Impact of Process Variations on Multicore Performance Symmetry

Eric Humenay, David Tarjan, Kevin Skadron
Dept. of Computer Science, University of Virginia
Charlottesville, VA 22904
humenay@virginia.edu, dtarjan@cs.virginia.edu, skadron@cs.virginia.edu

Abstract

Multi-core architectures introduce a new granularity at which process variations may occur, yielding asymmetry among cores that were designed—and that software expects—to be symmetric in performance. The chief source of this phenomenon are highly correlated, "systematic" within-die variations such as optical imperfections yielding variations across the exposure field. Per-core voltages can be used to bring all cores to the same performance level, but this compensation strategy also affects power, chiefly due to leakage power. Boosting a core's frequency may therefore boost its leakage sufficiently to engage thermal throttling. This sets up a tradeoff between static *performance asymmetry due to frequency variation versus* dynamic *performance asymmetry due to thermal throttling. This paper explores the potential magnitude of these effects.*

1. Introduction

The 2005 International Technology Roadmap for Semiconductors [14] projects that parameter variations will present critical challenges for manufacturability and yield. At the same time, multicore designs have become the dominant organization for future high-performance microprocessors. The inclusion of multiple cores allows continued exponential performance scaling for applications that exhibit a high degree of parallelism. Multi-core architectures, however, also multiply the ways in which parameter variations can affect a processor.

Parameter variations encompass *process* variations due to manufacturing phenomena, *voltage* variations due to manufacturing and runtime phenomena, and *temperature* variations due to varying activity levels and power dissipations. Process variations are static and manifest themselves as *die-to-die* (D2D), *within-die* (WID), and em wafer-to-wafer variations (W2W), while temperature and voltage variations are a dynamic phenomena.

This paper introduces a new granularity of particular interest to microarchitects, *core-to-core* (C2C) variations, which arise due to spatially correlated WID variation, for example due to non-uniformity in the lithographic exposure field. Individual cores are now small enough that the chief impact of many spatially correlated phenomena manifests across rather than within cores. This is a problem because multicore chips with non-uniform frequency or power characteristics from core to core create scheduling and thermal-management problems. This can cause reduced throughput [2], missed real-time deadlines, or excessive thermal throttling if more computationally intensive threads are mapped to higher-power cores.

Of course, these problems can always be rectified by slowing all the cores down to the frequency of the slowest core, but this reduces the yield of premium chips. Instead, cores that are initially slow can be sped up to in order to reduce C2C frequency heterogeneity. However, frequency compensation techniques entail additional power, chiefly due to leakage, so thermal constraints limit the symmetry obtained from frequency compensating techniques.

In order to demonstrate the importance and motivate work on hardware and software techniques to address the problem, this paper presents preliminary work to characterize the magnitude of C2C frequency variations and the extent to which per-core frequency compensation is limited by thermal throttling. After background and related work in Section 2, Section 3 describes our model for C2C variation, Section 4 describes our experimental setup, and Section 5 presents the results. Section 6 concludes the paper.

2. Background and Related Work

2.1. Process-Variation Background

Process variations cause maximum clockable frequency and power dissipation of a high-performance chip to vary from the target frequency and from chip to chip. Post-manufacture testing is used to characterize chips and identify the best operating frequency for each. Unfortunately, faster chips usually have higher sub-threshold leakage currents, because the main contributor to frequency variations, L_{eff}, also affects sub-threshold leakage. In fact, the fastest chips often cannot operate at their peak sustainable frequency because the excessive leakage causes the chip to overheat, and a suitable cooling solution may be too expensive. Slower chips must increase their frequency or be sold at a lower profit. Per-chip adaptive body biasing (ABB)

and adaptive voltage scaling (AVS) can reduce these spreads and boost the yield of high-quality parts at the cost of some additional test-time circuitry [15, 16].

As feature sizes become smaller with technology scaling, WID variations become relatively more important. While it was once sufficient to deal with them by adding some error margin and by binning, at future technology nodes, these techniques will no longer suffice to obtain satisfactory yield of premium parts. Furthermore,

WID variations can be sub-divided into two main categories: random and systematic. *Random* variations are small changes from transistor to transistor typically modeled with a normal distribution. *Systematic* WID variations, on the other hand, exhibit high degrees of spatial correlation. This paper argues that these correlated phenomena are important in an era of multicore chips.

2.2. Prior Treatments of Systematic WID

Very little work has considered the impact of systematic WID variations. Zhang et al.[18] present a modeling methodology for determining chip-wide subthreshold leakage and show the importance of considering systematic WID variation. Ashouei et al. [1] propose a model to addresses WID systematic leakage variation at the circuit level, treating systematic variations as circular areas on the die with highly correlated L_{eff} values. These circular areas may vary in their area, location, and magnitude. Our modeling methodology differs since our pattern of variation is based on measured data from [6, 13].

A die's pattern of systematic WID variation is highly dependent upon the fabrication process, and can be deterministic or stochastic in nature. Deterministic systematic variations can be mitigated with a combination of optimal proximity correction, phase-shift masking, as well as other mask-level techniques. Since masks cost are already burdensome and increasing with every technology node, design-for-manufacture techniques that simplify mask complexity with variation-tolerant designs are desired.

The main advantage of modeling a measured deterministic systematic pattern is to better understand at what granularity the systematic change will occur, and how this will affect multicore architectural decisions. This paper shows that systematic WID variation can cause large performance, power, and thermal variations among cores intended to be identical.

2.3. Architectural Implications

Very little work to date has considered how variations affect the *microarchitecture*.

An important basis for much work in this area is the modeling work by Bowman et al.[3]. This paper proposed an analytical model to capture the maximum clockable frequency (FMAX) distribution. A generic critical path model is derived from a canonical NAND gate. The NAND's delay is derived from the RC delay equation and the delay distribution is determined by Monte Carlo simulation. Two basic parameters then suffice to illustrate the way circuit and microarchitecture choices determine sensitivity to variations: number of independent critical paths, N_{cp}, and critical path depth, n_{cp}. The ratio of variance to mean, σ/μ, decreases with both N_{cp} and n_{cp}. This paper is the only work we are aware of that reports actual measured WID frequency distribution. The WID σ/μ of three different critical paths is shown to be roughly 3%. The WID FMAX σ/μ will be considerably less than 3% since a max operation must be performed across N_{cp} critical paths. Bowman et al. conclude that WID_{rand} variations only affect the processor's mean frequency, and D2D variations then determine frequency variance.

Marculescu and Talpes [11] apply the FMAX model in the microarchitecture domain by assuming that N_{cp} is proportional to the stage's device count. The authors show that a GALS architecture can mitigate the impact of process and temperature variations, because a globally asynchronous processor does not require that the global frequency be dictated by the worst-case delay of all critical paths. Rather, each clock domain's frequency is determined only by the slowest path in the domain, and buffering limits the impact of the slowest domain.

In our prior work [9], we also used FMAX as a starting point and showed that N_{cp} is not simply proportional to a stage's device count, because array structures such as register files and caches have many short critical paths (essentially guaranteeing many instances of the worst case), while datapath logic has longer critical paths. This means that SRAM structures will likely experience the worst within-core unit-to-unit variations. Other work, e.g. [17], has also demonstrated the severity of WID variations in SRAMs.

Core-to-core variations are important if software has been designed assuming symmetric core performance. Balakrishnan et al. [2] considered the impact of asymmetric multiprocessor performance on multithreaded commercial workloads. They observed highly variable and generally suboptimal performance because the operating system and or application could unwittingly assign too much work to slow cores and too little to fast cores. This calls for an interface to expose core asymmetry to the software, but also for hardware techniques to mitigate the asymmetry, which is the focus of this paper.

3. Model

For frequency binning and marketing, C2C frequency variations present an interesting problem for vendors. Should the chip be marketed by the frequency of the slowest core, an average of all cores, or some alternative method?

ABB and AVS have been proposed for reducing bin

splits, and could also be used within a chip to mitigate C2C variation. However, both techniques also increase power dissipation. This may cause the compensated cores to more easily overheat, incurring thermal throttling and trading static, predictable asymmetry for dynamic, unpredictable throttling.

We have previously argued that only systematic WID variations are likely to induce C2C variations. In order to estimate the possible magnitude of C2C asymmetry, we develop a model based on the pattern and magnitude measured by Cain [6]. We assume that the chief source of systematic WID variation is variability in effective gate length (L_{eff}) due to optical variations across the exposure field. While these variations are spatially correlated, over the large exposure field required for large multicore chips, the magnitude can impact cores which are sufficiently far apart (as they may be to minimize thermal coupling).

The optical component that we model is chiefly due to lens aberrations and can be modeled as a simple polynomial function of position within the exposure field. Assuming a 28mm X 28mm exposure field divided into four 14mm X 14mm chips, the cross-chip systematic variation in L_{eff} (in nm), Δ_{sys}, for the die positioned in the lower-left hand quadrant of the reticle can be approximated by:

$$\Delta_{sys} = a{\cdot}x^2 + b{\cdot}y^2 + c{\cdot}x + d{\cdot}y + e{\cdot}xy + Intercept \qquad (1)$$

We have scaled the coefficients in this model in order to model variations at the 45nm technology node. A 2D contour map of the average WID systematic pattern for the chip located in the lower-left hand quadrant of the reticle is shown in Figure 1. This was derived using Eqn. 1 and the baseline constants in Table 1. Note that if the systematic variation is stochastic (in which case it is not properly called systematic, but rather random, spatially correlated varation), each chip will have a unique distribution.

Parameter	Value
a	5.37×10^{-4}nm/mm^2
b	1.829×10^{-3} nm/mm^2
c	-1.06×10^{-2}nm/mm
d	-.458 nm/mm
e	-1.67×10^{-3}nm/mm
Intercept	3.0 nm

Table 1. Constants for the 2nd order polynomial modeling WID systematic variations.

Systematic L_{eff} variations chiefly affect gate delay. Orshansky et al. [13], propose the following equation for modeling the dependency between L_{eff} and delay:

$$D \sim L_{eff}^{1.5} \cdot V_{dd}/(V_{dd} - V_{th})^{\alpha} \qquad (2)$$

where V_{dd} is supply voltage, V_{th} is threshold voltage, and α is velocity saturation. As channel lengths become shorter

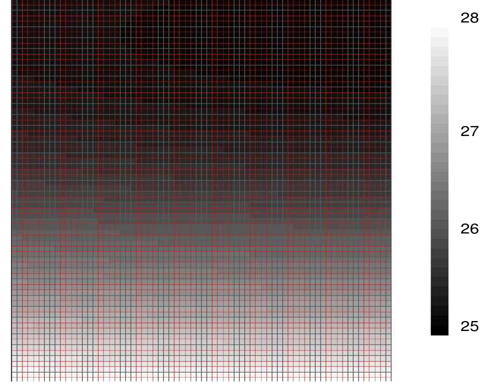

Figure 1. 14mm X 14mm 2D contour map of cross-chip variation in L_{eff} (in nm).

this value will approach 1. For 45nm devices we judged $\alpha = 1.3$ to be an appropriate value.

Because of drain induced barrier lowering (DIBL), V_{th} and L_{eff} are also related according to [7]:

$$V_{th_{eff}} = V_{th0} - V_{dd} \cdot exp(-\alpha_{DIBL} \cdot L_{eff}) \qquad (3)$$

where V_{th0} is the threshold voltage for long channel transistors, 0.22; α_{DIBL} is the DIBL coefficient, 0.15; and V_{dd} is the supply voltage, 1V in this study. The default values for V_{th0} and α_{DIBL} were provided in [5].

Because subthreshold leakage is an exponential function of V_{th}, the variations in L_{eff} also cause C2C leakage variation. We treat the smallest L_{eff} due to systematic variations as our nominal target, so these leakage variations simply mean that slow cores are also less leaky.

4. Experimental Methodology

To evaluate the magnitude of performance asymmetry in a high-performance multicore chip, we consider a POWER4-like core scaled to a 45nm technology node and nominally operating at 3.0 GHz and 1.0V. Assuming constant scaling, the core area (including first-level cache) will be 2.0mm by 2.25mm. The baseline floorplan we model consists of 9 cores evenly distributed across the chip with each core being surrounded by L2 cache as shown in Figure 2. To consider the tradeoff between grouping cores together to minimize the impact of the exposure-field variation versus the higher temperatures resulting from thermal coupling, we later compare this this distributed floorplan to one with all the cores adjacent to each other.

While this is a fairly arbitrary choice of core count and placement, the multicore design space and associated floorplan design space are staggeringly large and beyond the scope of this study. The focus here is simply to show the impact of systematic WID variation on C2C performance variation, and this simple floorplan suffices to illustrate the potential problems.

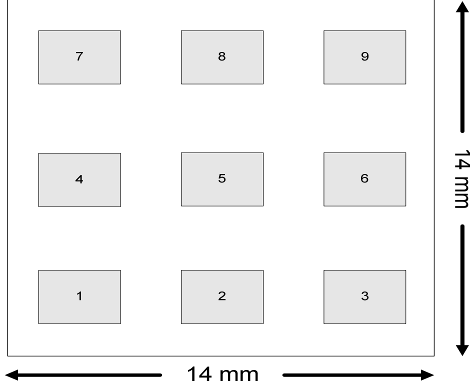

Figure 2. Multi-core floorplan.

HotLeakage [10] was used for subthreshold leakage modeling. Dynamic performance and power data was gathered from the Turandot/PowerTimer/Hotspot simulation environment [4, 12, 8]. At 45nm, sub-threshold leakage contributes to roughly 28% of the total core power with nominal V_{dd} and a temperature of 360K.

The two most natural techniques for compensating core frequencies to obtain symmetric performance are adaptive voltage scaling (AVS) and adaptive body biasing (ABB). Both techniques, however, have an exponential impact on leakage power. AVS also has a cubic impact and ABB a linear impact on dynamic power.

Performance asymmetry can always be eliminated without negative power consequences by slowing down all cores to the speed of the slowest core, but this wastes potential performance. In this paper, we consider how much performance can be reclaimed and how much asymmetry can be eliminated by boosting slow cores. To achieve the desired frequency boost, AVS requires a much smaller change (percentage-wise) in supply voltage than ABB requires in threshold voltage. As a result, AVS has a much milder impact on leakage and is a more power-efficient and thermally compatible solution than ABB. Figure 3 illustrates the total (static + dynamic) increase in power required to achieve a desired frequency boost. To boost frequency by 10% requires a 16% change in V_{dd} but a 30

Implementing AVS requires the ability to provide each core with a different supply voltage, as well as a way to measure each core's maximum clockable frequency during testing and then compute the necessary V_{dd} scaling.

5. Symmetrical Core Performance

5.1. Magnitude of C2C Variation

In our model, the variation in L_{eff} is stronger in the Y dimension than in the X dimension. If we therefore calculate the resulting frequency distribution for the floorplan shown in Figure 2, the frequencies break intro groups corresponding to rows of cores, as shown in Table 2. Within a row, the frequency variation is minimal.

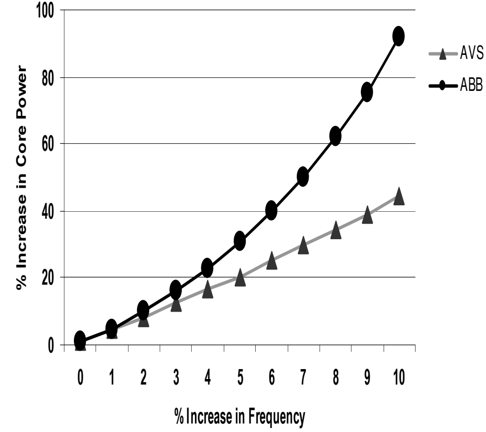

Figure 3. Comparison of the performance/power tradeoff for voltage scaling and ABB at 360K. To achieve 10% improvement in frequency requires a 16% increase in V_{dd} (from 1.0 to 1.16V) but a 30% decrease in V_{th} (from 0.2 to 0.14V).

	Mean norm. freq.	Mean norm. power
Row 1 (Cores 7–9)	0.995 ± 0.005	1.000 ± 0.002
Row 2 (Cores 4–6)	0.952 ± 0.004	0.950 ± 0.004
Row 3 (Cores 1–3)	0.826 ± 0.002	0.814 ± 0.002

Table 2. Pre-compensation frequency and power distribution (normalized to nominal) due to systematic WID variation for the sample floorplan shown in Figure 2.

5.2. Impact of Thermal Throttling

Increased V_{dd} from AVS will, however, cause a core's power density to be greater than nominal, resulting in higher temperatures on those cores. For the most affected core, with an initial frequency at 82.4% and power at 81.0% of nominal, boosting its frequency to 100% boosts its power to 166% of nominal! Depending on the workload, these cores will periodically engage thermal throttling. AVS therefore eliminates static performance asymmetry (frequency) at the cost of dynamic performance asymmetry (thermal throttling), or requires a more expensive cooling solution.

Dynamic voltage and frequency scaling (DVFS) is a commonly accepted form of thermal throttling, because its reduction in power density is roughly cubic relative to the performance loss. Figure 4 shows the potential slowdown from thermal throttling using DVFS for different degrees of frequency compensation. These results were obtained using *gcc*, the hottest of the SPECcpu2000 benchmarks. These results assume that the package and cooling solution are the

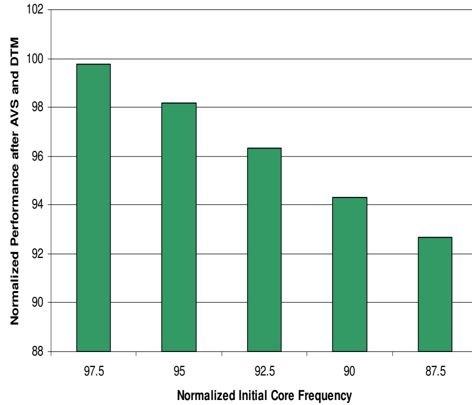

Figure 4. Performance degradation due to thermal throttling when AVS is used to increase core frequency to the nominal speed. The x-axis presents the initial core frequency prior to AVS compensation.

minimum required to never require thermal throttling for a nominal core, and assume that exactly the necessary voltage and frequency can be selected. We chose the worst case benchmark, *gcc*, in order to illustrate the potential magnitude of the effect.

5.3. Impact of Floorplan

In the presence of systematic WID variation, the closer to each other that cores are placed, the more correlated their frequencies will be. For this reason, the baseline floorplan (Figure 2) will be more susceptible to C2C performance variations than a floorplan that has a denser distribution of cores such as the one shown in Figure 5. On the other hand, in the absence of frequency compensating techniques, the floorplan with a denser core layout will be much hotter than the floorplan with a distributed layout: the majority of the die's power is dissipated in a small concentrated area, while the cache surrounding each core in the distributed floorplan provides some thermal buffering. Core proximity therefore offers a tradeoff between performance asymmetry due to frequency variation and thermal throttling due to high core temperatures after compensation, versus reduced (but possibly more uniform) performance due to higher temperatures with densely placed cores. The better choice depends on how severe the cross-chip systematic WID variation is.

In order to evaluate this tradeoff, post-AVS steady state temperatures were calculated for both floorplans as a function of the magnitude of systematic WID variation. Each core is running *gcc*, our hottest benchmark program. Each floorplan's hottest chip temperatures is shown in Figure 6. The X axis of the graph shows the magnitude of systematic variation across the die as a percentage of nominal L_{eff} (25nm); 10% corresponds to the value calculated in Eqn. 1 and shown in Figure 1.

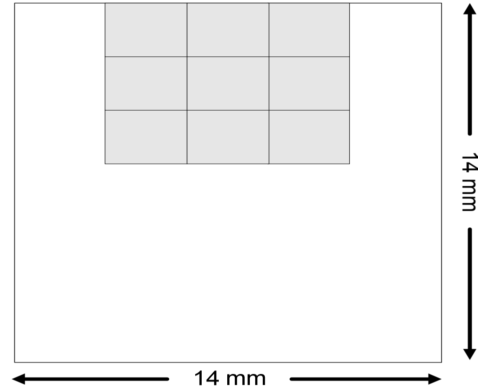

Figure 5. Floorplan with dense core layout.

When cross-die variation is small, frequency variation is small and only minimal core compensation is required, so the distributed floorplan will be cooler than the dense floorplan. As the amount of cross-die variation increases, frequency variation increases and AVS must be applied more aggressively. In contrast, in the dense floorplan, the cores' frequencies are tightly correlated even with large cross-die variation, so the dominant factor is that the dense placement reduces lateral heat transfer. The distributed floorplan's maximum temperature is therefore heavily dependent on the magnitude of systematic WID variation, while the dense floorplan is relatively immune.

Overall, these results indicate that the floorplan must be designed with the likely magnitude of systematic WID variations in mind. These considerations therefore need to be explored before the floorplan is fixed.

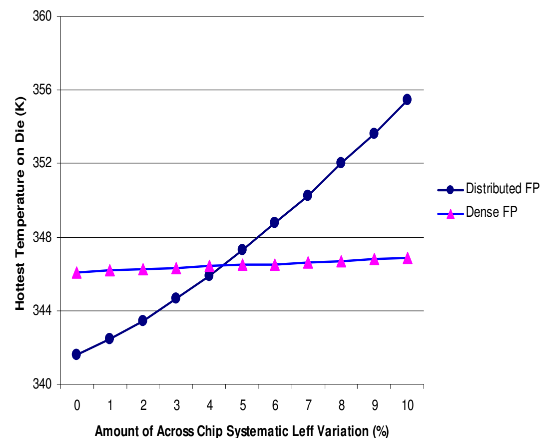

Figure 6. Post-AVS chip temperatures for both floorplans when different amounts of systematic variations are considered

6. Conclusions and Future Work

This paper analyzes the performance impact of systematic within-die (WID) parameter variations for multicore

chips. The main contributions are:

- Cores are becoming sufficiently small with technology scaling that spatially correlated phenomena like optical-field variations can introduce significant "systematic" WID variations that produce significant core-to-core (C2C) frequency asymmetry.
- Adaptive voltage scaling (AVS) can improve yield and the software impact of C2C asymmetry by reducing the frequency spread among cores.
- For substantial C2C frequency asymmetry, AVS raises leakage too much, and performance homogeneity will be unattainable without a more expensive cooling solution. Otherwise thermal throttling occurs, thus trading static performance asymmetry (frequency) for dynamic (throttling).
- The choice of floorplan has an important effect on core-to-core asymmetry. When cores are distributed across a large die, they are vulnerable to WID systematic variations. When cores are placed close to each other, the increased power density incurs a greater risk of thermal throttling. This creates a multidimensional tradeoff space among core power, floorplan, magnitude of cross-chip variation, and cooling cost.

Both hardware and software techniques are needed to address the problems created by C2C asymmetry. In addition to hardware techniques to mitigate the asymmetry, algorithms are needed to find the optimal frequency that balances performance loss against asymmetry. New instruction-set architecture mechanisms are needed to expose C2C asymmetry to software, and new scheduling techniques are needed to allow software to adapt to the asymmetry of each unique chip.

Acknowledgments

This work has been supported in part by NSF grant nos. CCR-0133634 (CAREER), CCF-0429765, Army Research Office grant W911NF-04-1-0288, and a research grant from Intel MTL. We would like to thank Wei Huang and Mircea Stan for their assistance and the anonymous reviewers for their helpful comments.

References

[1] M. Ashouei, A. Chatterjee, A. D. Singh, V. De, and T. M. Mak. Statistical estimation of correlated leakage power variation and its application to leakage-aware design. In *Proc. VLSI Design 2006*, Jan. 2006.

[2] S. Balakrishnan, R. Rajwar, M. Upton, and K. Lai. The impact of performance asymmetry in emerging multicore srchitectures. In *Proc. ISCA 2005*, June 2005.

[3] K. Bowman, S. Duvall, and J. Meindl. Impact of Die-to-Die and Within-Die Parameter Fluctuations on the Maximum Clock Frequency Distribution for Gigascale Integration. *IEEE J. Solid State Electronics*, 37(2):183–90, Feb. 2002.

[4] D. Brooks, P. Bose, V. Srinivasan, M. K. Gschwind, P. G. Emma, and M. G. Rosenfield. New methodology for early-stage, microarchitecture-level oower-performance analysis of microprocessors. *IBM J. Res. Dev.*, 47(5-6):653–670, Sep. 2003.

[5] Berkeley predictive technology model. http://www-device.eecs.berkeley.edu.

[6] J. Cain. Characterization of spatial variability in photolithography. Master's thesis, Univ. of California, Berkeley EECS Dept., Nov. 2002.

[7] Y. Cao and L. T. Clark. Mapping Statistical Process Variations Toward Circuit Performance Variability: An Analytical Modeling Approach. In *Proc. 42nd DAC*, June 2005.

[8] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, and S. Ghosh. Hotspot: A compact thermal modeling method for CMOS VLSI systems. *IEEE Trans. VLSI Systems*, 14(5):501–513, May 2006.

[9] E. Humenay, D. Tarjan, and K. Skadron. Impact of parameter variations on multi-core chips. In *Proc. Wkshp. on Architecture Support for Gigascale Integration*, June 2006.

[10] Y. Li, D. Parikh, Y. Zhang, K. Sankaranarayanan, K. Skadron, and M. Stan. State-preserving vs. non-state preserving leakage control in caches. In *Proc. DATE 2004*, Feb. 2004.

[11] D. Marculescu and E. Talpes. Variability and energy awareness: a microarchitecture-level perspective. In *Proc. 42nd DAC*, June 2005.

[12] M. Moudgill, J. Wellman, and J. Moreno. Environment for PowerPC microarchitectural exploration. *IEEE Micro*, 19(3):15–25, May/June 1999.

[13] M. Orshanksy, L. Milor, and C. Hu. Characterization of spatial intrafield gate CD variability, its impact on circuit performance, and spatial mask-level correction. *IEEE Trans. Semiconductor Manufacturing*, 17(1):2–11, Feb. 2004.

[14] SIA. *International Technology Roadmap for Semiconductors*, 2005. http://public.itrs.net.

[15] J. Tschanz, K. Bowman, and V. De. Variation-tolerant Circuits: Circuit Solutions and Techniques. In *Proc. 42nd DAC*, June 2005.

[16] J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De. Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage. *IEEE J. Solid-State Electronics*, 37(11):1396–1402, Nov. 2002.

[17] H. Wang, M. Miranda, W. Dehaene, F. Catthoor, and K. Maex. Systematic analysis of energy and delay impact of very deep submicron process variability effects in embedded sram modules. In *Proc. DATE 2005*, Mar. 2005.

[18] S. Zhang, V. Wason, and K. Banerjee. A probabilistic framework to estimate full-chips subthreshold leakage power distribution considering within-die and die-to-die P-T-V variations. In *Proc. ISLPED 2004*, Aug. 2004.