Physical Mechanism of HfO_2-based Bipolar Resistive Random Access Memory

Huan-Lin Chang[1], Hsuan-Chih Li[1], C. W. Liu[1], F. Chen[2], and M.-J. Tsai[2]
[1]Department of Electrical Engineering and Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.
[2]Electronics and Optoelectronics Research Laboratory, Industrial Technology Research Institute, Hsinchu, Taiwan, R.O.C.

I. Introduction

The (anode) TiN/Ti/HfO_2/TiN (cathode) resistive random access memory (RRAM) (Fig. 1) has shown yield ~ 100% [1]. Its simple metal-insulator-metal (MIM) structure exhibits great potential for an embedded BEOL memory compatible with the high-k/metal gate CMOS process. There have been many theories of RRAM physical mechanism in the literature. This paper focuses on HfO_2-based RRAM and describes a complete physical mechanism from forming, SET/RESET, current conduction, to explanations of various observed phenomena including multilevel, cell size scaling, resistance fluctuation, soft error, and non-abrupt RESRT process. Finally, suggestions for device optimization are given based on the physical model.

II. Physical Mechanism

The resistance switching between high resistance state (HRS) and low resistance state (LRS) of HfO_2-based RRAM is based on the formation and rupture of Hf filaments via an electrochemical redox process in a region of a few nm near the anode.

Oxygen Vacancy

Oxygen vacancies play a crucial role in the resistance switching and the undesired resistance fluctuation of RRAM. There are two types of oxygen vacancies in HfO_2: (1) HfO^{2+} (V_O^{2+}) if one oxygen atom is missing, and (2) Hf^{4+} if two oxygen atoms are out of place (Fig. 2). The positively charged oxygen vacancies are mobile under high field and serve like donor dopants to make HfO_2-based RRAM n-type semiconductor.

Forming

The forming process is to produce sufficient amount of oxygen vacancies in HfO_2 to initiate the resistance switching. The Ti layer serves as a reservoir of O (~ 10 % solid solubility). Under high positive voltage, the O^{2-} ions move into the Ti layer by diffusion and/or drift under high field, and then oxidize at anode while the Hf^{4+} ions (a type of oxygen vacancies) reduce at cathode. One or many Hf filaments grow from cathode to anode, making RRAM exhibit LRS after forming (Fig. 3).

RESET/SET

Under negative voltage, reverse redox process happens near the anode. The O in Ti layer reduces, releasing O^{2-} ions back into HfO_2 while Hf filaments oxidize to have Hf^{4+} ions. The recombination of Hf^{4+} and O^{2-} causes the filaments to rupture partially (a few nm) near anode, making RRAM exhibit HRS (Fig. 4) after RESET. Once positive voltage (< forming voltage) is applied, the filaments form again in the rupture region, switching RRAM back to LRS after SET. Because the rupture and reformation of the Hf filaments occur in a few nm near anode, this explains the reported high-speed (< 5 ns) switching of RRAM [2].

LRS/HRS Conduction

The conduction of LRS might be quasi-ballistic because the filaments cannot sustain such high current density as predicted by Joule heating. The thickness of RRAM (10 nm) is just one to two times of the mean free path of electrons in Hf, so there are just a few scattering events during carrier transport, leaving temperature mostly absorbed by the metal electrodes (Fig. 5). The conduction of HRS is by direct tunneling near anode, which is verified by the temperature insensitivity and exponential dependence on the applied voltage (Fig. 6).

Multilevel

The LRS versus current compliance (CC) (Fig. 7) can be understood using Faraday's electrolysis law, which states the mass altered at an electrode during electrolysis is proportional to charge transferred at the electrode. Higher CC means more charge transferred at a given time period and more Hf atoms are produced. More Hf filaments and/or larger filament cross section cause lower LRS. On the other hand, multilevel HRS is related to the tunnel gap, modulated by the magnitude of negative RESET voltage. Larger RESET voltage forces more O^{2-} ions to recombine with Hf^{4+}, causing larger tunnel gap and higher HRS (Fig. 8).

LRS/HRS Scaling

The LRS (same SET voltage) is independent on the cell size because the filament cross section is much smaller than the device area. On the other hand, the HRS (same RESET voltage) increases as RRAM is scaled (Fig. 9). Though same amount of O^{2-} ions are driven from the Ti layer per unit area, smaller device receives more O^{2-} ions from peripheral SiO_2 used for RRAM isolation because Hf shows stronger attraction to O than Si. More O^{2-} ions mean longer tunnel gap and higher HRS (Fig. 10)

Cycling Fluctuation

One reliability issue is that resistance fluctuation during cycling tests is proportional to the resistance value (Fig. 11). Since the total amount of oxygen vacancies is determined after forming, the more Hf^{4+} ions turn into filament, the fewer HfO^{2+} (V_O^{2+}) remain as Frenkel-Poole (F-P) defects. The resistance fluctuation is due to randomly distributed current contributed from F-P defects at each cycle. Because lower LRS (higher CC) means more Hf atoms and less V_O^{2+}, the cycling fluctuation is less severe (Fig. 12). The impact of F-P defects is insignificant during HRS cycling.

HRS Retention

Another reliability issue is the relatively poor HRS retention (Fig. 13). The observed soft error can be understood as the read disturb. Since read is carried out by applying a small positive voltage across RRAM, this is equivalent to an accumulated SET process by gradually extend the broken filament via aforementioned redox process. After RRAM has been read sufficient times, HRS suddenly drops to LRS because the filament has reached to the anode.

Non-abrupt RESET

The asymmetrical *I-V* curve shows an abrupt SET process and a gradual RESET process. Two mechanisms are competing with each other in the RESET process. As RESET voltage turns more negative, tunnel gap increases, but more V_O^{2+} will accumulate at anode and effectively shorten the gap (Fig. 14).

III. Conclusion

Based on the physical model, device optimization can be performed by (1) designing the thickness of Ti layer for control of oxygen vacancies, and (2) using Si_3N_4 instead of SiO_2 for isolation to keep HRS scaling trend flat and avoid undesirable large resistance fluctuation.

References

[1] H.-Y. Lee et al., *IEDM Tech. Dig.*, pp. 297-300, 2008.
[2] S.-S. Sheu et al., *VLSI Symp. Tech. Dig.*, pp. 82-83, 2009.

978-1-4244-8492-8/11/$26.00 ©2011 IEEE

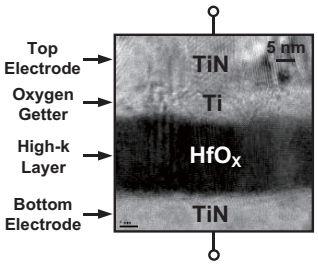


Fig. 1. TEM image of TiN/Ti/HfO_X/TiN RRAM.

$$HfO_2 \rightleftharpoons HfO^{2+}(V_O^{2+}) + O^{2-}$$

$$HfO^{2+}(V_O^{2+}) \rightleftharpoons Hf^{4+} + O^{2-}$$

Fig. 2. Two types of oxygen vacancies in HfO_2.

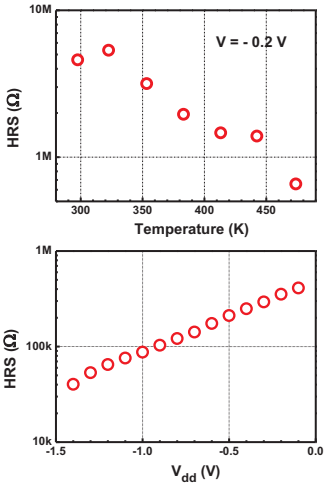


Fig. 6. HRS temperature and voltage dependence.

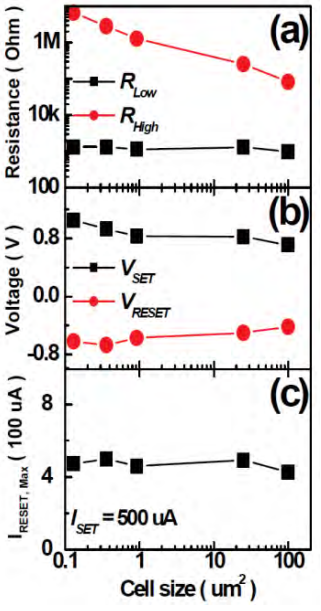


Fig. 9. HRS/LRS scaling effect.

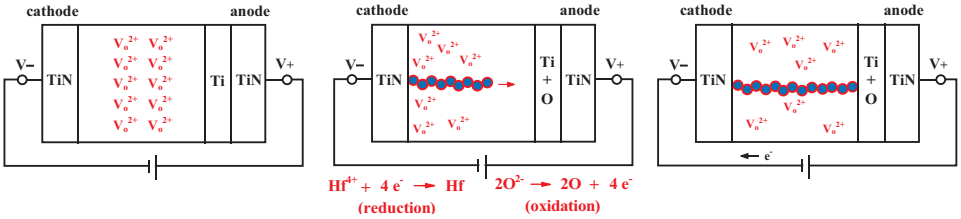


Fig. 3. Forming (switching to LRS) and the electrochemical redox process.

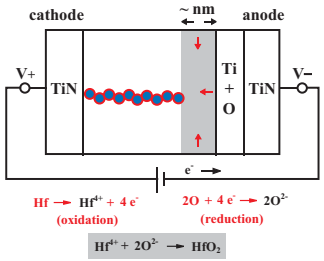


Fig. 4. RESET process (switching to HRS).

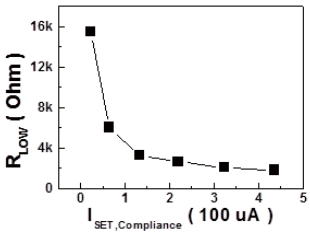


Fig. 7. Multilevel LRS versus CC.

Fig. 8. Multilevel HRS versus RESET voltage.

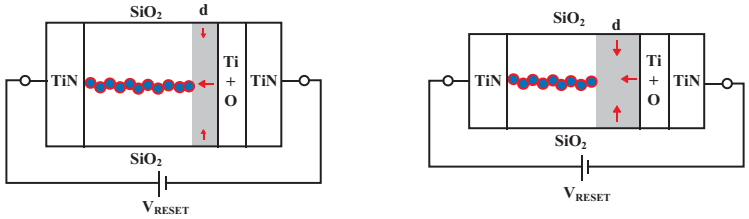


Fig. 10. HRS increases with scaled device area.

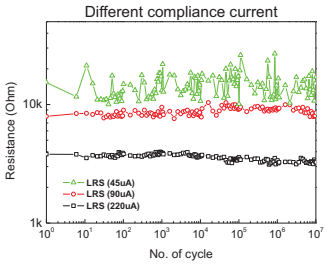


Fig. 11. Cycling fluctuation.

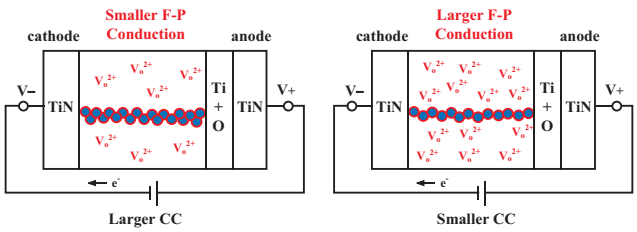


Fig. 12. Lower LRS (higher CC) shows less resistance fluctuation.

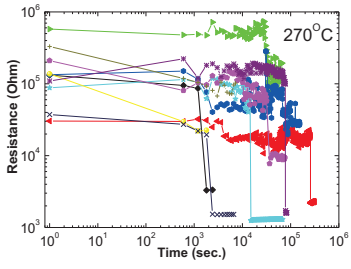


Fig. 13. HRS retention problem (soft errors).

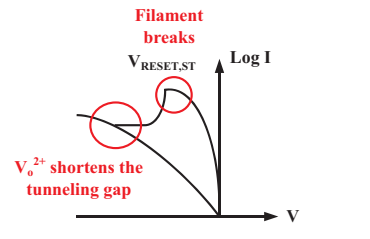
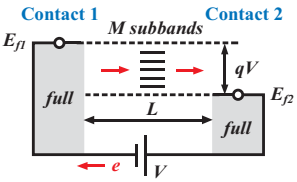


Fig. 14. Two mechanisms compete in the gradual RESET process.



$$I = \frac{2q}{h}\int_{E_{f1}}^{E_{f2}} T(E)M(E)(f_1 - f_2)dE$$

Fig. 5. Quasi-ballistic transport of LRS.