Analysis and Mitigation of NBTI-Induced Performance Degradation for Power-Gated Circuits

Kai-Chiang Wu[1], Diana Marculescu[1], Ming-Chao Lee[2], and Shih-Chieh Chang[2]
[1]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[2]Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan
{kaichiaw, dianam}@ece.cmu.edu, chao@nthucad.cs.nthu.edu.tw, scchang@cs.nthu.edu.tw

ABSTRACT

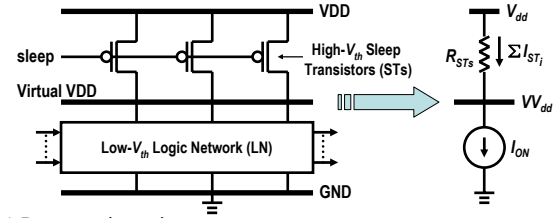
Device aging, which causes significant loss on circuit performance and lifetime, has been a main factor in reliability degradation of nanoscale designs. Aggressive technology scaling trends, such as thinner gate oxide without proportional downscaling of supply voltage, necessitate an aging-aware analysis and optimization flow in the early design stages. Since PMOS sleep transistors in power-gated circuits suffer from static NBTI during active mode and age very rapidly, the aging of power-gated circuits should be explicitly addressed. In this paper, for power-gated circuits, we present a novel methodology for analyzing and mitigating NBTI-induced performance degradation. Aging effects on both logic networks and sleep transistors are jointly considered for accurate analysis. By introducing 25% redundant sleep transistors with reverse body bias applied, the proposed methodology can significantly mitigate the long-term performance degradation and thus extend the circuit lifetime by 3X.

Keywords: Aging, Leakage, NBTI, Power gating, Reverse body bias

I. INTRODUCTION

In the recent decade, the design and manufacturing of semiconductor devices have undergone dramatic innovations, many of them coming at the price of decreased reliability of nanometer integrated circuits. Some of the major challenges driving the design of reliable systems (*i.e.*, design for reliability) encompass soft errors, process variations, and aging phenomena. With the continuous scaling of transistor dimensions, device aging, which causes temporal performance degradation and potential wear-out failure, is becoming increasingly dominant for lifetime reliability concerns because of limited timing margins. Therefore, an early-stage analysis and optimization flow considering aging effects is necessitated as a key factor in guaranteeing consistently reliable operation over a desired lifespan.

Among various aging mechanisms, *negative bias temperature instability* (NBTI) [1] is known for being particularly crucial due to current scaling trends such as shrinking thickness of gate oxide. NBTI is a PMOS aging phenomenon that occurs when PMOS transistors are stressed under negative bias ($V_{gs} = -V_{dd}$) at elevated temperature. As a result of the dissociation of *Si-H* bonds along the $Si\text{-}SiO_2$ interface, NBTI-induced PMOS aging manifests itself as an increase in the threshold voltage (V_{th}) and decrease in the drive current (I_{ds}) [2], which in turn slow down the rising propagation delay of logic gates. Experiments on PMOS aging [3] indicate that NBTI effects grow exponentially with thinner gate oxide and higher operating temperature. If the thickness of gate oxide shrinks down to 4nm, the circuit performance can be degraded by as much as 15% after 10 years of stress and lifetime will be dominated by NBTI [4].



(a) Power gating using PMOS sleep transistors **(b)** Equivalent RC model

Figure 1. A header-based power gating structure

In addition to the oxide thickness and operating temperature, NBTI-induced performance degradation strongly depends on the amount of time during which a PMOS transistor is stressed. The increase in threshold voltage has been shown in [5] to be a logarithmic function of the corresponding stress time/probability. This parameter of each transistor is distributed non-uniformly across a logic circuit, leading to 2-5X difference in the degradation rate of threshold voltage [6]. In contrast, when the stress condition is relaxed ($V_{gs} = 0$), the aging mechanism can be recovered partially and the threshold voltage decreases toward the nominal value by more than 75% if the recovery phase lasts sufficiently long [6].

In order to minimize static power dissipation which accounts for a large portion of total power consumption in the 90nm technology or below, high-V_{th} sleep transistors [7] are employed as switches to disconnect a circuit from VDD (see Figure 1(a)) or GND when the circuit is inactive, *i.e.*, in standby mode (sleep = "1"). A PMOS/NMOS sleep transistor is referred to as a header/footer inserted between VDD/GND and the circuit. Despite smaller size required for the same driving strength, a footer has to be placed in an isolated p-well, which involves a twin-well manufacturing process and, for cell-based design, re-modeling the cell library. Generally, the header-based style of using PMOS is fairly popular due to its ease of manufacturing and library design. This technique, called *power gating* (PG), is a coarse-grained application of multi-threshold CMOS (MTCMOS) and widely used for reducing sub-threshold leakage current [8]. However, in a header-based PG design, the PMOS sleep transistors suffer continuous NBTI stress during active mode (sleep = "0") and age very rapidly. The relentless aging impact on the headers will aggravate the performance degradation of the logic circuit in a PG structure. As a result, not only the NBTI effects on logic networks but also those on sleep transistors need to be addressed when header-based PG is exploited. In this paper, for power-gated circuits, we present an integrated NBTI degradation model for accurate analysis of the long-term performance behavior. Afterwards, an optimization methodology is proposed to mitigate the overall performance degradation for a longer period of reliable operation.

978-1-61284-660-6/11/$26.00 © 2011 IEEE

The rest of this paper is organized as follows: Section II gives an overview of related work and outlines the contribution of our paper. In Section III, we introduce the NBTI degradation model used to evaluate the performance loss of logic circuits in the power-gating design style. Section IV presents our methodology of mitigating NBTI-induced performance degradation for power-gated circuits. In Section V, the experimental results for a set of standard benchmarks are demonstrated. Finally, we conclude our work in Section VI.

II. Related Work and Paper Contribution

A. Previous Work on Aging-Aware Optimization

Traditional design methods add guard-bands or adopt worst-case margins to deal with aging phenomena, which in practice imply over-design and may be expensive. To avoid overly conservative design, the mitigation of aging-induced performance degradation can be formulated as a timing-constrained area minimization problem with consideration of aging effects. The authors of [9] proposed a gate sizing algorithm based on Lagrangian relaxation. An average of 8.7% area penalty is required to ensure reliable operation for 10 years. Other methods related to gate or transistor sizing can be found in [10][11].

A novel technology mapper considering signal probabilities for NBTI was developed in [12]. On average, 10% area recovery and 12% power saving are accomplished, as compared to the most pessimistic case assuming static NBTI on all PMOS transistors. A framework using joint logic restructuring and pin reordering [13] can mitigate NBTI-induced performance degradation with no gate area overhead. In [14], a reconfigurable flip-flop design based on time borrowing is introduced for aging detection and correction.

Instead of reducing NBTI effects during active mode as described above, an idea of NBTI-aware optimization during standby mode was presented in [15]. Input vectors for minimum standby-mode leakage are selected to minimize PMOS aging. Furthermore, for gates that are deep in a large circuit and cannot be well controlled by primary input vectors, internal node control [16] intrusively assigns logic "1" to those gates if they are on the critical paths. The logic "1" relaxes the stress condition and can thus relieve the NBTI impact. In [17], power gating is exploited for aging optimization by shutting off the power supply to a circuit. However, the continuous V_{th} degradation of sleep transistors during active mode in the case of header-based PG design is ignored in [15][16][17].

The first work addressing the aging of sleep transistors was outlined in [18]. The authors proposed to realize NBTI-aware power gating through (*i*) sleep transistor over-sizing, (*ii*) forward body-biasing, and (*iii*) stress time reduction. As opposed to [15][16][17], the aging of logic networks is not considered in [18]. In the sequel, we will show the interdependence between the degradation effects on logic networks and sleep transistors. We have also experimentally verified that, without joint modeling of these interdependent effects, the overall performance degradation of power-gated circuits cannot be precisely estimated.

B. Paper Contribution

Based on the characterization of NBTI effects on both logic networks (LNs) and sleep transistors (STs), we present an analysis and optimization methodology for header-based power-gated circuits in terms of performance-centric lifetime reliability. Existing work does not model both LN aging and ST aging in a unified manner with their interdependency formulated. The contributions and advantages of this work are threefold:

- ***Joint modeling of interdependent degradation effects on logic networks and sleep transistors*:** Due to the increasing V_{th} of sleep transistors during active mode, the voltage level (denoted by VV_{dd} as depicted in Figure 1) at which the logic network operates gradually decreases, therefore imposing additive performance loss. On the other hand, the decrease in VV_{dd} can be offset, to a certain extent, by the smaller current required for normal operation of the logic network due to its own degradation. These two effects are interdependent and should not be treated separately. In this paper, for the first time, a joint model considering the interdependency is developed for accurate analysis of aging behavior for power-gated circuits.
- ***Exploration of ST redundancy and NBTI recovery*:** We introduce redundant STs and implement a scheduling architecture such that the original STs can be shut off periodically during active mode. The proposed methodology explores the recovery mechanism by taking STs' turns recovering from NBTI. Hence, the VV_{dd} decrease is slowed down, which mitigates the long-term performance degradation and extends the circuit lifetime.
- ***Significant lifetime extension while retaining the purpose of power gating – leakage saving*:** To minimize the additional leakage current flowing through those redundant STs, reverse body bias is applied to increase their fresh V_{th} values (at time 0). Based on the observation in [6] that a high-V_{th} transistor ages slower than a low-V_{th} transistor, the use of redundant STs with reverse body bias can achieve significant lifetime extension for power-gated circuits without incurring much overhead in leakage power. This is in contrast to using forward body bias (as in [18]) which can increase leakage power by 197%.

III. Aging Analysis for Power-Gated Circuits

A. NBTI Degradation Model for Logic Networks

Before discussing the integrated analysis framework for power-gated circuits as a whole, we briefly introduce the NBTI degradation model for logic gates/networks [6][19] used in our paper. This model enables us to analyze the long-term behavior of NBTI-induced PMOS degradation, with both aging and recovery mechanisms taken into account. First, the degradation of threshold voltage at a given time t can be predicted as:

$$\Delta V_{th} = \left(\frac{\sqrt{K_v^2 \cdot T_{clk} \cdot \alpha}}{1-\beta_t^{1/2n}} \right)^{2n} \quad \textbf{(1)}$$

where K_v is a function of temperature, electrical field, and carrier concentration, α is the stress probability, and n is the time exponential constant, 0.16 for the used technology. The detailed explanation of each parameter can be found in [6].

Next, the authors of [19] simplify this predictive model to be:

$$\Delta V_{th} = b \cdot \alpha^n \cdot t^n = b \cdot (\alpha \cdot t)^n \quad \textbf{(2)}$$

where $b = 3.9\times10^{-3}$ V·s$^{-1/6}$.

Finally, the rising propagation delay of a gate through the degraded PMOS can be derived as a first-order approximation:

$$\tau_p' = \tau_p + a \cdot (\alpha \cdot t)^n \quad \textbf{(3)}$$

where ι_p is the intrinsic delay of the gate without NBTI degradation and a is a constant.

We apply Equation (3) to calculate the delay of each gate under NBTI, and further estimate the performance of a logic circuit. The coefficient a in Equation (3) for each gate type and each input pin is

extracted by fitting HSPICE simulation results in 65nm, Predictive Technology Model (PTM). The simplified long-term model successfully predicts the PMOS degradation, with less than 5% loss of accuracy against cycle-by-cycle simulations [6].

B. NBTI Degradation Model for Sleep Transistors

To analyze the performance degradation of power-gated circuits due to the NBTI impact on sleep transistors, the voltage level of virtual VDD should be the main focus. Virtual VDD, which supplies the logic circuit in a PG structure with required operating voltage, is a virtual bus connecting the drain terminals of all sleep transistors [7]. Because of the resistance between VDD and virtual VDD, a voltage drop at virtual VDD can be observed. Typically, sleep transistors are sized such that a tradeoff among voltage drop, leakage saving, and area overhead is obtained [8].

In the presence of NBTI, the effective resistance between VDD and virtual VDD increases due to the increasing V_{th} of sleep transistors and thus, the voltage drop is becoming larger with NBTI stress, which will impose additive performance loss beyond that on the logic itself. The model for performance degradation as a result of the increasing voltage drop is described as follows.

Consider the example of header-based power gating in Figure 1(a). An equivalent RC model is shown in Figure 1(b) where the resistor characterizes the network of sleep transistors (between VDD and virtual VDD) and the current source characterizes the logic network (between virtual VDD and GND). Note that a finer-grained RC model with various resistors and current sources can be employed for more realistic analysis if the detailed information about physical implementation is available.

The increase in V_{th} of sleep transistors can be determined by Equation (1). Given the degraded threshold voltage ($V_{th}' = V_{th} + \Delta V_{th}$), we update the current flowing through a sleep transistor ST_i using the MOSFET current equation:

$$I'_{ST_i} \approx \mu_p C_{ox} \frac{W_{ST_i}}{L}(V_{gs} - V'_{th})V_{ST} \text{ as } V_{ST} \text{ is small} \quad \textbf{(4)}$$

where μ_p is the hole mobility, C_{ox} is the oxide capacitance, W_{ST} is the width of the sleep transistor, and V_{ST} is its drain-to-source voltage, simply the voltage drop at virtual VDD and supposed to be small (*e.g.*, 5% of V_{dd}).

Next, the effective resistance of the network of sleep transistors under aging can be derived as:

$$R'_{STs} = \frac{V_{dd} - VV_{dd}}{\sum_i I'_{ST_i}} = \frac{V_{ST}}{\sum_i I'_{ST_i}} \approx \frac{L}{\sum_i W_{ST_i}} \cdot \frac{1}{\mu_p C_{ox}(V_{gs} - V'_{th})} \quad \textbf{(5)}$$

where VV_{dd} is the voltage level of virtual VDD.

We can then calculate the new (lower) VV_{dd}:

$$VV'_{dd} = V_{dd} - I_{ON} \cdot R'_{STs} \quad \textbf{(6)}$$

where I_{ON} is the active (turned-on) current drained by the logic network, which is the maximum cumulative switching current of a set of gates that switch simultaneously.

Finally, the propagation delay of each gate in the power-gated circuit can be estimated based on the alpha-power law:

$$\tau_p \propto \frac{VV_{dd}}{(VV_{dd} - V_{th})^{\alpha_f}} \quad \textbf{(7)}$$

where α_f is the technology-dependent velocity saturation factor.

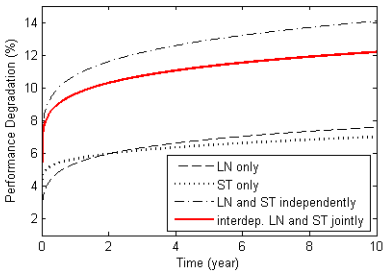


Figure 2. Analysis results of the proposed model for power-gated circuits

In prior art, only the NBTI degradation effect of VV_{dd} on the logic network has been examined, and the aging of the logic itself, which leads to a decreasing I_{ON}, is not included. It is evident from Equation (6) that the performance degradation of power-gated circuits will be overestimated without taking the I_{ON} decrease into account. The dependence of I_{ON} on the degradation of logic networks is based on the charge-current formula:

$$I = \frac{dQ}{dt} = C\frac{dV}{dt} \Rightarrow I_{ON(Gate)} \propto \frac{\Delta VV_{dd}}{\tau_p} \quad \textbf{(8)}$$

According to Equation (8), we can trace the change in the current drained by a gate and further derive the degraded I_{ON} by summing up the current of those gates that switch simultaneously. In terms of the VV_{dd} degradation (see Equation (6)), the decrease in I_{ON} is actually beneficial since it partially (but not fully) offsets the increase in R_{STs}.

Here, we summarize the interdependence between the degradation effects on sleep transistors (STs) and logic networks (LNs):

(i) The effect of ST aging (*i.e.*, decreasing VV_{dd}) aggravates the performance degradation of LNs.

(ii) The effect of LN aging (*i.e.*, decreasing I_{ON}) alleviates the effect of ST aging.

The interdependent effects are particularly important for accurate analysis of NBTI-induced performance degradation for power-gated circuits and have been incorporated in our analysis framework. Figure 2 shows the analysis results of the proposed framework in 65nm PTM for an industrial benchmark *AES* assuming that it is in active mode 60% of the time. As it can be seen, considering LN aging only ("LN only") or considering ST aging only ("ST only") underestimates the performance degradation, while ignoring the interdependency ("LN and ST independently") overestimates the performance degradation. Note that power gating can remove the stress condition for logic devices by pulling down VV_{dd} toward 0V several clock cycles after the circuit goes standby [17]. Therefore, the 10-year performance loss of "LN only" is smaller than that reported in the literature, where circuits are not power-gated. In the case of joint modeling of interdependent LN and ST aging ("interdep. LN and ST jointly"), the results exhibit an in-between degradation trend. A chain of inverters simulated by HSPICE (see Figure 3) indicates that the normalized error of the proposed NBTI degradation model is always within 1.5% over the 10-year performance prediction. Note that the error tends to saturate after 7 years even though it is increasing from the 2nd to 7th years.

IV. Lifetime Extension for Power-Gated Circuits

It has been demonstrated in Section III-B that the overall performance degradation of a power-gated circuit can reach significant levels (>12%). If the timing margin of a design is 10%, the design under power gating will likely wear out within two years, as shown

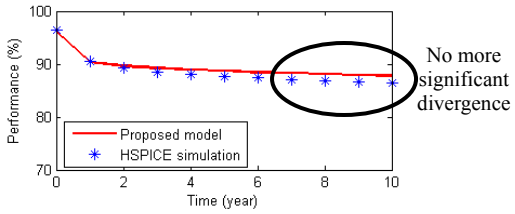


Figure 3. HSPICE validation with a chain of inverters

in Figure 2. This is definitely unacceptable for most state-of-the-art applications of power gating. In this section, we propose to introduce redundant STs and develop a scheduling framework such that the original STs can be shut off periodically during active mode. The ultimate goal of our methodology is to maximize the lifetime of power-gated circuits while retaining the purpose of power gating, *i.e.*, leakage saving.

A. Problem Formulation

The proposed methodology is formulated as an area-constrained optimization problem for concurrent lifetime extension and leakage saving. Given an allowable percentage p% on the total width of redundant STs, the objective is to determine an optimal value of reverse body bias such that, when applied on the redundant STs, the lifetime of a power-gated circuit can be significantly extended with minimal leakage overhead. The lifetime is measured as the duration of time during which the circuit can operate with its performance loss not exceeding 10% (wear-out if exceeding 10%). The problem formulation is given as:

$$\texttt{Maximize} \quad w \cdot \frac{Lifetime(C_{r\%}, V_b, d) - Lifetime(C_{0\%}, V_b, d)}{Lifetime(C_{0\%}, V_b, d)} + (1-w) \cdot \frac{Leakage(C_{0\%}, V_b, d) - Leakage(C_{r\%}, V_b, d)}{Leakage(C_{0\%}, V_b, d)}$$

$$\texttt{Subject to} \quad \begin{matrix} r \le p \\ V_{dd} < V_b \le V_{\max} \end{matrix} \qquad \textbf{(9)}$$

where w ($0 < w < 1$) is the weight for lifetime extension, $C_{r\%}$ is the circuit with r% ST redundancy introduced (thus $C_{0\%}$ is the original power-gated circuit), V_b is the bulk voltage assigned to redundant STs (for reverse body-biasing), and d is the duty cycle of the circuit (defined as the ratio of active time to total time).

B. Exploring NBTI Recovery via ST Redundancy

Given a power-gated circuit with the number and total width of STs optimally determined, a certain number of STs are introduced as redundant STs to combat NBTI-induced performance degradation. Since the current circuit has more STs than necessary, not all of them need to be turned on for normal operation during active mode, especially before the STs experience significant aging. With the existence of ST redundancy, we can explore the recovery mechanism by shutting off STs by turns, giving them extra time to recover from NBTI. Hence, the VV_{dd} decrease due to ST aging is slowed down, which mitigates the long-term performance degradation and extends the circuit lifetime.

Figure 4 shows the hardware architecture of our NBTI-aware power gating design [20] (the width of each ST specified below it) where ST_1-ST_4 and ST_6-ST_9 are the original STs, and ST_5 and ST_{10} (highlighted) are the redundant STs, *i.e.*, 25% ST redundancy in terms of total ST width (4W/16W). Shift registers (SRs) are deployed to drive groups of STs such that, during active mode, one or more of the

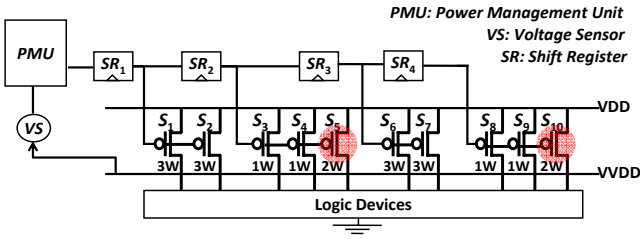


Figure 4. NBTI-aware power gating design

ST groups can be shut off by intermittent "sleep" signals (logic "1") sent from the power management unit (PMU). ST grouping is pre-determined based on the wakeup scheduling [21] and the redundant STs are evenly distributed to the groups in which the subtotal widths of STs are smaller. By doing so, every group has more balanced subtotal ST width and subsequently, we will have more flexibility in exploring NBTI recovery by switching STs on and off. After introducing redundant STs, the wakeup scheduling can be further refined for better behavior during power mode transition. The refinement of wakeup scheduling is beyond the scope of this work and not particularly addressed in this paper. The voltage sensor (VS) compares VV_{dd} with a reference value and outputs a signal on which the PMU decides whether to adjust the "sleep" patterns.

In this example, where 25% redundant STs have been placed, the logic circuit can operate properly under the 10% performance bound with part of the original and redundant STs turned on, as long as the total width of turned-on STs is sufficient. To this end, round-robin scheduling is adopted in the PMU to assert a "sleep" signal (logic "1") every five cycles during active mode, thus rendering a duty cycle of 80% for each ST while satisfying the requirement on the total width of turned-on STs. Once the VS detects a significant voltage drop, the PMU will assert "sleep" signals less frequently to realize a higher duty cycle and on average, more STs can be turned on for guaranteeing reliable operation. In this hardware configuration with the support of SRs, PMU, and VS, the stress probability of each ST is as low as 80%, meaning that the STs no longer suffer continuous NBTI stress during active mode and can recover from NBTI within the 20% time intervals.

Not shown in Figure 4, to avoid excessive glitches on the virtual VDD resulting from the switching of STs when logic devices are draining current, the clock signals to SRs are delayed and/or frequency-divided, so that STs will not be triggered at the same time as the logic devices. Without affecting the circuit behavior (timing and functionality) and cumulative amount of time for NBTI recovery, this strategy effectively diminishes the likelihood of excessive glitches occurring on the virtual VDD.

C. Applying Reverse Body Bias

The major drawback of introducing ST redundancy is the additional leakage current flowing through those redundant STs, which is proportional to the total width of redundant STs in a power-gated circuit. In order to not incur too much leakage overhead, *reverse body bias* (RBB) is applied on the redundant STs to increase their fresh V_{th} values. It is well-known that sub-threshold leakage decreases exponentially with higher V_{th}:

$$I_{\text{sub}} \approx I_0 \cdot e^{\frac{(V_{gs} - V_{th})q}{nkT}} \qquad \textbf{(10)}$$

where I_0 is the current at $V_{gs} = V_{th}$, n is the sub-threshold slope factor,

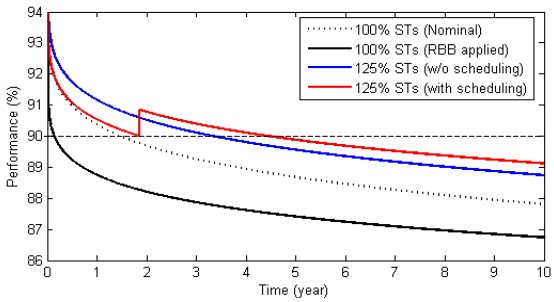


Figure 5. Comparison of aging behaviors with various settings

k is the Boltzmann constant, T is the absolute temperature, and q is the electron charge.

Therefore, the use of reverse body-biasing greatly reduces the overhead in leakage power. Meanwhile, the benefit of ST redundancy along with the proposed scheduling scheme, which is influenced only marginally, is discussed in the following.

As previously indicated [6], the variation in V_{th} can be compensated by the NBTI effect: the transistor with a higher (lower) fresh V_{th} ages at a lower (higher) rate and thus, the high V_{th} and the low V_{th} tend to converge toward the nominal case as the stress of NBTI continues. For example, the high, nominal, and low V_{th} values at time 0 are 210mV, 180mV, and 150mV, respectively. The difference between the high V_{th} and the low V_{th} remarkably shrinks from 60mV at time 0, to 11.9mV at 10 years.

As a result, we can minimize the leakage overhead due to the introduction of ST redundancy, by assigning an optimal value of bulk voltage (V_b) greater than V_{dd} to the redundant STs. On the other hand, based on the aforementioned fact that a high-V_{th} transistor ages slower than a low-V_{th} transistor, the mitigation of "long-term" performance degradation and the extension of circuit lifetime are still comparable to the case where RBB is not applied. The comparison will be demonstrated later in Section V.

By determining the optimal V_b value which maximizes the cost as a joint function of *Lifetime* and *Leakage* (see Equation (9)), we can achieve significant lifetime extension for a power-gated circuit without incurring too much leakage overhead. Typically, w is chosen to be smaller than 0.5 because the lifetime extension is always larger than the leakage overhead. Due to the efficiency of our analysis framework, we can afford to exhaustively search for the optimal V_b with a discrete step of 50mV from V_{dd} to $V_{\max}$.

V. Experimental Results

We have implemented the proposed methodology of mitigating NBTI-induced performance degradation for power-gated circuits. Experiments are conducted on an industrial circuit (*AES*) and a set of benchmarks from the ISCAS and MCNC suites. The technology used is 65nm, Predictive Technology Model (PTM). The supply voltage is 1.0V and the operating temperature is assumed to be 300K. Our framework aims at extending circuit lifetime with less leakage overhead. For each benchmark, logic simulation with 10,000 random patterns, assuming that the 0-probabilities of all primary inputs are 0.5, is applied to calculate the probability of each signal. Given signal probability α of the input to a PMOS (a ST or in the logic), the V_{th} degradation of the PMOS can be predicted by Equation (2). For each gate type and input pin (PMOS) in the logic, HSPICE simulations with its nominal and degraded V_{th} values are performed for a discrete set of signal probabilities from 0 to 1. We fit these

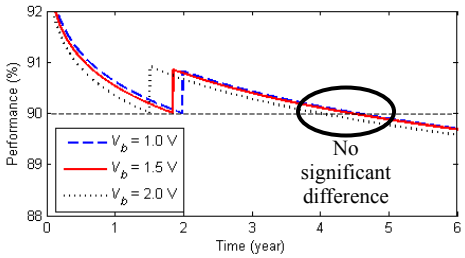


Figure 6. Lifetime vs. V_b (bulk voltage)

HSPICE results to obtain coefficients a's for Equation (3). Therefore, the gate delay and circuit timing under NBTI can be estimated based on the degradation model presented in Section III.

Figure 5 depicts the normalized aging behaviors of circuit *AES* with various settings: (*i*) the nominal PG design of *AES*, (*ii*) applying RBB on all STs in the nominal design (no ST redundancy), (*iii*) introducing 25% ST redundancy (RBB applied) with ***no*** ST scheduling implemented, *i.e.*, all STs (original and redundant) on during active mode, and (*iv*) 25% ST redundancy (RBB applied) with round-robin scheduling implemented. As it can be seen, the nominal design (dotted line) has a performance degradation of more than 10% after 1.47 years (lifetime = **1.47** yrs). If 25% body-biased redundant STs are introduced (blue line), the lifetime becomes **3.33** yrs. If the proposed ST scheduling architecture is incorporated (red line), the lifetime is further extended to **4.45** yrs. The upward bounce of red line around year 2 is because we discard the scheduling scheme when it first reaches the margin of 10% performance loss. By doing so, all original and redundant STs are constantly (rather than periodically) on during active mode for redeeming the PG design from wear-out failure. The case of applying RBB on all STs in the nominal design (black line) is to demonstrate that, even though the aging "rate" of STs with RBB is slower, the overall performance degradation is still larger than the other cases due to its lower VV_{dd} at time 0. Accordingly, it does not make much sense to use RBB alone.

Figure 6 shows the relationship of the circuit lifetime versus the amount of bulk voltage (V_b) applied. By assigning a V_b greater than 1.0V (RBB) to redundant STs, the aging curve shifts toward the left slightly, implying a shorter lifetime. However, the difference is not significant so as to provide perfect opportunities of trading lifetime extension for leakage saving, which is more sensitive to V_b with an exponential relation. This is the key motivation of our methodology exploiting RBB to reduce the leakage current flowing through redundant STs.

Table 1 tabulates the experimental results of our NBTI-aware power gating methodology, where columns 2-4 correspond to the aforementioned 1st case (dotted line), the 3rd case (blue line), and the 4th case (red line) in Figure 5, respectively. Column 5 shows the leakage overhead incurred by redundant STs. Note that, if RBB is not applied, the leakage overhead will be approximately equal to the percentage of ST redundancy (25% in this experiment). To realize RBB, the bulk voltage (V_b) shown in the last column is assigned to all redundant STs. For example, as also depicted in Figure 5, the nominal PG design of circuit *AES* has a lifetime of **1.47** yrs. It is extended to **3.33** yrs if 25% ST redundancy is introduced but no scheduling is used. By employing the proposed scheduling framework based on round robin, the lifetime of power-gated *AES* can be further extended to **4.45** yrs, where the leakage overhead is **5.32%** with V_b = 1.5V assigned. On average across all benchmarks considered, we can achieve **3.04X**

Table 1. Optimization results of lifetime and leakage

Circuit	Lifetime (year)			Leakage overhead	V_b (V)
	Original power gating design	25% ST redundancy with RBB, without scheduling	25% ST redundancy with RBB & scheduling (proposed)		
AES	1.468	3.333	4.446	5.32%	1.50
alu2	1.083	2.509	3.314	8.46%	1.35
alu4	1.719	3.941	5.267	6.21%	1.45
C2670	2.835	6.407	8.675	4.56%	1.55
C7552	2.022	4.641	6.223	6.21%	1.45
s832	3.381	7.307	9.997	2.45%	1.75
s9234	1.259	2.919	3.867	8.46%	1.35
AVG.	**1.967**	**4.437**	**5.970**	**5.95%**	
	1.00X	**2.26X**	**3.04X**		

lifetime extension with only **5.95%** leakage overhead. In contrast to existing work [18] where 20-200% overhead in leakage power is incurred for 1.85X lifetime extension, our methodology reveals superior benefits by jointly (*i*) exploring NBTI recovery via ST redundancy (Section IV-B) and (*ii*) applying reverse body bias (Section IV-C). The area overhead, which comes from redundant STs, SRs, and VS (see Figure 4), is also small as compared to the whole circuit. Consider *AES* again, about **5%** area overhead is needed for the lifetime extension from 1.47 yrs to 4.45 yrs.

Finally, the impact of increasing ST redundancy on lifetime extension is demonstrated in Figure 7. Despite notable improvements (right shifts) in the circuit lifetime, the overhead in area and power is a major concern when a higher degree of ST redundancy is deployed. The use of 25% ST redundancy is a good tradeoff since it brings sufficient lifetime extension while keeping the leakage overhead below an acceptable extent.

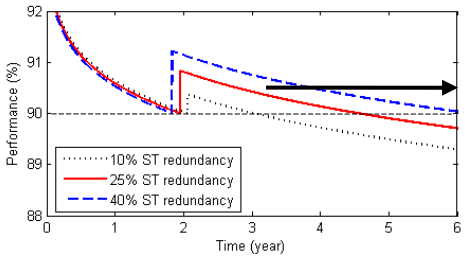


Figure 7. Lifetime vs. ST redundancy

VI. Conclusion

In this paper, we present an analysis and mitigation methodology in terms of NBTI-induced performance degradation for power-gated circuits. For the first time, joint modeling of interdependent degradation effects on logic networks and sleep transistors is included. Based on exploring NBTI recovery, redundant STs are introduced for mitigating the aging of STs and thus extending the lifetime of power-gated circuits. Furthermore, we formulate an optimization problem for concurrent lifetime extension and leakage saving, by applying RBB on redundant STs. Experiments demonstrate that the proposed methodology can accurately analyze the performance degradation and effectively extend the circuit lifetime.

References

[1] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing," *Journal of Applied Physics*, vol. 94, no. 1, pp. 1-18, July 2003.

[2] J. H. Stathis and S. Zafar, "The negative bias temperature instability in MOS devices: a review," *Microelectronics Reliability*, vol. 46, no. 2-4, pp. 270-286, Feb.-April 2006.

[3] S. Chakravarthi *et al.*, "A comprehensive framework for predictive modeling of negative bias temperature instability," in *Proc. of Int'l Reliability Physics Symp.* (IRPS), pp. 273-282, April 2004.

[4] N. Kimizuka *et al.*, "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," in *Proc. of Symp. on VLSI Technology*, pp. 73-74, June 1999.

[5] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "An analytical model for negative bias temperature instability," in *Proc. of ICCAD*, pp. 493-496, Nov. 2006.

[6] W. Wang *et al.*, "The impact of NBTI effect on combinational circuit: modeling, simulation, and analysis," *IEEE TVLSI*, vol. 18, no. 2, pp. 173-183, Feb. 2010.

[7] C. Long and L. He, "Distributed sleep transistor network for power reduction," in *Proc. of DAC*, pp. 181-186, June 2003.

[8] D.-S. Chiou, S.-H. Chen, and S.-C. Chang, "Timing driven power gating," in *Proc. of DAC*, pp. 121-124, July 2006.

[9] B. C. Paul *et al.*, "Temporal performance degradation under NBTI: estimation and design for improved reliability of nanoscale circuits," in *Proc. of DATE*, pp. 780-785, March 2006.

[10] K. Kang *et al.*, "Efficient transistor-level sizing technique under temporal performance degradation due to NBTI," in *Proc. of ICCD*, pp. 216-221, Oct. 2006.

[11] X. Yang and K. Saluja, "Combating NBTI degradation via gate sizing," in *Proc. of ISQED*, pp. 47-52, March 2007.

[12] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "NBTI-aware synthesis of digital circuits," in *Proc. of DAC*, pp. 370-375, June 2007.

[13] K.-C. Wu and D. Marculescu, "Joint logic restructuring and pin reordering against NBTI-induced performance degradation," in *Proc. of DATE*, pp. 75-80, April 2009.

[14] H. Dadgour and K. Banerjee, "Aging-resilient design of pipelined architectures using novel detection and correction circuits," in *Proc. of DATE*, pp. 244-249, March 2010.

[15] Y. Wang *et al.*, "Temperature-aware NBTI modeling and the impact of input vector control on performance degradation," in *Proc. of DATE*, pp. 546-551, April 2007.

[16] D. R. Bild, G. E. Bok, and R. P. Dick, "Minimization of NBTI performance degradation using internal node control," in *Proc. of DATE*, pp. 148-153, April 2009.

[17] A. Calimera, E. Macii, and M. Poncino, "NBTI-aware power gating for concurrent leakage and aging optimization," in *Proc. of ISLPED*, pp. 127-132, Aug. 2009.

[18] A. Calimera, E. Macii, and M. Poncino, "NBTI-aware sleep transistor design for reliable power-gating," in *Proc. of GLSVLSI*, pp. 333-338, May 2009.

[19] W. Wang *et al.*, "An efficient method to identify critical gates under circuit aging," in *Proc. of ICCAD*, pp. 735-740, Nov. 2007.

[20] M.-C. Lee *et al.*, "NBTI-aware power gating design," in *Proc. of ASP-DAC*, pp. 609-614, Jan. 2011.

[21] M.-C. Lee *et al.*, "An efficient wakeup scheduling considering resource constraint for sensor-based power gating designs," in *Proc. of ICCAD*, pp. 457-460, Nov. 2009.