A Bulk FinFET Unified-RAM (URAM) Cell for Multifunctioning NVM and Capacitorless 1T-DRAM

Jin-Woo Han, Seong-Wan Ryu, Sungho Kim, Chung-Jin Kim, Jae-Hyuk Ahn, Sung-Jin Choi, Jin Soo Kim, Kwang Hee Kim, Gi Sung Lee, Jae Sub Oh, Myeong Ho Song, Yun Chang Park, Jeoung Woo Kim, and Yang-Kyu Choi

***Abstract*—A bulk FinFET-based unified-RAM (URAM) cell technology is demonstrated for the fusion of a nonvolatile-memory (NVM) and capacitorless 1T-DRAM. An oxide/nitride/oxide layer and a floating-body are combined to perform a URAM operation in a single transistor. A buried n-well technology for NMOS allows hole accumulation for the 1T-DRAM operation in a p-type bulk substrate. The bulk FinFET URAM offers a cost-effective and fully compatible process with a conventional FinFET SONOS, and it also expedites heat dissipation. Highly reliable NVM and high-speed 1T-DRAM operation are confirmed, and it was also verified that there is no disturbance between the two memory functions.**

***Index Terms*—Bulk FinFET, capacitorless DRAM, FinFET, nonvolatile memory (NVM), SONOS, unified-RAM (URAM), 1T-DRAM.**

I. Introduction

INTEGRATION of various memory functions in a chip is attractive for embedded systems. Nonvolatile memory (NVM) and DRAM can be integrated in the memory block in a single chip. However, fusion of these memories at a system level is obstructed by poor process compatibility, high cost, and low yield. Accordingly, various device level fusion concepts have been presented [1]–[3]. Unified-RAM (URAM), which operates NVM and capacitorless 1T-DRAM in a single transistor, was proposed with a FinFET-based structure [3]. By implementation of oxide/nitride/oxide (O/N/O) as an electron trapped zone for NVM and a partially depleted floating-body as a hole storage zone for capacitorless 1T-DRAM, URAM operation was realized. This memory was fabricated on a silicon-on-insulator (SOI) wafer utilizing a conventional FinFET SONOS process with minor modifications. However, SOI has an inherent difficulty regarding heat dissipation, which degrades the sensing window for 1T-DRAM operation [4]. Additionally, SOI has a difficulty in cost-effective manufacturing.

Manuscript received February 25, 2008; revised March 24, 2008. This work was supported in part by the National Research Program for the 0.1-Terabit Nonvolatile Memory Development, sponsored by the Ministry of Knowledge Economy. The review of this letter was arranged by Editor S. Kawamura.

J.-W. Han, S.-W. Ryu, S. Kim, C.-J. Kim, J.-H. Ahn, S.-J. Choi, and Y.-K. Choi are with the Division of Electrical Engineering, School of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea (e-mail: ykchoi@ee.kaist.ac.kr).

J. S. Kim, K. H. Kim, G. S. Lee, J. S. Oh, M. H. Song, Y. C. Park, and J. W. Kim are with the National Nanofab Center, Daejeon 205-806, Korea.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2008.922142

In this letter, a bulk FinFET-based URAM is presented. In order to form the floating-body, a buried n-well layer is employed in the bulk substrate. Highly reliable NVM and high-speed 1T-DRAM operation are confirmed.

II. Device Design and Fabrication

A p-type (100) bulk wafer is used as a starting material. After n+ deep implantation is carried out for the buried n-well layer, the Si fin is patterned. High density plasma (HDP) SiO_2 is deposited and planarized. The HDP SiO_2 is then partially recessed until the Si fin is exposed to a height of 100 nm. Next, an O/N/O layer with a thickness of 4/6/4 nm and an *in situ* n+ poly-Si is deposited sequentially. After gate patterning, source/drain (S/D) implantation and activation steps are carried out. The device has a 30-nm fin width, a 100-nm fin height, and a 180-nm gate length. Fig. 1(a) shows images of the fabricated bulk FinFET SONOS. Fig. 1(b) shows the secondary ion mass spectrometry (SIMS) profiles at the S/D region.

III. Results and Discussion

Data states for NVM are identified by the existence of trapped charges in the nitride layer. The program/erase (P/E) can be carried out by Fowler–Nordheim (FN) tunneling or channel hot-electron injection. In this letter, FN tunneling is used for P/E. Fig. 2(a) shows the P/E transient characteristics of NVM. The P/E conditions $V_{\mathrm{G,PGM}} = 11$ V with $\tau_{\mathrm{G,PGM}} = 100\ \mu$s for programming and $V_{\mathrm{G,ERS}} = -15$ V with $\tau_{\mathrm{G,ERS}} = 10$ ms for erasing exhibit a 4.5-V threshold voltage window. During the P/E, S/D/substrate terminals are grounded ($V_{\mathrm{SUB}} = V_S = V_D = 0$ V). A retention period of ten years is obtained with a 3-V threshold voltage window, as shown in Fig. 2(b).

Data states of 1T-DRAM are stored as the floating-body capacitance. In programming, holes created by impact ionization are accumulated in the floating-body, and the channel potential is lowered. In erasing, accumulated holes are eliminated by the negative drain bias, and the channel potential is raised. As a result, the data states are identified by the presence of excessive holes. Generally, the silicon layer formed on the buried oxide becomes a floating-body in the SOI platform. The valence-band barrier height (at Si/SiO_2 interface) can retain holes, and thus, excess holes can be stored in the floating-body. However, as the SOI substrate does not readily dissipate heat, diffusion current is increased. Thus, the data retention time and the sensing window tend to be degraded. In order to solve this problem, a

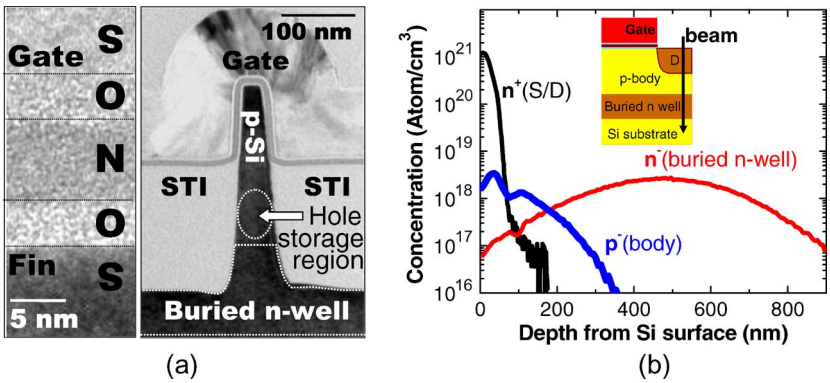

Fig. 1. (a) Transmission electron microscopy images of a bulk FinFET SONOS structure. (b) SIMS profiles at the S/D region. O/N/O gate dielectric is used as a charge trapping layer for an NVM operation, and a p-n junction barrier built with a p-body and a buried n-well is used as a hole accumulation region for capacitorless 1T-DRAM.

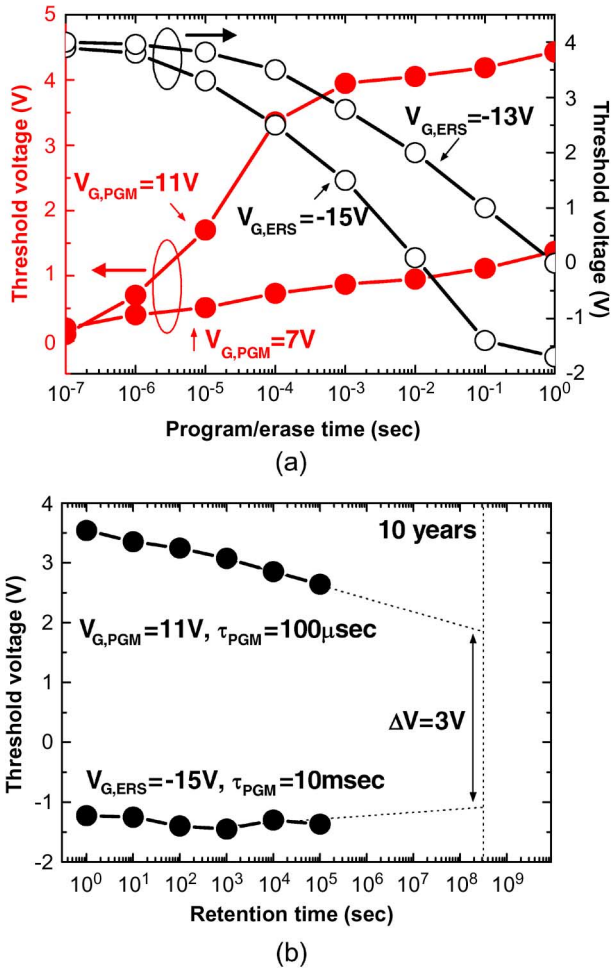

Fig. 2. NVM characteristics. (a) Transient threshold voltage behavior for P/E. (b) Retention characteristic.

buried n-well was employed in a bulk substrate as a hole barrier layer [5], [6]. The body to buried n-well forms a p-n junction potential barrier. Thus, the holes created by impact ionization can be retained inside the body.

Fig. 3 shows the hole concentration biased at the programming voltage, i.e., the impact ionization condition ($V_G = 1$ V and $V_D = 2$ V). As the electron–hole pairs are created by the impact ionization process, electrons are swept out to the drain, and holes move toward the minimum potential region.

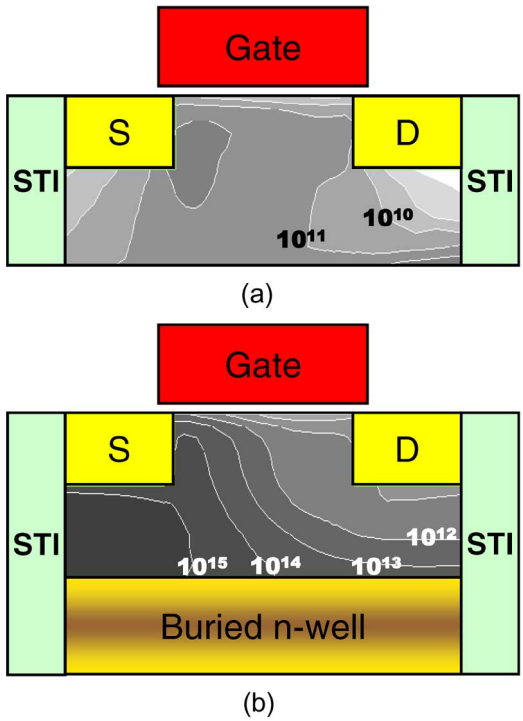

Fig. 3. Silvaco simulation profile of the hole concentration biased at read condition ($V_G = 1$ V, $V_D = 0.4$ V) 1 ns after impact ionization ($V_G = 1$ V, $V_D = 2$ V). The profile is cut-plane along the S/D direction at the center of the fin. (a) Conventional bulk FinFET. (b) Buried n-well bulk FinFET. In contrast with the case of the conventional bulk FinFET, holes created by impact ionization are accumulated in the body region in the buried n-well bulk FinFET.

As shown in Fig. 3(a), most holes are collected by the substrate contact in general case. However, if the buried n-well forms a p-n junction barrier in the substrate, the created holes are confronted with the junction barrier. Thus, holes can be accumulated in the body, as shown in Fig. 3(b), which implies that a capacitorless 1T-DRAM can be realized in the bulk substrate. As a result of hole accumulation, a kink will appear in the I_D–V_D characteristics. Fig. 4 shows the measured I_D–V_D characteristics. The kink observed in the measured data verifies that the holes were accumulated in the bulk substrate. Fig. 5 shows the P/E characteristics for the 1T-DRAM operation. The tunable substrate voltage is an important feature in bulk devices. Because the hole barrier (p-n junction) height can be modulated by the substrate voltage, the 1T-DRAM operation characteristic

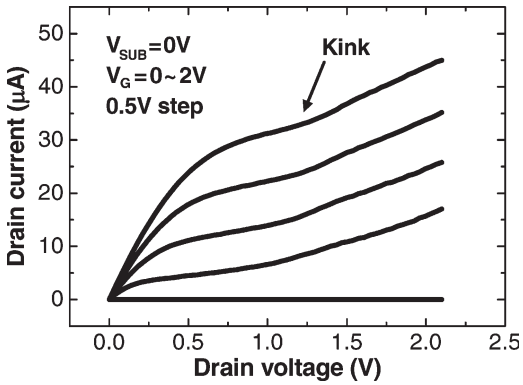

Fig. 4. Measured drain–current (I_D)–drain-voltage (V_D) characteristics. A kink appears as a result of excessive hole accumulation.

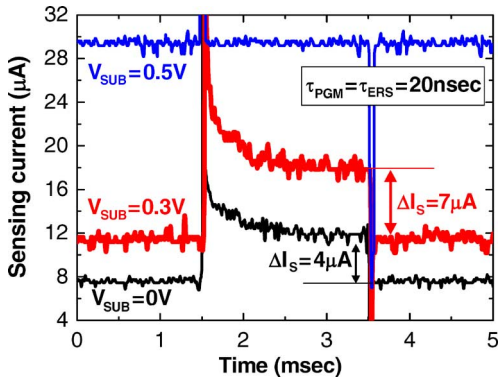

Fig. 5. Measured source current (I_S) for 1T-DRAM operation (read at $V_G = 1$ V, $V_D = 0.4$ V). Whereas two data states are clearly distinguished according to the existence of the holes in the body, the difference between these two states is increased as a small amount of positive voltage is applied to the substrate terminal. However, as forward p-n junction diode turns on, the 1T-DRAM is not working.

would be improved. By applying a weak positive voltage (e.g., $0.1 < V_{\rm SUB} < 0.3$ V) to the substrate, an enlarged hole barrier height enhances the sensing window, as shown in Fig. 5. The sensing windows are $\Delta I_S = 4$ μA and $\Delta I_S = 7$ μA, and the extrapolated data retentions (defined as time at $\Delta I_S = 3$ μA) are 8 and 30 ms for $V_{\rm SUB} = 0$ V and $V_{\rm SUB} = 0.3$ V, respectively. It means that the slight positive substrate voltage can improve the performance of the 1T-DRAM operation. However, in the case of strong positive voltage (e.g., $V_{\rm SUB} = 0.5$ V), 1T-DRAM cannot be operated because the forward-biased source to the body junction diode is turned on.

It is important to note that the sensing window of SOI device will be larger than that of bulk device because the higher barrier height of hole (= valence band offset) can retain more excess holes in the body. In addition, the data retention of SOI device will be longer than that of bulk because the diffusion and recombination can be limited due to the small junction area. Even though the performance of bulk might not be superior to that of SOI, the bulk can be still useful for 1T-DRAM in terms of cost-effective manufacturing and heat dissipation with moderate sensing window.

IV. Conclusion

A FinFET-based URAM is demonstrated in a bulk substrate. The combination of an O/N/O layer and a floating-body provides URAM operation in a single transistor. The floating-body in a bulk substrate is implemented by employing a buried n-well layer. The bulk platform for URAM is an attractive approach as it is cost effective. In addition, bulk facilitates heat dissipation, which suppresses sensing window degradation for 1T-DRAM operation. URAM is expected to see widespread use for fusion memory and future embedded chip applications.

References

[1] C. W. Oh, N. Y. Kim, H. J. Song, S. I. Hong, S. H. Kim, Y. L. Choi, H. J. Bae, D. U. Choi, Y. S. Lee, D.-W. Kim, D. Park, and B.-I. Ryu, "Floating body DRAM characteristics of silicon-on-ONO (SOONO) devices for system-on-chip (SoC) application," in *VLSI Symp. Tech. Dig.*, 2007, pp. 168–169.

[2] C. W. Oh, S. H. Kim, N. Y. Kim, Y. L. Choi, Y. S. Lee, W. J. Jang, H. S. Lee, H. S. Park, D.-W. Kim, D. Park, and B.-I. Ryu, "A novel multifunctional silicon-on-ONO (SOONO) MOSFETs for SoC applications: Electrical characterization for high performance transistor and embedded memory applications," in *VLSI Symp. Tech. Dig.*, 2006, pp. 58–59.

[3] J.-W. Han, S.-W. Ryu, C. Kim, S. Kim, M. Im, S. J. Choi, J. S. Kim, K. H. Kim, G. S. Lee, J. S. Oh, M. H. Song, Y. C. Park, J. W. Kim, and Y.-K. Choi, "A unified-RAM (URAM) cell for multi-functioning capacitorless DRAM and NVM," in *IEDM Tech. Dig.*, 2007, pp. 929–932.

[4] A. Villaret, R. Ranica, P. Malinge, P. Masson, B. Martinet, P. Mazoyer, P. Candelier, and T. Skotnicki, "Further insight into the physics and modeling of floating-body capacitorless DRAMs," *IEEE Trans. Electron Devices*, vol. 52, no. 11, pp. 2447–2454, Nov. 2005.

[5] R. Ranica, A. Villaret, P. Malinge, P. Mazoyer, D. Lenoble, P. Candelier, F. Jacquet, P. Masson, R. Bouchakour, R. Foumel, J. P. Schoellkopf, and T. Skotnicki, "A one transistor cell on bulk substrate (1T-Bulk) for low-cost and high density eDRAM," in *VLSI Symp. Tech. Dig.*, 2004, pp. 128–129.

[6] R. Ranica, A. Villaret, P. Malinge, G. Gasiot, P. Mazoyer, P. Roche, P. Candelier, F. Jacquet, P. Masson, R. Bouchakour, R. Fournel, J. P. Schoellkopf, and T. Skotnicki, "Scaled 1T-Bulk devices built with CMOS 90 nm technology for low-cost eDRAM applications," in *VLSI Symp. Tech. Dig.*, 2005, pp. 38–39.