POWER REDUCTION IN MODERN VLSI CIRCUITS – A REVIEW

Debopam Ghosh, Jyotirmoy Guha, Arka De, Anirban Mukherjee
Electronics and Communication Engineering
Heritage Institute of Technology Kolkata, India
debopamghosh2010@gmail.com, jyotirmoy1991@gmail.com, korebio@gmail.com
anirban.mukherjee24@gmail.com

Abstract

This paper focuses on the importance of reducing power consumption by VLSI circuits using different power reduction techniques. In this paper, various sources of power dissipation in modern VLSI circuits along with various power reduction techniques as adopted in industry today are discussed.

1. INTRODUCTION

With ever increasing demand of System On Chip (SOC), more and more transistors are getting added in modern VLSI chip to incorporate new architectural features leading to the risk of very high power dissipation in a die [1][2]. On the other hand, nowadays, customers look for portable, handheld devices (like smart phone, tablet PC) having high battery life time. Therefore, low power dissipation by VLSI circuits has become a key design parameter. Considering energy crisis being the utmost concern in modern world, manufacturers cannot afford to allow excessive power leakage from a single chip. More power dissipation might unnecessarily heat up the devices thereby reducing performance, reliability and durability. Hence, the need for low power VLSI circuit techniques arises. The world now needs ways to efficiently reduce the dissipation of huge amount of power from the circuits while keeping pace with ever increasing demands of more and more features in a single die. In the following sections, the different ways in which power leaks from VLSI circuits and different techniques of how power dissipation is reduced in industry is discussed.

2. POWER DISSIPATION IN VLSI CIRCUITS

Power dissipation in VLSI circuits is of three types [1][3]:-

i) **Dynamic power** (≈ 40 - 70% of die power) ii**) Leakage power** (≈ 20 – 50 % of die power and increasing) iii**) Short- circuit power** (≈ 10 % today)

2.1 DYNAMIC POWER

CMOS digital circuits dissipate power while charging various

load capacitances (C_L) (receiving gate capacitance, wire capacitance and drain capacitance of driver gate) whenever they are switched (Fig 1). In one complete cycle of CMOS logic, current flows from V_{DD} to the load capacitance to charge it and then flows from the charged load capacitance to ground during discharge. The switching power [4] dissipated by a CMOS gate is given by:

$$P = C_L V_{DD}^2 f . \text{--------(1)}$$

where, C_L-Load Capacitance, V_{DD} – Supply Voltage,

f – Frequency of switching

Since most gates do not operate/switch at every clock cycle, they are often accompanied by a factor (α), called the activity factor. Hence, the dynamic power dissipation can be re- written as:

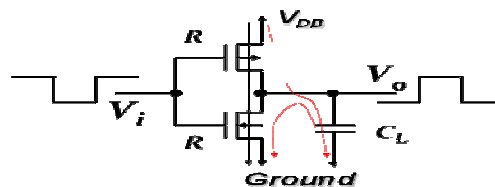$$P = \alpha C_L V_{DD}^2 f \text{--------------------(2)}$$

Fig 1: Dynamic power

Thus dynamic power dissipation depends on four parameters namely Load Capacitance (C_L), Supply Voltage (V_{DD}), Activity factor (α) and frequency (f). Their impacts and way of reducing power dissipation are discussed below.

i. Load Capacitance: In modern SOC applications, large load capacitance may appear due to high fanout net or long wire distance between driver and receiver gate resulting in increase in dynamic power dissipation.

Reduction Technique: To improve power performance downsizing driver and receiver gate is popular technique. By downsizing a transistor, the design can be more compact. Proper placement and routing is also necessary so that the length of the wire can be minimized to reduce load capacitance.
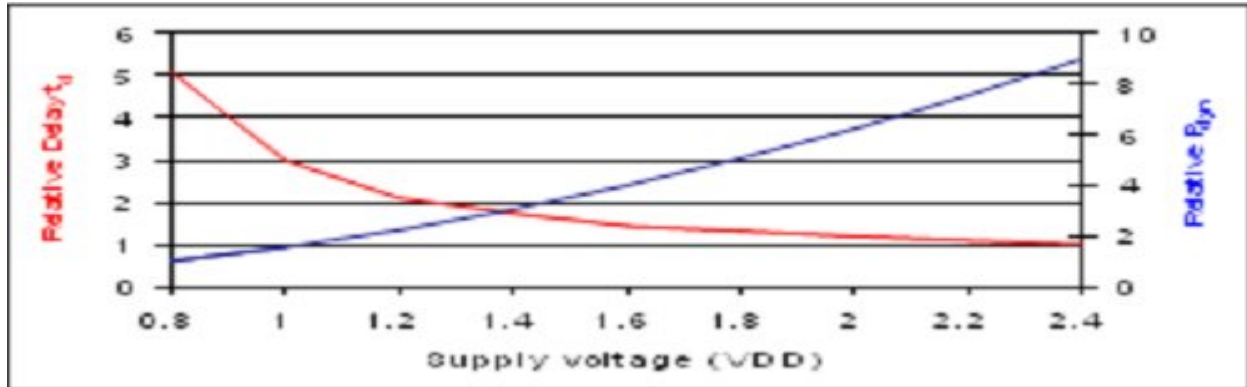
ii. Supply Voltage:-

Fig 2: Relationship between supply voltage, delay and dynamic power performance [5]

It is observed that with the increase of supply voltage the gate delay decreases but dynamic power increases (fig. 2).

Reduction Technique :- To strike a balance between speed and power, the concept of multiple VDD in Data Path comes into picture. In Data Path there are alternate combinations of Sequential and Combinational circuits. Frequency of operation depends on the path with largest delay. Data path can be broadly classified into two categories **Timing Critical** and **non-Timing Critical.** Here, **Critical path means the path having largest time delay**.

To reduce the dynamic power dissipation, a Timing Critical path is fed with High supply voltage so that it doesn't suffer from large delay and the non critical paths are fed with Low supply voltage. It is observed that in CPU and DSP type circuits as used in industry, 80% of the paths are non-timing critical, hence most of the paths can be supplied with low voltage as a result of which dissipation due to dynamic power can be reduced.

iii. Activity factor: The toggling rate of a signal with respect to clock is termed as Activity factor (α). Dynamic power dissipation is directly proportional to α.

Reduction Technique: To reduce Activity factor, most popular technique used in industry is **Clock Gating** [5][6]. In this method, the clock signal to different functional block is gated by enabling signals. In most cases, AND gates are preferred for gating clock signals.
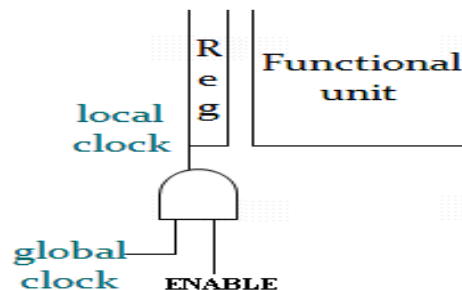

Fig 3: Clock Gating

The above diagram (Fig 3) illustrates the concept of clock gating. When ENABLE=1, clock can be applied to functional block through AND gate. However to avoid unnecessary switching when the functional block is inactive, the ENABLE input can be made 0 so as to disable the clock propagation when not required. This causes the clock to become totally inactive when the functional block is in non-active mode, thus eliminating any sort of activity in the circuit. Clock Gating can reduce power leak due to dynamic power by as high as 70 to 80%.

2.2 LEAKAGE POWER

Leakage power (Fig 4) [7], $P_{leakage}$ is the static power consumption when circuit is not switching. It is observed that in deep sub micron technology, leakage contributes up to 50%
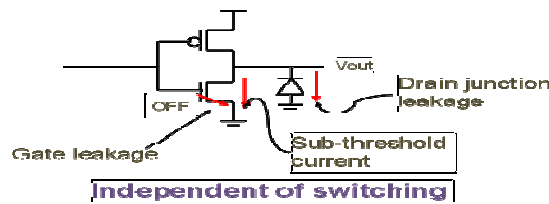
of the total power consumption.
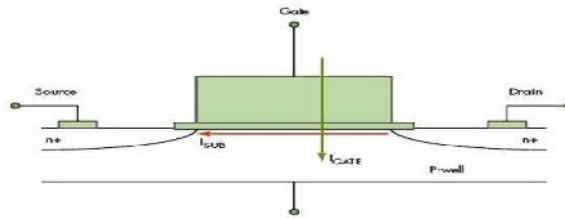

Fig 4: Leakage power

Fig 5: Subthreshold and Gate Leakage [8]

There are two major components of leakage power in deep sub micron technology:

i. Subthreshold leakage [8] (indicated by red arrow) (Fig 5)

is weak inversion conduction current between source and

drain in a MOS transistor that occurs when the gate voltage in below Threshold voltage (V_T). Subthreshold leakage is defined as:

$I=I_0*exp((V_{GS}-V_T)/ \eta V_T)*(1-exp(-V_{DS}/\eta V_T)$ ----------(3) where, **I_0 α $\mu*C_{ox}*(W/L)$, η**=ideality factor, **V_{DS}**=Drain to source voltage, **V_{GS}**= Gate to source Voltage, **V_T**= Threshold voltage, **μ**= mobility of electrons/holes, **C_{ox}**= oxide capacitance, **W**= width of the transistor, **L**=channel length.

In deep sub micron technology, sub threshold leakage can be as high as 80-90% of total circuit leakage

power.

ii. Gate oxide leakage (indicated by green arrow) (Fig 5) is the oxide tunneling current due to the low oxide thickness and high electric field.

Leakage Reduction Techniques:

i. Variation of the Supply Voltage(V_{DD}):

Delay in VLSI circuits can be expressed as:

$\tau_{PHL}= C_{load}/K_n(V_{DD}-V_{T,n})*[2V_{t,n}/(V_{DD}-V_{T,n})+ln(4*(V_{DD}- V_{T,n})/V_{DD}-1)]$ -------------------------------(4)

$\tau_{PLH}= C_{load}/K_p(V_{DD}-V_{T,p})*[2V_{t,p}/(V_{DD}-V_{T,p})+ln(4*(V_{DD}- V_{T,p})/V_{DD}-1)]$ -------------------------------(5)

where, τ_{PHL} and τ_{PLH} are high to low and low to high delay respectively, $k_n=\mu_n*C_{ox}*(W/L)$, $k_p= \mu_p*C_{ox}*(W/L)$,

C_{load}=load capacitance, $V_{t,n}$ and $V_{t,p}$ are the threshold voltages on NMOS and PMOS respectively, μ_n = mobility of electrons,

μ_p = mobility of holes, C_{ox}= oxide capacitance, W= width of the transistor, L=channel length.

From expressions of dynamic power and sub-threshold leakage power, it is clear that these powers can be reduced significantly by lowering V_{DD}. However from the delay expression it is evident that reduction in V_{DD} increases delay which affects performance as it reduces switching speed. Hence a balance has to be maintained between speed and power while selecting V_{DD} for new Technology and Design.

ii. Variation of Threshold Voltage (V_T):

There are two significant techniques of modifying threshold voltage-

 a. **Multiple threshold CMOS (VTCMOS)** ,

b. V**ariable Threshold CMOS (MTCMOS)**.

a.   Multiple Threshold CMOS circuit (MTCMOS): In MTCMOS technique [8][9][10], MOS transistors with low threshold voltage (V_T) are placed in the critical path while MOS transistors with high threshold voltage (V_T), are placed in the non-critical path (Fig 6). The objective of such a circuit design is to maximize performance and minimize leakage. In the critical path gate delay time should be as low as possible. Hence delay reduction is traded off against leakage by placing low threshold transistors causing faster gate transitions. However, in the non timing critical path, high threshold transistors are placed to reduce leakage as speed/delay has less importance here. It is observed that in modern VLSI circuits as used in industry, 80% of the paths are non-timing critical. Hence for most of the paths, high threshold

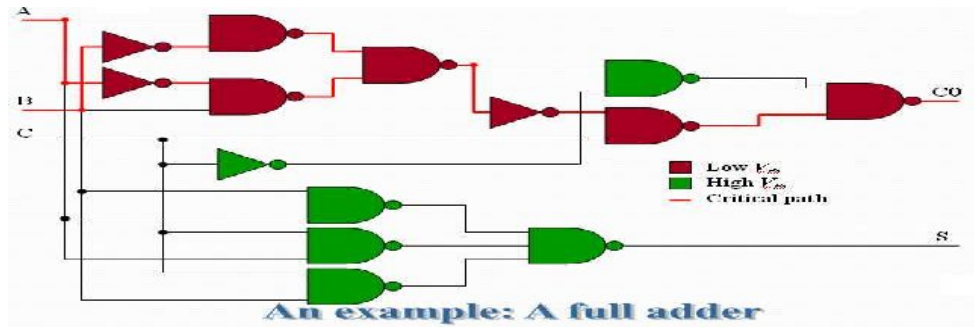transistors can be used to enable aggressive leakage reduction.

Fig 6: MTCMOS Circuit

b. Variable Threshold CMOS circuit (VTCMOS): The equation for Threshold voltage (**V$_T$**) is defined as:

V$_T$ =V$_{T0}$ + γ(√(|2Ø$_F$|+V$_{SB}$)- √(|2Ø$_F$|)) ------------ (6)

where **γ = (t$_{ox}$/ε$_{ox}$) √(2qε$_{si}$N$_A$) ----------------- (7)** V$_{T0}$ =Threshold voltage with no body effect, Ø$_F$= Fermi potential, **V$_{SB}$**= source to body voltage/bias, **t$_{ox}$** = Thickness of the oxide, **ε$_{ox}$**= permittivity of the oxide, **q**= charge of electron/hole, **ε$_{si}$**= permittivity of silicon, **N$_A$**= Channel doping concentration.

In VTCMOS technique [8][10] the threshold voltage is varied by a suitable control circuit (Fig 7), according to the circuit requirements. As is evident from the expression above, threshold voltage is a function of the source to body bias. Hence variation of this bias voltage using suitable circuitry can vary threshold voltage. Practically for critical paths threshold voltage is lowered as switching speed is of primary concern for such paths. But for non-critical paths threshold voltage is increased to reduce the leakage power dissipation, as higher threshold prevents any sort of unwanted leakage power flow.
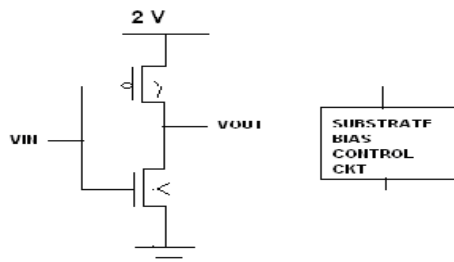

Fig 7: VTCMOS Technique

In active mode when the circuit is switching (Fig 7), the following substrate bias voltages V$_{BP}$=2V and

V_{BN}=0V are produced for PMOS and NMOS respectively. Due to this, the circuit has high switching speed (due to low V_T) at the cost of leakage power. In standby mode the control circuit generates a lower substrate voltage for NMOS and higher substrate voltage for PMOS .This causes the threshold voltage of both the NMOS and PMOS to increase, thus reducing the leakage current. Besides changing the supply voltage (V_{DD}) and the threshold voltage (V_T), additional techniques to reduce leakage power as adopted in industry are

iii. Aggressive device size reduction on non-timing critical path as well as

iv. Enabling 'Stacking Effect' in CMOS gate during standby mode. When 2 or more transistors in series are OFF, sub threshold leakage reduces exponentially known as 'Stacking Effect'.

v. Gate Leakage Prevention:

In deep sub-micron technology, gate oxide thickness reached 10-20A^0 causing exponential increase in tunneling current through extremely thin gate oxide. In 45nm technology, gate leakage got increased by 10 times with respect to previous process technology. Hence, thicker insulator is used to reduce gate leakage. Thicker insulator (increase in t_{ox}) reduces Cox, i.e the gate oxide capacitance per unit area = (**K/t_{ox}**, where **K**= permitivitty of gate oxide) causing increase in threshold voltage (**V_T**) and thus reducing device speed. Hence, to keep Cox constant, different insulating materials having higher permittivity than SiO$_2$ is used from 45nm technology onwards. Since insulating material is changed, polysilicon gate is also replaced by new metal gate to maintain low work function.

2.3. SHORT CIRCUIT POWER

The component of power caused due to a leakage current flowing between V_{DD} and ground is called short circuit power[7][11]. In a CMOS gate, nMOS conducts when

$V_{GS}>V_{T, n}$ and pMOS conducts when $V_{GS}>|V_{T, p}|$. But as the input voltage makes a transition from logic 0 to logic 1, ideally pMOS should be turned OFF and nMOS should turn ON. Practically pMOS remains ON till $V_{GS}=V_{DD}-|V_{T, p}|$. Hence, during the period from $V_{T, n}$ to $V_{DD}-|V_{T, p}|$, both the transistors remain ON, thus, providing a short circuit path from V_{DD} to ground. A similar scenario is observed when the input voltage is making a negative transition (Fig 7). Assuming rise and fall time to be equal the short circuit current (I_{SC}) and Power (P_{SC}) can be expressed as:

I_{sc}=1/12 ß *τ*f_{clk}/V_{DD}*(V_{DD}-2V_T)^3 ----------------(8)

P_{sc}=V_{DD}*I_{sc} = (ß/12)*(V_{DD}-2V_T)^3*(t_{rf}/t_p) ---------(9)

Where **τ, t_{rf}** = input rise/fall time, **t_p**=Waveform period,

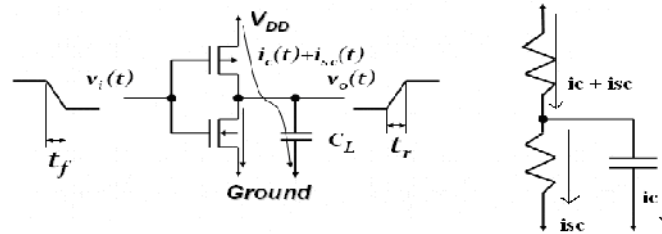ß=Gain factor, **f$_{clk}$**=clock, **V$_T$**=Threshold Voltage

Fig 8: Short Circuit Power

Reduction Techniques:

The short circuit power is reduced by **i.** device size reduction and **ii.** increasing threshold voltage both of which is also used for leakage reduction. In addition, the input rise and fall times for a digital gate needs to be less than an upper limit which can be taken care of during gate sizing. It is important to note that short circuit power dissipation happens only during switching (Dynamic power dissipation). Hence reduction of Activity factor (α) helps both power reductions simultaneously.

3.   CONCLUSION

The above discussions provide an insight into the different types of power dissipation encountered in any VLSI circuit and the various power reduction techniques as adopted in industry. Modern VLSI circuits are integrated with a separate power management unit whose function is to manage the power distribution in the die efficiently. There is a need for industry to develop more of such methods to maintain chip performance while continuously improving power consumption.

4.   REFERENCES

I.   Venkat Rao, Gaurav Singhal *Power Issues In Vlsi Design*

II.   Massoud Pedram DESIGN TECHNOLOGIES FOR LOW POWER VLSI

III.   www.ee.ncu.edu.tw/~jfli/vlsi21/lecture/ch04.pdf by L Benini and G De Micheli

IV.   http://www.ee.ncu.edu.tw/~jfli/vlsi21/lecture/ch04.pdf by Jin-Fu Li; Advanced Reliable Systems (ARES) Lab. D

V.   epartment of Electrical Engineering, National National Central University Central University, Jhongli, Taiwan

VI.   MICRO TRANSDUCTORS '08 OW POWER VLSI DESIGN 2 by Dr.-Ing. Frank Sill; Department of Electrical Engineering, Federal University of Minas Gerais, Av. Antônio Carlos 6627, CEP: 31270-010, Belo Horizonte (MG), Brazil franksill@ufmg.br, ttp://www.cpdee.ufmg.br/~frank/

VII. LOW POWER DESIGN: CLOCK GATING by M. Ohashi, Matsushita

VIII. LOW-POWER MULTI-THRESHOLD CMOS CIRCUITS OPTIMIZATION AND CAD TOOL DESIGN

IX.   A Thesis Presented to The Faculty of Graduate Studies of The University of Guelph by WENXIN WANG In partial fulfillment of requirements for the degree of Master of Science May, 2004 ,Wenxin Wang, 2004 Advisors: Professor Shawki Areibi, Mohab Anis

X.   LEAKAGE POWER REDUCTION IN CMOS VLSI CIRCUITS BY INPUT VECTOR CONTROL by Afshin Abdollahi, Farzan Fallah, and Massoud Pedram IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, VOL. 12, NO. 2, FEBRUARY 2004

XI.   TRANSISTOR AND CIRCUIT DESIGN OPTIMIZATION FOR LOW-POWER CMOS By M.Chang, C.Chang,C.Chao, K.Goto,M.Leong, L.Lu,and C.Diaz Presented By Mozammel Haque For the Low-Power High- Sped VLSI ELEC 5705Y W-2009

XII. CMOS DIGITAL INTEGRATED CIRCUITS by Sung- Mo Kang and Yusuf Lablebici Tata McGraw-Hill Education, 2003

XIII. SHORT CIRCUIT POWER REDUCTION BY USING HIGH THRESHOLD TRANSISTORS by Arkadiy Morgenshtein Received: 15 October 2011; in revised form: 13 February 2012 / Accepted: 21 February 2012 /Published: 1 March 2012 Journal of Low Power Electronics and Applications ISSN 2079-9268 www.mdpi.com/journal/jlpea/