Improving Multi-core Processor Energy Efficiency and Lifetime by Embracing Variability and Wearout

Mehmet Basoglu and Mattan Erez

Department of Electrical and Computer Engineering
The University of Texas at Austin
e-mail: {mbasoglu,mattan.erez}@mail.utexas.edu

Abstract – Negative Bias Temperature Instability (NBTI) is a reliability challenge due to circuit degradation, which has only received attention after its appearance in the last decade. The problem leads to an increase in the threshold voltage and a decrease in the drive current of p-channel transistors when they are stressed over extended periods of time. In this paper, we use an advanced model for estimating NBTI degradation in order to minimize its impact on a multi-core processor's lifespan through automatic workload management and consequently save power at the same time. We show that our approach saves approximately 9.25% power and increases life expectancy by 3 years while avoiding any negative performance impact.

Index Terms — degradation, delay, dynamic voltage and frequency scaling (DVFS), multi-core processor, negative bias temperature instability (NBTI), reaction-diffusion (R-D) model, threshold voltage

I. Introduction

Negative Bias Temperature Instability (NBTI) is a key reliability issue of immediate concern in p-channel MOSFET devices stressed with negative gate voltages as it deteriorates circuits and leads to chip failure. NBTI manifests itself as an increase in the threshold voltage and consequent decrease in the drain current and transconductance. The increase in threshold slows the transistor down, which eventually may cause a timing violation in the critical path. Currently, processors are designed by keeping built-in timing margins to compensate for the worst threshold degradation in order to meet the designated frequency over the entire lifetime.

The timing margins introduced at design time squander performance and energy efficiency during the early life of the processor as the supplied voltage is higher than that is needed to operate the system. Consequently, to cover for the margin in the early life of the processor, power is consumed unnecessarily. One can exploit these margins at run-time to operate the processor either at a faster frequency than the specifications to increase performance or with a lower-than-nominal supply voltage to save power. In this paper, we chose to implement the latter as it in return also slows down the temperature-dependent circuit degradation rate, thus improving the life expectancy of the chip. We describe a technique that can reduce the energy consumption of the processor while at the same time improving its expected lifetime, both without impacting its specified performance. Our method relies on a detailed analytical model for NBTI-induced degradation, which we use to guide optimizations and, for the first time, evaluate long-term potential and impact of NBTI-aware process scheduling.

We believe that this research can impact all general-purpose multi-core processors and applications where power savings and life expectancy is a major concern while having no negative impact on system performance. At present, DVFS attempts to achieve some degree of such power savings. However, we show that for every DVFS voltage selected by the system, we can choose a lower voltage, which we call NBTI voltage, to take advantage of the built-in margins and save more power. In return, we stress the transistors on the system less, and therefore increase the lifetime without any performance compromises.

Another aspect considered for this research is that different cores on multi-core processors witness process variations and that these cores start with different threshold voltages. Consequently, some cores on the same chip start with shorter life expectancy than others. If no precaution is taken based on the degradation status of each core, the most probable outcome is that the processor will die early as one core in the system fails before others. Such a processor, however, cannot be utilized in most scenarios today. The proposed technique can help alleviating this problem through core-level DVFS control and OS-controlled workload management based on core utilization. If the OS is informed of the degradation status of each core, it could schedule the workload cleverly so that sturdy cores (those with low threshold voltages) work more than the frail cores (those with high threshold voltages). This will lead to the lifetime equalization of cores and thereby extend overall processor life while saving more power.

Other research groups have proposed different solutions to tackle the NBTI-related aging of processors. The closest solution to ours was proposed by Tiwari and Torrellas through workload management based on application temperature and chip-wide changes to supply and threshold voltages without considering the needs of DVFS. As a result, the system performance is compromised to increase the lifetime of the processor [1]. Another approach is to place spare parts at manufacturing time as proposed by Srinivasan *et al*. The spare parts are either utilized at the beginning to boost system performance, and as parts die, they are turned off, which reduces system performance gradually, or they are kept off until a part dies and used for replacing the failing component to maintain processor functionality [2]. Finally, a transistor-level solution was proposed by Abella *et al.*, in which bits are flipped to even out the degradation across the processor and undo some damage through relaxation [3].

Our paper is organized as follows: In the next section, we discuss the NBTI problem at the transistor level and the

attempts to measure NBTI-related degradation. Then, section III explains the system we are considering for this research. In section IV, our NBTI-aware DVFS model used for estimating power and lifetime savings is presented. We briefly describe our simulator in section V, which was used to generate the results, followed by our evaluation. Finally, in section VI, we conclude with our observations and give an idea of where this research is headed.

II. Background

While most of the miniaturization problems in CMOS technology are more or less related to doping issues, the continued reduction of the gate oxide thickness has necessitated the incorporation of nitrogen into silicon dioxide, which in turn has aggravated NBTI, especially in technology nodes smaller than 90 nm. The key device parameters of p-MOSFETs, such as threshold voltage and saturation current, show a rapid shift under negative bias at elevated temperatures due to the build-up of positive interface charges, which essentially slows the operation of the circuit down [4].

Although several NBTI models have been developed to explain the physics of interface trap generation based on electrochemical reactions and activation energies, only the reaction-diffusion (R-D) model can explain the power-law time dependence of the NBTI degradation. The R-D model states that the NBTI induced shift in p-MOSFET parameters is driven by breaking of hydrogen-passivated silicon bonds at the substrate interface and the subsequent diffusion of hydrogen into the gate oxide [5]. The formation and movement of these interface traps is shown in Figure 1 below.

In this research, we attempt to utilize Zhang and Orshansky's R-D model for NBTI degradation and determine the status of the degradation in a multi-core processor instead of monitoring it constantly through hardware structures placed on the chip. There are a few research groups working on measuring the NBTI degradation right on the chip; however, most methods make the degradation worse and therefore are not very accurate. Our model approximates the degradation, thus requiring minimal assistance from hardware measurements for only verification and adjustment purposes. Consequently, the degradation does not require constant hardware monitoring. We use the R-D model to predict an upper bound on processor degradation, which will allow the system be monitored infrequently and decrease the negative impact of these hardware measurements to a minimum.
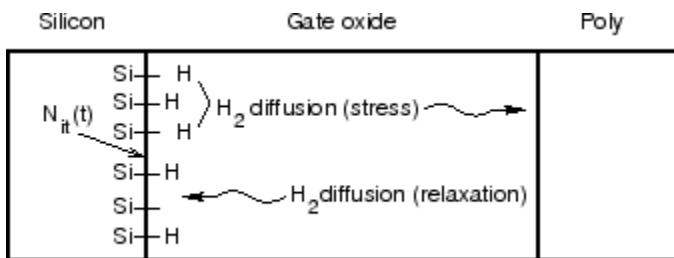
III. System

As we are looking at a multi-core chip, process variation across the processor is a naturally expected phenomenon, which necessitates the monitoring of degradation at different parts on the die since process variation may cause the state and progress of degradation in each core be distinct. In other words, different cores will have different initial V_T values and will degrade at a different rate. For this purpose, to account for process variation across the die, we allow the cores in our system to vary by up to $\pm10\%$ of a base V_T [6].

Due to the different threshold voltages present on the chip, the timing margin in each core is different, thus demanding the operation of each core with a distinct supply voltage in order to meet the critical timing requirements. Furthermore, as the degradation progresses during the lifetime of the processor, our system needs to adjust itself to compensate for the decaying critical path in each core. Depending on the current status of degradation, we can select a different NBTI voltage for every core individually at any instant, which is lower than the DVFS voltage selected by the processor to sustain its operation at the desired frequency. This allows our research to save power beyond any prior study and extend lifetime.

We implemented our system in two steps, where we first looked at the effects of core scheduling at a very coarse-grained level. For this initial study, 8 cores were used and scheduled by the OS based on the typical workloads observed in Google's servers [7]. The cores were considered either "ON" or "OFF" and scheduled with day time granularity. Then, we added DVFS to our model, where we no longer assumed that cores were just either "ON" or "OFF". This required that we change the amount of degradation depending on the utilization of the core, and allowed the OS to not only schedule whether a core should be "ON" or "OFF" but how much it should be utilized. Core utilization was reflected in the DVFS voltage the core was run at, and depending on the voltage, the core would see different amounts of degradation.

IV. Model

Using the R-D model along with our system assumptions presented in the previous section, we developed a way of splicing curves to get degradation based off of DVFS effects and coded a simulator capable of following multiple cores using DVFS on varying workloads and core utilizations.

Fig. 1. Movement of hydrogen atoms during stress and relaxation phases [4].
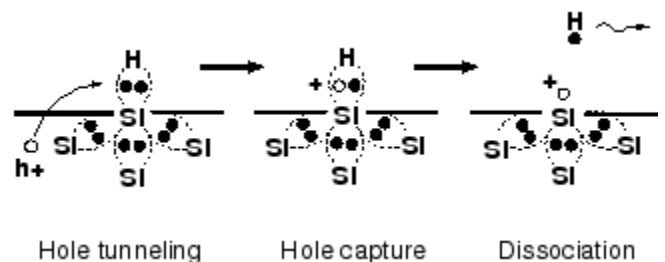
Moreover, in our simulator, we assumed that we were investigating the behavior of the weakest transistor in each core; thus, other transistors did not matter for the purposes of our experiment. This implied that our simulator could detect failure as soon as it occurred. Particularly, we assumed we were able to measure the current state of degradation for all cores.

For our simulation, we used a curve splicing technique where degradation curves were generated for each voltage DVFS was allowed to run a core at, and the current V_T was used to find the starting point for the degradation calculation. Figure 2 shows approximately how our curve splicing technique would proceed. For example, assume a core on a processor is run at voltage V1 for one day and V2 for the next day. The V_T of the core is found after the first day using the degradation curve for V1, and this point is then located on the V2 degradation curve. Next, the core would run for another day at this voltage, and the terminal V_T is taken off of the V_T curve for V2 one day later.

V. Evaluation & Results

To facilitate easy calculation of core lifetimes and power savings, we wrote a simulator that uses our model with core scheduling and DVFS in C. The simulator follows 8 cores over a 10-year period. It randomly selects an initial V_T for each core within ±10% of a nominal value (0.20V in our case). Then, it generates results by assigning the demanding work to the sturdier cores to facilitate OS scheduling. The simulator schedules workloads at the granularity of a single day based on a distribution graph generated at Google. Below Table I shows all of the parameters used in our simulator.

For every day during the lifetime of the processor, we determine how many cores are required to handle the tasks that will be run on the processor based on our workload model. Then, we select that many sturdy cores in our system to run these tasks. Next, depending on processor utilization, we select the appropriate DVFS voltage for each "ON" core that can handle the workload with no degradation in

performance, which are selected from the set of 1.00V, 1.10V, and 1.20V. Since the number of cores and the DVFS voltage for each core is determined by the workload itself, our approach does not degrade the performance by any means. For example, if all cores are needed for any given day with 100% utilization in each core, our model allows the entire processor to be used to the full extent. On the other hand, if only one core is needed, then only the sturdiest core in the system is utilized. This approach therefore maximizes the lifetime of the whole processor by equalizing the degradation of the cores over their lifetime, which was beyond 10 years in our simulations. In contrast, random scheduling led to core failure after only 4 years of service on average.

To estimate the power savings of the system, we compared the voltage selected by our model against that normally selected by DVFS alone. The power savings were computed by the following formula in equation (1).

$$Power\ savings = 1 - \frac{(NBTI\ voltage\ )^2}{(DVFS\ voltage\ )^2} \qquad (1)$$

As a consequence of examining a multi-core processor, the NBTI-related power savings needed to be estimated for each core independently. In the absence of DVFS during our initial study, an average power savings of about 13.5% at the end of a processor's 10-year lifetime was achieved if proper scheduling was attained. On the other hand, our simulation estimates with DVFS showed 9.25% power savings over 10 years on average. Consequently, we have observed a decrease in the power savings attributed to our NBTI model when we introduced DVFS into the system. The reduction comes from the fact that the voltage margins get smaller and smaller with lower DVFS voltages. As DVFS has become the baseline for most processors today, we believe our estimate of the power savings is therefore more accurate in the presence of DVFS.

VI. Conclusion & Future Work

In this report, we have presented an estimation of power and lifetime savings by incorporating DVFS into prior NBTI
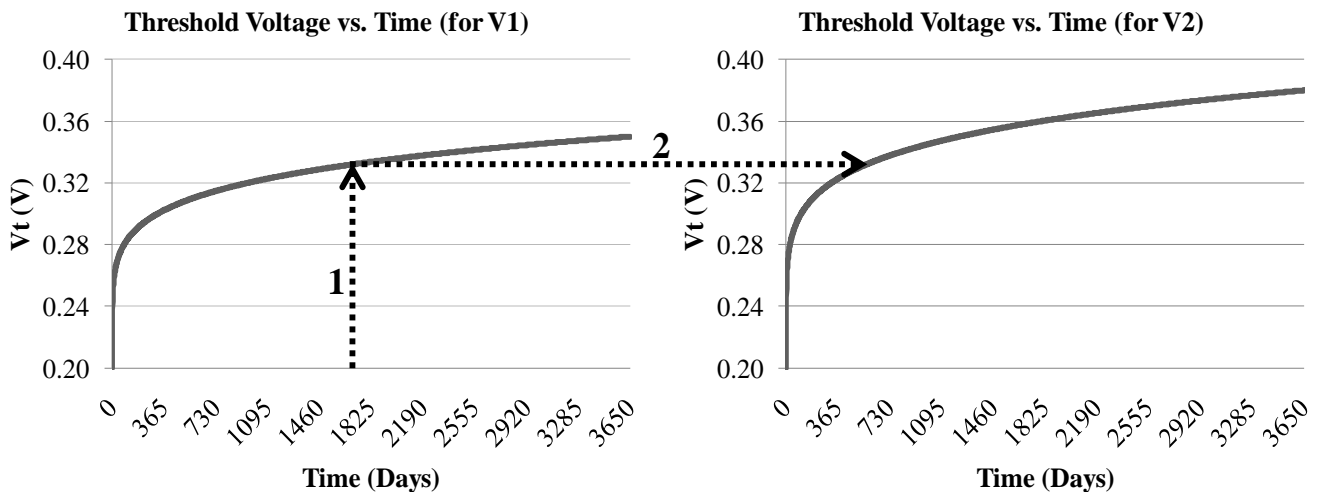

Fig. 2. Threshold voltage is found for the first given supply voltage and then mapped onto the curve for the second supply voltage.

TABLE I
Simulation Parameters

Parameter	Value
Number of cores/processor	8
Process variation method	Random (Linear distribution)
Process variation amount	±10%
Scheduling method	Lifetime maximization
Workload method/amount	Google workload curve
Voltage step intervals	1 day (86,400 seconds)
Starting base threshold voltage	0.20V
Failure threshold voltage	0.34V
Operating (supply) base voltage	1.20V
DVFS voltages	0.00V, 1.00V, 1.10V, and 1.20V
Operating temperature	335K (62^0C)

degradation model. Our approach does not degrade system performance at all while taking advantage of the existing time margins of any kind of processor over its lifetime and does not require any manual user intervention. We show that the actual power savings is approximately 9.25% over 10 years in the presence of DVFS and 13.5% in its absence. On the other hand, lifetime savings lies around 3 years in both cases.

We believe that our work can be extended to pinpoint the exact amount of lifetime savings with simple modifications in our simulator, which currently only gives a rough estimate. Furthermore, we take advantage of the fact that we have coded a simulator to estimate the degradation in a system with a priori knowledge of frail and sturdy cores. Such knowledge may not be readily available and most likely requires some hardware structures to measure the initial process variation of the system. Even then, the rate of degradation in our NBTI model is based on approximation presented in Intel Technology Journal – June 2008 on a 45nm processor and may differ from one processor to another. Consequently, our approach needs to be supported by on-chip structures to measure the degradation during the lifetime of the processor and correct our model when need be.

Another important evaluation step to perform is to model NBTI degradation under real system workload. To accomplish this, our simulator can be augmented to accept utilization traces from real systems and estimate degradation in the given system based on the real workload. For this purpose, SPEC and distributed benchmarks can be employed, which would give us more standardized results than our current Google workload model.

Acknowledgements

References

[1] A. Tiwari and J. Torrellas, "Facelift: Hiding and Slowing Down Aging in Multicores," In *Proceedings of the 2008 41st IEEE/ACM international Symposium on Microarchitecture - Volume 00* (November 08 - 12, 2008). International Symposium on Microarchitecture. IEEE Computer Society, Washington, DC, 129-140.

[2] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Exploiting Structural Duplication for Lifetime Reliability Enhancement," In *Proceedings of the 32nd Annual international Symposium on Computer Architecture* (June 04 - 08, 2005). International Symposium on Computer Architecture. IEEE Computer Society, Washington, DC, 520-531.

[3] J. Abella, X. Vera, and A. Gonzalez, "Penelope: The NBTI-Aware Processor," In *Proceedings of the 40th Annual IEEE/ACM international Symposium on Microarchitecture* (December 01 - 05, 2007). International Symposium on Microarchitecture. IEEE Computer Society, Washington, DC, 85-96.

[4] R. Wittmann, "NBTI Reliability Analysis," *Miniaturization Problems in CMOS Technology: Investigation of Doping Profiles and Reliability*, sec. 5, Apr. 2009. http://www.iue.tuwien.ac.at/phd/wittmann/node10.html

[5] B. Zhang, and M. Orshansky, "Modeling of NBTI-Induced PMOS Degradation under Arbitrary Dynamic Temperature Variation," In *Proceedings of the 9th international Symposium on Quality Electronic Design* (March 17 - 19, 2008). International Symposium on Quality Electronic Design. IEEE Computer Society, Washington, DC, 774-779.

[6] M. Royd *et al.*, "System power management support in the IBM POWER6 microprocessor," *IBM J. Res. & Dev.*, v. 51, n. 6, pp. 733-746, Nov. 2007.

[7] X. Fan, W. D. Weber, and L. A. Barroso, "Power provisioning for a warehouse-sized computer," *Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA '07)*, pp. 13–23, 2007.