STATIC AND DYNAMIC POWER CONSUMPTION OF ARITHMETIC CIRCUITS IN MODERN TECHNOLOGIES

Bojan Jovanović, Milun Jevtić, *University of Niš, Faculty of Electronic Engineering, {bojan,milun.jevtic}@elfak.ni.ac.rs*

Abstract – *This paper gives an overview of technology parameters influencing static and dynamic power consumption in modern arithmetic circuits. Also, some techniques for power minimization are presented. As an example, this paper presents the results of power consumption of binary dividers implemented in FPGAs with various technological properties.*

1. INTRODUCTION

In today's world of numerous high-volume battery powered portable electronic devices low power consumption becomes obvious need and dominant design goal. While historically for CMOS circuits there has always been a strong relationship between power and performance, the power of the chip remained within the allowable power envelope since large packages, cooling fins, and fans have been capable of dissipating the generated heat. In this scenario, designer focused primarily on achieving the needed performance. However, as the density and size of the chips and systems continue to increase according to Moore's low, the difficulty in providing adequate cooling might either add significant cost to the system or provide a limit on the amount of functionality. This new relationship between peak achievable performance and energy efficiency change the way one tends to think about design. Starting from 0.18μm technologies, static power consumption due to leaky "off" transistors, is now a non negligible source of power dissipation even in running mode. Thus, the total power consumption (i.e. dynamic plus static power) has to be optimized instead of simply reducing dynamic power. Design methods that explore true power optimization need to work in a large dimension search space, where power and performance of different solutions are compared. This includes system architecture optimization (outer loop), block-level optimization (intermediate loop), and fixed topology optimization (inner loop).

2. THE SOURCES OF POWER CONSUMPTION

The two main sources of power dissipation in CMOS VLSI's are the dynamic power dissipation due to charging and discharging of load capacitance, and the power dissipation due to subthreshold leakage. There may be short-circuit power dissipation (V_{DD} to ground) as the third source of power dissipation. This power source is due to non-zero rise and fall time of input waveforms and it is less than 10% in total power dissipation.

The expression for dynamic power consumption is widely known. It depends on squared power supply, V_{DD}, operating frequency, load capacitance of the node, and the average number of $0 \rightarrow 1$ transitions within one clock cycle.

Leakage currents consists of two main components as shown in Fig. 1.: subthreshold leakage (I_2) and gate leakage current (I_3). There are some other leakage current components that have started to gain interest recently due to an excessive scaling of the transistor dimensions. They occur due to the short channel-lenght: injection of hot carriers from substrate to gate oxide (I_4) and punchthrough leakage (I_6), due to the thinner oxide thickness: gate-induced drain leakage (I_5), and, due to high doping concetracions: junction reverse-biased current (I_1) [1].

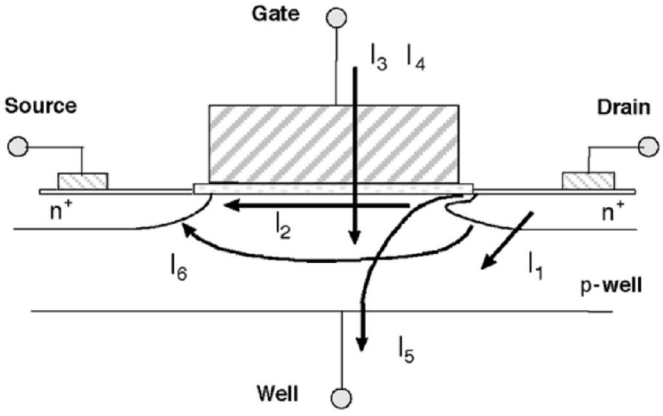

Fig. 1. Leakage current components

However, the largest amount of static power is still owed to subthreshold leakage current. It is the most temperature-dependent leakage component, and thus, every increase in dynamic power, produces an increment of the chip temperature, which in turn, increase the leakage current. This leakage component is also one of the main reasons why the scaling process is facing difficulties. Subthreshold leakage current can be expressed as follows:

$$I_{sub} = \frac{\mu W C_{ox}}{L} V_T^{\,2} e^{\frac{|V_{GS}|-|V_{TH}|}{\eta V_T}} \left(1 - e^{\frac{-|V_{DS}|}{V_T}}\right) \qquad (1)$$

where μ is carrier mobility, W and L channel width and lenght, respectively, C_{ox} the oxide capacitance, V_T the thermal voltage (26 mV at 25 °C), η DIBL (Drain Induced Barrier Lowering) coefficient, V_{GS} and V_{DS} voltages of gate and drain related to the source, respectively, and V_{TH} threshold voltage.

Four tunneling mechanisms (the gate to channel, bulk, source, and drain) as well as analytical expressions for gate leakage current can be found in [1,2].

3. FIXED TOPOLOGY POWER OPTIMIZATION

Power optimization techniques in this level do not alter the circuit topology, so the principle variables they affect are transistor sizes, supply voltages, and the threshold voltages. Some of the authors investigate the impact of single variable on circuit power consumption and delay while other perform thorough analysis considering mutual influence of two or

even more variables on design power consumption. There are a few commonly used power minimization techniques: gate-sizing, variable supply-voltage, variable threshold-voltage, multi voltage design, power gating, clock gating, stack forcing, on-chip optical interconnect, nano devices etc.

3.1. GATE SIZING AND SUPPLY VOLTAGE IN POWER MINIMIZATION

Decreasing transistor sizes enables higher densities of transistors on a chip. In order to control the power of the circuit, the power supply voltage is also reduced with each transistor scaling. Due to quadratic relationship between dynamic power consumption and power supply, this supply voltage reduction is the most effective way to lower the dynamic power. For CMOS circuits, a lower supply voltage means lower performance. This problem is solved by reducing the threshold voltage (V_{TH}) of a transistor. V_{TH} is defined as a gate-source voltage of MOSFET transistor, above which, the transistor is turned on. Ideally, if the gate voltage is below the threshold voltage, the transistor is not conducting any current. However, in practice there is still some current flowing from the drain to the source of a transistor. This is the subthreshold current. Its most important feature is that it increase exponentially with any V_{TH} decrease as shown in (1). That's why this leakage current is one of the main limiting factors to scaling process. SIA Roadmap [3] forecast supply voltage as low as 0.8 to 0.5 V in year 2018. Predicted threshold voltages are up to 0.1 V.

Fig. 2. depicts equi-speed and eque-power lines on V_{DD}-V_{TH} plane calculated from alpha-power MOSFET model [4].

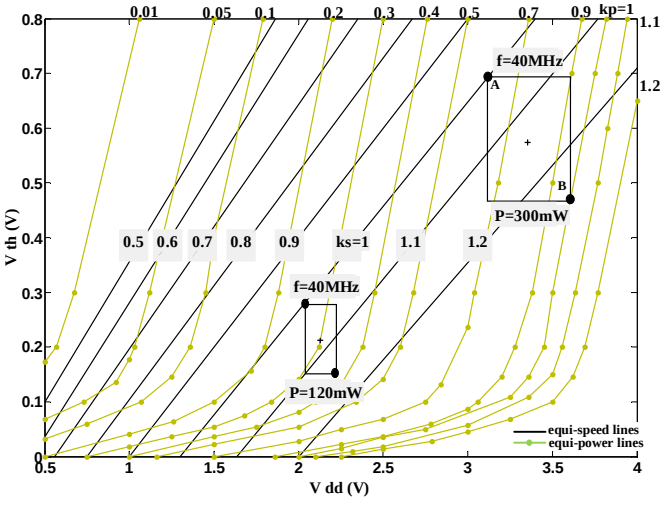

Fig. 2. Equi-power and equi-speed lines in V_{DD}-V_{TH} design space

If we, for example, have technology imposed constraints of V_{DD}=3.3 V ± 10% and V_{TH}=0.55 V ± 0.1 V, bigger rectangle in Fig. 2 defines design window. All the circuit specifications should be satisfied within the rectangle for yield-conscious designs. In the design window, circuit speed becomes slowest at the corner A while at the corner B power dissipation becomes the highest. Therefore, better tradeoffs between speed and power can be found by reducing fluctuations of V_{DD} and V_{TH} especially in low V_{DD}. The equi-speed and equi-power lines are normalized at the corners A and B by normalized factors *ks* and *kp*, respectively. Now, by sliding and sizing the design window on the V_{DD}-V_{TH} plane, it can be figured out how much speed and power dissipation are improved or degraded compared to the typical condition. For example, at V_{DD}=2.1 V ± 5% and V_{TH}=0.18 V ± 0.05 V power dissipation can be reduced to about 40% while maintaining the circuit speed.

Piguet et al. [5] conclude that between all the combinations of V_{DD}/V_{TH} guaranteeint the desired speed only one couple will result in the lowest power consumption. The location of this optimal working point and its associated total power consumption are tightly related to architectural and technology parameters. The same authors in [6] give the equation of total power consumption for circuits working at their optimal supply and threshold voltages (2). They use alpha-power law [4] and consider total power consumption as a sum of dynamic and sub-threashold leakage power.

$$P_{tot}^{\ opt} \cong \frac{aCNf}{(1-\chi A)^2}\left[nV_T\left(\ln\left(\frac{I_0}{2nV_T aCf}(1-\chi A)\right)+1\right)+\chi B\right]^2 \quad (2)$$

with *N* number of cells in the circuit; *a* average cell activity (i.e. the number of switching cells in a clock cycle over total number of cells); *C* equivalent cell capacitance; *f* operating frequency; I_0 avarage off-current per cell for $V_{GS}=V_{TH}$; *n* slope in weak inversion. *A* and *B* are two fitting variables that depend of α from alpha power law. Variable χ is equal to:

$$\chi^{\alpha} = \frac{\xi \cdot f \cdot LD}{I_0\left(\frac{e}{\alpha n V_T}\right)^{\alpha}} \quad (3)$$

with ζ (measured in Farad) a fitting parameter, which also includes the switching gate capacitance and *LD* the delay on critical path or logical depth.

The equation (2) is very important because it permits to analytically estimate the optimal total power directly from architectural parameters like activity (*a*), number of cells (*N*), frequency (*f*), logical depth (*LD*, included in χ) and technology parameters like average off-current (I_0), weak inversion slope (*n*), alpha power law coefficient (*α*, included in *A* and *B*) and delay coefficient (*ζ*, included in *χ*). Thus, starting from (3), it is possible to understand the impact of common architectural transformations, and to compare the performance of different technologies for a given architecture.

In [7] authors present closed-form formula for optimum supply and threshold voltages that minimize power dissipation when technology parameters and required speed are given. These formulas take the temperature into account.

Kuroda et al. [8] minimize supply voltage by applying variable supply-voltage (VS) technique on a 32-b RISC core processor developed in a 0.4μm CMOS technology. From an external supply, the VS scheme automatically generates minimum internal supply voltages by feedback control of a buck converter, a speed detector and a timing controller. Minimum internal supply voltage is determined so that critical-path delay is not changed. Performance in MIPS/W

is improved by a factor of more than two while area penalty because of VS scheme is smaller than 1%. The same authors in [9] introduce circuit technique for dynamically varying threshold voltage (VT) in order to reduce power dissipation of processor for portable multimedia equipment with HDTV-resolution. VT scheme consists of leakage current monitors (LCMs), self substrate bias (SSB) circuits, and a substrate charge injector (SCI) circuit. In the active mode, the SBB controls V_{BB} to compensate V_{TH} fluctuation. In standby mode, the SBB applies deeper V_{BB} to increase V_{TH} and cut off leakage.

Instead of using variable supply or threshold voltages many authors propose power minimization techniques with dual or multiple V_{DD} and/or V_{TH} [10,11]. The gates on critical paths operate at the higher V_{DD} or lower V_{TH}, while those on non-critical paths operate at the lower V_{DD} or higher V_{TH}, thereby reducing overall power consumption without performance degradation. Hamada et al. [12] derived a set of practical expressions for optimal number and values of discrete supplies and thresholds. They concluded that no more than three discrete values are needed for each tuning variable.

Among the leakage power reduction techniques power gating is commonly used to disconnect idle logic blocks from power network to curtail sub-threshold leakage [13]. The similar clock gating technique is used to prevent clock signal to give a pace to non-active gates. Stack forcing is another technique to tackle the ever-increasing leakage power. It has been shown that the stacking of two off transistors can significantly reduce leakage power than a single off transistor [14]. Stack arrangement of P-Channel MOS is preferred over N-Channel one because value of leakage current in PMOS is lesser as compared to NMOS. It results as the mobility of holes in PMOS is lesser than mobility of electrons in NMOS.

More recently, researchers have looked at doing multiple optimizations at once. So Brodersen et al. [15] search for tuning variable (among Vdd, Vth and gate sizing) with the largest capability for energy reduction and conclude that, to achieve the most energy-efficient design, the energy reduction potentials of all tuning variables must be balanced.

In analysis restricted to simple logic gates and inverter chains authors in [16,17] show that parasitic capacitances and velocity saturation of submicron technologies favor wider than minimum transistor sizes. Increasing the transistor size allow additional supply voltage reduction resulting in more substantial power savings. This is fruitful until the optimal sizing factor is reached because further increase in the device sizes will only deteriorate the performance and will consequently require an increase in supply voltage.

In 90nm technologies and beyond gate oxide leakage current has become comparable to subthreshold leakage. It is, therefore, necessary to develop methods for oxide leakage reduction, which unlike subthreshold leakage, occurs only in transistors that are ON as shown in Fig. 3. Increasing the oxide thickness will decrease gate oxide leakage current but this will be payed with substantial transistor delay. So, thick-oxide transistors in non-critical path will not speed down the circuit but will reduce static power consumption. Authors in [18] propose a new method that uses dual oxide-thickness process to minimize gate oxide leakage current reducint total leakage current more than 5 times with just a 5% delay penalty.

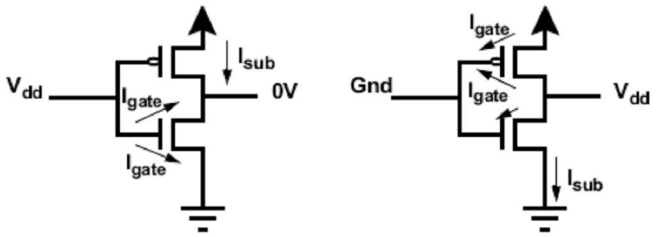

Fig. 3. Gate oxide leakage in CMOS inverter

3.2. ON-CHIP OPTICAL INTERCONNECTIONS FOR POWER MINIMIZATION

Global interconnect performance required for future generations of ICs cannot be achieved with metal. Using optical instead of electrical interconnections lead to decrease in power consumption, enormous bandwidth increase, immunity to electromagnetic noise, and reduced sensitivity to temperature variations. However, there are some difficulties in obtaining a large enough optical-electrical conversion efficiency. In [19] authors apply optical interconnection technology in clock distribution networks by replacing the electrical clock distribution tree with optical one. It can be seen that power dissipated by the electrical system is highly dependent on the operating frequency, while in the optical system, it remains almost the same. Also, power consumption is 5 times lower in optical than in an electrical network (at a frequency of 5GHz).

Another power minimization technique is on-chip wavelength division multiplexing. For example, a single waveguide could be used to replace a 64-bit bus, where each individual signal makes use of a distinct wavelength.

A possible solution for power reduction is to go towards nano-scale devices where a lower amount of charge is needed to code a bit. In these purposes single electron transistor is developed and used along with MOSFET's in building low-power gates [19].

4. TECHNOLOGY INFLUENCE ON TOTAL POWER CONSUMPTION

In order to practically demonstrate the influence on design's technology parameters on total power consumption the 12-bit binary divider logic circuit is described in VHDL and implemented in Xilinx FPGA devices from different families (Virtex-4, Virtex-5, Virtex-6, and Virtex-6 Lower Power). For binary division Radix-2 non-restoring algorithm with non-fractional remainder is used [20].

XPower CAD tool (within ISE 12.4 software) was utilized for power consumption measurements. Divider inputs were generated in MATLAB as signals with Gaussian distribution (1000 values for both dividend and divisor). Mean value, auto-correlation and cross-correlation of these signals are all equal to zero. Implementation results are presented in Fig. 4.

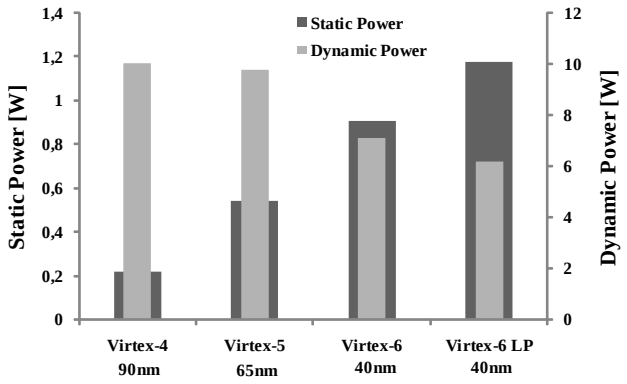

Fig. 4. Power consumption of the same logic design implemented in FPGAs with various technology properties

Technology process variations (from 90nm to 40nm), supply voltage variations (from 1.2 to 0.9 V), threshold voltage and transistor sizes variations obviously influence both static and dynamic power consumption. There is a clear increase in static power consumption when moving toward newer generation FPGA families. Static consumption increases due to increase in leakage currents as a consequence of shrink in transistor sizes. The shorter channel lengths and thinner gate oxides generally used at the new process node make it easier for current to leak, either across the channel region or through the gate oxide of the transistor. Concerning dynamic power, the core FPGA supply voltage and node capacitance generally reduce with each new process node, providing substantial dynamic power savings over previous generation FPGAs. The new 28nm Xilinx-7 series FPGAs will enable a 50% overall power reduction compared to previous 40nm generation [21].

5. CONCLUSION

Some power-aware system design methods are presented in this paper. These methods take into account static power consumption as more and more dramatic issue in very deep submicron technologies. The influence of technology parameters on design power consumption is demonstrated through implementation of binary dividers in FPGAs with the different technological properties. With the improvements of the technology parameters, there is an obvious trend in decreasing dynamic and increasing static power consumption of the design.

ACKNOWLEDGEMENT

This paper is supported by Project Grant TR33051 financed by Serbian Ministry of Education and Science.

REFERENCES

[1] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-submicrometer CMOS circuits," *in Proc. of the IEEE*, vol. 91, pp. 305-327, February 2003.

[2] R. Inagaki, N. Sadachika, D. Navarro, M.M. Mattasch, and Y. Inoue, "A Gate-Current Model for Advanced MOSFET Technologies Implemented in HiSIM2," ICCCAS, pp. 1057-106, July 2007.

[3] http://public.irts.net

[4] T. Sakurai, A.R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journ. of SSC*, vol. 25, pp. 584-549, April 1990.

[5] C. Piguet, C. Schuster, J.L. Nager, "Static and Dynamic Power Reduction by Architecture Selection," PATMOS, pp. 659-668, July 2006.

[6] C. Schuster et al.,"Architectural and Technology influence on the Optimal Total Power Consumption," DATE, pp. 13-19, March 2006.

[7] K. Nose, T. Sakurai, "Optimization of V_{DD} and V_{TH} for Low-Power and High-Speed Applications," ASP-DAC, pp. 469-474, June 2000.

[8] T. Kuroda et al., "Variable Supply-Volgate Scheme for Low-Power High-Speed CMOS Digital Design," *IEEE Journ. of SSC*, vol. 33, pp. 454-462, March 1998.

[9] T. Kuroda et al., "A 0.9V 150MHz 10mW 4mm2 2-D discrete cosine transform core processor with variable threshold-voltage scheme," *IEEE Journ. of SSC*, vol. 31, pp. 1770-1779, February 1996.

[10] F. Gao, J.P. Hayes,"Total Power Reduction in CMOS Circuits via Gate sizing and Multiple Threshold Volgates," ASP-DAC, pp. 31-36, June 2005.

[11] W. Hung et al., "Total Power Optimization through Simultaneously Multiple-Vdd Multiple-Vth Assignments and Device Sizing with Stack Forcing," ISLPED, pp. 144-149, May 2004.

[12] M. Hamada, Y. Otaguro, T. Kuroda, "Utilizing Surplus Timing for Power Reduction," CICC, pp. 89-92, 2001.

[13] S.Y. Chen et al.,"Power gating design for standard-cell-like structured ASICs," DATE, pp. 514–519, November 2010.

[14] S. Narendra et al.,"Scaling of Stack Effect and its Application for Leakage Reduction," ISLPED, pp. 195–200, 2001.

[15] R. Brodersen et al.,"Methods for True Power Minimization," ICCAD pp. 35–42, 2002.

[16] J. Rabey, A. Chandrakasan, B. Nikolic, Digital Integrated Circuits: A Design Perspective, second edition, New York, Prentice Hall, 2003.

[17] R. Rogenmoser, H. Kaeslin, "The Impact of Transistor Sizing on Power Efficiency in Submicron CMOS Circuits," IEEE Journ. of SSC, vol. 32, pp. 1142-1145, 1997.

[18] D. Lee et al., "Simultaneous State, Vt and Tox Assignments for Total Standby Power Minimization," DATE pp. 494–499, 2004.

[19] C. Piguet et al.,"Extremely Low-Power Logic," DATE, pp. 656–661, 2004.

[20] B. Jovanović, M. Jevtić,"FPGA implementation of throughput increasing techniques of the binary dividers," UNITECH, pp. 397-401, 2010.

[21] Xcell Journal, Fourth Quarter 2010, Issue 73, pp. 65, www.xilinx.com/xcell