Intrinsic nanofilamentation in resistive switching
Xing Wu, Dongkyu Cha, Michel Bosman, Nagarajan Raghavan, Dmitri B. Migas, Victor E. Borisenko, Xi-Xiang Zhang, Kun Li, and Kin-Leong Pey

Citation: Journal of Applied Physics 113, 114503 (2013); doi: 10.1063/1.4794519
View online: http://dx.doi.org/10.1063/1.4794519
View Table of Contents: http://scitation.aip.org/content/aip/journal/jap/113/11?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in
Intrinsic SiOx-based unipolar resistive switching memory. II. Thermal effects on charge transport and characterization of multilevel programing

Quantum-size effects in hafnium-oxide resistive switching
Appl. Phys. Lett. 102, 183505 (2013); 10.1063/1.4802265

Resistive switching mechanisms relating to oxygen vacancies migration in both interfaces in Ti/HfOx/Pt memory devices
J. Appl. Phys. 113, 064510 (2013); 10.1063/1.4791695

Robust unipolar resistive switching of Co nano-dots embedded ZrO2 thin film memories and their switching mechanism

Non-volatile high-speed resistance switching nanogap junction memory
Intrinsic nanofilamentation in resistive switching

Xing Wu,1,2 Dongkyu Cha,3 Michel Bosman,4 Nagarajan Raghavan,5 Dmitri B. Migas,6 Victor E. Borisenko,6 Xi-Xiang Zhang,3 Kun Li,3,a) and Kin-Leong Pey1,a)

1Singapore University of Technology and Design (SUTD), Singapore 138682
2SEU-FEI Nano-Pico Center, Key Laboratory of MEMS of Ministry of Education, Southeast University, Nanjing 210096, China
3Imaging and Characterization Core lab, 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia
4Institute of Materials Research and Engineering, A*STAR (Agency for Science, Technology and Research, 3 Research Link, Singapore 117602
5Division of Microelectronics, School of Electrical and Electronics Engineering, Nanyang Technological University (NTU), Singapore 639798
6Belarusian State University of Informatics and Radioelectronics, P.Browka 6, Minsk 222013, Belarus

(Received 7 January 2013; accepted 21 February 2013; published online 15 March 2013)

Resistive switching phenomena in metal-insulator-metal structures have been observed many decades ago.1 Correlated electrons in resistance random access memories (RRAM) based on transition metal oxides (TMO) exhibit contrasting electrical properties, from insulator to conductor, and vice versa.2,3 Recently, this phenomenon of resistive switching has gained renewed interest for nonvolatile data storage and reconfigurable electronics, considering the demand for exponential increase in data storage capacity.4–6 Electrically programmable RRAM has multiple advantages as compared to current FLASH memory technology which include fast writing/erasing times (<100 ns), low retention lifetime (>10 yr), low power consumption, simplified structure, ease of design, scalability and Si-based complementary metal oxide semiconductor (CMOS) process technology compatibility.7,8 In addition, RRAM is a good candidate for three-dimensional stacking, which further enhances data storage capability into the terabyte regime.9 RRAM is based on the reversible dielectric breakdown of TMO. From a microscopic point of view, resistive switching in various materials can be classified into three mechanisms: metallization, valence change, and fuse-antifuse.8

First, in electrochemical metallization mechanism, the migration of the highly mobile electrode metal forms a conductive filament during SET and can be electrochemically dissolved upon reversed polarity applied voltage. For example, it is convincingly demonstrated in Refs.10 and 11 that the electrochemically active electrode metal Ag⁺ ions can punch through the insulating layer, setting the device to ON state. The metal filament undergoes dissolution at high electrochemical current (also called Faradaic current) when a reverse voltage bias is applied, resetting the system to OFF state. The insulating layer for this mechanism is usually a solid electrolyte layer without oxygen ions, i.e., Ag₂S and Ag-Ge-Se. Second, in valence change mechanism, specific TMO materials are chosen as the insulating layer (such as TiO₂, NiO, HfO₂).3,12–15 The creation and migration of the oxygen vacancies leads to a redox reaction expressed by the valence state change of the cations. Upon SET, dislodged oxygen species from the TMO migrate to the electrode material (typically made of Ni, Ti, Ta, etc.), which serves as a good reservoir for oxygen in the ionic form.16 Under a reverse bias, the O₂⁻ ions stored in the gate are driven back to the oxygen-deficient conductive path. The devices driven by the electrochemical metallization mechanism and the RRAM devices of valence change mechanism usually show bipolar behavior, in which SET and RESET voltages are of opposite bias polarities. Third, in the fuse-antifuse mechanism, the conductive filament fusion is triggered by an electric field assisted thermal runaway process. It is well documented that in most insulating and semiconducting materials, the conductivity of the material is dependent exponentially on the temperature.17 In the antifuse process, the conductive filament is

---

I. INTRODUCTION

Resistive switching phenomena in metal-insulator-metal structures have been observed many decades ago.1 Correlated electrons in resistance random access memories (RRAM) based on transition metal oxides (TMO) exhibit contrasting electrical properties, from insulator to conductor, and vice versa.2,3 Recently, this phenomenon of resistive switching has gained renewed interest for nonvolatile data storage and reconfigurable electronics, considering the demand for exponential increase in data storage capacity.4–6 Electrically programmable RRAM has multiple advantages as compared to current FLASH memory technology which include fast writing/erasing times (<100 ns), long retention lifetime (>10 yr), low power consumption, simplified structure, ease of design, scalability and Si-based complementary metal oxide semiconductor (CMOS) process technology compatibility.7,8 In addition, RRAM is a good candidate for three-dimensional stacking, which further enhances data storage capability into the terabyte regime.9 RRAM is based on the reversible dielectric breakdown of TMO. From a microscopic point of view, resistive switching in various materials can be classified into three mechanisms: metallization, valence change, and fuse-antifuse.8 First, in electrochemical metallization mechanism, the migration of the highly mobile electrode metal forms a conductive filament during SET and can be electrochemically dissolved upon reversed polarity applied voltage. For example, it is convincingly demonstrated in Refs.10 and 11 that the electrochemically active electrode metal Ag⁺ ions can punch through the insulating layer, setting the device to ON state. The metal filament undergoes dissolution at high electrochemical current (also called Faradaic current) when a reverse voltage bias is applied, resetting the system to OFF state. The insulating layer for this mechanism is usually a solid electrolyte layer without oxygen ions, i.e., Ag₂S and Ag-Ge-Se. Second, in valence change mechanism, specific TMO materials are chosen as the insulating layer (such as TiO₂, NiO, HfO₂).3,12–15 The creation and migration of the oxygen vacancies leads to a redox reaction expressed by the valence state change of the cations. Upon SET, dislodged oxygen species from the TMO migrate to the electrode material (typically made of Ni, Ti, Ta, etc.), which serves as a good reservoir for oxygen in the ionic form.16 Under a reverse bias, the O₂⁻ ions stored in the gate are driven back to the oxygen-deficient conductive path. The devices driven by the electrochemical metallization mechanism and the RRAM devices of valence change mechanism usually show bipolar behavior, in which SET and RESET voltages are of opposite bias polarities. Third, in the fuse-antifuse mechanism, the conductive filament fusion is triggered by an electric field assisted thermal runaway process. It is well documented that in most insulating and semiconducting materials, the conductivity of the material is dependent exponentially on the temperature.17 In the antifuse process, the conductive filament is
ruptured by Joule heating.\textsuperscript{18,19} In this case, resistive switching can be achieved with the same bias polarity, and is thus called unipolar switching behavior. Both bipolar and unipolar switching behaviors have been observed in HfO$_2$ based dielectric stacks.\textsuperscript{20} Several materials also show both types of switching behaviors and are termed as "non-polar."\textsuperscript{21,22}

In this study, we use an asymmetric electrode stack (metal-insulator-semiconductor) to study the effects of the electrode materials on switching. Careful \textit{in-situ} electrical stressing was performed in a transmission electron microscope (TEM) on a Ni-electrode/HfO$_2$/n$^{-}$Si structure. Our findings capture the \textit{real-time} observation of the formation and rupture of the nanofilament, demonstrating reversible switching with more than three orders of magnitude, which is compatible to state-of-the-art RRAM devices. Knowledge of the composition, structure, and dimensions of the nanofilament reveals the mechanism of resistive switching in thin oxide films, and provides better understanding of the scalability of such structures for future advanced electronic applications.

The device used in this study is a unipolar RRAM device based on an asymmetric metal-insulator-semiconductor structure: Ni as top electrode, HfO$_2$ as insulator, and n$^{-}$Si substrate as bottom electrode (inset in Fig. 1). Ni has been widely used in the main-stream CMOS technology as a source/drain contact material.\textsuperscript{23} The motivation to replace the bottom metal electrode with a highly doped n$^{-}$-type Si ($N_D \sim 10^{19}$ cm$^{-3}$) is to study the role of the metal electrode on switching. In Fig. 1, the RRAM device shows well-behaved unipolar switching current-voltage ($I$–$V$) characteristics under DC sweep, confirming the switching mechanism to be due to fuse-antifuse process. The demonstrated HfO$_2$-based RRAM shows a high ON/OFF resistance ratio of $\sim 10^4$ with a very low reverse current. The device shows good switching performance after 500 DC cycles.

For the \textit{in-situ} TEM experiment, a two-step electrical stress methodology was developed to capture the real-time evolution of nanofilament by limiting the compliance current at different levels. The unique feature of this methodology is its ability to precisely control the evolution of the nanofilament formation, while the conventional one-step sweep method fails to capture the kinetic evolution of such phenomenon. First, a voltage sweep from 0 to 5 V at a step size of 300 ms/V was applied to the device with a leakage current limited at 10 $\mu$A to avoid catastrophic irreversible damage to the device due to high current flux. During the voltage sweep, when a large current step increase occurs of about two or three times greater than the previous current level, the insulator has been subjected to a soft breakdown stage, as seen in Fig. 2(a). Figures 2(b) and 2(c) show the dynamic observation of a sequence of morphological changes during real-time electrical stressing of the metal-insulator-semiconductor sample in TEM. The Si-substrate was viewed along the (110) direction and the flat surface corresponds to the (100) plane of the conventional Si wafer. At the surface, the (100) planes of the surface face off to the (111) planes at an angle of 54.7°. There was no obvious physical or morphological change at this soft breakdown stage, especially at the interface due to low spatial resolution of the real-time TEM analysis setting. The same sample stressed in real-time was then put into a double tilt holder to take high resolution TEM micrograph. Figure 2(d) shows the high resolution micrograph of a virgin sample for comparison. The thickness of the HfO$_2$ layer is $\sim 3.1$ nm, while the interfacial layer of SiO$_2$ in-between HfO$_2$ and Si substrate is $\sim 0.7$ nm. For conventional RRAM devices reported in the literature, the thickness of the oxide is usually $\sim 10$ nm.\textsuperscript{24,25} The ultrathin oxide layer used in this study sheds light on the future scaling feasibility of RRAM devices. Furthermore, a key advantage of the Si-based bottom electrode is the ultrasmooth interface between Si and oxide, allowing us to study the “intrinsic” behavior of the conductive filament and switching phenomenon without process induced thickness/roughness variations. Figure 2(e) shows the device which suffered a soft breakdown during a real-time stress in TEM, where white dots appear at the edge of the inter-surface (as indicated by the triangular markers). We also observe the presence of a crystalline Si hillock, punching into the SiO$_2$ layer, which is also called the dielectric breakdown induced epitaxy (DBIE).\textsuperscript{26,27} From our previous experiments, it was established that the chemistry of the percolation path responsible for high electrical conduction in the broken-down gate dielectrics is an O-deficient conductive path, where oxygen atoms are washed out from the core of the percolation path.\textsuperscript{20,28–30} It is important to note that in these former cases, the transistor gate electrode is polycrystalline silicon not a metal.

After the voltage sweep, a constant voltage stress was applied at 3.3 V, about 0.7 V lower than the breakdown voltage, with the set leakage current limited at 100 $\mu$A, as shown in Fig. 3(a). This approach of lowering 0.7 V with reference to the breakdown voltage is commonly used in dielectric breakdown studies.\textsuperscript{31} The stressing was halted once defects such as filaments were observed from the real time $I$–$V$ curve and the real-time TEM image monitoring. Figures 3(b)–3(e) show the micrographs taken at different breakdown stages from a video. Heavy atoms from the top metal electrode (i.e., gate) suddenly migrated into the HfO$_2$ dielectric and Si-substrate, forming a unique geometrical defect which was an

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Current-voltage of a device with metal/insulator/semiconductor structure: Ni/HfO$_2$/SiO$_2$/Si-diode. Note that the reverse current is very low. The inset shows the schematic diagram for RRAM integration with a Si-substrate as selector. The device shows good switching performance after 500 DC cycles (red lines).}
\end{figure}
inverted pyramid if viewing it from the Si substrate. On top of the triangular defect in the Si-substrate, a “depleted” region with a lower image contrast in the top electrode could be seen. This is an indication that the Ni atoms from the anode migrated down. A typical metal migration process requires several tens of transformation steps. Each of the transformations should occur in the range of pico- or nano-seconds. It is therefore reasonable that no intermediate stage during the migration could be observed in these experiments, because this far beyond the time resolution of our image-recording device. A high resolution TEM micrograph in Fig. 3(f) further confirms the shape of the filament. The filament inside the dielectric is of a bowl-shape with a wider top of ~20 nm near the anode and shorter bottom of ~10 nm. Figure 3(g) shows the high angle annular dark field (HAADF) STEM of the filament, where the bright contrast regions in the HfO2 and SiOx layers indicate that heavy metal atoms from the top metal electrode have diffused and/or drifted into the dielectric layer.

In order to study the chemistry of the filament, we employ the Z-contrast STEM analysis (contrast dominated by atomic number), EELS, and energy dispersive x ray (EDX) analysis techniques. EELS and EDX signals were collected under 80kV e-beam rather than 300kV to avoid possible sample damage under strong electron irradiation. Figure 3(h) shows the EELS mapping results of Ni, Hf, and Si. It reveals that the nanofilament is Ni-element-rich. Ni atoms diffused into the insulator layers and reached the Si-substrate, forming a triangular shaped defect. (i) EDX signals from the center of the filament (solid line) and the intact area (dashed line). The results prove that Ni atoms from the anode were the main source for the conductive filament formation constituting to the chemical composition of the nanofilament during SET.
Si. It is clear that Ni atoms diffused in the dielectric from the top electrode, which is the anode side. Ni atoms continued to diffuse along the \(\{111\}\) direction of the silicon substrate due to the largest inner-plane distance. The collective effect of many Ni stacking faults finally results in the triangular shaped morphology. Such needle-shaped Ni defects due to metal diffusion along \(\{111\}\) planes have also been reported as a “process issue” in the conventional NiSi-based CMOS process for the 10 nm technology node. From a physical analysis perspective, the Ni in the silicon substrate forms a good marker indicating the location of the nanofilament inside the dielectric. The EDX line profiles of various elements of interest were further collected from the center of the filament (solid line) and the intact area (dashed line), as shown in Fig. 3(i). The relative difference in the signals confirms the diffusion of Ni from the top electrode into the dielectric and the silicon substrate.

Figure 4(a) shows the real-time electrical data during RESET. The ON/OFF ratio is more than 1000. The device was significantly “switched-off,” remaining at a very low current level. Figures 4(b) and 4(c) show the real-time TEM micrographs wherein the physical changes were too small to be resolved and observed. High resolution TEM and STEM micrographs in Figs. 4(d) and 4(e) clearly show that the filament has ruptured. The filament is disconnected in the dielectric, and hardly any contrast is detectable inside the HfO\(_2\) or SiO\(_x\) layer. One of the major driving forces that causes the “switch-off”/rupture of the nanofilament is the Joule heating at high current prior to the instant of RESET. There were still some Ni residuals in the dielectric, but their number is much less than after the forming/SET transition. EELS and EDX results in Figs. 4(f) and 4(g) show that Ni has been depleted from the Ni-rich filament, as observed previously in Fig. 3(e). The whole process, including electrical SET and RESET, could be repeated many times on the same filament, and the switching voltage and current are well reproduced during the cycles which we observed on other real-time TEM analyzes.

The process of SET and RESET of the filament is illustrated in Figs. 5(a)–5(e). Starting from a virgin unstressed device (Fig. 5(a)), the formation of the conductive filament is a two-stage process. The first stage is soft breakdown of the dielectric (Fig. 5(b)) as established in various logic gate stack reliability studies (i.e., with low compliance current of approximately 1–10 \(\mu\)A). In this stage, oxygen vacancies are
the physical defects responsible for the formation of the percolation path. Oxygen atoms are washed away from the percolated core, followed by the formation of DBIE. The formation of DBIE is the finger-print of the existence of the breakdown path. In the second stage (Fig. 5(c)), as the compliance current level is raised to a higher level (i.e., ∼10 μA-1 mA), metal atoms from the top electrode (anode) migrate along the oxygen deficient breakdown path, driven by the high current density and temperature enhanced metal atom diffusion.34,35 The diameter of the metal nanofilaments ∼10-20 nm. The schematic of the filament rupture and meltdown due to Joule heating during the unipolar RESET process is presented in Fig. 5(d). Some Ni atoms might remain in the insulator layers but the metal fragment dielectric is no longer present in Fig. 5(d). Some Ni atoms might remain in the insulator layers but the metal fragment dielectric is no longer conductive (Fig. 5(e)).

It is obvious that the presence of Ni fragments in the dielectric is critical for the switching behavior, as it determines the dielectric resistivity and connectivity between the top electrode and the bottom electrode. These fragments play a prominent role in the reversible switching transitions for many endurance cycles. To develop a microscopic model for the filament, we performed first-principle calculations of different HfOxNiy stoichiometries. As shown in Fig. 6(a), the perfect HfO2 (x = 2, y = 0) has a bandgap value of ∼4.6 eV in this case by means of the projector-augmented wave method implemented in the VASP code.36–38 When oxygen atoms are removed from the lattice, oxygen vacancy defects form. In the SET state, Ni atoms diffuse and fill in the oxygen vacancy sites (take x = 1.5, y = 0.5 as an example), forming a connected nanofilament path. The bandgap of this chemical compound HfO1.5Ni0.5 collapses to zero (Fig. 6(b)), indicating a good ohmic conductor. In the RESET state, Ni atoms dislodge from the oxygen vacancy sites (y < 0.5 as an example), raising the correlated hopping barrier for electrons. This phenomenon has also been found in nickel oxide based RRAM device.39 Sub-bands subsequently appear in the collapsed bandgap region (Fig. 6(c)), and the conductor transforms into an insulator.

In summary, the real-time, in-situ TEM analysis on a metal-insulator-semiconductor structure allows the study of the dynamics, evolution, and underlying physics governing the switching mechanism which was previously at best speculated from electrical measurements without much direct evidence. We present results that convincingly show the conduction path formation in the SET stage and rupture in the RESET state in RRAM devices driven by non-polar Joule heating, which can only be attributed to the fuse-antifuse mechanism. The use of an asymmetric metal-insulator-semiconductor structure clearly identifies the presence of metal-rich conductive filaments that originate from the anode electrode. The nickel fragments present in the dielectric after filament rupture (or RESET) helps in localizing and triggering further switching events at the same location, which enhances controllability and endurance of switching, and reduces variability. This quantitative study on the filament’s chemistry, morphology, and its dynamical behavior lends strong support to the feasibility of scaling future RRAM technologies further, paving way for very high density data storage.

ACKNOWLEDGMENTS

This research work was funded by the Singapore University of Technology and Design (SUTD) Research Grant No. SRG ASPE 2010 004.