Emerging trends in the developments of low power design technique

Dimov Stojce Ilcev *

Space Science Centre (SSC), Durban University of Technology (DUT), Durban, South Africa
*Corresponding author E-mail: ilcev@dut.ac.za

Abstract

In this paper is introduced a low power design technique for developing more reliable, functional, and more cost-effective handheld cellular telephones, portable computers, and peripherals. The portability requirements of handheld computers and other portable devices have placed tremendous pressure on electronic equipment designers, who need to deal with restrictions in the size of electronic units and power consumption. Even though battery technology is continuously improving, including reduced power consumption of processors and displays, extensive and continuous use of network services aggravates these issues. Now the onus is on the research and industrial communities to extend battery life and reduce weight. Equally, research on new techniques and technologies continues, to carefully manage energy consumption in mobile devices, while still providing continuous and fast connections to services and applications. This paper also discusses the novel trends in the developments and advancements in the area of low power Very Large Scale Integration (VLSI) design, dynamic power dissipation static power loss in Complementary Metal Oxide Semiconductor (CMOS), and advanced low power technique. Though low power as a well-established domain that undergone lots of developments from transistor sizing, process shrinkage, voltage scaling, clock gating, etc., to adiabatic logic are elaborated.

Keywords: VLSI; CMOS; CoC; RTL; GDSII; TTL; PMOS; NMOS; ICG; MOSFET.

1. Introduction

The power technique is presently emerging as the most critical issue in System-on-Chip (SoC) design. Power management is becoming an increasingly urgent problem for almost every category of design, as power density, measured in watts per square millimeter, rises at an alarming rate. From a chip-engineering perspective, effective energy management for an SoC must be built into the design starting at the architecture stage, thus low-power techniques need to be employed at every stage of the design, from Register Transfer Level (RTL) to Graphic Data System I (GDSII).

The growing market of mobile battery-powered electronic systems, such as cellular phones, laptops, or personal digital assistants requires the design of microelectronic circuits with low power dissipation. Thus, the size and complexity of the chips continue to increase, the difficulty in providing adequate cooling might either add significant cost or limit the functionality of the computing systems that make use of those integrated circuits. The necessity for personal electronic devices that today offer greater functionality and performance at lower costs and smaller sizes have increased rapidly. This market trend is driving the need for efficient SoC designs, where the power arises as one of the biggest problems. The microprocessor design has traditionally focused on dynamic power (active power) consumption as the limiting factor in system integration. As technology has shrunk to 90 nm and below, static power (leakage power) is posing new challenges to low power design [1 - 3].
New projected CMOS electronic technology has dissipated much less power than earlier technologies such as transistor-transistor and emitter-coupled logic devices, when not switching, CMOS transistors consumed negligible power. With the increase in device speed and chip density, the power of CMOS increased dramatically. According to technology trends, the dynamic power per device decreases over time. However, if it is assumed that the number of on-chip devices doubles every two years, total dynamic power increases on a per-chip basis. Additionally, the shrinking of feature sizes makes static power dissipation grows exponentially. Consequently, the packaging and cooling costs as well as the limited power capacity of batteries become unsustainable.

The reduction of power growth below the predicted numbers has become one of the most important tasks for designers. To cope with the power challenges, a number of new power saving techniques have been developed that can efficiently target both static and dynamic power loss. Therefore, the selection of power saving strategy becomes one of the most important challenges for designers. This presentation starts by covering basic sources of power consumption in CMOS, including both static and dynamic power loss. Furthermore, the presentation gives a detailed overview of existing low power techniques including standard techniques developed for mature technologies, process level techniques, and advanced low power techniques used in deep-submicron CMOS. Also, some attention is given to the selection of power saving strategy for design and early power estimation. Finally, future trends in low power design are discussed as well [1, 4 - 6].

2. Increased power consumption

Any design engineer will cite power dissipation as the main challenge in portability. Mobile device users demand more features and longer battery life at a lower cost; some 70% demand longer talk and standby times as primary mobile phone features. Equally, devices are getting smaller and sleeker, which requires a high level of silicon integration in advanced processes, which inherently bring higher leakage currents. This in turn requires reducing that leakage current which then helps reduce power consumption.

The trends for power needs versus expenditure in modern devices are shown in Figure 1, where the widening gap is abundantly clear. The power consumption of an CMOS transistor can be divided into three main groups: dynamic, also known as switching power, static or leakage power; and short-circuit power. In geometry values smaller than 90nm, leakage has become the dominant consumer of power, whereas in larger geometries, switching power is the larger contributor. In fact, power reduction strategies can be successfully used to minimize both types.

In both logic and memory, static and dynamic power are growing very fast, fuelled by high degree of semiconductor integration; an SoC often integrates all the components of an electronic system that contains digital, analogue, mixed-signal and often RF functions. In past, designers were primarily concerned with improving the performance of their designs, such as latency, throughput and frequency, and reducing silicon area to lower manufacturing costs. Now power is replacing performance as the key competitive metric for SoC design, which affects almost all SoC designs. In such a way, with the explosive growth of wireless system and home electronics, comes the demand for high-speed computation and complex functionality for competitive reasons. Today’s portable products are expected not only to be small, cool and lightweight, but also to provide extremely long battery life. Even wired communications systems must pay attention to heat, power density, and low power requirements. Among the products requiring low-power management are:

1) Consumer, wireless, and handheld devices: cell phones, Personal Digital Assistants (PDA), MP3 players, Global Positioning System (GPS) receivers and digital cameras.
2) Home electronics, such as game consoles for DVD/VCR players, digital media recorders, cable and satellite television set-top boxes and network and telecom devices; and
3) Tethered electronic equipment, such as servers, routers and other communication products bound by packaging costs, cooling costs, and energy star requirements supporting the Green movement to combat global warming.

For most designs being developed today, the emphasis on active low-power management, as well as on performance, area, and other concerns, is increasing [4], [6 - 9].

3. Low power design

Electronics devices are more reliable during their exploitation if they power consumption is reduced. Namely, the need for electronic devices that consume a minimum amount of power is the significant reason behind the development of CMOS technologies for low power design. Thus, electronic devices with CMOS input are the best-known electronic components for low power consumption. In fact, these devices may use less power than equivalent devices from other technologies, such as devices with Transistor–Transistor Logic (TTL) input. Accordingly, here will be introduced an important relation of power versus energy in low power design. In Figure 2 (Top) is illustrated diagram for power in Watts with two types of curve height approaches, and in Figure 2 (Bottom) is determined diagram for energy with two approaches of area under curves. The possible choice for diagram of power shown in Figure 2 (Top) is to sets battery life in hours and to provide packaging limits, while dynamic power has to be proportional to frequency. Besides, comparing the power of two designs can be misleading, so a lower power design could simply be slower. A good advice is to analyze the total energy per operation rather than the power to complete it. Unfortunately, to fix above problems with the energy metric and with change the operating frequency will not realize the answer.
In Figure 2 (Bottom) are shown two approaches of energy is the area under the curve. Although this metric looks promising, during an operation, the problem of energy reduction can be solved by deceleration of the chip by lowering the supply voltage or using small transistors. Since lower energy solutions might simply be lower performance, it will be necessary to have a metric that includes both energy and performance. When designing for low power, it is necessary to understand the basic tradeoffs. The system will need some metrics for low power designs, such as energy delay metric, which is helpful in making these trade-offs.

The designers of low power components have to realize not only how to calculate power consumption, but is also necessary to understand how input voltage level, input rise time, power-dissipation capacitance and output loading will affect the power consumption of a device. The high frequencies used in computer systems generally impose a strict limit on energy consumption, so it is necessary to minimize consumption for each device on the board. When calculating the power of the device, it is possible to determine the power size, the current requirements and it is also possible to determine the maximum reliability of the operating frequency. Besides, in the process of low power research and design is very important to determine power dissipation. In fact, for the most recent CMOS feature sizes (e.g., 90nm and 65nm), leakage power dissipation has become the biggest concern for VLSI circuit designers. In addition, power dissipation has become a very critical design metric with miniaturization and the growing trend towards wireless communication. The exploding market of portable electronic appliances fuels the demand for complex integrated systems that can be powered by lightweight batteries with long times between re-charges. Additionally, system cost must be extremely low to achieve high market penetration. Both battery lifetime and system cost are heavily impacted by power dissipation.

In addition to performance, power is becoming a significant concern for designers of CMOS elements that was originally a low power technology. However in same cases the low power design is not representing low power any more, because the CMOS chips can dissipate 50 Watts. This power comes from charging and discharging capacitors and is fundamental to all circuits that drive wires. In such a way, the circuit designer will look at power dissipation in CMOS circuits and provide proposed methods for reducing the power. The key areas in low power design are technology and algorithm, which both can make a huge difference in power consumption.

The CMOS power in total is cumulative sum of dynamic and static powers. The dynamic power is causing power dissipation when logic gates are switching, which is associated with active mode of operation and consists of two components, such as switching and internal power. The static power results from leakage currents, which also dissipates when transistors are turned off and increases with device shrinking, i.e. technology scaling [1], [10 - 12].

4. Dynamic power dissipation in CMOS

Dynamic power is consumed when electronic transistors are switching from one state to another, namely this power consumption is due to the charging and discharging of output load capacitance of a transistor. It consists of switching power consumed while charging and discharging the loads on a device, and internal (short circuit) power consumed internal to the device while it is changing state. Thus, the dynamic power consumption of a CMOS IC is calculated by adding the transient power consumption and capacitive-load power consumption. There are three major sources of power dissipation in a CMOS circuit:

$$ P_t = P_{SW} + P_{SC} + P_t $$

from which: $P_{SW} = C_{EFF} \cdot V_{DD}^2 \cdot f_{CLK}$

where $P_t$ = total power dissipation of a CMOS circuit, $P_{SC}$ = short circuit power and $P_t$ = leakage power, $f_{CLK}$ = clock frequency and $C_{EFF}$ = effective capacitance. The switching power ($P_{SW}$) is the result of charging and discharging parasitic capacitances in the circuit, which can be expressed as:

$$ P_S = C_L \cdot V_{DD} \cdot \Delta V \cdot f_{CLK} $$

from which: $C_{EFF} = C_L \cdot \alpha$

where value $\alpha$ = node transition factor, which is the effective number of power-consuming voltage transitions experienced per clock cycle, $C_L$ = output node capacitance, $V_{DD}$ = power supply voltage or rail-to-rail voltage and $\Delta V$ is the voltage swing.

There are two solutions showing dynamic power dissipation that can occur in CMOS Integrated Circuits (IC), such as Positive MOS (PMOS) and Negative MOS (NMOS) transistors, which is most widely adopted in current of VLSI systems. In Figure 3 (Left) is depicted power consumption due to charge of load capacitance and in Figure 3 (Right) is shown depicted power consumption due to discharge of load capacitance.

Dynamic power can be lowered by reducing switching activity and clock frequency, which affects performance and reduces capacitance and supply voltage. It can also be reduced by cell selection, when faster slew cells consume less dynamic power. Second dynamic power loss of internal power is consumed during the short period of time when the input signal is at an intermediate voltage level, during which PMOS and NMOS transistors can be conducting. This condition results in a nearly short-circuit conductive path from voltage to ground. In Figure 4 is shown the short-circuit power when input signal is at intermediate voltage level. A relatively large current, called the crowbar current, flows through the transistors for a brief period of time. Lower threshold voltages and lower transitions result in more internal power consumption.
Consider the inverter circuit in Figure 4, when input signal changes from high to low, the output node voltage makes a full transition from 0 to VDD and one half of the energy drawn from the power supply is dissipated as heat in the conducting PMOS M1. When the input signal changes from low to high, the energy stored in the output node capacitance is dissipated as heat in the conducting NMOS point. Short circuit power is due to the finite rise time and fall time of the input signal as shown in Figure 4 [1], [6], [12 - 14].

\[ P_{SC} = I_{SC} \cdot V_{DD} = t_{SC} \cdot V_{DD} \cdot I_{PEAK} \cdot f_{CLK} \]  

(3)

Where \( I_{SC} \) is the short circuit current, \( t_{SC} = \) time circuit current and \( I_{PEAK} = \) current peak.

The short circuit power can be kept less than 15% of the switching power with careful design. In such a way, the reduction of the power supply voltage is one of the most effective ways to achieve low power design, which can be provided at the circuit level using multiple VDD or at the system level using dynamic VDD control according to system workload. The reduction of node transition factor \( \alpha \) requires a detailed analysis of signal transition probabilities and the use of various circuit and system level techniques such as logic optimization, the usage of a gated clock, and the prevention of glitches.

The load capacitance can be reduced at the circuit level with novel circuit design or transistor sizing. The total switched load capacitance can also be reduced at the system level by clock gating or stopping certain units from useless transitions. The voltage swing \( \Delta V \) can be reduced at circuit level using novel circuits. The clock frequency \( f_{CLK} \) can be reduced at the logic and architecture level by using parallel architecture to achieve the same throughput at lower clock frequency. The leakage power \( (P_L) \) and dynamic power \( (P_{DYN}) \) can be expressed as:

\[ P_L = I_{L} \cdot V_{DD} \quad \text{and} \quad P_{DYN} = P_{SW} \cdot P_{SC} = C_{EFF} \cdot V_{DD}^{2} \cdot f_{CLK} + t_{SC} \cdot V_{DD} \cdot I_{PEAK} \cdot f_{CLK} \]  

(4)

where \( I_{L} = \) total leakage current in a CMOS circuit caused by 6 short channel leakage mechanisms: the reverse bias P-N junction leakage, subthreshold leakage, oxide tunneling and gate current due to hot carrier injection, gate induced drain leakage and the channel punch through current [6], [13 - 16].

5. Static power loss in CMOS

Leakage current was negligible in earlier CMOS technologies, however with shrinking device geometries and reduced threshold voltages, leakage power is becoming increasingly significant, sometime approaching the levels of dynamic power dissipation. The main causes of leakage power are reverse-bias P-N junction diode leakage, subthreshold leakage, and gate leakage. These leakage paths in a CMOS inverter resulting from leaking currents in transistors, which P-N junction leakage to substrate is shown in Figure 5 (Left) and P-N junction leakage from N-well is shown in Figure 5 (Right). Where \( I_1 = \) reverse-bias P-N junction diode leakage; \( I_2 = \) subthreshold leakage; \( I_3 = \) gate leakage through the oxide; and \( I_4 = \) gate induced drain leakage.
The static power components become important when the circuits are at rest, i.e. when there is no activity in the circuits and they are all biased to a specific state. Electronic CMOS devices have very low static power consumption, which is the result of leakage current, but when switching at a high frequency, while dynamic power consumption can contribute significantly to overall power consumption. Thus, charging and discharging a capacitive output load further increases this dynamic power consumption.

The subthreshold current (Isub) is a leakage between source and drain when the gate bias is below VTH value, which is mainly dominated by the diffusion current. It is enhanced by the Drain Induced Barrier Lowering (DIBL) effect. Thus, the DIBL effect accounts for the threshold voltage reduction depending on the drain bias, which parameter is dependent on the type of gate dielectric, the gate length and the doping. The subthreshold current is calculated with the following equation:

\[ I_{\text{SUB}} = \mu \cdot C_{\text{OX}} \cdot \left( V_{\text{TH}} - V \right) \cdot W/L \cdot e^{V_{\text{GS}}/nV_T} \]

where \( \mu \) = bias or denotes carrier mobility, \( C_{\text{OX}} \) = gate oxide capacitance per unit area, \( V_{\text{TH}} \) = threshold voltage, \( W \) & \( L \) = effective transistor width and length, respectively, \( e \) = electron charge, \( V \) = sum of the voltage coefficient, \( V_{\text{GS}} \) = gate-to-source voltages, \( V_{\text{TH}} \) = threshold voltage, \( V_T \) = thermal voltage, and \( n \) = subthreshold swing coefficient constant.

The static power (\( P_{\text{ST}} \)) value is dissipated due to non-ideal characteristic of the transistor and totally the leakage power (\( P_{\text{LE}} \)). But there is a true potential of reducing unwanted and unnecessary power consumptions via other power reduction techniques based on the static power. Total power (\( P_T \)) is mainly composed of static power and dynamic power (\( P_{\text{DOV}} \)), such as relation:

\[ P_T = P_{\text{DOV}} + P_{\text{ST}} \]

where \( f \) = operating frequency and other values are already stated above. Static power dissipated when the device is sitting ideal and it’s totally a leakage power. At this state the semiconductor devices are powered on but switched off and so it will leads to a leakage of the devices [2], [17], [18].

6. Standard low power technique

Static power dissipation becomes an issue when the circuit is inactive or in a power-down mode focusing on reducing supplies voltage, switching activity and load capacitance. The common standard low power technique are supply voltage scaling (multi-Vdd), multi-threshold voltage, clock gating and operand isolation.

1) Clock Gating Technique – Clock gating is a popular technique used in many synchronous circuits for reducing dynamic power dissipation, by removing the clock signal when the circuit is not in use. It saves power by pruning the clock tree, at the cost of adding more logic to a circuit. In many cases, data is loaded into registers infrequently, but the clock signal continues to switch at every clock cycle, driving a large capacitive load. The registers get evaluated to the same value every clock cycle. The clock may be shut off to these registers using a gating circuit, which basically prevents the clock from triggering the registers. Clock gating can result in 30% of the power savings compared to the design without clock gating, which typical clock gating circuit using AND gate is shown in Figure 6 (Left), and using Integrated Clock Gating (ICG) cell in Figure 6 (Right).

Pruning the clock disables portions of the circuitry so that the flip-flops in them do not have to switch states. Switching states consumes power. When not being switched, the switching power consumption goes to zero, and only leakage currents are incurred. The main aim of the clock gating technique is to cut off the clock during the idle cycles of flip-flop. This technique is implemented for three different cell types: 1) Latch based cell, 2) Flip flop based cell, 3) Gate based cell. It is inferred that gate based clock-gating cell is preferable than latch based cell or flip flop based cell when power is the major constraint.

2) Operand Isolation Technique – This technique is similar to clock gating, which reduces switching activity and initiatives datapath blocks. Sometimes functional units can remain “on” (evaluating) even though the results are not used in the subsequent stages. These functional units may be switched “off” when their results are not needed. This is provided by gating the input to that combinational
logic when the output is not in use. In Figure 7 (Left) is illustrated before operand isolation and Figure 7 (Right) shows after operand isolation. In the other words, in electronic low power digital synchronous circuit design, operand isolation is a technique for minimizing the energy overhead associated with redundant operations by selectively blocking the propagation of switching activity through the circuit. This technique isolates sections of the circuit (operation) from "seeing" changes on their inputs (operands) unless they are expected to respond to them. This is usually done using latches at the inputs of the circuit. The latches become transparent only when the result of the operation is going to be used. Thus, the operand isolation technique reduces dynamic power dissipation, namely, when the enable is inactive, the datapath inputs are disabled so that unnecessary switching power is not wasted in the datapath.

3) Gate Level Optimization Technique – This technique, such as logic restructuring, illustrated in Figure 8 (Left) can reduce three stages to two stages through logic equivalence transformation, so the circuit has less switching and fewer transitions. It reduces dynamic power up to 5% and had no significant impact on any other aspects of the design flow. Logic (transistor) resizing is upsizing improves slew times, reducing dynamic current, downsizing reduces leakage current and making high-activity nets internal to the cell. To be effective, sizing operations must include accurate switching information. In Figure 8 (Right) is shown pin swapping technique, which is rewiring of low capacitance gate pins to high-activity nets, by swapping gate pins, switching occurs at gates and pins with lower capacitive loads [2], [7], [19 - 21].

7. Advanced low power technique

Advanced Low Power Techniques is developed to deal with the increasing contribution of leakage currents in deep-submicron CMOS, which is divides in two groups. The first process-level techniques are retrograde and halo doping (Bulk CMOS), Silicon-On-Insulator (SOI) and multiple-gate Metal Oxide Semiconductor Field Effect Transistor (MOSFET), while the second circuit-level techniques are multi-voltage design, voltage and frequency scaling, power gating, body biasing and stacked transistor.

1) Process-Level Techniques – This technology is requiring retrograde and halo doping techniques to reduce leakage and improve short channel characteristics. In Figure 9 is illustrated retrograde and halo doping (Bulk CMOS) as different aspects of well engineering. Its retrograde channel doping gives low surface channel concentration followed by a highly-doped subsurface region, improves short-channel effects, increase surface channel mobility and increase of threshold voltage. Besides, halo doping introduces highly doped positive regions at the edges of the channel, reduces width of depletion area in drain and source and also reduces of threshold voltage degradation. Compared to Bulk CMOS devices, SOI CMOS Devices can reduce power supply voltage while maintaining operating performance and can greatly reduce power consumption, which large area to substrate and noise coupling is shown in Figure 10 (Left). Small area for leakage and noise coupling is shown in Figure 10 (Right). In general, the Partially Depleted SOI (PDSOI) device is very similar to the bulk one in the design for depressing Short Channel Effect (SCE) controls. The retrograde and halo (pocket) doping, two typical body doping ways of the Nano-scale bulk MOSFET transistors, can also be applied in the PDSOI device. From a circuit designer's point of view, the bulk and SOI CMOS input behavior is quite identical. The gate capacitance is formed in a similar way, being practically equal between devices of the same geometry in SOI and bulk technology. There are slightly different capacitances in SOI and bulk technologies that load the gate. Still, the differences seem so small that they are not measurable by standard RF device layouts. Specifically tailored layouts are required for this purpose.
Circuit-Level Techniques – Here will be introduced just one solution of circuit-level techniques as a concept of power islands (voltage islands, power domains), known as multiple voltage design. The most basic form of this new approach is to partition the internal logic of the chip into multiple voltage regions or power domains, each with its own supply called multi-voltage design. This technique is based on the realization that in a modern SoC design, different blocks have different performance objectives and constraints. A processor, for instance, may need to run as fast as the semiconductor technology will allow. In this case, a relatively high supply voltage is required. An Universal Serial Bus (USB) block, on the other hand, may run at a fixed, relatively low frequency dictated more by the protocol than the underlying technology. Thus, a lower supply rail may be sufficient for the block to meet its timing constraints and a lower supply rail means that its dynamic and static power will be lower. In Figure 11 is depicted solution of multiple voltage design with the following categorization of multi-voltage strategies:

- **Static Voltage Scaling (SVS):** different blocks or subsystems are given different, fixed supply voltages;
- **Multi-level Voltage Scaling (MVS):** an extension of the static voltage scaling case where a block or subsystem is switched between two or more voltage levels. Only a few, fixed, discrete levels are supported for different operating modes;
- **Dynamic Voltage and Frequency Scaling (DVFS):** an extension of MVS where a larger number of voltage levels are dynamically switched to follow changing workloads; and
- **Adaptive Voltage Scaling (AVS):** an extension of DVFS where a control loop is used to adjust the electrical voltage [2], [7], [12], [22].

8. Conclusion

In recent years, several techniques, methods and tools for designing low-power circuits have been presented in the scientific world. Thus, only a few of them have found their way in current design flows, so purpose of this work is to summarize, mainly by way of examples, what in experience of researchers are the most trustful approaches to low power design. The contribution of this work should not be intended as an exhaustive survey of the existing literature on low-power design; rather, it proposes insights that the designer can rely upon when power consumption is a critical constraint. In this sense, this technology is focused solely on digital circuits with modest attention to CMOS and MOSFET devices, which is most widely adopted in current VLSI technique.

References