

A Flexible Power Model for FPGAs [★]

Kara K.W. Poon, Andy Yan, and Steven J.E. Wilton

Department of Electrical and Computer Engineering,
University of British Columbia, Vancouver, BC, Canada
{karap,ayan,steve}@ece.ubc.ca
<http://www.ece.ubc.ca/~steve>

Abstract. This paper describes a flexible power model for FPGAs. The model estimates the dynamic, short circuit, and leakage power for a wide variety of FPGA architectures. Such a model will be essential in the design and research of next-generation FPGAs, where power will be one of the primary optimization goals. The model has been integrated into the VPR CAD flow, and is available to the research community for use in FPGA architectural and CAD tool experimentation.

1 Introduction

Power dissipation is becoming a major concern for semiconductor vendors and customers [1]. According to [2], if current design trends continue, a typical microprocessor (MPU) will consume 50 times more power than that can be supported by cost-effective packaging techniques by 2016. A study in [3] indicates that power will become one of the two most serious design concerns (along with design complexity) in coming process generations. FPGAs will not escape this trend; already, FPGA vendors report that power consumption is one of the primary concerns of their customers. Compared to ASICs and other custom chips, FPGAs contain long routing tracks with significant parasitic capacitance; during high speed operations, the switching activity on these long routing tracks causes significant power dissipation.

There have been several low-power architectures described in previous works [4] [5] [6]. However, these papers present "point solutions," in that each only considers a single architecture. In order to migrate these low-power techniques to commercial FPGAs, it is critical that researchers be able to estimate power for a wide variety of architectural parameters. To do this, a power model that is flexible enough to target many different FPGA architectures is required.

There have also been numerous CAD algorithms that target low power [7] [8] [9]. Often, these studies rely primarily on reducing switching activity to result in a low-power solution; although reducing switching activity does lower the power, power also depends on the architecture, the lengths of critical signal routes, the rise and fall times of the signals, and the amount of static power.

[★] This work was supported by Micronet, the British Columbia Advanced Systems Institute, and the Natural Sciences and Engineering Research Council of Canada

Though neglected in the past, static power is expected to become an increasingly important part of the total power. In order to adequately evaluate these new CAD algorithms and techniques, a detailed power model is needed to take all these factors into account.

This paper presents a detailed and flexible power model for FPGAs, which includes terms for dynamic power, short-circuit power, and leakage power. Although the techniques we employ have been used before, the integration of these techniques into a flexible power model for FPGAs is a novel approach. The model is flexible enough to target FPGAs with different look-up table (LUT) sizes, different interconnect strategies (segment length, switch block type, connection flexibility), different cluster sizes (for a hierarchical FPGA), and different process technologies. As described above, a model such as the one in this paper will become an essential part of any FPGA architect's and CAD tool designer's arsenal.

The power model has been integrated into the widely-used VPR CAD tool, which already contains detailed area and delay models [10]. Section 2 describes this framework. The model itself is described in Section 3. Section 4 shows how the model can be used to evaluate FPGA architectures. The model is publicly available for use by the FPGA research community (refer to appendix A).

2 Frameworks

The model has been integrated into the VPR CAD tool, which is commonly used by the FPGA research community [11]. As shown in figure 1, VPR contains a place and route tool, and detailed area and delay models. The place and route tool maps a benchmark circuit to an FPGA architecture, using estimates from the area and delay model to guide the tool. A description of the architecture is provided to the tool using an architecture file; the architecture file contains information such as segment length, connection topologies, logic block size and composition, and process parameters. The flexibility of this architecture file is key to VPR; the CAD tool is flexible enough to target any architecture that can be specified in the architecture file. The area and delay models are also used to estimate the area and critical path delay of the circuit after placement and routing has completed.

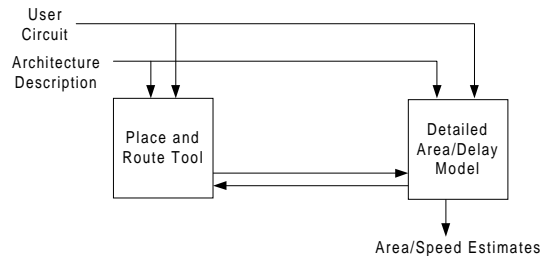


Fig. 1. Original VPR Framework

Figure 2 shows how we have included the power model in the VPR framework. An activity estimator was developed which estimates the switching frequencies of all nodes in the circuit; the activity estimator will be described in section 3.1. The detailed power model was integrated into the area/delay model. This model is described in sections 3.2 to 3.4. In our current implementation, the activity estimator and the power model are not used to guide the placement and routing; the model is only used to estimate the power after placement and routing has occurred. It is possible using this framework, however, to use power estimates during placement and routing to optimize the implementation for power.

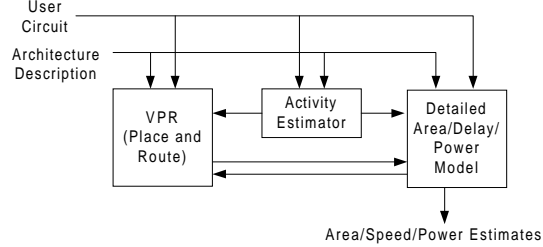


Fig. 2. VPR Framework with Power Model

3 Power Model

Our power model is aimed at the island-style FPGA architectures. Figure 3(a) shows a typical island-style FPGA architecture, which has logic blocks, switch blocks, connection blocks, and routing. We also assume that an H-tree clock network is employed, as illustrated in figure 3(b).

3.1 Activity Generation

Our power model uses the transition density signal model to determine signal activities within the FPGA [12] [13]. The transition density of a signal is the expected number of toggles for the signal in each clock cycle. It is calculated as follows:

For each LUT in the circuit, the function implemented by that LUT can be expressed as a function $f(x)$. For each input, x_i , two new boolean functions $f(x_i)$ and $f(\bar{x}_i)$ can be generated from $f(x)$ by setting input x_i to 1 and 0 respectively. To illustrate, consider the following example:

Example 1. If $f(x) = x_1x_2 + \bar{x}_1x_3$, then $f(x_1) = x_2$ and $f(\bar{x}_1) = x_3$. where x_1 , x_2 , and x_3 are inputs and $f(x)$ represents the output of the LUT.

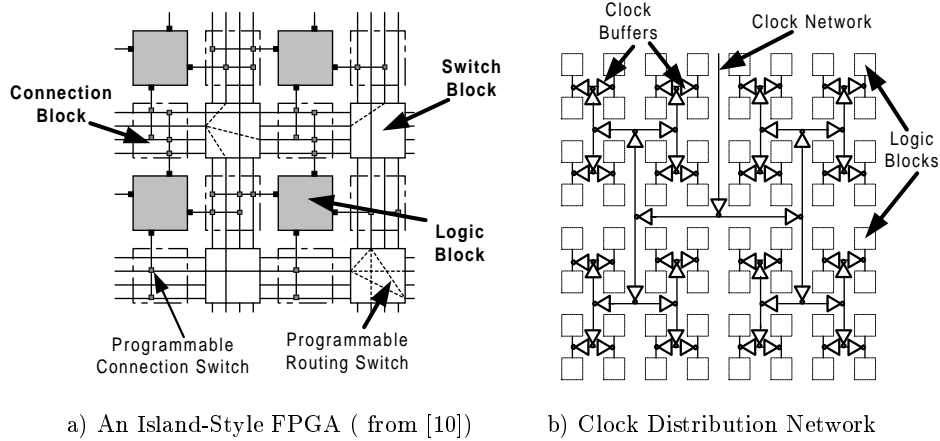


Fig. 3. Assumed Architectural Model for FPGAs

A *boolean difference* of the output with respect to an input, x_i , is then calculated [13]:

$$\frac{df(x)}{dx_i} = f(x_i) \oplus f(\bar{x}_i) \quad (1)$$

The probability of the *boolean difference*, $P(df(x)/dx_i)$, is then the static probability that a change in x_i causes a change at the output. Because of the uncorrelated input assumption of the transition density signal model, each input contributes a static probability, $P(df(x)/dx_i)$, and a transition density, $D(x_i)$, to the total density, $D(y)$, at the output [13].

$$D(y) = \sum_{\text{all inputs}} P\left(\frac{df(x)}{dx}\right) D(x) \quad (2)$$

We assume that the primary inputs to our circuits are random; all primary inputs have a static probability of 0.5 and a transition density of 0.5. These values can be adjusted to reflect other user inputs.

The extension of the transition density model to sequential circuits is straightforward. For each D flip-flop, the output probability is the same as the input probability. The transition density of the output, $D(y)$, of the flip-flop can be written as:

$$D(y) = 2P(x)(1 - P(x)) \quad (3)$$

where $P(x)$ represents the input probability. For the sequential feedback loops, we first determine the static probability at the output of the LUT by iterations, and then apply (3) for the transition density at the output.

3.2 Dynamic Power

Dynamic power is dissipated every time when a signal changes due to the charging and discharging of the load and parasitic capacitance associated with the

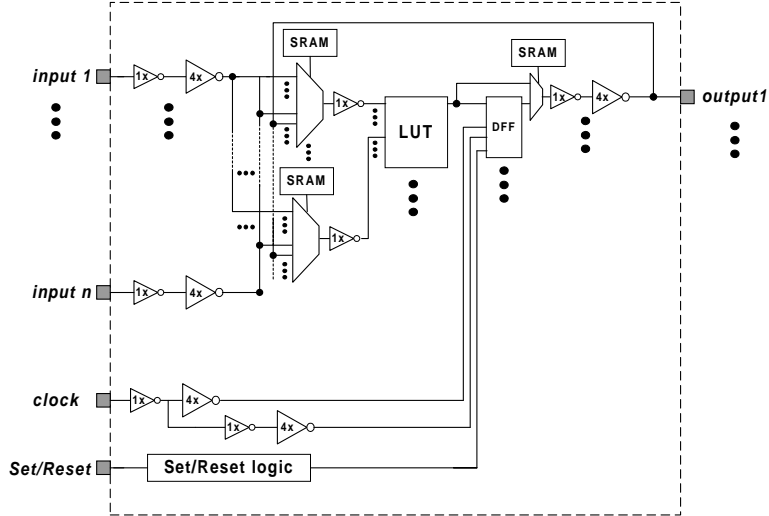


Fig. 4. Schematic of a logic block (from [10])

signal. The dynamic power is the dominant portion of the total power in today’s technologies. Using the transition density model [13], the total power dissipation is:

$$\text{Dynamic Power} = \sum_{\text{all nodes}} 0.5C_y V^2 D(y) f_{\text{clk}} \quad (4)$$

where V is the swinging voltage of each node; C_y is the capacitance being charged and discharged during each transition; $D(y)$ is the individual transition density at each node, and f_{clk} is the clock frequency of the circuit. For each signal, we can apply this equation using the transition density calculated in section 3.1 and the node capacitance estimated by VPR. The clock frequency can be determined by the critical path of the circuit.

We estimate the power dissipated both within the logic blocks and in the global routing network. The assumed logic block structure is shown in figure 4 [10]; the model is flexible enough to account for any number of LUTs and any number of internal routing wires. The length of each internal routing wire is calculated based on the number of LUTs within the logic block; this length is then used to estimate the parasitic capacitance of each internal routing wire. The structure of the LUTs and multiplexers significantly affects the power estimates; we assume the LUTs and multiplexers are implemented using a tree of 2-input multiplexers, as shown in figure 5. We estimate the transition density and parasitic capacitance of each node in figures 5(a) and 5(b) using the transistor sizes in [10] and layout assumptions in [14], and use these results to estimate the power within each LUT and multiplexer.

Note that the internal nodes of each local routing multiplexer swing between the power supply voltage and ground because the gate voltage of the pass transis-

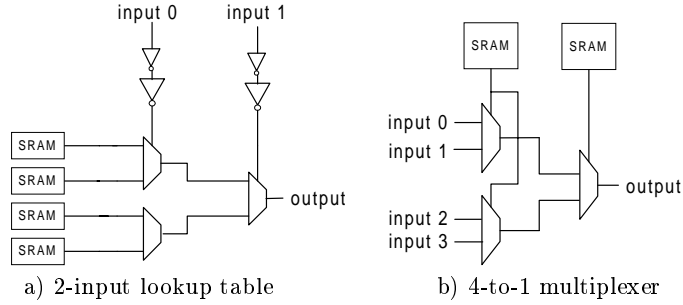


Fig. 5. Implementation of a LUT and MUX using 2-input multiplexers

tors in the SRAM cells has been boosted to overcome the body effect. However, the internal nodes of the LUT toggle at a smaller voltage range since the body effect is applied to the pass transistors inside the multiplexers. We verified our results using HSPICE by simulating different sizes of LUTs and multiplexers with inputs toggling from 5Mhz to 50Mhz. By shifting the delay of each input to model the inputs with different input correlation, we gathered the minimum, average, and maximum values of the average dynamic power dissipation for each size of LUT and multiplexer. From our analysis, the power dissipation calculated by our model matches the maximum values of our simulation results. This is because the transition density signal model that we used assumes uncorrelated inputs. The power dissipated in the global routing network was calculated using the transition densities of each network node, and the capacitance of each node from [10]. For the clock network, we assume a metal-5 H-tree network, as shown in figure 3(b). For very large FPGAs, this results in long wire segments between the clock buffers. FPGA vendors can add buffers to shorten wire segments. However, since we want our model to apply to FPGAs of any size, we need a way, within the model, to estimate an appropriate number of buffers, and to use this number of buffers in the power estimation. To do this, we use a distributed RC model, find the delay as a function of the number of clock buffers, and differentiate the function to find the optimum clock buffer size. This ensures that we have a realistic clock network for any given architecture.

3.3 Short-Circuit Power

Short-circuit power is the power dissipated through a direct current path formed between the power supply and the ground during the rise and fall times of each transition. Short circuit power is a function of the rise and fall time and the load capacitance [15]. Based on calculations using Altera and Xilinx datasheets [16] [17], we assume that short-circuit power is 10% of dynamic power.

3.4 Leakage Power

The leakage power dissipation comes from two sources: reverse-bias leakage power and sub-threshold leakage power. As the majority of the leakage power is

from the sub-threshold current [18], we assume the reverse bias leakage current to be negligible. To estimate the sub-threshold current, we use [18]:

$$I_{\text{drain}}(\text{weak inversion}) = I_{\text{on}} \exp \left[\frac{(V_{\text{gs}} - V_{\text{on}})q}{nkT} \right] \quad (5)$$

V_{on} is defined as the boundary between the weak and strong inversion regions. To calculate V_{on} , we use the following equations [18]:

$$V_{\text{on}} = V_{\text{t}} + \frac{nkT}{q} \quad (6)$$

where

$$n = 1 + \left(\frac{qN_{\text{FS}}}{C_{\text{ox}}} \right) + \left(\frac{C_{\text{d}}}{C_{\text{ox}}} \right) \quad (7)$$

I_{on} is the drain current at the boundary when V_{gs} is equal to V_{on} . We apply the velocity saturation model [20] to calculate I_{on} .

$$I_{\text{on}} = \frac{W v_{\text{sat}} C_{\text{ox}} (V_{\text{gs}} - V_{\text{t}})^2}{(V_{\text{gs}} - V_{\text{t}}) + E_{\text{c}} L_{\text{eff}}} \quad (8)$$

where W is the device width, v_{sat} is electron velocity, E_{c} is the piecewise carrier drift velocity, L_{eff} is the effective source-drain channel length, V_{gs} is the gate-source voltage, V_{t} is the threshold voltage, and C_{d} is the drain capacitance. The constants k and q are Boltzman's constant and elementary charge, respectively. T is the temperature in Kelvins.

N_{FS} is defined as the number of fast superficial states; it is a current fitting parameter that determines the slope of the sub-threshold current-voltage characteristic [18]. To determine the N_{FS} values of NMOS and PMOS, we have run HSPICE simulations for both NMOS and PMOS transistors which operate within the junction temperature range, from -40 to 100 °C [16] [17]. All the inactive transistors, including unused switches, SRAM cells, and inactive pass transistors in the LUTs and multiplexers, contribute to the leakage power for the FPGA. In our model, we assume the gate-source voltage of these inactive transistors to be half of their threshold voltages.

4 Experimental Results

To illustrate an application of our model, we investigated how the values of various architectural parameters affect the power dissipation. We used 36 MCNC benchmark circuits, and varied the segment length¹, the cluster size², and the LUT size. Energy rather than power is used for comparing our results to avoid bias from the circuits' clock frequencies.

¹ Segment length is defined as the number of logic blocks spanned by a wire segment.

² Cluster size is defined as the number of LUTs contained in a logic block.

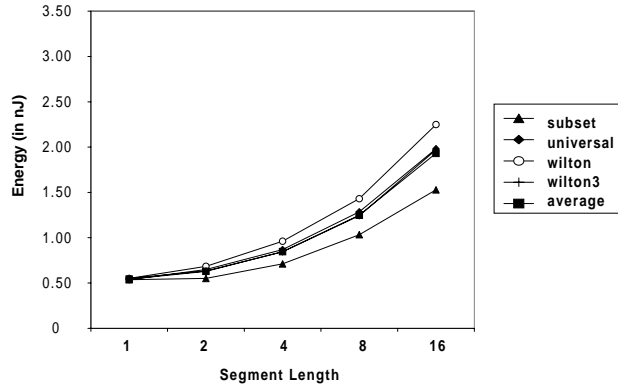


Fig. 6. Routing energy versus segment length

Figure 6 illustrates the relationship between the routing energy and the segment length, using four previously published switch blocks [19]. The graph shows that the longer the global routing wire, the more routing energy is dissipated. This result indicates that short wires are more favorable to reduce power consumption on FPGA interconnects.

Figures 7(a) and 7(b) show how the LUT and cluster sizes affect total energy. As figure 7(a) shows, increasing cluster size does not cause any significant changes on the routing energy, but it slightly increases the logic block energy and significantly decreases the clock energy dissipation. This is because as more LUTs are included in a logic block, more internal wires for local connections and longer routing wires for block-to-block connections are required, but each logic block can implement more complicated functions. As a result, the total number of logic blocks required decreases, and therefore, reduces the number of clock branches and routing branches.

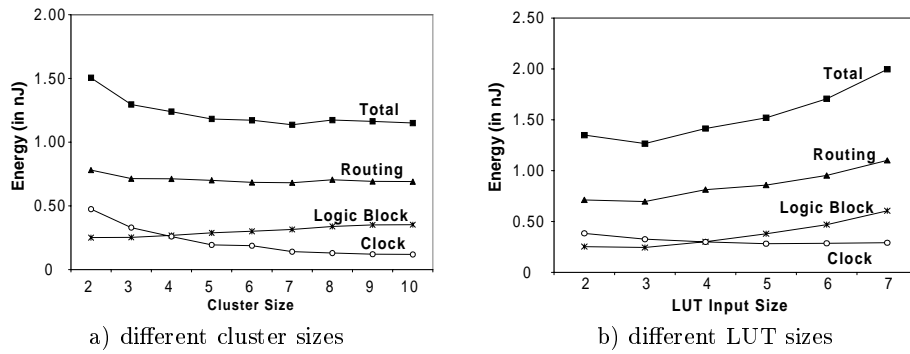


Fig. 7. Energy versus cluster and LUT sizes

As shown in figure 7(b), energy consumed by the logic blocks and routing fabric increase with the LUT size. This is because a larger LUT contains more internal nodes, and results in longer global wiring segments. The behavior of the clock energy is the result of two competing factors: smaller LUTs means more LUTs are required to implement a circuit, and hence, more clock branches are required; on the other hand, the length of each clock branch increases as the LUT size increases. Overall, a three-input lookup table appears to be a good choice.

Our results show that approximately 57% of the total energy consumption is due to the routing fabric, 24% is due to the logic blocks, and 19% is due to the clock distribution network.

5 Conclusions

In this paper, we have presented a detailed power model that is flexible enough to estimate the power dissipation on a wide variety of island-style FPGA architectures. The power model is available to the research community, and has been integrated into the popular VPR CAD tool suite. We have shown how the model can be used to investigate architectural tradeoffs; the model can also be used to estimate the effectiveness of CAD tools that attempt to minimize power.

Acknowledgments

The authors would like to thank Dr. Resve Saleh for his helpful discussions and Dr. Vaughn Betz for supplying us with the VPR place and route tool.

References

1. The International Technology Roadmap for Semiconductors, 2001 Edition, International Sematech, Austin, Texas, 2001.
2. A. Allan, D. Edenfeld, W. Joyner Jr, A. Khang, M. Rogers, Y. Zorian: "2001 Technology Roadmap for Semiconductors". Computer, Vol. 35, Issue 1, Jan 2002, pp. 42-53
3. G. Lim, R. Saleh, "Trends in Low Power Digital Systems on Chip Design", in International Symposium on Quality of Electronic Design, March 2002.
4. V. George, H. Zhang, J. Rabaey, "The design of a low energy FPGA", in proceedings of the Low Power Electronics and Design, August 1999.
5. L. Shang, A. S. Kaviani, K. Bathala, "Dynamic Power Consumption in Virtex-II FPGA Family", ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, February 2002.
6. F. G. Wolff, M. J. Knieser, D. J. Weyer, C. A. Papachristou, "High-level low power FPGA design methodology" National Aerospace and Electronics Conference, October 2000.
7. A. Garcia, W. Burlison, J. Danger, "Power Modelling in Field Programmable Gate Arrays (FPGA)", in Proceeding of Field-Programmable Logic and Applications, pp. 396-404, September 1999.

8. L. Shang, N. K. Jha, "High-level power modeling of CPLDs and FPGAs", International Conference on Computer Design, September 2001.
9. K. Roy, "Power-Dissipation Driven FPGA Place and Route Under Timing Constraints", IEEE Transactions on Circuits and Systems: Fundamental Theory and Applications, Vol 46, No 5 May 1999
10. V. Betz, Architectures and CAD for Speed and Area Optimizations of FPGAs. PhD thesis, University of Toronto, 1998.
11. V. Betz, VPR and T-VPack User's Manual. version 4.30, March 2000.
12. F. N. Najm, "A Survey of Power Estimation Techniques in VLSI Circuits", IEEE Trans. on Very Large Scale Integration (VLSI) Systems, Vol 2, No 4, pp 446-455, December 1994.
13. G. Yeap, Practical Low Power Digital VLSI Design. Kluwer Academic Publishers, 1998
14. S. J. E. Wilton, N. P. Jouppi, "CACTI: An Enhanced Cache Access and Cycle Time Model", in IEEE Journal of Solid-State Circuits, Vol 31, No 5, pp 677-687, May 1996.
15. D. Eckerbert, P.L. Edefors, "Interconnect-Driven Short-Circuit Power Modeling", in Proceedings of Euromicro Symposium on Digital Systems, September 2001.
16. Altera. APEX 20K Programmable Logic Device Family Data Sheet. version 4.1. Altera Corporation. September 2001.
17. Xilinx. Virtex-E 1.8V Field Programmable Gate Arrays Data Sheet. version 2.2. Xilinx Corporation. November 2001.
18. S.M. Kang, Y. Leblebici, CMOS Digital Integrated Circuits: Analysis and Design. 1999
19. M. I. Masud, S. J. E. Wilton, "A New Switch Block for Segmented FPGAs", in Proceeding of Field-Programmable Logic and Applications, pp. 396-404, September 1999.
20. K.Y. Toh, P.K. Ko, R.G. Meyer, "An Engineering Model for Short-Channel MOS Devices", IEEE Journal of Solid-State Circuits, Vol. 23, No. 4, August 1988.

Appendix A

The model can be downloaded free for non-commercial use by the research community. Instructions and a download link are available from <http://www.ece.ubc.ca/~stevew/powermodel.html>