SAN098-1640C

2K Nonvolatile Shadow RAM and 256K EEPROM SONOS Nonvolatile Memory Development

R. D. Nasby, J. R. Murray, S. D. Habermehl, R. S. Bennett, B. C. Tafoya-Porras, P. R. Mahl,
J. L. Rodriguez, R. V. Jones, M. G. Knoll
Sandia National Laboratories, Albuquerque, NM 87185

SAND--98-1640C
CONF-980674--

Abstract

This paper describes SONOS nonvolatile memory development at Sandia National Laboratories. A 256K EEPROM nonvolatile memory and a 2K nonvolatile shadow RAM are under development using an n-channel SONOS memory technology. The technology has 1.2 μm minimum features in a twin well design using shallow trench isolation.

1. Introduction

The 256K EEPROM was designed to operate in high radiation environments. The circuit was designed for fabrication in a CMOS/SONOS (Complementary Metal Oxide Semiconductor / Silicon Oxide Nitride Oxide Semiconductor) process since it offers maximum radiation hardness for nonvolatile circuits [1-6]. The 256K (32K x 8) EEPROM is organized as 512 pages with 64 bytes in each page. Data is programmed into the memory one page at a time. Each memory cell contains two SONOS transistors with the difference in their threshold voltages representing the binary value of the bit. For the data in the cell to be read correctly, one or both of the SONOS transistors must be depletion mode.

One disadvantage of an EEPROM is the slow write time (i.e. 10ms programming). A solution is to combine a fast CMOS static RAM with a SONOS backup in a shadow RAM configuration. The static RAM will allow fast writing and reading of the data. The SONOS shadows will retain the data through loss of power. The 2K shadow RAM consists of a 2K (256 x 8) static RAM with two nonvolatile shadows. The memory cell also includes a latch to hold the data being programmed. This allows full access to the static RAM while the shadow is being programmed. The NV shadows are used to back-up the data in the static RAM allowing the circuit to be powered down without losing critical data. Two shadows are required so that if power is removed while one shadow is being programmed, corrupting the data storage; the second shadow will contain valid data.

2. EEPROM Circuit Description

The general specification goals for the designs are listed in Table 1.

Table 1.Specification Goals

Power Supply:	10V Pgm (-5V - +5V)
	5V (i/o),
Clear/Write Time:	7.5ms/2.5ms (EE)
	450us/150us (SR)
Endurance:	1E4 Write Cycles (EE)
	1E7 Write Cycles (SR)
	Infinite Read Cycles
Read Access Time:	<200ns (EE)
	<50ns (SR)
Retention:	10 Yr. (80C) (EE)
	100 Yr. (80C ROM)
	24 Hrs (SR)
Temperature:	-55C - +125C
Radiation	
Total Dose:	1Mrad(Si)
Transient	
Read/Write:	>5e8 rad(Si)/s
Data Upset:	>1e12 rad(Si)/s
Single Event	
Soft Error:	LET > 35 MeV-cm^2/g
Hard Error:	>Ar

A block diagram of the 256K EEPROM is shown in Figure 1. The three control signals (i.e. chip enable bar (CEB), write enable bar (WEB), and output enable bar (OEB)) determine the memory's mode of operation. Writing data into the memory is a two step process: loading 64 bytes into the data-in latches and then programming the latched data into a page of the

DISCLAIMER

memory. The page address is latched on chip when the first of the 64 bytes is written. The programming time for the EEPROM is controlled by an internal counter and the externally supplied clock input. After the last byte is written, the internal timer waits 546 clock cycles before starting the programming. The clock frequency for the counter is specified to be in the range of 1 to 2 MHz to ensure 10 year data retention.

The DATA polling mode can be used to verify the completion of programming. To monitor for completion of programming the user can read the last address written until the correct data is read.

3. EEPROM Memory Cell Operation

Each memory bit consists of two SONOS memory transistors and two n-channel access transistors. The data is stored as the difference in the threshold voltages of the two SONOS transistors. Storing data in a memory cell is a two step process. First both SONOS transistors are

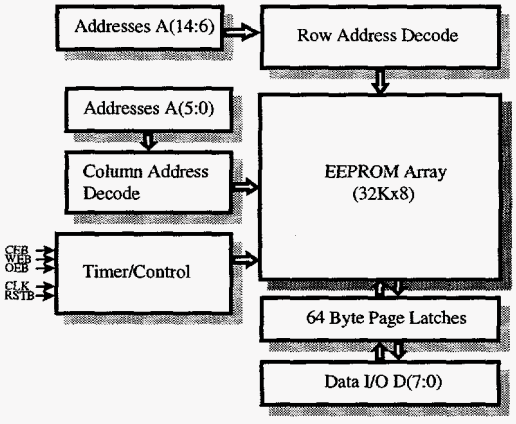

Figure 1. EEPROM Block Diagram

erased to the most depletion threshold voltage (Vth). Then one of the SONOS transistors is programmed to the most enhancement Vth while the other SONOS transistor's Vth remains in the most depletion state.

To read the cell, the n-channel access transistors are turned on while the SONOS gates and drains are grounded. The sense amplifier detects which of the two SONOS transistors can sink more current. For the circuit to operate correctly either one or both SONOS transistors must be depletion mode.

4. 2K Shadow RAM Operational Description

The shadow RAM was designed for fabrication in a double level metal SONOS/CMOS process. A block diagram of the circuit is shown in Figure 2. The static RAM portion of the circuit operates in the standard

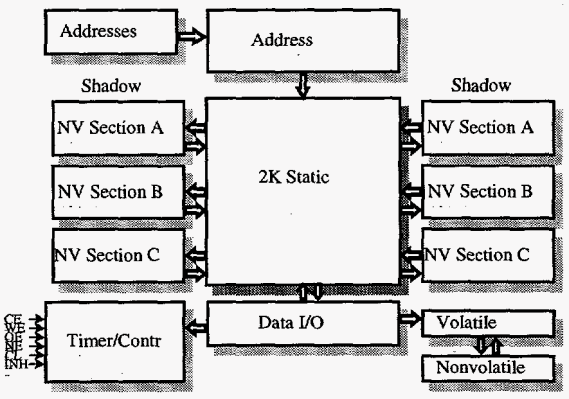

Figure 2. Shadow RAM Block

fashion with either asynchronous or synchronous addressing (address latches are included for synchronous operation). The NV shadows cannot be accessed directly, but only through the static RAM. The NV shadows are divided into three sections, which can be programmed independently. In addition to the static RAM and NV shadows, three single bit volatile flags with non-volatile backup are included. The flags are used to indicate which shadow should be used for recall and storage. The flag bit is bit 0 (D0) of the eight data bits (D0-D7).

There are four basic NV commands; erase, store, recall, and load erase/store time. The erase and store commands initiate programming the shadow. Programming terminates when the on-chip timer times out. Erase is a preconditioning command that is required before the data can be copied into a shadow by the store command. The timer is a counter which is started at an initial count that is set by the load erase/store time commands.

Each bit of the non-volatile shadows consists of two SONOS transistors and two NMOS access transistors. The difference in the thresholds of two SONOS transistors determines if the bit is a one or a zero. Two SONOS transistors per bit improve the retention and reliability of the NV shadows. To program the SONOS transistors, a -5 volt supply is needed. The part includes a circuit to generate this voltage internally or it can be supplied externally.

5. Shadow RAM Memory Cell Design

The Shadow RAM cell consists of four memory elements: a static RAM cell, a latch, and two non-volatile shadows. Eight transistors are used to form the static RAM cell instead of the usual six. The extra transistors allow the cell to be disconnected from the negative supply rail during recall operations.

The latch is used to hold the data being programmed into one of the shadows. During a write operation, the data is first shifted into the latch. Then the latch is connected to the NV shadow, controlling the data programming. This allows the static RAM to be accessed during programming operations.

Four transistors (two SONOS and two NMOS) form each of the two NV shadows. The n-channel transistors are used only during the recall operation to select which shadow is recalled. The data is stored in a shadow as a difference in threshold voltage (Vth) between the two SONOS transistors.

During static RAM operation the latch transistors are turned on and the negative rail voltage is at 0 volts. The cell then functions as a standard six transistor static RAM cell.

6. Flag Circuit Design

The device includes three flag bits which are used to record which of the shadows has the valid data. Each flag bit consists of a volatile latch that can be read and written directly and a non-volatile backup bit. When reading or writing the flags, the flag bit is bit 0 (D0) and all the rest of the bits are held low.

The flag bit consists of three parts: the volatile latch, the non-volatile bit, and the sense amplifier. The non-volatile memory bit consists of two SONOS transistors (S1, S2) which are programmed to opposite states. During a flag store, the gate of S2 and the backgate of S1 are connected to the output of the volatile latch. The not output of the latch is connected to S1's gate and S2's backgate. This results in one transistor's threshold voltage being increased while the threshold voltage of the other is reduced. This difference in threshold voltages represents the data from the volatile latch.

To recall the data from the non-volatile bit, the SONOS transistors are connected to the inputs of the sense amplifier and the output of the sense amplifier is fed into the volatile latch. The SONOS transistors are connected as source followers. The differential amplifier senses the difference in the output voltages of the two SONOS source followers. An output buffer drives the input of the volatile latch.

7.Memory and SONOS Stack Processing

The memory is processed using a twin well with trench isolation and single level poly with two level metal. This is the first SONOS memory processed at Sandia using trench isolation. A memory gate first flow is used. The wells are first formed and the trench isolation process ends with a CMP planarized surface. The SONOS stack is formed and then is removed by dry etching from CMOS areas. The CMOS gates are grown and the single poly deposited and defined. The source/drain regions are formed and then the two level metal backend using CMP planarization completes the process.

The SONOS memory transistors investigated are 1.2 μ n-channel devices with memory stacks for 10 V operation. Test transistors were fabricated using shallow trench isolation in separate short development lots and in complete memory fabrication lots. Several approaches for stack fabrication were investigated to meet requirements for the two applications with write cycle times of 10 msec and ~ 500 μsec. The stack consists of a tunnel oxide, oxynitride layer and a top oxide with process variations in both the growth of the tunnel oxide and in the manner of forming the top oxide. Three distinct SONOS stacks are presented. The first (1) consists of a 1.2 nm tunnel oxide (grown at 600°C in O_2 for 10 minutes), with a 11.5 nm LPCVD deposited Si oxynitride charge trapping layer (800°C, 200 mTorr, 30 sccm $SiCl_2H_2$, N_2O:NH_3 ratio=1.85), capped with a 3.5 nm LPCVD deposited SiO_2 layer (800°C, 200 mTorr, N_2O:$SiCl_2H_2$ ratio=3). No anneal of the top oxide was performed after deposition. The second (2) stack consists of the same tunnel oxide, with a 14.0 nm deposited oxynitride, which is then pyrogenically oxidized to form a 3.5 nm top SiO_2 layer (900°C, 1 ATM, H_2:O_2=1.75). The third stack (3) is identical to the first, with the exception that the tunnel oxide is formed in a nitrous oxide ambient (850°C in N_2O at 1 Torr for 15 minutes).

7. Non Volatile Memory Transistor Characteristics

Figures 3-5 show the retention and cycling behavior of these three stacks. These data were obtained using a write of 150 μs and an erase of 450 μs, both at 10V. To simulate "worse case" device operation, the write pulse occurred after 10 erase pulses to represent relative erase saturation. The erase pulse occurs after 10 cycles of E/W to mimic device operation. Both states are measured as a function of retention time after having been cycled as indicated, up to 1E7 cycles. For the first stack the window thresholds shift strongly positive after 1E7 cycles, such that the lower state would no longer be in depletion.

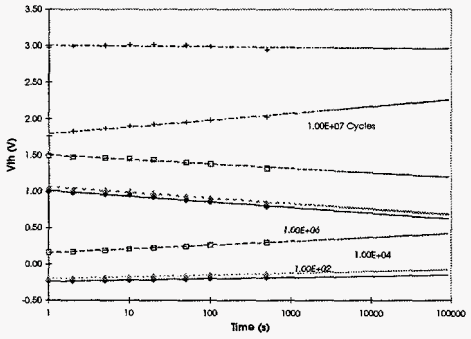

Figure 3. Retention /Cycling Shadow Ram Characteristics for 150 μs write, 450 μs erase, 10V, room temperature. (1): Tunnel oxide grown in O_2 and a deposited top oxide.

The second device shows much less positive shift such that an initial threshold adjust would allow for operation after 1E7 cycles.

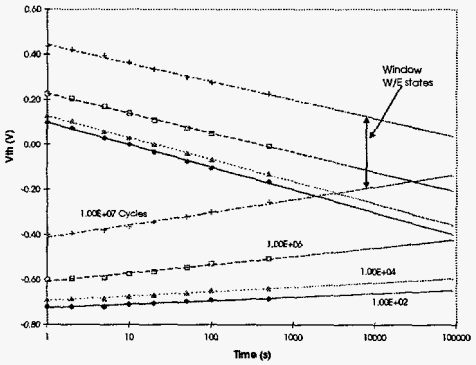

Figure 4. Retention/Cycling Shadow Ram Characteristics for 150 μs write, 450 μs erase, 10V, room temperature. (2): Oxidation of tunnel oxide in O_2 and steam oxidation for top oxide.

The third sample again shows a decreased shift. A change in the decay rate is also apparent, but both latter devices have sufficient retention for the 24 hr. retention requirement.

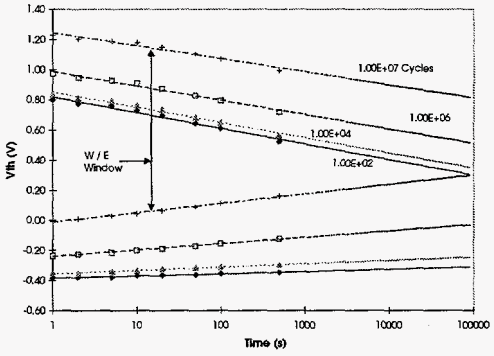

Figure 5 Retention/Cycling Shadow Ram Characteristics for 150 μs write, 450 μs erase, 10V, room temperature. (3): Oxidation of tunnel oxide in N_2O, deposited top oxide

In Figure 6 the stack was written using the EEPROM write conditions and with cycling to 1E4, a typical specification for EEPROM memories. Simple extrapolation using the last four points of the data indicates good retention with a window of greater than 0.5 volts at 10 years. Such extrapolation is of course, risky; longer-term tests and temperature tests are necessary to further characterize the behavior.

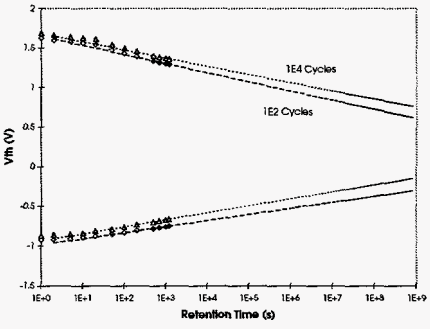

Figure 6. Retention/Cycling EEPROM Characteristics for 2.5 ms write, 7.5 ms erase, 10V, room temp. Oxidation of Tunnel oxide in N_2O and deposited top oxide.

The next figure summarizes the threshold shifts for the Shadow RAM conditions. Figure 7 shows the threshold voltage shift, δV_{th}, as a function of the number of W/E cycles, for transistors in saturation. Above 10^4 cycles, $\delta V_{\tau\eta}$ for the first sample increases dramatically over

$\delta V_{\tau\eta}$ for the N_2O formed tunnel oxide stack and the steam treated sample.

Shifts in the threshold voltage can be attributed to changes in the amount of fixed charge or interface trapped charge retained during cycling. The data in figure 8 elucidates the mechanism behind the observed threshold voltage behavior. The channel mobility, μ_n, and the sub-threshold swing, S, are monitored as a function of cycles. Changes in S can be directly correlated to changes in the Si/SiO_2 interface state density, D_{it}, while changes in μ_n are dependent upon changes in D_{it}. Both μ_n and S are observed to significantly degrade in transistors prepared with the conventional SiO_2 tunnel oxide, while transistors with the nitrided tunnel oxide exhibit less degradation. The observed degradation and resulting threshold voltage shifts are believed to be a result of an increase in interface state density with cycling. This conclusion is supported by evidence that thermal oxidation in N_2O environments leads to a buildup of N at the Si/SiO_2 interface [7]. Nitrogen occupies defect sites at the interface with activation energy sufficient to resist interface state creation.

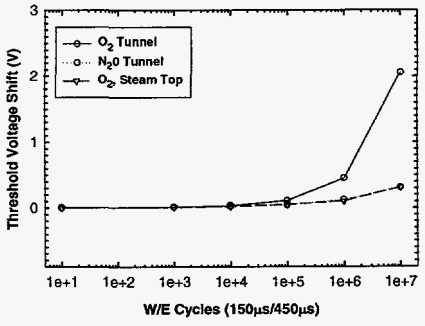

Figure 7. Threshold shift as a function of cycling for 3 samples: Shadow RAM W/E.

As shown in Figure **7,** the threshold voltage shift for stack 2, prepared with O_2 tunnel oxidation and a pyrogenically grown top oxide, shows a similar cycling response to that observed with the N_2O grown tunnel oxide stack. Saks, et. al., have proposed a mechanism for the oxidation of oxynitrides whereby N is released as a by-product [8]. Although more analysis is warranted, such as anneals after deposition of the top oxide, comparison to the results for the N_2O formed

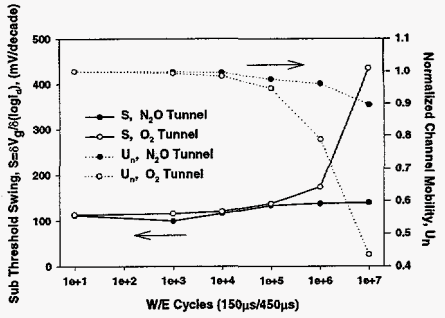

Figure 8. The mobility, μ, and sub-threshold swing, S, as a function of cycling for samples with tunnel oxides grown in oxygen and nitrous oxide

tunnel oxide indicates that diffusion of the evolved N to the interfacemay be responsible for the enhanced cycling endurance of this stack.

8. Summary

Data on test transistors indicate adequate nonvolatile behavior for further development and optimization of the Shadow RAM memory and EEPROM. The first Shadow RAM lot has been completed and the design verified.

9. Acknowledgments

We thank the staff of Sandia's Microelectronics Development Laboratory for providing the process support for this work and M. J. Rightley and J. B. Stanley for the electrical test support.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the Department of Energy under Contract DE-AC04-94AL85000.

10. References

[1] P. J. McWhorter, S. L. Miller, and T. A. Dellin, "Radiation Response of SONOS Nonvolatile Transistors", IEEE Trans. Nucl. Sci. NS-33, pp. 1414-1419 (1986).

[2] P. J. McWhorter, S. L. Miller, T. A. Dellin and C. A. Axness, "Retention Characteristics of SONOS Nonvolatile Devices in a Radiation Environment", IEEE Trans. Nucl. Sci. NS-34, pp. 1652-1657 (1987).

[3] E. S. Synder, P. J. McWhorter, T. A. Dellin and J. D. Sweetman, "Radiation Response of Floating Gate EEPROM Memory Cells", *IEEE Trans. Nucl. Sci.* NS-36, pp. 2131-2139 (1989).

[4] T. F. Wrobel "Radiation Characterization of a 28C256 EEPROM", IEEE Trans. Nucl. Sci. NS-36, pp. 2247-2251 (1989).

[5] J. R. Schwank, R. D. Nasby, S. L. Miller, M. S. Rodgers, and P. V. Dressendorfer, "Total-Dose Radiation-Induced Degradation of Thin Film Ferroelectric Capacitors", IEEE Trans. Nucl. Sci. NS-37, pp. 1703-1712 (1990).

[6] J. M. Benedetto, W. M. DeLancey, T. R. Oldham, J. M. McGarrity, C. W. Tipton, M. Brassington, and D. E. Fisch, "Radiation Evaluation of Commercial Ferroelectric Nonvolatile Memories", IEEE Trans. Nucl. Sci. NS-38, pp. 1410-1414 (1991) [7]

[7] S. Singhvi and C. Takoudis, J. Appl. Phys. 82(1), 442 (1997).

[8] N. Saks, D. Ma and W. Fowler, Appl. Phys. Lett. 67(3), 374 (1995).