# A Review on Power Reduction Technique for VLSI Designs

**Priyanka S. Patel[1] Dharmistha T. Patel[2]**
[1,2]Masters of Electronics and Communication Engineering
[1]SVIT, VASAD, Anand, Gujarat, India [2] HGCE, VAHELAL Ahmedabad, Gujarat, India

*Abstract*--- Low power has emerged as a necessity in portable electronics and communication devices. As a result, power dissipation has become an important design parameter for digital circuit designs. In this paper a survey is made regarding some existing techniques for power optimization in CMOS designs. Power optimization at logic and circuit levels is considered.
**Keywords**: Power reduction, low power, optimization, digital designs.

## I. INTRODUCTION

Now a days low power devices has become necessary for any electronic portable device to be successful in the market. In this paper have surveyed various techniques for power reduction in the CMOS designs.

## II. POWER ANALYSIS OF CMOS DESIGNS

CMOS technology has become successful because of its intrinsic low power consumption. As a result designers can concentrate on maximizing circuit performance and minimizing circuit area. Another feature of CMOS technology is its scaling property which has permitted a steady decrease in the feature size. This allowss highly complex systems on a single chip. Sources of power dissipation in the CMOS circuit consist of leakage current, power loss due to charging and discharging of the capacitance at the gate output and short circuit power loss during the transition of output line from one voltage level to other, there is a period of time when both the PMOS and NMOS transistors are on, thus creating a path from Vdd to ground. This is summarized in the following expressions:

$$P_{AVG} = P_{DYNAMIC} + P_{SHORTCIRCUIT} + P_{STATIC} \quad (2.1)$$

Where $P_{AVG}$ = Average power

$P_{DYNAMIC}$ = Dynamic power

$P_{SHORTCIRCUIT}$ = Short circuit power

$P_{STATIC}$ = Static power

### A. *Dynamic power*

The first term in eq. (2.1) represent the power required to charge and discharge circuit nodes. Node capacitances are represented by C. The factor N is the switching activity, i.e., the number of gate output transitions per clock cycle.

For dynamic power dissipation there are two components one is switching power due to charging and discharging of load capacitance. The other is the short circuit power due to the nonzero rise and fall time of input waveforms. The switching power of a single gate can be expressed as

$$P_D = \alpha C_L V_{DD}^2 f \quad (2.2)$$

Where,

$\alpha$ is the switching activity,

f is the operation frequency,

$C_L$ is the load capacitance,

$V_{DD}$ is the supply voltage.

Short circuit power

The second term in eq. (2.1) represent the power dissipation during output transitions due to current flowing from the supply to ground. This current is often called short-circuit. The factor QSC represents the quality of charge carried by the short-circuit current per trasition.

The short circuit power of an unloaded inverter can be approximately given by

$$P_{SC} = \beta(V_{DD} - V_{th})^3 \tau/12T \quad (2.3)$$

Where,

$\beta$ is the transistor coefficient,

$\tau$ is the rise/fall time,

T(1/f) is the delay.

### B. *Leakage power*

There are three dominant components of leakage in a MOSFET in the nanometer regime:

(1) Sub-threshold leakage, which is leakage current from drain to source($I_{sub}$).
(2) Direct tunning gate leakage which is due to the tunning of electron (or hole) from the bulk silicon through the gate oxide potential barrier into the gate.
(3) The source/substrate and drain/substrate reverse-biased p-n junction leakage.

The third term in eq. (2.1) represents static power dissipation due to leakage current $I_{leak}$ .device source and drain diffusions from parasitic diodes with bulk regions. Reverse-biase currents in these diodes dissipate power. Subthreshold transistor currents also dissipate power. In the sequel, we will refer to the three terms above as switching activity power, short-circuit power and leakage current power.

## III. CIRCUIT LEVEL

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

### A. *Transistor sizing*

Transistor sizing is the process of designing gates that can implement a given delay by changing the dimensions of the transistors. The problem of transistor sizing in a CMOS layout to minimize the short-circuit power dissipation was described by Borah et al. [1995]. It was shown that the power optimal size of the transistor in a gate that is driving a given load can be larger than minimum size. The authors derived the optimal sizes considering the power and delay constraints

 

for these transistors and presented an algorithm for calculating the optimal power sizing for all gates in a circuit. Those researchers have reported about 15–20% reduction in total power dissipation as a result of transistor sizing.

### B. Voltage scaling

In voltage scaling technique, the supply voltage of the CMOS circuit is dynamically adjusted as the workload varies with time, without degrading the performance of the circuit. It is effective technique to reduce power dissipation in CMOS integrated circuits at the circuit level.

The voltage scaling technique exploits variation in the computational workload by dynamically modifying the supply voltage and hence the clock frequency of the sequential system. The main objective of voltage scaling technique is to provide more throughput for the execution of the more computation intensive tasks while saving energy for the rest of the time by reducing the supply voltage and also lowering the operating speed of the sequential system.

### C. Voltage island

Of the different approaches to reduce power, the use of multiple voltages is an attractive solution in core-based SoC design [2]. Performance critical blocks such as processor cores usually require the highest supply voltage level while other functions, such as memories or control logic, can operate on lower supply voltages [3]. Multi-voltage designs involve the partitioning of a chip into areas called "voltage islands" that can be operated at different voltage levels, or be turned off when idle. Each voltage island occupies a contiguous physical space and operates at a supply voltage that meets the performance requirement. With the use of voltage islands, the chip design process is becoming more complicated. One of the crucial steps of voltage island design is grouping cores with similar supply voltages into small number of islands.

### D. Multiple threshold voltage

Multiple threshold voltage technique uses both high and low threshold voltage transistors within the same digital integrated circuit. The main goal of multiple threshold voltage technique circuits is to selectively scale the threshold voltages along with the supply voltage to increase the operating speed without increasing the sub threshold leakage current.

The multiple threshold voltage technique selectively puts the low threshold voltage transistors on the speed critical paths in a circuit to increase the speed while working at a lower supply voltage, reducing the power consumption of the integrated circuit. A single threshold voltage circuit waste power in the form of leakage current on many non critical delay paths. The multiple threshold voltage technique utilizes this characteristic by selectively scaling the threshold voltage only for the speed critical paths. The main goal of the multiple threshold voltage technique is to minimize the number of low threshold voltage transistors required to satisfy a target clock frequency while maximizing the number of high threshold voltage transistors to achieve the lowest sub threshold leakage current. The multiple threshold voltage technique provides an opportunity to further scale the threshold voltage without violating any limitation in the total sub threshold leakage power. Therefore the target clock frequency can be satisfied within the limited power budget by only scaling the threshold voltages of those portion of a circuit where a threshold voltage transistor is required to achieve a specific propogation delay at a reduced supply voltage.

### E. Power gating

Power gating is one of the promising technique to reduce the leakage power. In power gating the leakage power is saved by cutting off the( voltage supply of the blocks of the circuit that are idle. As power gating is based on prediction method, care has to be taken that mispredictions does not occur frequently as turning off and on the circuit block requires energy.

## IV. Logic Level

### A. Clock gatin

Large VLSI circuits such as processors contain register files, arithmetic units and control logic. The register file is typically not accessed in each clock cycle. Similarly, in an arbitrary sequential circuit, the values of particular registers need not be updated in every clock cycle. If simple conditions that determine the inaction of particular registers can be determined, then power reduction can be obtained by gating the clocks of these registers. When these conditions are satisfied, the switching activity within the registers is reduced to negligible levels. The same method can be applied to “turn off” or “power down” arithmetic units when these units are not in use in a particular clock cycle. For example, when a branch instruction is being executed by a CPU, a multiply unit may not be used. The input registers to the multiplier are maintained at their previous values, ensuring that switching activity power in the multiplier is zero for this clock cycle.

### B. Logic factorization

A primary mean of technology independent optimization is factoring of logical expressions. For example, the expression $(a \cdot c + a \cdot d + b \cdot c + b \cdot d)$ can be factored into $(a + b) \cdot (c + d)$ considerably reducing transistor count. Common sub expressions can be found across multiple functions and can be reused. Kernel extraction is a commonly used algorithm to perform multi level logic optimization for area optimization. In this algorithm, kernels of the given expressions are generated and kernels that reduce the literal count maximally are selected. When targeting power dissipation, the cost function is not the literal count but the switching activity. Modified kernel extraction methods that target switching activity power are described in [10] by K.Roy.

### C. Don’t care optimization

Generally any gate in a combinational circuit has an associated controllability and observability don’t-care set. The controllability don’t-care set corresponds to those input combinations that never occur at the gate inputs. The observability don’t-care set corresponds to collections of those input combinations that produce the same values at the circuit outputs. Methods to reduce circuit area and improve the delay exploiting don’t care sets have been presented. The

 

power dissipation of a gate is dependent on the probability of the gate evaluating to a 1 or a 0. This probability can be changed by utilizing don't care sets. A method of don't-care optimization to reduce switching activity and therefore power dissipation was presented in [11]. This method was improved upon by S. Imam and Massoud Pedram in [12] where the effect of don't-care optimization of a particular gate on the gates in its transitive fan out is considered.

### *D. Path balancing*

Spurious transitions in logical circuit account for between 10% and 40% of the switching activity power in typical combinational logic circuits [13]. In order to reduce this spurious switching activity, the delays of paths that converge at each gate in the circuit should be equal roughly. The delay of all paths in the circuit can be made equal by selectively adding unit delay buffers to the inputs of gates in a circuit. This addition will help in not increasing the critical delay of the circuit and will effectively eliminate spurious transitions. However, the addition of buffers increases capacitance which may offset the reduction in switching activity and hence reduction in dynamic power. Methods which reduce rather than completely eliminate spurious switching activity, while adding a minimal number of unit delay buffers have been proposed. The design of a multiplier with transition reduction circuitry that proposes glitch reduction by path balancing is described in [14].

### *E. Technology mapping*

After the optimized logic equations have been obtained, the task remain is to map the equations into a target library that contains optimized logic gates in the chosen technology. A typical target library will contain hundreds of gates with different transistor sizes. Modern technology mapping methods use a graph covering formulation, originally presented in [15], to target area and delay cost functions. The graph covering formulation of [15] has been extended to the power cost function. Under the zero delay model, the optimal mapping of a tree can be determined in polynomial time, by extending the algorithm of [15]. Different approaches to technology mapping that assumes different delay models and target minimal power dissipation have been described [16] [17] [18].

### *F. Static encoding*

State encoding is a well-researched problem for minimal area. These methods have to be modified to target a power cost functions like weighted switching activity. If a state s has a large number of transitions to state q, then the two states should be given unidistant codes, so as to minimize switching activity at flip-flop outputs. However, the complexity of combinational logic resulting from a state assignment should not be ignored. Methods to encode State Transition Graphs to produce two level and multi level implementations with minimal power are described in [10] and [19]. A method to re-encode logic-level sequential circuits to minimize power dissipation is presented in [20]. Encoding to reduce switching activity in data path logic has also been the subject of interest. A method to minimize the switching on buses is proposed in [21]. In this technique, an extra line E is added to the bus which signifies if the value is being transferred is the true value or needs to be bitwise complemented upon reception. Depending on the value transferred in the previous cycle, a decision is made to either transfer the true current value or the complemented current value, so as to minimize the number of transitions on the bus lines. For example, if the previous value transferred was 0000, and the current value is 1011, then the value 0100 is transferred instead and the line E is asserted to signify that the value 0100 has to be complemented at the other end. Other methods of bus coding are also proposed in [21]. Methods to implement arithmetic units other than in standard two's complement arithmetic are also being researched. A method of one hot residue coding to minimize switching activity of arithmetic logic is presented in [22].

### *G. Retiming*

Retiming is a process of re-positioning the flip-flops in a pipelined digital circuit to either minimize the number of flip-flops or to minimize the delay through the longest pipeline stage. The idea is to identify circuit nodes with high hazard activity and high load capacitance as candidates for adding a flip-flop. The technique does not produce the optimal retiming solution because the retiming of a single node can dramatically change the switching activity of many other nodes in the circuit.

The authors report that the power dissipated by the 3-stage pipelined circuits obtained by retiming for low power with a delay constraint is about 8% less than that obtained by retiming for minimum number of flip-flops given a delay constraint.

## V. CONCLUSION

In this paper an attempt has been made to discussed various power reduction technique for digital system design. This survey will be of use to the new comer researchers in the field of low power VLSI.

## REFERENCES

[1] T.D. Burd and R. W. Broadersen, "Energy Efficient Microprocessor Design," Kluwer Academic Publishsers, 2002.
[2] K.J. Nowka et al., "A 32-bit PowerPC System-on-a-Chip With Support for Dynamic Voltage Scaling and Dynamic Frequency Scaling," IEEE Journal of Solid State Circuits, Vol. 37, No. 11, pp. 1441-1447, November 2002.
[3] T.D. Burd, T. A. Pering, A. J. Stratakos and R. W. Broadersen, "A Dynamic Voltage Scaled Microprocessor System," IEEE Journal of Solid-State Circuits, Vol. 35, No. 11, pp. 1571-1580, November 2000.
[4] D. Takashima et al., "Standby/Active Mode for Sub-1-V Operating ulsi Memory," IEEE journal of Solid-State Circuits, Vol. 29, No. 4, pp. 441-447, April 1994.
[5] L. Su et al., "A High- Performance Sub-0.25 µm CMOS Technology with Multiple Thresholds and opper Interconnects," Proceedins of the IEEE International Symposium on VLSI Technology, pp. 18-19, June 1998.
[6] T. McPherson et al., "706 Mhz G6 S/390 Microprocessor Exploiting Multiple Vt and Copper

 

Interconnects," Proceedings of the IEEE International Solid-State Circuits Conference, pp. 96-97, February 2000.

[7] Volkan Kursun, "Supply and Threshold Voltage Scaling Techniques in CMOS Circuits," Ph.D. thesis, University of Rochester, Rochester, New York.

[8] R. Puri, D. Kung, L. Stok, "Minimizing Power with Flexible Voltage Islands," in Proc. ISCAS, May 2005, pp. 21-24.

[9] W. Mak and J. Chen, "Voltage Island Generation under performance Requirement for SoC Designs," in Proc. ASP-DAC, Jan. 2007, pp. 798-803

[10] K. Roy and S. Prasad, "SYCLOP: Synthesis of CMOS logic design for power applications in processing of the tnt;1conference on computer design," VLSI in computers design: VLSI in computers and processors,pages 464-467, October 1992.

[11] A.Shen, S. Devadas, A. Ghosh, and K. Keutzer,"On Average Power Dissipation and Random Pattern Testability of Combinational Logic Circuits," In Proceedings of the Int'l Conference on Computer-Aided Design, pages 402–407, November 1992.

[12] S. Iman and M. Pedram, "Multi-Level Network Optimization for Low Power," In Proceedings of the Int'l Conference on Computer- Aided Design, pages 371–377, November 1994.

[13] A.Ghosh, S. Devadas, K. Keutzer, and J. White, " Estimation of Average Switching Activity in Combinational and Sequential Circuits," In Proceedings of the 29th Design Automation Conference, pages 253–259, June 1992.

[14] C. Lemonds and S. S. Mahant Shetti, " A Low Power 16 by 16 Multiplier Using Transition Reduction Circuitry," In Proceedings of the Int'l Workshop on Low Power Design, pages 139–142, April 1994.

[15] K.Keutzer, "DAGON: Technology Mapping and Local Optimization," In Proceedings of the 24th Design Automation Conference, pages 341–347, June 1987.

[16] C-Y. Tsui, M. Pedram, and A. M. Despain, "Technology Decomposition and Mapping Targeting Low Power Dissipation," In Proceedings of the 30th Design Automation Conference, pages 68–73, June 1993.

[17] V.Tiwari, P. Ashar, and S. Malik, "Technology Mapping for Low Power," In Proceedings of the 30th Design Automation Conference, pages 74–79, June 1993.

[18] B. Lin, "Technology Mapping for Low Power Dissipation," In Proceedings of the Int'l Conference on Computer Design: VLSI in Computers and Processors, October 1993.

[19] C-Y. Tsui, M. Pedram, C-A. Chen, and A. M. Despain, " Low Power State Assignment Targeting Two- and Multi-level Logic Implementations," In Proceedings of the Int'l Conference on Computer-Aided Design, pages 82–87, November 1994.

[20] G. D.Hachtel, M. Hermida, A. Pardo, M. Poncino, and F. Somenzi, "Re-Encoding Sequential Circuits to Reduce Power Dissipation," In Proceedings of the Int'l Conference on Computer- Aided Design, pages 70–73, November 1994.

[21] M. Stan and W. Burleson, "Limited-weight codes for low-power I/O," In Proceedings of the Int'lWorkshop on Low Power Design, pages 209–214, April 1994.

[22] W. A. Chren, "Low Delay-Power Product CMOS Design Using One Hot Residue Coding," In Proceedings of the International Symposium on Low Power Design, April 1995.

 