Solid-State Electronics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

Improvement of performances HfO_2-based RRAM from elementary cell to 16 kb demonstrator by introduction of thin layer of Al_2O_3

M. Azzaz [a,b,*], A. Benoist [a], E. Vianello [b], D. Garbin [b], E. Jalaguier [b], C. Cagli [b], C. Charpin [b], S. Bernasconi [b], S. Jeannot [a], T. Dewolf [b], G. Audoit [b], C. Guedj [b], S. Denorme [a], P. Candelier [a], C. Fenouillet-Beranger [b], L. Perniola [b]

[a] STMicroelectronics, 850 Rue Jean Monnet, 38920 Crolles, France
[b] CEA, LETI, Minatech Campus, 17 Rue des martyrs, 38054 Grenoble, France

ARTICLE INFO

Article history:
Available online xxxx

The review of this paper was arranged by Jurriaan Schmitz

Keywords:
HfO_2
Al_2O_3
RRAM
Demonstrator
HRS variability
Trap assisted tunneling

ABSTRACT

In this article, the reliability of HfO_2-based RRAM devices integrated in an advanced 28 nm CMOS 16 kbit demonstrator is presented. In order to improve the memory performance, a thin Al_2O_3 layer is inserted in the HfO_2-based memory stack (TiN/Ti/HfO_2/Al_2O_3/TiN). Thanks to extensive electrical characterizations on both single layer HfO_2 and bilayer HfO_2/Al_2O_3 memory stacks at device and array levels, the potential of the bilayer is put forward. From the experimental results, the thin Al_2O_3 layer has allowed to improve the endurance (memory window of about one decade after 1 M cycles) and data retention (both the low and the high resistance states are stable after 6 h at 200 °C). Finally, thanks to our 3D model based on calculation of the Conductive Filament resistance using trap assisted tunneling (TAT) the role of Al_2O_3 as series resistance is highlighted.

1. Introduction

Oxide-based Resistive Random Access Memories (RRAM) featuring fast programming and reading times (few 10's of ns, even <1 ns), good endurance (higher than Flash), good scalability ($10 \times 10\ nm^2$) and compatibility with the conventional CMOS back-end-of-line (BEOL) process [1–4] are promising candidates for future non-volatile memory embedded applications [5,6]. However, one of the main limitations for the industrial fabrication of this technology is the variability of its resistance states, especially the high resistance state (HRS) [7,8].

Among, the different dielectrics proposed in RRAM memory stacks, HfO_2 and Al_2O_3-based stacks attracted growing interest because of their compatibility with typical BEOL bulk CMOS processing [9]. Recent research has demonstrated the improvement of both switching voltage uniformity and dispersion of HRS resistances of the HfO_x/AlO_x bilayer RRAM devices as compared to the single-layer HfO_x ones [10]. Other studies showed that inserting a thin Al_2O_3 layer in-between the HfO_2 switching layer and the metallic bottom electrode allows to improve the data retention [11] and the memory window as well as the operating current (down to the sub-μA current operating regime) [12]. However, all these results have been demonstrated on few single devices and a clear performance assessment obtained on both large range of devices and integrated in a state of the art advanced CMOS technology, is still missing.

In our previous study we have investigated the bipolar switching properties of HfO_2 and Al_2O_3/HfO_2 based resistive switching memory cells integrated into 28 nm design rules CMOS bulk process. The study has been completed by performance analysis at the array-level on 1 kb [4]. In this paper, we investigate the performance of both HfO_2 and Al_2O_3/HfO_2 in term of endurance at a large statistic scales on a16 kb test chip and also the data retention before and after endurance (10^5 cycles) on 1 kb. Thanks to extensive electrical characterizations on 16 kb test chip and device modeling, the interest of the Al_2O_3/HfO_2 bilayer is highlighted.

2. Experimental

(A) RRAM 1T1R bitcell

1T1R structure RRAM bitcell used in this work is composed by an RRAM resistor in series with a 1 mA access NMOS transistor (Fig. 1a). Two different RRAM stacks integrated in BEOL are studied (Fig. 1b). The first one is composed by a 5 nm thick HfO_2 resistive

* Corresponding author at: STMicroelectronics, 850 Rue Jean Monnet, 38920 Crolles, France.
E-mail address: mourad.azzaz@st.com (M. Azzaz).

http://dx.doi.org/10.1016/j.sse.2016.07.007
0038-1101/

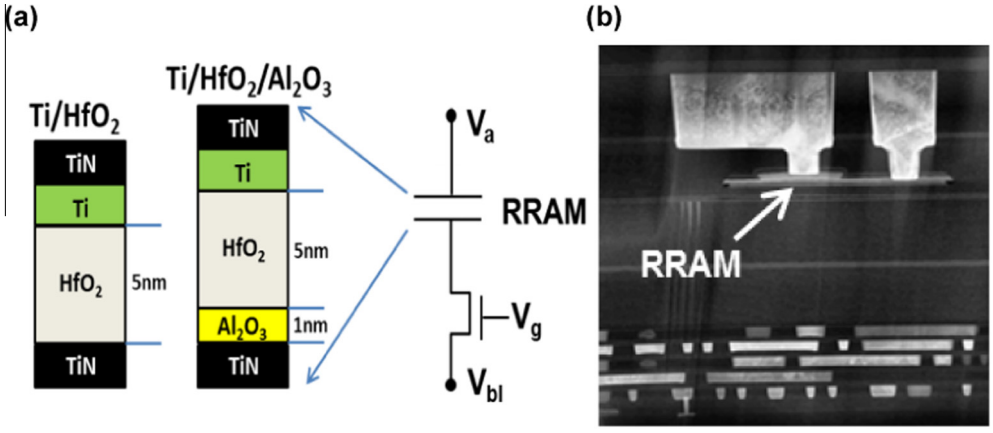

Fig. 1. (a) Schematics of the studied RRAM cells integrated into a 1T1R structure and (b) TEM cross section of a 1T1R device.

switching layer embedded in-between a TiN/Ti Top Electrode (TE) and a TiN Bottom Electrode (BE). The second stack consists in an additional 1 nm Al_2O_3 layer inserted between the TiN BE and the HfO_2 layer (Fig. 1a). Both the resistive switching layers HfO_2 and Al_2O_3 are deposited by Atomic Layer Deposition (ALD) while the metallic electrodes are deposited by Physical Vapor Deposition (PVD).

(B) 28 nm CMOS 16 kb RRAM demonstrator

This bitcell has been embedded in a 28 nm CMOS digital test chip comprising 16 CUTs of 1 kbit coupled with its digital controller (Fig. 2a and b). 1 CUT corresponds to 1kb array. Designing and manufacturing such test chip on an advanced CMOS technology provides key features to evaluate the capability of this solution at a large scale.

3. Results and discussion

In a first part an in-depth electrical characterization was performed on the HfO_2 and HfO_2/Al_2O_3 1T1R bitcells in order to highlight the first trend. In a second part, these results are compared to larger statistical values obtained from our digital test-chip.

(A) 1T1R bitcells performances

For both HfO_2 and HfO_2/Al_2O_3 stacks, 10^6 endurance cycles tests, without both algorithm correction and preconditioning operations, were carried out on 15 and 25 cells respectively. Several series of 100 consecutive readings were performed at low field (0.1 V) during the test (Figs. 3a and 4a). The operating current of 230 μA used in the endurance test was identified as the best trade-off between cycling endurance and the memory window. Indeed lower operating currents reduce drastically the memory window.

For the HfO_2 sample, a degradation of the Low Resistance State (LRS) as well as a decrease of High Resistance State (HRS) are observed after 10^5 cycles. The closing of the memory window after cycling is put forward by the shift of the LRS and HRS distributions for the 100 consecutive cycles performed before and after endurance on a single cell, as shown in Fig. 3b. A memory window of more than 1 decade is maintained after 10^6 cycles for the HfO_2/Al_2O_3 stack. No shift of the LRS and HRS distributions for the 100 consecutive cycles before and after endurance on a single cell can be observed, as shown in Fig. 4b. The role of the Al_2O_3 layer is highlighted by maintaining both the HRS and LRS resistance distributions stable during endurance (Fig. 4a and b). Test conditions are similar to the ones used for the HfO_2 stack except the forming voltage that is higher (4 V instead of 2.5 V) and the RESET voltage that is slightly lower (−1.3 V instead of −1.4 V).

In order to evaluate the thermal stability, Fig. 5a shows LRS and HRS retention evolution for the HfO_2/Al_2O_3 sample versus time at 200 °C. The devices are programmed at 230 μA which has been identified as the best trade-off between endurance and memory window [3]. The LRS state is stable up to 2.1×10^4 s (~6 h) of bake. Fig. 5b compares the retention of HfO_2 (dashed lines) with HfO_2/Al_2O_3 (continuous lines) samples. The introduction of the Al_2O_3 layer greatly improves the thermal stability. The enhanced LRS retention could be explained by the increase of the oxygen vacancies diffusion barrier thanks to the incorporation of Al into the HfO_2 matrix estimated. Estimation has been done with ab-initio calculations [13].

(B) RRAM demonstrator performance

The interest of HfO_2/Al_2O_3 highlighted at the 1T1R device level in the previous section is now evaluated at a larger scale on a 1 kbit demonstrator in terms of endurance and data retention.

(1) Endurance

The evolution of the LRS and HRS states distributions on 1 kb and 16 kb during the endurance up to 10^5 cycles is evaluated for both HfO_2 and HfO_2/Al_2O_3 memory stacks.

Fig. 6 shows the LRS and HRS distributions obtained for the HfO_2/Al_2O_3 memory stack after 100 cycles. The programming conditions are the same as the ones used for the single cell tests excepted for the programming current increased up to 300 μA. Discrete steps in the experimental distributions are given by discrete thresholds in the read current sensing. The experimental distribution is intentionally stopped at ~1 MΩ to cover a wide range of discrete thresholds in the read current sensing.

The evolution of the LRS and HRS distributions is evaluated through endurance tests. Fig. 7a reports the LRS and HRS distributions achieved after 1, 10 and 10^5 cycles for the HfO_2 sample. A small memory window of around 1.5 (between the maximum value of LRS and the minimum value of HRS) is obtained after the first cycle (red[1] line). However, the *pre-conditioning* of 10 cycles, allows to increase the HRS resistance to get the best memory window case around 3 (blue line). The HRS resistance is slightly decreased after 10^5 cycles (black line). These results are in good agreement with the data obtained on single cells (Fig. 3).

[1] For interpretation of color in Figs. 7 and 13, the reader is referred to the web version of this article.

Please cite this article in press as: Azzaz M et al. Improvement of performances HfO_2-based RRAM from elementary cell to 16 kb demonstrator by introduction of thin layer of Al_2O_3. Solid State Electron (2016), http://dx.doi.org/10.1016/j.sse.2016.07.007