Power Dissipation in VLSI Circuits and Low Power VLSI Design Strategies –A Review

ManukrishnaV R, Assistant Professor, ECE Department, Rajadhani Institute of Engineering and Technology, Attingal, Kerala
manukrishnavr@gmail.com

Abstract—Earlier times VLSI designers focused on increasing the speed and reducing the power of the system. However the growing market for battery powered portable electronic systems demands the design of microelectronic circuits with low power dissipation. Low power reduces the cooling cost and increases the reliability for high density systems. It also reduces the weight and size of the portable system. Low power has emerged as a principal theme in today's electronics industry. The need for low power has caused a major paradigm shift where power dissipation has become as important a consideration as performance and area. This article reviews various strategies and methodologies for designing low power circuits and systems. It describes the many issues facing designers at architectural, logic, circuit and device levels and presents some of the techniques that have been proposed to overcome these difficulties .This paper presents some most trustful approaches to low power design. The aim of this paper is to provide an insight to the design ,when power consumption is a major constraint. This paper focus on the CMOS based digital circuits.

Keywords— Portable electronic devices; CMOS based digital circuits; Existing techniques; Advanced techniques .

I. INTRODUCTION

During the last five decades the electronics industry has evolved tremendously, and the last ten years of aggressive scaling have moved integrated circuits from the micrometer regime down to the nanoscale regime . In the late 1950s, putting more than one transistor on a piece of semiconductor device was considered cutting edge. The concept of integrated circuits with even as little as tens of devices was unheard of. To obtain a 50% probability of functionality for a 20-transistor circuit, the probability of individual device functionality had to be $(0.5)^{1/20} = 96.6\%$, which was considered optimistic well beyond anything imaginable. Nevertheless, ongoing innovations in technology and integration have continued to overcome the predicted limits,and today transistors are manufactured with gate lengths well below 100 nm, and integrated circuits contains over a billion transistors per chip .

In 1965 Gordon Moore published his famous paper , in which he predicted that the number of components per integrated circuit for minimum cost would increase by two every year. This prediction was updated ten years later, predicting that the number of devices should double every second year from then on, which is popularly refereed to as "Moore's Law" . These predictions have since then inspired the microelectronic industry to strive for increased complexity and lower fabrication costs of integrated circuits. Up until now Moore's predictions have been quite accurate, as a result of vast improvements in circuit capabilities, enabled by dimensional scaling. This can be illustrated in the form of the microprocessor evolution in the last four decades seen in Figure1. From the first Intel® 4004 microprocessor (Figure 2(a)) with 2300 transistors clocked at a frequency of 108 kHz to the present Core™ 2 Quad (Figure 2(b)) with 820 million transistors and clocked at frequencies above 3 GHz, the number of transistors has roughly doubled every two years.

This article reviews the existing low power reduction techniques and suggests low power design techniques begin from process stage to circuit stage . The last section discuss the the use of Carbon Nano Tubes as interconnects for sub-threshold VLSI devices which is one of the most advanced trends in low power VLSI design.

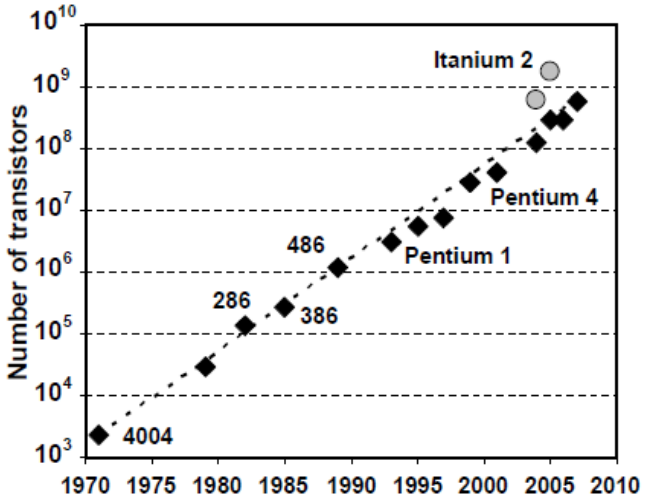

Figure 1. 40 years of evolution of Intel Microprocessor

Figure 2.Intel® 4004 in 10 µm

Figure 3 . Intel® 2 Quad in 45 nm

II. POWER DISSIPATION IN CMOS CIRCUITS

There are two types of power dissipation in CMOS circuits

A. DYNAMIC POWER DISSIPATION

Dynamic power is primarily caused by the current flow from the charging and discharging capacitance.[8] It consists three components; switching power, short circuit power and glitching power. The value of each of this component is a function of logic style used and topology of the circuit. These components are mentioned below.

1) Switching Power Dissipation

Switching power is defined as the power consumed as the power consumed by the logic circuit to charge from low level "0" to high level "1". In a well designed circuit switching power is a very dominant circuit. Usually, it is independent of the logic function of the circuit.

$$P_{switching} = F_{switching} \cdot V_{dd}^2 \cdot C_L \qquad (1)$$

Where $F_{switching}$ is the switching frequency, V_{dd} is the supply voltage and C_L is the net node capacitance. The loading capacitance C_L consists of net gate capacitances of subsequent gates inputs; interconnect capacitance and diffusion capacitances of drains of inverter transistors. Tests have shown that the total capacitance have split almost equally between these three. As the minimum gate length scales down the interconnect capacitance become dominant. Fig. 4 shows basic capacitive elements of an inverter.

Equation 1 indicates that supply voltage is a dominant factor in switching power dissipation. Thus reducing supply voltage is the most effective technique in reducing power dissipation. This equation is valid only for logic families with rail to rail output. Some logic families like CPL (Complementary Pass logic) has reduced voltage swing.

Switching power can be expressed as

$$P_{switching} = F_{switching} \cdot V_{dd} \cdot V_{swing} \cdot C_L \qquad (2)$$

Where V_{swing} is the logic swing of digital output.

The switching frequency is given as

$$F_{switching} = F_{operating} \cdot \alpha \qquad (3)$$

α is the switching activity factor. This factor determines the probability of having a $0 \rightarrow 1$ transition at the output.

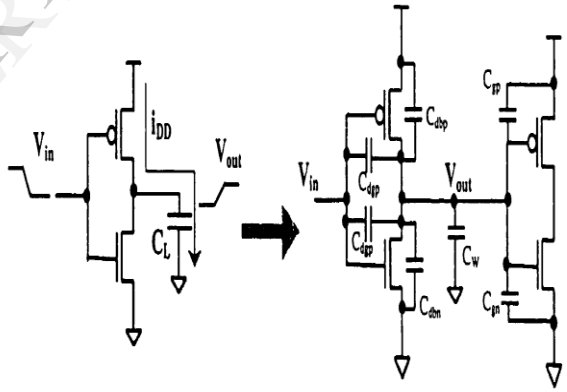

Figure 4.It shows the basic capacitive elements of an inverter

2) Short circuit Power Dissipation

It is the power transmitted from supply to ground during the transition from logic 0 to logic1 and from logic 1 to logic 0. Unlike switching power it is a function of number of 0 to 1 transitions, the short circuit power is a function of toggling frequencies .During this transition both PMOS and NMOS Transistor become ON creating a short circuit path from V_{dd} to ground. The figure below the output voltage of a standard CMOS at short circuit current.

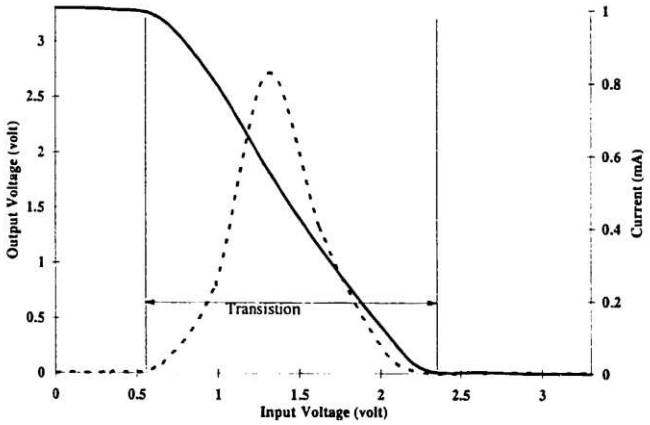

Figure 5. Output voltage of a standard CMOS at short circuit current

The short circuit power can be calculated as follows:

$$P_{sc} = I_{sc}.V_{dd} \quad (4)$$

Where I_{sc} is the average short circuit current.

3) Glitching power dissipation

Glitching power is the power dissipated in intermediate transitions during the evaluation of logic function of the circuit. When the inputs to the logic gates are not synchronized, erroneous output occur at the output node until all inputs settle down to final values. These intermediate erroneous outputs lead to a power loss in charging and discharging the output node capacitance. It is difficult the glitching power because it is a function of topology of the circuit, layout, delay, previous inputs, new inputs and gate type. Glitching power dissipation can be expressed as follows:

$$P_{Glitch} = V^2_{dd}.C_L.F_{Glitch} \quad (5)$$

Where F_{Glitch} is the average frequency of glitches. Glitching power consumed 40% of the total power dissipation of the circuit, especially architectures with large logic depth.

To avoid this power loss designers must use synchronous circuits in which all the outputs are either latched or gated to synchronize the inputs to the next stage. Dynamic circuits also avoid this problem by synchronizing the circuit with the clock signal.

B. STATIC POWER DISSIPATION

Static power is usually a small portion of the total power dissipation.[8] Unfortunately as the threshold voltage V_{th} decreases and the number of transistors per chip increases, the static power dissipation become more important. In digital circuits there are three main sources of static power dissipation: Diode leakage current, sub threshold current and biasing current.

1) Diode Leakage Current

Diode leakage current occurs, when one transistor is turned OFF and another transistor charges up/down the drain with respect to the former bulk potential. In the case of an inverter with high input voltage, the output become "0", because the NMOS transistor is ON. The PMOS transistor is turned OFF ,when its drain to bulk voltage is equal to the supply voltage ($-V_{dd}$).The resulting diode leakage current is approximately:

$$I_L = A_d \, . \, J_s \quad (6)$$

Where Ad is the area of drain diffusion and Js is the leakage current density set by the technology. Since the diode reaches the maximum reverse bias current for a relatively small reverse biased potential, the leakage current is roughly independent of the supply voltage. The leakage current is proportional to the diffusion area and perimeter of the drain. Therefore it is preferred to minimize the diffusion area and perimeter of the layout.

2) Sub threshold Leakage Current

Sub threshold leakage occur in the circumstances similar to diode leakage. In the case of an inverter, PMOS is turned OFF. Even for V_{gs}=0V there is current flowing through the channel, because V_{ds} of the PMOS almost equal to $-V_{dd}$ The I_d vs V_{ds} characteristics has an exponential relation in the sub threshold region ($V_{gs} < V_{th}$). The Figure below shows the sub threshold current at V_{gs}=0 .

The magnitude of sub threshold current is a function of process ,size and supply voltage. Also threshold voltage predominantly affects the current ,because reducing V_{th} increases the sub threshold current exponentially .

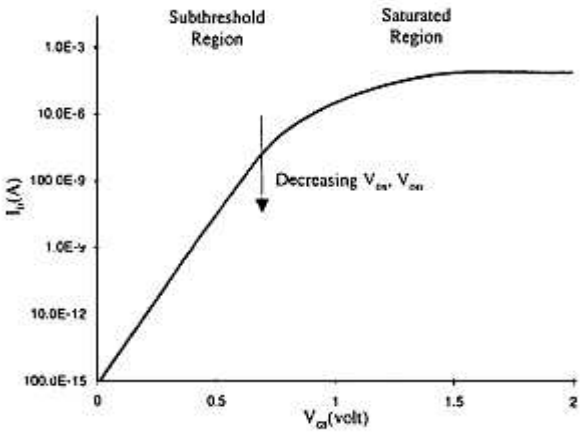

Figure 6. Sub threshold Leakage Current

III. EXISTING LOW POWER REDUCTION TECHNIQUES

A. Scaling of supply voltage

Average switching power is proportional to the square of supply voltage. Therefore the reduction of V_{dd} reduces the power consumption.

B. Source biasing

The concept of source biasing refers to the application of a positive voltage to the source terminal of the NMOS transistor during standby mode which rises the threshold voltage of the transistor. This reduce the sub threshold leakage current.

C. Multi threshold CMOS

The principle of MTCMOS is the employment of low threshold voltage transistors for the design of logic gates where switching speed is necessary and high threshold voltage transistors are used to isolate the stand by state and reduce dissipation.

Table 1:Some Low power Design Techniques Used today

Tradutional Power Reduction Techniques	**Dynamic Power Reduction Techniques**	**Leakage Power Reduction Techniques**	**Other Power Reduction Techniques**
Clock Gating	**Clock Gating**	**Minimize use of all low V_t cells**	**Multi Oxide Devices**
Power Gating	**Power Efficient Techniques**	**Power Gating**	**Minimize capacitance by custom design**
Variable Frequency	**Variable Frequency**	**Back Biasing**	**Power Efficient circuits**
Variable supply Voltage	**Variable Supply Voltage**	**Use Oxide thickness**	
Variable Device Threshold	**Variable Island**	**Use Fin FET**	

IV. ADVANCED LOW POWER DESIGN TECHNIQUES

In the context of the growing importance of low power designs for portable electronic devices, it is necessary to develop strategies to significantly reduce the power dissipation of the clock network, since this will lead to a major reduction in overall power dissipation of the chip. The power is reduced at different levels of VLSI design. Levels of optimization are shown in Fig. 7.

Here we see that as we move from lower level to higher level of design, ie from layout level to system level the number of counted errors increases. It means that power saving capability is more at higher level, compared to the lower stages

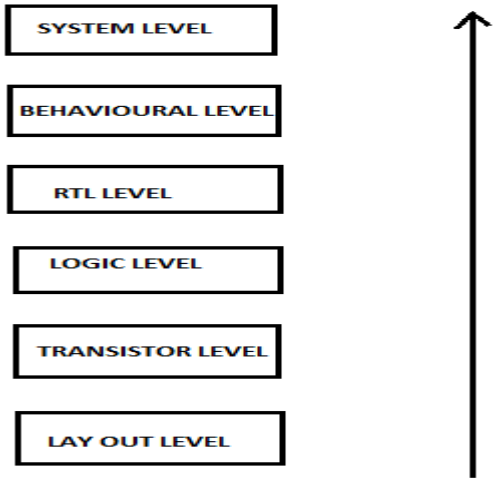

Figure 7:Power optimization at different levels of design

A. Algorithm Level Reduction

Power consumption at algorithm level is related to properties of that particular algorithm technique. So it should be carefully selected for lowering power consumption. For lowering the power, algorithm should be such that it should have minimum number of switching requirements. The algorithm used must have minimum number of operation because it takes less hardware. By increasing concurrency we can increase the efficiency of the device.

B. Architectural Level Techniques

The basic building blocks of this level are registers, buses, multipliers, memories, state machines etc. Each block perform high level function. At architectural level, it is important to perform power analysis because it is difficult to analyze each and every gate since chips become complex now days. The architectural level is the design entry point for the large majority of digital designs and design decisions at this level can have dramatic impact on the power budget design.

Power is as the function of their operating frequency and number of bits of components. For example, if n is the number of bits and f is the frequency, the power in the adder is given by

$$P= (n.k1+k2).f \quad (7)$$

1) Power management Techniques

Firstly power management is done by different saving modes such as DOZE, NAP and SLEEP modes by deactivating different levels, function levels.[5] At Doze mode cache memory is active except it all are off. Thus coherency is made by cache memory. In Nap mode processor wakes up after some interval of time or by external interrupt. In sleep mode all thing is done by reset.

2) Performance management technique

Performance management is done by adaptive technique by sensing the load of input.

3) Parallel and pipelined architecture

It is another aspect which is also taken into consideration. In parallel architecture frequency is scaled down by factor n, number of blocks. In parallel architecture area required is more but operation is become faster. In pipelining mode frequency remain same but voltage is scaled down.

Table 2: Comparison between parallel and pipelined architectures.

Factors	N-stage parallel processor	N-stage Pipelined processor
Capacitance	N*C	C
Voltage	V/N	V/N
Frequency	F/N	F
Dynamic power	$C*V^2*F/N^2$	$C*V^2*F/N^2$
Chip area	N times	Increase 10% or more

C. Carbon Nanotube Interconnects for Low power circuits

The sub-threshold circuits are very promising for ultra-low power applications. However, operating circuits at very low values of supply voltages raises robustness issues such as sensitivity to process, voltage and temperature (PVT) variations. High performance carbon nanotube structures that can withstand very high 'on'currents have been reported in literature and are a promising solution for VLSI requirements. This Carbon Nanotubes can be used as interconnect for VLSI circuits. The switching time of sub-threshold circuits are low and hence the interconnect delay should be kept minimum for sub-threshold operations.CNT serve this purpose .

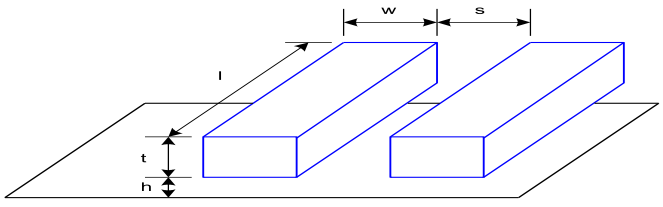

Figure 8: Typical structure of Interconnect

Pitch=w+s
Aspect ratio=w/s

1) Interconnect in an Inverter chain

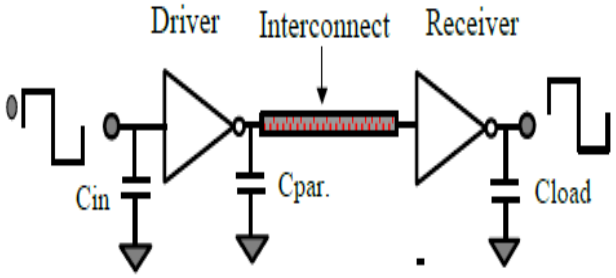

Figure 9: Typical Interconnect in a Inverter chain Circuit

The propagation delay of an RLC line driven by CMOS driver with driver resistance R_{tr} and a load capacitance C_L can be expressed as follows.

$$\acute{\Gamma}_d = R_{driv}(C_{driv} + C_{load}) + 0.4R_wC_wL^2 + (R_{driv}.C_{driv} + R_w.C_{load})L \quad (8)$$

R_{driv} is driver resistance, C_{driv} is driver capacitance, R_w and C_w are interconnects resistance and capacitance, and C_{load} is the load capacitance. R_{driv} increases exponentially as V_{DD} scales below the V_{th} and it is very high as compared to interconnect resistance (RW) of global interconnect. However, C_{driv} is very small as compared to global interconnects capacitance(CW). Hence from (1) and (2), in sub-threshold region global interconnect delay is dominated by driver resistance and interconnect capacitances.

2)SWCNT Interconnect

An isolated SWCNT on ground plane is shown in Fig.10. The separation between the nanotube and the ground is y and the diameter of the SWCNT is d. Assuming it to be in cylindrical form on the basis of Luttinger Liquid Theory, Burke developed an electrical equivalent of the structure as shown in Fig. 11. If a 1-D system has N conducting channels in parallel then its resistance is h/Ne^2T. Where h is Planks constant, e is electron charge and T is electron transmission coefficient. Due to spin and sub lattice degeneracy of electrons there are 4 parallel conducting channel in SWCNT (N=4). Thus assuming perfect contacts (T=1), the resistance of an SWCNT is $h/4e^2$. With the values of the physical constants substituted the resistance assumes the fairly large value of 6.45 KΩ. In the equivalent circuit (Fig.11) this resistance (Rf) is equally divided between the contacts at the two ends of the nanotube.

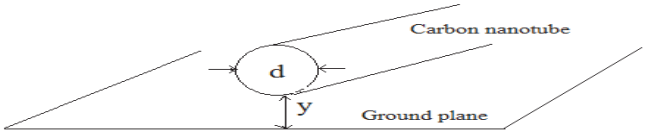

Figure10. Carbon nanotube, of diameter 'd', distance 'y' below it.

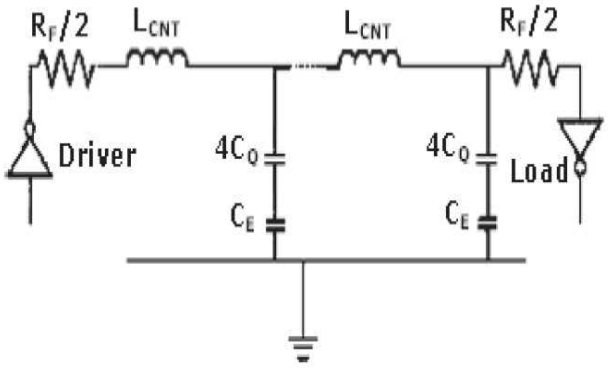

Figure 11: Equivalent circuit of SWCNT interconnect

$$R_f = h/4e^2 \quad (9)$$

where h is Planck's constant and e is the charge of electron. R_f is the resistance due to parallel channels and is divided equally between two ends.

The electrostatic capacitance (C_E) is due to charge stored by the CNT ground plane system (Fig.9) and is given by

$$C_E = 2\pi\varepsilon / \ln [\, y/d] \quad (10)$$

Where y is the distance of tube from ground and d is the diameter.

C_Q is the energy effective capacitance

$$C_Q = 2e^2/(hv_f) \quad (11)$$

Where v_f is the Fermi velocity L_{CNT} is the Inductance contributed by magnetic inductance and kinetic inductance

V. CONCLUSIONS

Electronic design aims at striking a balance between performance and power efficiency. Designing low power applications is a multi-faceted problem, because of the plurality of embodiments that a system specification may have and the variety of degrees of freedom that designers have to cope with power reduction. Here we reviewed the different low power techniques at each level of VLSI Design and the last section discussed one of the most advanced low power design technique using Carbon nanotube interconnects for sub-threshold devices. We use a particular technique according to given specification.

VI. REFERENCES

[1] J. Wang, S. Fang, W. Feng, "New Efficient Designs for XOR and XNOR Functions on the Transistor Level", *IEEE Journal of Solid-State Circuits,* **29(7),** 1994, 780–786.

[2] R. Zimmermann, W. Fichtner, "Low-power Logic Styles: CMOS Versus Pass-transistor Logic", *IEEE Journal of Solid-State Circuits*, **32** (7),1079–1089,1997

[3] Shams A. M., Darwish T. K. and Bayoumi M. A., "Performance analysis of low-power 1-bit CMOS full a adder cells," *IEEE Trans. Very Large Scale Integration. (VLSI) Syst.*, Vol. 10, No. 1, pp. 20–29, Feb. 2002.

[4] Radhakrishnan D., "Low-voltage low-power CMOS full adder," *IEEE Proc. Circuits Devices Syst.*, Vol. 148, No. 1, pp. 19–24, Feb. 2001

[5] Massoud Pedram., Design Technologies For Low Power VLSI*, Encyclopedia for Computer Science and Technology ,1995. pp 19-27*

[6] J. Montanaro, *et al.*, "A 160-MHz, 32-b, 0.5-W CMOS RISC Microprocessor", *IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, pp. 1703–1714, November 1996.

[7] Goel S., Kumar A. and Bayoumi M. A., "Design of Robust, Energy-Efficient Full Adders for Deep-Sub micrometer Design Using Hybrid- CMOS Logic Style," *IEEE Trans. Very Large Scale Integration. (VLSI) Syst., Vol. 14, No. 12, pp. 1309–1321, Feb. 2006.*

[8] Gary Yeap, "Practical Low Power Digital VLSI Design" *Kulwer Academic publishers,101 Philip Drive, Assinippi, Norwell, Massachusetts 02061, pp 2-22, 1998.*

[9] Manyak Kumar Rai and Sankar Sarkar,"Carbon Nanotubes as VLSI Interconnects "*InTech Open 2011,475-494*

[10] Rohit Dhyman,Rajeevan Chandel and R P Agarwal "Sub-Threshold Circuit Design for Ultra Low Power Logic". *International Journal of Contemporary Research in Engg. And Tech. Vol. 1, No. 1, 2011*

[11] Kanika Kaur and Arti Noor"Strategies and Methodologies for Low Power VLSI Design-A Review". International Journal for Advances in Engineering and Technology May -2011