A Unified-RAM (URAM) Cell for Multi-Functioning Capacitorless DRAM and NVM

Jin-Woo Han, Seong-Wan Ryu, Chungjin Kim, Sungho Kim, Maesoon Im, 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*

School of EECS, Division of EE, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea
*National Nanofab Center, Daejeon 305-806, Korea
E-mail: ykchoi@ee.kaist.ac.kr, Phone: +82-42-869-3477, Fax: +82-42-869-8565

Abstract

A novel partially-depleted (PD) SONOS FinFET is demonstrated for unified function of a high speed capacitorless 1T-DRAM and non-volatile memory (NVM). A floating body and O/N/O layer are combined in a single FinFET to provide multi-functional unified-RAM (URAM) operation. The fabricated URAM shows a V_T window of 3V with a retention time exceeding 10 years for NVM operation and a sensing margin of 9μA with a program/erase time of 10nsec for 1T-DRAM operation in a single memory cell transistor.

Introduction

For multi-functional memory applications, integration of various types of memories in one cell has become important in terms of device level fusion. As an example, while non-volatile memory (NVM) and DRAM can be integrated in one chip, their hybrid integration is inevitably complicated in conventional approaches. If a single cell could perform NVM and DRAM, slow programming/erasing (P/E) speed, a drawback of flash memory, would be significantly improved by the nature of fast DRAM operation, and the weakness of volatility in DRAM could be essentially countervailed by the nature of NVM operation. In this work, we propose a multi-functional SONOS FinFET structure to perform NVM and DRAM characteristics in a single transistor. A schematic of this concept is shown in **Fig. 1**. By implementation of O/N/O as an electron trapped zone for NVM and a partially depleted body as a hole storage zone for 1T-DRAM in SOI FinFETs [1], unified-RAM (URAM) operation is realized.

In conventional fully-depleted (FD) 1T-DRAM, the back-gate bias has to be used to form a deep potential well to accumulate holes [2]. However, the newly developed partially-depleted (PD) FinFET allows excess holes to be accumulated in the extremely narrow fin without a back-gate. The distinct functions of NVM and 1T-DRAM are demonstrated without disturbance to either operation mode in a single transistor.

Device Fabrication

The process sequence of PD SONOS FinFET is summarized in **Fig. 2**. A p-type (100) SOI wafer was used as a starting material. After Si_3N_4 deposition, a fin was patterned and HDP SiO_2 was deposited. By utilization of CMP and HF etch-back, the fin was partially recessed until 55% of the upper region Si was exposed. The remaining 45% of the lower region Si is covered by HDP SiO_2. 3/8/3nm of O/N/O was stacked sequentially, and n+ *in-situ* poly-Si was deposited and patterned. Finally, S/D implantation and an activation process were carried out. The fabricated PD FinFET is shown in **Fig. 3**. The device has 30nm fin width, 110nm fin height, and 180nm gate length. As a control group, a conventional fully-depleted SONOS FinFET was also fabricated, as shown in **Fig. 4**.

Results and Discussion

Table 1 presents the operation principle of the URAM. The essence of the proposed PD FinFET is that the upper part of the fin stays fully-depleted and the lower part of the fin, which is uncovered by the gate, remains partially-depleted so as to hold excess holes generated by impact ionization. Therefore, floating body effects can be generated even for sub-30nm fin width. Thus, suppression of short-channel effects as well as 1T-DRAM operation can be achieved.

NVM characteristics - **Fig. 5** shows P/E characteristics of I_D-V_G and electrical device parameters. Despite the thick O/N/O stack, superior short-channel characteristics were obtained due to elimination of leakage paths by narrowing fin width. **Fig. 6** shows the speed response of P/E and the threshold voltage (V_T) window. A 3V V_T window with 80μs pulse at 11V programming and -11V erasing is achieved. Retention behavior longer than 10 years and an endurance characteristic of more than 10^7 P/E cycles are observed along with a 2V V_T window, as shown in **Fig. 7**. In terms of NVM characteristics, the FD SONOS FinFET exhibits very similar performance to the PD SONOS FinFET.

1T-DRAM characteristics – Data in 1T-DRAM are stored at the floating-body. In programming, the holes generated by impact ionization are accumulated in the body, and the channel potential is lowered. In erasing, accumulated holes move into the negative biased drain and vanish. Also, the channel potential is raised. Therefore, the data states are identified by the presence of excess holes. In a fully-depleted 1T-DRAM, a negative biased second gate is needed in order to accumulate holes at the back interface [2]. In this work, we propose a novel structure that performs 1T-DRAM operation

1-4244-0439-X/07/$25.00 © 2007 IEEE

without a back-gate bias. A simulated potential biased at 1T-DRAM programming conditions is shown in **Fig. 8**. The entirely protruded fin is fully depleted in the FD FinFET (**Fig. 8(a)**) while the fin uncovered by the gate is partially depleted in the PD FinFET (**Fig. 8(b)**). Therefore, the PD FinFET can retain more holes and thereby provide increased retention time relative to the FD FinFET. The simulation results, presented in **Fig. 9**, clearly show a potential difference according to data states. Whereas little potential difference before and after programming in the FD FinFET is expected, the potential is significantly lowered after programming in the PD FinFET, i.e., the two states are distinguished by P/E. As evidence of hole accumulation, the kink effect is observed only in the PD FinFET, as shown in **Fig. 10**.

Fig. 11 shows the P/E speed response of the 1T-DRAM. The device programmed at $V_{G,PGM}$=1V & $V_{D,PGM}$=1.5V and erased at $V_{G,ERS}$=1V & $V_{D,ERS}$=−1V with 100ns of P/E time exhibits a sensing margin greater than 9μA. In addition, P/E speed less than 10nsec is achieved with a reasonable sensing margin. **Fig. 12** shows the impact of P/E voltages on the sensing margin. An abrupt increase of the program and erase efficiency is observed at $V_{D,PGM}$=1.5V & $V_{D,ERS}$=−1V, respectively. This implies that there exists a minimum voltage for impact ionization and forward junction current. Since the impact ionization event is inversely proportional to the gate length, 1.5V of $V_{D,PGM}$ can be reduced as the device is scaled down. **Fig. 13** show the P/E characteristics set by these P/E voltages with 100nsec. P/E states are clearly distinguished in the PD FinFET while the FD FinFET exhibits less difference between P/E states. One of the important advantages of 1T-DRAM is that excess holes can be retained during P/E operation. This non-destructive readout saves refresh time, which is required in 1T/1C DRAM. The fabricated PD FinFET showed non-destructive readout longer than 70msec, as presented in **Fig. 14**. Together with fast P/E speed, this non-destructive readout will allow faster operation compared to a conventional 1T/1C DRAM.

The schematics of the memory cell array and conditions of gate/drain disturbance are presented in **Fig. 15**. It should be noted that the cell layout of URAM is exactly the same as that for NOR-type flash memory. Therefore, URAM performs multi-functions without sacrificing chip area compared to a conventional NOR flash and can be fabricated by a standard flash memory process with minor modifications. In the array cell, gate and drain disturbance inevitably arise. The measured gate and drain disturbance are characterized in **Fig. 16**. The gate disturbance at cell (b) in **Fig. 15(b)** is negligible because the gate voltage (i.e. 1V) is sufficiently small to suppress hole tunneling. Regarding the drain disturbance, slight degradation of the erased state is observed due to soft impact ionization at cell (c) in **Fig. 15(b)**; however, this can be alleviated if the operation voltage is scaled.

NVM–1T-DRAM interference – Two crucial conditions should be satisfied for reliable URAM operation. First, V_T of NVM should not be disturbed during 1T-DRAM operation. The operational gate voltage in 1T-DRAM is in a range above 0V. Therefore, soft erase in NVM does not take place, but the programming gate voltage for 1T-DRAM should be sufficiently small so as to avoid soft programming in NVM. Second, 1T-DRAM should be operated after the threshold voltage of NVM is set to ~0.2V. If the initial V_T is too high, low V_G-V_T reduces the impact ionization efficiency for the 1T-DRAM program. If the initial V_T is too low, the erased state current will be increased. As a result, identifying data states of 1T-DRAM becomes difficult for both cases. From the measured results, the most moderate initial V_T is found to be ~0.2V.

In **Fig. 17**, available bias ranges are mapped from the measured results. Above V_G=6V, charge trapping occurs for NVM programming. Up to V_G=6V, undesired charge trapping is prohibited, and therefore the 1T-DRAM operated without disturbing the NVM states. Even though $V_{G,PGM}$ of 1T-DRAM can range from 0.2V (initial V_T) to 6V, $V_{G,PGM}$ should be carefully selected with consideration of the trade-off between program efficiency in 1T-DRAM and soft programming in NVM. In this work, the same gate voltage (i.e. 1.2V) is used for the P/E and read operation as an optimized condition. For the same reason, $V_{D,PGM}$ is optimized, and is slightly larger than the minimum voltage for impact ionization so as to avoid undesired charge trapping in the O/N/O stack. In this work, $V_{D,PGM}$ is set to 1.5V.

In order to verify that there is only negligible interference, I_D-V_G curves after DC stress biased at 1T-DRAM program voltages are shown in **Fig. 18**. V_T shift is not observed up to 10^4 sec. This implies that charge trapping is negligible during 1T-DRAM operation. This result approves that there is distinctive operation between both memory modes.

Conclusions

A multi-functional unified-RAM (URAM) that combines a non-volatile memory and capacitorless 1T-DRAM is demonstrated. URAM performs multi-functions with low cost and a standard process. The newly developed URAM ensures negligible disturbance between non-volatile memory and fast DRAM operation. The URAM is a promising structure for next generation fusion memory and high performance SoC applications.

References

[1] T. Tanaka *et al.*, *IEDM Tech. Dig.*, p. 573, 2006.
[2] I. Ben *et al.*, *IEDM Tech. Dig.*, p. 573, 2006.
[3] R. Ranica *et al.*, *IEDM Tech. Dig.*, p. 277, 2004.
[4] Silvaco International, ATLAS user's Manual, 2002.

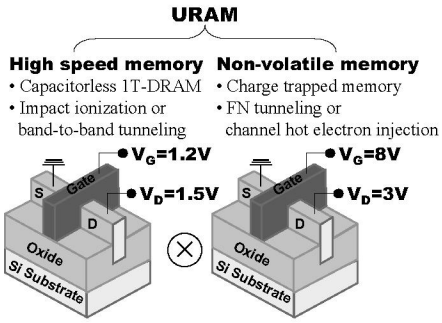


Fig. 1 Schematic of URAM: SONOS NVM and capacitorless 1T-DRAM.

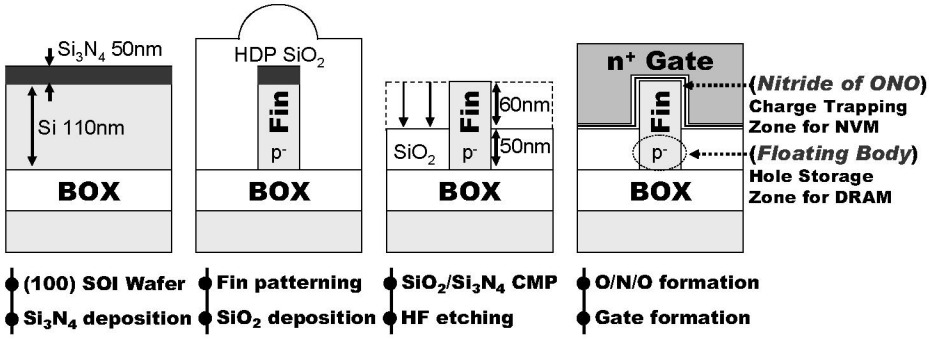


Fig. 2 Process flow of the multi-functional URAM. Due to a partial recess of SiO_2 by CMP and HF etch-back, the bottom part of the fin is covered by SiO_2. Thus it is partially depleted (PD).

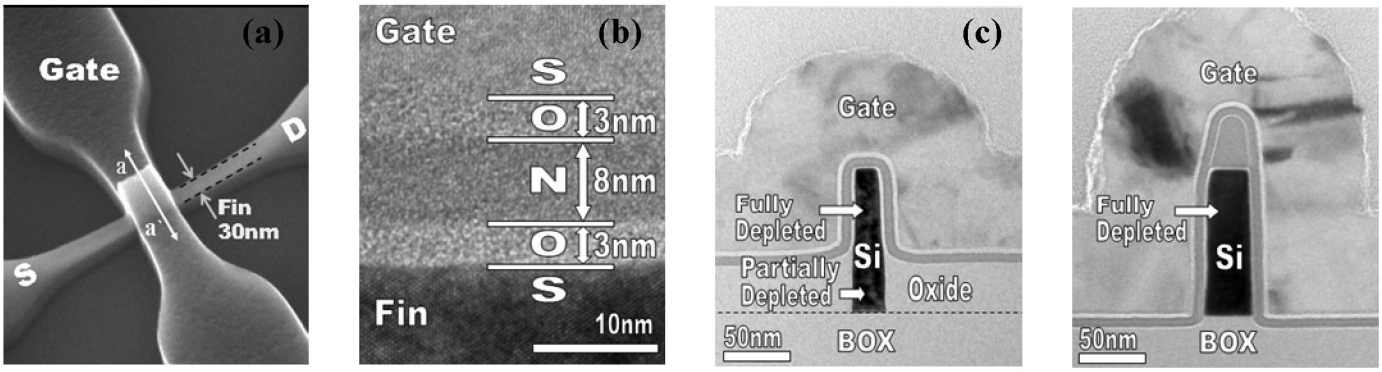
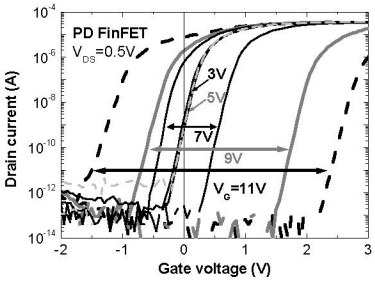


Fig. 3 A fabricated PD SONOS FinFET for URAM : (a) tilted SEM image of URAM, (b) TEM image of O/N/O layer, (c) TEM image across a-a' direction in Fig. 3(a). The upper 60nm of the fin (FD) is wrapped by a gate, and the lower 50nm of the fin (PD) remains out of the gate field (partially depleted).

Fig. 4 TEM image of conventional FD FinFET (control group). The fin is entirely surrounded by the gate, and thus is fully depleted (FD).

Table 1 Comparison of NVM and 1T-DRAM characteristics in terms of physical mechanisms and bias conditions for each operation mode.

Unified-RAM (URAM)		
NVM (Non-Volatile Memory)	**Operation type**	**Capacitorless 1T-DRAM**
Nitride trap in ONO	Storage node	Floating body
Charge trapping	Mechanism	Hole accumulation
FN tunneling / Channel hot-electron injection	Program	Impact ionization / Band-to-band tunneling
FN tunneling / Hot-hole injection	Erase	Forward biased pn junction
Non-volatile	Volatility	Volatile
Slow	Speed	Fast
Program / Erase	Schematic	Program / Erase



Parameter	Value
W_{Fin}	30nm
L_G	180nm
H_{Fin}(FD)/H_{Fin}(PD)	60/50nm
V_T	0.2V
SS	95mV/dec
DIBL	32mV/V
I_{on}	~300μA/μm
I_{off}	~10^{-11}A/ μm

Fig. 5 Measured programming and erasing characteristics of I_D-V_G. The NVM memory characteristics are clearly shown with a wide V_T window.

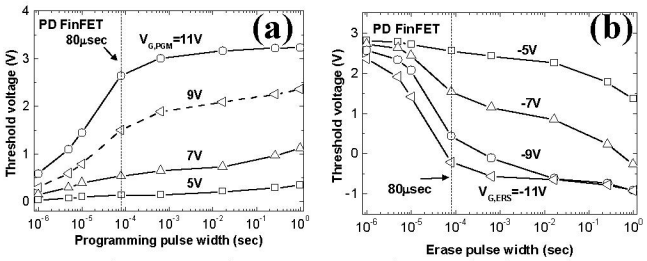


Fig. 6 Speed response of program (a) and erase (b) for various gate voltages for NVM application. P/E is carried out by F-N tunneling. 3V of V_T window is observed at 80μsec and ±11V of P(+)/E(-).

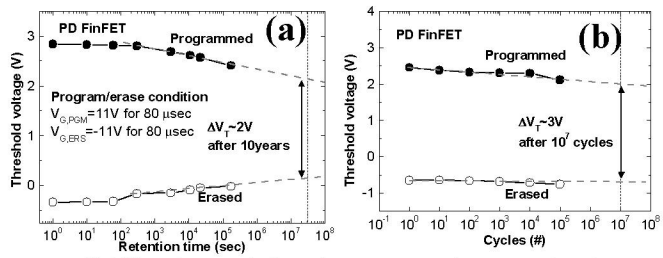


Fig. 7 Reliability characteristics of NVM: retention (a) and endurance (b). Retention time is longer than 10 years with 2V of V_T window and endurance cycle is more than 10^7 P/E with a 3V of V_T window.

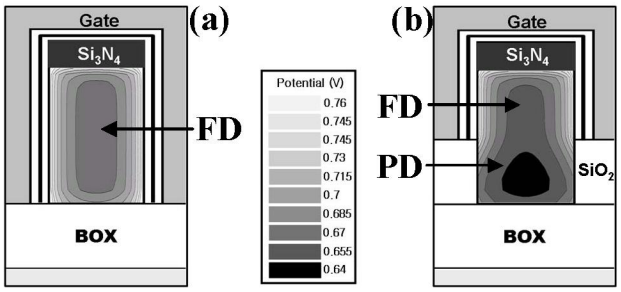


Fig. 8 Simulated body potential biased at V_G=1.2V and V_D=1.5V in SONOS FinFET: FD (a) and PD (b). The PD fin shows lower potential, which allows holes to be accumulated in this designed zone.

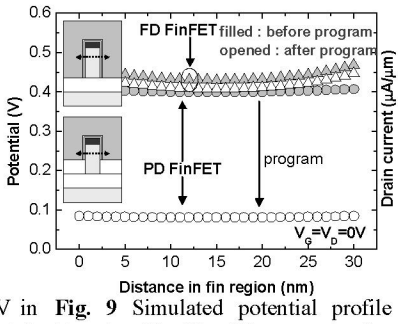


Fig. 9 Simulated potential profile biased at V_G=V_D=0V across the fin. The potential is lowered in PD after programming.

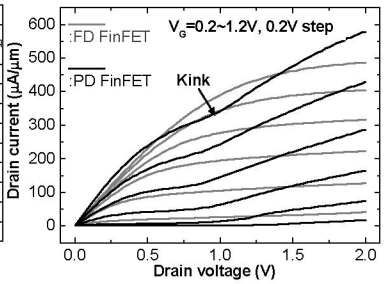


Fig. **10** Measured I_D-V_D. A kink effect caused by excess hole accumulation is observed only in the PD FinFET.

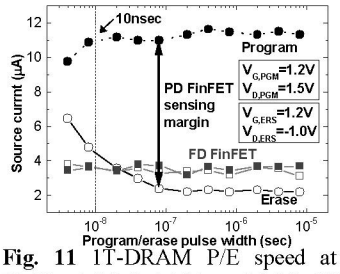


Fig. 11 1T-DRAM P/E speed at V_G/V_D=1.2/1.5, 1.2/-1, and 1.2/0.4V pulse for program, erase, and read respectively. 10nsec fast P/E is ensured with proper sensing margin.

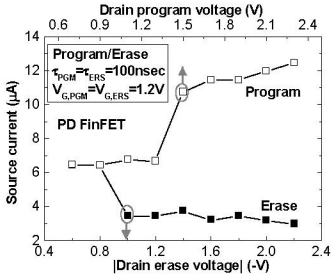


Fig. 12 The drain voltage dependence of P/E speed at 100nsec of pulse width. V_D=1.5V for program and V_D=-1V for erase are proper in terms of P/E efficiency.

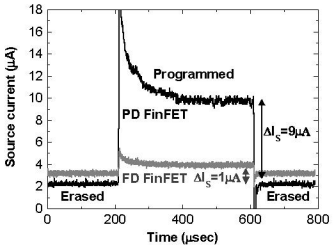


Fig. 13 Measured time dependence of I_S. Distinctive two states: "0" and "1" are observed only for the PD FinFET with a 9μA of sensing margin.

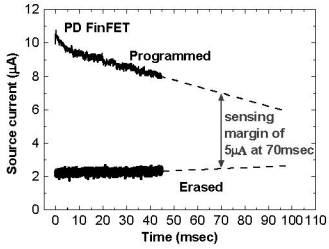
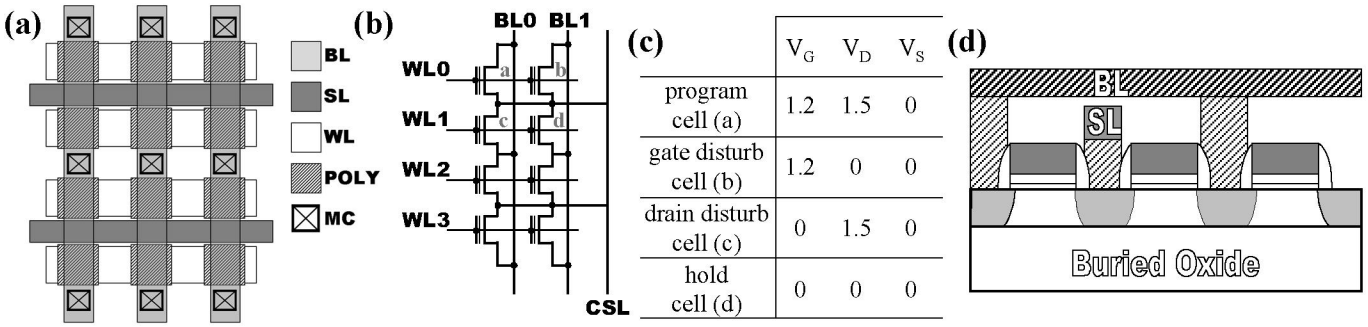


Fig. 14 Reading current evolution. The non-destructive readout guarantees 5μA of sensing margin until 70msec.



(c)	V_G	V_D	V_S
program cell (a)	1.2	1.5	0
gate disturb cell (b)	1.2	0	0
drain disturb cell (c)	0	1.5	0
hold cell (d)	0	0	0

Fig. 15 Cell array structure: layout (a), circuit schematic (b), a summary of disturbance biases (c), and cross-sectional schematic of array cell (d).

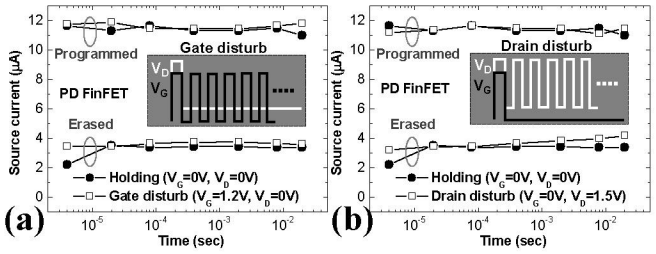


Fig. 16 Gate disturbance (a) and drain disturbance (b). A 100nsec pulse is applied every 1μsec. At drain disturbance, slight degradation of '0' state is observed due to soft impact ionization. But, no gate disturbance is observed at gate disturbance.

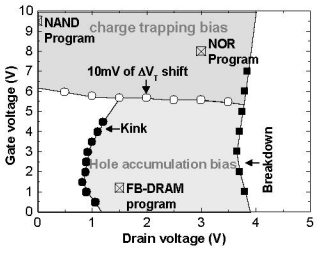


Fig. 17 Program bias map for NVM and 1T-DRAM. The distinctive operation can be guaranteed by each P/E bias condition.

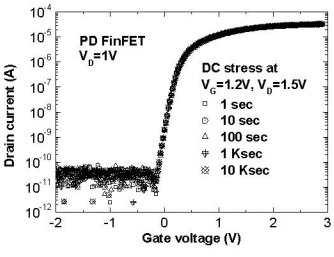


Fig. 18 I_D-V_G curves after DC stress bias at 1T-DRAM program voltage. There is no programming disturbance in ONO for NVM operation.