Reliable CMOS VLSI Design Considering Gate Oxide Breakdown

Kyung Ki Kim
School of Electronic and Electrical Engineering
Daegu University
Daegu, South Korea
e-mail: kkkim@daegu.ac.kr

***Abstract*—As technology scales down into the nanometer region, the reliability mechanism caused by time dependent dielectric breakdown (TDDB) has become one of the major reliability concerns. TDDB can lead to performance degradation or logic failures in nanoscale CMOS devices, and can cause significant increase of leakage power in the standby mode. In this paper, the TDDB effects on the delay and power of the nanoscale CMOS circuits are analyzed using ISCAS85 benchmark circuits that are designed using a 45-nm CMOS predictive technology model. Based on the TDDB analysis, a reliable CMOS VLSI design methodology using a redundancy system has been proposed.**

Keywords-TDDB; Reliability; Gate oxide breakdown; Time dependent dielectric breakdown; Aging effect

I. INTRODUCTION

As MOSFET [1] technology is scaled down more aggressively, various aging phenomena (or reliability mechanisms) such as negative bias temperature instability (NBTI), hot carrier injection (HCI), and time-dependent dielectric breakdown (TDDB) on the MOSFET device have become one of the most important issues in the nanoscale MOSFET technology [1]. These mechanisms lead to device aging, resulting in performance degradation and eventually design failures during the expected system lifetime [2]-[4]. Recently, the oxide thickness of less than 2nm is common in state-of-the-art technologies, and TDDB becomes one of the key challenges among these reliability mechanisms [5][6].

Moreover, the saturating trend for supply voltage scaling causes a large electric field in the gate oxide, which generates gate tunneling currents. The lifetime of a particular gate oxide thickness is determined by the total amount of charge that flows through the gate oxide by the tunneling current. Therefore, nanometer devices are more prone to oxide breakdown compared to micrometer devices. The oxide breakdown is categorized into hard breakdown (HBD) and soft breakdown (SBD): HBD causes a catastrophic failure of the device and the entire circuits. On the other hand, SBD leads to parametric variations such as energy, delay, and noise margin of a gate although it does not bring about functional failures [5].

Figure 1 shows the variation of leakage path current depending on the wear-out and breakdown phase for thin gate oxides. At the end of the phase I, the traps start to be increased. During Phase II, moving traps lead to random fluctuation in the leakage current, power and delay, where the device is still functional. In phase III, the conduction path is created in the oxide, the leakage current is exponentially increased, and then, finally, the conduction path causes a catastrophic failure of the device [6].

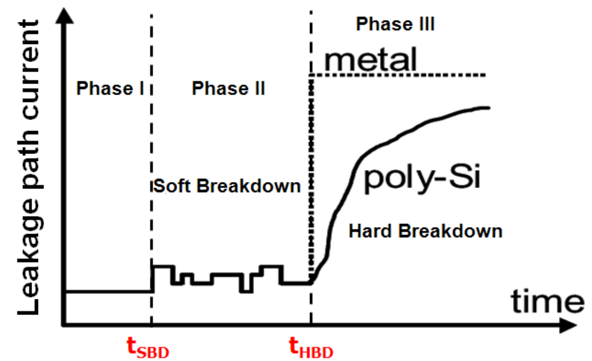

Figure 1. Wear-out and breakdown model for thin gate oxides[6].

TDDB is typically treated statistically using the Weibull distribution [7]. For this reason, usually a large gate oxide area has to be used in order to be able to detect the breakdown. Gate oxide breakdown manifests itself as an increase of gate current while retaining its insulating properties (soft break down-SBD or hard break down-HBD) that happens when I-V curve becomes linear, i.e., manifests a resistor-like behavior. For thinner gate oxides ($\leq 30\times10^{-10}$m), SBD is the most likely event. In this paper, TDDB effect only in NMOS devices is considered because the probability of PMOS gate-oxide breakdown is at least an order less than that of NMOS breakdown [8].

Recently, many researches on the TDDB effect have been proposed, but they have mainly focused on device physics not circuit-level analysis. Although a few researches on the circuit-level analysis of TDDB in nanoscale CMOS circuits have been proposed, they have concentrated on mathematical modeling and performance degradations only in simple digital circuits. In this paper, we propose a new simple reliable CMOS VLSI design methodology using a redundancy system.

The remaining of the paper is organized as follows. Section II introduces the TDDB modeling in nanoscale

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology. (2011-0014255)

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-213-4

CMOS circuits. Section III shows our proposed reliable system, followed by the conclusion in Section IV.

II. TDDB MODELING

At current operating voltages for MOSFETs, the degradation mechanism can be approximated by an exponential function of voltage applied across gate oxide and Arrhenius function of temperature as follows:

$$TDDB = C \cdot \exp(\frac{E_a}{kT}) \cdot \exp(-\beta V_g) \quad (1)$$

where V_g is the voltage across gate oxide, E_a is the activation Energy, k is the Boltzman's constant, T is the junction temperature, C and β are the technology specific constants.

Gate oxide breakdown is sensitive to:

- Voltage across the gate oxide (V_g): The higher V_g and the higher V_g duty cycles the shorter TDDB. For this reason, worst case condition for gate oxide occurs when the device is operated in DC mode (100% duty cycle).
- Junction temperature
- Gate oxide thickness: The thinner the gate oxide the more difficult it is to get good quality oxides and interfaces.
- Gate oxide area: The bigger the gate oxide area the higher the gate oxide breakdown induced failure rate. SRAMs typically have intensive gate oxide area and can be good vehicles to test gate oxide lifetime.

In this paper, for post-breakdown analysis at the circuit-level, a MOSFET with the oxide breakdown is modeled using two breakdown resistors (R_{BD}) as shown in Figure 2 [8]. The R_{BD} value ranges from GΩ (no TDDB) to a few hundreds of KΩ (HBD). The time-dependent gate to source/drain resistor model was experimentally verified in Lombardo et al. [9].

The worst stress for NMOS and PMOS gate oxide happens when NMOS and PMOS are in ON state. For this reason, in calculating product the gate oxide FIT, we make the assumptions that duty cycle of the voltage across gate oxide is 50%. However, when PMOS and NMOS are off, respectively gate/drain and gate/source overlaps are stressed, i.e, gate oxide in overlap region is continuously stressed. The FIT contribution from this stress condition is small due to the relatively smaller gate oxide area in the overlap regions.

Based on time dependent dielectric breakdown data (TDDB), there is a maximum voltage across gate oxide that a given technology can support, for a given gate oxide area and failure rate. The oxide thickness, for a technology, is determined so that this voltage is not exceeded, in a dc sense, for any circuit design. However, when signals are driven, there is a certain amount of overshoot/undershoot in the waveform that can result in accelerated gate oxide wear-out and lower reliability, if appropriate limits are not established.

In Figure 3, it is assumed that the NMOS (MN1) of the 1st inverter is stressed by the gate oxide breakdown; while Input signal is changed from logic “0” to “1’, the NMOS of

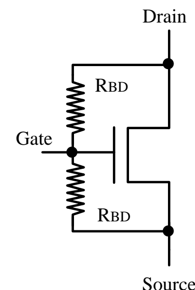

Figure 2. TDDB model for NMOS

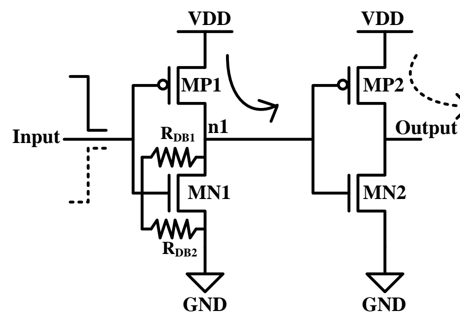

Figure 3. TDDB resistance in NMOSFET

the 1st inverter is turned on; the node n1 capacitor starts to be discharged through MN1, but the discharge time is longer than the normal discharging time because of the breakdown resistor. As a result, this makes the charging time at Output node longer, which means the propagation time from Input to Output will be longer comparing to the normal case without TDDB. The propagation time depends on the breakdown resistor value: if the resistor is large, the charging time at Output node will be a little increased; however, if the resistor is small, the charging time at Output node will be much more increased due to the small voltage-swing at node n1 caused by the small breakdown resistor.

On the other hand, while Input signal is changed from logic “1” to “0’, the PMOS of the 1st inverter is turned on; the node1 capacitor starts to be charged through MP1; however, the rising time and rising voltage at node n1 are affected by the breakdown resistor. As a result, this makes the charging time at Output node shorter, which means the propagation time from Input to Output will be shorter comparing to the normal case without TDDB. In this case, if the resistor is large, the discharging time at Output node will be much more decreased because the a small voltage-swing at node n1 (caused by the breakdown resistor) moves up the high-to-low transition at Output node; however, if the resistor is too small, the charging time at Output node will be a little decreased because less voltage-swing (compared to the voltage-swing when the resistor is large) at node n1 makes the current driving force of the 1st inverter weak.

Putting the abovementioned two transitions together, if the breakdown resistor is small, the propagation time from Input to Output will be increased; if the breakdown resistor is

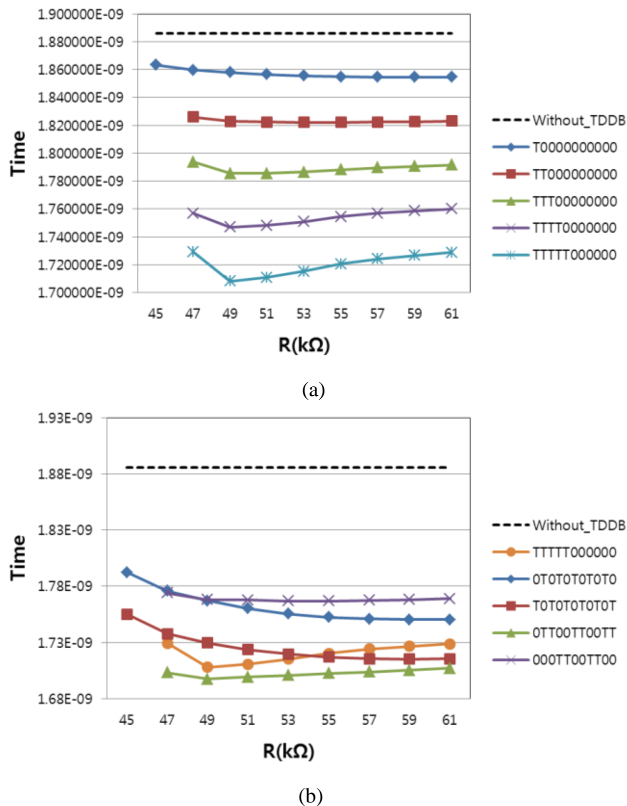

Figure 4. Delay vs. breakdown resistor considering the spatial correlation and the number of inverters with the breakdown resistor, where inverter-chain has 11 inverters; "T" means the inverter with the breakdown resistor; and "0" means the inverter without the breakdown resistor: (a) Delay vs. resistor size vs. the number of inverters with the resistor, (b) Delay vs. resistor size vs. the spatial correlation.

large, the propagation time from Input to Output will be decreased. This means that the propagation time in the case with SBD might be shorter than the time in the normal case without TDDB; and as a gate oxide goes to HDB, the propagation time might be longer than the time in the normal case without TDDB.

In addition, the total delay of the inverter chain depends on the spatial correlation between stressed devices and un-stressed devices; that is, the total delay can increase or decrease depending on the location of the stressed device and the number of stressed device [10]. Figure 4 shows the breakdown resistor impact on the Nand chain circuit delay. As expected, as the breakdown resistor size decreases, the delay of the Nand chain becomes decreased around up to 49KΩ, then the delay gets increased and finally the inverter chain goes to a functional failure. Figure 4 (a) shows the effect of the number of Nand gates with the breakdown resistor on the Nand chain delay: as the number of inverters with breakdown resistor increases, the total delay gets shorter. In Figure 4 (b), it is presented that the total delay of the inverter chain depends on the spatial correlation between stressed devices and un-stressed devices. Although the spatial correlation can change the total delay, the total number of the stressed inverter has a great influence on the total delay.

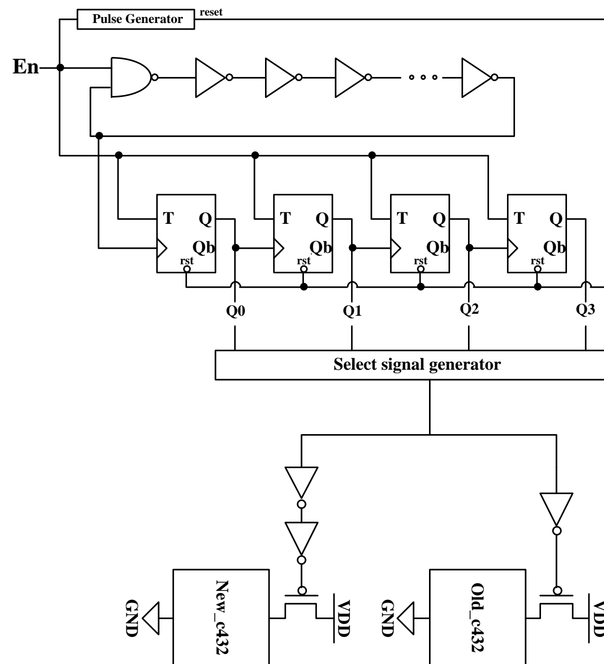

Figure 5. Proposed reliable C432 circuit

III. PROPOSED RELIABLE CMOS VLSI DESIGN

Figure 5 shows a reliable C432 ISCAS benchmark circuit using a redundancy system. The targeting circuit consists of old_c432 and new_c432. The proposed reliable circuit turns off the old_c432 circuit when the test device circuit suffers from HBD and leads to functional failure and turns on the new_c432 circuit with no TDDB according to the indication of the TDDB monitoring circuits consisting of a ring oscillator and T-flip flops. At the beginning point, both of the old and new circuits do not suffer from TDDB effects and "Select" signal output is logic '0'. At this time, the PMOS header of the old_c432 circuit is turned on, but the PMOS header of the old_c432 circuit is turned off. The primary inputs and outputs of the c432 circuits are connected to flip-flops that are controlled by the "Select" signal output.

In the proposed circuit, the ring oscillator operates as a replica circuit of a designed digital system. In the stress mode, the external signal En asserts logic '0' in order to turn off the ring oscillator. On the other hand, in the sensing mode, to monitor the TDDB effect on the ring oscillator, the external signal En asserts logic '1' in order to turn on the ring oscillator. During the sensing mode, when a soft breakdown presents at the ring oscillator, the gate to source/drain leakage current gets increased due to the decreasing of the breakdown resistor. Therefore, the counter consisted of four T flip-flops generates different numbers of pulses depending on the total delay of the ring oscillator with the breakdown resistor. When the ring oscillator goes to functional failure owing to the HBD effect, the ring oscillator finally stops oscillating. The proposed HBD sensor circuit detects this moment using an NOR gate. When the ring

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-213-4

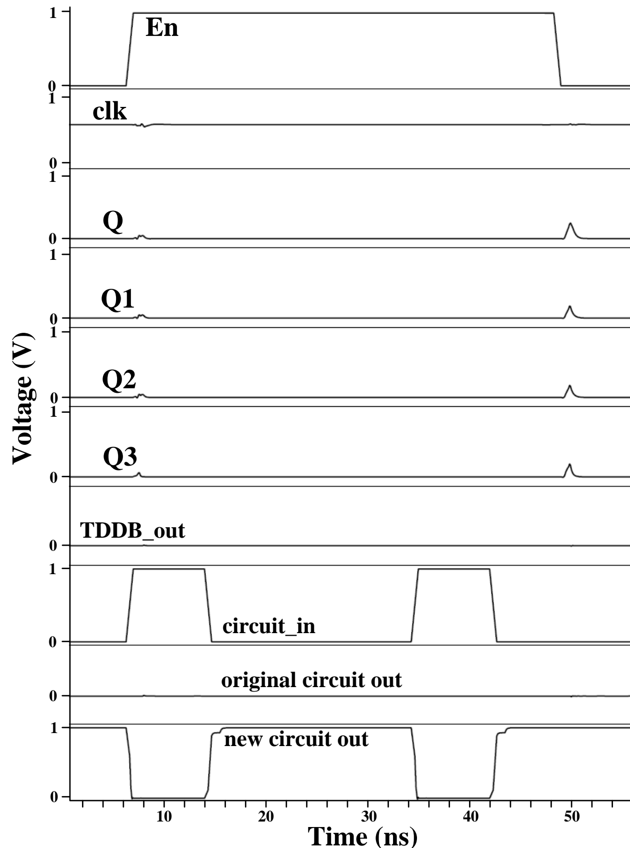

Figure 6. The transition waveform for the proposed monitoring circuits when HBD is generated.

oscillator has HBD resistors on inverters, the "Sensor_ Output" signal is changed to logic "0".

When "Select" signal output is logic '1', the primary inputs and outputs are connected to new_c432. At the initial time, the select signal output is logic '0', and the primary inputs and outputs are connected to old_c432. When HBD is generated in the old_c432 circuit, "Select" signal will be changed to logic '1' due to the output of the TDDB monitoring circuit. The PMOS header of the old_c432 circuit will be turned off, and the PMOS header of the new_c432 circuit will be turned on. Therefore, the impact of the HBD on digital circuit can be completely avoided using the proposed design methodology. Figure 6 shows transition waveforms, where HBD in the C432 circuit brings about functional failures. When the circuit leads to functional failures, the binary counter generates '0' outputs, which is different from the circuit with no TDDB or SBD. On the other hands, the monitoring circuit generates meaningful output only when the HBD is generated in the replica circuit (ring oscillator) after long time. When the replica circuit has the HBD effect, the outputs of monitoring circuit (Q0, Q1, Q2, and Q3 are all '0') are asserted to the TDDB signal generator circuit. The TDDB_out signal from the signal generator circuit is changed to logic '0', and then old_c432 circuit is blocked, but the new_c432 circuit will work instead of old_c432.

IV. CONCLUSION AND FUTURE WROK

In this paper, we proposed a reliable CMOS design methodology considering the gate oxide breakdown in the 45nm CMOS technology. The proposed design is based on the TDDB monitoring circuit, the simulation results shows that the impact of TDDB due to severe HBD on the digital circuit can be compensated by the proposed circuit design. Finally, we can extend the life time of the targeting circuit and prevent from performance degradation as well as functional failures due to TDDB. Future efforts aim at applying our proposed monitoring circuit to real systems. Moreover, a reliable system considering all the aging effects (NBTI, PBTI, HCI, and TDDB) will be designed and implemented in the nanoscale technology.

REFERENCES

[1] R. J. Baker, "CMOS-Circuit design, layout, and simulation," Wiley and IEEE Press, 2010.

[2] A. B. Kahng, "Design challenges at 65nm and beyond," Design, Automation and Test in Europe (DATE) Conference, pp. 1-2, Nice Acrolis, France, March 2007.

[3] J. W. McPherson, "Reliability challenges for 45nm and beyond," Proceeding of IEEE Design Automation Conference (DAC), pp. 176-181, San Francisco, USA, July 2006.

[4] X. Li, J. Qin, and J. B. Bernstein, "Compact modeling of MOSFET wearout mechanisms for circuit-reliability simulation," IEEE Transactions on Device and Materials Reliability, Vol. 8, Issue 1, pp. 98-121, March 2008.

[5] M. Choudhury, V. Chandra, K. Mohanram, and R. Aitken, "Analytical model for TDDB-based performance degradation in combinational logic", Design, Automation & Test in Europe (DATE) Conference, pp. 423–428, , Dresden, Germany, March 2010.

[6] H. Wang, M. Miranda, F. Cattoor, and D. Wim, "Impact of random soft oxide breakdown on SRAM energy/delay drift," IEEE Transactions on Device and Materials Reliability, Vol. 7, Issue 4, pp. 581-591, Dec. 2007.

[7] B. Kaczer, R. Degraeve, M. Rasras, K. Van de Mieroop, P. J. Roussel, and G. Groeseneken, "Impact of MOSFET gate oxide breakdown on digital circuit operation and reliability," IEEE Transactions on Electron Devices, Vol. 49, No. 3, pp. 500–506, 2002.

[8] R. Rodriguez, R. V. Joshi, J. H. Stathis, and C. T. Chuang, "Oxide breakdown model and its impact on SRAM cell functionality," International conference on Simulation of Semiconductor Processes and Devices (SISPAD), pp. 283–286, Cambridge, USA, September 2003.

[9] S. S. Lombardo, J.H. Stathis, B.P. Linder, K.L. Pey, F. Palumbo, and C.H. Tung, "Dielectric breakdown mechanisms in gate oxides," Journal of Applied Physics, 98(12), pp. 1-36, 2005.

[10] H. Luo, X. Chen, J. Velamala, Y. Wang, Y. Cao, V. Chandra, Y. Ma, and H. Yang, "Circuit-level delay modeling considering both TDDB and NBTI," IEEE International Symposium on Quality Electronic Design (ISQED), pp. 14-21, Santa Clara, USA, March 2011.