Technology and Circuit Optimization of Resistive RAM for Low-Power, Reproducible Operation

D. C. Sekar, B. Bateman, U. Raghuram, S. Bowyer, Y. Bai[+], M. Calarrudo, P. Swab, J. Wu, S. Nguyen, N. Mishra, R. Meyer, M. Kellam, B. Haukness, C. Chevallier, H. Wu[+], H. Qian[+], F. Kreupl*, G. Bronner
Rambus, [+]Tsinghua University, *Technische Universität München
E-mail: dsekar@rambus.com, Phone: 650-804-8415

Abstract

Low-power, reproducible operation of Resistive RAM (RRAM) requires control of capacitive surge currents during write. We propose a fab-friendly TiN/conductive TaO_x/HfO_2/TiN RRAM with a built-in surge current reduction layer. It reduces worst case write current by 33% and fail bit count by 23x compared to conventional RRAM. A novel circuit to control surge current is demonstrated that improves write current by 40% and endurance by 63%. Switching, endurance and retention data for a 256kb chip with these concepts is presented.

Introduction

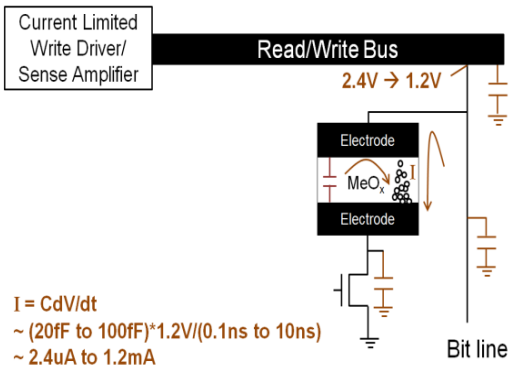


Figure 1: Capacitive surge currents can be high for RRAM chips.

RRAMs are actively pursued for embedded non-volatile memory due to their low power [1] and ease of integration into a logic process [2] compared to flash memory. An important issue for RRAM chips is control of capacitive surge currents during switching. If capacitive surge currents are high during FORM and SET operations (Fig. 1), filaments are known to become stronger and require higher RESET current to break [2][3]. Reproducibility of switching, which is crucial for building commercial products, is also sensitive to capacitive surge currents.

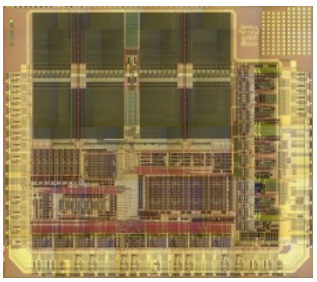


Figure 2: Die photo of 256kb 1T-1R RRAM chip.

In this work, we present results from a 256kb RRAM chip in in which memory devices, circuits and programming algorithms were developed to lower capacitive surge currents (Fig. 2).

TiN/Conductive TaO_x/HfO_2/TiN RRAM

Our proposed RRAM device uses *fab-friendly* materials such as TiN, HfO_2 and conductive TaO_x. Fig. 3 shows the device structure and its measured I-V curve. The conductive metal oxide electrode made of TaO_x plays an important role in device operation, as will be shown later in this paper.

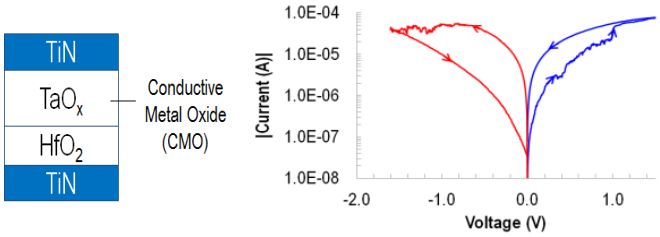
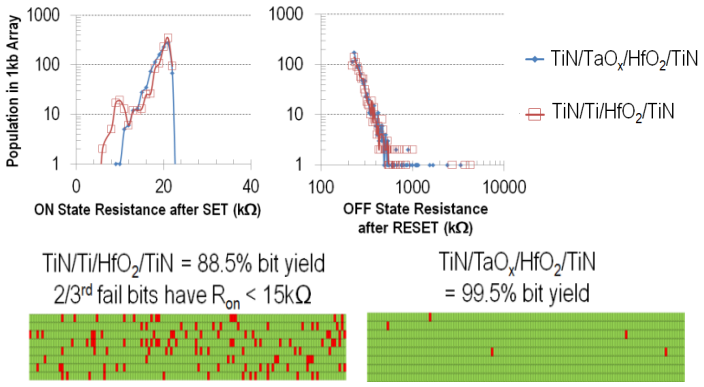


Figure 3: Our proposed device (left), and its measured I-V curve (right).

Fig. 4 shows a comparison of our proposed RRAM with well-studied TiN/Ti/HfO_2/TiN RRAM [4]. 1kb 1T-1R arrays were used for experiments. Median ON state resistance of both RRAM devices is similar in Fig. 4. However, for the same programming algorithm, the TiN/conductive TaO_x/HfO_2/TiN RRAM has 8.6kΩ worst case ON resistance (R_{on}) compared to 5.8kΩ for TiN/Ti/HfO_2/TiN RRAM - a 50% increase. This is important since RESET current is normally a function of R_{on} ($I_{reset} \sim V_{reset}/R_{on}$). Thus, the RRAM with the conductive metal oxide electrode has approximately 5.8kΩ/8.6kΩ = 33% lower worst case RESET current. Cell size is a function of the worst case RESET current, so this is valuable. Fig. 4 shows that the TiN/Ti/HfO_2/TiN RRAM has ~88.5% bit yield, compared to ~99.5% bit yield for our proposed RRAM. This represents a 23x difference in fail bit count. A majority of the failed bits in Fig. 4 come from cells having R_{on} < 15kΩ that have higher RESET currents and see higher temperatures on their oxygen vacancy filaments. Voltages and data retention for the two RRAM devices are quite similar (Fig. 4).

The comparison between these two device types was kept as fair as possible – both device types had their structure and process flow optimized rigorously and the programming algorithm was kept the same, as previously noted. By changing the programming algorithm for TiN/Ti/HfO_2/TiN, improved bit yields were obtained. However, such algorithms degraded power or write time of TiN/Ti/HfO_2/TiN further compared to TiN/conductive TaO_x/HfO_2/TiN, or resulted in write times that were unrealistic for practical applications.



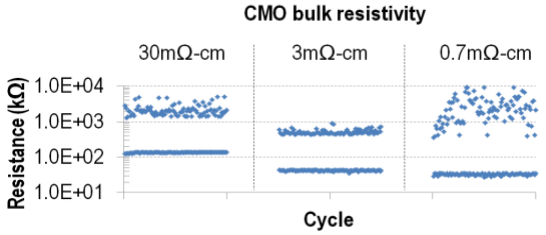
1kb arrays, 32 pulse write	Well-studied RRAM TiN/Ti/HfO_2/TiN [4]	Our proposal TiN/TaO_x/HfO_2/TiN
FORM Voltage	4V	4V
SET Voltage	1.75V	1.75V
RESET Voltage	1.8V (median)	1.6V (median)
Retention	Similar	

Figure 4: Comparison of the proposed device to well-studied TiN/Ti/HfO_2/TiN [4] RRAM. 1kb 1T-1R arrays are used for experiments. Bit yield is defined as percentage of cells in the 1kb array that pass 5 cycles. In the bit maps, = failed bits, = passed bits.

Understanding Switching Improvement for RRAMs with Conductive Metal Oxide Electrodes

We observed in Fig. 4 that for a typical programming algorithm, RRAMs with a conductive metal oxide (CMO) electrode provided ~33% lower worst case RESET current and 23x less fail bit count compared to conventional RRAM devices. The following discussion explains the reasons.

A. Reduction in Capacitive Surge Current



CMO bulk resistivity	30 mΩ-cm	3 mΩ-cm	0.7 mΩ-cm
ON state resistance	166kΩ	44kΩ	35kΩ
Retention test (20min @ 110°C)	Failed	Passed	Passed

Figure 5: Higher Conductive Metal Oxide (CMO) resistance gives higher ON state resistance, indicating CMO resistance reduces surge current during SET.

Fig. 5 shows ON and OFF state resistances measured during cycling for three different CMO resistivity values. A higher resistivity CMO causes higher ON resistance. This indicates that a higher CMO resistance reduces current flowing through the RRAM during SET, leading to weaker filaments that can switch well but have degraded retention (Fig. 5). A device designer can therefore pick an optimal CMO resistivity that gives a good tradeoff between switching and retention properties. Simulations in Fig. 6 indicate that significant voltage can drop across the CMO during SET. This confirms that the CMO resistance and the IR drop across it play an important role in device operation. Tighter ON state resistance distributions in Fig. 4 are also likely due to the CMO reducing capacitive surge currents during SET.

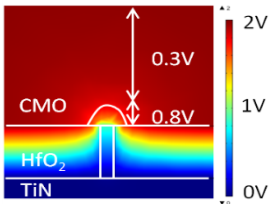


Figure 6: Simulations show current crowding at the CMO-filament interface impacts CMO voltage drop significantly.

B. Improved Thermal Efficiency

Fig. 6 shows an interesting property. A substantial amount of the CMO voltage drop occurs due to current crowding resistance as current spreads from the narrow filament to the wide top electrode. This current crowding causes Joule heating and temperature increases in the RRAM (Fig. 7). In addition, the higher thermal resistance of CMOs compared to metal electrodes prevents heat flow to the surroundings and leads to higher temperatures. RRAMs with conductive metal oxide electrodes can therefore provide several times higher temperature than conventional RRAM devices for the same write current, as shown in Fig. 7. This is advantageous, since ion motion and RRAM switching require high temperatures on filaments, and one can get these high temperatures with lower write currents when RRAMs have conductive metal oxide electrodes. This could be one reason why the RESET voltage in Fig. 4 does not increase much despite the IR drop one would expect from the CMO resistance.

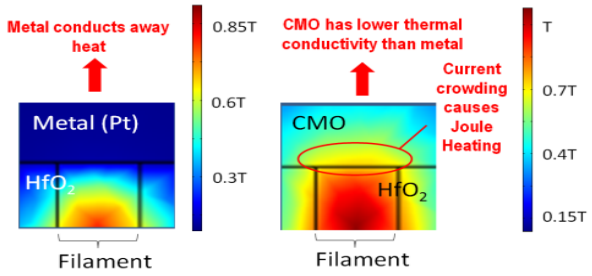


Figure 7: Simulations show that a device with a CMO has higher temperature for the same write current, allowing easier ion motion.

To understand conduction mechanism, data from 128 TiN/TaO_x/HfO_2/TiN RRAM cells was analyzed. Fits of the data to Schottky conduction models [5], Nearest Neighbor Hopping [5], Variable Range Hopping [6] and Frenkel Poole conduction [5] were attempted. Nearest Neighbor Hopping theory provided the best explanation for OFF state conduction while Variable Range Hopping theory provided the best explanation for ON state conduction, as shown in Fig. 8. The observation that hopping is the conduction mechanism indicates HfO_2 film optimization is crucial for improving our RRAM device.

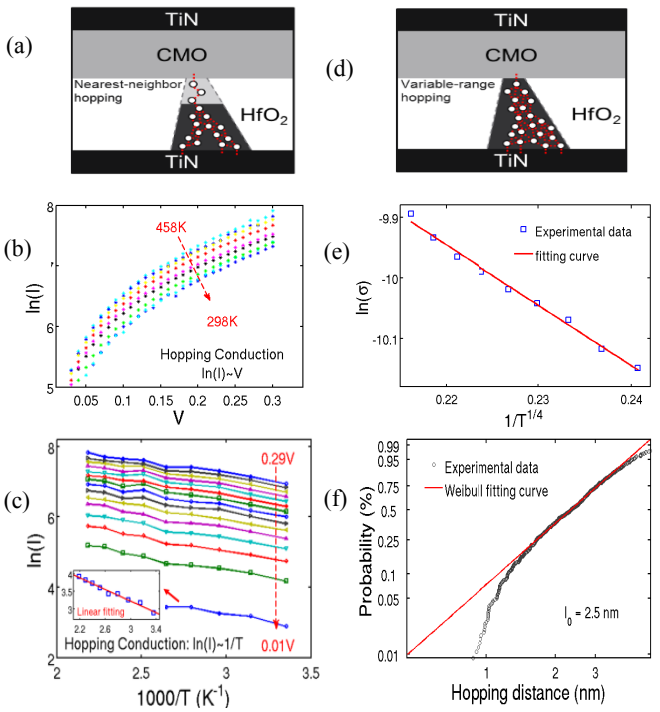


Figure 8: (a) Nearest Neighbor Hopping (NNH) model for OFF state. (b) Measured OFF state I-V curves fit NNH model (c) Measured temperature trends fit NNH model (d) Variable Range Hopping (VRH) model for ON state (e) Measured temperature trends for ON state conductivity fit VRH model (f) Extracted hopping distance is 2.5nm.

Circuits to Reduce Capacitive Surge Currents in RRAMs

A common circuit technique to reduce capacitive surge current is shown in Fig. 9(a). Gate voltage of the select transistor is incremented so that the select transistor resistance limits surge current during SET. Fig. 9(b) reveals the challenges with this approach. To get a write current of ~100µA, the select transistor needs to be biased near its threshold voltage (V_t). V_t variation and exponential change of current near V_t make current control difficult.

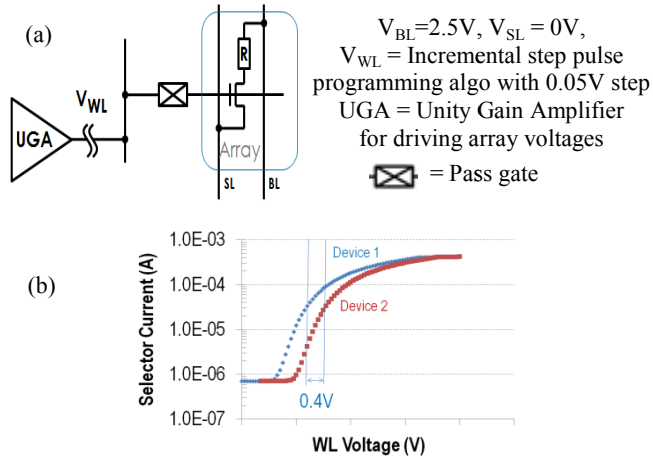
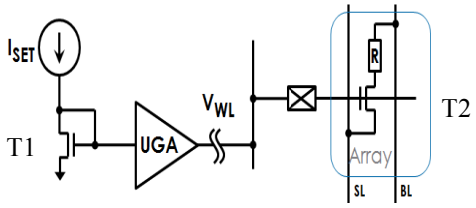


Figure 9: (a) A common circuit approach to reduce capacitive surge currents during SET and FORM (b) Current of different transistors in an array can be substantially different near the threshold voltage.

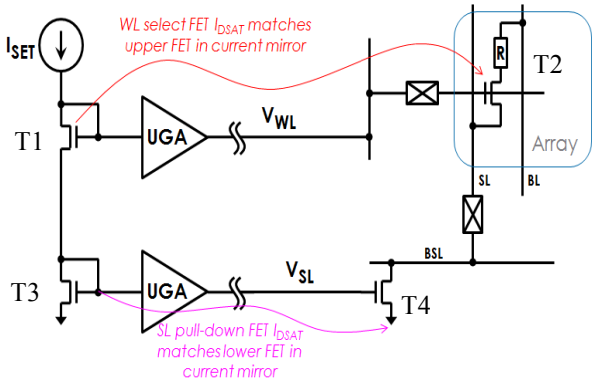
To resolve this, a novel circuit was developed. One component of the circuit is shown in Fig. 10. A cell-level current mirror is used to limit current flow and provide immunity to temperature variations. Any die-to-die or wafer-to-wafer V_t variation impacts both transistors T1 and T2 in the current mirror, so their impact is minimized compared to Fig. 9(a). However, since T2 is a small-size memory FET, V_t mismatch between T1 and T2 could be a challenge.



V_{BL}=2.5V, V_{SL} = 0V, I_{SET} = Incremental step pulse programming algorithm with 10µA step

Figure 10: Current mirror circuit to limit T2's current to I_{SET}.

To tackle mismatch between T1 and T2, transistors T3 and T4 are added to form the capacitive surge current reduction circuit in Fig. 11. The Source Line transistor T4 is shared between all cells on a Source Line (SL) and can be sized big to reduce mismatch effects. T2 reduces surge currents due to BL capacitance.



V_{BL}=2.5-3V, V_{SL} = 0V, I_{SET} = Incremental step pulse programming algorithm with 10µA step

Figure 11: Proposed capacitive surge current reduction circuit for RRAM.

Fig. 12 indicates a 40% reduction in RESET current is obtained with a circuit implementation of Fig. 11 for a 256kb array. In addition, reduction of capacitive surge currents improves control of oxygen vacancy filaments and gives a 63% increase in endurance for the 256kb array (Fig. 13).

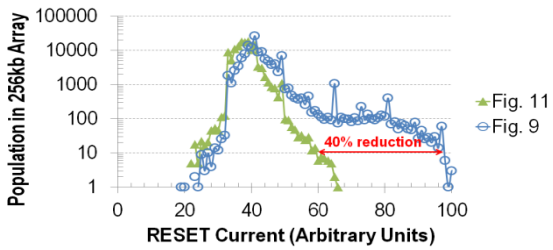


Figure 12: Approximate RESET current distributions for a 256kb array with two current compliance circuits.

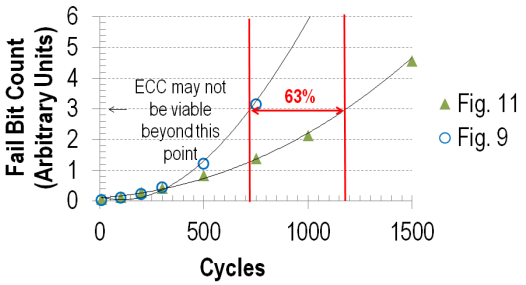


Figure 13: Endurance for a 256kb array with two current compliance circuits.

256kb Chip Results

Our 256kb 1T-1R RRAM chip (Fig. 2) has the memory device shown in Fig. 3 and the capacitive surge current reduction circuit shown in Fig. 11. In addition, the 256kb chip has state machines, sense amplifiers and various analog blocks. The chip's proper operation is demonstrated by the written bit patterns shown in Fig. 14.

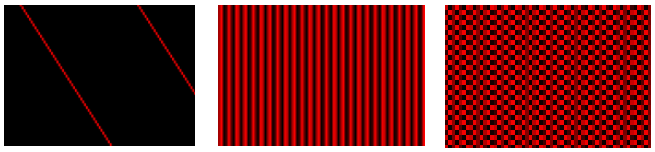


Figure 14: Various bit patterns demonstrated in the 256kb chip, indicating it works well. ='1', ='0'.

As Fig. 15(a) indicates, SET voltages are less than 2.25V, with bits in the 256kb array requiring an average of fourteen 40ns SET pulses. RESET voltages are less than 3V, with a large percentage of bits switching in the first 2V pulse, as shown in Fig. 15(b)

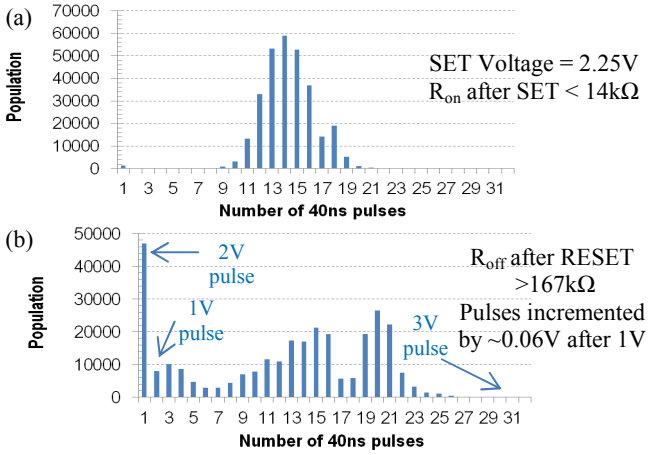


Figure 15: Number of pulses after which bits in the 256kb array finish (a) SET (b) RESET.

Endurance is greater than 1000 cycles, as shown in Fig. 16. This is sufficient for many applications such as code storage. Some degradation is occurring during switching since the number of pulses required for write increases with cycling. Fig. 16 also indicates tail bits in the 256kb array meet 70°C 10 year data retention specs.

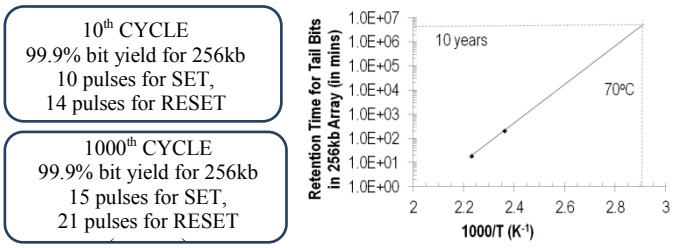
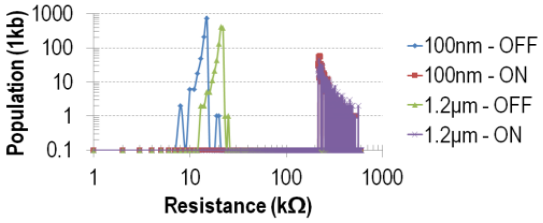


Figure 16: 256kb chip results as a function of Cycling (left) and Temperature of Retention Bake Test (right).

Scaling Trends

Fig. 17 shows scaling behavior of the TiN/conductive $TaO_x/HfO_2/TiN$ RRAM by analyzing two very different device sizes. The switching behavior is similar or improved at small dimensions.



1kb arrays	1.2 μm	~100 nm
RESET Voltage (median)	1.7V	1.75V
SET Voltage (median)	<3V	<3V
ON-OFF Ratio (worst case)	10x	15x
Retention	x	6x

Figure 17: Scaling trends for 1kb arrays.

Summary

To build large RRAM arrays suitable for practical applications, it is important to have reproducible RRAM switching and low write power. A fab-friendly TiN/conductive $TaO_x/HfO_2/TiN$ RRAM was proposed and demonstrated in this paper. For the same programming algorithm, it improved worst case write current by 33% and fail bit count by 23x compared to well-studied $TiN/Ti/HfO_2/TiN$ RRAM. Our studies show this improvement is because the conductive metal oxide layer serves as a built-in capacitive surge current reduction layer and because it improves thermal efficiency. A circuit to reduce capacitive surge current was proposed, which improved write current by 40% and endurance by 63%. The proposed RRAM device and surge current reduction circuit were demonstrated on a 256kb chip. Results in this paper indicate technology and circuit optimizations can reduce capacitive surge currents in RRAM circuits and improve device operation quite significantly.

References

[1] R. Aitken, et al., Symp. on VLSI Technology, 2014.
[2] H. S. P. Wong, et al, Proceedings of the IEEE, 2012.
[3] Y. Sato, et al., IEEE Transactions on Electron Devices, vol.55, no.5, pp.1185,1191, May 2008
[4] H. Y. Lee, et al., Intl. Electron Device Meeting, 2008.
[5] F-C. Chiu, Advances in Materials Science and Engg., 2014.
[6] N. F. Mott, E. Davis, Oxford University Press, 1979.