

Process Variation Characterization of Chip-Level Multiprocessors

Lide Zhang[†]

Lan S. Bai[‡]

Robert P. Dick[‡]

Li Shang^{*}

Russ Joseph[†]

[†]EECS Department
Northwestern University
Evanston, IL, USA

{lzh228@u,rjoseph@ece}.northwestern.edu

[‡]EECS Department
University of Michigan
Ann Arbor, MI, USA

{dickrp@eecs.,lanbai}@umich.edu

^{*}ECE Department
University of Colorado
Boulder, CO, USA

li.shang@colorado.edu

ABSTRACT

Within-die variation in leakage power consumption is substantial and increasing for chip-level multiprocessors (CMPs) and multiprocessor systems-on-chip. Dealing with this problem via conservative assumptions is sub-optimal. Instead, operating systems may adapt task assignment and power management decisions to the variable characteristics of cores, improving system-wide power consumption and performance. Researchers have proposed such adaptation techniques. However, they rely on knowledge of CMP process variation (PV) maps. These maps are not provided by processor vendors, providing them would impose additional cost during the testing process, and static maps would not permit adaptation to aging effects. Further progress on developing and validating PV aware control techniques for CMPs requires access to PV maps for real processors. We present an online technique to extract the PV maps of CMPs. Potentially automatic temperature measurements with built-in on-die sensors during the execution of characterization workloads are used to determine variation in leakage power consumption. The proposed technique is applied to real CMPs, and the resulting PV maps are used within a PV aware task assignment and scheduling algorithm.

Categories and Subject Descriptors: B.8 [Performance and Reliability]: Performance Analysis and Design Aids

General Terms: Design, Verification, Performance

Keywords: Process variation, characterization, software

1. Introduction and Motivation

Process variation (PV), the deviation of process parameters from their nominal values, can be divided into three categories: die-to-die variation, within-die variation, and wafer-to-wafer variation. Ongoing technology scaling has a tendency to increase PV [1]. In this paper, we will focus on spatially-correlated within-die variation and die-to-die variation. More specifically, we focus on the within-die variation among CMP and multiprocessor system-on-chip cores.

PV has received substantial attention from researchers [2, 3], who have developed techniques that adapt to the characteristics of individual cores to achieve better power consumption and performance [2, 3]. Such techniques require a PV

This work was supported in part by the SRC under awards 2007-HJ-1593 and 2007-TJ-1589 and in part by the NSF under awards CCF-0702761, CNS-0347941, CNS-0720820, and CNS-0720691.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2009, July 26 - 31, 2009, San Francisco, California, USA.

Copyright 2009 ACM ACM 978-1-60558-497-3 -6/08/0006 ...\$10.00.

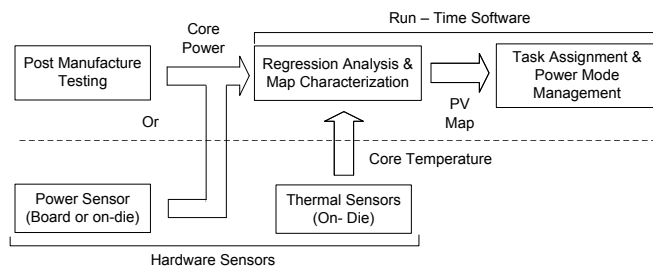


Figure 1: Generating a PV map and applying it in on-line power consumption and performance optimizations.

map as input. However, these maps are not provided by processor vendors, providing them would impose additional cost during the testing process, and in-factory characterization results might be invalidated by aging effects. Furthermore, there are not published post-testing methods for researchers or end users to derive PV maps. Borkar et al. discuss power, voltage, and temperature variations and their impact on circuits and microarchitecture [1]. They also present measurements of the leakage power consumptions and maximum frequencies of numerous microprocessors. However, these data are not provided with individual processors and their work suggests no way for end users or OSs to derive them.

We present an automatic on-line software-based technique to characterize the threshold voltages and leakage power consumptions of CMP cores, i.e., derive PV maps. These parameters are used to optimize CMP power consumption and performance. Our work makes the following main contributions: (1) it is the first on-line software-based technique for characterizing the variation in leakage and threshold voltage among CMP cores using built-in sensors; (2) we present a technique that predicts the power and thermal profiles of a given workload when run on a CMP with a particular PV map; and (3) we formulate and solve the time-constrained CMP task assignment and power management mode selection problem.

The proposed PV characterization technique derives a PV map based on temperature differences among cores. Due to PV, core leakage power coefficients (and also leakage power) may differ, resulting in temperature variation given identical workloads. We extract the PV map by performing multi-variable regression analysis based on measured results, a thermal model, and a leakage model. The temperature readings come from on-die sensors present in commercially available processors. Figure 1 outlines our approach for producing a PV map, and explains how this map can be used to perform run-time optimization for task assignment and power mode selection. The characterization technique can be implemented within the OS, firmware, or hardware. It imposes low computational overhead, eliminates the need to measure the PV map during testing, and can be used to adapt to aging effects.

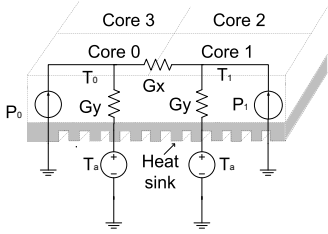


Figure 2: CMP chip-package thermal model.

2. Thermal and Leakage Models

Our approach to generating PV maps is based on two physical models: a compact CMP thermal model and a leakage power model, which are described in this section. These models will be used to derive and validate the PV map as explained in Section 3.

2.1 Thermal Model for Multi-Core Processors

To analyze heat flow, we use a discretized Fourier thermal model [4], in which heat flow is analogous to electrical current and temperature is analogous to voltage. As shown in Figure 2, multicore processors can be divided into several blocks, each representing a core. T_a is the ambient temperature of the processor. T_0 and T_1 are the temperatures for Core 0 and Core 1. G_y is the thermal conductance from the core to the ambient and G_x is the thermal conductance between cores. We assume a symmetric cooling solution for which all core-to-ambient thermal conductances are identical. P_0 and P_1 represent the power dissipations of the cores and hence correspond to the heat generated in the core active layers.

Heat flow can be modeled as follows:

$$\mathbf{C}dT(t)/dt + \mathbf{G}T(t) = \mathbf{P}u(t), \quad (1)$$

where N is the number of thermal elements (with at least one element per core), \mathbf{C} is a diagonal $N \times N$ matrix containing the core heat capacities as the diagonal elements, $T(t)$ is an N -element vector, each element of which represents the temperature of a core as a function of time, and \mathbf{G} is an $N \times N$ thermal conductance matrix where G_{ij} represents the thermal conductance between core i and core j . $u(t)$ is the $t = 0$ unit step function. Section 3.2 explains one use of heat capacity and dynamic thermal effects to assist characterization.

2.2 Leakage Power Consumption Model

Leakage power consumption is a first-order concern in CMP management, and is sensitive to the effects of inter-core PV. Subthreshold leakage power consumption is presently the major component of leakage power, and the introduction of high- κ gate dielectrics means this is likely to remain true in the near future. According to the BSIM model [5], transistor-level subthreshold leakage power can be approximated as follows:

$$P^{leak} = \eta \cdot T^2 e^{\frac{-q(V_{th}(T) + V_{off} - V_{GS})}{nkT}}, \quad (2)$$

where n ($=1.5$) is the subthreshold swing coefficient for the transistor [6], $V_{off} = 0.08$ V is offset voltage, T is the temperature, and η is a technology-dependent parameter.

3. Process Variation Map Characterization

In this section, we formulate the PV map characterization problem based on the thermal and leakage power models described in Section 2. Our objective is to determine the threshold voltages and leakage power consumptions of individual cores from the temperature readings of all cores. For this purpose, we produce a number of different thermal profiles by using workloads with carefully controlled levels of processor utilization. For each thermal profile, the dynamic power profile is held uniform among cores. Leakage power is then estimated by doing regression analysis with the dynamic power profiles as independent variables and the thermal profile as dependent variables. The thermal model (Equation 4) and leakage power model (Equation 2) are used as regression equations.

This regression process requires several parameters including the nominal leakage power coefficient, the dynamic power of each workload, and thermal conductance characteristics of the processor. We explain our characterization technique by first developing regression model with the assumption that these parameters are known, then describe how the parameters may be obtained.

3.1 Regression Analysis

Assuming knowledge of dynamic power, the nominal leakage power coefficient η , and the thermal conductance matrix, the primary input of our regression model is a series of thermal profiles. These thermal profiles are readings from temperature sensors when all the cores are stressed with different workloads with known CPU usage. T_{ij} is the temperature of the i th processor with workload j . Given knowledge of the thermal conductance matrix \mathbf{G} , power and steady-state temperature are related as follows:

$$\mathbf{G} \times T_j = P_j + T_a \times G_y, \quad (3)$$

where T_j is the j th column of the T matrix, P_j is the power profile of all the cores with the j th workload, and T_a is the ambient temperature as shown in Figure 2. The power consumption is composed of dynamic power and leakage power, the equation for which follows:

$$P_j^{leak} = \mathbf{G} \times T_j - P_j^{dyn} - T_a \times G_y. \quad (4)$$

Similarly, P_j^{leak} and P_j^{dyn} are the leakage power profile and dynamic power profile of the CMP with the j th workload. ξ is an $N \times W$ matrix, in which P_j^{leak} is the j th column. W is the set of workloads available for use. Therefore, ξ_{ij} is the leakage power of the i th core at temperature T_{ij} , which is approximated based on Fourier thermal model.

Given the nominal leakage power coefficient η , we can also estimate the leakage power profile using the model shown in Equation 2. This leakage power profile is subject to the threshold voltage ($V_0^{th}, V_1^{th}, \dots, V_N^{th}$). Γ_{ij} is the leakage power for the i th core with T_{ij} using the leakage model. For a set of data (T_j, P_j^{dyn}), the threshold voltages can be estimated by minimizing the sum of squared errors: $\sum_{j \in W} |\xi_j - \Gamma_j|^2$.

3.2 Parameter Derivation

The proposed regression approach constructs a leakage variation map for a CMP based on knowledge of the temperature, CMP total power consumption, and the thermal conductances of the system. In practice, these quantities are not readily available. In this subsection, we describe methods for determining the vertical thermal conductance G_Y , the lateral thermal conductance G_X , nominal leakage, and core dynamic power consumption P_j^{dyn} . The nominal leakage power is the average leakage power of all chips that the manufacturer measures during testing process [1]. The dynamic power of each workload is determined by subtracting the nominal leakage power from the total power. As these values may be associated with a particular processor model, they can potentially be stored in non-volatile memory or fuses by the chip manufacturer. However, this is not now common, so we have developed a characterization technique.

Existing and upcoming processor technologies for power management support on-die power measurement. Some, such as Intel's Foxtan technology, support independent power sensors on each processor core [7]. The required power measurements could instead be provided by motherboards. Even if the total power consumption of each workload on an individual core is known, separating leakage and dynamic power is challenging. To do this, we use workloads designed to control the utilization, and therefore power consumptions, of cores. We also take advantage of the dynamic thermal effect resulting from heat capacities (\mathbf{C}) in Equation 1 to set dynamic power and temperature independently.

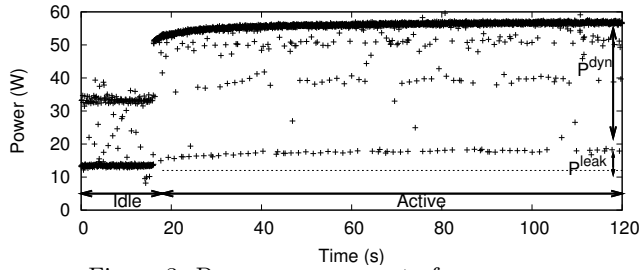


Figure 3: Power measurement of processor.

Figure 3 shows power measurements when all cores transition between idle and active workloads. Note that due to the heat capacity, the temperatures and therefore leakage power consumptions of all cores increase gradually. By stressing the processor to alternate CPU utilization between 0% and 100%, we isolate the impact of the workload on dynamic (\mathbf{P}^{dyn}) and change in leakage power consumption (\mathbf{P}^{leak}). With the knowledge of the temperature-dependent component (change in leakage power), we use regression on Equation 2 to determine the values of η and the temperature-independent component of the power consumption associated with the measured temperature-dependent component of the power consumption, temperature, and the process-dependent parameters.

In addition to the nominal leakage power and dynamic power, the vertical and lateral thermal conductances are also required. Starting from the processor power consumptions, we get the relationship between total power consumption, thermal profile, and the vertical thermal conductance from Equation 3. To give an example, we apply Equation 3 to the four-core processor shown in Figure 2 to derive the following equations. The same technique can be applied to general multi-core processors. By adding the four equations in the extension of Equation 3, we have

$$G_y \times \sum_{i=1}^4 (T_{ij}) - 4 \times T_a \times G_y = \sum_{i=1}^4 (P_{ij}) = P_j^{total}, \quad (5)$$

where P_j^{total} is the total power, measured using power sensors. Hence, by taking advantage of the linear relationship between $\sum_{i=1}^4 (T_{ij})$ and P_j^{total} , we set parameters to minimize the sum of squared errors. The vertical thermal conductance and ambient temperature can thus be derived. Similarly, we get the lateral thermal conductance by doing linear regression on thermal and power profiles. To get the relationship between the lateral thermal conductance, thermal profile, and power profile, we subtract the $i + 1$ th equation from the i th in the extension of Equation 3. For example, subtracting the 2nd equation from the 1st yields the following equation:

$$(3G_x + G_y)T_{0j} - (3G_x + G_y)T_{1j} + G_x T_{2j} - G_x T_{3j} = P_{0j} - P_{1j}.$$

Hence, G_x is set to minimize the sum of squared errors in this relationship between thermal profile and its corresponding power profile. This regression is easy if per-core power sensors are available. However, if only a single sensor or single power supply network is available for the entire CMP, it still remains a challenge to isolate the power consumptions of individual cores. To achieve this, we use the nominal power coefficient and thermal profile to estimate the leakage power of each core. By adding the dynamic power profile and the leakage power profile, we approximate the total power profile yielding an initial G_x . This dynamic power profile is obtained by measuring the difference between the power consumption immediately after the power transition in Figure 3. As a result the heat capacity \mathbf{C} in Equation 1, the temperature, and therefore leakage power consumption, change only gradually after the power consumption change, allowing the isolation of dynamic and leakage power consumption. This is not yet

accurate because we assume the leakage power coefficients are the same for each core. To solve this, we use the initial G_x as input in the multi-variable regression and get the initial leakage power map. The initial leakage power map is then used in the estimation of G_x , iterating until G_x convergence.

4. Process Variation Aware Task Assignment and Power Management Mode Selection

In this section, the time-constrained CMP task assignment and power management mode selection problem is used to demonstrate the importance of knowing a PV map when attempting to optimize system characteristics such as power consumption and performance.

Problem statement: Given a deadline for a set of tasks, determine the assignment of tasks to cores and power management configurations to tasks. The objective is to complete all tasks within a time constraint and with minimal energy consumption. Tasks are independent. The core configuration parameters, such as supply voltage, frequency, and active cache size, are controlled independently for each core. Many of these features are supported on existing multi-core processors, e.g., the AMD’s Quad-Core Opteron.

Let T be the set of tasks, C be the set of cores, P be the set of power management configurations, and B be the time constraint. Let E_{ijk} be the energy consumption of running task i on core j with configuration k . Let δ_{ijk} be the execution time of task i on core j with configuration k . Let binary variable A_{ijk} indicate whether task i is assigned to core j with configuration k . The task assignment and core configuration problem is formulated as an integer linear program (ILP).

$$\begin{aligned} & \text{minimize} && \sum_{i \in T, j \in C, k \in P} E_{ijk} \times A_{ijk} \\ & \text{subject to} && \max_{j \in C} \sum_{i \in T, k \in P} \delta_{ijk} \times A_{ijk} \leq B \\ & && \forall i \in T \sum_{j \in C, k \in P} A_{ijk} = 1 \end{aligned}$$

To evaluate the impact of considering PV when solving this problem, we compare optimal solutions for formulations that consider, and neglect, PV. From the perspective of problem input, the PV-unaware approach bases its decisions on the mean energy–performance relationships, while the PV-aware approach considers the different energy–performance tradeoffs of different cores. Note that we are not proposing to use the optimal ILP solver within the OS.

5. Experimental Results and Validation

This section explains our experimental setup and results, the approach used for preliminary validation of the proposed technique, and indicates the impact of considering PV during task assignment and power state control.

5.1 Experimental Setup

To validate our idea, we use Intel Core 2 Duo E6420 processors and a Shuttle SD32G2B motherboard as our experimental platform. This processor is equipped with one thermal diode per core, which were calibrated by the manufacturer to within 1°C error [8]. To reduce the impact of asymmetric cooling, we replace the original fan and heat sink with a symmetric cooling solution. Intel E6420 processors lack on-die power sensors. We therefore use a current clamp to measure CPU power consumption. Note that this is unnecessary for processors or motherboards with built-in power sensors (see Section 3.2). The voltage regulator power offset does not influence the result because it affects both ambient temperature and dynamic power estimates. The workloads used to stress the processor are controlled CPU utilization programs. By adjusting the duty cycle of the workload, we indirectly control processor temperature.

Table 1: Fitted Parameters of Processor

	Processor 0		Processor 1	
	Core 0	Core 1	Core 0	Core 1
V_{th}	0.2176 V	0.2233 V	0.2672 V	0.2404 V
Thermal cond.	Vert. G_y	Lat. G_x	Vert. G_y	Lat. G_x
	0.54 W/K	0.28 W/K	0.37 W/K	0.18 W/K
Leakage const. η	0.0478 V · (eV/K) ²		0.0272 V · (eV/K) ²	

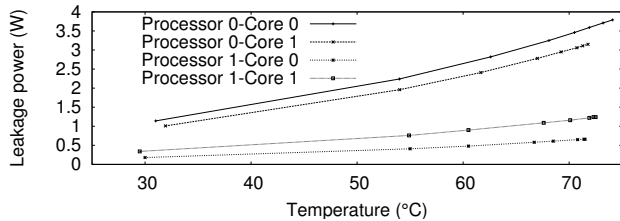


Figure 4: Leakage power of four cores.

5.2 Results and Analysis

Figure 4 illustrates the leakage power map characterized through regression. Consider processor 0. The largest difference between leakage power for two cores is 0.6 W, 20.28% of the leakage power of core 1. The total leakage power is 6.93 W at 72.9°C, 12.57% of the total power. When the processor is idle, the leakage power is responsible for up to 16.71% of the total power. All the parameters including the threshold voltage and thermal conductances are shown in Table 1. The estimated variation in threshold voltage is 2.62%.

5.3 Validation of Characterization Technique

In the interest of validation, we use another set of workloads with different dynamic power consumptions and predict temperatures based on the extracted PV map (see Section 3). We start from dynamic power consumption and iterate until Equation 4 converges. The predicted temperatures are compared with the measured values for four cores. The average difference between the predicted and measured temperatures is 1.1°C and the maximum is 2.14°C. These results are a step toward validating the proposed technique. This prediction technique is also a potential use of the extracted PV maps.

5.4 Task Assignment and Power Management

We used the M5 instruction set architecture simulator running a number of SPEC CPU2000 benchmarks to generate input energy–delay relationships for different tasks. We consider the following configuration parameters: instruction cache (one-way or two-way), L1 data cache (one-way or two-way), L2 data cache (one-way to eight-way), frequency (1.0, 0.95, 0.9, 0.85 × maximum frequency). The M5 simulation results were used as the input for different tasks on cores with mean PV parameters, i.e., without any PV.

After obtaining the energy–delay relationships with mean variation parameters, we applied the variation to each curve to emulate cores with PV. Based on the switching and leakage power scaling trends reported by Keshavarzi [9], the leakage power is approximately half of the total power for a 45 nm process. We use this leakage power proportion. We obtained the PV map (Figure 4) by using the characterization technique described in Section 3 on two Intel Core 2 Duo processors. The frequency variation is derived from the leakage power variation according to a function fitted to the frequency–leakage variation plot (see Borkar et al.’s Figure 1 [1]). In addition to solving the instance based on our characterized PV maps, we also solved the problem for a number of synthetic PV maps using frequency and leakage power variation distributions based on large-scale measurements [1]. The problem instances were solved with the CPLEX ILP solver [10]. Each problem instance is an assignment of 12 tasks to four cores with 128 power management modes.

The solutions yielded by the PV-unaware formulation have two disadvantages compared with those from the PV-aware

Table 2: Penalty of PV-Unaware Formulation

Time constraint (ms)	4	5	6	7	8
Energy overhead (%)	2.2	5.3	7.2	8.8	10.7
Deadline violation (%)	34.0	21.5	20.5	8.5	15.5

formulation: (1) they have higher energy consumption and (2) they sometimes violate their deadlines. The second row of Table 2 shows the energy consumption overhead of the PV-unaware formulation relative to the PV-aware formulation with various time constraints for the synthetic PV maps based on the distributions of Intel measurements [1]. They are computed by averaging the energy overheads of the 20 simulated chips, neglecting infeasible solutions. As the time constraint is relaxed, the PV-aware technique is increasingly able to adapt to CMP characteristics. On average, the PV-unaware formulation imposes a 6.8% energy penalty. For the measured PV map, the PV-unaware formulation imposes a 26.0% energy overhead. The third row of Table 2 shows the probability of PV-unaware solutions violating time constraints. Averaging over all task sets and constraints, the PV-unaware technique violates the deadline for 20% of the problems while the PV-aware technique meets all deadlines; the use of the PV-unaware formulation would require tightening timing constraints, thereby increasing energy consumption or making the problem unsolvable.

In summary, considering PV during optimization of task assignment and power mode selection can substantially improve system quality. However, for such an approach to be used, the PV maps of individual processors are required, motivating the characterization technique described in Section 3.

6. Conclusions

This paper presented a technique that uses built-in sensors to derive the leakage power PV maps for processor cores in a CMP or multiprocessor system-on-chip. This technique makes use of the interdependence between temperature and leakage power to isolate dynamic and leakage power consumption for different regions of the CMP, and uses dynamic thermal effects to independently control the temperature and dynamic power consumption of the CMP being characterized. Our characterization results are validated by comparing measured and predicted temperatures for CMP workloads. We also present a novel PV aware technique for controlling CMP task assignment and power management state in order to optimize energy consumption under hard deadlines.

7. References

- [1] S. Borkar, et al., “Parameter variation and impact on circuits and microarchitecture,” in *Proc. Design Automation Conf.*, June 2003, pp. 338–342.
- [2] R. Teodorescu and J. Torrellas, “Variation-aware application scheduling and power management for chip multiprocessors,” in *Proc. Int. Symp. Computer Architecture*, June 2008.
- [3] P. Ndaï, et al., “Within-die variation-aware scheduling in superscalar processors for improved throughput,” in *IEEE Trans. Computers*, vol. 57, no. 7, July 2008.
- [4] J. Fourier, *The Analytical Theory of Heat*, 1822.
- [5] “BSIM4,” <http://www-device.eecs.berkeley.edu/~bsim4/bsim4.html>.
- [6] Z. Chen, et al., “Estimation of standby leakage power in CMOS circuits considering accurate modeling of transistor stacks,” in *Proc. Int. Symp. Low Power Electronics & Design*, Aug. 1998, pp. 239–244.
- [7] C. Poirier, et al., “Power and temperature control on a 90 nm Itanium-family processor,” in *Proc. Int. Solid-State Circuits Conf.*, Feb. 2005, pp. 304–305.
- [8] “Core 2 Quad and Duo temperature guide,” <http://www.tomshardware.com/forum/221745-29-core-quad-temperature-guide>.
- [9] A. Keshavarzi, “Technology scaling and low-power circuit design,” in *The VLSI Handbook*, W.-K. Chen, Ed. CRC Press, 2007, ch. 21, p. 21.12.
- [10] “CPLEX,” ILOG, Inc., <http://www.ilog.com/products/cplex>.