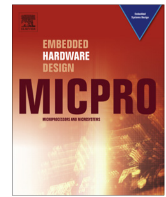
Microprocessors and Microsystems 38 (2014) 22–30

Contents lists available at ScienceDirect

Microprocessors and Microsystems

journal homepage: www.elsevier.com/locate/micpro



Design configuration selection for hard-error reliable processors via statistical rules

Ying Zhang [a,*], Lide Duan [b], Bin Li [c], Lu Peng [a], Xin Fu [d]

[a] Division of Electrical & Computer Engineering, School of Electrical Engineering and Computer Science, Louisiana State University, Baton Rouge, LA 70803, United States
[b] Advanced Micro Devices, Inc., Austin, TX 78735, United States
[c] Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA 70802, United States
[d] Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, United States

ARTICLE INFO

Article history:
Available online 11 November 2013

Keywords:
Hard-error reliability
Modeling
Process variation

ABSTRACT

Lifetime reliability is becoming a first-order concern in processor manufacturing in addition to conventional design goals including performance, power consumption and thermal features since semiconductor technology enters the deep submicron era. This requires computer architects to carefully examine each design option and evaluate its reliability, in order to prolong the lifetime of the target processor. However, the complex wear-out mechanisms which cause processor failure and their interactions with varying microarchitectural configurations are still far from well understood, making the early optimization for chip reliability a challenging problem. To address this issue, we investigate the relationship between processor reliability and the design configuration by exploring a large processor design space in this paper. We employ a rule search strategy to generate a set of rules to identify the optimal configurations for reliability and its tradeoff with other design goals.

In addition to the wear-out effects, the ever-shrinking feature size of modern transistors makes process variation a significant issue in the chip fabrication. Process variation results in unexpected distributions of key design parameters, thus remarkably impacting important features of the target processor. Therefore, we also extend our investigation to identify the optimal configurations in the presence of process variation.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The unceasing downscaling of the semiconductor technology makes hard-error reliability a first-order concern in modern high-performance processor design. Due to the substantial transistors integrated on a chip, the processor power density and runtime temperature keep rising, which tend to largely impair the processor lifetime reliability. For instance, Negative Bias Temperature Instability (NBTI), caused by the continuous increase in processor threshold voltage, becomes a key reliability issue when the manufacturing technology reaches 90 nm. The situation is further exacerbated when the on-die temperature increases.

Although several techniques including power balancing and hotspot elimination have been proposed to improve the lifetime reliability for a given processor [8], designing hard-error resilient processors at the pre-silicon stage remains an open topic. Considering that processors with different microarchitectural configurations are prone to show distinct behavior such as various power dissipations, it is reasonable to infer that microarchitectural configurations selections will heavily impact the hard-error reliability of a processor [10,28]. However, traditional design space explorations mainly concentrate on performance, power, and thermal features of the target processor [17], leaving the relationship between hard-error reliability and the underlying design option far from obvious. This indicates that selecting the most reliable configurations at the early stage of processor manufacturing is of great importance for computer architects.

In this paper, we aim to address this issue by exploring a vast design space. We model four important and well-studied failure mechanisms including electromigration (EM), stress migration (SM), thermal cycling (TC), and the aforementioned NBTI. The processor failure rate (FIT, or failures in 10^9 h) is used to interpret the reliability, where a smaller FIT value indicates a longer MTTF (mean time to failure), i.e., higher reliability. To correlate configuration parameters to the design objective, we employ an advanced statistical technique called Patient Rule Induction Method (PRIM) to facilitate our work. The PRIM model can extract a few simple "rules" or conditions by which the processor satisfies a preset design goal for a certain response. Specifically in our work, the optimal design configurations for improving hard-error reliability and

* Corresponding author.
E-mail addresses: yzhan29@lsu.edu (Y. Zhang), lide.duan@amd.com (L. Duan), bli@lsu.edu (B. Li), lpeng@lsu.edu (L. Peng), xinfu@ittc.ku.edu (X. Fu).

0141-9331/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.micpro.2013.11.002

its tradeoff with other design goals are generated by training PRIM models.

On the other hand, process variation (PV) is gradually becoming a significant issue during processor manufacturing in the deep submicron era. Process variation manifests itself in a variability of transistor process parameters (e.g., threshold voltage, or V_{th}) from their design specifications, which is caused by the difficulty in consistently controlling the transistor manufacturing in small dimensions. Process variation can impose significant impact on the performance and power consumption of the target processor. For example, the varying V_{th} among integrated transistors results in substantial delay variability in critical paths, thus greatly degrading the processor frequency and hence the overall performance as it is determined by the slowest critical path across the chip. In one word, process variation is becoming a major challenge in the design of future high-performance microprocessors.

Moreover, the interaction between varying microarchitectural configurations and process variation introduces more complexities to the designers. For instance, the design options can impact the number and distribution of critical paths, so processors with various architecture designs exhibit different capability in tolerating the PV effect. Therefore, in the early design stages, it is essential for designers to consider the PV effect when selecting the optimal design parameters. Taking this into consideration, we apply the PRIM based approach to circumstances where process variation is present, in order to seek the optimal configurations in this scenario.

In summary, the main contributions of this paper are as follows:

- *Optimal design configurations for hard-error reliability*: We demonstrate that different design choices will result in distinct reliability behaviors. Based on the investigation, we identify the most reliable microarchitectural configurations.
- *Optimal configurations for other metrics*: We show that different metrics including performance, power, temperature, and hard-error reliability prefer disparate design configurations. Therefore, we also investigate optimal configurations for other metrics.
- *Balancing reliability, power, and performance*: We generate rules to filter out configurations that yield optimal tradeoffs among multiple metrics.
- *Design stage optimization under process variation*: We take the widely-acknowledged process variation effect into consideration, and aim to identify the most promising microarchitectural configurations in the deep submicron era when PV is present.

2. Related work

The hard-error defects that degrade processor reliability [9,22] are first noticed in the circuit design process. Srinivasan et al. propose a reliability-aware microprocessor (RAMP) model [24] to assist overcoming the unnecessary high cost in manufacturing an overestimated processor. The authors further extend their study in [25] and demonstrate the impact of fabrication technology on hard-error reliability. At the architectural level, Bower et al. [5] introduce an online diagnosis technique to detect the hard faults. On the other hand, since the processor hard-error reliability is highly related to its runtime power and temperature, many techniques are proposed to improve the processor reliability by removing the on-die hotspot at runtime. Coskun et al. [8] discuss the impact of different job scheduling algorithms on processor lifetime. They demonstrate that the processor lifetime can be effectively prolonged when smart temperature-aware scheduling mechanisms are engaged. Hsu and Feng [14,15] concentrate on the heating and consequent reliability issues in high performance computing (HPC) systems and propose power management techniques by using dynamic voltage and frequency scaling, in order to reduce the failure rates.

The process variation phenomenon has gained much attention in recent years due to its increasing significance in modern chip fabrication. Borkar et al. [4] investigate the impact of parameter variations on circuits and microarchitecture and propose the Body Bias Control Techniques to reduce the negative impact by PV. Teodorescu et al. [26] propose using dynamic fine-grain body biasing to mitigate the process variation effect. The analysis on leakage power in the face of process variation is given by Chang and Sapatnekar in [7]. Garg et al. [13] develop a model to describe the system performance under PV.

Our work deviates from the above studies in that we investigate the relationship between microarchitectural configurations and hard-error reliability and extract the promising configurations at early design stages.

3. Methodology

3.1. Patient Rule Induction Method (PRIM)

PRIM is an advanced statistical model [12], the objective of which is to find a region in the input space (composed of configuration parameters in this work) that gives relatively low values for the output response, e.g. the FIT value. The selected region (or "box") is described in an interpretable form involving a set of "rules" depicted as $B = \cap_{j=1}^{p}(x_j \in s_j)$, where x_j represents the jth input variable and s_j is a subset of all possible values of the jth variable. In other words, the identified region B is the intersection of p subsets, each of which is from one of the p input variables.

Fig. 1 illustrates the construction of the "optimal" region, which is composed of two phases: (1) patient successive top-down peeling process and (2) bottom-up recursive pasting process. The top-down peeling starts from the entire space (box B) that covers all the data. In each iteration, a small subbox b within the current box B is removed; we calculate the output mean for the elements remaining in $B - b = \{x \in B \;\&\; x \notin b\}$, performing this operation in each dimension (i.e., try removing a different subbox from each input variable); finally we choose the one which yields the smallest output mean value for the next box $B - b$. This procedure is applied iteratively until the proportion of the data points remaining in the current box (termed the support) is below a preset threshold β. Note that for a categorical variable, an eligible subbox b contains only one element of the possible values of the variable in the current box B. To give an example, let us assume that three values for the L1 data cache size, which is an architectural parameter in the design space, exist in the current box, i.e., s_1 = {8 KB, 16 KB, 32 KB}.

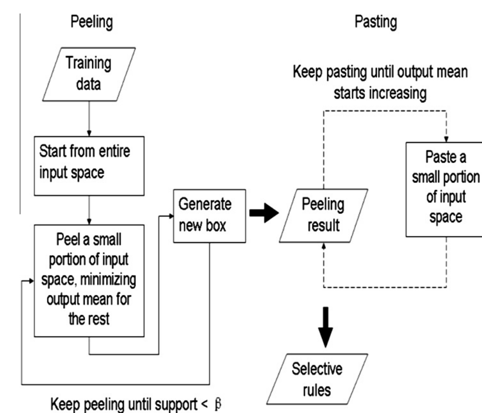


Fig. 1. PRIM training procedure, including peeling and pasting.

Download English Version:

https://daneshyari.com/en/article/462773

Download Persian Version:

https://daneshyari.com/article/462773

Daneshyari.com