

Research Summary

David Brooks
School of Engineering and Applied Sciences
Harvard University

November 18, 2008

Advances in computational capabilities have driven the information technology revolution, which in turn has driven advances in nearly all fields of science, medicine, and business. At the same time, computing devices have become ubiquitous, and today's cell phones pack an enormous amount of computational power into small form-factor devices. These advances have largely been driven by Moore's Law, or the doubling of the number of transistors available for a given cost every two years. Advanced semiconductor technologies provide a wealth of transistor resources that computer architects exploit to provide improved performance and other advanced computing capabilities. However, as manufacturing process technologies shrink feature sizes to the nanoscale, this growth is threatened by severe technology-related scaling challenges. These scaling challenges include minimizing energy consumption, mitigating process variations, and optimizing power delivery and voltage control [1]. Addressing these scaling challenges through both hardware and software design mechanisms, and hence striving to maintain growth in computational capabilities, is the primary activity of my research group.

Although incredibly powerful computing devices are available today, the single-minded pursuit of performance has led to power consumption emerging as one of the main bottlenecks for nearly all types of computing systems. Power dissipation is a limiting factor in battery-operated portable devices. As society evolves to a highly connected world of cellular phones and portable computers, reducing power consumption in these devices is of the utmost importance and therefore the center of active research and development. However, issues related to power dissipation are not restricted to battery-powered systems. Energy has been identified as a first-class design constraint for systems ranging from embedded microcontrollers in tiny sensor devices to high-performance servers powering Internet data centers. In recent years thermal limits and power delivery have become key constraints in the design of high-performance microprocessors. The thermal and power delivery constraints heavily depend on the temporal and spatial characteristics of power dissipation within the microprocessor, and hence are tightly linked to the computer architecture and application workloads. Another key technology challenge that designers must face is uncertainty in transistor device characteristics due to manufacturing process variations. This challenge threatens to derail Moore's Law scaling by severely disrupting memory device scaling, increasing standby power dissipation, and causing stagnation in chip-level performance growth.

My research focuses on each of these design issues across a range of computing application domains spanning low-power embedded and high-performance computing systems. This research is connected by several common themes. First, understanding the challenges and opportunity of underlying trends in computer implementation technology is critical to future computer system design. Understanding these future trends allows me to choose research problems that become increasingly significant as these challenges emerge. A basic tenet of my research is that many of these low-level, technology related problems are more easily solved when one considers the impact of the problem at the system level. A problem that may seem

incredibly challenging when considered at the scale of a single transistor, may become significantly easier when one understands how many such transistors are connected into a system and the needs of software as it runs on these systems. Another theme of my research is that in order to make headway on the design of novel solutions to these low-level problems, we must develop understanding of the system-level impact of these challenges. Examples include constructing models to understand the ramifications of architectural decisions on power dissipation and developing models to analyze how transistor device variability impacts systems as a unit. Finally, the temporal and spatial characteristics of these technology challenges indicate that we must consider solutions that span multiple disciplines in order to find approaches that operate at the appropriate granularity. My research reflects this with inter-disciplinary work between computer architecture and circuits, VLSI-CAD, compilers, virtual machines, networking, and application design.

My Ph.D. thesis research at Princeton was one of the first efforts to study power dissipation and thermal issues at the chip architecture level [2–4]. Part of my thesis research involved development of the Wattch power-performance architectural simulator [3], which is heavily used by researchers worldwide to study power issues. Spending a year at IBM Research after completing my Ph.D. allowed me to apply these lessons to the early-stage design of IBM’s Cell processor and IBM’s POWER6 processor. During my time at IBM I also developed the PowerTimer research infrastructure, which allows accurate power modeling of high-performance PowerPC microprocessors [5, 6]. This infrastructure allowed us to perform a detailed power analysis of one of the most fundamental parameters in microprocessor design – the number of pipeline stages in a microprocessor [7, 8]. This research influenced the computing industry by highlighting the critical importance of architectural design on power-performance efficiency.

After I brought the PowerTimer infrastructure to Harvard, we extended the modeling toolkit to analyze the impact of multi-threading and multi-core design on power efficiency and thermal hotspots [9–11]. We also demonstrated that multi-core designs that utilize simple cores provide inherent energy and thermal advantages compared to complex single-core designs [12]. These efforts, as is the case with much microarchitectural analysis, involves an ad-hoc approach that is not scalable to large design spaces. To address this problem, we proposed a new approach to power-performance design space exploration utilizing regression-based statistical inference [13]. This analysis approach has been extremely powerful, allowing detailed studies of Pareto-optimal design spaces and multiprocessor heterogeneity [14]. We have also applied the approach to tuning of application parameters [15], understanding the roughness of microarchitectural design topologies [16], and quantifying the power-performance efficiency of fine-grained architectural reconfigurability [17]. In collaboration with Intel, we demonstrated the technique within the context of an evolutionary product design cycle and demonstrated a composable multi-core simulation approach [18]. Strategies based on statistical inference and other machine-learning approaches are likely to have a significant role to play in many areas of constrained architecture and system design.

Computer designs are increasingly bumping up against the *power wall* – hard caps on the amount of power that can be consumed in a system. This power wall may be set by power-source limitations (e.g., battery) or thermal limitations, and has the potential to limit technology scaling [19]. For the largest power-efficiency gains (100-1000×), we have begun to consider composable specialization. For example, we have explored a domain-specific architecture for wireless sensor nodes that leverages the event-driven, regular processing characteristics of the sensor node domain [20–22]. The proposed architecture has been evaluated through a detailed simulation analysis [23] and through a prototype implementation in 130nm CMOS. We have also explored heterogeneous, domain-specific instruction set architectures for general-purpose microcontrollers [24, 25]. Our recent efforts have looked to expand the composable specialization approach to high-performance systems that are increasingly facing the power wall.

Another challenge of nanoscale design is variation in device characteristics due to difficulties in manufacturing transistors with shrinking feature size. While there have been efforts to address variations at the

circuit and VLSI-CAD levels, our work is among the first to address these challenges at the system-level. Our first effort explored tradeoffs between microarchitectural parameter selection and sensitivity to variations [26]. We then explored techniques to mitigate variations in microprocessor register files and execution units, with concepts including port-switching register files and variable-latency functional units [27]. We also developed a technique called *voltage interpolation* which allows fine-grained spatial voltage control in a microprocessor to alleviate variations. We developed a prototype floating-point unit in 130nm CMOS that demonstrates both the variable-latency and voltage interpolation techniques [28], and we performed a system-level study that explores tradeoffs between the two techniques for a 16-core CMP machine [29, 30]. Our recent efforts explore the optimal selection of voltage cut-points for the interpolation technique when considering tradeoffs in terms of power overheads and tunability [31].

Process variations are especially problematic for memory designs, which rely on the smallest feature-size transistors. Recent results from Intel show that traditional 6T memory cells are not scaling with technology nodes which is alarming because a large fraction of chip area is devoted to memories. We have been exploring the radical approach of replacing 6T static memory cells with dynamic 3T memory cells. Our analysis shows that in 3T memories, all effects of variations can be captured by the simple data retention time parameter, which can be managed with architectural solutions including cache refresh and replacement policies [32, 33]. We are currently working on prototyping this idea in a 90nm test chip, and studying other benefits of 3T memories.

Fluctuations in the delivery of power to a microprocessor result in different types of variations. Power delivery variations result in dips or spikes in the supply voltage which could lead to transient faults. In order to avoid these *voltage emergencies*, designers often impose wide voltage margins to protect against worst-case operating conditions, resulting in large performance penalties or power overheads in the common case. Our initial efforts in this area explored using formal feedback-control mechanisms to guide a voltage sensor and throttling mechanism [34]. One of the observations of this work was that certain sequences of instruction code can lead to these voltage emergencies, and we explored this idea in more depth in subsequent analysis [35–37]. This ultimately led to our development of a simple, highly-accurate voltage emergency predictor, which can be used to proactively avoid voltage transients [38]. We have also explored a variety of architectural and software solutions for these problems, including fine-grained architectural state checkpointing [39] and compiler-inserted instructions that can smooth processor current load [40].

Dynamic runtime systems provide many opportunities for energy savings due to the potential for exploiting slack within program execution by applying techniques such as dynamic voltage/frequency scaling (DVFS). For example, memory-bound loops within workloads provide an opportunity to reduce the frequency and voltage of the CPU with minimal performance impact. We explored the use of a software optimization layer that samples performance counters to determine appropriate DVFS settings, and we applied the technique to an Intel test-board with a Pentium M processor [41, 42]. Conventional DVFS schemes are inherently limited by the speed of voltage transitions, and in the multi-core era, by having a single global voltage knob for all cores in a microprocessor. We are exploring integrated on-chip voltage regulators that can address these challenges. Integrating the regulator onto the chip can drastically reduce voltage response (potentially providing nanosecond-scale voltage switching times), and the regulator can be distributed across multiple cores allowing per-core voltage control [43]. We are now in the process of building a prototype of the on-chip regulator which will allow us to validate our existing simulation-based analysis.

Highlights of my research can be found in the following five research papers [7, 13, 23, 29, 32].

References

- [1] D. Brooks, R. Dick, R. Joseph, and L. Shang, "Power, thermal, and reliability modeling in nanometer-scale microprocessors," *IEEE Micro*, May 2007.
- [2] D. Brooks and M. Martonosi, "Dynamically exploiting narrow width operands to improve processor power and performance," in *International Conference on High-Performance Computer Architecture (HPCA-5)*, Jan. 1999.
- [3] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A framework for architectural-level power analysis and optimizations," in *International Symposium on Computer Architecture (ISCA-27)*, June 2000.
- [4] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *International Conference on High-Performance Computer Architecture (HPCA-7)*, Jan. 2001.
- [5] D. Brooks, P. Bose, V. Srinivasan, M. Gschwind, P. Emma, and M. Rosenfield, "New methodology for early-stage, microarchitecture-level power-performance analysis of microprocessors," *IBM J. Research and Development*, vol. 47, no. 5, 2003.
- [6] X. Liang, K. Turgay, and D. Brooks, "Architectural power models for SRAM and CAM structures based on hybrid analytical/empirical techniques," in *International Conference on Computer-Aided Design (ICCAD)*, Nov. 2007.
- [7] V. Srinivasan, D. Brooks, M. Gschwind, P. Bose, V. Zyuban, P. N. Strenski, and P. G. Emma, "Optimizing pipelines for power and performance," in *International Symposium on Microarchitecture (MICRO-35)*, Nov. 2002. Selected as one of the four Best IBM Research Papers in Computer Science, Electrical Engineering and Math published in 2002.
- [8] V. Zyuban, D. Brooks, V. Srinivasan, M. Gschwind, P. Bose, P. Strenski, and P. Emma, "Integrated analysis of power and performance for pipelined microprocessors," *IEEE Transactions on Computers*, vol. 53, pp. 1004–1016, August 2004.
- [9] Y. Li, D. Brooks, Z. Hu, K. Skadron, and P. Bose, "Understanding the energy efficiency of simultaneous multithreading," in *International Symposium on Low-Power Electronics and Design (ISLPED)*, August 2004.
- [10] Y. Li, M. Hempstead, P. Mauro, D. Brooks, and Z. H. K. Skadron, "Power and thermal effects of SRAM vs. latch-mux design styles and clock gating choices," in *International Symposium on Low-Power Electronics and Design (ISLPED)*, August 2005.
- [11] Y. Li, D. Brooks, Z. Hu, and K. Skadron, "Performance, energy, and thermal considerations for SMT and CMP architectures.," in *International Conference on High-Performance Computer Architecture (HPCA-11)*, Feb. 2005.
- [12] Y. Li, B. Lee, D. Brooks, Z. Hu, and K. Skadron, "CMP design space exploration subject to physical constraints," in *International Conference on High-Performance Computer Architecture (HPCA-12)*, Feb. 2006.
- [13] B. Lee and D. Brooks, "Accurate and efficient regression modeling for microarchitectural performance and power prediction," in *International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, October 2006.

- [14] B. Lee and D. Brooks, "Illustrative design space studies with microarchitectural regression models," in *International Symposium on High-Performance Computer Architecture (HPCA-13)*, Feb. 2007.
- [15] B. Lee, D. Brooks, B. de Supinski, M. Schulz, K. Singh, and S. McKee, "Methods of inference and learning for performance modeling of parallel applications," in *Symposium on Principles and Practice of Parallel Programming (PPoPP)*, March 2007.
- [16] B. Lee and D. Brooks, "Roughness of microarchitectural design topologies and its implications for optimization," in *International Symposium on High-Performance Computer Architecture (HPCA-14)*, Feb. 2008.
- [17] B. Lee and D. Brooks, "Efficiency trends and limits from comprehensive microarchitectural adaptivity," in *International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, March 2008.
- [18] B. Lee, J. Collins, H. Wang, and D. Brooks, "CPR: composable performance regression for scalable multiprocessor models," in *International Symposium on Microarchitecture (MICRO-41)*, December 2008.
- [19] M. Hempstead, G.-Y. Wei, and D. Brooks, "Architecture and circuit techniques for low throughput, energy constrained systems across technology generations," in *International Conference on Compilers, Architecture, and Synthesis for Embedded Systems*, Oct. 2006.
- [20] M. Hempstead, D. Brooks, and M. Welsh, "TinyBench: The case for a standardized benchmark suite for TinyOS based wireless sensor network devices.," in *IEEE Workshop on Embedded Networked Sensors (EmNets'04)*, Nov. 2004.
- [21] M. Hempstead, M. J. Lyons, D. Brooks, and G.-Y. Wei, "Survey of hardware systems for wireless sensor networks," *ASP Journal of Low Power Electronics*, 2008.
- [22] M. Hempstead, G.-Y. Wei, and D. Brooks, "System design considerations for sensor network applications (Invited Paper)," *International Symposium on Circuits and Systems*, 2008.
- [23] M. Hempstead, N. Tripathi, P. Mauro, G.-Y. Wei, and D. Brooks, "An ultra low power system architecture for sensor network applications.," in *International Symposium on Computer Architecture (ISCA-32)*, June 2005.
- [24] L. Strozek and D. Brooks, "Efficient architectures through application clustering and architectural heterogeneity," in *International Conference on Compilers, Architecture, and Synthesis for Embedded Systems*, Oct. 2006.
- [25] L. Strozek and D. Brooks, "Efficient architectures through application clustering and heterogeneity," *ACM Transactions on Architecture and Code Optimization*, Accepted for publication. 2009.
- [26] X. Liang and D. Brooks, "Microarchitecture parameter selection to optimize system performance under process variation," in *International Conference on Computer-Aided Design (ICCAD)*, Nov. 2006.
- [27] X. Liang and D. Brooks, "Mitigating the impact of process variations on cpu register file and execution units," in *International Symposium on Microarchitecture (MICRO-39)*, December 2006.
- [28] X. Liang, D. Brooks, and G.-Y. Wei, "A process-variation-tolerant floating-point unit with voltage interpolation and variable latency," in *International Solid-State Circuits Conference (ISSCC)*, Feb. 2008.

- [29] X. Liang, G.-Y. Wei, and D. Brooks, "ReVIVaL: A variation tolerant architecture using voltage interpolation and variable latency," in *International Symposium on Computer Architecture (ISCA-35)*, June 2008.
- [30] X. Liang, G.-Y. Wei, and D. Brooks, "ReVIVaL: Variation tolerant microarchitecture," *IEEE Micro Top Picks*, Feb. 2009.
- [31] K. Brownell, G.-Y. Wei, and D. Brooks, "Evaluation of voltage interpolation to address process variations," in *International Conference on Computer-Aided Design (ICCAD)*, Nov. 2008.
- [32] X. Liang, R. Canal, G.-Y. Wei, and D. Brooks, "Process variation tolerant 3T1D-based cache architectures," in *International Symposium on Microarchitecture (MICRO-40)*, Dec. 2007. Nominated for CACM special issue consideration by SIGMICRO. Selected as one of the Top Picks in Computer Architecture in 2007.
- [33] X. Liang, R. Canal, G.-Y. Wei, and D. Brooks, "Replacing 6T SRAMs with 3T1D DRAMs in the L1 data cache to combat process variability," *IEEE Micro Top Picks*, Feb. 2008.
- [34] R. Joseph, D. Brooks, and M. Martonosi, "Control techniques to eliminate voltage emergencies in high performance processors.," in *International Conference on High-Performance Computer Architecture (HPCA-9)*, February 2003.
- [35] K. Hazelwood and D. Brooks, "Eliminating voltage emergencies via microarchitectural voltage control feedback and dynamic optimization," in *International Symposium on Low-Power Electronics and Design (ISLPED)*, August 2004.
- [36] M. S. Gupta, J. L. Oatley, R. Joseph, G.-Y. Wei, and D. Brooks, "Understanding voltage variations in chip multiprocessors using a distributed power-delivery network," in *Design, Automation, and Test in Europe Conference (DATE-10)*, April 2007.
- [37] M. S. Gupta, K. K. Rangan, M. D. Smith, G.-Y. Wei, and D. Brooks, "Towards a software approach to mitigate voltage emergencies," in *International Symposium on Low Power Electronics and Design (ISLPED)*, Aug. 2007.
- [38] V. Reddi, M. Gupta, G. Holloway, M. D. Smith, G.-Y. Wei, and D. Brooks, "Voltage emergency prediction: A signature-based approach to reducing voltage emergencies," in *International Conference on High-Performance Computer Architecture (HPCA-15)*, Feb. 2009. Nominated for Best Paper Award.
- [39] M. S. Gupta, K. K. Rangan, M. D. Smith, G.-Y. Wei, and D. Brooks, "DeCoR: A delayed commit and rollback mechanism for handling inductive noise in processors," in *International Symposium on High-Performance Computer Architecture (HPCA-14)*, Feb. 2008.
- [40] V. Reddi, M. Gupta, G.-Y. Wei, and D. Brooks, "An event-guided approach to handling inductive noise in processors," in *Design, Automation, and Test in Europe Conference (DATE-12)*, April 2009.
- [41] Q. Wu, V. Reddi, Y. Wu, J. Lee, D. Connors, D. Brooks, M. Martonosi, and D. W. Clark, "Dynamic compilation framework for controlling microprocessor energy and performance," in *International Symposium on Microarchitecture (MICRO-38)*, Nov. 2005. Best Paper Award. Selected as one of the Top Picks in Computer Architecture in 2005.
- [42] Q. Wu, V. Reddi, Y. Wu, J. Lee, D. Connors, D. Brooks, M. Martonosi, and D. W. Clark, "Dynamic compiler driven control for microprocessor energy and performance," *IEEE Micro Special Issue: Top Picks from Computer Architecture Conferences*, Jan/Feb 2006.

- [43] W. Kim, M. S. Gupta, G.-Y. Wei, and D. Brooks, “System level analysis of fast, per-core DVFS using on-chip switching regulators,” in *International Symposium on High-Performance Computer Architecture (HPCA-14)*, Feb. 2008.