A novel 2 T P-channel nano-crystal memory for low power/high speed embedded NVM applications*

Zhang Junyu(张君宇), Wang Yong(王永), Liu Jing(刘璟), Zhang Manhong(张满红), Xu Zhongguang(许中广), Huo Zongliang(霍宗亮), and Liu Ming(刘明)†

Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China

Abstract: We introduce a novel 2 T P-channel nano-crystal memory structure for low power and high speed embedded non-volatile memory (NVM) applications. By using the band-to-band tunneling-induced hot-electron (BTBTIHE) injection scheme, both high-speed and low power programming can be achieved at the same time. Due to the use of a select transistor, the "erased states" can be set to below 0 V, so that the periphery HV circuit (high-voltage generating and management) and read-out circuit can be simplified. Good memory cell performance has also been achieved, including a fast program/erase (P/E) speed (a 1.15 V memory window under 10 μs program pulse), an excellent data retention (only 20% charge loss for 10 years). The data shows that the device has strong potential for future embedded NVM applications.

Key words: P-channel; select transistor; nano-crystal memory; embedded
DOI: 10.1088/1674-4926/33/8/084006 **PACC:** 8730G

1. Introduction

Embedded non-volatile flash memory is becoming one of the key components in system-on-chip (SOC) products due to its data storage capability without a power supply. High speed and low power are required to meet the increasing demand for high performance SOC applications such as in mobile units. P-channel flash memory, introduced by Hsu *et al.* in 1992, was proposed as an alternative for producing high-density and high-speed flash memory[1].

Compared to the conventional N-channel flash memory, the P-channel flash memory can achieve a large programming gate current at low power consumption and provides high-speed programming[2]. As very large scale integration (VLSI) technologies progress and battery-powered electronics products prevail, the advantages of P-channel flash memory become more and more important. During the past two decades, several solutions have been proposed, such as P-channel floating gate (FG) flash memories with different P/E methods[3−5] and array architectures[6], and single-poly P-channel SONOS cells[7−9] for high-density and low cost. However, there is still quite a lot of work required to further optimize memory devices for future advanced NVM applications.

In this paper, for the first time, we introduced a 2 T P-channel nano-crystal (NC) memory for high density and low power nonvolatile memory application. By inserting a select transistor between the memory cell and the source connection, zero standby current and an enhanced read speed of the memory core can be achieved[10, 11]. It is shown that this type of NVM has advantages of process simplicity and can reduce the periphery circuit complexity.

2. Experimental setup

Figure 1 illustrates the schematic cross-section of the 2 T p-channel NC memory structure. The two P-channel transistors are formed in the N-well (NW) and connected in series; the one which has an inserting nano-crystal trapping layer is the memory transistor while the other with a thick gate-oxide is the select transistor.

Such 2 T NC memory devices are fabricated by using 1-poly/4-metal 0.13 μm CMOS technology. The fabrication of the 2 T NC memory followed a conventional logic process. The integration of the nano-crystal charge storage layer into conventional CMOS flow can be accomplished with the addition of only four non-critical masks over the baseline logic process. After trench isolation and well implantation, a charge storage stack is formed consisting of a 3-nm-thick thermal grown tunneling oxide (bottom oxide), a nano-crystal charge trapping

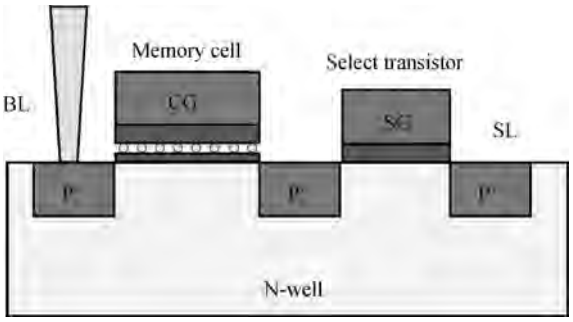

Fig. 1. Schematic of the embedded SONOS memory cell.

* Project supported by the National Basic Research Program of China (Nos. 2010CB934204, 2011CBA00600), the National Natural Science Foundation of China (Nos. 60825403, 60676008, 60676061), and the Hi-Tech Research and Development Program of China (Nos. 2008AA031403, 2009AA03Z306).

† Corresponding author. Email: liuming@ime.ac.cn

Received 19 January 2012, revised manuscript received 5 March 2012

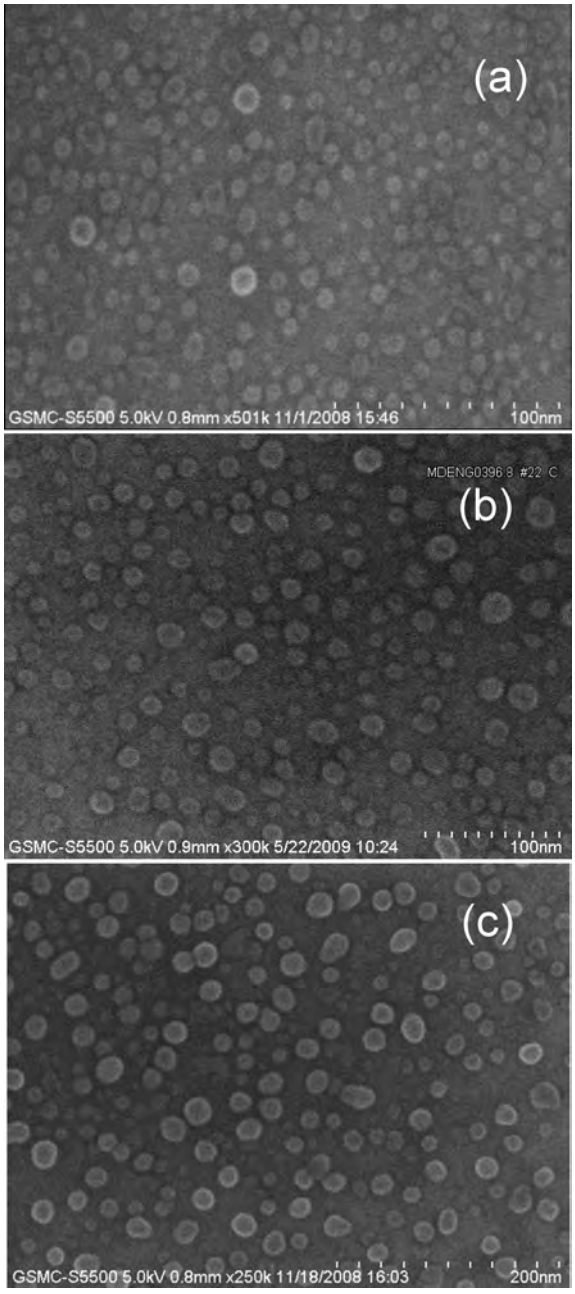

Fig. 2. SEM cross-sectional view of Si-NCs: (a) 5 nm (1 min growth time); (b) 10 nm (3 min growth time); (c) 20 nm (5 min growth time).

layer and 11-nm-thick high-temperature oxidation (HTO) as the top blocking oxide (top oxide). A two-step low pressure chemical vapor deposition (LPCVD) process was adopted for Si-NCs growth: nucleation at high temperatures (580–600 ℃) for 1–3 s, and growth at low temperatures (550 ℃) for 1–5 min. By adjusting the growth time from 1 min, 3 min, and 5 min, different nano-crystal sizes of 5 nm, 10, and 20 nm can be achieved, separately. Figures 2(a)–2(c) give scanning electron microscope (SEM) images of Si-NCs with different diameters of 5 nm, 10 nm and 20 nm.

Next, a mask is used to define the NC-transistor area and the three stacked layers in other places are etched away. After cleaning, the select transistor gate oxide is grown. Then, a 180-nm-thick gate poly-silicon layer is deposited and followed by a dry etch to form the whole transistor gate stack. Finally, spacer formation and source/drain implantation were performed to complete the device. The size of our 2 T memory cell is: $W/L = 0.32\ \mu$m/0.17 μm (selector transistor) and 0.32 μm/0.24 μm (memory transistor). The total area of the 2 T nano-crystal memory cell is 0.475 μm^2.

Table 1. Operation conditions for the NC memory cell.

Parameter	Sub (V)	SL (V)	BL (V)	SG (V)	CG (V)
Program	0	0	−6	FL	6
Erase	0	0	0	FL	−11
Read	0	0	−0.2	−2.5	Vread

3. Results and discussion

For conventional N-channel flash memories, the program schemes are usually channel hot electron injection (CHEI) or Fowler–Nordheim (FN) tunneling. By using the CHEI method, a high programming speed (typically 5–10 μs) can be guaranteed, but the power consumption is very large. For the FN tunneling, the program current is small but the program time would be several micro-seconds. In this work, we introduce P-channel technology, which adopts a band-to-band tunneling-induced hot-electron (BTBTIHE) injection programming scheme to solve this problem.

As for BTBTIHE programming, the programming is performed by: (1) floating the source voltage to shut off the channel current; (2) applying a high drain voltage (e.g., –5 to –6 V to form a strong-enough electrical-field near the drain junction); and (3) setting a suitable gate voltage to pull the generated-electrons into the stack dielectric. A typical operation table is shown in Table 1.

The cross section of BTBTIHE programming is shown in Fig. 3(a). Due to the high voltage applied at drain side, the band-to-band tunneling at the drain/gate overlap regions generate electron–hole pairs, so-called gate-induced drain leakage (GIDL). The generated electrons flow through the large electric field at the drain–gate overlap region and some of the electrons gain sufficient energy (i.e., 3.2 eV) to inject into the nano-crystal trapping layers (Fig. 3(b)).

During the program operation, the source side is floating, so there is no channel current, which can limit the program power, and since a higher gate current can be obtained[3], the programming time can be improved to around 5–20 μs[5], which is comparable with the CHE programming scheme.

Figure 4 shows the initial/program/erase states of the 2T P-channel NC memory cell. This drain current–gate voltage (I_d–V_g) curve is extracted at $V_d = 0.2$ V. For the device with 20 nm NCs, the initial threshold voltage (V_t) is about –0.6 V, after a 10 μs programming pulse, the programmed V_t shifts to 0.65 V, then a 10 ms –11 V pulse was applied to the memory transistor control gate to form the FN erase, the erased V_t is close to initial one, a memory window of about 1.3 V was achieved. The memory window (MW) of devices with different-diameter-NCs are compared in the inset picture, at the same program/erase condition, the device with larger NCs shows bigger MW due to an increased capture cross-section[12].

Such well-designed P/E states can reduce the complexity

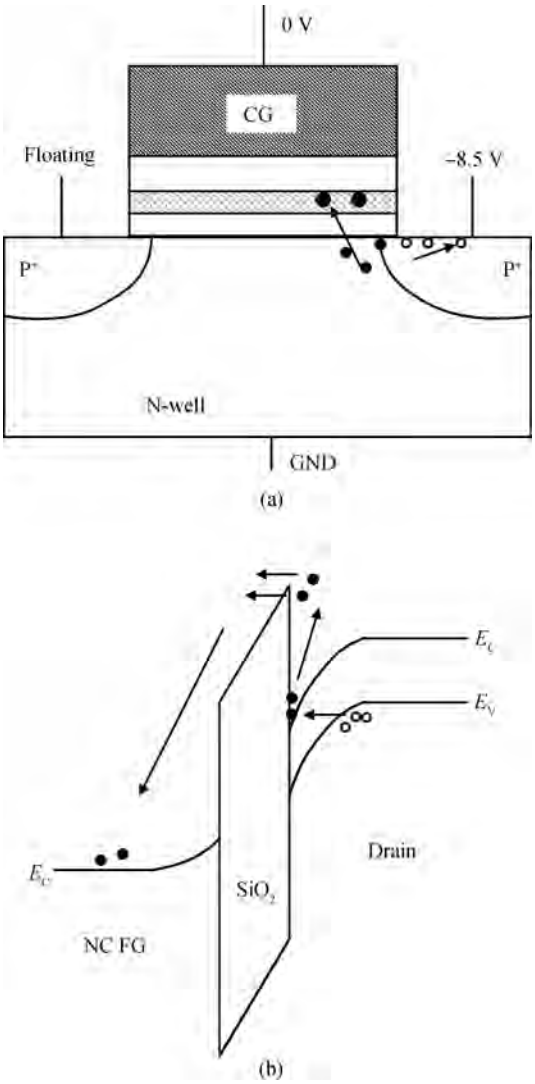

Fig. 3. Band-to-band tunnel-induced hot-electron (BTBTIHE) programming. (a) Cross section. (b) Band diagram.

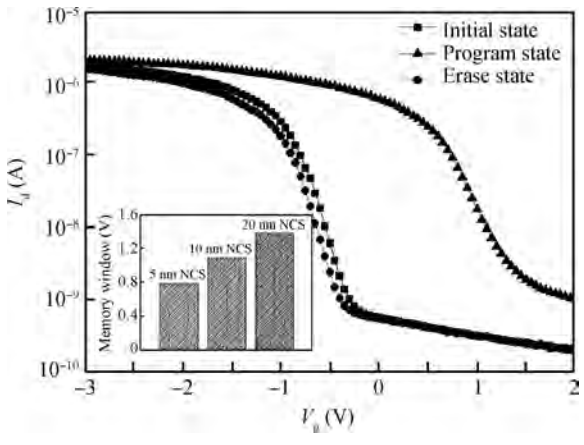

Fig. 4. I_d–V_g curve of NC memory. Inset shows the memory window of devices with nanocrystals of different diameters.

of peripheral circuit design. In a conventional design, both the erase and program V_t were set to be above 0 V, and usually a high word-line voltage (above V_{DD}) needed to be maintained

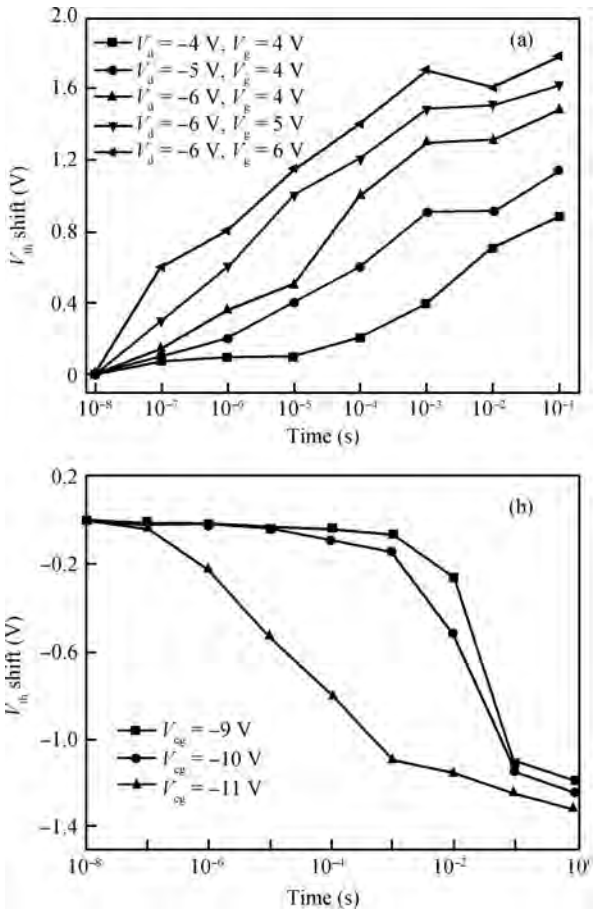

Fig. 5. P/E speed characteristics of NC memory cell.

during read access. To ensure the high voltage, a boosting circuit is needed, which consumes more power and introduces access latency. By using this design (programmed $V_t > 0$ V, erased $V_t < 0$ V), the word-line read voltage (i.e. 0 V) without boosting can be employed while producing a high-enough cell current to tip the balance of the sense amplifier in the specified access time budget.

Of course, to allow reliable and expedient sensing of a selected bit, all other unselected bits sharing the same bit-line must be shut off. This is achieved by the select transistor between the NC memory transistor and the source connection. This also ensures zero standby current from the memory core.

The program speed is compared in Fig. 5(a). The V_g is set to 4 V and the V_d changed from −4 to −6 V. With the increase of drain voltage, the program speed was also improved due to the increasing amount of generated electron-hole pairs and the higher vertical electric field. While $V_d = -6$ V, a larger gate voltage increased the program voltage. It is found that at the condition $V_d = -6$ V & $V_g = 6$ V, a 1.15 V V_t shift can be obtained under a 10 μs program pulse, which can meet the high-speed NVM requirement. The erase characteristics are shown in Fig. 5(b), FN tunneling is used for the cell erase operation, so with the higher gate voltage, faster erasing can be gained. When the $V_g = -11$ V and pulse width = 1 ms, a 1.12 V V_t shift is achieved, which shows a good agreement with the program condition.

Data retention characteristics at room temperature for the

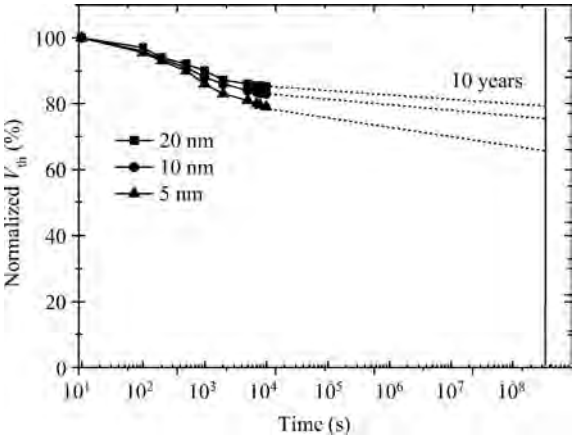

Fig. 6. Memory cell retention characteristics at room temperature.

NC memory cells are shown in Fig. 6. As is known, one of the main advantages of NC NVM is the potential good data retention. Due to discrete charge storage, it is expected that a current leakage path in the tunneling oxide of one NC does not cause the charges in other NCs to leak. Thus, an NC memory cell can still maintain good retention even when the tunnel oxide thickness is very thin. We can find that a 20 nm-NCs-device has only 20% charge lost for ten years retention time, which guarantees the high performance of NVM applications. Effect of nano-crystal size on data retention is also shown in Fig. 6, the devices with smaller NCs show a worse retention characteristic because of the more serious quantum confinement effect[13].

4. Conclusion

In summary, a 2 T P-channel NC memory structure is demonstrated in this work. The device process is simple and fully compatible with conventional CMOS technology. The BTBTIHE injection programming scheme is adopted to realize high-speed and low power programming. As a select transistor has been employed, high read speed can be achieved by using a simplified periphery circuit design. This device shows good performance in terms of program/erase speed and data retention, and also with an acceptable storage widow of more than 1.5 V, which make the 2 T P-channel nanocrystal memory a good candidate for the next generation of high performance embedded NVM applications.

References

[1] Hsu C C H, Acovic A, Dori L, et al. A high speed low power p-channel flash EEPROM using silicon rich oxide as tunneling dielectric. The Conference on Solid State Devices and Materials, 1992: 140

[2] Ohnakado T, Ajika N, Hayashi H, et al. 1.5 V operation sector-erasable flash memory with bipolar transistor selected p-channel cells. Symp VLSI Technol Dig Tech Papers, 1998: 14

[3] Ohnakado T, Mitsunaga K, Nunoshita M, et al. Novel electron injection method using band-to-band tunneling induced hot electron (BBHE) for flash memory with a p-channel cell. IEDM Tech Dig, 1995: 279

[4] Shen S J, Yang C S, Wang Y S, et al. Novel self-convergent programming scheme for multilevel p-channel flash memory. IEDM Tech Dig, 1997: 287

[5] Chung S S, Kuo S N, Yih C M, et al. Performance and reliability evaluations of p-channel flash memories with different programming schemes. IEDM Tech Dig, 1997: 295

[6] Ohnakado T, Takada H, Hayashi K, et al. Novel self-limiting program scheme utilizing n-channel select transistors in p-channel DINOR flash memory. IEDM Tech Dig, 1996: 181

[7] Sim J S, Park I H, Park H, et al. BAVI-cell: novel high-speed 50 nm SONOS memory with band-to-band tunneling initiated avalanche injection mechanism. Symp VLSI Technol Dig Tech Papers, 2005: 122

[8] Shukuri S, Ajika N, Mihara M, et al. A 60 nm NOR flash memory cell technology utilizing back bias assisted band-to-band tunneling induced hot electron injection (B4-flash). Symp VLSI Technol Dig Tech Papers, 2006: 15

[9] Lee H M, Woo S T, Chen H M, et al. Neoflash-true logic single poly flash memory technology. IEEE NVSMW, 2006: 15

[10] Liu W, Chang K T, Cavins C, et al. A 2-transistor source-select (2TS) flash EEPROM for 1.8 V-only applications. IEEE NVSM Workshop, 1997: 4.1.1

[11] Yang Xiaonan, Wang Yong, Zhang Manhong, et al. A novel 2-T structure memory device using a Si nanodot for embedded application. Journal of Semiconductors, 2011, 32(12): 124007

[12] Rao R A, Gasquet H P, Steimle R F, et al. Influence of silicon nano-crystal size and density on the performance of non-volatile memory arrays. Solid-State Electron, 2005, 49: 1722

[13] Guan W, Long S, Liu M, et al. Modeling of retention characteristics for metal and semiconductor nano-crystal memories. Solid-State Electron, 2007, 51: 806