Energy Band Engineered Unified-RAM (URAM) for Multi-Functioning 1T-DRAM and NVM

Jin-Woo Han[1], Seong-Wan Ryu[1], Sungho Kim[1], Chung-Jin Kim[1], Jae-Hyuk Ahn[1], Sung-Jin Choi[1], Kyu Jin Choi[2], Byung Jin Cho[1]
Jin Soo Kim[3], Kwang Hee Kim[3], Gi Sung Lee[3], Jae Sub Oh[3], Myong Ho Song[3], Yun Chang Park[3], Jeoung Woo Kim[3], and Yang-Kyu Choi[1]

[1]KAIST, Daejeon, Korea, [2]Jusung Engineering, Gwangju, Korea, [3]National Nanofab Center, Daejeon, Korea
E-mail: ykchoi@ee.kaist.ac.kr, Phone: +82-42-350-3477, Fax: +82-42-350-8565

Abstract

A novel fusion memory is proposed as a new paradigm of silicon based memory technology. An O/N/O gate dielectric and a floating body are combined with a FinFET, and the non-volatile memory (NVM) and high speed capacitorless 1T-DRAM are performed in a single transistor. A nitride trap layer is used as an electron storage node for NVM, and hetero-epitaxially grown $Si/Si_{1-x}Ge_x$ energy band engineered bulk substrates allow excess hole storage for 1T-DRAM. Highly reliable 1T-DRAM and NVM are demonstrated.

Introduction

In semiconductor memory technology, 'scaling' has unequivocally allowed explosive IT market growth. However, semiconductor technology is approaching its physical and technological end, indicating that revenue obtained from downscaling will decrease as scaling slows. Therefore, an entirely new concept should be required in order for silicon memory technology to remain competitive. According to this requirement, we have exploited a new paradigm that can create value and continue the use of the infrastructure of the semiconductor industry.

In an important breakthrough, URAM has been proposed as a new type of fusion memory [1-2]. URAM combines 1T-DRAM and NVM function in a single memory cell by implementation of O/N/O as an electron trapped zone and a floating body as holes storage zone; operational domains are identified by the bias conditions, as shown in **Fig. 1**. In order to utilize the floating body effect, the capacitorless 1T-DRAMs have been implemented on SOI [1], buried n-well [2], and buried $Si_{1-y}C_y$ [2] substrates. While hole charging in the floating body is allowed by a *quantum barrier* on the three aforementioned substrates, a *quantum well* structure can also store holes as illustrated in **Fig. 2**. The quantum barrier uses its body as both a current conduction path and a hole storage area simultaneously whereas the proposed structure (quantum well) separates the current conduction path and hole storage area.

In this work, URAM fabricated on buried $Si_{1-x}Ge_x$ substrates is presented as a type of quantum well device. The electrical characteristics including NVM and 1T-DRAM are confirmed.

Device Fabrication

The process sequence is summarized in **Fig. 3**. *Si* and $Si_{1-x}Ge_x$ layers are epitaxially grown on a (100) bulk Si wafer, and the subsequent process are similar to those of a conventional bulk FinFET SONOS. The device dimensions and splits are listed in **Table 1**. π- and Ω-gate are fabricated on various substrates with different *Ge* content (*x*) in $Si_{1-x}Ge_x$. As a nominal device structure, dimensions of L_g=180 nm/W_{fin}=50nm are used. TEM images of the band engineered substrates are shown in **Fig. 4**. The thicknesses of the buried $Si_{1-x}Ge_x$ and top *Si* are 55 nm and 85 nm. $Si/Si_{1-x}Ge_x$ lattices are matched at x=0.3. Images of the fabricated device are shown in **Fig. 5**.

Results and Discussion

I_D-V_G is presented in **Fig. 6**. Because of the two conduction paths at the *Si* and $Si_{1-x}Ge_x$ layers, Ω-gate shows double subthreshold slopes. V_T-roll off in **Fig. 7** shows that the Ω-gate exhibits inferior short-channel properties to the π-gate due to the thicker oxide on $Si_{1-x}Ge_x$. The valence band offset (ΔE_v) between the top *Si* and buried $Si_{1-x}Ge_x$ is examined via a capacitance-voltage (C-V) analysis, as shown in **Fig. 8**. While a single slope is shown at the control Si MOS capacitor, double slopes are found at the $Si_{1-x}Ge_x$ one. Because the maximum valence band energy of $Si_{1-x}Ge_x$ is higher than that of *Si*, transition of accumulated holes from the buried $Si_{1-x}Ge_x$ layer to the top *Si* surface according to V_G decrement results in double slopes in the C-V plot. The clearer hump in C-V implies a larger ΔE_v. Simulated ΔE_v was first fitted to the measured data. The calibrated ΔE_v was then compared to theoretical value and plotted in **Fig. 9**. ΔE_v linearly increases as Ge content *x* increases. In addition to the double slopes in C-V, as a signature of hole accumulation in the floating body, a kink effect can also assure ΔE_v. I_D-V_D plotted in **Fig. 10** clearly show the kink effect. Consequently, the double slope in C-V and the kink in I_D-V_D verify that the bulk substrate can be utilized as a floating body memory cell if energy band engineering is applied.

A simulated distribution of holes at the hold condition is shown in **Fig. 11**. Whereas the holes accumulate inside the body of the SOI substrate, they prefer to stay inside $Si_{1-x}Ge_x$ in the energy band engineered substrate. Hence the $Si_{1-x}Ge_x$ quantum well can serve as a data storage node to distinguish '0' (without

holes) and '1' (with holes). The stored hole concentration increases as ΔE_V increases, and it saturates at ΔE_V=0.24eV, which corresponds to x=0.4. This implies that the use of a $Si_{1-x}Ge_x$ quantum well with x>0.4 would be ineffective in terms of hole storage efficiency.

During the impact ionization process for 1T-DRAM programming, undesirable electron trapping into the O/N/O in NVM, *i.e.* soft programming in **Fig. 12**, should be restricted for distinctive operation. Therefore, the operational bias domain should be set to avoid disturbance between NVM and 1T-DRAM operation as shown in **Fig. 12.** If the program voltage for 1T-DRAM is too high, soft programming into O/N/O can disturb the NVM state. Also, when program voltage is too low, it reduces program efficiency. With consideration of the trade-off relation between program efficiency and soft programming, the operational biases for 1T-DRAM are optimized, and these conditions are summarized in **Table 2**. It should be noted that V_T of NVM should be adjusted to ~0.3V before the URAM is set for 1T-DRAM mode. If the initial V_T is adjusted such that it is large, larger V_G will be necessary to enable sufficient gate overdrive voltage (V_G-V_T). This increased V_G can trigger to program into NVM during 1T-DRAM operation. In contrast, if the initial V_T is too low, programming can take place even at an unselected memory cell, because the carrier energy is sufficiently high to cause impact ionization.

The operational conditions are summarized in **Table 2**, and P/E 1T-DRAM characteristics for π- and Ω-gate are shown in **Fig. 13**. The π-gate exhibits a higher sensing current window than the Ω-gate because the partially-depleted (PD) volume is larger in the π-gate. Consequently, the π-gate is more desirable for 1T-DRAM application. **Fig. 14** shows the P/E of 1T-DRAM at a π-gate for two different *Ge* content values (x=0.3 and x=0.5). In the case of x=0.5, a wider sensing current window is observed, but faster charge loss is observed at the '1' state as well as the '0' state. This fast charge loss causes significant degradation of the data retention time. **Fig. 15** shows the sensing current window and retention time for various x. Despite the larger sensing window at x>0.4, the retention is found to be inferior to that at x=0.3. This is due to the fact that the higher defect density caused by lattice mismatch degrades the '1' retention time through recombination while the stored charge density is saturated at x>0.4. In addition, a deeper quantum well also degrades the '0' retention time through hole-to-hole repulsion and its diffusion, as described in **Fig. 16**. As a result of the trade-off between ΔE_V and defect density, the retention time is maximized at x=0.3. While the retention time (on the order of micro-seconds) appears to be insufficient to warrant application for 1T-DRAM at present, refinement of the epitaxial process and dimension optimization will enhance the retention time.

An advantage of bulk over a SOI substrate is that it allows modulation of quantum barrier/well energy band by changing V_{Sub}. A positive V_{Sub} increases the hole potential barrier and supports prolonged retention of holes in the quantum well. As a result, both the sensing current window and the retention time are improved at a V_{Sub} of 0.4V, as shown in **Fig. 17**. However, when V_{Sub} is larger than 0.6V, the forward-biased S/D to body diode obstructs 1T-DRAM operation.

In the case of a cell array, P/E disturbance among adjacent cells should be verified. The schematics of the array cell and the measured gate- and drain-disturbance characteristics are shown in **Fig. 18.** The absence of interference by charge trapping (soft program in **Fig. 12**) during 1T-DRAM operation should be investigated to ensure reliable operation. In order to verify this, V_T shift is measured after cyclic 1T-DRAM operation. It is found that V_T variation is negligible after 1T-DRAM operation up to 10^6 P/E cycles, as shown in **Fig. 19**.

Fig. 20 shows the NVM speed response of P/E. P/E voltage of ±12V with a 100μs pulse exhibits a V_T window of 3V. A V_T window of 2.5V ensures a 10 year lifetime, as shown in **Fig. 21**. In terms of NVM characteristics, URAM exhibits similar performance at various *Ge* content. As an example of URAM, **Fig. 22** shows versatile memory chip applications according to customization. If the memory core region consists of four banks, each bank can be assigned according to the designer's demands. Thus, a portion of the embedded memory chip can be allocated for satisfying the requirements of high speed and non-volatility.

Conclusions

An energy band engineered unified-RAM (URAM) is demonstrated in conjunction with hetero-epitaxially grown $Si/Si_{1-x}Ge_x$ substrates. The combination of O/N/O and a floating body provides non-volatile memory and high-speed capacitorless 1T-DRAM function in a single transistor. The π-gate structure is found to be superior to the Ω-gate in terms of 1T-DRAM characteristics. According to the trade-off relation between quantum well depth and defect density, *Ge* content x in a buried $Si_{1-x}Ge_x$ layer is optimized at x=0.3. As the consumer demands higher performance as well as multipurpose products, URAM will allow customized device architecture, increased performance, and reduced bit cost while enabling products to be smaller with more functionality. The paradigm shift from 'downscaling' to a 'multi-function-cell' beyond the multi-bit-cell can continue the evolution of silicon memory technology.

Acknowledgment

This work was supported by the National Research Program for the 0.1-Terabit Nonvolatile Memory Development, sponsored by the Ministry of Knowledge Economy.

References

[1] J.-W. Han *et al.*, *IEDM Tech. Dig.*, p. 929, 2007.
[2] J.-W. Han *et al.*, *Symp. VLSI Tech. Dig.*, p. 200, 2008.
[3] C. G. Van de Wall *et al.*, PR-B, p. 8154, 1987.
[4] M. E. Kurdi *et al.*, PR-B, p. 195327, 2006.

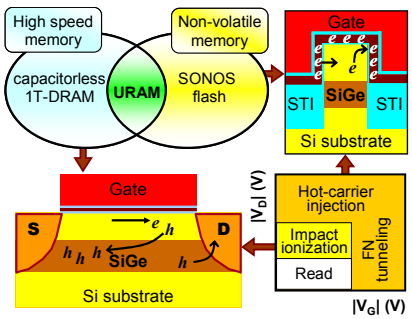


Fig. 1 Schematics of Unified-RAM (URAM). Traps in O/N/O hold electrons for NVM, and potential well at buried $Si_{1-x}Ge_x$ store holes for 1T-DRAM.

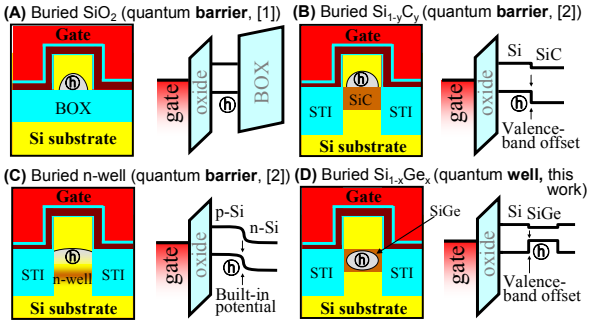


Fig. 2 Various substrate types for 1T-DRAM operation. (A)~(C) quantum barrier is used to retain excess holes. In this work (D), a quantum well is formed to store excess holes in the buried $Si_{1-x}Ge_x$.

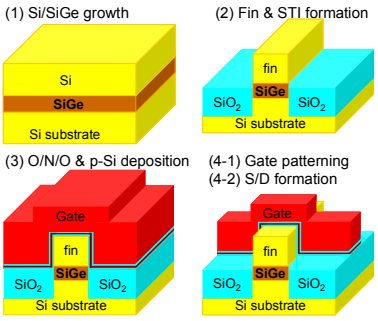


Fig. 3 Process flow of URAM. After the buried $Si_{1-x}Ge_x$ substrate is prepared, subsequent processes are similar to bulk FinFET SONOS.

Device dimensions	
L_g	180 nm
W_{fin}	50 nm
$T_{O/N/O}$	3/6/4 nm
Device splits	
H_{fin}	80 (π-), 100 nm (Ω-gate)
x ($Si_{1-x}Ge_x$)	0.1, 0.2, 0.3, 0.4, 0.5 at. %

Table 1 Device dimensions and splits. Different gate shape (π- or Ω-gate) and Ge content (x) are compared. The reference device is a π-gate with x=0.3.

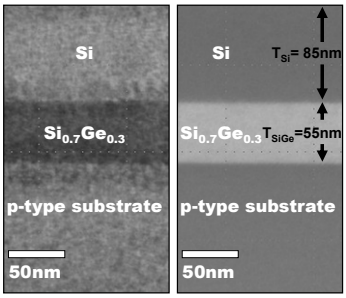


Fig. 4 TEM / STEM - HAADF images of buried $Si_{0.7}Ge_{0.3}$ band engineered substrate formed by hetero-epitaxial growth. T_{Si}=85nm, T_{SiGe}=55nm.

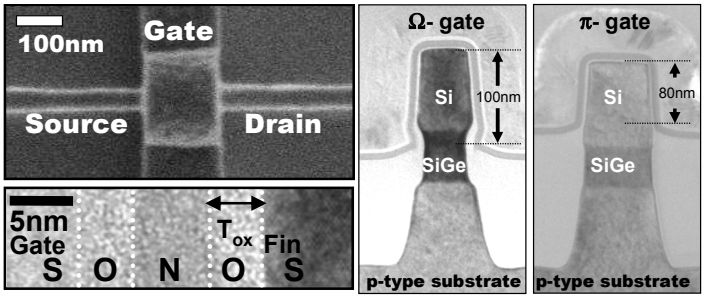


Fig. 5 Top view of the device (L_g=180nm and W_{fin}=50nm). TEM image of the O/N/O gate dielectric shows 3nm tunneling oxide, 6nm nitride, and 4nm control oxide, respectively. TEM cross-sections of the URAM show phi- (π) / omega-gate (Ω) structure depending on gate coverage over H_{fin}.

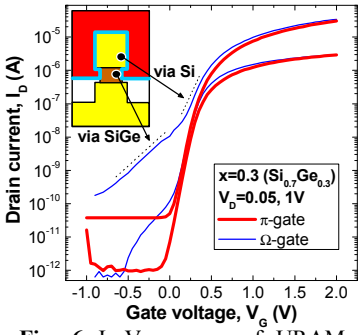


Fig. 6 I_D-V_G curves of URAM fabricated on buried $Si_{0.7}Ge_{0.3}$. Due to thicker T_{ox} at $Si_{0.7}Ge_{0.3}$, Ω-gate shows double slopes in I_D-V_G.

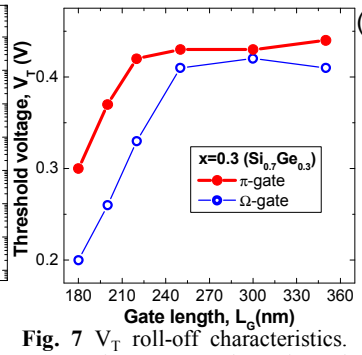


Fig. 7 V_T roll-off characteristics. Ω-gate shows worse short channel characteristics due to thicker T_{ox} and narrow band-gap in $Si_{0.7}Ge_{0.3}$.

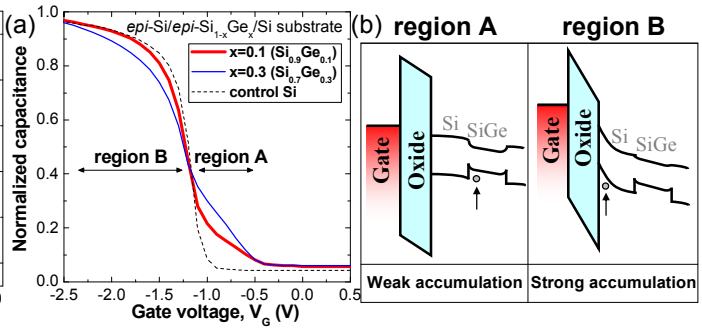


Fig. 8 (a) C-V characteristics measured from MOS capacitor and (b) band diagrams for different conditions. The hole transition from the bottom $Si_{1-x}Ge_x$ to the top Si results in double slopes in the C-V plot. Also, double slopes become clearer at higher x, which implies a deeper quantum well.

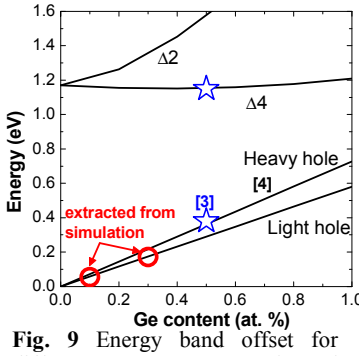


Fig. 9 Energy band offset for different Ge content x. Estimated valence band offset from calibrated C-V by simulation and theoretical data reported in [3-4] are plotted.

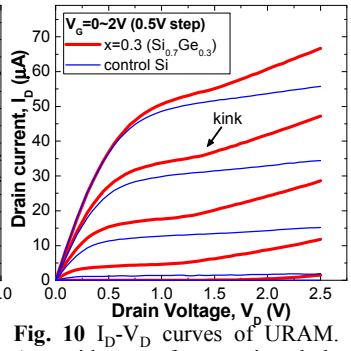


Fig. 10 I_D-V_D curves of URAM. As evidence of excessive hole accumulation, the kink effect is observed only in the band engineered quantum well substrate.

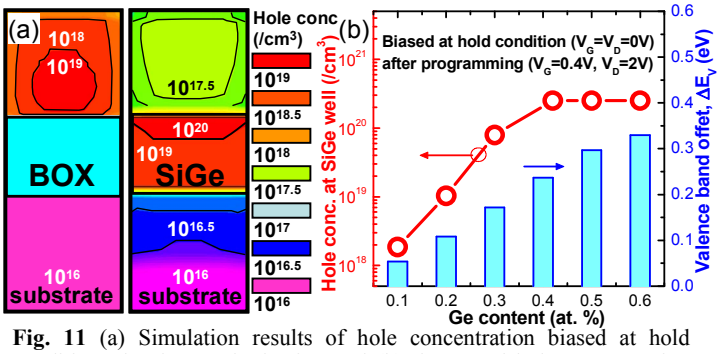


Fig. 11 (a) Simulation results of hole concentration biased at hold condition after impact ionization and (b) the stored hole concentration for various Ge content x. Excessive holes generated by impact ionization are accumulated in the $Si_{0.7}Ge_{0.3}$ quantum well. The stored hole concentration starts to be saturated from x=0.4 (ΔE_V=0.24eV).

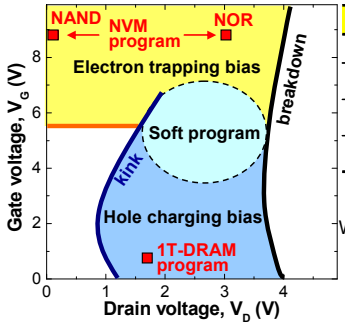


Fig. 12 Bias map for 1T-DRAM and NVM. Program voltage of 1T-DRAM is set as low as possible to avoid soft program into NVM.

	V_G	V_D	V_S	V_{Sub}
Write '1'	0.5	2	0	0
Write '0'	0.5	-2	0	0
Read	0.5	0.4	0	0
Hold	0	0	0	0

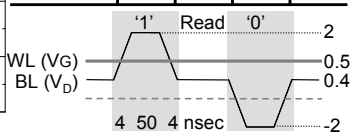


Table 2 Optimized operational bias conditions and pulse wave form for 1T-DRAM. 50 nsec pulse is used for writing.

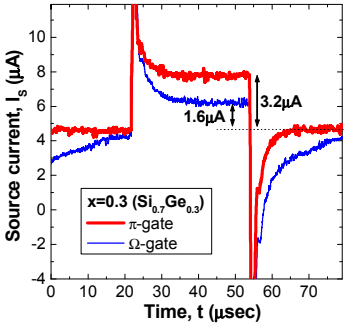


Fig. 13 1T-DRAM characteristics of π- and Ω-gate. Ω-gate is shown to be inferior to π-gate due to the reduced partially depleted volume.

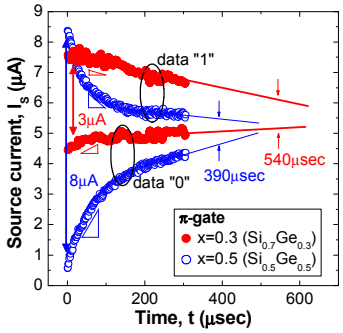


Fig. 14 1T-DRAM characteristics for different x in π-gate. x=0.5 exhibits a wider sensing window but shorter retention time.

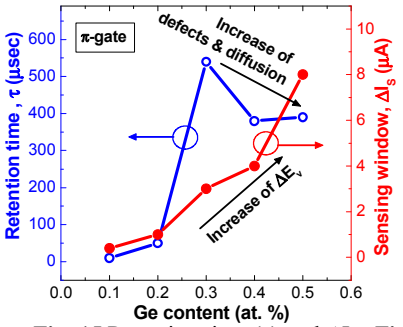
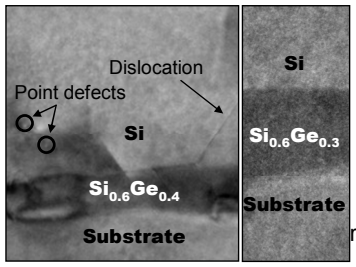


Fig. 15 Retention time (τ) and ΔI_S for different x. As x is increased, ΔI_S is accordingly enhanced, but τ is maximized at x=0.3



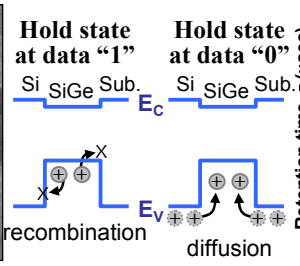


Fig. 16 (a) TEM images of x=0.4 and 0.3 and (b) schematics for retention degradation mechanisms. The defects reduce '1' retention as a result of recombination, and a deeper quantum well degrades '0' retention due to hole-to-hole repulsion and its diffusion mechanism.

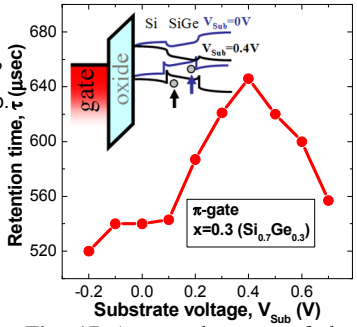
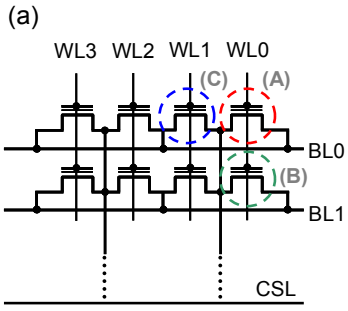


Fig. 17 As an advantage of the bulk type floating body, V_{Sub} modulates the energy band and enhances the data retention time.



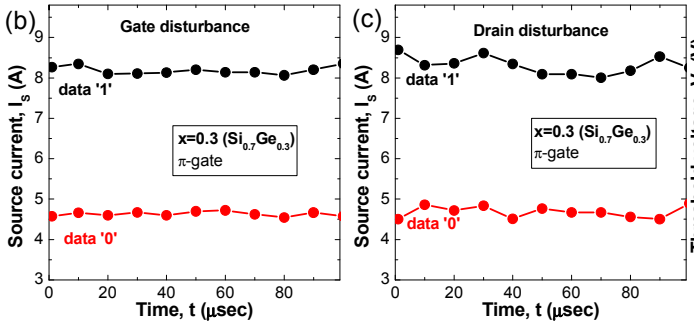


Fig. 18 (a) Schematic of array and (b) gate- (c) drain-disturbance characteristics. A 50 nsec pulse is applied every 1 μsec for disturbance. When cell (A) is programmed, cell (B)/cell (C) can experience gate- and drain-disturbance. This disturbance is found to be negligible for both cases.

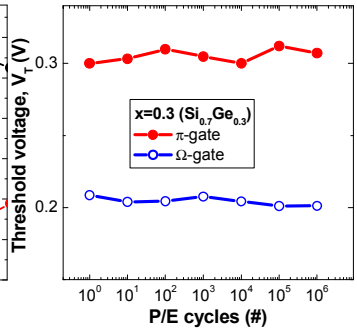
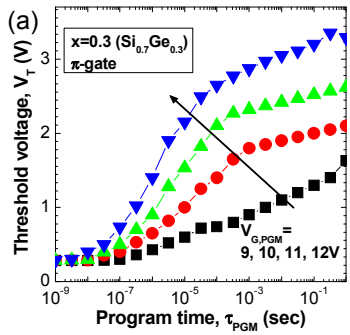


Fig. 19 V_T change during P/E cycles of 1T-DRAM. Even after 10^6 pulse cycles, charge trapping into O/N/O is negligible.



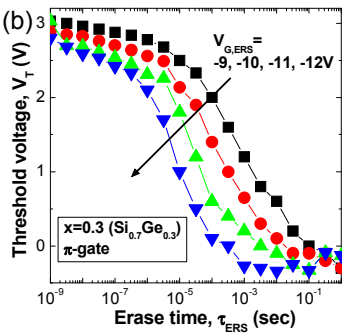


Fig. 20 (a) Program / (b) erase speed response of NVM operation. P/E is carried out by a FN tunneling mechanism. Threshold voltage (V_T) window is ~3 V at 12 V/100 μsec program and -12 V/100 μsec erase.

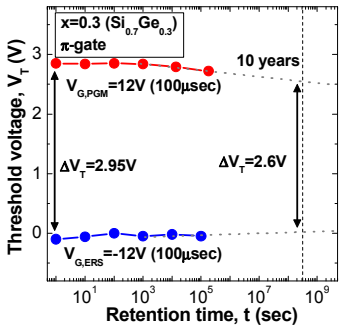


Fig. 21 Retention characteristics of NVM. V_T window of 2.6V ensures 10 year data retention.

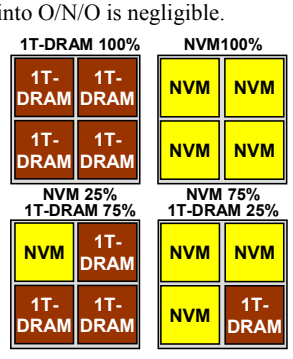


Fig. 22 Example of URAM for versatile memory chip application. The configuration of URAM can be adapted to the designer's demands.