Impacts of Contact Resistance and NBTI/PBTI on SRAM with High-κ Metal-Gate Devices

Hao-I Yang, Ching-Te Chuang, and Wei Hwang

Department of Electronics Engineering and Institute of Electronics
National Chiao-Tung University, Hsinchu, Taiwan, R. O. C.

ABSTRACT

The contact resistance of CMOS device increases sharply with technology scaling, especially in SRAM cells with minimum size and/or sub-groundrule devices. Meanwhile, V_T drifts caused by Negative-Bias Temperature Instability (NBTI) and Positive-Bias Temperature Instability (PBTI) degrade stability, margin, and performance of nanoscale SRAM with high-κ metal-gate devices over the lifetime of usage. In this work, we comprehensively analyze the impacts of contact resistance and the combined effects with NBTI and PBTI on SRAM cell stability, margin, and performance. The effect of contact resistance on power-gated SARM is also investigated.

INTRODUCTION

With technology scaling, the device contact and series resistance of the channel/source/drain increase sharply, thus severely degrading the transistor performance [1, 2]. At the same time, Negative-Bias Temperature Instability (NBTI) causes V_T drift of scaled PMOS, while Positive Bias Temperature Instability (PBTI) causes V_T drift of high-κ metal-gate NMOS over the lifetime of usage. Previous works have shown that NBTI/PBTI degrades SRAM Read Static Noise Margin (RSNM) but improves Write Margin (WM) [3]. As such, it is crucial to understand the effects of contact and series resistances on the SRAM cell, and the combined impacts with NBTI/PBTI on SRAM stability, margin, and performance. In this paper, we present a comprehensive analysis based on BSIM 32 nm high-κ metal-gate predictive model [4]. The effects of contact and series resistances on power-gated SRAM with footer and header are also analyzed.

SERIES RESISTANCE AND NBTI/PBTI

As shown in Fig. 1(a), the Source/Drain (S/D) series resistance can be divided into overlap resistance (R_{OV}), extension resistance (R_{EXT}), deep resistance (R_{DP}), and silicon-contact diffusion resistance (R_C), where all resistance are in units of Ω/□. Conventionally, R_{OV}, R_{EXT}, and R_{DP} are included in the device model, but R_C isn't. We model the R_C of a transistor as shown in Fig. 1(b). With technology scaling, the sum of R_{OV}, R_{EXT}, and R_{DP} decreases, but R_C increases. The formula for silicon-contact diffusion resistance is given by:

$$R_C = \sqrt{R_S \rho_C}\cosh\left(\frac{L_{silicide}}{l_t}\right) \quad (1)$$

where R_S is the sheet resistance per square of the underlying heavily doped silicon layer, in unit of Ω, ρ_C is the specific contact resistivity between the metal and the diffusion layer in unit of Ω/□, and l_t is the transfer length, which is defined as $l_t = \rho_C / R_S$ [2]. When $L_{silicide}$ is larger than l_t, the contact resistance is slightly dependent on the contact region; however, when $L_{silicide}$ is smaller than l_t, the contact resistance increases sharply if $L_{silicide}$ is further scaling down. According to [2], the contact resistance would be larger than the sum of R_{OV}, R_{EXT}, and R_{DP}, and increases sharply beyond 45nm technology node. As diffusion contact resistance dominates the short channel resistance, we focuses on its impacts on SRAM array in the following analysis.

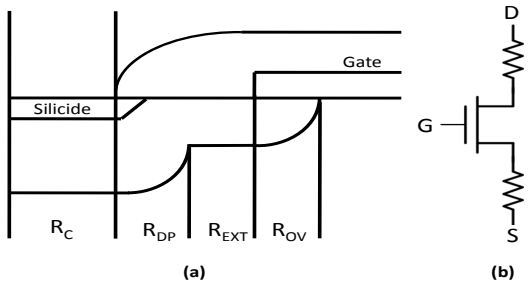

Fig. 1. (a) Series resistance components of S/D, and (b) Schematic of NMOS with S/D diffusion contact resistances.

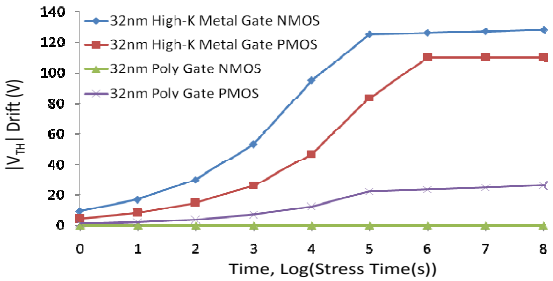

Fig. 2. V_{TH} drifts of high-κ metal-gate devices induced by NBTI and PBTI.

NBTI causes V_{TH} of PMOS to increase with time when V_{GS} of PMOS is negative as shown in Fig. 2. After removing the negative V_{GS}, V_{TH} drift of PMOS decreases (partially recovered). The recovery mechanism makes the life-time of PMOS longer than the prediction based on DC stress. Recently, with the introduction of high-k metal-gate technology, the Positive Bias Temperature Instability (PBTI) has emerged to be a major reliability concern for NFETs as well due to V_T instability caused by charge trapping at the interface. The V_{TH} drifts can be described by AC Reaction Diffusion model when the stress signal of PMOS and NMOS changes with time [3].

This work is based on BSIM 32-nm high-κ metal-gate predictive model. The contact resistance data are based on the ITRS Road-map [5] and published data [2]. Moreover, V_{TH} drift of PMOS and NMOS are calculated based on AC Reaction Diffusion model and calibrated against published data [6]. In following sections, we analyze Read and Write operation of SRAM cell with contact resistance and NBTI/PBTI induced V_T drift. We also investigate the impacts of contact resistance on power-gated SRAM with footer and header.

978-0-7695-3797-9/09 $26.00 © 2009 IEEE
DOI 10.1109/MTDT.2009.25

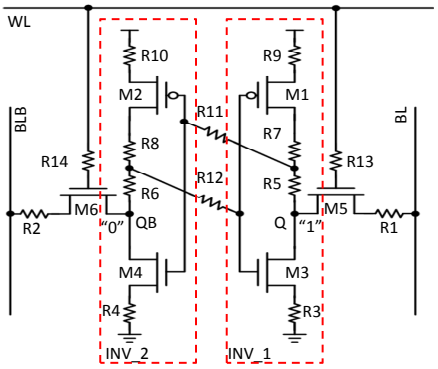

Fig. 3. SRAM cell structure with contact resistances.

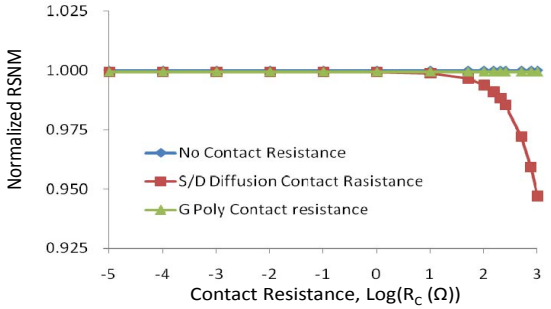

Fig. 4. Normalized RSNM vs. contact resistance, RSNM is ormalized with respect to the case with no contact resistance.

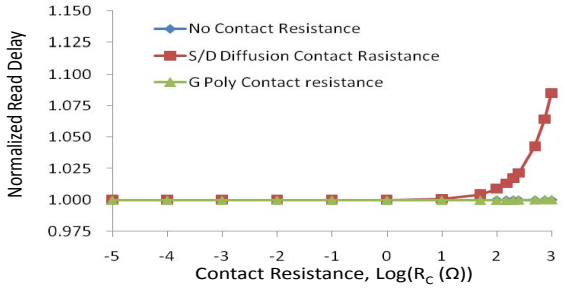

Fig. 5. Normalized Read delay vs. contact resistance. Read delay is normalized with respect to the case with no contact resistance.

READ OPERATION

Read Static Noise Margin (RSNM) can be defined as the minimum trip voltage of the SRAM inverter pairs minus the maximum Read disturb during Read cycles. Fig. 3 shows a SRAM cell with contact resistances. We also assume the cell stores "logic 1" (Q = 1). When the diffusion contact resistances, R1, R5, R7, and R9 increase, the trip voltage of INV_1 decreases and RSNM degrades. If the diffusion contact resistance, R3, increases, the trip voltage of INV1 increases and RSNM improves. On the other hand, M4 and M6 form a voltage divider and induce Read disturb during Read cycle. The Read disturb voltage increases with larger R4 but decreases with larger R2. All gate-poly contact resistances affect neither the trip voltage because they are in series with the "infinite" gate resistance, nor the RSNM as they are not on the Read current paths.

RSNM decreases with increasing contact resistance as shown in Fig. 4. When the contact resistance approaches 1kΩ, RSNM would

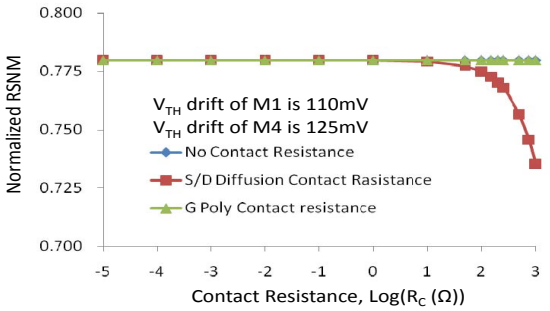

Fig. 6. Normalized RSNM under NBTI and PBTI vs. contact resi-stance. RSNM is normalized with respect to the case with no contact resistance and no NBTI/PBTI stress.

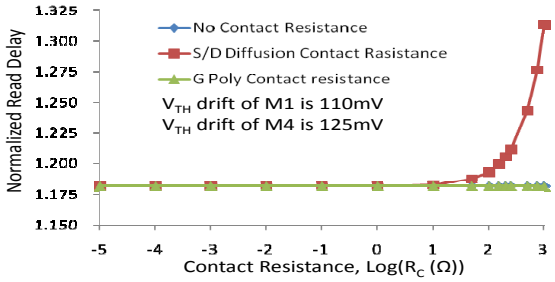

Fig. 7. Normalized Read delay under NBTI and PBTI vs. contact resistance. Read delay is normalized with respect to the case with no contact resistance and no NBTI/PBTI stress.

degrade 6%. The reason is that R5, R7, and R9 form a series resistance chain, causing the trip voltage to decrease. Although R3 increases the trip voltage, its effect is smaller than the R5/R7/R9 resistance chain. Notice that R2 compensates the Read disturb increase caused by R4, thus the Read disturb voltage remains almost unchanged. Fig. 4 also shows the RSNM is not impacted by the increase in gate-poly contact resistance as discussed in previous section.

Notice that when the diffusion contact resistance increases, Read delay becomes longer as shown in Fig. 5. This is due to increased R2 and R4 on the Read current (bit-line discharge) path. As a result, the discharge time of BLB increases with increasing diffusion contact resistance. Moreover, Read delay is insensitive to the gate-poly contact resistance because there are not on the Read current path of a SRAM cell.

When NBTI and PBTI are considered, in the worst case, V_{TH} of M1 and M4 increase while V_{TH} of M2 and M3 remain unchanged. Because access transistors, M5 and M6, are stressed only during WL turning on period, the V_{TH} drifts of access transistors are negligible. V_{TH} drift of M1 lowers the trip point of INV_1, and V_{TH} drift of M4 causes increase of the Read disturb voltage, resulting in RSNM degradation with usage time. By using AC Reaction-Diffusion model, V_{TH} drifts induced by NBTI and PBTI for 10^6s of M1 and M2 are calculated to be 110mV and 125mV respectively in the worst case, leading to RSNM degradation of about 22% without considering contact resistance effect as shown in Fig. 6. Fig. 6 also shows that RSNM degradation becomes more serious when the cell is impacted by both NBTI/PBTI and the diffusion contact resistance. Furthermore, the Read delay increases when the cell is impacted by NBTI and PBTI according to Fig. 7. The reason is that M4 is on the BLB discharging path, leading to longer Read delay with larger V_{TH} of M4. Fig. 7 also shows that NBTI/PBTI and the diffusion contact resistance degrade SRAM Read performance cumulatively.

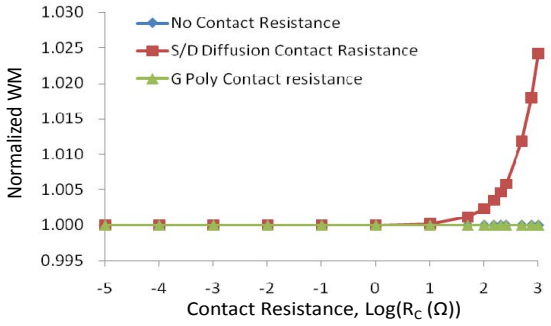

Fig. 8. Normalized WM vs. contact resistance. WM is normalized with respect to the case with no contact resistance.

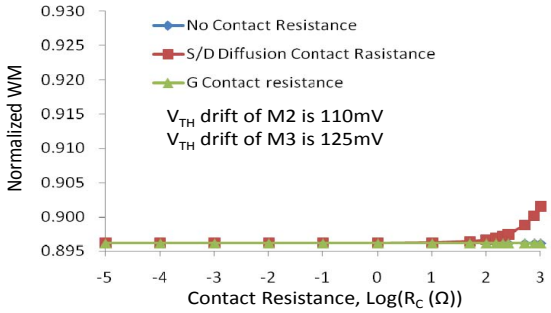

Fig. 9. Normalized WM under NBTI and PBTI vs. contact resistance. WM is normalized with respect to the case with no contact resistance and no NBTI/PBTI stress.

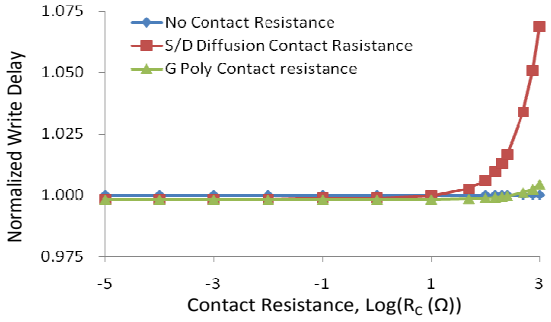

Fig. 10. Normalized Write delay vs. contact resistance. Write delay is normalized with respect to the case with no contact resistance.

WRITE OPERATION

Write Margin (WM) can be defined as the BL voltage level below which the cell will flip during Write cycles. Referring to Fig. 3, larger R5, R7, and R9 reduce the holding strength of PMOS M1, thus facilitating pull-down of the storage node Q. Larger R1 impedes BL to discharge Q through M5. Larger R6, R8, and R10 impede M2 to charge up node QB, and larger R2 also prevents BLB to charge up QB through M6. Thus, larger R5, R7, and R9 improve WM, while larger R1, R2, R6, R8, and R10 degrade WM. Nevertheless, charging up QB is the second order effect during Write, and WM is mainly impacted by R1, R5, R7, and R9. As shown in Fig. 8, WM is improved by larger diffusion contact resistance, but is relatively insensitive to the gate-poly contact resistance because gate-poly

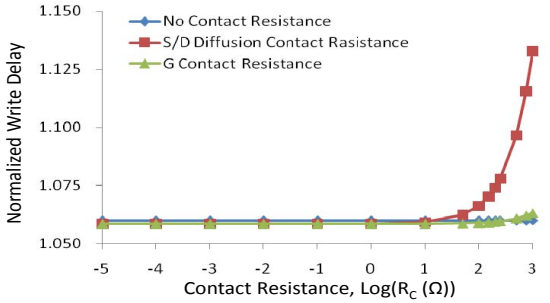

Fig. 11. Normalized Write delay under NBTI and PBTI vs. contact resistance, Write delay is normalized with respect to the case with no contact resistance and no NBTI/PBTI stress.

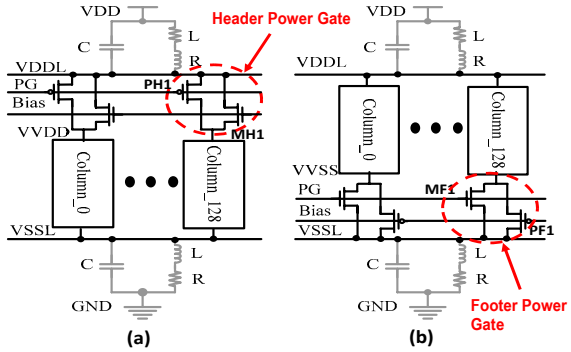

Fig. 12. Power-gating structure with (a) header, and (b) footer.

contacts are not on the access paths of Q and QB. When the diffusion contact resistance approaches 1kΩ, WM improves by about 2.5%.

When the cell is also impacted by NBTI and PBTI, in the worst case, V_{TH} of M2 and M3 increase, while V_{TH} of M1 and M4 remain unchanged. The V_{TH} drifts of M5 and M6 are negligible. Weak M2 slows down the charging of QB, and weak M3 slightly impedes the discharging of Q. Consequently, WM of SRAM cell under NBTI and PBTI degrades in the worst case. Fig. 9 shows the relation between WM and contact resistance when V_{TH} of M2 and M3 are 110mV and 125mV, respectively. As can be seen, WM degrades about 10% due to NBTI and PBTI. In contrast with RSNM, larger diffusion contact resistance improves WM slightly (about 0.5%), as the current charging QB is limited by M2 under NBTI effect.

Write delay is defined as the latency between the time WL rises to half V_{DD} and the time Q and QB cross each other. Write delay normally tracks WM, and better (higher) WM would improve Write delay in general. However, Write delay is also affected by the RC time constant, and larger diffusion contact resistances lead to longer Write delay as shown in Fig. 10. Additionally, Fig. 11 shows the relation between Write delay and the contact resistance when the cell is under NBTI and PBTI stress. The Write delay can be seen to degrade about 6% with NBTI and PBTI. The Write delay also increases sharply when the diffusion contact resistance is larger than 100Ω.

SRAM POWER-GATING STRUCTURE

Most of state-of-the-art low power SRAM designs use the power gating technique to suppress the leakage power during Standby/Sleep mode. The power gating structures can be divided to Header and Footer as shown in Fig. 12. Fig. 12(a) is a column-based header-gated SRAM, where PH1 is the power switches for leakage reduction

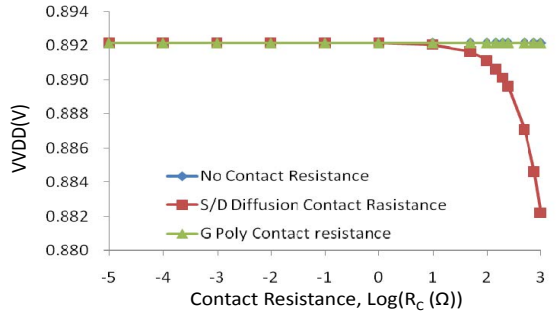

Fig. 13., Active mode VVDD of a header power-gating structure.

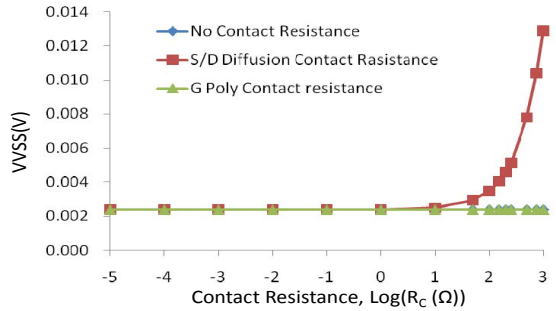

Fig. 14. Active mode VVSS of a footer power-gating structure.

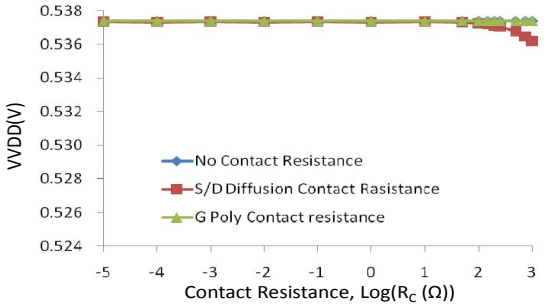

Fig.15. Standby mode VVDD of header power-gating structure.

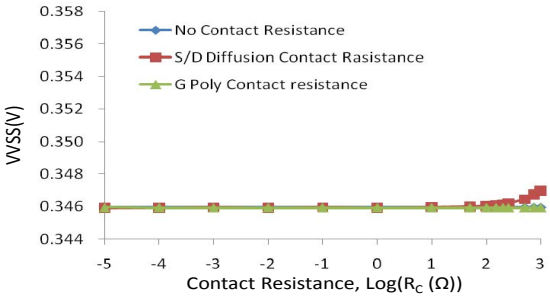

Fig. 16. Standby mode VVSS of footer power-gating structure.

in Standby mode and MH1 is the clamping device for data retention. Fig. 12(b) is a column-based footer-gated structure, where PF1 is the power switches and MF1 is the clamping device. In the power-gating structure, the power switches should fully turn on to provide sufficient currents and voltage to maintain performance and margin of SRAM cells in Active mode, while the clamping devices provide proper voltage levels to maintain SRAM stability for data retention in Standby mode.

When diffusion contact resistances increase, VVDD of a header-gated structure decreases, and VVSS of a footer-gated structure increases as shown in Fig. 13 and Fig. 14, respectively. Consequently, the voltage across the SRAM array reduces, and RSNM degrades while WM improves. On the other hand, larger diffusion contact resistances reduce the leakage during Standby mode. It also reduces Standby VVDD of Header-gated structure, and increases Standby VVSS of Footer-gated structure shown in Fig. 15 and Fig. 16, respectively. However, the changes in virtual supply/GND voltage during Standby mode are smaller than those during Active mode. Since the current flowing through the SRAM during Standby is significantly smaller than that during Active mode.

When the power switch turns on during wake-up transition, large wake-up current flows through the package parasitic capacitors, inductors, and resistance, resulting in VVDD bounce in header-gated structure or VVSS bounce in footer-gated structure. As the diffusion contact resistance increases, the wake-up current reduces and virtual supply/GND bounce is mitigated. However, due to reduced wake-up current, the wake-up time becomes longer.

CONCLUSIONS

In this work, we presented a comprehensive analysis on the impact of contact resistances and NBTI/PBTI on SRAM stability, margin, and performance based on BSIM 32nm high-κ metal-gate predictive model. We show that when the diffusion contact resistance is larger than 100Ω, the impacts on RSNM, WM, and Read/Write delay become significant and can't be neglected anymore. However, the gate-poly contact resistances are not on the SRAM access path, and their impacts are negligible. Diffusion contact resistances and NBTI/PBTI cumulatively degrade RSNM, Read delay, and Write delay. On the other hand, WM degradation induced by NBTI/PBTI is offset by larger diffusion contact resistances. Finally, in power-gated SRAM, larger diffusion contact resistances reduce the voltage across SRAM array in Active mode, leading to RSNM degradation. Virtual supply/GND bounce during wake-up transition is reduced with increasing contact diffusion resistances, but wake-up time becomes longer.

ACKNOWLEDGEMENT

The authors would like to thank National Science Council, R.O.C. for their support.

REFERENCES

[1] F. Babarada, et al., "MOSFET Distorsion Analysis Including Series Resistance Modelling Aspects," *Semiconductor Conference*, pp. 307-310, Oct. 2004.
[2] Lan Wei, et al., "Selective Device Structure Scaling and Parasitic Engineering: A Way to Extend the Technology Roadmap," *IEEE TED,* vol. 56, No. 2, pp. 312-330, Feb. 2009.
[3] K. Kang, et al., "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," *IEEE TCAD*, vol. 26, No. 10, pp. 1770-1781, Oct. 2007.
[4] http://www.eas.asu.edu/~ptm/.
[5] http://www.itrs.net/.
[6] S. Zafar et al., "A Comparative Study of NBTI and PBTI (Charge Trapping) in SiO2/HfO2 Stacks with FUSI, TiN, Re Gates," *IEEE Symp. on VLSI Tech. Digest of Tech. Papers*, pp.23-25, 2006.