A MECHANISM STUDY OF HIGH-K DIELECTRIC QUALITY AND METAL GATE AL DIFFUSION AFFECTING PPU TRANSISTOR THRESHOLD VOLTAGE

Weiwei Ma[1], Ran Huang[1], Yamin Cao[1], Wei Zhou[1]*

[1]Shanghai Huali Integrated Circuit Corporation, Shanghai, China
Corresponding Author's Email: maweiwei@hlmc.cn

ABSTRACT

High-k /metal replace SiO_2/polysilicon as gate stack enables transistor size continuously scaling down. In this paper, the Vt (threshold voltage) instability mechanism of 28 nm PPU (p-type pull up) transistors in HKMG SRAMs (static random access memory) is investigated. A defect-assisted Al diffusion and dipole formation model is proposed to explain this phenomenon. Vt up-shift level of PPU transistor is dominated by the formation amount of dipoles, which are relied on both Al diffusion amount and defect density of high-k layer. Furthermore, reducing PNA (post nitridation anneal) temperature is demonstrated to be effective to suppress the Vt up-shift of PPU transistors. And this Vt adjustment is much more convenient than altering the metal gate stack to block Al diffusion.

INTRODUCTION

With the continuously scaling down of transistors, high-k dielectrics and metal gates were introduced as new gate stacks for solving intolerable tunnel leakage problems [1-4]. Two different metal gate integration approaches, namely gate first and gate last, were implemented in high-volume production. The later one was also known as replacement gate approach, which had become mainstream solution in industry.

In a typical planar gate last approach, gate electrode including work function metals were deposited after dummy silicon poly was removed. To ensure a low threshold voltage the work function difference between gate electrode and channel material should be small [1], which required different work function metals for NMOS and PMOS. Al with a work function of 4.1 eV was a suitable work function metal for NMOS. And TiN with a work function around 5 eV was suitable for PMOS.

One typical gate stack of NMOS transistors was IL (interfacial oxide)-high-k dielectric-TiN-TaN-TiAl-Al electrode. And for PMOS, the gate stack was IL-high-k dielectric-TiN-TaN-TiN-TiAl-Al electrode, which had one more layer of P-type work function metal (TiN) under TiAl layer when compared with NMOS. Ideally, these two complex gate stacks would deliver appropriate effective work functions for NMOS and PMOS respectively. However, when two adjacent NMOS and PMOS transistors shared one gate, especially in SRAM, these two adjacent different gate electrode stacks would affect each other through their boundary area, as shown in figure 1. For one situation, the threshold voltages (absolute value) of PMOS transistors would increase due to undesirable Al diffusion from NMOS side, which decreased the effective work function of the whole PMOS gate stack. What's worse, such Vt shift could be different among different transistors throughout the whole wafer, which resulted in serious transistor mismatch issue. Therefore, understanding the mechanism of how this boundary (N-type and P-type metal gate boundary) affecting transistor performance and how to minimize this undesirable phenomenon were important.

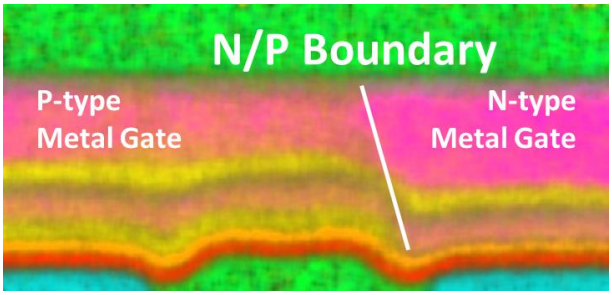

Figure 1. Boundary of N-type and P-type Metal Gates

In this study, two kinds of boundaries were designed, one located exactly in the middle of two NMOS and PMOS transistors, and the other one located on AA (active area) of NMOS transistor far away from PMOS. A third kind of PMOS transistor without boundary was also designed as a contrast.

The threshold voltages of PMOS transistors in these three designs were tested and compared. It was found that threshold voltages of PMOS transistors with or without boundaries reacting to PNA temperature differently. A defect-assisted Al diffusion and dipole formation model was proposed to explain these results. Reducing PNA temperature was demonstrated to be effective to suppress the Vt up-shift of PPU transistors in SRAM and achieve better transistor match.

EXPERIMENT

The layouts of three designed transistors with or without boundary were illustrated in figure 2. All wafers analyzed in this study were manufactured based on HLIC 28 nm HKMG technology.

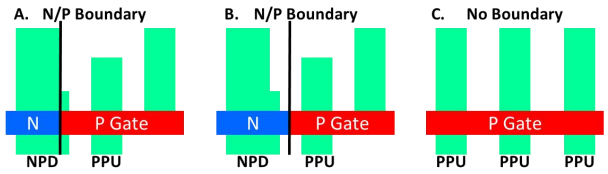

Figure 2. Boundary Designs: A. Boundary on NPD, B. Boundary in The Middle of NPD and PPU, C. Pure P-type Metal Gate without Boundary

Two different PNA temperatures were applied to all these three different designs. Process temperature of PNA1 was much lower than PNA2. The threshold voltage differences of

PMOS were tested and analyzed. IL and high-k film quality were also characterized by XPS (X-ray photoelectron spectroscopy).

RESULTS AND DISCUSSION

Threshold Voltage Reacting to PNA Temperature

It was found that threshold voltages of PMOS transistors with or without boundaries reacting to PNA temperature differently. And these results were shown in figure 3. Firstly, the threshold voltages of PMOS with boundary reduced when process by a lower PNA temperature. Secondly, the threshold voltage of PMOS without boundary remained the same. Thirdly, the farther the PMOS was away from the boundary, the less the threshold voltage shift was. The threshold voltage reduction ratios were 3% and 6% for boundary on NMOS AA and boundary at center respectively.

These results suggested that the Vt shift phenomenon was related to metal gate boundary. However, the reason why threshold voltage had a dependence on PNA temperature was not clear. Therefore, quality characterization of IL and High-k dielectric was carried out.

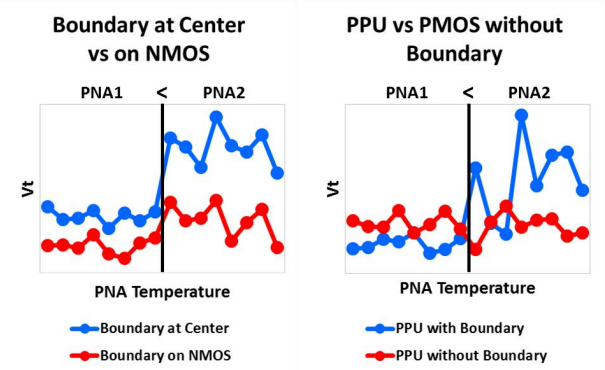

Figure 3. Threshold Voltage Results of PMOS with or without Boundary Processed by Different PNA Temperature

IL and HK Quality Characterization

The thickness results of IL and high-k dielectric were tested by XPS, and shown in figure 4. It was interesting that IL thickness increased along with the increase of PNA temperature and high-k dielectric thickness was just the opposite.

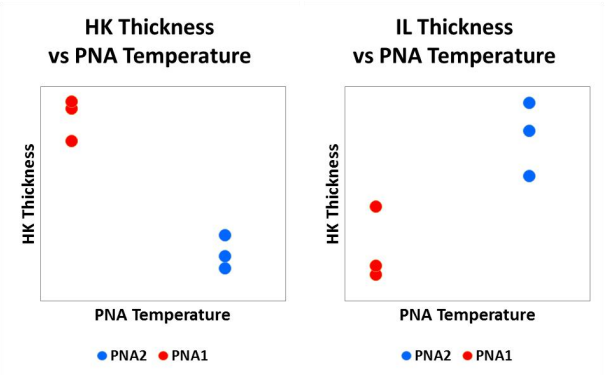

Figure 4. Thickness results of IL and HK Processed by Different PNA Temperature

It suggested that during high temperature (around 900°C) PNA process IL took oxygen atoms from high-k layer, which resulted in the increase of IL thickness. This final resulted in the increase of EOT (equivalent oxide thickness), which was detrimental to transistor performance. Furthermore, no Oxygen was used in PNA process, so this IL re-growth would increase defects in high-k layer, especially Oxygen vacancies. Figure 5 illustrated the phenomenon of high-k layer loss Oxygen when IL re-growth. PNA2 had a higher PNA temperature than PNA1 and resulted in thicker IL and more Oxygen vacancies.

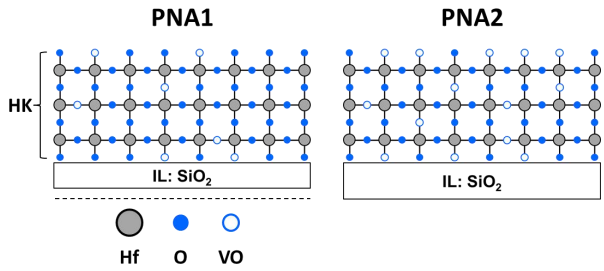

Figure 5. PNA Temperature Affects IL Thickness and High-k Oxygen Vacancies

Al Diffusion at Boundary

Prior to this study, we had studied the diffusion of Al at metal gate boundary area. The EDS (energy disperse spectroscopy) image of metal gate boundary was shown in figure 6. One less work function metal layer resulted in higher Al electrode portion in NMOS metal gate. It was clear that the step coverage of barrier layer under Al gate at the boundary area was not so good. Thus, Al diffusion would occur from NMOS Al gate to PMOS Al gate, which had been detected by EDS. These extra Al would not only alter the EWF (effective work function) of PMOS directly, but also diffuse further into high-k layer and then effect the threshold voltage of the adjacent PMOS.

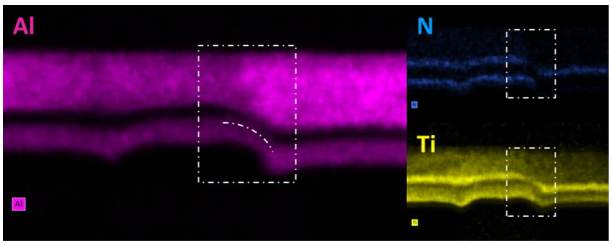

Figure 6. Al Diffusion at Metal Gate Boundary

Proposed Model

Dipoles contained Al accumulated at the interface layer between the gate and high-k layer would lower the EWF of PMOS metal gate stack and then the threshold voltage would increase as a result [5]. Figure 7 presented the schematic diagram illustrating how the interface dipoles lowered the EWF.

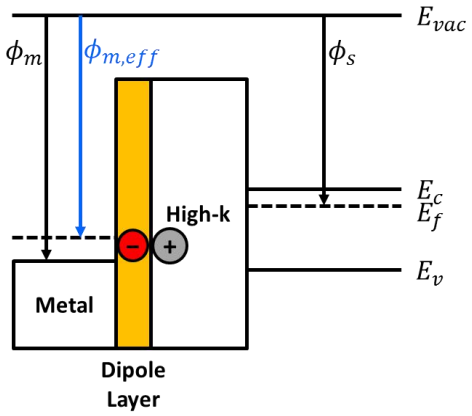

Figure 7. Schematic Diagram: Dipole Lower EWF of PMOS Metal Gate

Therefore, we supposed that PNA temperature should have affected the amount of dipole formation. Under lower PNA temperature fewer dipoles formed. And that was suitable for PMOS with boundary.

However, the threshold voltage of PMOS without boundary had no reaction to PNA temperature. This could be explained by two factors, Al and Oxygen vacancy in high-k layer, that affecting diploe formation. Because PMOS without boundary had much fewer Al diffused into interface layer, while Oxygen vacancies in high-k layer were abundant. So the amount of formed dipoles was determined by the amount of diffused Al. When PMOS with boundary, Al diffused into interface layer became abundant and they could not form dipoles fully. And when PNA temperature decreased, the high-k quality increased with fewer Oxygen vacancies, reducing the formation of dipoles. Furthermore, Al diffusion from NMOS side to PMOS side through high-k layer would be more difficult when the quality of high-k layer increased. Through the above analysis, we proposed a defect-assisted Al diffusion and dipole formation model to explain the benefit of reducing PNA temperature on reducing the uncontrollable influence of boundary on PMOS threshold voltage. Figure 8 illustrated the model proposed. PNA1 had a lower PNA temperature than PNA2. Fewer dipoles formed due to fewer Oxygen vacancies available in the PMOS with metal gate boundary.

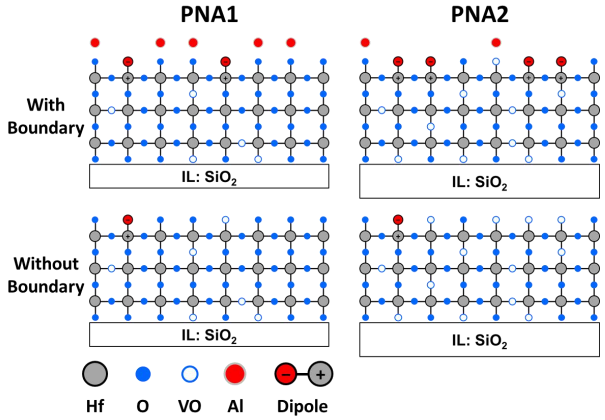

Figure 8. Defect-assisted Al Diffusion and Al-Hf Dipole Formation Model

CONCLUSION

The mechanism of the Vt instability of 28 nm PPU transistors in HKMG SRAMs was studied. A defect-assisted Al diffusion and dipole formation model was proposed to explain this phenomenon. Vt up-shift level of PPU transistor was determined by the formation amount of dipoles, which were relied on both Al diffusion amount and defect density of high-k layer. Furthermore, reducing PNA process temperature was demonstrated to be effective to reduce the Vt up-shift of PPU transistors. The Vt reducing ratio was 3%-6%.

ACKNOWLEDGEMENTS

I would like to appreciate senior engineer Ran Huang, Yamin Cao, Wei Zhou for their support and advice on this experiment.

REFERENCES

[1] "Work Function Setting in High-k Metal Gate Devices," Complementary Metal Oxide Semiconductor Chaper 3, 2018.

[2] "Challenges for The Integration of Metal Gate Electrodes," International Electron Devices Meeting, 2004.

[3] "Characteristics and Mechanism of Tunable Work Function Gate Electrodes Using a Bilayer Metal Structure on SiO_2 and HfO_2," Electron Device Letters, 2005.

[4] "Dipole-induced Modulation of Effective Work Function of Metal Gate in Junctionless FETs," AIP Advances, 2020.

[5] "Interface Dipole Engineering in Metal Gate/High-k Stacks," Chin Sci Bull, 2012.