Design of Highly Stackable Charge Trap-Based 3D DRAM

Hyeongyu Kim*
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Jeonbuk, Rep. of Korea
rhtmaehcl0327@jbnu.ac.kr

Dabok Lee*
Electrical Engineering
Gyeongsang National University
Jinju, Gyeongnam, Rep. of Korea
ddabok11@gnu.ac.kr

Hyun-Sik Choi
Electronic Materials Engineering
Kwangwoon University
Seoul, Rep. of Korea
99choihs@gmail.com

Yoojin Seol
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Jeonbuk, Rep. of Korea
tjf010309@jbnu.ac.kr

Jonghyeon Ha
Electrical Engineering
Gyeongsang National University
Jinju, Gyeongnam, Rep. of Korea
jh990502@gnu.ac.kr

Kihyun Kim
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Jeonbuk, Rep. of Korea
kihyun.kim@jbnu.ac.kr

Jungsik Kim
Electrical Engineering
Gyeongsang National University
Jinju, Gyeongnam, Rep. of Korea
jungsik@gnu.ac.kr

Won-Ju Cho
Electronic Materials Engineering
Kwangwoon University
Seoul, Rep. of Korea
chowj@kw.ac.kr

Zvi Or-Bach
MonolithIC3D
United States
zvi@monolithic3d.com

Sung-Il Chang
MonolithIC3D
United States
sung.il@monolithic3d.com

*These authors contributed equally to the work.

TENCON 2025 - 2025 IEEE Region 10 Conference (TENCON) | 979-8-3315-3772-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/TENCON66050.2025.11375042

Abstract*—In this work, we propose a highly stackable Charge Trap-based 3D DRAM (CT 3D DRAM) structure that addresses key challenges in future memory scaling, including 3D integration, power consumption, and thermal management. Unlike conventional DRAM architectures that rely on complex capacitor structures, the proposed CT 3D DRAM utilizes a simple 1T memory cell with a poly-Si channel and Schottky barrier source/drain (S/D) contacts formed by metal silicide. Hot carrier injection (HCI) from the source side enables fast program operations through an ultrathin tunnel oxide. Key device parameters were optimized using 3D TCAD simulations, and planar CT DRAM devices were fabricated to validate the concept. The fabricated devices exhibited a program/erase window larger than 1 V under a 20 ns pulse, excellent retention characteristics exceeding 10 seconds at 85 °C, and endurance up to 10^{15} cycles with a remaining threshold voltage window of approximately 0.32 V. Moreover, the use of metal S/Ds significantly enhances heat dissipation and enables superior thermal management, critical for highly stacked 3D memories. The vertical integration of metal bit lines (BLs) and horizontal poly-Si channels results in lower RC delays, making the CT 3D DRAM scalable even beyond a thousand layers while maintaining effective cell area comparable to conventional $4F^2$ DRAMs. Through the optimized design of the word line (WL) and bit line (BL) structures, as well as control of key materials such as the tunnel oxide and charge trap nitride, we demonstrate that CT 3D DRAM can achieve both high speed and reliability. This architecture offers a promising solution for next-generation 3D DRAM applications requiring high density, low power, and efficient thermal management, particularly in emerging memory platforms like Compute Express Link™ (CXL™) memory.

Keywords*—*3D DRAM, charge trap, Schottky barrier

I. Introduction

The current technology of DRAM has reached the limit of miniaturization due to various reasons, such as RC delay and reliability issues [1-2]. While various 3D DRAMs-including new-material approaches based on IGZO and two-transistor (2T) DRAM are considered for the next-generation DRAM devices, stacking more than several hundred memory layers is quite difficult due to the complex 1T1C or 2T cell structures, leading to challenges in manufacturing cost and space efficiency. [1, 3]. Consequently, a cell structure as simple as 1T DRAM is essential to ensure the continued scalability of 3D DRAM [4].

In this paper, we present a novel charge trap-based 1T 3D DRAM (CT 3D DRAM) cell structure as a next-generation DRAM. Extremely high capacity and comparable operation speed to the conventional DRAM make it suitable for Compute Express Link^{TM} (CXL^{TM}) memory, which allows more latency and needs higher memory capacity as shown in Fig. 1. We designed device structure with TCAD simulation and fabricated 2D devices to validate the proposed device mechanism and feasibility of stacking it in a 3D architecture.

II. Device design of CT 3D DRAM

A. Structure design

Fig. 2(a) shows the schematic diagram of the CT 3D DRAM. It has multi-layered horizontal poly-Si channels between a bit line (BL) and a source line (SL). Staircases for the interconnection of WLs to row decoder circuits are formed in a way used for 3D NAND. BLs and SLs are located alternately on the top of the cells. Fig. 2(b) shows the cross-sectional CT 3D DRAM structure. BLs and SLs are formed with metal and metal silicide layers which have extremely high electric and thermal conductivity compared to semiconductor materials such as poly-Si and SiO_2. There are blocking/trap/tunnel layers (ONO layers) between a WL and a channel poly-Si. The ONO layers are also used to electrically separate BLs and SLs from WLs. The metal silicide is formed at the interface of the metal BLs/SLs and the channels. Fig. 2(c) shows its program operation. NOR flash memory usually utilizes hot carrier injection at its drain-side for programming. However, it is quite difficult to generate hot carriers in the poly-Si channel because there are

Authorized licensed use limited to: Gyeongsang National Univ. Downloaded on March 03,2026 at 08:22:40 UTC from IEEE Xplore. Restrictions apply.

plenty of grain boundaries and traps that dramatically shorten the mean free path of the carriers. We adopted the metal silicide source and drain (S/D) instead of heavily doped n+ S/D to form Schottky barrier (SB) between S/D and the channel. Electrons tunneling through SB from the source are accelerated rapidly toward the drain when a proper drain bias is applied and the accelerated hot electrons are injected to the trap layer at the source-side [5].

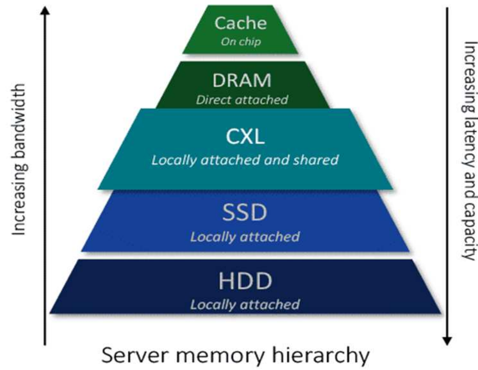


Fig. 1. Server memory hierarchy (from Rambus)

Fig. 3 summarizes an example of the main process flow. The holes for BLs, SLs and channels are etched together. Then, the holes for BLs and SLs are filled with a sacrificial material while the holes for channels are open and the mold nitride layers are partly indented by dry or wet etching. Poly-Si is deposited and etched away to form donut-type channels inside the indented nitride mold. Then, oxide is filled in the channel holes and metal is filled in the holes for BLs and SLs. After proper annealing process, metal silicide is formed between the metal and the poly-Si channel. Finally, the mold nitride is replaced with gate stack, which consists of ultrathin tunnel oxide, trap nitride, blocking layer, and gate material. The poly-Si channel indentation process is already verified in the floating-gate type 3D NAND devices [6-7].

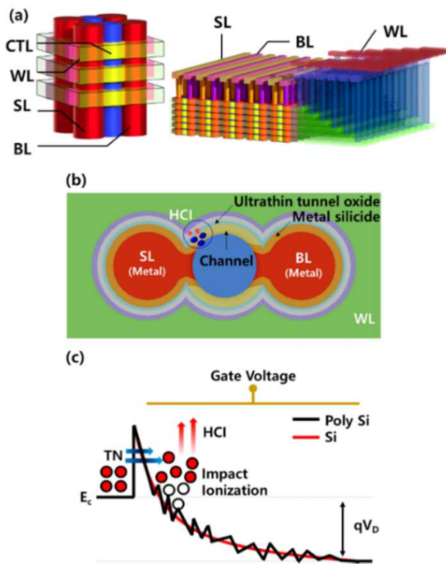


Fig. 2. (a) Schematic illustration of the CT 3D DRAM (b) Cross-sectional view of the CT 3D DRAM with Schottky barrier (SB) S/D and ultrathin tunnel oxide (c) HCI at the source-side was used for the program operation at the device having a poly-Si channel.

The proposed structure has strong advantages in power consumption and heat dissipation, which are the most important aspects for high-performance DRAM like HBM. Fig. 4(a) shows its power advantage. It consumes less power than the conventional 1T1C DRAM cell because it does not lose charges during the read operation and does not need to be restored. Also, it has better retention than a conventional DRAM and we can save power by increasing refresh period. The CT 3D DRAM without capacitors can achieve extremely high integration density and metal S/Ds are very effective to dissipate heat from the underlying logic die as shown in Fig. 4(b).

Vertical metal BLs and plate-type WLs have advantages on RC delay through optimization of design rules and material properties, as shown in Fig. 5(a). Considering that 88L 3D DRAM with a channel hole pitch of 200 nm has the same cell area with a conventional $4F^2$ DRAM with F = 15 nm, it is apparent that vertical BLs have lower RC delay even up to over a thousand layers. For WL RC delay, we have to consider the number of cells per WL and WL pad area for WL contacts together. If the number of cells is decreased, we can achieve lower WL RC delay than that of conventional 2D DRAM but we need more WL pad area, which may increase overall chip area. Therefore, it is essential to optimize both the number of cells per WL and the design of the WL pad area to ensure superior performance and a more effective cell area than that achieved by planar DRAM, as shown in Fig. 5(b).

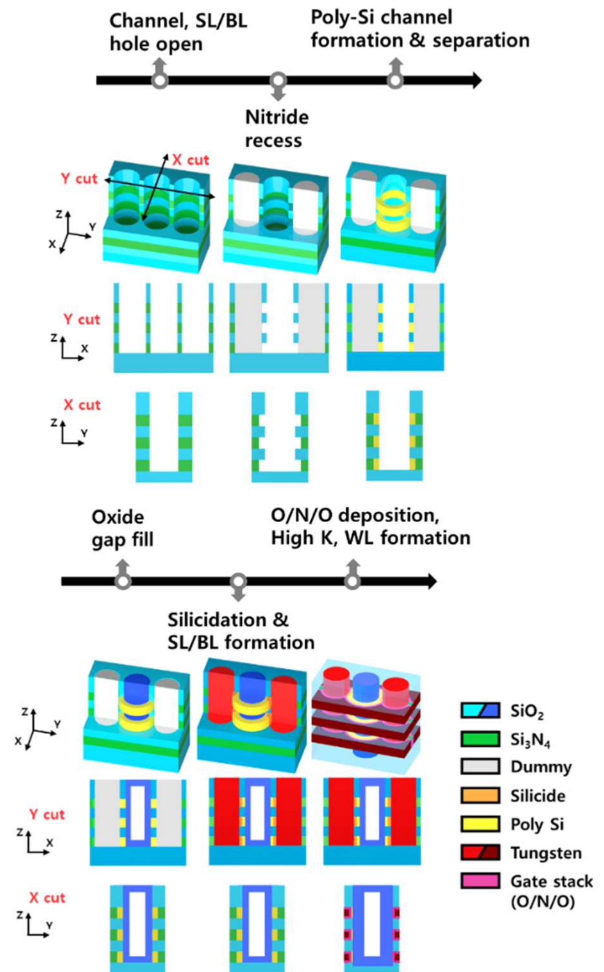


Fig. 3. Process integration of the CT 3D DRAM with SB S/D

Authorized licensed use limited to: Gyeongsang National Univ. Downloaded on March 03,2026 at 08:22:40 UTC from IEEE Xplore. Restrictions apply.

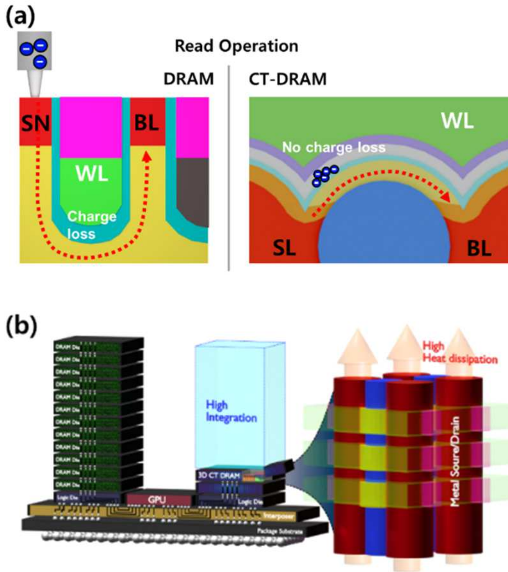


Fig. 4. Advantages of the CT 3D DRAM compared to the conventional DRAM in view of (a) power and (b) 3D integration/heat dissipation.

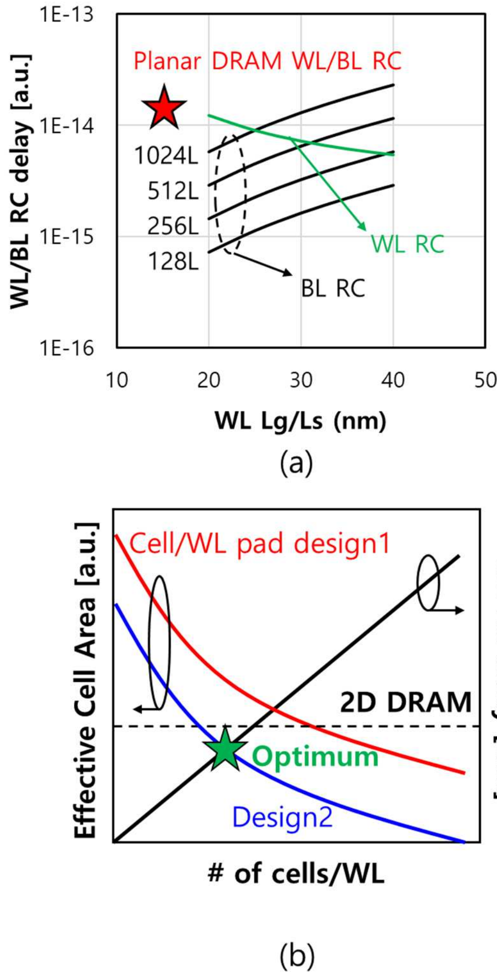


Fig. 5. (a)WL/BL RC delay of CT 3D DRAM according to WL Lg/Ls. Ch. hole pitch of 200 nm was assumed. 88L 3D DRAM has the same cell area with a conventional 4F2 DRAM with F = 15 nm. (b) WL RC delay and effective unit cell area according to the number of cells per WL.

B. Optimization of device parameters

We optimized the parameters of the CT 3D DRAM with 3D TCAD simulation. Fig. 6(a) and (b) show the simulated I_d-V_g curves and the energy band diagrams along with various SB heights, respectively. Considering the program characteristics and the on-current (I_{on}), SB of around 0.25 eV is good for the device. For tunnel oxide thickness (T_{oxb}) between 0.9 nm and 1.5 nm, the variation in threshold voltage (V_{th}) and electrostatic potential is small as shown in Fig. 6(c), which means that the difference in vertical and lateral E-fields required for hot carrier injection (HCI) is also negligible as shown in Fig. 6(d).

Fig. 7(a) and (b) shows the retention time as a function of the trap density and the trap energy level in charge trap layer, respectively. A higher trap density and energy level improve retention time by increasing initial V_{th} and reducing detrapping probability. Although programming speed is hardly affected by the thickness, the thicker T_{oxb} results in improved retention time. A T_{oxb} thickness of approximately 1 nm is good enough considering the proper retention characteristics for DRAM. The target parameters based on the simulation results were summarized in Table 1.

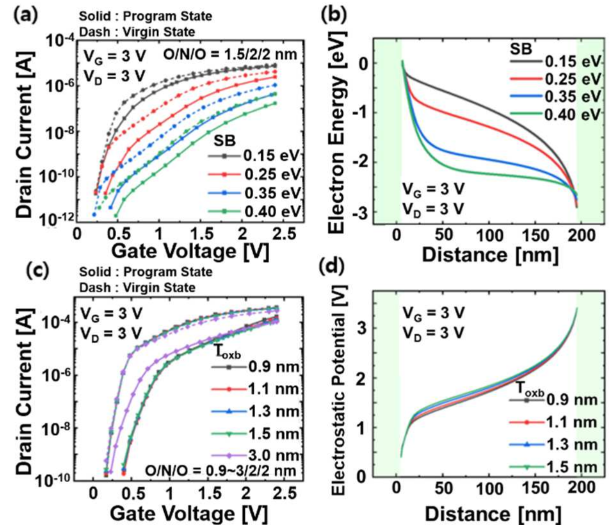


Fig. 6. I_d–V_g characteristics at different (a) SB (0.15 eV ~ 0.40 eV). (b) The SB characteristics that vary with programmed energy band. (c) I_d–V_g characteristics at the different T_{oxb}. (d) The electrostatic potential along the channel with different T_{oxb}.

Table 1. Optimized Device parameters

Description (parameter)	Values
Blocking oxide thickness (T_{oxt})	2 nm
CTL(Si_3N_4) thickness (T_{CTL})	2 nm
Tunnel oxide thickness (T_{oxb})	1 nm
Channel length (L_g)	150 ~ 300 nm
Channel thickness (T_{ch})	7 nm
Schottky barrier height (SB)	0.25 eV
Nitride Trap density (N_{trap})	7×10^{19} cm^{-3}
Substrate doping concentration	Undoped
Gate work function	4.65 eV
Trap energy level (EMC)	2.25 eV

Authorized licensed use limited to: Gyeongsang National Univ. Downloaded on March 03,2026 at 08:22:40 UTC from IEEE Xplore. Restrictions apply.

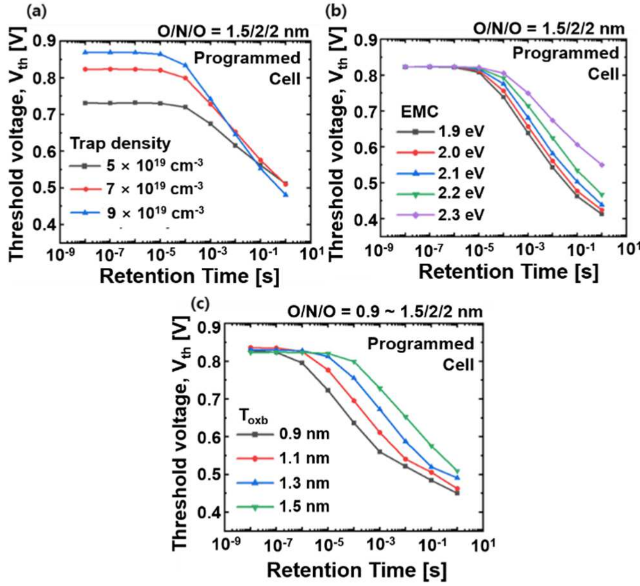


Fig. 7. The retention characteristics at (a) different trap density, (b) trap energy level (EMC), and (c) tunnel oxide thickness (T_{oxb}).

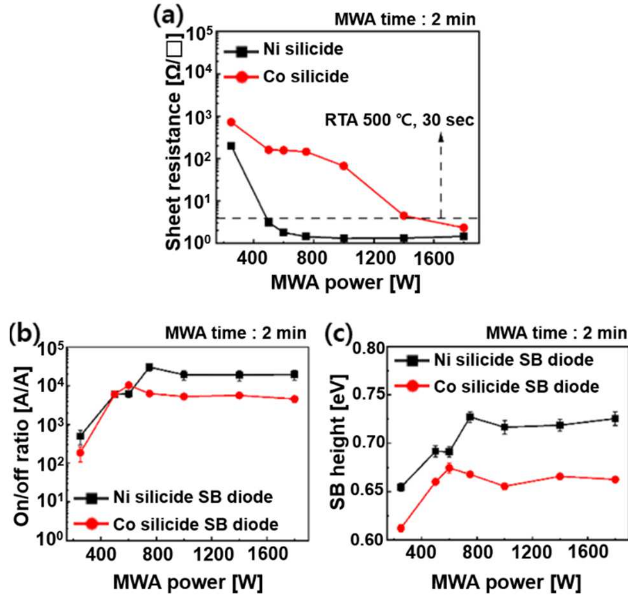


Fig. 8. (a) Sheet resistance, (b) on/off current ratio, and (b) SB heights of Ni and Co-silicide SB according to MWA power.

III. Device Fabrication

A. Schottky barrier silicide formation

The feasibility of the CT 3D DRAM was verified by planar CT DRAM. We formed silicides by using Microwave annealing (MWA). MWA utilizes electromagnetic radiation to heat the materials, offering the advantage of achieving uniform volume heating [8]. The sheet resistance (R_s) and junction characteristics of the silicide were examined as a function of MWA power. Fig. 8(a) depicts the sheet resistance of silicides according to MWA conditions. In this study, we applied Ni having higher the I_{on}/I_{off} ratio and lower R_s. Fig. 8(b) and (c) show the I_{on}/I_{off} ratio and SB heights (ϕ_b) of the SB diodes, respectively. The I_{on}/I_{off} were extracted from $V_F = 5$ V and averaging reverse current, respectively. ϕ_b was extracted using the following equation (1).

$$\boldsymbol{\phi_b} = \frac{kT}{q}\ln\left(\frac{A^{**}T^2}{J_0}\right) \qquad (1)$$

Where q, k, T, A^{**}, and J_0 are the unit electron charge, Boltzmann's constant, absolute temperature, equivalent Richardson's constant, and reverse saturation current density, respectively.

B. Device integration

Fig. 9 illustrates the fabrication process of the planar CT DRAM. A buried oxide and active poly-Si layers were formed on the substrate. After the active region was formed, ONO layers and n^+ poly-Si gate were formed. Then, ON gate spacer was formed to isolate gate and S/D regions during following self-aligned silicidation process. Finally, Ni was deposited using an electron beam evaporator, and MWA was performed to form a silicide with low contact resistance. The unreacted Ni was removed.

IV. Results and Discussions

Fig. 10(a) shows the cross-sectional TEM image of the fabricated planar CT DRAM. The S/D regions were fully silicided and its thickness is 20 nm. Fig. 10(b) illustrates the thickness of the CT layer (ONO = 1/2.3/4 nm). Fig. 10(c) illustrates initial I_d–V_g curves of the planar devices utilizing MWA at different time and power levels.

- Buried oxide formation (2000 Å)
- Undoped poly deposition (300 Å)
- Poly-Si channel formation
- O/N/O deposition (10/20/20 Å)
- N+ gate formation (1000 Å)
- Oxide/Nitride deposition (150/150 Å)
- Spacer formation
- Nickel deposition (300 Å)
- Microwave annealing (600/700 W, 1/2 min.)

Fig. 9. Process integration of the planar devices with SB S/D.

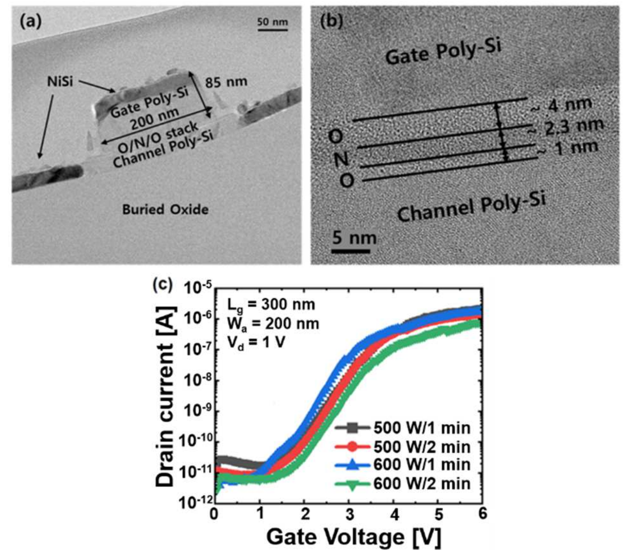


Fig. 10. TEM photographs of the fabricated 2D device with SB S/D (a) Cross-sectional TEM (b) Magnified TEM of ONO layers. (c) Measured initial I_d-V_g characteristics with $L_g/W_a = 300$ nm/200 nm.

Authorized licensed use limited to: Gyeongsang National Univ. Downloaded on March 03,2026 at 08:22:40 UTC from IEEE Xplore. Restrictions apply.

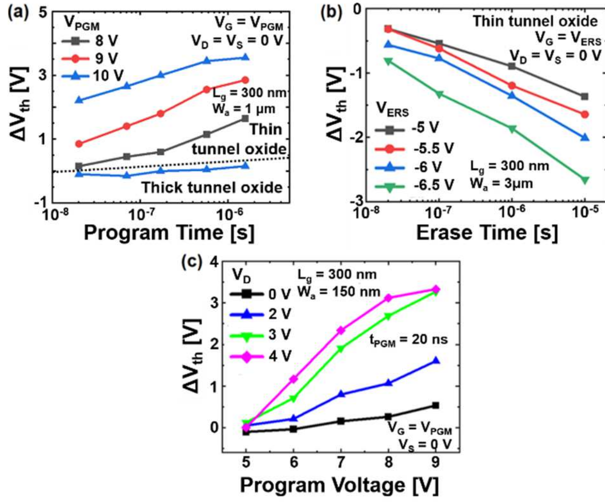
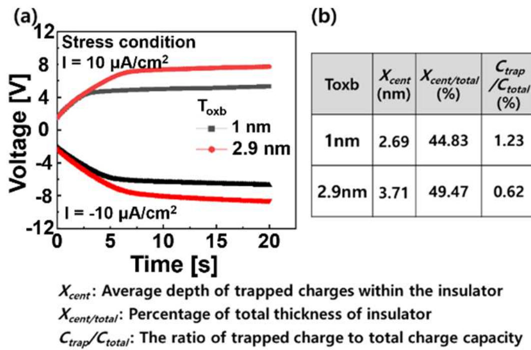


Fig. 11. Measured (a) program and (b) erase characteristics of the devices with thin and thick tunnel oxides. (c) Measured program operation characteristics using hot carrier injection with varying drain voltages.

Fig. 11(a) compares the program characteristics of the devices with a T_{oxb} of 1 and 2.9 nm. Fig. 11(b) shows the erase characteristics of the device as a function of erase time and voltage. We used 600 W/1 min of MWA condition considering the lateral growth of NiSi as well as I_{on} of the devices. The device with T_{oxb} of 1 nm showed program/erase V_{th} window larger than 1 V even at the operation time of 20 ns@V_{pgm} = 9 V and V_{ers} = - 6 V. Fig. 11(c) shows more effective program characteristics utilizing HCI at a short program time of 20 ns and lower gate biases. As the drain voltage increases, a larger E-field is generated at the source end, resulting in significantly faster program characteristics.

Fig. 12(a) and (b) show reliability of CT layer and extracted trapping characteristics. We can extract average depth of the trapped charges within the insulator (X_{cent}), percentage of total thickness of insulator ($X_{cent/total}$), and the ratio of trapped charge to total charge capacity from the time-dependent dielectric breakdown (TDDB) characteristics of the device [9]. The device with thinner oxide showed a shorter X_{cent} and better trapping efficiency (C_{trap}/C_{total}).



Toxb	X_{cent} (nm)	$X_{cent/total}$ (%)	C_{trap}/C_{total} (%)
1nm	2.69	44.83	1.23
2.9nm	3.71	49.47	0.62

Fig. 12. (a) TDDB characteristics of CT layers with stress current density of +/-10 uA/cm² for T_{oxb} = 1 and 2.9 nm. (b) Trapping characteristics extracted from the TDDB characteristics.

Fig. 13(a) shows the retention characteristics at 85°C of the device with a thin tunnel oxide. Even with the tunnel oxide as thin as 1 nm, more than 40% of the initial V_{th} window was retained after 10 seconds, which is sufficient for DRAM applications. Fig. 13(b) shows the measured endurance of the device up to 10^9 cycles. Although there is a fluctuation in V_{th} that caused by native-oxide -induced nonuniformity in the tunnel oxide and fast current reading, V_{th} window of 0.32 V is expected even after 10^{15} P/E cycles, ensuring sufficient reliability for the lifetime of DRAM.

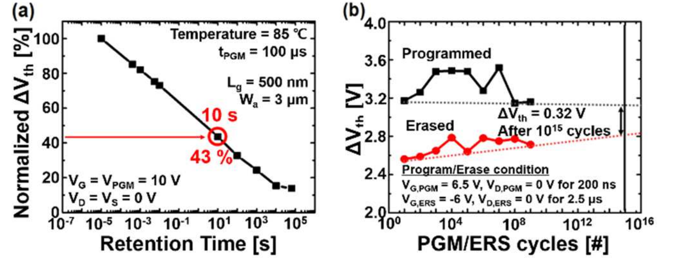


Fig. 13. (a) Measured retention characteristics of the planar CT DRAM with thin tunnel oxide at 85°C. (b) Measured endurance characteristics of the device with thin tunnel oxide.

V. Conclusion

This study investigated the CT device and its potential as a next-generation 3D DRAM solution. We showed that CT DRAM was promising when it was combined with ultrathin tunnel oxide and HCI from metal silicide S/D. Moreover, the CT 3D DRAM is one of the most promising candidates for 3D DRAM due to its simple process integration, power advantage, and easy heat dissipation through the metal S/D.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-02653773). The authors would like thank H.-C. Hwang/B.-R. Yoon at NNFC for helping the device fabrication.

References

[1] J. W. Han et al., "Ongoing Evolution of DRAM Scaling via Third Dimension-Vertically Stacked DRAM," *IEEE Symposium on VLSI Technology*, Kyoto, Japan, 2023.

[2] M. Suh et. al., "Design Guideline of Saddle-Fin-Based DRAM for Mitigating Rowhammer Effect," *IEEE Transactions on Electron Devices*, pp. 2417-2422, vol. 71, 2024.

[3] A. Belmonte et al., "Tailoring IGZO-TFT architecture for capacitorless DRAM, demonstrating > 10^3s retention, < 10^{11} cycles endurance and Lg scalability down to 14nm," *IEEE IEDM*, San Francisco, CA, USA, 2021.

[4] J. Kim et al., "First Demonstration of Innovative 3D AND-Type Fully-Parallel Convolution Block with Ultra-High Area- and Energy-Efficiency," *IEEE IEDM*, San Francisco, CA, USA, 2023.

[5] S. -J. Choi et al., "Performance breakthrough in NOR flash memory with dopant-segregated Schottky-barrier (DSSB) SONOS devices," *IEEE Symposium on VLSI Technology*, pp. 222-223, Kyoto, Japan, 2009.

[6] A. Khakifirooz et al., "A 1.67Tb, 5b/Cell Flash Memory Fabricated in 192-Layer Floating Gate 3D-NAND Technology and Featuring a 23.3Gb/mm2 Bit Density ," *IEEE ISSCC*, 2023

[7] K. Parat et al., "A floating gate based 3D NAND technology with CMOS under array," *IEEE IEDM*, 2015

[8] Min, J.G., et. al., "Implementation of ambipolar polysilicon thin-film transistors with nickel silicide Schottky junctions by low-thermal-budget microwave annealing," *Nanomaterials*, pp. 628, vol. 12, no. 4, 2022.

[9] M.-H. Jung et al., "Dependence of charge trapping and tunneling on the silicon-nitride (Si_3N_4) thickness for tunnel barrier engineered nonvolatile memory applications," *Appl. Phys. Lett.*, vol. 94, 053508, 2009

Authorized licensed use limited to: Gyeongsang National Univ. Downloaded on March 03,2026 at 08:22:40 UTC from IEEE Xplore. Restrictions apply.