Full text available at: http://dx.doi.org/10.1561/1000000022

Power Modeling and Characterization of Computing Devices: A Survey

Power Modeling and Characterization of Computing Devices: A Survey

Sherief Reda

Brown University
Providence 02912
USA
sherief_reda@brown.edu

Abdullah N. Nowroz

Brown University
Providence 02912
USA
Abdullah_nowroz@brown.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Electronic Design Automation

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is S. Reda and A. N. Nowroz, Power Modeling and Characterization of Computing Devices: A Survey, Foundations and Trends® in Electronic Design Automation, vol 6, no 2, pp 121–216, 2012

ISBN: 978-1-60198-560-6

Foundations and Trends® in Electronic Design Automation
Volume 6 Issue 2, 2012

Editorial Board

Editorial Scope

Foundations and Trends® in Electronic Design Automation will publish survey and tutorial articles in the following topics:

- System Level Design
- Behavioral Synthesis
- Logic Design
- Verification
- Test
- Physical Design
- Circuit Level Design
- Reconfigurable Systems
- Analog Design

Information for Librarians
Foundations and Trends® in Electronic Design Automation, 2012, Volume 6, 4 issues. ISSN paper version 1551-3939. ISSN online version 1551-3947. Also available as a combined paper and online subscription.

Foundations and Trends® in
Electronic Design Automation
Vol. 6, No. 2 (2012) 121–216

DOI: 10.1561/1000000022

now
the essence of knowledge

Power Modeling and Characterization of Computing Devices: A Survey

Sherief Reda[1] and Abdullah N. Nowroz[2]

[1] *Brown University, 182 Hope st, Providence 02912, USA, sherief_reda@brown.edu*
[2] *Brown University, 182 Hope st, Providence 02912, USA, Abdullah_nowroz@brown.edu*

Abstract

In this survey we describe the main research directions in pre-silicon power modeling and post-silicon power characterization. We review techniques in power modeling and characterization for three computing substrates: general-purpose processors, system-on-chip-based embedded systems, and field programmable gate arrays. We describe the basic principles that govern power consumption in digital circuits, and utilize these principles to describe high-level power modeling techniques for designs of the three computing substrates. Once a computing device is fabricated, direct measurements on the actual device reveal a great wealth of information about the device's power consumption under various operating conditions. We describe characterization techniques that integrate infrared imaging with electric current measurements to generate runtime power maps. The power maps can be used to validate design-time power models and to calibrate computer-aided design

tools. We also describe empirical power characterization techniques for software power analysis and for adaptive power-aware computing. Finally, we provide a number of plausible future research directions for power modeling and characterization.

Contents

1

Introduction

In the past decade power has emerged as a major challenge to computing advancement. A recent report by the National Research Council (NRC) of the National Academies highlights power as the number one challenge to sustain historical improvements in computing performance [39]. Power is limiting the performance of both mobile and server computing devices. At one extreme, embedded and portable computing devices operate within power constraints to prolong battery operation. The power budgets of these devices are about tens of milli-Watts for some embedded systems (e.g., sensor nodes), 1–2 W for mobile smart phones and tablets, and 15–30 W for laptop computers. At another extreme, high-end server processors, where performance is the main objective, are increasingly becoming hot-spot limited [46], where increases in performance are constrained by a maximum junction temperature (typically 85°C). Economic air-based cooling techniques limit the total power consumption of server processors to about 100–150 W, and it is the spatial and temporal allocation of the power distribution that leads to hot spots in the die that can comprise the reliability of the device. Because server-based systems are typically deployed in data centers, their aggregate performance becomes power

limited [6], where energy costs represent the major portion of total cost of ownership. The emergence of power as a major constraint has forced designers to carefully evaluate every architectural and design feature with respect to its performance and power trade-offs. This evaluation requires pre-silicon power modeling tools that can navigate the rich design landscape. Furthermore, runtime constraints on power consumption require power management tools that control a number of runtime knobs that trade-off performance and power consumption. Power management techniques that seek to meet a power cap, e.g., as in the case of servers in data centers [6, 37], require either direct power measurements when feasible, or alternatively, runtime power modeling techniques that can substitute direct characterization. In addition, software power characterization can help tune and restructure algorithms to reduce their power consumption.

The last decade has seen a diversification in possible computing substrates that offer different trade-offs in performance, power, and cost for different applications. These substrates include application-specific custom-fabricated circuits, application-specific circuits implemented in field-programmable logic arrays (FPGAs), general-purpose processors whose functionality is determined by software, general-purpose graphical processing units (GP-GPUs), digital signal processors (DSPs), and system-on-chip (SoC) substrates that combine general-purpose cores with heterogeneous application-specific custom circuits. None of these substrates necessarily dominate the other, but they rather offer certain advantages that depend on the target application and the deployment setting of the computing device. For instance, custom fabricated circuits outperform their FPGA counterparts in performance and power, but they are more expensive. SoCs offer higher performance/Watt ratio for a range of applications than general-purpose processors; however, general-purpose processors offer higher throughput for scientific applications. GPGPUs are also emerging as a strong contender to processors and FPGAs; however, the relative advantage of each of these substrates differs by the application [39, 54, 76]. Sorting out the exact trade-offs of all these substrates across different application domains is an active area of research [4, 29, 76]. While power modeling and characterization for these substrates share common concepts, each of these substrates

has its own peculiarities. In this survey we will discuss the basic power modeling and characterization concepts that are shared among these substrates as well as the specific techniques that are applicable for each one.

Pre-silicon power modeling and post-silicon power characterization are very challenging tasks. The following factors contribute to these challenges.

(1) Large die areas with billions of transistors and interconnects lead to computational difficulties in modeling.
(2) Input patterns and runtime software applications trigger large variation in power consumption. These variations are computationally impossible to enumerate exhaustively during modeling.
(3) Spatial and temporal thermal variations arising from power consumption trigger large variations in leakage power, which lead to intricate dependencies in power modeling.
(4) Process variabilities that arise during fabrication lead to intra-die and inter-die power leakage variations that are unique to each die. These deviations recast the modeling results to be educated guesses, rather than exact estimates.
(5) Practical limitations on the design of power-delivery networks make it difficult to directly characterize the runtime power consumption of individual circuit blocks.

The objective of this survey is to describe modern research directions for pre-silicon power modeling and post-silicon power characterization. Pre-silicon power modeling tools estimate the power consumption of an input design, and they can be used to create a power-aware design exploration framework, where different design choices are evaluated in terms of their power impact in addition to traditional design objective such as performance and area. Post-silicon power characterization tools are applied to a fabricated design to characterize its power consumption under various workloads and environmental variabilities. The results of power characterization are useful for power-related debugging issues, calibration of design-time power modeling tools, software-driven power analysis, and adaptive

power-aware computing. Our technical exposition reviews power modeling and characterization techniques of various computing substrates, while emphasizing cross-cutting issues. We also connect the dots between the research results of different research communities, such as circuit designers, computer-aided design (CAD) developers, computer architects, and system designers. Our discussions reveal the shared concepts and the different research angles that have been explored for power modeling and characterization.

1.1 Computing Substrates

1.1.1 General-Purpose Processors

A general-purpose processor is designed to serve a large variety of applications, rather than being highly tailored to one specific application or a class of applications. The design of a general-purpose processor has to be done carefully to lead to good performance within the processor's thermal design power (TDP) limit under different kinds of workloads. The TDP limit has forced a significant change in the design of processors. At present, designers aim to increase the processor's total throughput rather than improving the single-thread performance. This throughput increase is achieved by using more than one processing core per chip.

Figure 1.1 gives an example of a quad-core processor based on Intel's Core i7 Nehalem architecture. The 64-bit processor features four cores that share an 8 MB of L3 cache. The cores can run up to 3.46 GHz in a 130 W TDP. Each core has a 16-stage pipeline and includes a 32 KB L1 instruction cache, a 32 KB L1 data cache, and a 256 KB of L2 cache. The front-end of the pipeline can fetch up to 16 bytes from the L1 instruction cache. The instructions in the fetched 16 bytes are identified and inserted into an instruction queue. The decoder unit receives its inputs from the instruction queue, and it can decode up to four instructions per cycle into micro-ops. A branch prediction unit with a branch target buffer enables the core to fetch and process instructions before the outcome of a branch is determined. The back-end of the pipeline allocates resources for the micro-ops and renames their source and destination registers to eliminate hazards and to expose instruction-level

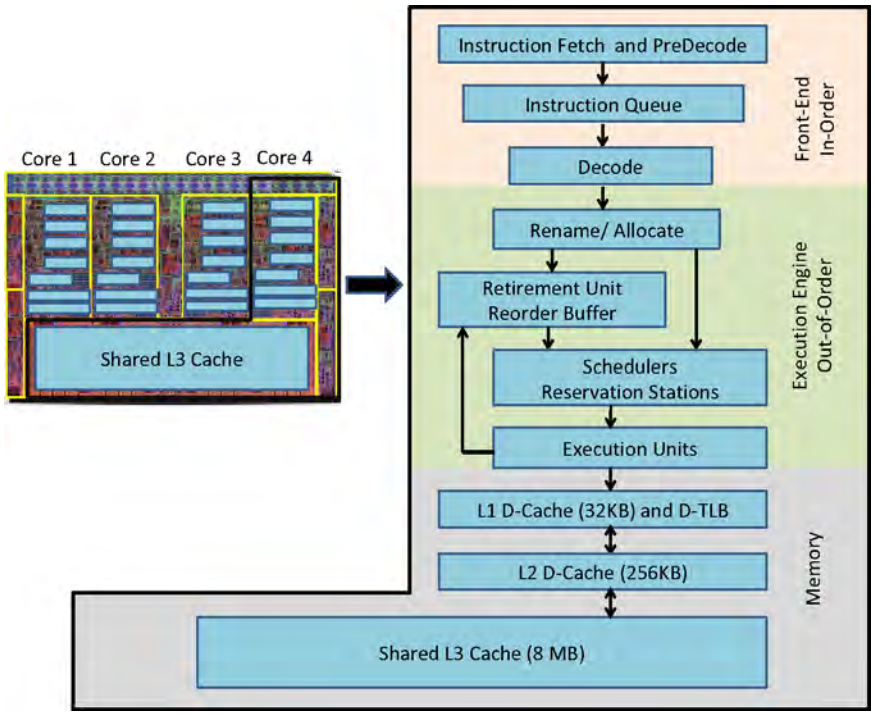

Fig. 1.1. High-level diagram of Intel Core i7 processor (Nehalem architecture).

parallelism. The micro-ops are then queued in the re-order buffer until they are ready for execution. The pipeline can dynamically schedule and issue up to six micro-ops per cycle to the execution units as long as the operands and resources are available. The execution units perform loads, stores, scalar integer or floating-point arithmetic, and vector integer or floating-point arithmetic. The results from the execution of micro-ops are stored in the re-order buffer, and results are committed in-order only for correct instruction execution paths.

1.1.2 Embedded SoC

SoCs are computational substrates that are targeted for embedded systems and mobile computing platforms for a certain niche of applications. An SoC for a smart phone or a tablet typically consumes less than 1–2 W of power, while delivering the throughput required for

applications that include video and audio playback, internet connectivity, and games. In contrast to a general-purpose processor, an SoC includes, in addition to the general-purpose core(s), application-specific custom hardware (HW) components that can provide the required throughput for the target applications within the power envelope of the embedded system. Because total die area is constrained by cost and yield considerations, the inclusion of application-specific custom HW components must come at the expense of the functionality of the general-purpose core. SoC general-purpose cores are less capable than the ones used in general-purpose processors. They are usually less aggressively pipelined with limited instruction-level parallelism capabilities and smaller cache sizes.

Figure 1.2 gives an example of an SoC based on nVidia's Tegra platform that has a total power budget of about 250 mW. The SoC features a 32-bit ARM11 general-purpose core that runs up to 800 MHz. The ARM11 core has an 8-stage pipeline, with a single instruction issue and support for out-of-order completion. The L1 data and code cache memory sizes are 32 KB each, and the size of the L2 cache is 256 KB. The performance specifications of the core are clearly inferior compared to the specifications of the Core i7. To compensate for the lost general-purpose computing performance, the SoC uses a number of application-specific components to deliver the required performance within its

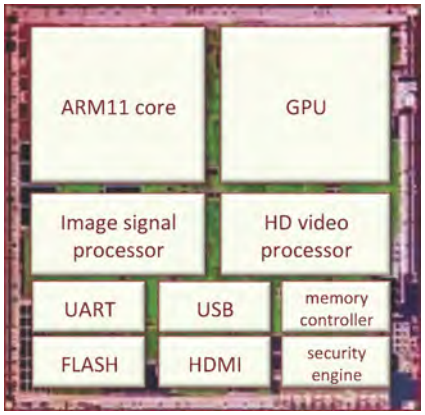

Fig. 1.2. Example of nVidia Tegra SoC.

power budget. These include an image signal processor that can provide image processing functions (e.g., de-noising, sharpening, and color correction) for images captured from embedded cameras. The SoC includes a high-definition audio and video processor for image, video and audio playback, and a GPU to deliver the required graphics performance for 3-D games. The SoC supports an integrated memory controller, an encryption/decryption accelerator component, and components for communication, such as Universal Asynchronous Receiver/Transmitter (UART), Universal Serial Bus (USB), and High-Definition Multimedia Interface (HDMI). All SoC components communicate with each other using an on-chip communication network, which can take a number of forms, including shared and hierarchical busses, point-to-point busses, and meshes.

1.1.3 Field-Programmable Gate Arrays

Soaring costs associated with fabricating computing circuitry at advanced technology nodes have increased the interest in programmable logic devices that can be configured after fabrication to implement user designs. The most versatile programmable logic currently available is Field Programmable Gate Arrays (FPGAs). The basic FPGA architecture is an island-style structure, where programmable *logic array blocks (LABs)* are embedded in a reconfigurable wiring fabric that consists of wires and switch blocks as illustrated in Figure 1.3. The inputs and outputs of the LABs are connected to the routing fabric through programmable switches. When programmed, these switches determine the exact input and output connections of the LABs. In addition, 10s–100s of programmable I/O pads are available in the FPGA. In many occasions, FPGAs also host heterogeneous dedicated computing resources, such as digital signal processors to implement multiplications, memory blocks to store runtime data, and even full light-weight processor cores.

Each LAB is composed of several *basic logic elements (BLEs)*, where a BLE is made up of a 4-, 5-, or 6-input look-up table (LUT) together with an associated flip-flop. A 4-input LUT can be used to implement any 4-input Boolean function. Figure 1.4(a) illustrates the structure

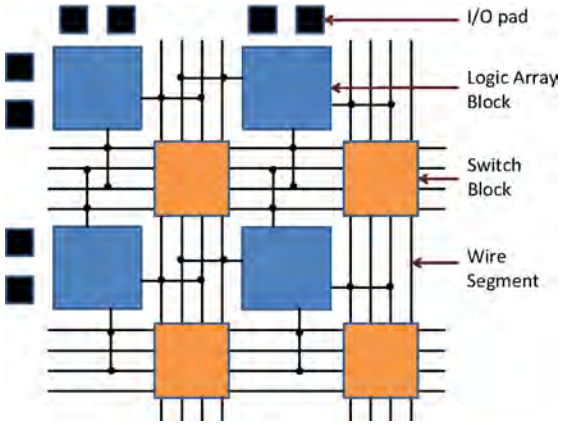

Fig. 1.3. Island-style FPGA.

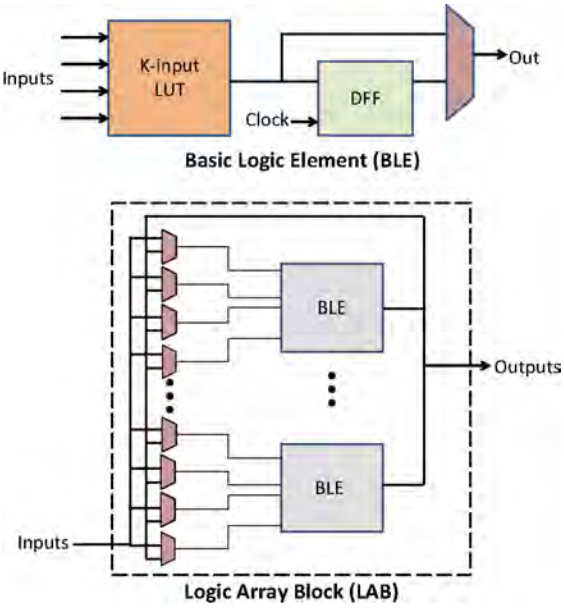

Fig. 1.4. Typical design of a Logic Array Block (LAB) and a basic logic element (BLE).

of a BLE and, Figure 1.4(b) illustrates the structure of a LAB. Each BLE can receive its inputs from other BLEs inside its LAB or from other LABs through the reconfigurable wiring fabric. Additional wiring structures in the LAB enable it to propagate arithmetic carry outputs

in a fast and efficient way. To implement a computing circuit into an FPGA, it is first necessary to synthesize the input circuit by breaking it up into subcircuits, where each subcircuit is mapped to a BLE. These BLEs are then clustered into groups, where the size of each group is determined by the number of BLEs in a LAB. These clusters are then mapped and placed at the LABs. Finally, routing is conducted to determine the exact routes and switches of the routing fabric used by the circuit. The configuration bits for the logic and routing are stored in SRAM or FLASH memory cells.

While FPGAs are very attractive to computer-system designers due to their post-silicon flexibility, this flexibility comes at the expense of higher design area and power consumption compared to custom circuits that perform the same computing tasks. For example, Kuon and Rose report almost a 35× overhead for using programmable logic over custom logic [68]. However, for low to mid-volume fabrication, programmable logic is the only economically feasible technology. Along with performance and area, power is also an important factor that must be considered during architectural design exploration of FPGAs. FPGA architectural parameters include segment length, switch block topology, cluster size, BLE/LAB designs. Choices for these parameters lead to different power, performance, and area trade-offs. Thus, proper evaluation of power consumption is required to help designers and users make correct choices for the FPGA's architecture and programmed designs.

1.2 Survey Overview

The basic techniques for circuit-level power modeling are discussed in Section 2. The power consumption of computing circuits can be des cribed by two components: dynamic power and static power. The section includes discussions on how to estimate each of these components when the design's circuit is available. We will also discuss the various factors that impact these power components, which include, circuit design and layout, input patterns, fabrication technology, process variability, and operational temperature. The discussions in Section 2 will form the basis for the techniques discussed in Sections 3 and 4.

In Section 3 we discuss techniques for *pre-silicon power modeling techniques*. Historically, performance and area were the two main criteria during the design of computing devices. In the past 10–15 years, power has emerged as a third criterion that has to be considered during design. Every architectural feature has to be judged in terms of its performance, area and power. A typical design space has an exponential number of possible combination of settings for the various features. Thus, there is a strong need for power modeling methods that enable designers to efficiently explore the design space and to evaluate the impact of various high-level system architectural choices and optimizations on power consumption. These architectural features and choices vary by the medium of the computing substrate. For multi-core processors, the choices include, for example, pipeline depth, instruction issue width, and cache sizes. For SoC-based embedded systems, the choices include, the functionality of the custom blocks and the on-chip communication architecture (e.g., network topology, buffer sizes and transfer modes). In some embedded systems, the boundary between hardware (HW) and software (SW) is fluid, where the choice of the implementation (SW or HW) of every component could be decided based on its impact on performance, power, and area. In embedded design environments, it is necessary to have power co-modeling tools that can effectively explore the possible HW/SW implementation choices of every design component, and guide designers to the correct choice. FPGA power modeling is also challenging as the user's design is not known during the design and fabrication of the FPGA. Furthermore, users do not have direct access to the internal circuits of the FPGA. Thus, precharacterized power models for the different FPGA structures must be estimated during the design of the FPGA and then bundled with the vendor's tools to be used by the end user.

Once a design is implemented and a physical prototype is available for direct measurements, new opportunities become possible. In Section 4, we discuss a number of techniques for *post-silicon power characterization*. We describe techniques that integrate infrared imaging and direct electric current measurements to develop power mapping techniques, that reveal the true power consumption of every design structure. These true power maps can be used to validate pre-silicon

design estimates, to calibrate power-modeling CAD tools, and to estimate the impact of variabilities introduced during fabrication. We also discuss power characterization techniques for adaptive power-aware computing, where power models based on lumped power measurements are used by power management systems to cut down operational margins and to enforce runtime power constraints. Another discussed topic is SW power characterization using instruction-level, architectural-level and algorithmic-level power models. SW power characterization helps software developers and compiler designers to cut down the power consumption of their applications.

1.3 Summary

In this section we have highlighted the importance of power modeling and characterization techniques for modern computing devices. Future computing systems will be constrained by power, and the choices for design features and runtime settings have to be guided by the impact on power consumption as well as traditional objectives such as performance and implementation area.

Computing substrates can come in a number of forms, including custom circuits with fixed functionality, general-purpose processors whose functionality is determined by software applications, SoCs that combine general-purpose processing cores with application specific custom circuits, and programmable logic that can be used to implement computing circuits in a cost-effective way. These computing forms share some basic power modeling techniques; however, their unique architectural features enable them to utilize efficient large-scale modeling and characterization methods.

Pre-silicon power modeling and post-silicon characterization techniques will be discussed in the remaining sections of this survey. The basic circuit-level power modeling techniques are discussed in Section 2. High-level power modeling techniques for various computing substrates will be discussed in Section 3. In Section 4 we overview different techniques for post-silicon power characterization through physical measurements on a fabricated device. Finally, a number of future research directions are outlined in Section 5.

References

[1] E. Acar, A. Devgan, R. Rao, Y. Liu, H. Su, S. Nassif, and J. Burns, "Leakage and leakage sensitivity computation for combinational circuits," in *International Symposium on Low-Power Electronics*, pp. 96–99, 2003.

[2] A. Agarwal, S. Mukhopadhyay, A. Raychowdhury, K. Roy, and C. H. Kim, "Leakage power analysis and reduction for nanoscale circuits," *IEEE Micro*, vol. 26, no. 2, pp. 68–80, 2006.

[3] Y. Alkabani, F. Koushanfar, N. Kiyavash, and M. Potkonjak, "Trusted integrated circuits: A nondestructive hidden characteristics extraction approach," in *Information Hiding*, pp. 102–117, 2008.

[4] K. Asanovic, R. Bodik, J. Demmel, T. Keaveny, K. Keutzer, J. Kubiatowicz, N. Morgan, D. Patterson, K. Sen, J. Wawrzynek, D. Wessel, and K. Yelick, "A view of the parallel computing landscape," *Communications of the ACM*, vol. 52, no. 10, 2009.

[5] N. Bansal, K. Lahiri, A. Raghunathan, and S. T. Chakradhar, "Power monitors: A framework for system-level power estimation using heterogeneous power models," in *VLSI Design*, pp. 579–585, 2005.

[6] L. A. Barroso and U. Holzle, *The Datacenter as a Computer*. Morgan and Claypool Publishers, 2009.

[7] L. Benini, D. Bruni, M. Chinosi, C. Silvano, V. Zaccaria, and R. Zafalon, "A power modeling and estimation framework for VLIW-based embedded systems," *ST Journal of System Research*, pp. 2.1.1–2.1.10, 2002.

[8] R. A. Bergamaschi1 et al, "SEAS: A system for early analysis of SoCs," in *CODES+ISSS*, pp. 150–155, 2003.

[9] R. Bertran, M. Gonzalez, X. Martorel, N. Navarro, and E. Ayguade, "Decomposable and responsive power models for multicore processors using performance counters," in *International Conference on Supercomputing*, pp. 147–158, 2010.

[10] C. Bienia and K. Li, "PARSEC 2.0: A new benchmark suite for Chipmultiprocessors," in *Proceedings of the Annual Workshop on Modeling, Benchmarking and Simulation*, 2009.

[11] W. L. Bircher, M. Valluri, J. Law, and L. K. John, "Runtime identification of microprocessor energy saving opportunities," in *International Symposium on Low Power Electronics and Design*, pp. 275–280, 2005.

[12] A. Boliolo and L. Benini, "Robust RTL power macromodels," *IEEE Trans on VLSI Systems*, vol. 6, no. 4, pp. 578–581, 1998.

[13] A. Bona, V. Zaccaria, and R. Zafalon, "System level power modeling and simulation of high-end industrial network-on-chip," in *Design, Automation and Test in Europe*, pp. 318–323, 2004.

[14] D. Boning and S. Nassif, "Models of process variations in device and interconnect," in *Design of high-performance microporcessor circuits*, (A. Chandrakasan, W. J. Bowhill, and F. Cox, eds.), pp. 98–115, IEEE Press, 2001.

[15] S. Borkar, "Thousand core chips — a technology perspective," in *Proceedings of Design Automation Conference*, pp. 746–749, 2007.

[16] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *Proceedings of Design Automation Conference*, pp. 338–342, 2003.

[17] O. Breitenstein, W. Warta, and M. Langenkamp, *Lock-In Thermography: Basics and Use for Functional Diagnostics of Electronic Components*. Springer Verlag, second ed., 2010.

[18] D. Brooks, M. Martonosi, J.-D. Wellman, and P. Bose, "Power-performance modeling and tradeoff analysis for a high end microprocessor," in *Power Aware Computing Systems Workshop at ASPLOS-IX*, pp. 126–136, 2000.

[19] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A framework for architectural-level power analysis and optimizations," in *Proceedings of International Symposium on Computer Architecture*, pp. 83–94, 2000.

[20] D. C. Burger and T. M. Austin, "The simplescalar tool set, version 2.0," *ACM Computer Architecture News*, vol. 25, no. 3, pp. 13–25, 1997.

[21] K. M. Buyuksahin and F. N. Najm, "Early power estimation for VLSI circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 7, pp. 1076–1088, 2005.

[22] H. Chang and S. S. Sapatnekar, "Full-chip analysis of leakage power under process variations, including spatial correlations," in *Design Automation Conference*, pp. 523–528, 2005.

[23] N. Chang, K. Kim, and H. G. Lee, "Cycle-accurate energy measurement and characterization with a case study of the ARM7TDMI," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 10, no. 2, pp. 146–154, 2002.

[24] D. Chen, J. Cong, Y. Fan, and L. Wan, "LOPASS: A low-power architectural synthesis system for FPGAs with interconnect estimation and optimization," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 18, no. 4, pp. 564–577, 2010.

[25] D. Chen, J. Cong, Y. Fan, and Z. Zhang, "High-level power estimation and low-power design space exploration for FPGAs," in *Proceedings of Design Automation Conference*, pp. 529–534, 2007.

[26] R. Y. Chen, M. J. Irwin, and R. S. Bajwa, "Architecture-level power estimation and design experiments," *ACM Transactions on Design Automation of Electronic Systems*, vol. 6, no. 1, pp. 50–66, 2001.

[27] X. Chen and L.-S. Peh, "Leakage power modeling and optimization in interconnection networks," in *International Symposium on Low-Power Electronics*, pp. 90–95, 2003.

[28] R. Cochran, A. Nowroz, and S. Reda, "Post-silicon power characterization using thermal infrared emissions," in *International Symposium on Low Power Electronics and Design*, pp. 331–336, 2010.

[29] J. Cong, V. Sarkar, G. Reinman, and A. Bui, "Customizable domain-specific computing," *IEEE Design & Test of Computers*, vol. 28, no. 2, pp. 6–14, 2011.

[30] D. Crisu, S. D. Cotofana, S. Vassiliadis, and P. Liuha, "High-level energy estimation for arm-based socs," in *Systems, Architectures, Modeling, and Simulation*, pp. 168–177, 2004.

[31] J. A. Darringer et al, "Early analysis tools for system-on-a-chip design," *IBM Journal of Research and Development*, vol. 46, no. 6, pp. 691–707, 2002.

[32] J. A. Davis, V. K. De, and J. D. Meindl, "A stochastic wire-length distribution for gigascale integration (GSI) — Part I. Derivation and validation," *IEEE Transactions on Electron Devices*, vol. 45, no. 3, pp. 580–589, 1998.

[33] R. P. Dick, G. Lakshminarayana, A. Raghunathan, and N. K. Jha, "Power analysis of embedded operating systems," in *Design Automation Conference*, pp. 312–315, 2000.

[34] M. Q. Do, M. Drazdziulis, P. Larsson-Edefors, and L. Bengtsson, "Parameterizable architecture-level SRAM power model using circuit-simulation backend for leakage calibration," *International Symposium on Quality Electronic Design*, pp. 557–563, 2006.

[35] D. Economou, S. Rivoire, C. Kozyrakis, and P. Ranganathan, "Full-system power analysis and modeling for server environments," in *In Proceedings of Workshop on Modeling, Benchmarking, and Simulation*, pp. 70–77, 2006.

[36] J. Emer, P. Ahuja, E. Borch, A. Klauser, C.-K. Luk, S. Manne, S. S. Mukherjee, H. Patil, S. Wallace, N. Binkert, R. Espasa, and T. Juan, "Asim: A performance model framework," *Computer*, vol. 35, pp. 68–76, 2002.

[37] X. Fan, W. Weber, and L. Barroso, "Power provisioning for a warehouse-sized computer," *International Symposium on Computer Architecture*, pp. 13–23, 2007.

[38] P. Friedberg, Y. Cao, J. Cain, R. Wang, J. Rabaey, and C. Spanos, "Modeling within-die spatial correlation effects for process-design co-optimization," in *International Symposium on Quality Electronic Design Automation*, pp. 516–521, 2005.

[39] S. H. Fuller and L. I. Millett, *The Future of Computing Performance: Game Over or Next Level?* The National Academies Press, 2011.

[40] T. Givargis, F. Vahid, and J. Henkel, "A hybrid approach for core-based system-level power modeling," in *Asia and South Pacific Design Automation Conference*, pp. 141–145, 2000.

[41] T. Givargis, F. Vahid, and J. Henkel, "Instruction-based system-level power evaluation of system-on-a-chip peripheral cores," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 10, no. 6, pp. 856–862, 2002.

[42] W. Gosti, A. Narayan, R. K. Brayton, and A. L. Sangiovanni-Vincentelli, "Wireplanning in Logic Synthesis," in *Proceedings of International Conference on Computer-Aided Design*, pp. 26–33, 1998.

[43] P. Gupta, A. B. Kahng, and S. Muddu, "Quantifying error in dynamic power estimation of CMOS circuits," *Analog Integrated Circuits and Signal Processing*, vol. 42, no. 3, pp. 253–264, 2005.

[44] S. Gupta and F. N. Najm, "Power modeling for high-level power estimation," *Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 1, pp. 18–29, 2000.

[45] S. Gurumurthi, A. Sivasubramaniam, M. J. Irwin, N. Vijaykrishnan, and M. Kandemir, "Using complete machine simulation for software power estimation: The softwatt approach," in *International Conference on High-Performance Computer Architecture*, pp. 141–150, 2002.

[46] H. Hamann, A. Weger, J. Lacey, Z. Hu, and P. Bose, "Hotspot-limited microprocessors: Direct temperature and power distribution measurements," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 1, pp. 56–65, 2007.

[47] B. Hargreaves, H. Hult, and S. Reda, "Within-die process variations: How accurately can they be statistically modeled?," in *Proceedings of Asia and South Pacific Design Automation Conference*, pp. 524–530, 2008.

[48] K. R. Heloue, N. Azizi, and F. N. Najm, "Full-chip model for leakage current estimation considering within-die correlation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, no. 6, pp. 847–887, 2009.

[49] S. Hong and H. Kim, "An Integrated GPU Power and Performance Model," in *International Symposium on Computer Architecture*, pp. 280–289, 2010.

[50] J. Howard et al, "A 48-Core IA-32 Processor in 45 nm CMOS using on-die message-passing and DVFS for performance and power scaling," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 173–183, 2011.

[51] C. Hsieh, C. Wu, F. Jih, and T. Sun, "Focal-plane-arrays and CMOS readout techniques of infrared imaging systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 4, pp. 594–605, 1997.

[52] C.-T. Hsieh, M. Pedram, G. Mehta, and F. Rastgar, "Profile-driven program synthesis for evaluation of system power dissipation," in *Design Automation Conference*, pp. 576–581, 1997.

[53] C.-W. Hsu et al, "PowerDepot: Integrating IP-based power modeling with ESL power analysis for multi-core SoC designs," in *Design Automation Conference*, pp. 47–52, 2011.

[54] S. Huang, A. Hormati, D. Bacon, and R. Rabbah, "Liquid metal: Object-oriented programming across the hardware/software boundary," in *n Proceedings of the European Conference on Object-Oriented Programming*, pp. 76–103, 2008.

[55] W. Huang, K. Skadron, S. Gurumurthi, R. J. Ribando, and M. R. Stan, "Differentiating the roles of IR measurement and simulation for power and temperature-aware design," in *International Symposium on Performance Analysis of Systems and Software*, pp. 1–10, 2009.

[56] C. Isci, G. Contreras, and M. Martonosi, "Live, runtime phase monitoring and prediction on real systems with application to dynamic power management," in *International Symposium on Microarchitecture*, pp. 359–370, 2006.

[57] C. Isci and M. Martonosi, "Runtime power monitoring in high-end processors: Methodology and empirical data," in *Proceedings of International Symposium on MicroArchitecture*, pp. 93–104, 2003.

[58] V. Jimenez et al, "Power and thermal characterization of POWER6 system," in *International Conference on Parallel Architectures and Compilation Techniques*, pp. 7–18, 2010.

[59] R. Joseph and M. Martonosi, "Run-time power estimation in high performance microprocessors," in *International Symposium on Low Power Electronics and Design*, pp. 135–140, 2001.

[60] N. Julien, J. Laurent, E. Senn, and E. Martin, "Power consumption modeling and characterization of the TI C6201," *IEEE Micro*, vol. 23, no. 5, pp. 40–49, 2004.

[61] A. B. Kahng, B. Li, L.-S. Peh, and K. Samadi, "ORION 2.0: A fast and accurate NoC power and area model for early-stage design space exploration," in *Design, Automation and Test in Europe*, pp. 423–428, 2009.

[62] A. B. Kahng and S. Reda, "A tale of two nets: Studies of wirelength progression in physical design," in *System-Level Interconnect Prediction Workshop*, pp. 17–34, 2006.

[63] T. Kemper, Y. Zhang, Z. Bian, and A. Shakouri, "Ultrafast temperature profile calculation in IC chips," in *International Workshop on Thermal inverstigations of ICs and Systems*, pp. 133–137, 2006.

[64] N. S. Kim, T. Austin, T. Mudge, and D. Grunwald, *Power Aware Computing*. Kluwer Academic Publishers Norwell, 2002.

[65] N. S. Kim, K. Flautner, D. Blaauw, and T. Mudge, "Drowsy instruction caches: Leakage power reduction using dynamic voltage scaling and cache sub-bank prediction," in *International Symposium on Microarchitecture*, pp. 219–230, 2002.

[66] F. Koushanfar and A. Mirhoseini, "A unified framework for multimodal submodular integrated circuits trojan detection," *IEEE Transactions on Information Forensic and Security*, 2011.

[67] A. Kumar and M. Anis, "An analytical state dependent leakage power model for FPGAs," in *Design, Automation and Test in Europe*, 2006.

[68] I. Kuon and J. Rose, "Measuring the gap between FPGAs and ASICs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits*, vol. 26, no. 2, pp. 203–215, 2007.

[69] K. Lahiri and A. Raghunathan, "Power analysis of system-level on-chip communication architectures," in *CODES+ISSS*, pp. 236–241, 2004.

[70] M. Lajolo, A. Raghunathan, S. Dey, and L. Lavagno, "Cosimulation-based power estimation for system-on-chip design," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 10, no. 3, pp. 253–265, 2002.

[71] J. Lamoureux and S. Wilton, "FPGA clock network architecture: Flexibility vs. area and power," in *International Symposium on Field Programmable Gate Arrays*, pp. 101–108, 2006.

[72] J. Lamoureux and S. Wilton, "On the trade-off between power and flexibility of FPGA clock networks," *ACM Transactions on Reconfigurable Technology and Systems*, vol. 1, no. 13, 2008.

[73] P. Landman, "High-level power estimation," in *International Symposium on Low-Power Electronics and Design*, pp. 29–35, 1996.

[74] B. Lee and D. Brooks, "Accurate and efficient regression modeling for microarchitectural performance and power prediction," in *Architectural Support for Programming Languages and Operating Systems*, pp. 185–194, 2006.

[75] I. Lee et al, "PowerViP: SoC power estimation framework at transaction level," in *Proceedings of Asia and South Pacific Design Automation Conference*, pp. 551–558, 2006.

[76] V. Lee et al, "Debunking the 100X GPU vs. CPU Myth: An evaluation of throughput computing on CPU and GPU," *International Symposium on Computer Architecture*, pp. 451–460, 2010.

[77] F. Li, D. Chen, L. He, and J. Cong, "Architecture evaluation for power-efficient FPGAs," in *International Symposium on Field Programmable Gate Arrays*, pp. 175–184, 2003.

[78] F. Li, D. Chen, L. He, and J. Cong, "Power modeling and characteristics of field programmable gate arrays," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 11, pp. 1712–1724, 2005.

[79] T. Li and L. K. John, "Run-time modeling and estimation of operating system power consumption," in *SIGMETRICS*, pp. 160–171, 2003.

[80] Y. Li and J. Henkel, "A framework for estimating and minimizing energy dissipation of embedded HW/SW systems," in *Design Automation Conference*, pp. 188–193, 1998.

[81] X. Liang, K. Turgay, and D. Brooks, "Architectural power models for SRAM and CAM structures based on hybrid analytical/empirical techniques," *International Conference on Computer Aided Design*, pp. 824–830, 2007.

[82] M. Y. Lim, A. Porterfield, and R. Fowler, "SoftPower: Fine-grain power estimations using performance counters," in *International Symposium on High Performance Distributed Computing*, pp. 308–311, 2010.

[83] M. Loghi, L. Benini, and M. Poncino, "Power macromodeling of MPSoC message passing primitives," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 31, pp. 1–22, 2007.

[84] M. Loghi, M. Poncino, and L. Benini, "Cycle-accurate power analysis for multiprocessor systems-on-a-chip," in *Great Lakes Symposium on VLSI*, pp. 401–406, 2004.

[85] X. Ma, M. Dong, L. Zhong, and Z. Deng, "Statistical power consumption analysis and modeling for GPU-based computing," in *Workshop on Power Aware Computing and Systems (HotPower)*, 2009.

[86] E. Macii, M. Pedram, and F. Somenzi, "High-level power modeling, estimation, and optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 11, pp. 1061–1079, 1998.

[87] M. Mamidipaka and N. Dutt, "eCACTI: An enhanced power estimation model for on-chip caches," Technical report, In Technical Report TR-04-28, CECS, UCI, 2004.

[88] R. Marculescu, D. Marculescu, and M. Pedram, "Probabilistic modeling of dependencies during switching activity analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 2, pp. 73–83, 1998.

[89] E. Marin, "The role of thermal properties in periodic time-varying phenomena," *European Journal of Physics*, vol. 28, no. 3, pp. 429–445, 2007.

[90] H. Mehta, R. M. Owens, and M. J. Irwin, "Energy characterization based on clustering," in *Design Automation Conference*, pp. 702–707, 1996.

[91] F. J. Mesa-Martinez, E. Ardestani, and J. Renau, "Characterizing processor thermal behavior," in *Architectural Support for Programming Languages and Operating Systems*, pp. 193–204, 2010.

[92] F. J. Mesa-Martinez, M. Brown, J. Nayfach-Battilana, and J. Renau, "Measuring performance, power, and temperature from real processors," in *Proceedings of International Symposium on Computer Architecture*, pp. 1–10, 2007.

[93] J. Monteiro, S. Devadas, A. Ghosh, K. Keutzer, and J. White, "Estimation of average switching activity in combinational logic circuits using symbolic simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, no. 1, 1997.

[94] J. Monteiro, R. Patel, and V. Tiwari, "Power analysis and optimization from circuit to register-transfer levels," in *EDA for IC Implementation, Circuit Design, and Process Technology*, vol. 2, (L. Scheffer, L. Lavagno, and G. Martin, eds.), Taylor & Francis, 2006.

[95] M. Moudgill, J.-D. Wellman, and J. H. Moreno, "Environment for PowerPC microarchitecture exploration," *Micro*, vol. 19, pp. 15–25, 1999.

[96] A. Muttreja, A. Raghunathan, S. Ravi, and N. K. Jha, "Automated energy/performance macromodeling of embedded software," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 3, pp. 542–552, 2007.

[97] H. Nagasaka, N. Maruyama, A. Nukada, T. Endo, and S. Matsuoka, "Statistical power modeling of gpu kernels using performance counters," in *Proceedings of the International Conference on Green Computing*, pp. 115–122, 2010.

[98] F. Najm, "Transition density: A new measure of activity in digital circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, no. 2, pp. 310–323, 1993.

[99] F. Najm, "A survey of power estimation techniques in VLSI circuits," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 2, no. 4, pp. 446–455, 1994.

[100] F. Najm, "Power estimation techniques for integrated circuits," in *International Conference on Computer Aided Design*, pp. 492–499, 1995.

[101] M. Nemani and F. N. Najm, "Towards a high-level power estimation capability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 6, pp. 588–598, 1996.

[102] A. N. Nowroz, R. Cochran, and S. Reda, "Thermal monitoring of real processors: Techniques for sensor allocation and full characterization," in *Design Automation Conference*, pp. 56–61, 2010.

[103] A. N. Nowroz and S. Reda, "Thermal and power characterization of field-programmable gate arrays," in *International Symposium on Field Programmable Gate Arrays*, pp. 111–114, 2011.

[104] A. N. Nowroz, G. Woods, and S. Reda, "Improved post-silicon power modeling using AC lock-in techniques," in *Design Automation Conference*, pp. 101–106, 2011.

[105] M. Onouchi, T. Yamada, K. Morikawa, I. Mochizuki, and H. Sekine, "A system-level power estimation methodology based on IP-level modeling, power-level adjustment, and power accumulation," in *Proceedings of Asia and South Pacific Design Automation Conference*, pp. 547–550, 2006.

[106] M. Orshansky and S. Nassif, *Design or Manufacturability and Statistical Design: A Constructive Approach*. Springer, 2007.

[107] T. Osmulski et al, "A probabilistic power prediction tool for the Xilinx 4000-series FPGA," in *Lecture Notes in Computer Science*, pp. 776–783, 2000.

[108] J. Ou and V. K. Prasanna, "Rapid energy estimation of computations on FPGA based soft processors," in *SOC Conference*, pp. 285–288, 2004.

[109] S. Pasricha and N. Dutt, *On-Chip Communication Architectures: System on Chip Interconnect*. Morgan Kaufmann Publishers, 2008.

[110] S. Pasricha, Y.-H. Park, N. Dutt, and F. J. Jurdahi, "System-level PVT variation-aware power exploration of on-chip communication architectures," *ACM Transactions on Design Automation of Electronic Systems*, vol. 14, no. 2, pp. 20:1–20:25, 2009.

[111] J. Peddersen and S. Parameswaran, "CLIPPER: Counter-based low impact processor power estimation at run-time," in *Asia and South Pacific Design Automation Conference*, pp. 890–895, 2007.

[112] M. Pedram and S. Nazarin, "Thermal modeling, analysis, and management in VLSI circuits: Principles and methods," *Proceedings of the IEEE*, vol. 94, no. 8, pp. 1487–1501, 2006.

[113] L. Peh and N. Jerger, *On-Chip Networks*. Morgan and Claypool Publishers, 2009.

[114] D. Ponomarev, G. Kucuk, and K. Ghose, "AccuPower: An accurate power estimation tool for superscalar microprocessors," in *Design, Automation and Test in Europe Conference and Exhibition*, pp. 124–129, 2002.

[115] K. Poon, S. Wilton, and A. Yan, "A detailed power model for field-programmable gate arrays," *ACM Transactions on Design Automation of Electronic Systems*, vol. 10, no. 2, pp. 279–302, 2005.

[116] K. K. W. Poon, "Power estimation for field programmable gate arrays," PhD thesis, The University of British Columbia, August 2002.

[117] E. Pop, S. Sinha, and K. Goodson, "Heat generation and transport in nanometer-scale transistors," *Proceedings of the IEEE*, vol. 94, no. 8, pp. 1587–1601, 2006.

[118] M. Powell, A. Biswas, J. Emer, and S. Mukherjee, "CAMP: A technique to estimate per-structure power at run-time using a few simple parameters," *International Symposium on High Performance Computer Architecture*, pp. 289–300, 2009.

[119] G. Qu, N. Kawabe, K. Usami, and M. Potkonjak, "Function-level power estimation methodology for microprocessors," in *Design Automation Conference*, pp. 810–813, 2000.

[120] A. Ragunathan, S. Dey, and N. K. Jha, "Register-transfer level estimation techniques for switching activity and power consumption," in *International Conference on Computer-Aided Design*, pp. 158–165, 1996.

[121] R. Rao, A. Srivastava, D. Blaauw, and D. Sylvester, "Statistical estimation of leakage current considering inter- and intra-die process variation," in *International Symposium on Low-Power Electronics and Design*, pp. 84–89, 2003.

[122] S. Reda, "Thermal and power characterization of real computing devices," *IEEE Journal on Emerging Topics in Circuits and Systems*, vol. 1, no. 2, 2011.

[123] S. Reda and S. Nassif, "Analyzing the impact of process variations on parametric measurements: Novel models and applications," in *Design, Automation and Test in Europe Automation*, pp. 375–380, 2009.

[124] S. Rivoire, P. Ranganathan, and C. Kozyrakis, "A comparison of high-level full-system power models," in *Conference on Power Aware Computing and Systems*, pp. 1–4, 2008.

[125] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 305–327, 2003.

[126] J. T. Russell and M. F. Jacome, "Software power estimation and optimization for high performance, 32-bit embedded processors," in *International Conference on Computer Design*, pp. 328–333, 1998.

[127] L. Shang, A. S. Kaviani, and K. Bathala, "Dynamic power consumption in virtex^{TM}-II FPGA family," in *International Symposium on Field Programmable Gate Arrays*, pp. 157–164, 2002.

[128] P. Shivakumar, N. P. Jouppi, and P. Shivakumar, "CACTI 3.0: An integrated cache timing, power, and area model," Technical report, 2001.

[129] T. Simunic, L. Benini, and G. De Micheli, "Cycle-accurate simulation of energy consumption in embedded systems," in *Design Automation Conference*, pp. 867–872, 1999.

[130] G. Sinevriotis et al, "SOFLOPO: Towards systematic software exploitation for low-power designs," in *International Symposium on Low-Power Electronics and Design*, 2000.

[131] K. Singh, M. Bhadauria, and S. A. McKee, "Real time power estimation and thread scheduling via performance counters," in *Proceedings of the Workshop on Design, Architecture, and Simulation of Chip Multi-Processors*, pp. 46–55, 2008.

[132] A. Sinha and A. P. Chandrakasan, "JouleTrack — A Web based tool for software energy profiling," in *Design Automation Conference*, pp. 220–225, 2001.

[133] G. Spirakis, "Designing for 65 nm and beyond: Where is the revolution?," in *Electronic Design Process (EDP) Symposium*, 2005.

[134] S. Steinke, M. Knauer, L. Wehmeyer, and P. Marwedel, "An accurate and fine grain instruction-level energy model supporting software optimizations," in *Proceedings of the International Workshop Power and Timing Modeling, Optimization and Simulation*, 2001.

[135] D. Stroobandt, *A Priori Wire Length Estimates for Digital Design*. Kluwer Academic Publishers, 2001.

[136] H. Su, F. Liu, A. Devgan, E. Acar, and S. Nassif, "Full chip leakage estimation considering power supply and temperature variations," in *International Symposium on Low Power Electronics and Design*, pp. 78–83, 2003.

[137] C. Talarico, J. W. Rozenblit, V. Malhotra, and A. Stritter, "A new framework for power estimation of embedded systems," *IEEE Computer*, vol. 38, no. 2, pp. 71–78, 2005.

[138] T. K. Tan, A. Raghunathan, G. Lakshminarayana, and N. K. Jha, "High-level software energy macro-modeling," in *Design Automation Conference*, pp. 605–610, 2001.

[139] V. Tiwari, S. Malik, and A. Wolfe, "Power analysis of embedded software: A first step towards software power minimization," *IEEE Transactions on Very Large Scale Integration Systems*, vol. 2, no. 4, pp. 437–445, 1994.

[140] T. Tuan and B. Lai, "Leakage power analysis of a 90 nm FPGA," in *Custom Integrated Circuits Conference*, pp. 57–60, 2003.

[141] S. Vangal et al, "An 80-Tile Sub-100-W Teraflops Processor in 65-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 1, pp. 29–41, 2008.

[142] A. Varma, E. Debes, I. Kozintsev, P. Klein, and B. Jacob, "Accurate and fast system-level power modeling: An XScale-based case study," *ACM Transactions on Embedded Computing Systems*, vol. 7, no. 3, pp. 25:1–25:20, 2008.

[143] H.-S. Wang, X. Zhu, L.-S. Peh, and S. Malik, "Orion: A power-performance simulator for interconnection networks," in *International Symposium on Microarchitecture*, pp. 294–305, 2002.

[144] Z. Wang, N. Tolia, and C. Bash, "Opportunities and challenges to unify workload, power, and cooling management in data centers," *ACM SIGOPS Operating Systems Review*, vol. 44, no. 3, pp. 41–46, 2010.

[145] S. J. E. Wilton and N. P. Jouppi, "CACTI: An enhanced cache access and cycle time model," *IEEE Journal Solid-State Circuits*, vol. 31, no. 5, pp. 677–688, 1996.

[146] B. Wong, A. Mittal, Y. Cao, and G. W. Starr, *Nano-CMOS Circuit and Physical Design*. Wiley-Interscience, 2004.

[147] W. Wu, L. Jin, J. Yang, P. Liu, and S. X.-D. Tan, "A systematic method for functional unit power estimation in microprocessors," in *Design Automation Conference*, pp. 554–557, 2006.

[148] W. Ye, N. Vijaykrishnan, M. Kandemir, and M. J. Irwin, "The design and use of simplepower: A cycle-accurate energy estimation tool," in *Design Automation Conference*, pp. 340–345, 2000.

[149] Y. Zhang, D. Parikh, K. Sankaranarayanan, K. Skadron, and M. Stan, "HotLeakage: A temperature-aware model of subthreshold and gate leakage for architects," Technical report, 2003.