Int J Parallel Prog (2009) 37:417–431
DOI 10.1007/s10766-009-0104-y

New-Age: A Negative Bias Temperature Instability-Estimation Framework for Microarchitectural Components

Michael DeBole · Ramakrishnan Krishnan · Varsha Balakrishnan · Wenping Wang · Hong Luo · Yu Wang · Yuan Xie · Yu Cao · N. Vijaykrishnan

Received: 15 December 2008 / Accepted: 14 April 2009 / Published online: 7 May 2009

Abstract Degradation of device parameters over the lifetime of a system is emerging as a significant threat to system reliability. Among the aging mechanisms, wearout resulting from Negative Bias Temperature Instability (NBTI) is of particular concern in deep submicron technology generations. While there has been significant effort at the device and circuit level to model and characterize the impact of NBTI, the analysis of NBTI's impact at the architectural level is still at its infancy. To facilitate architectural level aging analysis, a tool capable of evaluating NBTI vulnerabilities early in the design cycle has been developed that evaluates timing degradation due to NBTI. The tool includes workload-based temperature and performance degradation analysis across a variety of technologies and operating conditions, revealing a complex interplay between factors influencing NBTI timing degradation.

Keywords Microprocessor reliability · Negative bias temperature instability (NBTI) · NBTI framework · Reliable systems

1 Introduction

With transistor dimensions continuously shrinking in the nanoscale regime, traditional assumptions that transistor parameters will remain bounded by a certain margin will

M. DeBole (✉) · R. Krishnan · Y. Xie · N. Vijaykrishnan
Department of Computer Science and Engineering, Microsystems Design Laboratory,
The Pennsylvania State University, State College, PA 16802, USA
e-mail: debole@cse.psu.edu

V. Balakrishnan · W. Wang · Y. Cao
Nanoscale Integration and Modeling Group, Arizona State University, Tempe, AZ 85287, USA

H. Luo · Y. Wang
Circuit and System Division, Tsinghua University, 100084 Beijing, China

no longer be adequate for the reliability requirements in future technology generations. One the most prominent and persistent reliability concern for future CMOS technology is NBTI [1–8], which has the potential to increase the threshold voltage (V_{th}) of the pMOS device by up to 50 mV over a 10-year period. This can result in a reduction of circuit speed by more then 20% or, in extreme cases, cause functional failure [9,10].

The traditional methodology for extending circuit lifetimes is to design the circuit to reliably operate under the worst case conditions. This has led to a variety of techniques to increase the lifetime of microprocessors such as speed binning, burn-in testing and employing the use of guardbands.[1] However, in future technologies, such techniques may not be suitable for ensuring lifetime reliability requirements in future devices [11].

The growing concern of device failure due to NBTI has prompted a significant effort on the part of the research community to model NBTI's short-term and long-term effects. As a result, many researchers have begun to analyze these effects at the gate and circuit level in order to enable the evaluation of larger circuit structures [2,8,12]. Nevertheless, an approach which can analyze NBTI at the microarchitectural level has yet to evolve. In this paper a comprehensive approach to high-level aging estimation, dubbed *New-Age*, is introduced to enable the assessment of a microarchitecture and to provide an indication of its lifetime in the presence of NBTI.

The *New-Age* framework ties together a functional simulator, gate-level simulator, analog simulator, and state-of-the-art device models, in conjunction with timing and power/temperature analysis tools, to perform a comprehensive analysis of a circuits timing after NBTI's effects have been considered for varying workloads. In this work the capabilities of the *New-Age* framework are demonstrated by evaluating the lifetime of two designs, (a) pipeline stages of an out-of-order superscalar processor and (b) sub-components found within arithmetic logic units (ALUs). The capabilities of the *New-Age* framework include workload-based power, temperature, and performance degradation analysis, detailed path analysis, and the ability to evaluate designs across upcoming technology nodes.

The remainder of this paper is organized as follows: a summary of related work is presented in Sect. 2, a background on NBTI is given in Sect. 3, the New-Age framework is introduced in Sect. 4, case studies are presented in Sect. 5, and the concluding remarks are given in Sect. 6.

2 Related Work

The physics behind NBTI and its underlying phenomena has been widely investigated by researchers for many years, with more recent work appearing in [1,3–8]. This includes the emergence of NBTI studies and modeling for combinational and sequential circuit elements, along with memory elements. In [2,12] Wang et al.

[1] Guardbanding is a technique where the operating frequency is reduced in order to overcome any degradation that might be incurred over the lifetime of the circuit.

performed a study of NBTI on both sequential and combinational circuits, developing gate-level models based on the Berkeley Predictive Technology Model (PTM) [13]. The impact of NBTI on the performance degradation and SNM loss of SRAM cells was studied in [14]. In addition to understanding the underlying causes of NBTI, and its effects at both the physical and circuit level, work has also been done to improve a circuit's lifetime when faced with NBTI degradation. Authors in [15] studied the impact of NBTI on logic circuits and proposed mitigation techniques. NBTI aware synthesis was proposed in [16] to reduce the effect due to NBTI. Finally, the authors in [17] proposed a method to identify the critical gates and strengthen them. In addition, at the micro-architectural level, authors in [18] introduce techniques to reduce NBTI effects of structures found within microprocessors.

3 Preliminaries

3.1 NBTI Overview

Negative-bias-temperature instability (NBTI) is a result of the continuous generation of traps at the Si/SiO_2 interface of the pMOS transistor. The interaction between inversion layer holes and hydrogen passivated Si atoms breaks Si–H bonds created during the oxidation process, creating interface traps and neutral H atoms. These new H atoms can then form H_2 molecules, which either diffuse away from the interface through the oxide or can anneal an existing trap (Fig. 1). The general physical mechanism of NBTI is explained quantitatively through the reaction-diffusion model.

From a circuit standpoint this translates to the degradation of the pMOS transistor when its gate voltage is negative, or its input is "0". This effect is exacerbated by the fact that the Si–H bonds begin to break more easily over time leading to the pMOS device becoming more vulnerable to these physical effects. Broken bonds that do not anneal act as interface traps and result in an increase in the threshold voltage (V_{th}) of the transistor. This effectively causes a slowdown in the speed of the pMOS transistor, leading to reduced performance over time within digital circuits. In addition, NBTI is of greater concern as technology scales due to higher operating temperatures and the use of thin-oxides.

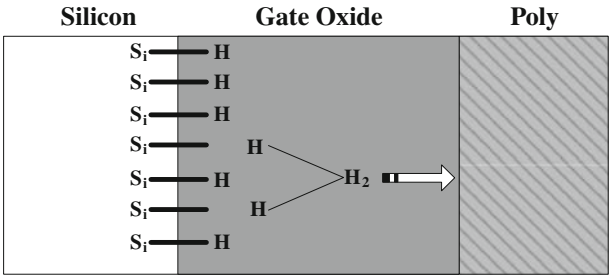

Fig. 1 The dissociation of Si–H bonds at the Si/SiO_2 interface

3.2 Device Model (Reaction-Diffusion)

NBTI can be described as the generation of interface charges (N_{it}) at the Si/SiO_2 interface. This process is described through the Reaction-Diffusion (R-D) model which consists of two critical steps:

(1) *Reaction*: The Si–H bonds at the Si/SiO_2 interface are broken under vertical electrical stress. Consequently, interface charges are induced, which cause an increase of V_{th}. Given the initial concentration of the Si–H bonds, i.e., N_0, and the concentration of the inversion carriers, i.e., P, the generation rate of N_{it} is given by [3]:

$$\frac{dN_{it}}{dt} = k_F\,(N_0 - N_{it})\,P - k_R N_{\mathrm{H}} N_{it}$$

With continued reaction, two H atoms combine to generate a H_2 molecule. The concentration of H_2, i.e., N_{H_2} is:

$$N_{\mathrm{H}_2} = k_{\mathrm{H}} N_{\mathrm{H}}^2$$

(2) *Diffusion*: The reaction generated species diffuse away from the interface toward the gate, driven by the gradient of the density. This process influences the balance of the reaction and is governed by:

$$\frac{dN_{\mathrm{H}}}{dt} = D_{\mathrm{H}} \frac{d^2 N_{\mathrm{H}}}{dx^2}$$

By integrating these two steps together, we can solve the models for V_{th} increases (ΔV_{th}), for both static and dynamic NBTI, as shown in Table 1, where A and K_v are functions of the vertical electrical field (T) and the carrier concentration ($C = \exp(-E_a/kT)/T_0$), respectively; δ, ξ_1 and ξ_2 are constant parameters.

There are also several fitting parameters in the model which are responsible for the degradation under various temperatures. These fitting parameters in the model may change from one technology to another, but they are relatively insensitive to local process variations at the transistor level. More details on the R-D model along with experimentally validated results can be found in [12].

Using V_{th}, as computed by the R-D model, a first-order approximation for the propagation delay of the gate (T_d) can be given by [19]:

$$T_d = a_0 + a_1 \cdot \Delta V_{th} + a_2 C_l$$

Table 1 ΔV_{th} model for both static and dynamic NBTI

Static		$A\left((1+\delta)\,t_{ox} + \sqrt{Ct}\right)^{2n}$
Dynamic	Stress	$\left(K_v\,(t - t_0)^{0.5} + \sqrt[2n]{\Delta V_{th0}}\right)^{2n}$
	Recovery	$\Delta V_{th0}\left(1 - \frac{2\xi_1 t_e + \sqrt{\xi_2 C(t-t_0)}}{2t_{ox} + \sqrt{Ct}}\right)$

where a_0 is the intrinsic delay of the gate without NBTI degradation, a_1 and a_2 are constants, and ΔV_{th} is the degradation due to NBTI in the pMOS transistor. Given that a gate can have multiple pMOS transistors, the NBTI degradation of each of the transistors can be different based on their inputs. Because characterizing the gate delay capturing these differences can become exceedingly time consuming two corner cases are considered. The first calculates the circuit's timing degradation where the ΔV_{th} is based on the least-most degraded PMOS transistor. The second considers the most degraded PMOS transistor in the gate to determine the circuit's timing. The constants in the above equation are obtained through HSPICE circuit simulations which are used to create a library of primitive cells. Each of the gates within the cell libraries have their nominal and degraded timing characterized for various temperatures, supply voltages, and across several technology nodes (65, 45, and 32 nm).

It is important to note, before going further, that the subsequent tool and results discussions is dependent upon the accuracy of the underlying models. Furthermore, even though the current models have been validated experimentally they may also change in the future. Therefore, another important aspect in the design of the framework was to allow for future integration of new device models.

4 Aging Assessment

The evaluation of NBTI stress on modern microprocessor architectures requires an integrated framework that consists of architectural simulation, gate level timing analysis, and pre-characterized gate libraries. In this section, an approach to synthesis-level evaluation is presented.

Figure 2 depicts the complete framework, including *New-Age*, when configured for microarchitectural studies, tying together the basic *New-Age* analysis tool and the microarchitectural simulation framework. The separation of the analysis toolflow from the microarchitectural design and simulator allows for other microarchitectural simulators to be integrated without requiring any major changes to the overall framework.

In order to simulate the aging due to NBTI, the *New-Age* analysis tool requires several parameters, such as operating conditions (temperature, voltage, and time ranges), the technology node, and the design files (HDL or a specified netlist of the design). In addition, input value probabilities are needed for determining a gate's input switching used to later determine degradation of a particular gate in the netlist.

Using the operating conditions and specified technology the tool begins by performing library characterization if needed. These custom cell libraries are created automatically for the various operating conditions using the first-order approximation to gate delay and NBTI models as described in Sect. 3. The cell libraries contain nominal cells with gate delays obtained by setting $\Delta V_{th} = 0$ in the first order gate delay model to serve as a baseline and degraded cells with gate delays based the appropriate ΔV_{th} resulting from the gates input probabilities. It is important to note that since library characterization can be time consuming it is only carried out with newly specified operating conditions, technologies, or changes to the degradation models. Therefore, for subsequent runs of a design, the libraries do not need to be re-characterized each time.

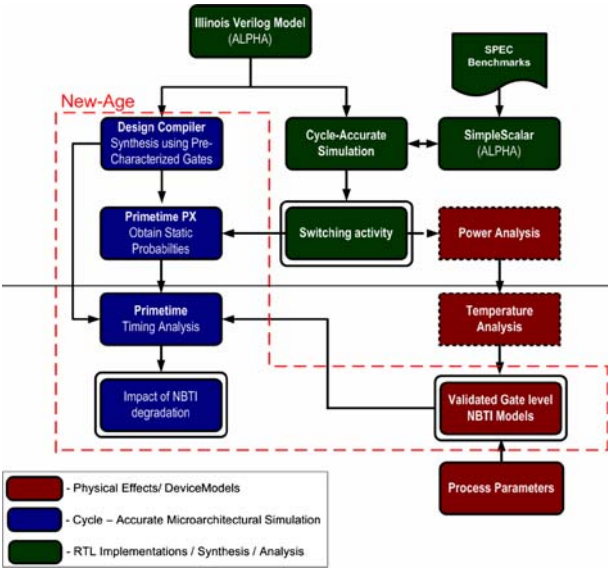

Fig. 2 NBTI microarchitectural assessment framework

Following cell library creation, the tool then performs synthesis if HDL of the component is received. Synthesis is performed using the cells with nominal delay to produce a netlist of the desired component. If a netlist is specified, synthesis is skipped and the tool proceeds immediately to propagating the top level switching probabilities to the internal nodes of the netlist. As a result of signal probability propagation, two netlists for aging analysis are created corresponding to the lower and upper bounds for NBTI vulnerability. For the lower bound analysis, the static probability for each of the inputs of a gate is set to correspond to that of its least stressful input, the input with the smallest static probability of 0. Similarly, for the upper bound analysis, each of the inputs of a gate corresponds to that of its most stressful input, the input with the highest static probability of 0.

This results in three final netlists corresponding to the baseline non-degraded design, the worst case design, and the best case design. The best-case netlist assigns probabilities to the gate's input which minimizes the delay through a particular gate. Conversely, the worst-case netlist assigns probabilities to the gate's input which maximizes the delay through a particular gate. The netlists corresponding to the baseline non-degraded design and the worst case and best case aged designs are fed to Synopsys Primetime to perform static timing analysis using the custom cell libraries. The timing analysis reveals the increase in critical path of the circuit as well as the reduction of slacks in the non-critical paths. In some cases our analysis revealed the emergence of a new critical path (a non-timing critical path changing to a critical path) due to aging.

As shown in Fig. 2 the static signal probabilities of the internal nodes and the gate temperature are obtained using a hybrid simulator consisting of the Illinois Verilog Model (IVM), a cycle-accurate register transfer level (RTL) model of the processor and Simplescalar [20], a functional simulator. The RTL model is used for monitoring

the static probabilities of the signals at each of the internal nodes of a circuit while the functional simulator is used for accelerating the architectural states to desired sample points for fast simulation speeds. The RTL model augments the IVM to capture the switching probabilities of all the nodes in the design. The switching activity is translated to input state probability as well as used for power estimation. The power estimates along with the architectural layout are fed to a thermal estimation tool, Hotspot 4.0 [21], to estimate the temperature. Once all the operating conditions and input state probabilities have been obtained they are then provided to *New-Age* to perform the final aging analysis of the design.

Separating the framework in this way enables the simulation of different architectural configurations and application suites while easily allowing for alternative simulators and architectures to be used. Further, it permits the exploration of changes in temperature and voltages. These features can be useful for evaluating the NBTI degradation in conjunction with circuit and architectural support designed for other constraints such as power consumption.

5 Experimental Results

In this section, *New-Age* is used to carry out NBTI aging analysis for two designs. First, the *New-Age* simulation framework is targeted at a processor architecture similar to that of the Alpha 21264 [22] and the AMD Athlon [23]. The processor architecture is based on the IVM, a synthesizable cycle-accurate register transfer level (RTL) model [24]. Within the IVM several pipeline stages are examined for NBTI vulnerabilities. In the study, workload driven analysis, including impacts of temperature, voltage, and technology scaling on NBTI-induced degradation are performed. Second, several sub-components found within a typical ALU have their lifetimes evaluated.

5.1 Evaluation of Processor Pipeline

This section provides the evaluation of the Illinois Verilog Model (IVM), a superscalar, dynamically-scheduled pipeline which executes a subset of the Alpha instruction set. The IVM consists of several pipeline stages which are detailed in [24]. Within the IVM, the NextPC, Fetch, Decode, and Rename stages are considered for vulnerability analysis. Within those stages memory structures were removed and replaced with the corresponding activity traces to overcome limitations due to the synthesis tools. Each of the stages was evaluated for NBTI under several conditions including temperature, supply voltage, and over technology generations. The SPEC2000 benchmark suite [25] was used for providing workloads for processor evaluation.[2]

Figure 3 depicts the average NBTI-induced delay degradation found through *New-Age* for the IVM pipeline stages at the 65 nm technology node after 10 years of operation. Also reported from the analysis are the best- and worst-case delay degradation ranges as described in Sect. 4. The results show that at 65 nm Fetch1 undergoes

[2] Traces from four representative benchmarks are presented for clarity: Art, Bzip2, Crafty, and Mesa.

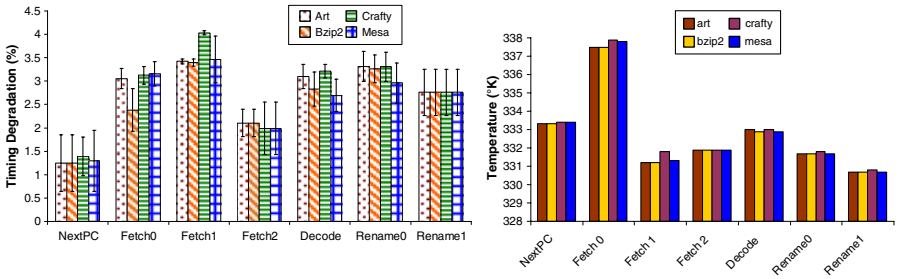

Fig. 3 NBTI-induced delays on pipeline stages within IVC for 65 nm technology and across temperature ranges with derived workload-dependent steady state temperature and NBTI-induced 10 year timing degradation. An ambient temperature of 300 K was used for Hotspot

3.5% degradation on average while Fetch2 (the highest temperature unit) sees only 3%. Average workload-dependent temperature results for these stages have also been reported and are shown in Fig. 3. This graph highlights that a ±2.5% variation in temperature leads to no visible correlation between a structures temperature and its corresponding degradation. For example, Fetch1 which reports the lowest temperature surprisingly has the highest amount of degradation. Alternatively, NextPC, which has an elevated temperature due to its location in the floorplanned design, experiences the least amount of degradation of all the structures. This implies that while temperature is an important parameter for contributing to the transistors V_{th} shifts, an accurate NBTI characterization tool must take into account factors, such as input switching, that also contribute to the degradation of the circuit.

In addition to performance degradation and temperature analysis the *New-Age* framework further allows for more in-depth analysis of a given circuit. This includes critical path analysis before and after degradation effects are considered. This is a fundamentally important aspect of the tool because it was recently found in [17] that critical and sub-critical paths are those which typically define the critical path after degradation effects are included in the delay measurements. Using *New-Age* it is possible to analyze these paths for the pipeline stages in a given microarchitecture and determine how structures are affected across multiple applications.

5.1.1 New-Age Predictive Analysis

New-Age also includes several features which can be used to assist circuit designers determine the vulnerability of their designs to various conditions. These features allow for degradation analysis of circuits under various supply voltages, technology nodes, and temperatures.

(1) *Effects Due to Supply Voltage*: A common design technique to reduce power consumption is using multi-V_{dd}, a technique where circuits operate at several lower supply voltages within the same chip. In addition to helping reduce temperatures, reducing the supply voltages also affects the amount of NBTI degradation a circuit undergoes. Figure 4a shows the maximal performance degradation due to NBTI for

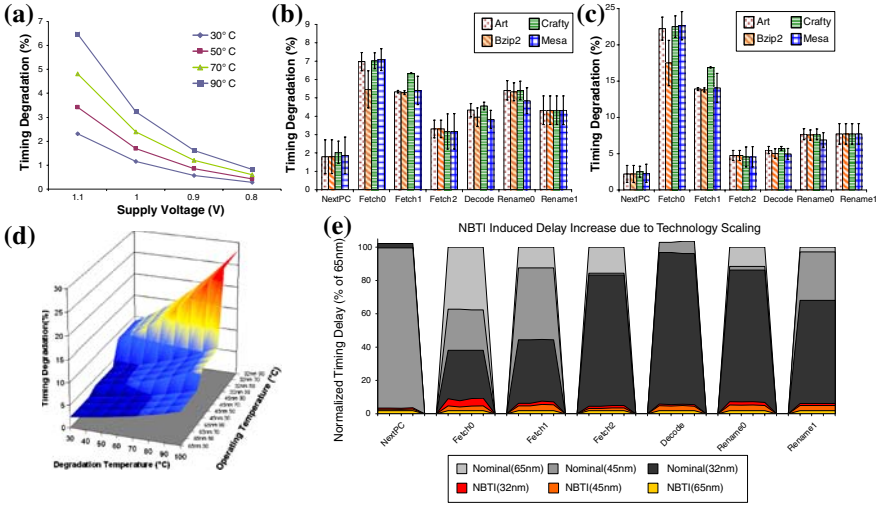

Fig. 4 Impact on NBTI degradation due to supply voltage and technology scaling. **a** Voltage scaling effects on NBTI degradation for Fetch1 stage in 65 nm technology. **b** NBTI-induced degradation for select stages within the IVM CPU core in 45 nm technology. **c** NBTI-induced degradation for select stages within the IVM CPU core in 32 nm technology. **d** Temperature effects on NBTI-induced degradation for Fetch1 stage across technologies. **e** Proportionality of NBTI-induced degradation to overall critical path delay across 65, 45, and 32 nm technology nodes

the Fetch1 stage as the supply voltage is reduced from 1.1 to 0.8 V operating at several temperatures, for the 65 nm node.

It can be observed that a decrease in supply voltage can help mitigate the performance degradation for a range of temperatures. In all pipeline stages there was approximately 30, 50, and 70% improvements in the amount of performance degradation as voltage scaled from 1.1 to 0.8 V.

(2) *Technology Scaling*: Figure 4b, c show the NBTI degradation results for the seven pipeline stages for 45 and 32 nm technologies respectively with technology dependent parameters scaled according to recent projections by ITRS [26]. The results from *New-Age* show that Fetch0 now exceeds the degradation of Fetch1, with performance degradation reaching 7.5 and 23.5% for 45 and 32 nm technologies respectively. In addition, the results also show a disproportionate increase in the amount of performance degradation for the Fetch0 and Fetch1 stages over the other stages as technology scales down from 45 to 32 nm. Figure 4e explains these results. Figure 4e depicts the change in critical path delay for each of the pipeline stages delay as technology scales across several of the benchmarks, with the delay being normalized to 65 nm. Included in the measurements are the effects of scaling on a stages power density and temperature, used in conjunction with updated cell libraries for each technology, to account for changes in the corresponding gate delays. The graph also shows the proportion of NBTI degradation superimposed onto the delay of each structure. This shows that as technology scales, the baseline critical path delays change non-uniformly from one stage to the next and, in most cases, decrease when compared to the previous technology. At the same time, the amount of delay due to NBTI increases as technology scales,

resulting in certain structures becoming more sensitive to NBTI degradation. This is what is observed for the Fetch0 and Fetch1 stages. As technology scaled these two structures had critical path delays which significantly reduced, while simultaneously they had NBTI delays which increased the most. This resulted in those two structure's performance degradations being increased up 23.5 and 18% respectively overall.

(3) *Effects Due to Temperature*: Figure 4d shows timing degradation of the Fetch1 stage across various degradation and operating temperatures, and with different technologies. This graph assumes that for the first 50% of the circuits lifetime (5 out of 10 years) the circuit is operating at one temperature undergoing degradation (along the *x-axis*, denoted degradation temperature) while for the remaining period of its lifetime it is operating at another temperature (along the *y-axis*, denoted operating temperature). This demonstrates how temperature can exacerbate NBTI-induced timing degradation when extremes in temperature are considered. This differs from the temperature results seen so far, where small variations in temperatures are shown across structures, but peek temperature and hotspot phenomena are absent. These results indicate that purely reducing peak operating temperatures may not be adequate for achieving longer lifetimes. For example, at 32 nm, even when the Fetch1 stage only reaches temperatures between 30 and 50°C during its lifetime, the circuit still undergoes significant timing degradation in the range of 10–15%. This implies that other techniques need to be explored which can further help reduce the effects of NBTI, such as supply voltage scaling and supply gating.

5.2 Evaluation of ALU Sub-Components

In this section, several of the common components which make up an ALU are investigated for NBTI vulnerabilities. Here we demonstrate the frameworks ability to quickly analyze different designs under a variety of operating conditions. For this experiment the main microarchitectural simulator is removed and the *New-Age* tool is supplied with HDL of custom implementations of several adders, multipliers, and shifters.[3] For demonstration purposes the operating temperature was set at 100°C, and results were gathered for 65, 45, and 32 nm technologies.

(1) *Performance Degradation*: The performance degradation of ALU components is depicted in Fig. 5. with results shown for a constant circuit operation of 10 years. Adders and multipliers have a similar amount of performance degradation, with less than a $\pm 3\%$ variation amongst them, while the log shifters did not follow this trend and instead had different amounts of degradation depending on the operand size (i.e. 16, 32, or 64 bits). The critical paths of the log shifter were further examined and it was found that this resulted from changes in the critical path and how it affected the input value probabilities at the internal nodes. In particular, the 16- and 32-bit shifter had the same gate sequence on the critical path. However, between these two circuits the switching at the internal nodes was not the same, ultimately leading to differences in performance degradation. After performing a more in-depth analysis of the paths

[3] 16/32-bit Brent Kung and Kogge Stone adders; 4/8/16 array, and 9×9 Booth multipliers; 4/8/16 parallel multipliers; 16/32/64 bit log-shifters.

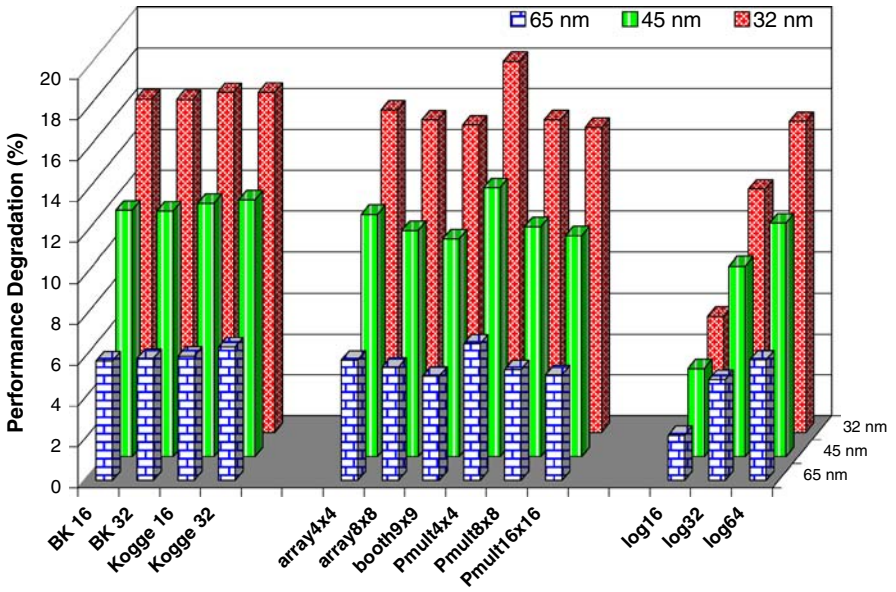

Fig. 5 NBTI-induced path delay across ALU components

within each log shifter it is possible to conclude that the 32-bit shifter is more vulnerable to NBTI-induced failure then the 16-bit shifter. The 64-bit log shifter proved to be even more vulnerable where the gate sequence, critical path length, and internal node switching changed.

(2) *Input-Vector Control*: One technique that has been frequently mentioned as a way to mitigate the effects from NBTI is input vector control (IVC). IVC is a technique where specific inputs are purposely applied to a circuit in order to minimize the number of internal nodes which will degrade. To this end three criteria are used to measure IVC's performance: *Capacity—The potential for reducing degradation*; *Effectiveness—The amount of degradation improvement*; *Component Idleness—How often IVC can be applied*.

Capacity: Figure 6 shows an alternate view for the performance degradation of the ALU components considering that they are idle for some portion of their lifetime. This idle period is identified as the active-to-standby ratio (RAS), which is defined as the fraction of the time the circuit spends in active mode to the time the circuit is idle. In this analysis a large idle period for RAS (1/9) was chosen, which is effectively 1 active year per 10 years of operation. During active-mode all inputs were assumed to have a .5 input probability of one, while in standby-mode two alternatives we considered. The first option is idle degradation analysis, defined to be when all internal nodes are set to '1'. The second possibility is always degrading degradation analysis, corresponding to all internal nodes being a '0'. Note that the timing degradation is not linear with stressing. It is most dominant in the initial stage [2].

To evaluate IVC's *capacity*, the difference between the amount of degradation for the best and worst cases is calculated. These results show that IVC does have the

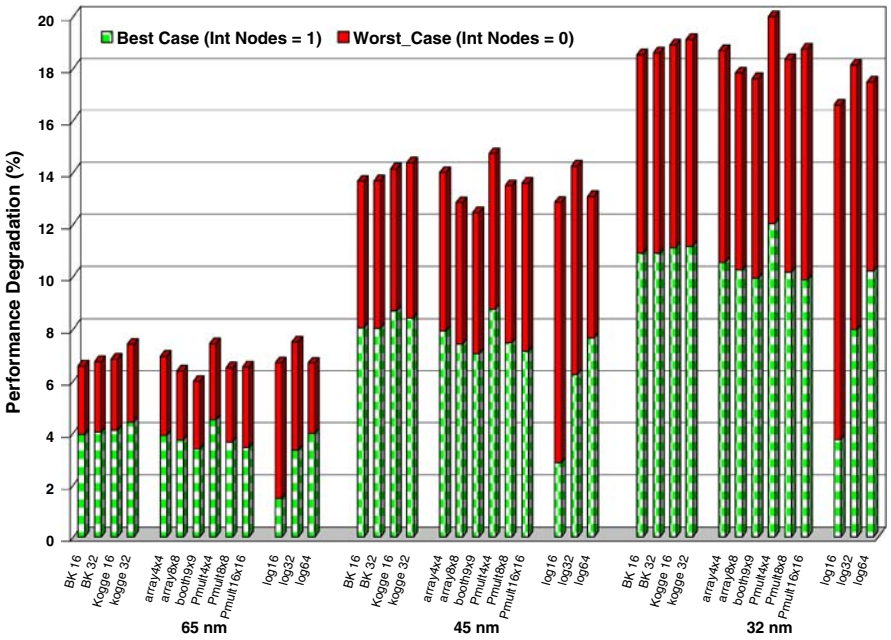

Fig. 6 NBTI-induced path delay across ALU components and technology with RAS

capability to cause improvement in the amount of performance degradation which stays relatively constant in most cases. More specifically, "perfect" IVC can result in up to 40% improvement for adders and multipliers and up to 70% improvement for the 16-bit log shifter. This consistency exists across modules and technologies because, under constant operating conditions, the active-to-standby ratio is effectively amplifying or reducing the effects of NBTI proportionally to the RAS value. This causes the improvement due to IVC to be very optimistic since the idle period's impact on a particular component's power and temperature is not taken into account.

Effectiveness: The performance of IVC is also impacted by the amount of degradation that still exists after IVC is applied. Referring back to Fig. 6, the ideal effectiveness of IVC can be associated to the best-case degradation. These results show that the amount of degradation in most circuits can be reduced to no less then 7% for 45 nm and 10% for 32 nm. In this case IVC proves to be effective, depending on the size of the initial guardband. For example, if large guardbands were used (tolerating 10% degradation) then the component would be able to function longer. However, if IVC was intended to be used as a way to reduce the guardband, the lifetime of the device will still suffer. More significantly, guardbands would still have to increase across technology generations since performance degrades more overall.

Component Idleness: Figure 7 shows results for the Kogge Stone 32-bit adder, 16 × 16 Parallel Multiplier, and 16-bit log shifter at 45 nm, as the active-to-standby ratio is varied from 1/9 to 9/1. The most immediate observation is that decreasing the amount of idle time decreases the amount of improvement. This is an expected result because if no time exists in which the input vector can be applied then there is no

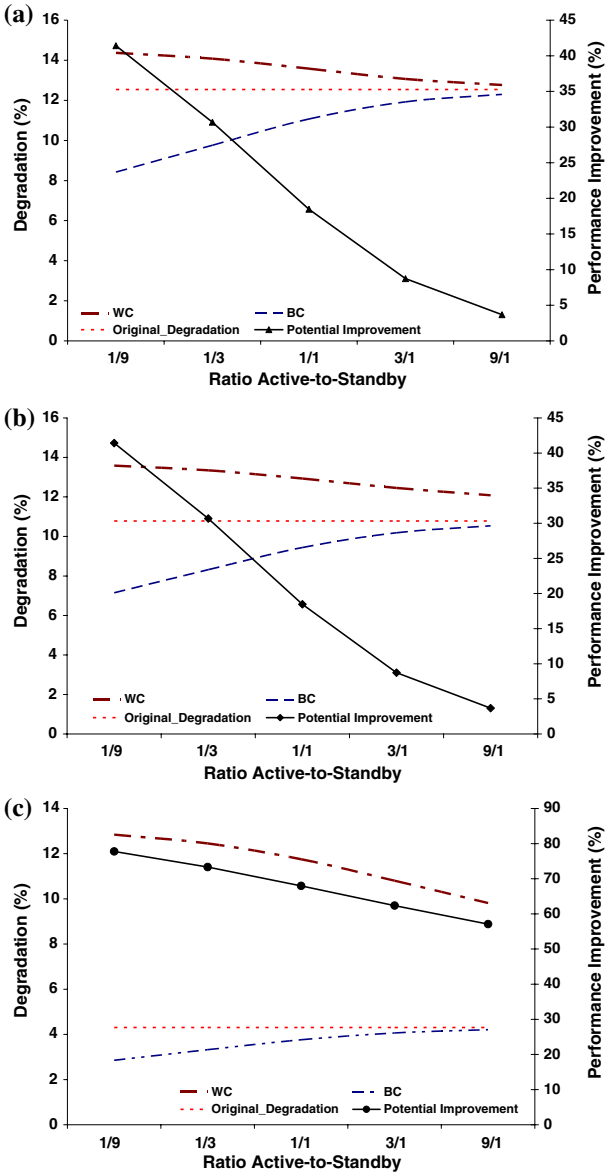

Fig. 7 Performance degradation with variation in RAS for **a** 32-bit Kogge stone Adder, **b** 16 × 16 parallel multiplier, **c** 16-bit log shifter

opportunity for improvement. More interesting is the point at which IVC will no longer yield any effective improvement. For the Kogge Stone adder and 16 × 16 parallel multiplier, this point appears to be reached at an active-to-standby ratio of 1/1. This corresponds to around 20% performance improvement, where best-case degradation decays to 10% degradation. In the case of the 16-bit log shifter, the break-even point is

different. First, observe that in general the best-case and worst-case degradation due to IVC grow towards the original degradation reported as the active-to-standby ratio decreases. In the case of the log shifter, the original degradation was more indicative of the best-case degradation, so the ratio between best and worst case remained large. Following trends on how the other components degrade it would be expected that most modules will see similar IVC improvement drop off as with the Kogge Stone adder and parallel multiplier.

Taking these three factors into consideration a conclusion on the performance of IVC can be discussed. On the one hand the capacity for improvement clearly exists. It was shown that at 45 and 32 nm that there can be a 40–70% improvement in a circuit's degradation if an ideal input vector had been applied. However, the resulting performance degradation was not sufficient to reduce guardbands, only sustain them for a particular technology node, given that they had a significant amount of idle time. In addition, the minimal guardbands from 65 nm technology could not be sustained across 45 and 32 nm technology generations. This makes the use of larger and larger guardbands necessary for smaller technology generations, defeating the goal of IVC, to reduce degradation and effectively, the required guardband.

Overall, IVC appears to be a short-term solution. In the cases where it could be applied a thorough knowledge of the workload being run along with a prior knowledge of path degradation must be known. Further, in cases where a critical component is not significantly idle, effort would need to be placed around increasing its idle periods using various hardware and software techniques. Even under these circumstances there is no guarantee that IVC will provide the savings necessary to meet lifetime requirements once the assumption that there exists a perfect input vector is removed. Finally, a potentially defeating factor of IVC as a long-term solution is that it is unable to reduce, or sustain, the amount of degradation from one technology generation to the next.

6 Conclusions

This work has introduced *New-Age*, an aging assessment framework which has the capability of examining arbitrary microarchitectural components and evaluating them for performance degradation. Its use was demonstrated on two-case studies. First, *New-Age* is used to examine several pipeline stages of a superscalar processor under various operating conditions and benchmark behaviors for NBTI-induced aging. Second, the effects of NBTI on common ALU components were investigated and evaluated.

Acknowledgements This project was supported by SRC's GSRC Focus Center and NSF Grants 0702617, 0454123, 90207002, and 60870001. The authors would also like to thank Toyota for a generous donation towards this project.

References

1. Kaczer, B., Arkhipov, V., Jurczak, M., Groeseneken, G.: Negative bias temperature instability (NBTI) in Si02 and SiON gate dielectrics understood through disorder-controlled kinetics. Microelectron. Eng. **80**, 122–125 (2005). doi:10.1016/j.mee.2005.04.054

2. Wang, W.: The impact of NBTI on the performance of combinational and sequential circuits. In: ACM/IEEE Design Automation Conference (DAC) (2007)
3. Alam, M.A., Mahapatra, S.: A comprehensive model of PMOS NBTI degradation. Microelectron. Reliab. **45**, 71–81 (2005). doi:10.1016/j.microrel.2004.03.019
4. Mahapatra, S., Kumar, P.B., Alam, M.A.: Investigation and of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs. IEEE Trans. Electron. Dev. **51**, 1371–1379 (2004). doi:10.1109/TED.2004.833592
5. Zafar, S., Lee, Y.H., Stathis, J.: Evaluation of NBTI in HfO2 gate-dielectric stacks with tungsten gates. IEEE Electron. Dev. Lett. **25**, 153–155 (2004)
6. Alam, M.A.: A critical examination of the mechanics of dynamic NBTI for pMOSFETS. In: IEDM Technical Digest, pp. 346–349 (2003)
7. Varghese, D., Mahapatra, S., Alam, M.A.: Hole energy dependent interface trap generation in MOSFET Si/SiO2 interface. IEEE Electron. Dev. Lett. **26**, 572–574 (2005). doi:10.1109/LED.2005.852739
8. Chakravarthi, S., Krishnan, A.T., Reddy, V., Machala, C.F., Krishnan, S.: A comprehensive framework for predictive modeling of negative bias temperature instability. In: Proceedings of the IEEE International Reliability Physics Symposium, pp. 273–282 (2004)
9. Borkar, S.: Electronics beyond nano-scale CMOS. In: ACM/IEEE Design Automation Conference (2006)
10. Schroder, D.K., Babcock, J.A.: Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing. J. Appl. Phys. **94**, 1–18 (2003)
11. Borkar, S.: Designing reliable systems from unreliable components: the challenges of transistor variability and degradation. IEEE Micro **25**, 10–16 (2005). doi:10.1109/MM.2005.110
12. Wang, W., Reddy, V., Krishnan, A.T., Krishnan, S., Cao, Y.: An integrated modeling paradigm of circuit reliability for 65nm CMOS technology. In: CICC (2007)
13. Zhao, Y.C.W.: New generation of predictive technology model for sub-45nm early design exploration. IEEE Trans. Electron. Dev. **53**, 2816–2823 (2006). doi:10.1109/TED.2006.884077
14. Kumar, S.V., Kim, C.H., Sapatnekar, S.: Impact of NBTI on SRAM read stability and design for Reliability. In: Proceedings of the IEEE International Symposium on Quality Electronic Design (ISQED'07) (2006)
15. Paul, B.C., Kang, K., Kufloglu, H., Alam, M.A., Roy, K.: Temporal performance degradation under NBTI: estimation and design for improved reliability of nanoscale circuits. In: Proceedings of the Design, Automation and Test in Europe (DATE) (2006)
16. Kumar, S.V., Kim, C.H., Sapatnekar, S.: NBTI-Aware Synthesis of Digital Circuits. In: Proceedings of the Design Automation Conference (DAC) (2007)
17. Wang, W., Wei, Z., Yang, S., Cao, Y.: An efficient method to identify critical gates under circuit aging. In: Proceedings of the International Conference on Computer Aided Design (ICCAD) (2007)
18. Abella, J., Vera, X., Gonzalez, A.: Penelope: the NBTI-Aware processor. In: International Symposium on Microarchitecture (MICRO' 07) (2007)
19. Sakurai, T., Newton, A.R.: Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. Solid State Circ. IEEE J. **25**, 584 (1990). doi:10.1109/4.52187
20. www.simplescalar.com
21. Huang, W., Sankaranarayanan, K., Ribando, R.J., Stan, M.R., Skadron, K.: An improved block-based thermal model in HotSpot 4.0 with granularity considerations. In: Proceedings of the Workshop on Duplicating, Deconstructing, and Debunking, in conjunction with the 34th International Symposium on Computer Architecture (ISCA), June (2007)
22. Gieseke, B.A.: A 600MHz superscalar RISC microprocessor with out-of-order execution. In: 1997 IEEE International Solid-State Circuits Conference Digest of Technical Papers (1997)
23. Meyer, D.: AMD-K7$^{(TM)}$. Technology Presentation. Advanced Micro Devices, Inc., Sunnyvale, CA (1998)
24. Wang, N.J., Quek, J., Rafacz, T.M., Patel, S.J.: Characterizing the effects of transient faults on a high-performance processor pipeline. In: Proceedings of the 2004 International Conference on Dependable Systems and Networks (DSN), Florency, Italy (2004)
25. www.spec.org
26. http://www.itrs.net/