
IMPACT OF PARAMETER VARIATIONS ON CIRCUITS AND MICROARCHITECTURE

Osman S. Unsal

Barcelona Supercomputing Center

James W. Tschanz

Keith Bowman

Vivek De

Circuits Research Lab, Intel

Microprocessor Technology Labs

Xavier Vera

Intel Barcelona Research Center

Antonio González

Universitat Politècnica de

Catalunya

& Intel Barcelona Research Center

Oguz Ergin

TOBB University of Economics and

Technology, Ankara

PARAMETER VARIATIONS, WHICH ARE INCREASING ALONG WITH ADVANCES

IN PROCESS TECHNOLOGIES, AFFECT BOTH TIMING AND POWER. VARIABILITY

MUST BE CONSIDERED AT BOTH THE CIRCUIT AND MICROARCHITECTURAL

DESIGN LEVELS TO KEEP PACE WITH PERFORMANCE SCALING AND TO KEEP

POWER CONSUMPTION WITHIN REASONABLE LIMITS. THIS ARTICLE PRESENTS

AN OVERVIEW OF THE MAIN SOURCES OF VARIABILITY AND SURVEYS

VARIATION-TOLERANT CIRCUIT AND MICROARCHITECTURAL APPROACHES.

..... Parameter variations have a great impact on maximum clock frequency. Designers set a processor's clock frequency to allow for the worst-case critical-path delay plus a safety margin. This process, known as guardbanding, is necessary because delays are not constant: Variations in process, voltage, temperature, and input values (PVTI) all contribute to the worst-case critical-path delay. To ensure correctness, designers must sum up the circuit-related worst-case delay and a safety margin for each PVTI element and then include this PVTI-related delay in the worst-case delay calculation. PVTI-related variability increases with technology scaling, especially beyond 90 nm, so safety margins are becoming an important path delay component.¹ That is, designers must make the clock cycle time much longer than actual delays to guarantee correctness.

Increased variability in the newer process technologies also has a serious cost implication: Companies must discard more low-per-

formance parts, which increases costs and decreases total revenues. Moreover, PVTI variations are chiefly responsible for leakage current fluctuations, which can vary by as much as a factor of 20 across dies.¹ In the future, excessive leakage currents and resulting extreme temperatures could phase out the standard burn-in test, necessitating alternative approaches to identifying infant mortality in manufactured processors.²

Thus, by affecting yield, increasing variability is becoming a reliability concern. The 2004 International Technology Roadmap for Semiconductors (<http://www.itrs.net>) identified variability as a key design challenge. The annual report identifies future challenges to IC performance and has been an excellent predictor of possible problem areas—for example, leakage. If alternative strategies are not developed, extreme device and circuit variability could stall the benefits of technology scaling. At Intel, within-die parameter variability causes gate delay

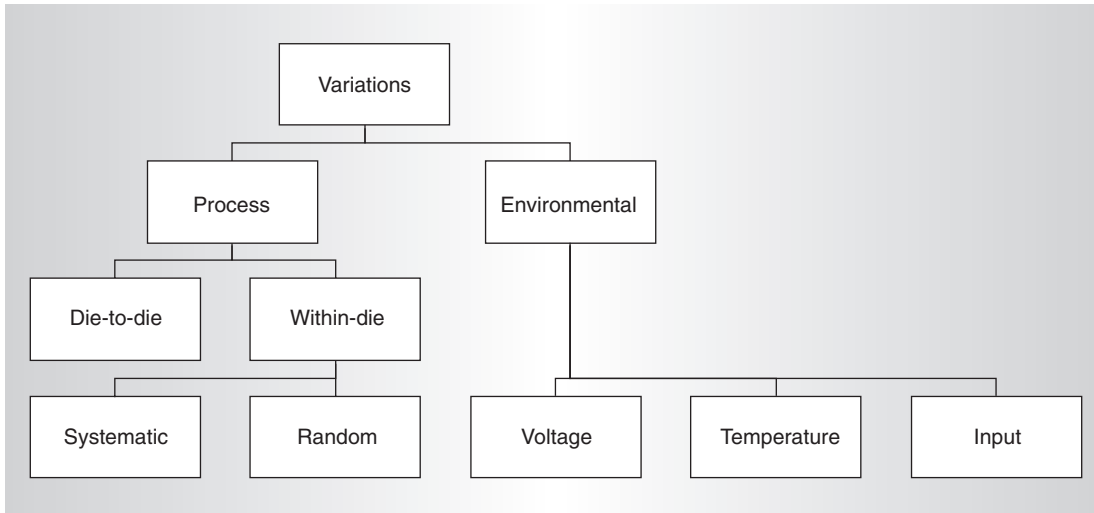


Figure 1. Parameter variations.

variations that account for a large percentage of the minimum (hold) and maximum (setup) delay margins at the 130-nm technology node.

The microarchitectural trend toward multicore processors will also contribute to the impact of variation in future processor designs. The projected benefit of a multicore design is that, in contrast to a uniprocessor design, it can achieve equivalent throughput with lower supply voltage (V_{DD}) and frequency. However, as V_{DD} scales relative to transistor threshold voltage (V_{TH}), the sensitivity of circuit delays to transistor parameter variations amplifies significantly. Thus, the impact of variations will increase as V_{DD} scales down.

In view of this growing threat, system design must take variability into account at all levels, including the microarchitecture, to keep pace with performance scaling. In this article, we present an overview of this problem and review variation-tolerant circuit and microarchitectural design approaches.

Introduction to variations

Figure 1 presents a general classification of the parameter variations we discuss in this article. Most process variations are static and stem from equipment processing. Environmental variations change with a part's use and workload.

Process variations

Die-to-die fluctuations (from lot to lot and wafer to wafer) result from factors such as pro-

cessing temperature and equipment properties.³ Conversely, within-die variations result from factors such as nondeterministic placement of dopant atoms and channel length variation across a single die. Traditionally, die-to-die fluctuations were the main concern in CMOS digital circuit design.⁴ However, within-die variations have become important as well, and their impact on frequency and power is becoming more pronounced.^{3,5}

Process dimensions have scaled by 30 percent each generation, resulting in a twofold increase in density and continuous improvement in transistor performance. However, future scaling will exacerbate process variation. This variability has several manufacturing-related causes. For example, process features printed today are smaller than the wavelength of light used to expose the mask, resulting in increased variability.

Classification

We classify process variations according to several characteristics:

- *source*—polishing, lithography, resist, etching, and doping;
- *granularity*—lot-to-lot or within-lot, wafer-to-wafer or within-wafer, and die-to-die or within-die;
- *manifestation*—random or systematic;
- *design parameter*—gate length, width, interconnect width, or threshold voltage; and

- *aging*—static or dynamic.

Source. Variations that come from chemical mechanical polishing variations can be traced back to nonuniform layout density. The chip's denser sections slow the polishing process. As a result, the dielectric in those sections is more highly polished than less dense sections, leading to differences in dielectric thickness across the die as great as thousands of angstroms.

As feature size has shrunk, variability due to lithography-related issues has become more pronounced. However, moving below the current 193-nm lithography wavelength won't solve the problem. A significant number of lithography-related variability problems stem from the stepper. The step-and-repeat process exposes each region of the wafer at different times. The wafer stepper holds the wafer and aligns the optics with each site. After each exposure, the stepper mechanism moves the equipment to the next site. Stepper lens heating, uneven lens focusing, and related aberrations cause variability.

After exposure, the wafer is coated with liquid plastic, using spin-on resist. The resist coating is uniform except at the wafer's edge, where surface tension causes beads, leading to thickness variation. After exposure and resist, the wafer is etched. Unevenness in etching power and density cause depth variations. Doping takes place after etching. The number of dopant atoms has decreased with scaling and is currently on the order of hundreds for the effective channel. Uniformly depositing the same small number of dopant atoms across billions of transistors on a die is impossible; thus, dopant concentration is becoming an important component of device variability.

Granularity. Die-to-die fluctuations can stem from lot-to-lot variations, wafer-to-wafer variations, or the type of within-wafer variations that affect every element on a chip equally.³ These variations are due to differences in lot- and wafer-scale processing characteristics such as temperatures, fab equipment properties, and wafer placement. In particular, within-wafer die-to-die variations can arise from oxide thickness fluctuations.

Within-die variations, on the other hand, stem from variations that create nonuniform design parameters (electrical characteristics)

across a single chip. These variations are the result of within-wafer processing differences such as resist thickness variation, stepper lens focus aberrations, and uneven doping.

Designers usually handle die-to-die variations with circuit techniques, whereas within-die variations are more amenable to architectural approaches. Therefore, in this article, we emphasize within-die variation sources.

Manifestation. Systematic variations are repetitive, and designers and microarchitects can characterize them, whereas random variations vary independently from device to device and are unpredictable. Because we can model systematic variations, they are amenable to elimination. However, some systematic variations are difficult to model, and designers treat them as random. Resist thickness fluctuation and lens aberrations are systematic, whereas dopant variations are random. Whether random or systematic variations will dominate systems in the future is open to debate. Microarchitects can construct optimistic or pessimistic scenarios, depending on the magnitude of variations. Scenarios in which both manifestations exist with equal intensity are also possible. The literature provides pointers on how to construct a suitable model of variations.^{3,6}

Design impact. Each variation source increases the variability of one or more key design parameters, such as channel length (also called the critical dimension), device and interconnect width, and V_{TH} . The major source of circuit performance variability is channel length variability, which is due to both die-to-die variations and within-die effects. Because transistor V_{TH} is a strong function of channel length for short-channel devices, this channel length variation affects both switching speed and static leakage current. Channel length variability results from wafer nonuniformity, lens focus and aberration, or line edge roughness. Device width variability is due to polishing or lithography issues such as poly and diffusion rounding. Interconnect width variability is due to etching, polishing, or lithography, and is important because it leads to erosion, dishing, trench depth, or via height variation. Causes of V_{TH} variability are oxide thickness and dopant fluctuations.

Aging. Most process effects, such as random dopant fluctuation and nonuniformity in etching and polishing, are static—they are determined by the fabrication process and don't change throughout the part's lifetime. In contrast, some process variation is dynamic and results in device performance changes as the fabricated part ages. An example of dynamic process variation for PMOS devices is the negative bias temperature instability effect, which causes the V_{TH} of PMOS transistors to gradually increase over time. The typical method of handling dynamic variations is margining processor voltage or frequency.

Circuit techniques

A useful technique for reducing the impact of static process variations at the circuit level is substrate or body biasing—applying a nonzero voltage between a transistor's body and source.⁷ Depending on the voltage applied, V_{TH} either increases (reducing leakage) or decreases (increasing the processor's shipping frequency, f_{max}). The adaptive body bias (ABB) technique compensates for the effects of process variations on a part-by-part basis after fabrication. Each die receives a unique bias voltage that maximizes the die's frequency subject to the power constraints.⁵ Dies that are slow because of process variations can be forward-biased, increasing their f_{max} , while dies with high leakage can be reverse-biased to meet the power constraint.

Figure 2 shows native leakage versus f_{max} distribution for a set of dies in 150-nm CMOS technology, as well as the distribution after application of ABB. In this example, all dies must meet a minimum frequency specification (shown as a normalized frequency of 1) as well as a maximum leakage limit. The leakage limit is a function of frequency; low-frequency dies have less switching power and thus can tolerate greater leakage for the same total power constraint. ABB reduces the sigma of the frequency variation by six times and moves 30 percent of the dies into the highest frequency bin.

ABB is effective at compensating for die-to-die variations, but we cannot handle within-die variations using only a single bias value per die. Instead, we can divide the die into multiple regions, each of which can potentially receive a different body bias voltage after fabrication. Figure 2 shows that this within-die ABB technique

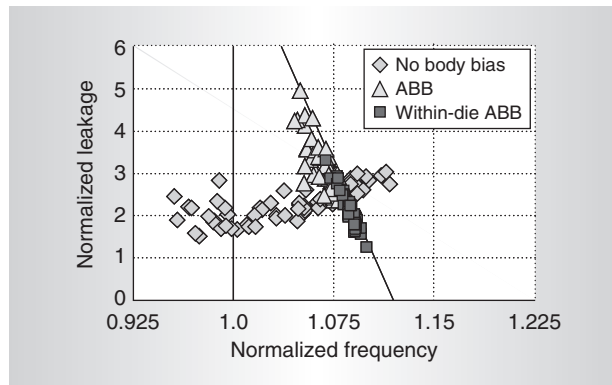


Figure 2. Leakage versus f_{max} distribution for dies without body bias, with ABB, and with within-die ABB.

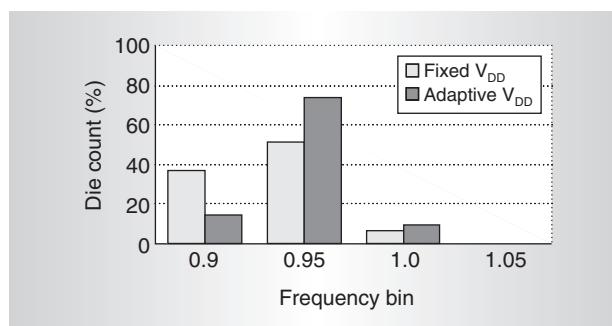


Figure 3. Binning improvement through adaptive V_{DD} .

further reduces frequency variation and moves most of the dies into the highest-frequency bin.

It is also possible to use supply voltage to reduce the impact of process variations.⁸ Both switching and leakage power have a superlinear dependence on supply voltage; therefore, the appropriate V_{DD} can modulate total power and frequency. Figure 3 demonstrates the binning improvement possible with the adaptive V_{DD} technique: The number of dies in the top two frequency bins improves by 45 percent over the standard fixed- V_{DD} case. Because switching power and leakage power respond differently to supply voltage and threshold voltage, combining ABB and adaptive V_{DD} is the most beneficial technique. As process variations increase, designers will likely include additional circuit features that can be tuned during the postsilicon phase to improve variation tolerance.

Microarchitectural techniques

Our previous research indicates that as the number of critical paths in the processor

increases, variability increases and maximum processor frequency suffers.³ This seems to point to using fewer critical paths per stage and deeper pipelining as the way to achieve process variability tolerance. However, consider the following: The impact of within-die random variations (as opposed to systematic variations) is increasing as technology scales. Moreover, random variations become responsible for a larger portion of f_{\max} loss as the number of pipeline stages increases (thus decreasing logic levels per stage). These facts point toward a microarchitecture with shorter pipelines for variability tolerance. These apparently contradictory conclusions underscore the need for careful, variability-aware design.

Process variations become more prominent for relatively large structures that occupy a large portion of the die. One such structure is obviously the cache. A recent example is the 24-Mbyte L3 cache in the Itanium 2 9000, a dual-core Intel processor. Of the 1.72 billion transistors in the design, 1.47 billion are reserved for the L3 cache, which occupies about 60 percent of the die area. To minimize process variation effects (and clock-related power consumption), the L3 cache has a self-timed asynchronous design style.⁹ Asynchronous design is challenging because few design automation tools exist for it, and validating the design is difficult. This challenge is manageable in the case of the Itanium 2, which is a relatively simple in-order design. In general, however, the cost of asynchronous design is high.

An alternative to asynchronous design for handling process variation is provisioning for nonuniform cache access times. The problem is as follows: Because of process variation, each cache block might return data in a different number of cycles. In each manufactured processor, the mapping of cache blocks to access time is different because each processor has a different process variation map. To measure process variations and get the mapping, current high-performance microprocessors employ test circuits.¹⁰ This special test circuitry records each cache block's access time, which is propagated to the microarchitecture in the form of a table of cache blocks and their access times. The microarchitecture uses this information to map frequently used cache lines to faster blocks.

Process variations in other processor blocks can also cause different entries in structures to have different operation latencies. For example, process variations in instruction-scheduling logic in contemporary microprocessors can cause different entries to wake up at different latencies. As a result of this phenomenon, the overall issue queue latency (and hence processor latency, if the instruction wake-up and select logic is the major critical path) is set to the selection latency of the issue queue's slowest entry. Similar operating frequency limitations apply to other structures, such as the register file, where the slowest register entry's access time limits the access time of the entire register file and the processor.

Designers can transform latency limitations due to random process variations into optimization opportunities by exposing these variations to the processor microarchitecture. Designing instruction-scheduling and register-renaming logic that are aware of variable component latencies is a way to leverage random variations in processor components.¹¹

Voltage variations

The demand for low power dissipation has translated into supply voltage scaling. V_{DD} is specified at two levels: Maximum V_{DD} is set as a reliability limit for a process, and minimum V_{DD} is set for the target performance. On the other hand, the variation of switching activity across the die and diverse logic cause uneven power dissipation. Voltage across an inductor is proportional to inductance and current change. This means that a big change in the current drawn by the processor will cause a voltage droop (ΔV_{DD}) across the inductance. Many factors affect inductance: traces on the motherboard, package routing, and chip pads. Packaging and platform technologies don't follow the scaling trends of CMOS processes, so voltage droops have become a significant percentage of V_{DD} .

Differences in transistor leakage current (due to process and temperature variations) as well as differences in active current demand across the die result in supply voltage variations. These voltage variations limit the processor's operating frequency and exacerbate temperature hot spots. For example, with a 10 percent V_{DD} variation, delay can vary as much as 20 percent.

Circuit techniques

Although adaptive V_{DD} reduces the impact of parameter variations and increases yield in high-frequency bins, it does not solve the voltage droop problem. One well-known technique for reducing ΔV_{DD} is to use on-die decoupling capacitors to supply the need for instantaneous charge.¹² An appropriate number of decoupling capacitors can reduce ΔV_{DD} by 50 percent, albeit with a cost in silicon area. Decoupling capacitors tend to increase gate oxide area, thus increasing gate oxide leakage in sub-90-nm technologies.

Microarchitectural techniques

One widely adopted technique for decreasing power consumption is clock gating. However, aggressive clock gating introduces voltage droop by generating large currents that can be a significant portion of the noise margin. This increases the required guardbanding. A possible solution is to gradually activate and deactivate gated blocks to limit voltage droop.¹³

Temperature variations

According to the 2004 ITRS, "A key form of dynamic variability is due to thermal effects during operation; this variation is on the time scale of billions of clock cycles and can strongly affect timing and noise phenomena." Heat management plays a vital role in the process of designing most electrical devices. Elevated chip operating temperatures impose constraints on the circuit's performance in several ways. Chip operating temperature has a direct impact on maximum reliable frequency and thus the IC's overall performance. Furthermore, higher operating temperatures restrict the permissible operating voltage and ambient temperature in the chip's environment.

Both spatial and temporal temperature variations affect a microprocessor's operation. Spatial variations occur when there is a temperature hot spot around a highly active unit (for example, a floating-point unit) adjacent to a region of relatively low temperature (for example, a cache). This temperature variation causes differences in transistor performance and leakage across the die and can also lead to functionality and reliability problems.

Temperature variations also occur with time, as the processor switches between idle and active periods. As the die's temperature rises

and falls as a function of the computing workload, power consumption and transistor performance change as well. A processor's cooling system is targeted to support a peak temperature, even though the processor spends most of the time running at far lower temperatures. Therefore, when the temperature is lower, the processor is running suboptimally.

Circuit techniques

Circuit designers attempt to mitigate hot spots by using low-threshold devices only where necessary, and limiting total switching capacitance by downsizing devices and interconnects. Chip designers have also relied on scaling down supply voltage to reduce power consumption. To counteract the negative effect of a lower V_{DD} on gate delay, they also scale down threshold voltage. However, lowering V_{TH} has a significant effect on leakage current. As leakage current increases, the die's temperature increases, further increasing leakage current. Therefore, there is a limit to how far threshold voltage can be scaled down.

Recently, researchers have proposed using low-temperature operation to optimize power-frequency trade-offs.¹⁴ Devices achieve low-temperature operation through refrigeration. (Refrigeration decreases the average temperature, making temperature variations a non-issue.) Researchers have tested refrigeration with different supply voltage selection, body bias, transistor sizing, and shorter channel length values in order to study the interaction between refrigeration with various circuit design parameters. When leakage power is substantial, refrigeration combined with a shorter channel length provides the best power-frequency trade-off.

Microarchitectural techniques

The most common temperature control technique, implemented in several commercial processors, is throttling.¹⁵ This technique decreases operating frequency when temperature exceeds a certain limit. Then, V_{DD} decreases, reducing power consumption and temperature. Once the processor cools, the process reverses. Voltage and frequency control can be implemented as an on-die microcontroller—effectively a separate simple core—as in the Itanium 2.¹⁶ This microcontroller has DSP-like capabilities, with embed-

ded firmware responsible for temperature control. The microcontroller reads the temperature from four on-die thermal sensors (two sensors located on each core's floating-point and integer units) to detect localized hot spots caused by unbalanced core workloads. Using those readings, the firmware implements a digital control system that maintains junction temperature below 90°C through closed-loop power control and a digital infinite impulse response (IIR) filter for system stability.

Usually, monolithic microarchitectures suffer from chronic power density and heat problems. In comparison, clustered architectures offer more possibilities to mitigate temperature-related variability, for two reasons: First, decoupling resources decreases power density and temperature. Second, some of the split resources can be clock- or V_{DD} -gated. Chaparro et al. applied these observations to the front end of a cluster.¹⁷ Applying the first observation, they split front-end structures such as the rename table and the reorder buffer. In comparison with a unified front end, splitting reduced the peak temperature of both structures with minimal slowdown. Applying the second observation, the researchers split the trace cache into two banks and then applied bank hopping, thus reducing average and peak trace cache temperatures.

Input variations

For synchronous systems, designers must know the maximum time it takes for the system to compute a function—also known as the worst-case delay. A circuit's depth, along with gate delays, determines its worst-case delay. A unit's actual delay depends on its inputs. Usually, a circuit finishes a computation before the worst-case delay elapses.¹⁸ For instance, the difference in lengths of the evaluation paths present in an ALU to calculate carry chains causes a large difference between average and worst cases. Nevertheless, circuits are designed to operate correctly in worst-case conditions.

Circuit techniques

To handle input variations, Lu proposes using extra hardware that mimics a logic function.¹⁹ The extra hardware provides correct outputs with a typical delay for a subset of inputs and raises a flag when the output is not

correct.¹⁹ This solution has a significant area penalty because both correct and approximate functions must be implemented on the die, and extra logic for error detection is also necessary. Suzuki, Jeong, and Roy use input variability to save power in a carry-select adder by exploiting delay differences according to input patterns.²⁰ Depending on the carry propagation length, this technique lowers supply voltage appropriately to finish the addition in the worst-case delay. Abdollahi, Fallah, and Pedram propose an approach that leverages the strong correlation between gate leakage current and input combinations by using the standby signal to shift in an input combination that minimizes leakage.²¹

Microarchitectural techniques

Some processor blocks' delays are especially sensitive to input values. For example, in the ALU, the adders' and the shifter's delays are expressed in terms of their input. The delay of a carry-propagate adder is on the order N , and that of a carry-lookahead adder is $\log(N)$, where N is the number of bits required for the addition. The adder critical path depends on the carry propagation; for most operations, the carry propagation chain is much shorter than the worst case. On the basis of this fact, Lu has proposed using fast adders with shorter carry propagation chains.¹⁹ The Pentium 4 also utilizes the adder's delay dependence on input size by using two staggered 16-bit adders instead of slower, full 32-bit adders.²² The 16-bit adders are clocked faster, thus feeding the low 16 bits of a dependent operation at the next fast clock cycle, while the other adder processes the higher-order 16 bits. For 32-bit additions that can be convoyed after each other, and additions that have operands of 16 bits or less, the effective addition latency is 1 fast clock cycle.

The shifter's delay also depends on data width. Frequently, effective shifting is restricted to a few bits. The Pentium 4 uses this property for a fast shifter (operating at twice the frequency) that operates on the operand's 8 low-order bits. The regular shifts occur on the slow port ALU and have a longer latency.²³

Combined microarchitectural techniques

Several microarchitecture approaches exploit the large gap between safe and opti-

mistic timing. The Razor processor saves power by eliminating voltage-related safety margins and introducing shadow flip-flops.²⁴ The shadow latches use a delayed clocking scheme to recover from timing errors introduced by the margin elimination. Another approach, timing-error avoidance (TEAtime), duplicates the critical paths and some of the safety margin delay.²⁵ This duplicated block becomes the feedback path in the frequency control system; the design then adapts to dynamic variability sources such as temperature by tracking fluctuations and changing frequency accordingly. Both the Razor and the TEAtime microarchitectures introduce extra hardware to detect possible errors; if a delay is longer than expected, the processor detects or corrects the error. However, errors can be expensive, so if the error rate is larger than a threshold, the processor can decrease clock frequency to maximize performance.

Marculescu and Talpes make the case that microarchitectures differ in their ability to mitigate variability.²⁶ They have developed a microarchitectural statistical-variability model based on the work of Bowman, Duvall, and Meindl,³ with the number of critical paths proportional to the microarchitectural blocks' total device count. Using this model, they show that a clustered, globally asynchronous, locally synchronous (GALS) microarchitecture is better in terms of variability than a monolithic synchronous microarchitecture. The intuition behind this conclusion is that clustered architectures have, overall, a smaller number of critical paths per clock domain (the classical monolithic architecture has only one domain) and therefore can be clocked faster. The net result is a performance gain if the increase in clock frequencies can offset the intercluster clock synchronization penalties.

Process and environmental variability will increase. Computer architects are already starting to develop techniques to mitigate or tolerate variability. To test those techniques and ideas, we must take an approach similar to that used to tackle the "power-wall" problem: We need to accurately model variability and incorporate variability as a design parameter at the architectural level. This implies that the variability models developed must be easily plugged into architectural simulators.

MICRO

References

1. S. Borkar et al., "Parameter Variations and Impact on Circuits and Microarchitecture," *Proc. Design Automation Conf. (DAC 03)*, IEEE Press, 2003, pp. 338-342.
2. S. Borkar, "Microarchitecture and Design Challenges for Gigascale Integration" (keynote speech), *Proc. 37th Int'l Symp. Microarchitecture (Micro 37)*, IEEE Press, 2004, p. 3.
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 Circuits*, vol. 37, no. 2, Feb. 2002, pp. 183-190.
4. S.G. Duvall, "Statistical Circuit Modeling and Optimization," *Proc. 5th Int'l Workshop Statistical Metrology*, IEEE Press, 2000, pp. 56-63.
5. J. Tschanz et al., "Adaptive Body Bias for Reducing Impacts of Die-to-Die and Within-Die Parameter Variations on Microprocessor Frequency and Leakage," *Proc. Int'l Solid-State Circuits Conf. (ISSCC 02)*, IEEE Press, 2002, vol. 1, pp. 422-478.
6. P. Friedberg, W. Cheung, and C.J. Spanos, "Spatial Variability of Critical Dimensions," *Proc. 22nd Int'l VLSI/ULSI Multilevel Interconnection Conf. (VMIC 05)*, Inst. Microelectronics Interconnection, 2005, pp. 539-546.
7. S. Narendra et al., "1.1V 1GHz Communications Router with On-Chip Body Bias in 150nm CMOS," *Proc. Int'l Solid-State Circuits Conf. (ISSCC 02)*, IEEE Press, 2002, vol. 1, pp. 270-271.
8. J. Tschanz et al., "Effectiveness of Adaptive Supply Voltage and Body Bias for Reducing Impact of Parameter Variations in Low Power and High Performance Microprocessors," *Proc. Symp. VLSI Circuits*, IEEE Press, 2002, pp. 310-311.
9. J. Wu et al., "The Asynchronous 24MB On-Chip Level-3 Cache for a Dual-Core Itanium-Family Processor," *Proc. Int'l Solid-State Circuits Conf. (ISSCC 05)*, IEEE Press, 2005, vol. 1, pp. 488-612.
10. J.C. Stinson and E.A. de la Iglesia, *Process Parameter Extraction*, US patent 6,533,535, Patent and Trademark Office, 2003.
11. S.M. Mueller, "On the Scheduling of Variable Latency Functional Units," *Proc. 11th*

- Ann. Symp. Parallel Algorithms and Architectures (SPAA 99)*, ACM Press, 1999, pp. 148-154.
12. T. Rahal-Arabi et al., "Design and Validation of the Pentium III and Pentium 4 Processors Power Delivery," *Proc. Symp. VLSI Circuits*, IEEE Press, 2002, pp. 220-223.
 13. M.D. Pant et al., "An Architectural Solution for the Inductive Noise Problem Due to Clock-Gating," *Proc. Int'l Symp. Low Power Electronics and Design (ISLPED 99)*, IEEE Press, 1999, pp. 255-257.
 14. A. Vassighi et al., "Design Optimizations for Microprocessors at Low Temperature," *Proc. 41st Design Automation Conf. (DAC 04)*, IEEE Press, 2004, pp. 2-5.
 15. D. Brooks and M. Martonosi, "Dynamic Thermal Management for High-Performance Microprocessors," *Proc. 7th Int'l Symp. High-Performance Computer Architecture (HPCA 7)*, IEEE Press, 2001, pp. 171-182.
 16. C. Poirier et al., "Power and Temperature Control on a 90nm Itanium-Family Processor," *Proc. Int'l Solid-State Circuits Conf. (ISSCC 05)*, IEEE Press, 2005, vol. 1, pp. 304-305.
 17. P. Chaparro et al., "Distributing the Front-end for Temperature Reduction," *Proc. 11th Int'l Symp. High-Performance Computer Architecture (HPCA 11)*, IEEE Press, 2005, pp. 61-70.
 18. G. Wolrich et al., "A High Performance Floating Point Coprocessor," *Proc. IEEE J. Solid-State Circuits*, vol. 19, no. 5, Oct. 1984, pp. 690-696.
 19. S.-L. Lu, "Speeding Up Processing with Approximation Circuits," *Computer*, vol. 37, no. 3, Mar. 2004, pp. 67-73.
 20. H. Suzuki, W. Jeong, and K. Roy, "Low-Power Carry-Select Adder Using Adaptive Supply Voltage Based on Input Vector Patterns," *Proc. Int'l Symp. Low Power Electronics and Design (ISLPED 04)*, IEEE Press, 2004, pp. 313-318.
 21. A. Abdollahi, F. Fallah, and M. Pedram, "Leakage Current Reduction in CMOS VLSI Circuits by Input Vector Control," *IEEE Trans. VLSI Systems*, vol. 12, no. 2, Feb. 2004, pp. 140-154.
 22. G. Hinton et al., "A 0.18- μm CMOS IA-32 Processor with a 4-GHz Integer Execution Unit," *IEEE J. Solid-State Circuits*, vol. 36, no. 11, Nov. 2001, pp. 1617-1627.
 23. D.J. Deiganes et al., "LVS Technology for the Intel Pentium 4 Processor 90nm Technology." *Intel Technology J.*, vol. 8, no. 1, Feb. 2004, pp. 43-53.
 24. D. Ernst et al., "Razor: A Low-Power Pipeline Based on Circuit-Level Timing Speculation," *Proc. 36th Ann. Int'l Symp. Microarchitecture (Micro 36)*, IEEE Press, 2003, pp. 7-18.
 25. A.K. Uht, *Achieving Typical Delays in Synchronous Systems via Timing Error Tolerance*, tech. report 032000-0100, Dept. of Electrical and Computer Eng., Univ. of Rhode Island, 2000.
 26. D. Marculescu and E. Talpes, "Variability and Energy Awareness: A Microarchitecture-Level Perspective," *Proc. 42nd Design Automation Conf. (DAC 05)*, IEEE Press, 2005, pp. 11-16.
- Osman S. Unsal** is a senior research associate at the Barcelona Supercomputing Center. His research interests include computer architecture, reliability, and programmer productivity. He performed the work described in this article while working for the Intel Barcelona Research Center. Unsal has a BS from Istanbul Technical University, an MS from Brown University, and a PhD from the University of Massachusetts, Amherst, all in electrical and computer engineering.
- James W. Tschanz** is a circuits researcher at Intel Laboratories, Hillsboro, Oregon. He is also an adjunct faculty member at the Oregon Graduate Institute in Beaverton, Oregon. His research interests include low-power digital circuits, design techniques, and methods for tolerating parameter variations. Tschanz has a BS in computer engineering and an MS in electrical engineering, both from the University of Illinois at Urbana-Champaign.
- Keith Bowman** is a senior researcher at Intel Circuit Research Labs, Hillsboro, Oregon. His current research focuses on the development of circuit design solutions to mitigate the impact of parameter variations on circuit performance and power. Bowman received a BS from North Carolina State University, and an MS and PhD from Georgia Institute of Technology, all in electrical engineering.
- Vivek De** is a senior principal engineer and a chief scientist in Intel's Circuits Research Lab,

Hillsboro, Oregon. He has been with Intel since 1996, leading long-term research in the area of low power circuit technology. De has a BTech from IIT Chennai, an MS from Duke University, and a PhD from Rensselaer Polytechnic Institute, all in electrical engineering.

Xavier Vera is a senior researcher at Intel Barcelona Research Center. His research interests include reliable and variation-aware microarchitectures. Vera has an MS in computer science from the Universitat Politècnica de Catalunya, Barcelona, and a PhD in Informatics from Mälardalens Högskola at Västerås, Sweden.

Antonio González is a professor in the Computer Architecture Department of the Universitat Politècnica de Catalunya (UPC). He is the founding director of the Intel-UPC Barcelona Research Center, which focuses on new microarchitecture paradigms and code generation techniques. González has an MS in informatics engineering and a PhD in com-

puter architecture from the Universitat Politècnica de Catalunya.

Oguz Ergin is an assistant professor in the Department of Computer Engineering of TOBB University of Economics and Technology, Ankara, Turkey. His research interests include computer architectures and VLSI design. He performed the work described in this article while working for the Intel Barcelona Research Center. Ergin has MS and PhD degrees in computer science from the State University of New York at Binghamton.

Direct questions and comments about this article to Osman S. Unsal, Barcelona Supercomputing Center, C/Jordi Girona, 29, Edificio Nexus II, 3^a planta, 08034 Barcelona Spain; osman.unsal@bsc.es.

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

IEEE/ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS

Stay on top of the exploding fields of computational biology and bioinformatics with the latest peer-reviewed research.

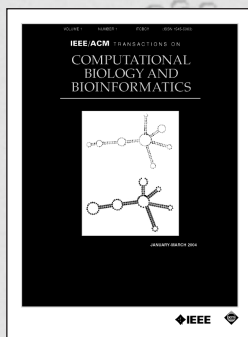
This new journal will emphasize the algorithmic, mathematical, statistical and computational methods that are central in bioinformatics and computational biology including...

- Computer programs in bioinformatics
- Biological databases
- Proteomics
- Functional genomics
- Computational problems in genetics

Publishing quarterly

Member rate: \$35

Institutional rate: \$385



Learn more about this new publication and become a subscriber today.

www.computer.org/tcbb



IEEE Computational Intelligence Society

Figure courtesy of Matthias Höchsmann, Björn Voss, and Robert Giegerich.