:
10.    if (Gj is at WLS and not marked yet )
11.      replace Gj temporarily;
12.      if (Total leakage reduces)
13.        Apply pin reordering;
14.          if (Total leakage reduces)
15.            Make the changes permanent and move to next gate
               Gj-1 at WLS;
16.          else go to next step;
17.      else mark and move to next gate in Gm till all gates in Gm are marked;
18.   else move to next gate in Gm until all gates in Gm are marked;
19. else move to next gate in Gi until all gates in Gi are marked;
```

Fig. 4.3 Pseudocode of the modified gate replacement algorithm

The higher value of T_{ox} considerably reduces the gate leakage while the overall propagation delay of the CMOS circuitry increases. One of the exciting solution to reduce the delay penalty is to increase the width of the transistor. Using this approach, a modified gate replacement technique can be applied after choosing some optimum value of T_{ox} on basis of the leakage and delay trade off. For an instance, the stacking effect of NMOS transistor in off state produces lower subthreshold leakage for a two-input NAND gate, whereas the PMOS transistors produces higher gate leakage if all the inputs are at higher potential. For this purpose, a 2-input NAND gate can be replaced using a 3-input NAND gate that substantially reduces the subthreshold leakage without considerable reduction in gate leakage.Therefore, it can be preferred to replace the PMOS transistors having lower T_{ox} by using the PMOS transistors having higher T_{ox}. It reduces the gate leakage as well as the subthreshold leakage of a two-input NAND gate.Apart from this, pin reordering is also referred as an effective technique for leakage reduction that can be applied after each replacement of the gate.

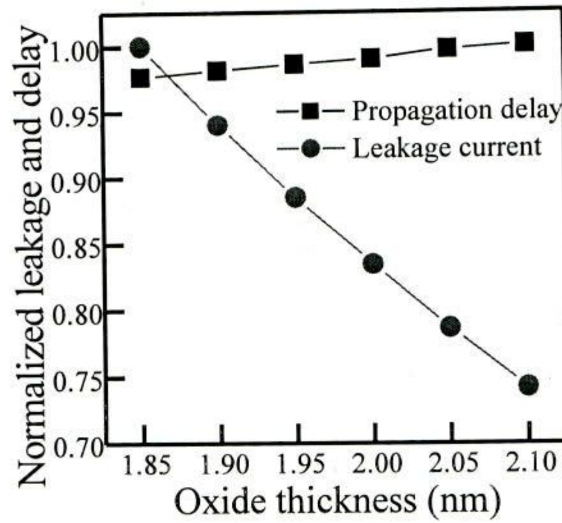
The conventional mechanism [7] used the technique to replace a gate at WLS in topological order. Using this approach, the outputs of the replaced gates have been changed that affects the leakages of fan out gates which are again considered for replacement. While moving towards the lower technology nodes, the gate leakage substantially dominates the subthreshold leakage that affects the leakage of fanin gates. Therefore, it is required to consider the leakage of fan out as well as fan in gates for replacement. Thus, using the modified algorithm as depicted in Fig. 4.3, the replacement has to be done for the gates from output towards the input by considering the effect of leakage on fan in gates.

4.2.1 Results and discussion

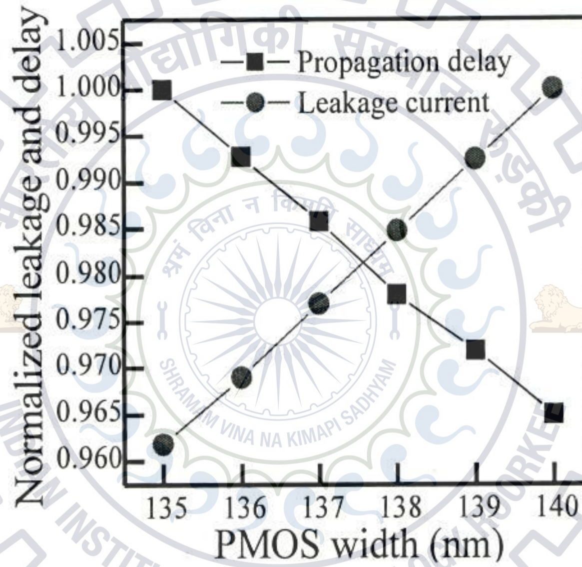
Using the above mentioned gate replacement algorithm, this section analyzes the reduction in leakage currents for different oxide thickness (T_{ox}), transistor widths (w) and benchmark circuits.

4.2.1.1 Analysis of leakage current for different oxide thickness

The gate oxide leakage is primarily referred as a strong function of gate oxide thickness. Therefore, the variation of oxide thickness results in tremendous impact on the gate leakage current. The higher oxide thickness in noncritical paths reduces the gate leakage along with the subthreshold leakage while the lower oxide thickness in critical paths maintains the performance of the circuit. The variation of normalized leakage and delay for different oxide thickness is shown in Fig. 4.4(a). To reduce the complexity of the approach, an appropriate value of T_{ox} is selected among the set of values as presented in Table 4.2.From these values, it is observed that the leakage current reduces with a slight delay penalty. For analysis, the value of T_{ox} for a PMOS transistor is selected as 2nm. To overcome delay penalty, width of the PMOS transistor is varied in the range of 135nm to 140nm for a fixed value of T_{ox} = 2nm. Figure 4.4(b) exhibits the normalized leakage and delay for different widths of PMOS transistors. Using the data presented in Table 4.3, the suitable value of the width for a PMOS transistor can be selected as 137nm. This quantitative value significantly reduces the delay with a negligible effect on the leakage of the circuit as presented in Fig. 4.4(b).



(a)



(b)

Fig. 4.4 Variation of propagation delay and leakage current with (a) oxide thickness and (b) width of PMOS transistor

Table 4.2 Variation of leakage current and propagation delay for different T_{ox} of a PMOS transistor

T_{ox} variation (nm)	Leakage current (nA)	Propagation delay (ns)
1.85	8.582	0.0987
1.90	8.069	0.0991
1.95	7.594	0.0996
2.00	7.156	0.1001
2.05	6.748	0.1007
2.10	6.370	0.1014

Table 4.3 Variation in leakage current and propagation delay for different widths of PMOS transistor at T_{ox} = 2nm

Width of PMOS transistor (nm)	Leakage current (nA)	Propagation delay (ns)
135	7.1557	0.1001
136	7.2129	0.0994
137	7.2702	0.0987
138	7.3275	0.0980
139	7.3847	0.0978
140	7.442	0.0967

4.2.1.2 Analysis of leakage current for gate replacement using Dual-T_{ox} transistors

Once the value of T_{ox} is chosen, the modified gate replacement algorithm is applied wherein the gates can be replaced with another gate. This replaced gate incorporates an extra sleep signal along with PMOS transistors with higher value of T_{ox}. As presented in Table 4.4 while the variation of T_{ox} is applied to a circuit, it will not reduce the leakage of the circuit for all input states. Thus, the modified gate replacement approach can be applied wherein a two input NAND gate is replaced with three input NAND gates having higher value of T_{ox} for PMOS transistors. It is observed that when the sleep signal is low, circuit is in the standby mode. Using this mode, a drastic reduction in leakage current is observed along with the functionality of the circuit that is maintained in active mode without significant increase in leakage current.

Table 4.4 Leakage current for different input states for a two input NAND gate using T_{ox} variation and stacking effect along with dual-T_{ox} transistors

Input states	Initial Leakage (nA)	Leakage using T_{ox} variation (nA)	Leakage using stacking along with dual-T_{ox} transistors (nA)	
			Sleep=0	Sleep=1
00	0.786	0.772	0.929	1.104
01	13.262	13.255	0.717	12.391
10	4.850	4.843	0.690	7.432
11	17.161	14.309	6.502	17.052

4.2.1.3 Analysis of leakage current for benchmark circuits

This sub section analyzes the minimum leakage finding technique by using the modified gate replacement along with dual-T_{ox} PMOS transistors. Different benchmark circuits of ISCAS and ITC'99 [24,25] series are used to find the minimum leakage current. It is observed that the overall leakage of the benchmark circuits is reduced by 39.9% in standby mode as presented in Table 4.5 and Fig. 4.5.

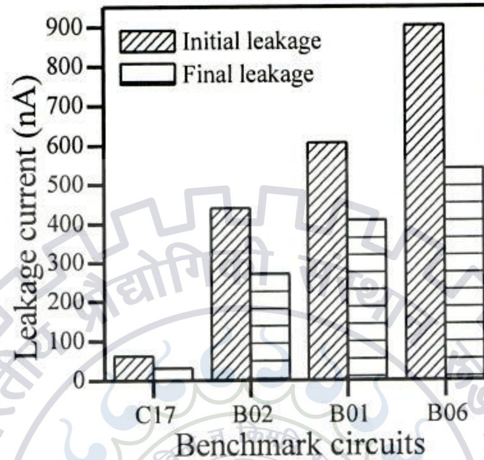


Fig. 4.5 Leakage current reduction for different Benchmark Circuits

Table 4.5 Leakage reduction using modified replacement technique for benchmark circuits

Benchmark circuits	Initial leakage (nA)	Final leakage (nA)	% reduction
C17	61.407	31.604	48.5
B01	606.140	407.66	32.7
B02	439.828	271.20	38.3
B06	902.301	540.77	40.1

4.3 Modified Gate Replacement Algorithm for Leakage Reduction using RBB variationin CMOS VLSI circuits

In current nanoscale CMOS VLSI circuits, subthreshold, gate and band-to-band tunneling leakage currents are introduced due to the adverse effect of scaling. This section presents a novel algorithm that utilizes the transistor stacking effect along with the variation in reverse body bias (RBB) [17-18]. This novel algorithm primarily accounts for the conventional gate replacement approach with reduced number of replaced gates. This approach applies the RBB variation for the gates at worst leakage state (WLS) while having

lower leakage than other gates at WLS. The RBB variation technique cannot be used for the gates with higher leakage as it requires larger body bias voltage for significant reduction in leakage current. Furthermore, increase in RBB aggravates short channel effects that degrade the circuit performance. The stacking effect along with a lower value of RBB helps in considerable reduction of leakage current for the gates producing higher leakage than others.

This section presents a detailed analysis of leakage reduction technique using RBB voltage variation and gate replacement technique in CMOS VLSI circuits. The major contributors of leakage current in deep nanometer regime are subthreshold leakage, gate leakage and reverse biased *pn* junction BTBT leakage current. The subthreshold leakage effectively reduces using stacking effect of transistors. Based on the stacking effect, the gates at posses worst leakage contributing stateare replaced by that of another gate with same functionality incorporating an extra sleep signal [7]. With the use of the sleep signal, the leakage current is reduced in standby mode along with maintaining circuit functionality in the active mode. If the number of replacements is higher than the gate replacement, this approach results in reduced leakage current at the cost of increament in total circuit's area and delay. To overcome this problem, it is preferred to use the RBB technique that increases the V_{th} of transistors in standby mode and reduces the subthreshold leakage [17, 18]. However, using this technique aggravates short channel effects. Beside this, the gate leakage posses a dependency on different input vectors applied to a CMOS circuit. With suitable choice of different input vectors, gate leakage can be effectively reduced. In comparison to the subthreshold and gate leakage, the BTBT leakage current can be ignored for the 45nm technology nodes. However, the reverse bias voltage variation increases the BTBT leakage current. Consequently, the value of reverse bias voltage is cautiously chosen that results in reduction of leakage without any significant increase in BTBT leakage current. Using the gate replacement along with the reverse body bias technique, a new approach for leakage reduction is introduced by considering the increase in area and short channel effects in CMOS VLSI circuits.

4.3.1 Modified gate replacement technique using RBB variation

The conventional gate replacement technique effectively reduces the leakage in nanoscale CMOS VLSI circuits [7]. However, if the number of gates at WLS is high, more gates are required to replace that accounts for increase in circuit area. For any logic gate, leakage current varies depending on different input vectors. Let us suppose for any gate, '*n*'

number of inputs states are regarded as WLS. Among these states, few states exhibits higher leakage than others. Hence, a higher body bias is required to apply for such cases and lower body bias for the remaining ones. With increase in reverse bias voltage, the short channel effects come into consideration. Thus, for the states with higher leakage than others, rather than applying higher reverse bias voltage, a gate replacement is applied that utilizes the stacking effect of transistors along with lower reverse bias voltage. This technique results in reduction of the number of replacements and short channel effects as compared to conventional gate replacement and reverse body biasing technique respectively.

In this approach, it is desirable to find the leakage in case of all possible combination of input states for the gate considered for leakage reduction in a CMOS VLSI circuit. The states which are at worst leakage states are considered for leakage reduction. Among these states, reverse body bias is applied to those states which have lower leakage than others. The states consisting of higher leakage are replaced by another gate incorporating an extra sleep signal with same functionality in active mode and reduced leakage at standby mode. For further reduction of leakage in these states, reverse body bias is used along with stack effect. Based on this approach, a novel algorithm is proposed in Fig. 4.6 that utilizes the gate replacement and reverse body biasing techniques for reducing leakage. Stacking is very effective technique for reduction of subthreshold leakage along with the gate leakages but it increases the circuit area. With lower reverse bias voltage, BTBT leakage current does not degrades the device performance and helps in reduction of subthreshold leakage. Hence, this approach effectively reduces leakage using stacking effect and reverse body bias techniques with reduction in number of replacements and thereby a lesser probability of short channel effects.

Input: $\{G_1, G_2, \ldots, G_{n-1}, G_n\}$: gates in circuit that are arranged in topological order

Output: a circuit that posses same operation in active mode as that of original along with the lower leakage in standby mode

Modified gate replacement algorithm using RBB variation:

```
1. for each gate G_i ∈ { G_1, G_2, ……… G_n-2, G_n-1, G_n}
2.   if (G_i is at WLS and not marked )
3.     apply reverse bias voltage;
4.       if (total leakage reduces)
5.         apply pin reordering;
6.           if (total leakage reduces)
7.             make the changes permanent and move to next gate G_i-1
               at WLS;
8.           else go to next step;
9.       else replace the gate G_i temporarily;
10.   if (total leakage reduces)
11.      apply pin reordering;
12.         if (total leakage reduces)
13.           make the changes permanent and move to next gate G_i-1
              at WLS;
14.         else go to next step;
15.   else mark and move to next gate in G_i till all gates in G_i are marked;
16. else move to next gate in G_i until all gates in G_i are marked;
```

Fig. 4.6 Pseudo code of the modified gate replacement technique using reverse body bias voltage variation

4.3.2 Results and discussions

Using the above mentioned approach of leakage reduction, this section analyzes the reduction in leakage current utilizing the transistor stacking effect and reverse body bias technique.

4.3.2.1Effect of RBB on leakage currents

Subthreshold leakage that dominates other leakages in nanometer regime, primarily depends exponentially on V_{th} [1]. With the increase in reverse bias body voltage, the V_{th} also increases. This higher V_{th} results in reduction of subthreshold current efficiently. However, with the further increment in reverse biasing voltage, the BTBT leakage increases that degrade the device performance in sub nanometer regime. Hence, limitation imposed on increased value of reverse bias applied [17, 18]. The effect due to applied reverse body bias on gate, subthreshold and BTBT leakage is also presented in Fig. 4.7(a).Consequently, the variation of total leakage for different reverse body bias (V_{bs}) voltages is also presented in Fig. 4.7(b). It is observed that the total leakage of the NMOS transistor reduces with increasing V_{bs} due to reduction in V_{th}. Moreover, there is no significant reduction in

भारतीय प्रौद्योगिकी संस्थान रुड़की श्रमं विना न किमपि साध्यम INDIAN INSTITUTE OF TECHNOLOGY ROORKEE SHRAMAM VINA NA KIMAPI SADHYAM

subthreshold leakage beyond reverse bias voltage (V_{bs}) of -0.5v as presented in Table 4.6. Also, beyond this value, subthreshold leakage becomes comparable to gate and BTBT leakage. If the body bias is further reduced, the BTBT and the gate leakage become dominant. To avoid the short channel effects, one can choose the value of V_{bs}= -0.5V.

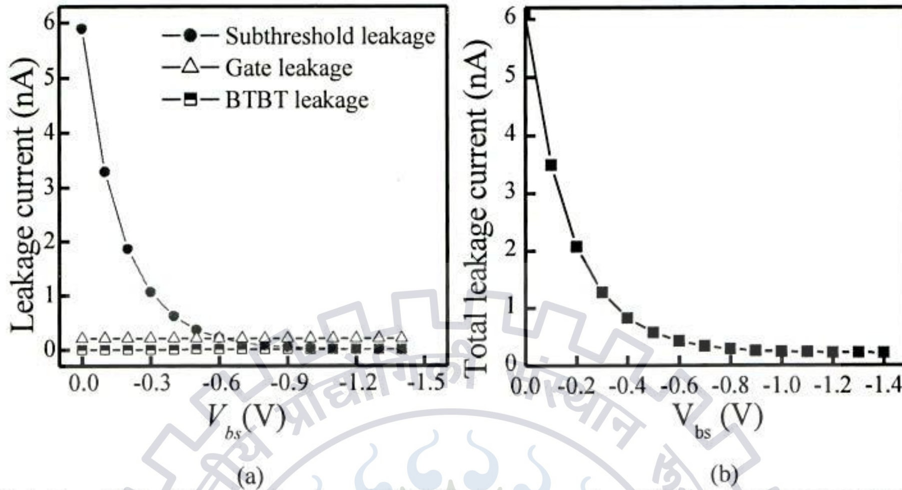


Fig 4.7 Variation of (a) subthreshold, gate and BTBT leakage current and (b) total leakage current for different reverse body bias voltages using a CMOS transistor at 45nm technology node

Table 4.6 Variation of different leakage currents with reverse body bias voltage at 45nm technology node for an NMOS transistor

Body bias voltage (V_{bs})	Gate leakage (nA)	Subthreshold leakage(nA)	BTBT leakage (nA)	Total leakage (nA)
0	0.209	5.890	0.0010	6.100
-0.1	0.209	3.285	0.0012	3.495
-0.2	0.209	1.858	0.0014	2.068
-0.3	0.209	1.065	0.0016	1.276
-0.4	0.209	0.618	0.0018	0.829
-0.5	0.209	0.363	0.0020	0.574
-0.6	0.209	0.215	0.0022	0.426
-0.7	0.209	0.129	0.0024	0.340
-0.8	0.209	0.078	0.0026	0.289
-0.9	0.209	0.047	0.0028	0.259
-1.0	0.209	0.029	0.0030	0.240
-1.1	0.209	0.017	0.0032	0.229
-1.2	0.209	0.010	0.0034	0.222
-1.3	0.209	0.006	0.0036	0.218
-1.4	0.209	0.003	0.0038	0.216

4.3.2.2 Effect of stacking on leakage current

The total leakage current for any gate in a CMOS transistor primarily depends on the different input vectors [6-8]. For a 2-input NAND gate, there are four possible combination of input states. The leakage current is significant for '11', '01' and '10' states compared to

'00' state. These three states are considered as worst leakage states, among which states '01' and '10' produces lower leakage than state '11'. On applying reverse body bias voltage V_{bs} (= -0.5v) to these two states, an effective reduction in leakage current is observed. However, to reduce the leakage of state '11' to a considerable level, reverse body bias voltage should be increased. With increase in body bias voltage, the short channel effect comes into consideration [10]. Thus, for this state stacking effect of transistor is applied in which 2-input NAND gate is replaced by a 3-input NAND gate that incorporates an extra *SLEEP* signal. While applying the V_{bs} = -0.5v along with stacking effect, the leakage current reduces to a great extend. The variation of leakage with different input vectors for 2-input and 3-input NAND gate is shown in Tables 7 and 8 respectively. The extra input is referred as *SLEEP* as '0' and '1' in standby and active modes, respectively. As shown in Tables 4.7 and 4.8leakage current reduces in standby mode and on application of reverse bias voltage, leakage further reduces to a maximum extend.

Table 4.7 Leakage current for 2-input NAND Gate at 45nm technology node

Input states	Leakage current (nA)
00	0.786
01	13.262
10	4.850
11	17.161

Table 4.8 Leakage current for 3-input NAND Gate in Standby and Active mode at 45nm Technology Node

Input states	Leakage current (nA) in standby mode	Leakage current (nA) in active mode	Leakage current (nA) in standby mode when Vbs = -0.5v
00	0.954	1.201	0.423
01	0.734	20.394	0.405
10	0.706	7.462	0.244
11	6.512	25.737	0.213

4.3.2.3 Effect of leakage on benchmark circuits

This sub section analyzes the leakage reduction technique employing stacking effect and reverse body biasing approach for nanometer CMOS VLSI circuits. Different circuits of ISCAS and ITC'99 series [24-25] are used to find the minimum leakage current as shown in Fig.4.8. It is observed that the overall leakage of benchmark circuits at 45nm technology

node is reduced by 21.5% and 53.27% by RBB variation and by stacking effect using the RBB variation respectively as presented in Table 4.9. Moreover, in comparison to the RBB technique, the proposed algorithm provides an overall reduction in leakage of 40.67%.

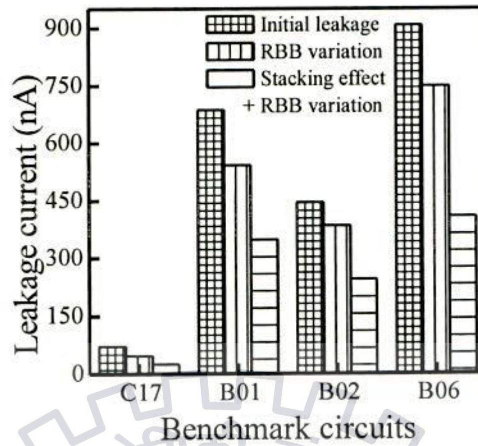


Fig. 4.8 Leakage current for different benchmark circuits by RBB variation and stacking effect using RBB variation

Table 4.9 Leakage Current for Different Benchmark Circuits by Reverse Body Bias (RBB) and Stacking Effect using RBB variation

Benchmark circuits	Initial leakage (nA)	Final leakage (nA)				Overall % reduction using both the approaches
		By RBB variation		By stacking effect using RBB variation		
		Leakage current (nA)	% reduction	Leakage current (nA)	% reduction	
C17	71.468	47.124	34.06	25.795	63.9	45.20
B01	686.139	542.03	21.01	348.06	49.3	35.78
B02	445.003	383.975	13.71	244.947	44.9	36.20
B06	907.469	748.942	17.47	408.297	55.0	45.49

This research work introduced a modified gate replacement algorithm that uses the concept of reverse body bias variation along with the stacking effect of transistors. Using this approach, one can reduce the number of replacements along with the lesser probability of short channel effects as compared to the gate replacement algorithm and the RBB variation technique. Moreover, the approach is applied to different ISCAS and ITC'99 [24, 25] benchmark circuits that exhibit an overall reduction in leakage current by 21.5% and 53.27% by RBB variation and by the proposed algorithm that utilizes the stacking effect along with RBB variation at 45nm technology node respectively. In comparison to RBB variation, this novel technique provides 40.67% of overall leakage reduction.

Chapter 5

Conclusion and Future Scope

With continuous scaling in the device size, the leakage current can be identified as a main contributor to the total power consumption of CMOS VLSI circuits. In current nanometer devices, gate and sub threshold leakage dominates other leakages such as reverse-biased *pn* junction leakage, band-to-band tunneling (BTBT) leakage, gate leakage due to hot carrier injection, gate induced drain leakage, and channel punch-through leakage. The research work primarily focuses on the reduction of the two major dominating leakages referred as subthreshold and gate leakage. It has been observed that on increasing number of transistors that are in 'off' state in a stack, the leakage current reduces effectively. Based on stacking effect of transistors the gates at worst leakage state are replaced by another gate incorporating an extra sleep signal. The replacement of gate at WLS results in reduced leakage in standby mode while maintaining device functionality in active mode. The drawback of this approach is that on replacement output of that gate may change. In this case if fan out of replaced gate is higher, the leakages of all fanout gates will be affected. This can lead to increase in overall circuit leakage. On further replacement of fan out gates, number of replacements increases. To reduce the number of replacements, the modified gate replacement technique is proposed which skip the gates with high fan out which on replacement increases overall circuit leakage and area. These skipped gates are further considered for replacement followed by AND/OR gate. But this approach becomes ineffective below 45nm technology node as gate leakage dominates other leakages. We have applied stacking effect along with the dual-T_{ox} approach which effectively reduces leakage of circuit. Furthermore, we have applied gate replacement algorithm from inputs towards outputs and vice versa. We have observed that there is more reduction in leakage in later case because gate leakage depends on fanin of a gate too. Hierarchical replacements from outputs to inputs results are a better choice of gate replacements. After this, leakage current of a circuit can be further reduced using pin reordering technique. This approach can be modified taking both leakage and delay as a reduction parameter and trade off can be achieved. We have applied this approach on C17 benchmark circuit at 32nm technology node. There is 37.8% reduction in leakage with slight delay penalty.

The subthreshold leakage effectively reduces using stacking effect of transistors. This approach results in reduced leakage current at the cost of increasing amount of area and the

delay of circuit. To overcome this problem, it is preferable to use the RBB technique that increases the V_{th} of transistors in standby mode and reduces the subthreshold leakage. On the other hand, with suitable choice of different input vectors, gate leakage can be effectively reduced. In comparison to the subthreshold and gate leakage, the BTBT leakage current can be ignored for the 45nm technology nodes. However, the reverse bias voltage variation increases the BTBT leakage current. Consequently, the applied reverse bias voltage is cautiously chosen that results in reduction of leakage without any significant increase in BTBT leakage current. Using the gate replacement along with the reverse body bias technique, a new approach for leakage reduction is applied by considering the increase in area and short channel effects in CMOS VLSI circuits.It is observed that the overall leakage of benchmark circuits at 45nm technology node is reduced by 21.5% and 53.27% by RBB variation and by stacking effect using the RBB variation.

As part of the future work, we can suggest the follwing improvements:

- This modified technique can be applied for lower technology nodes below 32nm. In this technology node, the BTBT leakage has a major impact alongwith the subthreshold and gate leakages.
- This approach can be applied to ISCAS benchmark circuits to analyze the leakage mechanisms in details.
- These approach can be further modified considering delay and area as critical parameters in VLSI design.

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
SHRAMAM VINA NA KIMAPI SADHYAM

REFERENCES

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

[2] S. Mukhopadhyay and K. Roy, "Accurate Modeling of Transistor Stacks to Effectively Reduce Total Standby Leakage in Nano-Scale CMOS Circuits," in *Proc. IEEE Int. Symp. VLSI Circuit Digest,* pp. 53-56, Jun. 2003.

[3] S. Yang, W. Wolf, N. Vijaykrishnan, Y. Xie, and W. Wang, "Accurate Stacking Effect Macromodelling of Leakage power in Sub-100nm Circuits," in *proc. IEEE 18th Int. conf. on VLSI Design,* pp. 165-170, Jan. 2005.

[4] Z. Chen, M. Johnson, L. Wei, and K. Roy, "Estimation of standby leakage power in CMOS circuits considering accurate modeling of transistor stacks" in *Proc. IEEE Int. Symp. Low Power Electronics and Design (ISLPED 1998),* pp. 239-244, Aug. 1998.

[5] R. Kumar and C. P. Ravikumar, "Leakage power estimation for deep submicron circuits in an ASIC design environment," in *Proc. IEEEDesign Automation Conference on VLSI Design (ASP-DAC 2002),* pp. 45 – 50, Jan. 2002

[6] A. Abdollahi, F. Fallah, and M. Pedram, "Leakage current reduction in CMOS VLSI circuits by input vector control," *IEEE Trans.Very Large Scale Integration Systems (VLSI),* vol. 12, no. 2, pp.140-154, Feb. 2004.

[7] L. Yuan and G. Qu, "A Combined Gate Replacement and Input Vector Control Approach for Leakage Current Reduction," *IEEE Trans. Very Large Integration (VLSI) systems,* vol.14, no.2, pp. 173-182, Feb. 2006.

[8] N. Jayakumar and S. P. Khatri, "A Simultaneous Input Vector Control and Circuit Modification Technique to Reduce Leakage with Zero Delay Penalty," *ACM Transactions on Design Automation of Electronic Systems (TODAES),* vol. 16, no. 1, pp. 9:1-9:20, Nov. 2010.

[9] S. Singh, B. Kaur, B. K. Kaushik, and S. Dasgupta, "Leakage Current Reduction using Modified Gate Replacement Technique for CMOS VLSI Circuit," in *Proc. IEEE CODIS 2012*, pp. 464-467 , Dec. 2012.

[10] F. Gao and J. P. Hayes, "Exact and heuristic approaches to input vector control for leakage power reduction" *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems,* vol. 25, no. 11,pp. 2561-2574, Nov. 2006.

[11] K. Kim, Y-Bin Kim, M. Choi, and N. Park, "Leakage Minimization Technique For Nano scale CMOS VLSI Based On Macro-Cell Modeling" in *proc. 19th IEEE Int. conf. on VLSI Design*, pp. 53-56, Jan. 2005.

[12] C.-H. Choi, Z. Yu, and R.W. Dutton, "Impact of gate direct tunneling on circuit performace: A simulation study," *IEEE Trans. Electron Devices*,vol. 47, no. 12, pp. 2823–2829, Dec. 2001.

[13] R. S. Guindi and F. N. Najm, "Design Techniques for Gate-Leakage Reduction in CMOS Circuits," in *Proc. 4th Int.Symp. on Quality Electronic Design (ISQED)*, pp. 24–26, Mar. 2003.

[14] N. Sirisantana, L. Wei, and K. Roy, "High-performace low-power CMOS circuits using multiple channel length and multiple oxide thickness," in *Proc. IEEE Int. Conf. Computer Design*, Sep. 2000, pp.227–232.

[15] K. A. Bowman, L. Wang, X. Tang, and J. D. Meindl, "A circuit-level perspective of the optimum gate oxide thickness," *IEEE Trans. Electron Devices*, vol. 48, no. 8, pp. 1800–1810, Aug. 2001.

[16] D. Lee, W. Cong, D. Blaauw, and D. Sylvester, "Analysis and minimization techniques for total leakage considering gate oxide leakage," in *Proc. ACM/IEEE Design Automation Conf.*, pp. 175–180,Jun. 2003.

[17] H. Jeon, Y. Kim Y, M. Choi, "Standby Leakage Power Reduction Technique for Nanoscale CMOS VLSI Systems," *IEEE Trans. Instrumentation and Measurements*,vol. 59, pp. 1127-1133, Jan. 2010.

[18] J. Yang, Y. Kim, "Self adaptive body biasing scheme for leakage power reduction in nanoscale CMOS circuit," in *Proc. Great Lakes symposium on VLSI (GLSVLSI'12)*, pp. 111-115, May 2012.

[19] R. Hossain, M. Zheng, and A. Albicki, "Reducing power dissipation in CMOS circuits by signal probability based transistor reordering," *IEEETrans. Computer-Aided Des. Integr. Circuits Syst.*, vol. 15, pp. 361–368,Mar. 1996.

[20] E. Musoll and J. Cortadella, "Optimizing CMOS circuits for low power using transistor reordering," in *Proc. Eur. Design Test Conf.*, Mar. 1996, pp. 219–223

[21] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits*, Prentice Hall, New Jersey, USA, 2002.

[22] S. –M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits Analysis and Design*, Tata Mc-graw Hill publishers, 3rd Edition, 2002.

[23] *http://www.ece.uci.edu/docs/hspice/hspice_2001_2-173.html*.

[24] *http://www.ece.uic.edu/~masud/resources.html*.

[25] *http://www.cad.polito.it/downloads/tools/benchmarks.html*.

LIST OF PUBLICATIONS

1. S. Singh, B. Kaur, B. K. Kaushik, and S. Dasgupta,"Leakage Current Reduction using Modified Gate Replacement Technique for CMOS VLSI Circuit," in *Proc. IEEE CODIS 2012*, Kolkata, pp. 464-467 , Dec. 2012.
2. S. Singh, B. K. Kaushik, and S. Dasgupta, "A Modified Gate Replacement Algorithm for Leakage Reduction using Dual-T_{ox} in CMOS VLSI circuits,"*17th VLSI Design and Test Symposium 2013 (VDAT2013),*Jaipur,Jul. 27-30,2013 (Accepted).
3. S. Singh, B. K. Kaushik, and S. Dasgupta, "Stacking Effect using Reverse Body Bias for Leakage Reduction in CMOS VLSI circuits," *IEEE International conference on advanced electronic systems (ICAES2013),* Pilani, Sep. 21-23, 2013 (Communicated).