

GUEST EDITORS' INTRODUCTION: POWER AND COMPLEXITY AWARE DESIGN

Pradip Bose
IBM T.J. Watson Research

David H. Albonesi
University of Rochester

Diana Marculescu
Carnegie Mellon
University

..... The quest for higher performance via deep pipelining (for high clock rate) and speculative, out-of-order execution (for high instructions per cycle) has yielded processors with increasing design complexity. The costs of higher complexity are manifold: increased verification time, higher power dissipation, reduced scalability in terms of microarchitectural resource size parameters and process shrinks, and so on. The recent industry-wide focus on processor power consumption as a major design constraint has forced companies to pay further attention to the underlying complexity issue.

It is important to understand that the two

issues, power and complexity, do not necessarily have complementary, causal effects, that is, an increase or decrease in one metric does not necessarily imply a corresponding increase or decrease in the other. To the extent that higher complexity generally translates into a higher transistor count, we can assert that complexity and area (and therefore leakage power, as well as *unmanaged* dynamic power) are positively correlated. On the other hand, power management through clock gating, V_{DD} gating, or adaptive resizing techniques usually *increases* verification complexity, even though these techniques reduce the net power consumption. Thus, in a sense, we must sep-

arately manage power and complexity, while remaining aware of the subtle interdependencies.

Future designs will require microarchitects, circuit designers, compiler developers, presilicon modelers, verification engineers, and system designers to cooperatively develop hardware and software techniques and tools that can break the power and complexity barriers. This is no less an issue today than the memory wall problem that architects have long worried about from a performance-only viewpoint.

With this key challenge in perspective, we decided to put together a special issue on this theme: power- and complexity-aware design (PCAD). We are pleased to present five excellent articles selected from 16 submissions. We based our selections on the *IEEE Micro* technical review process, which includes at least three reviews for each submission. These articles cover a broad range of topics related to the theme, but they do not address some aspects, such as verification complexity or design for “verifiability.” As such, before formally introducing each article, we provide a brief tutorial-style overview of the PCAD theme. This overview will hopefully help future researchers focus on PCAD topics not covered in this special issue.

Verification complexity

Perhaps the best-appreciated measure of design complexity is the cost of verification. Companies spend a major fraction (approximately 60 to 70 percent) of the net development cost of a processor or a system on chip (SoC) on verification and validation. As Moore’s law continues and it becomes possible to offer more and more transistors on a die, microarchitects invariably find ways of using all of those available devices. This has caused a steady escalation in chip verification costs. Can you estimate the verification complexity beforehand, during the microarchitecture definition phase? If so, can you choose a microarchitecture that bounds the verification cost to an affordable value? These are tough questions that remain unanswered in the PCAD work covered by articles in this special issue.

Informally, however, architects are aware of the following:

- Microarchitecture designs that employ ever-larger fractions of the die in implementing regular storage macros, like caches and register files, are more manageable in terms of verification than those containing mostly random logic.
- Cellular architectures that build processing power by replicating simple processing elements in grid-like structures promise sublinear growth in verification cost with advances in semiconductor technology.
- Modular design principles, in which designers use a single macro, unit, core, or even a control algorithm pervasively throughout the chip, usually enable modular verification strategies that reduce cost.
- Moving complexity from hardware to software (in other words, to the compiler) or firmware, when feasible, can reduce verification cost—at least in terms of hardware bring-up and time to market.

Many microprocessor chip (or chipset) designs employ a combination of these heuristics to manage the overall verification budget. A systematic and quantitative methodology of ensuring scalable verification complexity over the life of a product family is still a topic for future research.

Power dissipation

The power a processor consumes is a tangible metric measurable in familiar units (watts). Package and cooling costs increase with chip power, and these cost sensitivities are reasonably well known to those working on chip design teams. As such, much recent work proposes and evaluates power-efficient design ideas. The advent of microarchitecture-level power estimation tools such as Wattch (from Princeton University) and SimplePower (from Penn State) have facilitated academic research in this important aspect of PCAD research. Industrial R&D groups have also developed their own proprietary toolsets to enable early stage power-performance tradeoffs. Virtually all of the articles in this special issue touch on the energy-efficiency aspect of PCAD—either in terms of design or the underlying presilicon modeling issues.

Managing power by adding on-chip con-

trols can enhance overall energy efficiency but at the price of increased verification complexity. This type of design tradeoff that attempts to balance power reduction against verification cost is inadequately covered in current PCAD research.

Manufacturing cost, testability, and yield

Microarchitecture complexity can have a negative impact on testability, manufacturing cost, and effective chip yield. However, quantifying these aspects of complexity is not easy. Again, in qualitative terms, it seems intuitive that chip designs using complex, irregular microarchitectural constructs suffer from poor yield and testability metrics. These and other related factors could also lead to an increase in the microprocessor's effective manufacturing cost.

From the preceding discussion, defining quantitative metrics for evaluating complexity effectiveness is not easy, even if you interpret the measure of complexity from a single, restricted viewpoint, such as verification cost, power, testability, yield, or manufacturability. So far, the architecture community has found it easiest to view complexity effectiveness primarily in terms of power efficiency. But even here, the issue of using the right metric in the right context—whether millions of instructions per second (MIPS) per watt, versus MIPS² per watt, MIPS³ per watt, and so on—remains rather poorly understood in architectural research today. In putting together this issue, we expected to publish some new thinking on the topic of “metrics” that measure complexity effectiveness and power efficiency, but the submissions did not adequately cover this topic.

In this issue

Huang et al., in the first article of this issue, focus on the branch prediction logic within a current-generation superscalar microprocessor. This piece of modern microarchitecture has undergone progressive increases in relative area and control complexity in an attempt to attain ever-higher branch prediction accuracies. Clearly, not all applications require the large prediction tables and multilevel decision schemes to yield acceptably high prediction accuracy. It would therefore make sense to architect the branch prediction mechanism to be reconfigurable, where the effective (dynamic) complexity and power consumption is a

function of the workload's inherent predictability. This is the basic idea behind the research reported in this article. The authors demonstrate reductions in branch predictor energy consumption by up to 90 percent without appreciable loss in prediction accuracy and performance.

In their article on statistical simulation, Eeckhout et al. address the problem of deriving fast, early stage design tradeoff decisions. To explore the PCAD design space at high speed, it is fruitful to complement classical trace-driven, cycle-accurate simulation tools with more abstract statistical models. This article presents a comprehensive view of such statistical analysis methods.

Julien et al. present an article that describes a new approach to characterizing the power dissipation in complex digital signal processors (DSPs). They illustrate this methodology by applying it to the Texas Instruments C6201 DSP. After proposing a power model, they validate it against actual measurements, with the observed error margin being within 4 percent.

The issue queues within a current-generation, out-of-order superscalar processor are known hot spots in terms of power dissipation and also present a complex problem in terms of control. Recently, various research groups have focused on power- and complexity-aware designs for issue logic in general. In this issue, the article by Abella, Canal, and Gonzalez provides a sound survey of PCAD techniques applied to this key aspect of modern microprocessors.

Last, but not the least, Fryman et al. present an article that explores the energy and delay tradeoffs that occur when you move some or all of the local storage out of a given embedded device and on to a remote network server. The authors demonstrate that using the network to access remote storage in lieu of local DRAM results in significant power savings.

We hope you enjoy this theme issue on PCAD, consisting of a carefully reviewed selection of articles that touch on a broad range of topics. MICRO

Pradip Bose is research staff member and project leader at the IBM T.J. Watson Research Center; he is also editor in chief of *IEEE Micro*. His research interests include high-performance computer architectures, power-

and complexity-aware design, and computer-aided design. Bose has a PhD in electrical and computer engineering from the University of Illinois, Urbana-Champaign. He is a senior member of the IEEE and the Computer Society and a member of ACM.

David H. Albonesi is an associate professor of electrical and computer engineering at the University of Rochester. His research interests include power-aware computing, adaptive microarchitectures, and multithreaded architectures. Albonesi has a PhD in computer engineering from the University of Massachusetts at Amherst. He is a senior member of the IEEE and a member of the ACM.

Diana Marculescu is an assistant professor of electrical and computer engineering at Carnegie Mellon University, Pittsburgh, Pennsylvania. Her research interests include

energy-aware computing, CAD tools for low-power systems, and emerging technologies (such as electronic textiles or ambient intelligent systems). Marculescu has a MS in computer science from the Politehnica University of Bucharest, Romania, and a PhD in computer engineering from the University of Southern California. She is a member of the IEEE, the ACM, and the ACM Special Interest Group on Design Automation (SIGDA).


Direct questions and comments about this special issue to Pradip Bose, IBM T.J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598; pbose@us.ibm.com.

For further information on this or any other computing topic, visit our Digital Library at <http://computer.org/publications/dlib>.

WE'RE ONLINE

Submit your manuscript to Micro on the Web!

Our new Manuscript Central site lets you monitor your submission as it progresses through the review process. This new Web-based systems helps our editorial board and reviewers track your manuscript.

IEEE 
COMPUTER
SOCIETY

 **IEEE**

For more information, see us at <http://cs-ieee.manuscriptcentral.com>