US 20140349472A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2014/0349472 A1

(43) Pub. Date: Nov. 27, 2014

Chin et al.

(54) FLASH MEMORY

(71) Applicant: **National Chiao Tung University**, Hsinchu (TW)

(72) Inventors: **Albert Chin**, Hsinchu City (TW); **Chun-Yang Tsai**, Hsinchu City (TW)

(21) Appl. No.: **14/455,271**

(22) Filed: **Aug. 8, 2014**

Related U.S. Application Data

(62) Division of application No. 13/308,959, filed on Dec. 1, 2011, now Pat. No. 8,836,009.

Publication Classification

(51) **Int. Cl.**

H01L 29/51 (2006.01)

H01L 29/49 (2006.01)

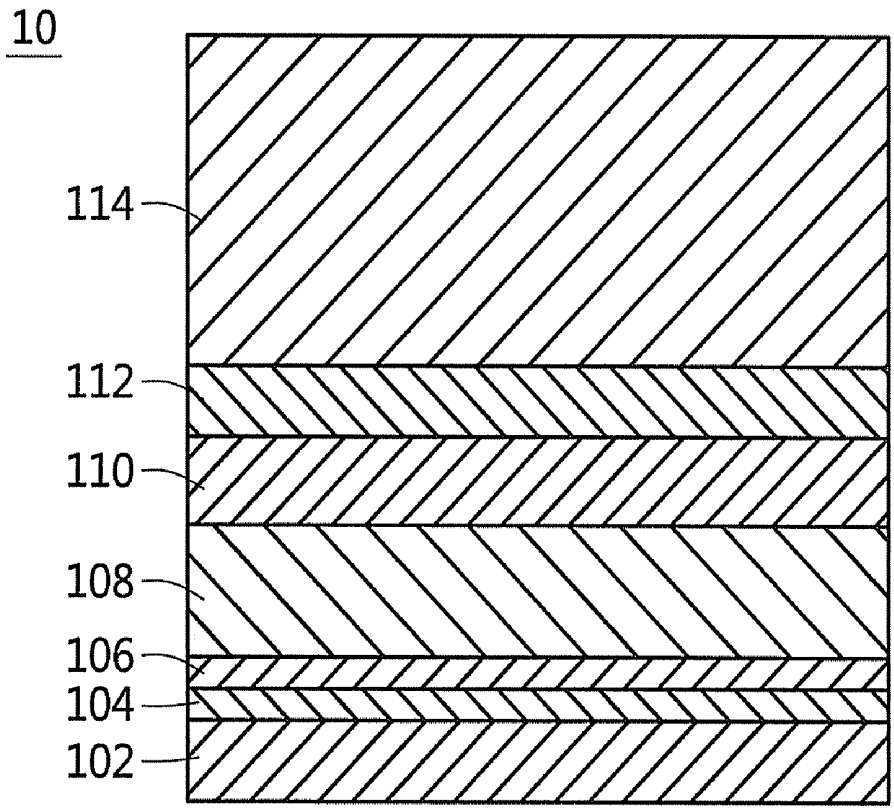
(52) **U.S. Cl.**

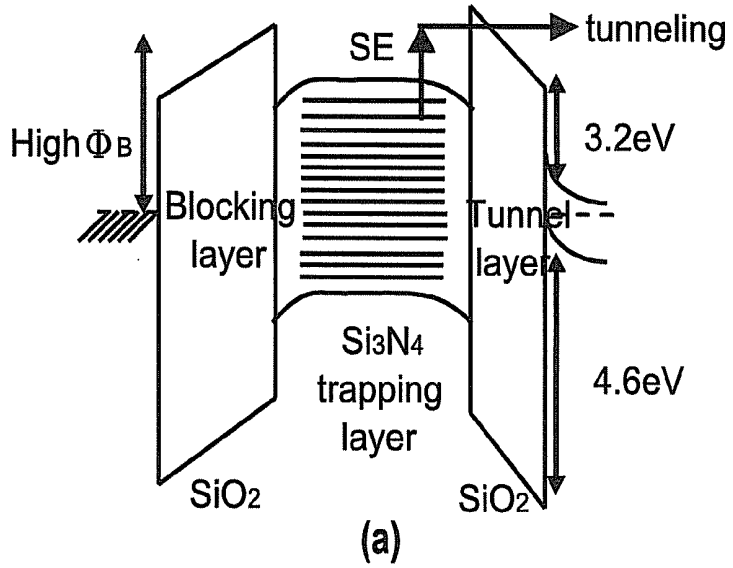
CPC *H01L 29/513* (2013.01); *H01L 29/4966* (2013.01); *H01L 29/518* (2013.01)

USPC **438/591**

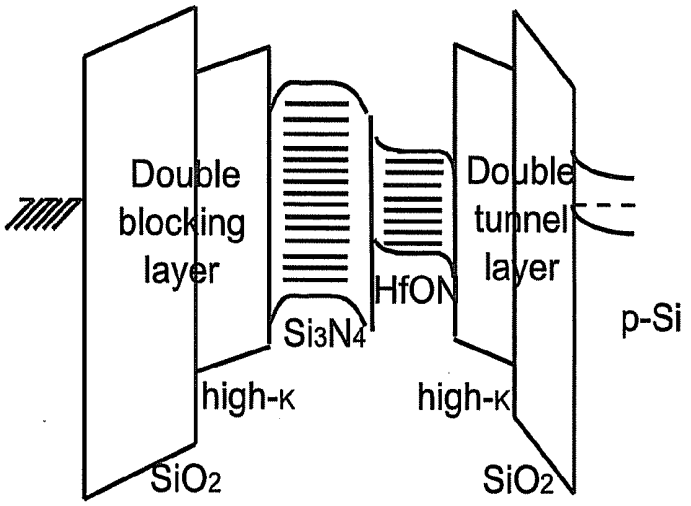
(57) **ABSTRACT**

A MONOS Charge-Trapping flash (CTF), with record thinnest 3.6 nm ENT trapping layer, has a large 3.1 V 10-year extrapolated retention window at 125° C. and excellent 10^6 endurance at a fast 100 μs and ±16 V program/erase. This is achieved using As^+-implanted higher κ trapping layer with deep 5.1 eV work-function of As. In contrast, the un-implanted device only has a small 10-year retention window of 1.9 V at 125° C. A MoN—[SiO_2—$LaAlO_3$]—[Ge—HfON]—[$LaAlO_3$—SiO_2]—Si CTF device is also provided with record-thinnest 2.5-nm Equivalent-Si_3N_4-Thickness (ENT) trapping layer, large 4.4 V initial memory window, 3.2 V 10-year extrapolated retention window at 125° C., and 3.6 V endurance window at 10^6 cycles, under very fast 100 μs and low ±16 V program/erase. These were achieved using Ge reaction with HfON trapping layer for better charge-trapping and retention.

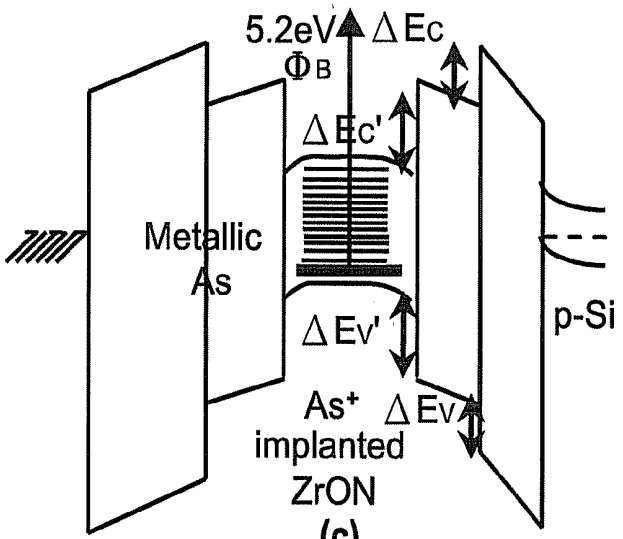



SE
tunneling
High ΦB
3.2eV
Blocking layer
Tunnel layer
Si3N4 trapping layer
4.6eV
SiO2
SiO2

(a)


Double blocking layer
Si3N4
HfON
Double tunnel layer
p-Si
high-κ
high-κ
SiO2
SiO2

(b)


5.2eV
ΔEc
ΦB
ΔEc'
Metallic As
p-Si
ΔEv'
As+ implanted ZrON
ΔEv

(c)

FIG.1

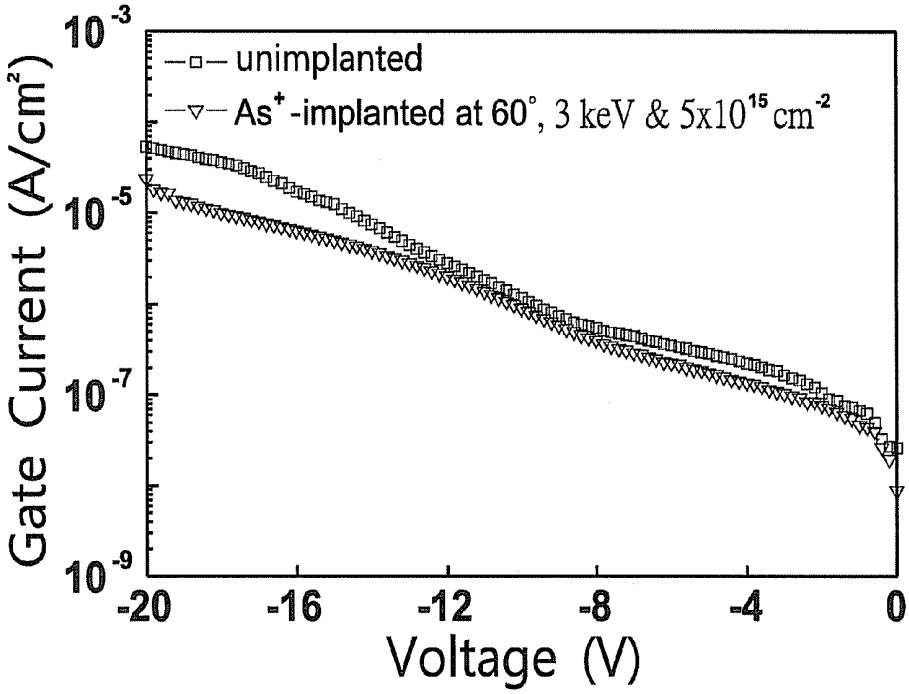

Gate Current (A/cm²)
10⁻³
10⁻⁵
10⁻⁷
10⁻⁹
unimplanted
As⁺-implanted at 60°, 3 keV & 5x10¹⁵ cm⁻²
-20
-16
-12
-8
-4
0
Voltage (V)

FIG. 2

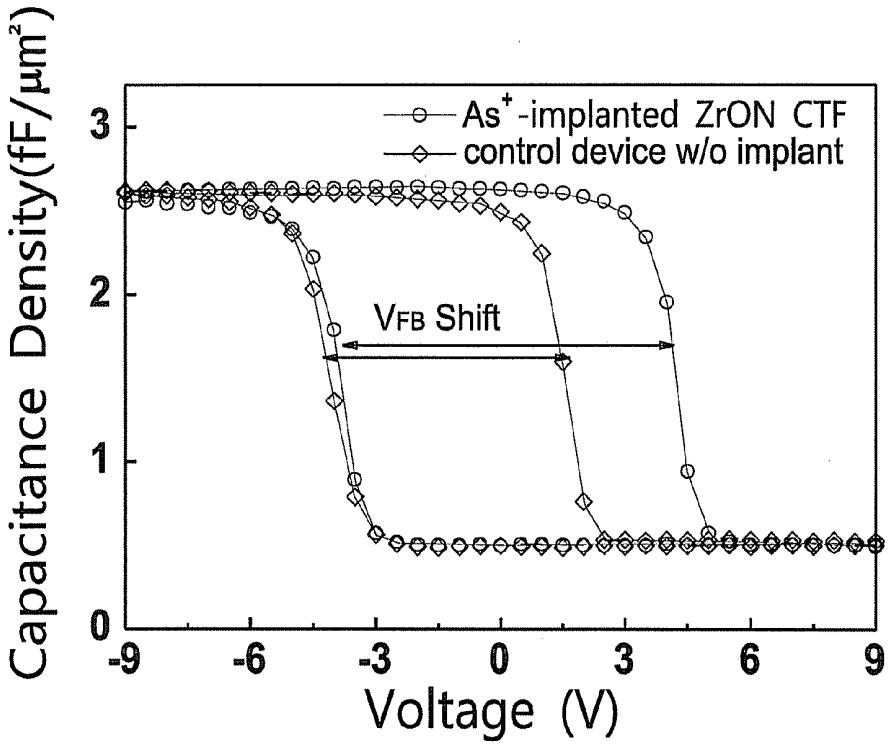

Capacitance Density(fF/μm²)
3
2
1
0
As⁺-implanted ZrON CTF
control device w/o implant
VFB Shift
-9
-6
-3
0
3
6
9
Voltage (V)

FIG. 3

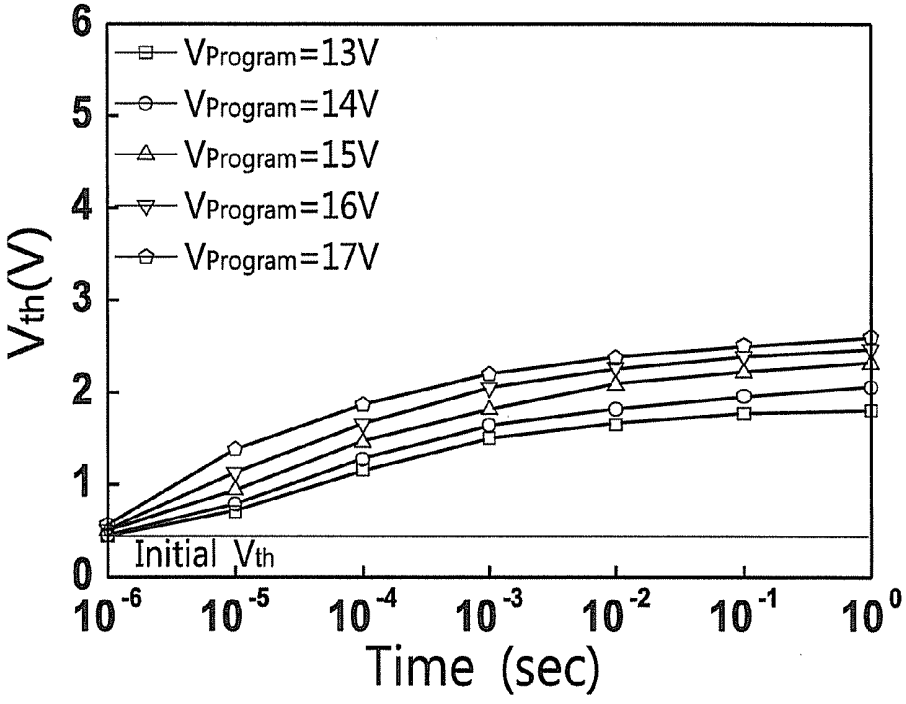

VProgram=13V
VProgram=14V
VProgram=15V
VProgram=16V
VProgram=17V
Initial Vth
Vth(V)
Time (sec)

FIG. 4

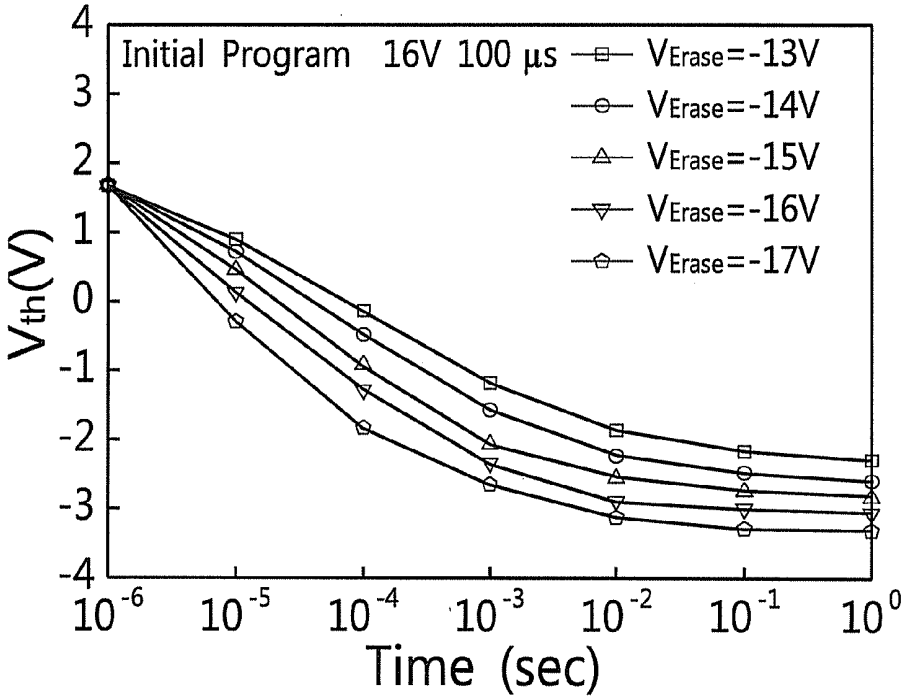

Initial Program 16V 100 μs
VErase=-13V
VErase=-14V
VErase=-15V
VErase=-16V
VErase=-17V
Vth(V)
Time (sec)

FIG. 5

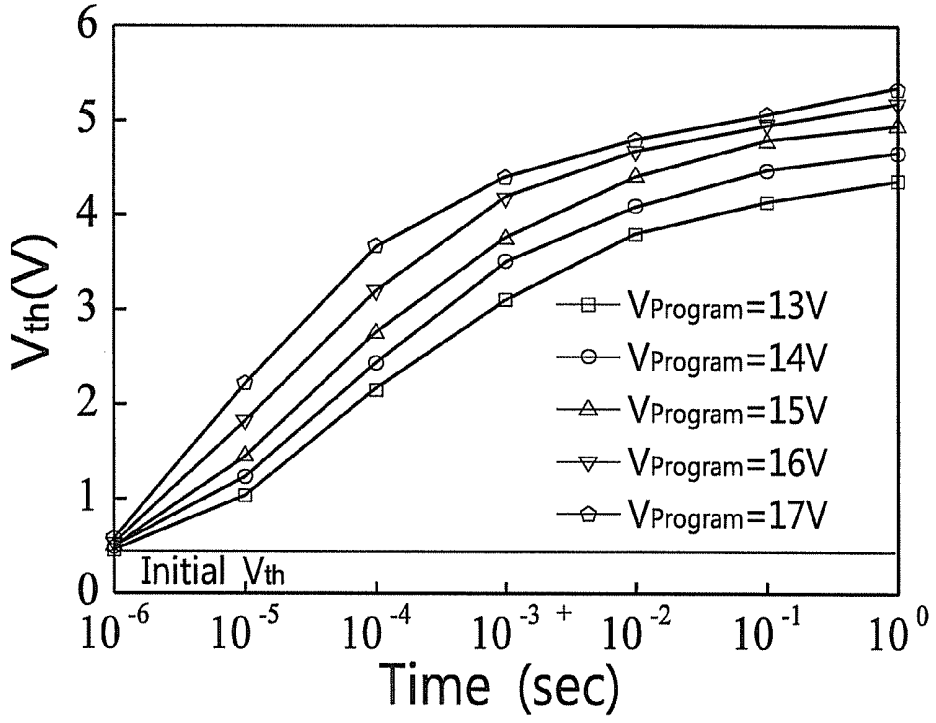

6
5
4
3
2
1
0
Vth(V)
VProgram=13V
VProgram=14V
VProgram=15V
VProgram=16V
VProgram=17V
Initial Vth
10-6
10-5
10-4
10-3
10-2
10-1
100
Time (sec)

FIG. 6

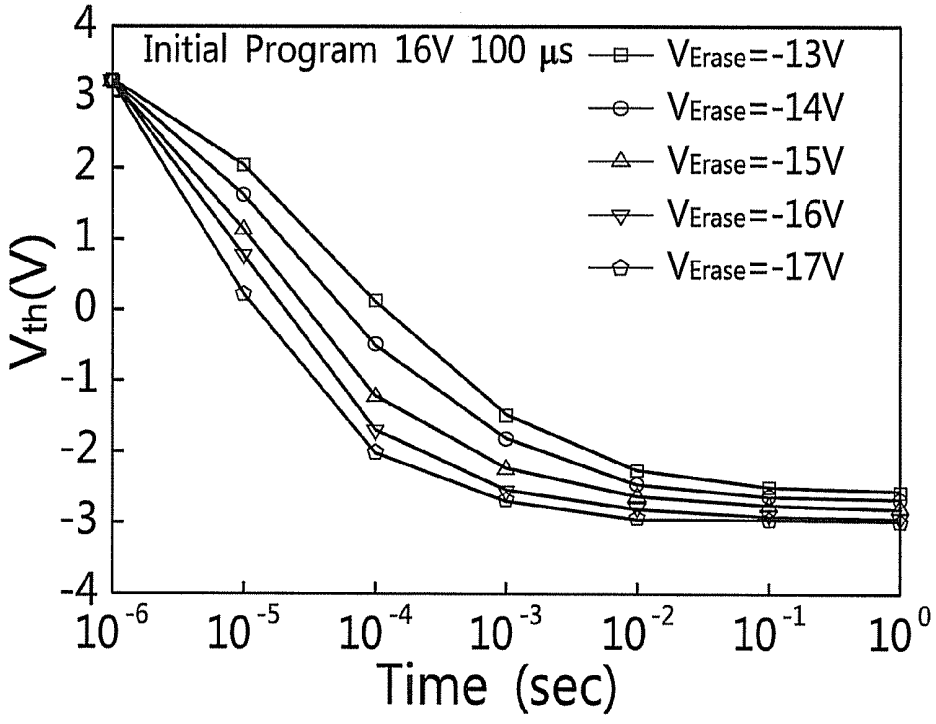

4
3
2
1
0
-1
-2
-3
-4
Vth(V)
Initial Program 16V 100 μs
VErase=-13V
VErase=-14V
VErase=-15V
VErase=-16V
VErase=-17V
10-6
10-5
10-4
10-3
10-2
10-1
100
Time (sec)

FIG. 7

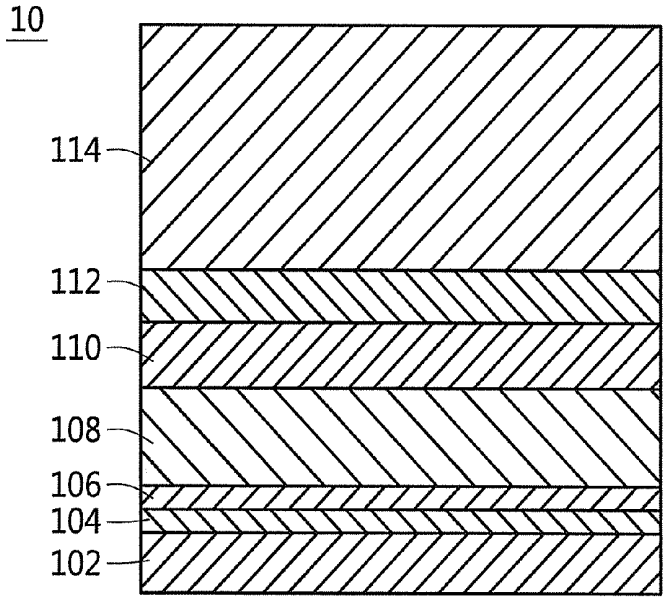

10
114
112
110
108
106
104
102

FIG. 8

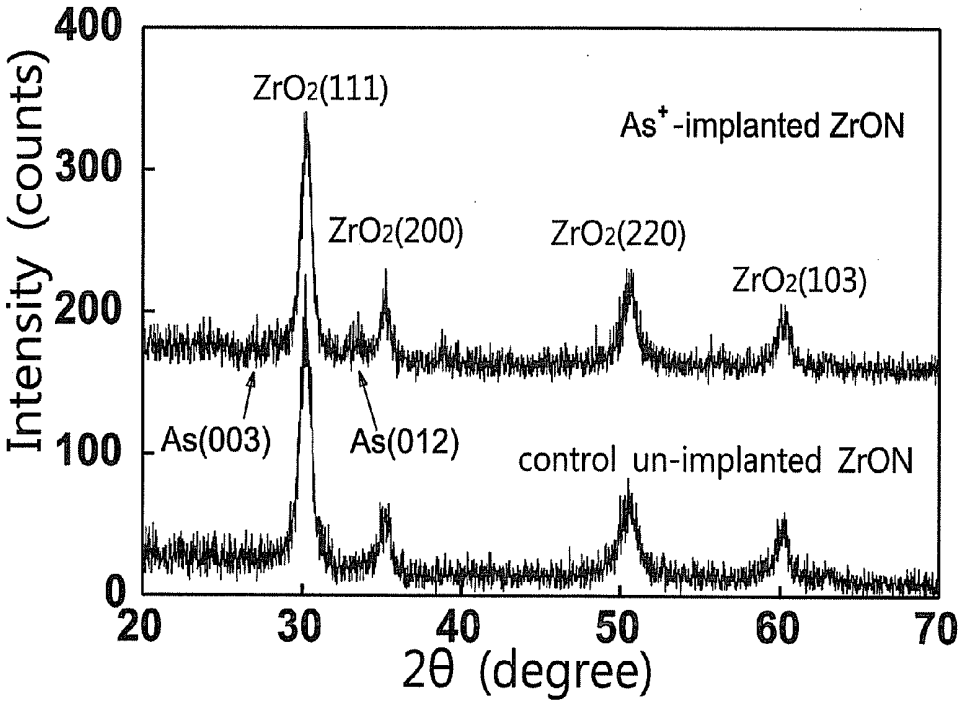

400
300
200
100
0
Intensity (counts)
ZrO2(111)
As+-implanted ZrON
ZrO2(200)
ZrO2(220)
ZrO2(103)
As(003)
As(012)
control un-implanted ZrON
20
30
40
50
60
70
2θ (degree)

FIG. 9

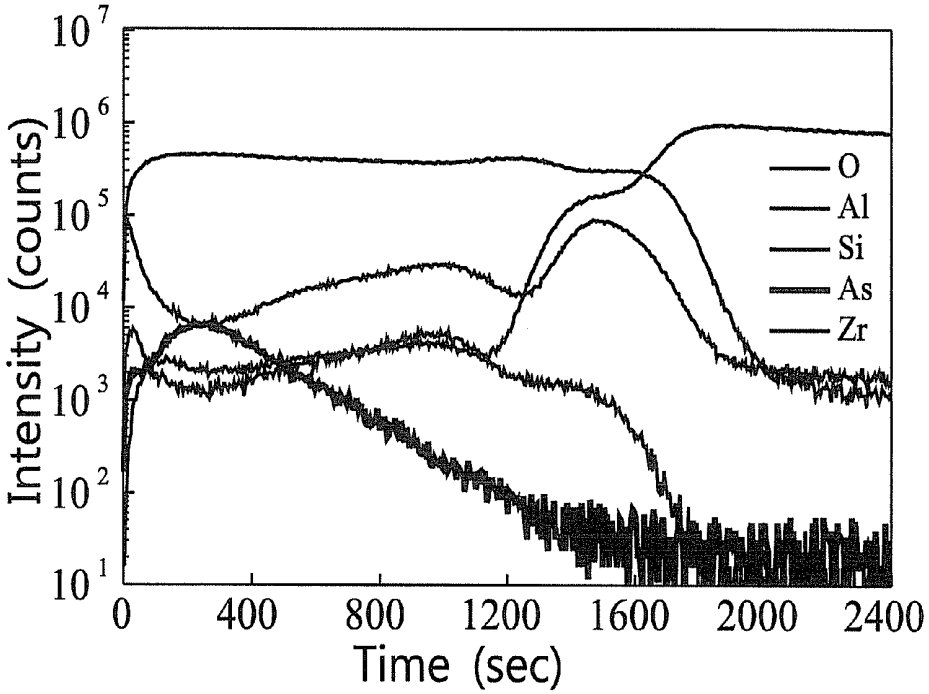

Intensity (counts)
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0
400
800
1200
1600
2000
2400
Time (sec)
O
Al
Si
As
Zr

FIG. 10

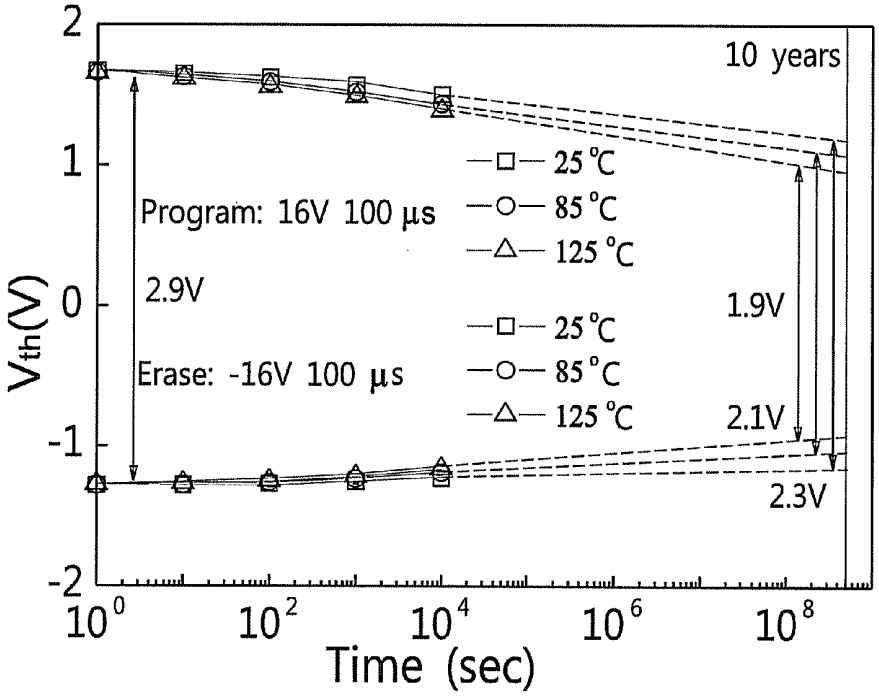

V_{th}(V)
2
1
0
-1
-2
10^0
10^2
10^4
10^6
10^8
Time (sec)
10 years
Program: 16V 100 μs
2.9V
Erase: -16V 100 μs
25 °C
85 °C
125 °C
25 °C
85 °C
125 °C
1.9V
2.1V
2.3V

FIG. 11

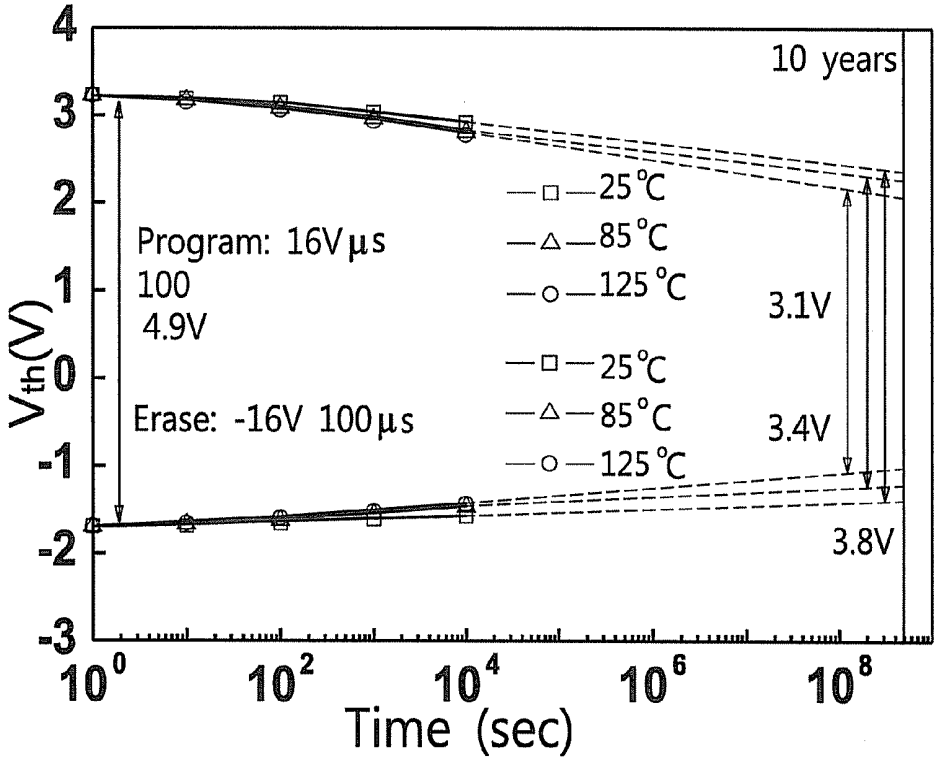

10 years
25°C
85°C
125°C
Program: 16V μs
100
4.9V
3.1V
Erase: -16V 100μs
25°C
85°C
125°C
3.4V
3.8V
Vth(V)
Time (sec)

FIG. 12

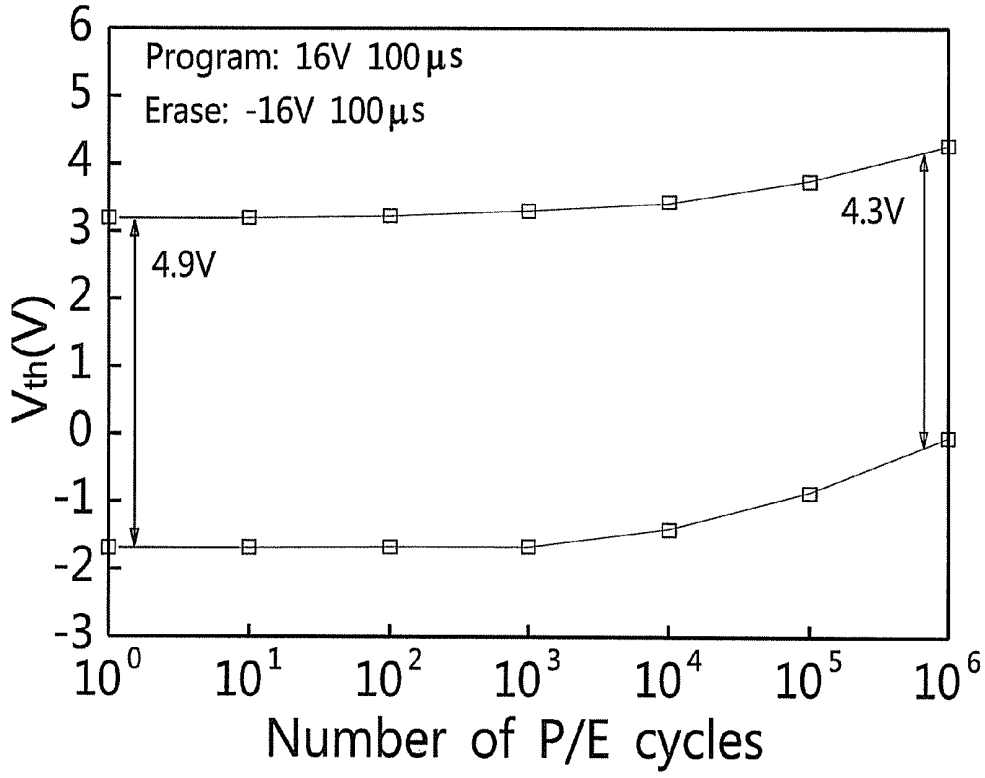
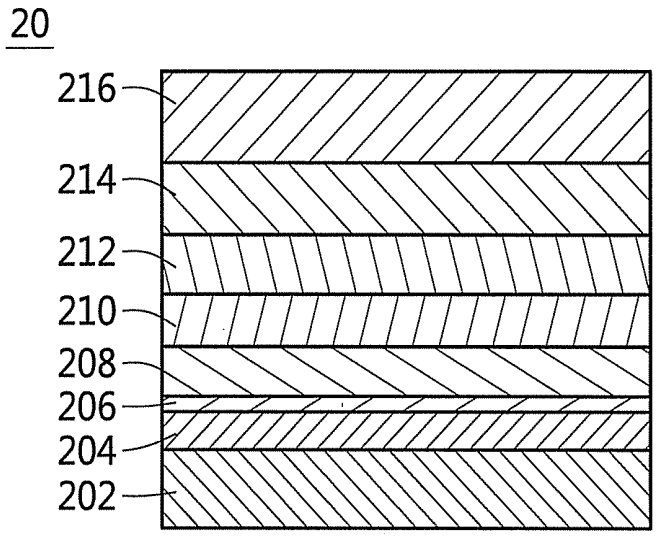
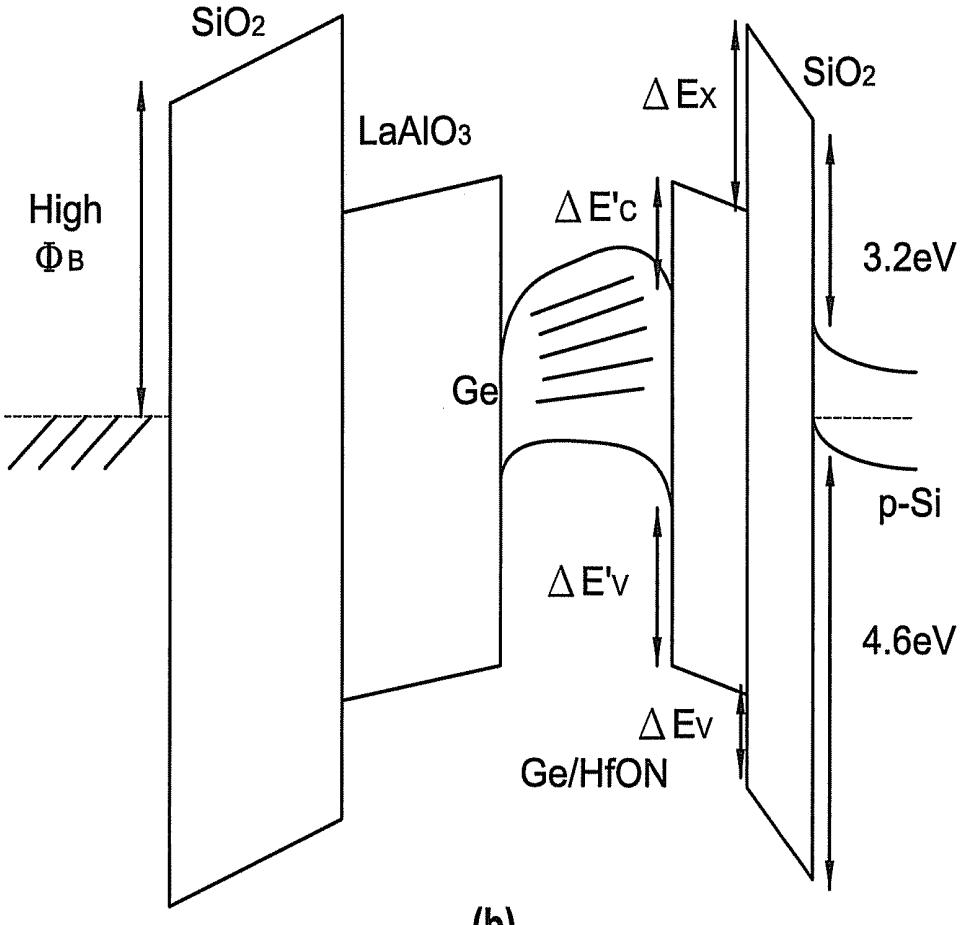

Program: 16V 100μs
Erase: -16V 100μs
4.9V
4.3V
Vth(V)
Number of P/E cycles

FIG. 13

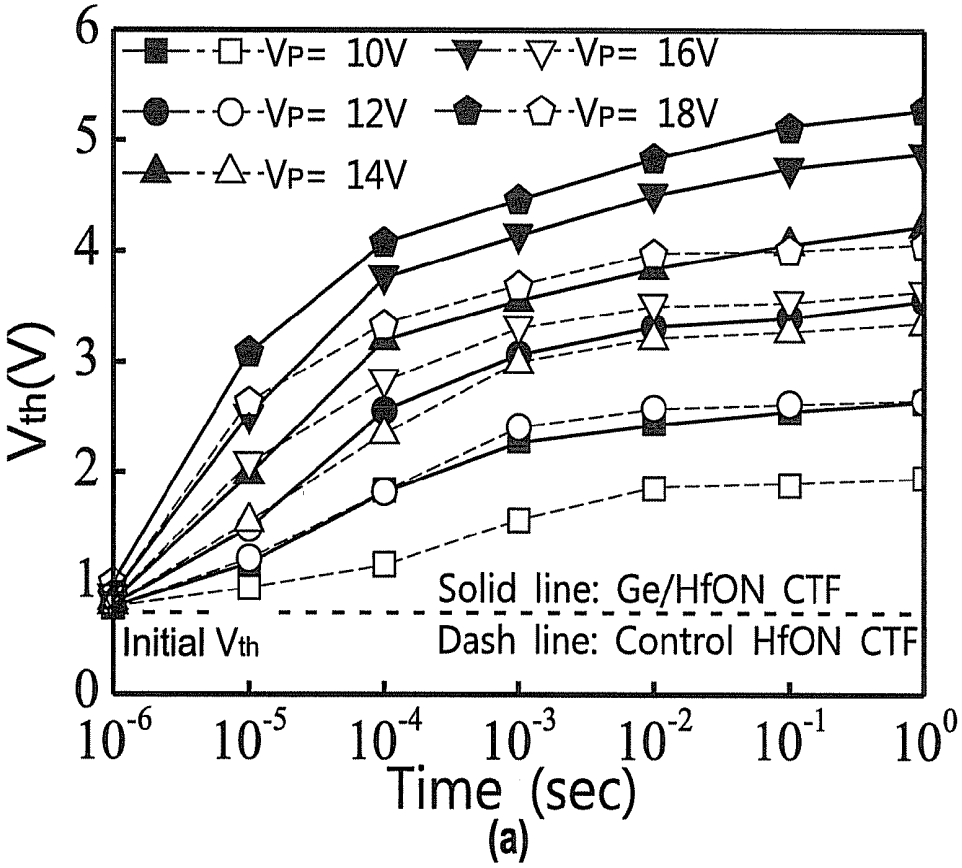

20
216
214
212
210
208
206
204
202

(a)


SiO2
LaAlO3
High
ΦB
Ge
ΔEx
ΔE'c
ΔE'v
ΔEv
Ge/HfON
SiO2
3.2eV
p-Si
4.6eV

(b)

FIG. 14


VP= 10V
VP= 16V
VP= 12V
VP= 18V
VP= 14V
Vth(V)
Solid line: Ge/HfON CTF
Initial Vth
Dash line: Control HfON CTF
Time (sec)
(a)

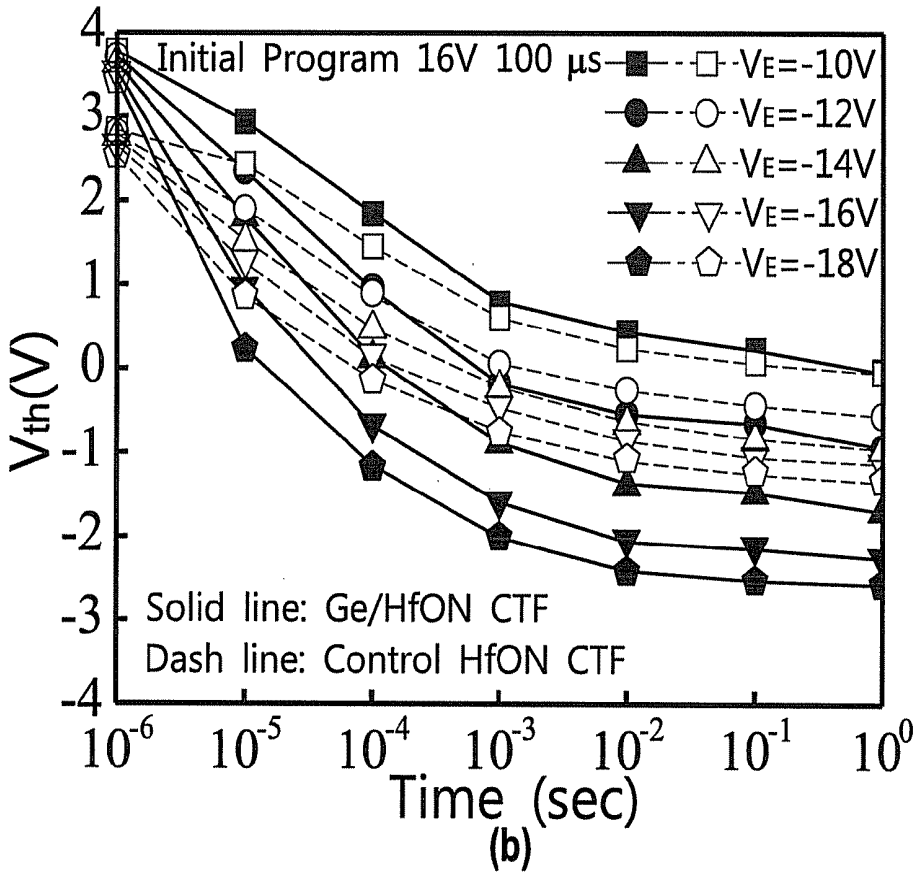
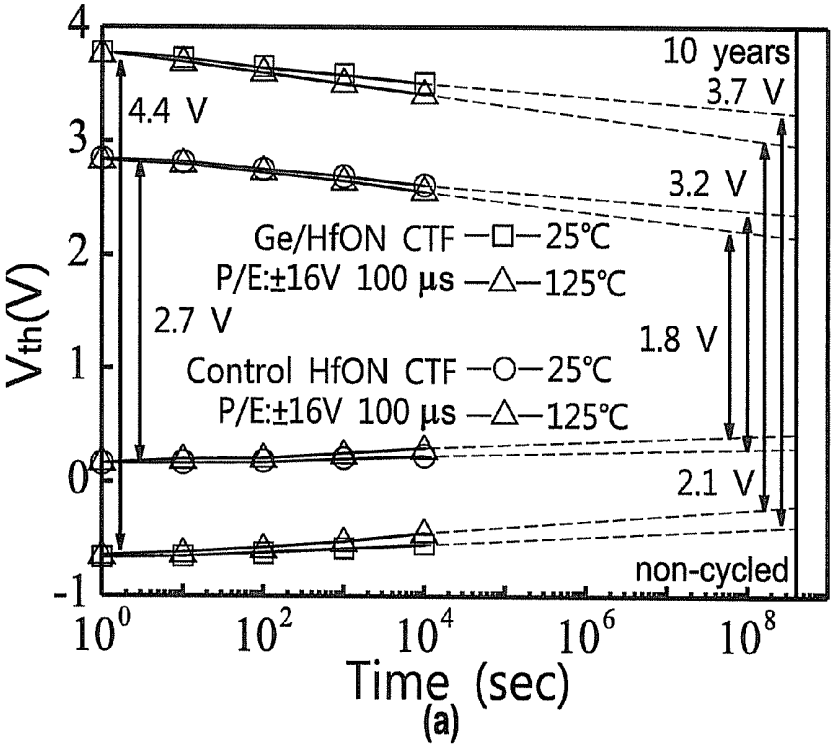

Initial Program 16V 100 μs
VE=-10V
VE=-12V
VE=-14V
VE=-16V
VE=-18V
Vth(V)
Solid line: Ge/HfON CTF
Dash line: Control HfON CTF
Time (sec)
(b)

FIG. 15


10 years
3.7 V
4.4 V
3.2 V
Ge/HfON CTF —□—25°C
P/E:±16V 100 μs —△—125°C
2.7 V
1.8 V
Control HfON CTF —○—25°C
P/E:±16V 100 μs —△—125°C
2.1 V
non-cycled
Vth(V)
Time (sec)
(a)

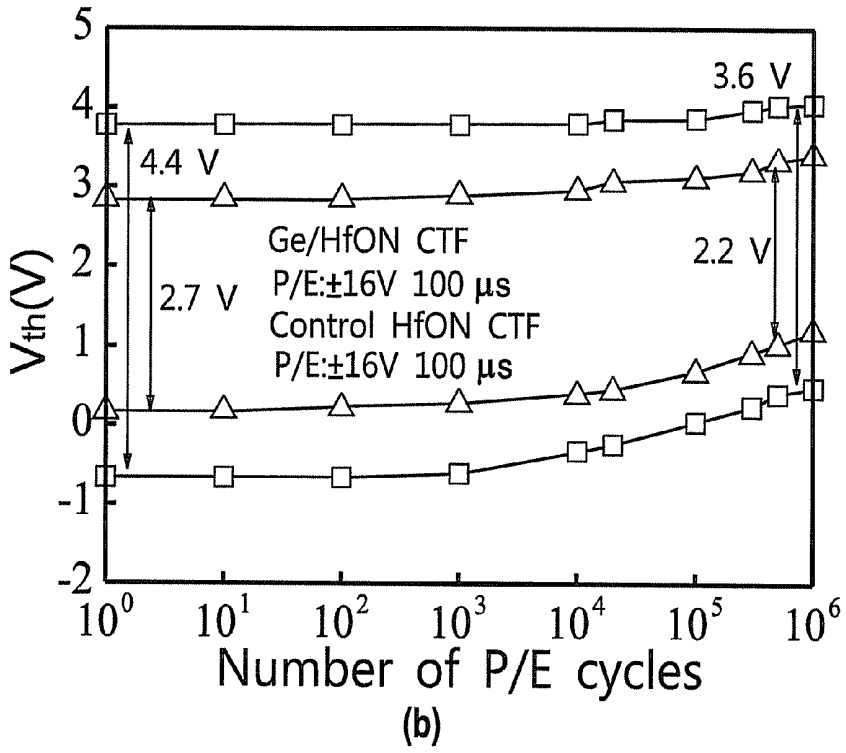

3.6 V
4.4 V
Ge/HfON CTF
P/E:±16V 100 μs
2.2 V
2.7 V
Control HfON CTF
P/E:±16V 100 μs
Vth(V)
Number of P/E cycles
(b)

FIG. 16

FLASH MEMORY

BACKGROUND OF THE INVENTION

[0001] 1. Field of the Invention

[0002] This invention relates to flash memories, and, more particularly, to a flash memory having a charge-trapping layer formed by implanting arsenic into ZrON.

[0003] 2. Description of Related Art

[0004] According to International Technology Roadmap for Semiconductors (ITRS), the degraded endurance and retention are the toughest challenges to further down-scaling the Charge-Trapping Flash (CTF), due to the fewer electrons stored in highly scaled device. On the other hand, scaling down the Si_3N_4 charge-trapping layer to 3-4 nm is needed in ITRS for continuous device scaling, but no proposed solution up to now. However, this worsens the retention and endurance due to the poorer trapping capability at thinner Si_3N_4, where nearly no charge trapping was found in 2 nm Si_3N_4. Although the retention is improved by using a thicker tunnel oxide, this yields reduced erase speed. Such retention and erase-speed trade-off is a basic limitation of CTF.

[0005] Previously we addressed this limitation with a deep trapping energy E_{vac}-E_C Al(Ga)N or HfON in a metal-oxide-nitride-oxide-Si (MONOS) device. The better retention of high-κ Al(Ga)N MONOS CTF was also listed in ITRS. One drawback of desired higher κ HfON is the lower trapping efficiency; thus, the double trapping HfON—Si_3N_4 CTF was used. Yet the scaling equivalent-Si_3N_4-thickness (ENT) is still limited to 7 nm.

SUMMARY OF THE INVENTION

[0006] In view of the above-mentioned problems of the prior art, it is a primary objective of the present invention to provide flash memory that has a charge-trapping layer formed by implanting arsenic into ZrON.

[0007] In an embodiment of the present invention, the flash memory includes: a substrate; a first SiO_2 layer formed on the substrate; a first high-κ layer formed on the first SiO_2 layer; a metal-implanted oxynitride layer formed on the first high-κ layer; a second high-κ layer formed on the metal-implanted oxynitride layer; a second SiO_2 layer formed on the second high-κ layer; and a gate layer formed on the second SiO_2 layer.

[0008] In another embodiment of the present invention, the flash memory includes: a Si substrate; a first SiO_2 layer formed on the Si substrate; a first high-κ layer formed on the first SiO_2 layer; a first HfON layer formed on the first high-κ layer; a first Ge/oxynitride layer formed on the oxynitride layer; a second high-κ layer formed on the Ge/oxynitride layer; a second SiO_2 layer formed on the second high-κ layer; and a gate layer formed on the second SiO_2 layer.

[0009] The present invention further provides a method of fabricating a flash memory, including: providing a substrate; forming a first SiO_2 layer on the substrate; forming a first high-κ layer on the first SiO_2 layer; forming a metal-implanted oxynitride layer on the first high-κ layer; forming a second high-κ layer on the metal-implanted oxynitride layer; forming a second SiO_2 layer on the second high-κ layer; and forming a TaN layer on the second SiO_2 layer.

[0010] In an embodiment, the As-implanted ZrON is formed by implanting ZrON with As at 60°-tilted angle, 3 KeV and 5×10^{15} cm^{-2} dose, and followed by 950° C. RTA.

BRIEF DESCRIPTION OF DRAWINGS

[0011] The invention can be more fully understood by reading the following detailed description of the preferred embodiments, with reference made to the accompanying drawings, wherein:

[0012] FIGS. **1**(*a*) to (*c*) are schematic band diagrams of a traditional MONOS CTF, a double Si_3N_4-HfON trapping CTF, and a As^+-implanted ZrON CTF, respectively;

[0013] FIG. **2** shows J_g-V_g curves of ZrON MONOS non-volatile memory (NVM) devices with and without As^+-implant;

[0014] FIG. **3** shows C-V hysteresis of ZrON MONOS NVM devices with and without As^+-implant;

[0015] FIG. **4** shows program characteristics of control ZrON MONOS NVM devices for different voltages and times;

[0016] FIG. **5** shows erase characteristics of control ZrON MONOS NVM devices for different voltages and times;

[0017] FIG. **6** shows program characteristics of As^+-implanted ZrON MONOS NVM devices for different voltages and times;

[0018] FIG. **7** shows erase characteristics of As^+-implanted ZrON MONOS NVM devices for different voltages and times;

[0019] FIG. **8** is a schematic structure of an As^+-implanted ZrON/$LaAlO_3$/SiO_2/Si of an embodiment according to the present invention;

[0020] FIG. **9** shows XRD of control and As^+-implanted ZrON/$LaAlO_3$/SiO_2/Si structure after 950° C. RTA;

[0021] FIG. **10** shows SIMS of the As^+-implanted ZrON/$LaAlO_3$/SiO_2/Si structure after 950° C. RTA;

[0022] FIG. **11** shows retention characteristics of control un-implanted ZrON MONOS NVM devices at 25-125° C.;

[0023] FIG. **12** shows retention characteristics of As^+-implanted ZrON MONOS NVM devices at 25-125 ° C.;

[0024] FIG. **13** shows endurance characteristics of As^+-implanted ZrON MONOS NVM devices;

[0025] FIGS. **14**(*a*) and (*b*) are a schematic structure and s schematic energy band diagram of a Ge/HfON CTF memory of an embodiment according to the present invention, respectively;

[0026] FIGS. **15**(*a*) and (*b*) show program and erase characteristics of HfON CTF memory with and without Ge for different voltages and times; and

[0027] FIGS. **16**(*a*) and (*b*) show retention characteristics of a flash memory and cycling characteristics of HfON CTF memory with and without Ge.

DETAILED DESCRIPTION OF THE INVENTION

[0028] The following illustrative embodiments are provided to illustrate the disclosure of the present invention, these and other advantages and effects can be apparently understood by those in the art after reading the disclosure of this specification. The present invention can also be performed or applied by other different embodiments. The details of the specification may be on the basis of different points and applications, and numerous modifications and variations can be devised without departing from the spirit of the present invention.

[0029] High performance MONOS CTF with highly scaled 3.6 nm ENT is reached and meets ITRS scaling target for the first time. At 125° C. and ±16 V program/erase (P/E), the device has fast 100 μs speed and large extrapolated 10-year

retention of 3.1 V. The excellent results were reached using metallic Arsenic (As) implant into higher κ ZrON (κ=35) as trapping layer. In contrast, small 10-year retention window of 1.9 V is found in control ZrON CTF. The improved memory window is due to the better electron trapping capability of As^+-implanted ZrON. The excellent 10^6 cycles and good 125° C. retention may be ascribed to the deep E_{vac}-E_C ZrON and 5.1 eV work-function (Φ_m) of metallic As, to minimize the Schottky emission and tunnel leakage. The excellent 10^6 cycles are vitally important to allow further endurance improvement in highly scaled CTF device with fewer electrons. These results compare well with other reported data listed in Table 1, with the smallest 3.6 nm ENT, fast 100 μs speed, large memory window, good retention at 125° C. and the best 10^6 endurance.

TABLE 1

	P/E conditions for retention & cycling	Initial ΔV^{th} (V)	ΔV_{th} (V) for 10-year retention @ 85° C.	ΔV_{th} (V) for 10-year retention @ 125° C.	ΔV_{th} (V) @ Cycles
This work (As^+-implanted)	16 V 100 μs/-16 V 100 μs	4.9	3.4	3.1	4.3 @ 10^6
This work (control un-implanted)	16 V 100 μs/-16 V 100 μs	2.9	2.1	1.9	—
TANOS $SiO_2/Si_3N_4/Al_2O_3/TaN$	13.5 V 100 μs/-13 V 10 ms	4.4	2.07	—	4 @ 10^5
Tri-gate $SiO_2/Si_3N_4/SiO_2$	11.5 V 3 ms/-11.5 V 100 ms	1.2	1.1 (@ 25° C.)	—	1.5 @ 10^4
FinFET $SiO_2/Si_3N_4/SiO_2$	13 V 10 μs/-12 V 1 ms	4.5	2.4	—	3.5 @ 10^4
SiO_2/AlGaN/$AlLaO_3$	11 V 100 μs/-11 V 100 μs	3.0	1.6	—	2.3 @ 10^5

[0030] The TaN-[SiO_2—$LaAlO_3$]—ZrON—[$LaAlO_3$—SiO_2]-Si CTF device has 2.5 nm thermal SiO_2, 2.5 nm $LaAlO_3$, 18 nm $ZrON_{0.2}$, 8 nm $LaAlO_3$, 6 nm LPCVD SiO_2, and 200 nm TaN. The $LaAlO_3$, $ZrON_{0.2}$, and TaN were deposited by physical vapor deposition (PVD). To improve the trapping capability, the ZrON was implanted by As at 60°-tilted angle, 3 KeV and 5×10^{15} cm^{-2} dose, followed by 950° C. RTA to reduce the ion-implanted damage. After gate definition, self-aligned As implant is applied and RTA is used to activate the dopants. The $LaAlO_3$ was from binary Al_2O_3 and La_2O_3, used for V_t tuning in 32 nm gate-first high-κ p- and n-MOSFETs, respectively. For comparison, control CTF device was also fabricated without the As^+-implant into ZrON.

A. P/E Characteristics:

[0031] FIG. **1** shows the traditional, double trapping Si_3N_4-HfON, and As^+-implanted ZrON MONOS CTF devices. The As^+-implanted ZrON has higher κ than HfON and deep Φ_m of As for trapping. FIG. **2** shows the gate current (J_g) of CTF. Close J_g of As^+-implanted device with control is reached. Larger C-V hysteresis of 8.1 V was obtained in As^+-implanted CTF than control device under ±16 V sweep (FIG. **3**). FIGS. **4** and **5** show the program and erase data in control devices. A small ΔV_{th} memory window of 2.9 V was measured at 100 μs at ±16 V P/E that is typical for metal-oxide-nitride trapping MONOS with low trapping efficiency. The ΔV_{th} is significantly larger for As^+-implanted CTF devices shown in FIGS. **6** and **7**, with a large ΔV_{th} window of 4.9 V at ±16 V 100 μs is P/E. This indicates the better trapping efficiency in As^+-implanted ZrON that may be due to metallic As atoms and implant-created defects. Since a 950° C. RTA is applied to lower the implanted defects, the As atoms may play a major role for better trapping. The fast 100 μs P/E speed is due to the existing ΔE_C and ΔE_V in $LaAlO_3$/ SiO_2 for easier tunneling, where the larger physical thickness improves the retention with only 3 nm EOT in tunnel oxide.

B. Characterization of As^+-implanted ZrON:

[0032] The As^+-implanted ZrON has a schematic structure shown in FIG. **8**, which was analyzed by XRD and SIMS shown in FIGS. **9** and **10**. As shown in FIG. **8**, a flash memory **10** of an embodiment according to the present invention comprises a substrate **102**, a SiO_2 layer **104** formed on the substrate **102**, a $LaAlO_3$ layer **106** formed on the SiO_2 layer **104**, an As-implanted ZrON layer **108** formed on the $LaAlO_3$ layer **106**, another $LaAlO_3$ layer **110** formed on the As-implanted ZrON layer **108**, another SiO_2 layer **112** formed on the another $LaAlO_3$ layer **110**, and a TaN layer **114** formed on the another SiO_2 layer **112**. In an embodiment, the SiO_2 layer **104** is 2.5 nm thick; the $LaAlO_3$ layer **106** is 2.5 nm thick; the As-implanted ZrON layer **108** is 18 nm thick; the another $LaAlO_3$ layer **110** is 8 nm thick; the another SiO_2 layer **112** is 6 nm thick; and the TaN layer **114** is 200 nm thick. The ZrON poly-grains are found in X-TEM that gives the higher κ value and smaller ENT. The XRD shows weak As peaks with the same angle of clustered As-dots in As-rich GaAs, which suggests the forming small As metal dots in ZrON although beyond our TEM resolution. Such metallic As with deep 5.1 eV Φ_m inside ZrON traps may also reduce the Schottky emission and tunnel leakage, in addition to the large E_{vac}-E_C of ZrON. From SIMS, the As concentration at 60° 3-keV implant reduces rapidly with thickness and mainly within ZrON that explains the close J_g with control device in FIG. **2**.

C. Retention & Endurance:

[0033] FIGS. **11** and **12** show the retention data of As^+-implanted and control ZrON CTF devices at 25, 85 and 125° C. A large 10-year extrapolated window of 3.1 V is measured at 125° C. in As^+-implanted devices at 100 μs and ±16 V P/E. This is significantly better than the 1.9 V 10-year window in control devices. Such large 10-year retention data allows multi-level cells (MLC) storage even at 125° C. The good retention is due to the extra ΔE_C confinement energy in $LaAlO_3$/ZrON/$LaAlO_3$ and also deep Φ_m of metallic As shown in FIG. **1**(*c*), while fast 100 μs erase is also reached

from the lowered hole energy barrier ΔE_V in $LaAlO_3/SiO_2$ tunnel oxide. The fast P/E speed and lowered hole tunnel barrier ΔE_V lead to excellent endurance: as shown in FIG. **13**, a still large 4.3 V window is obtained even at 10^6 P/E cycles. Such excellent cycling data are vitally important and allow further improving the endurance in highly scaled CTF device with fewer stored electrons. Table 1 compares various MONOS CTF devices. The novel device compares well with other devices, with the record thinnest 3.6 nm ENT trapping layer, large memory window, good 125° C. retention, fast 100 μs P/E speed, and the highest 10^6 endurance.

[0034] Using low energy As^+-implant into higher κ ZrON trapping layer, this novel CTF device shows excellent device performance of highly scaled 3.6 nm ENT, large 10-year extrapolated retention window of 3.1 V at 125° C. and 1 million times endurance, at a fast 100 μs and low ±16V P/E.

[0035] Among various types of NVM, the flash memory has irreplaceable merits of the lowest switching energy and excellent device distribution that are vital for high-density sub-Tb memory arrays. To continue downscale into sub-20-nm, the MONOS CTF devices are proposed to replace the poly-Si floating-gate (FG) flash memory according to ITRS. This is due to the discrete charge-trapping property, simple planar structure, and small cell-cell disturbance that are needed for three-dimensional (3D) flash memory.

[0036] One challenge for CTF is the small confinement energy between Si_3N_4 trapping layer and SiO_2 barrier that degrades the high temperature retention. To address this issue, deep E_C high-κ AlGaN and HfON were used to replace the Si_3N_4, which was listed in ITRS for continuous downscaling. In addition, we improved the charge-trapping efficiency at an ENT of 3.6-nm, by using As^+-implanted high-κ trapping layer to reach a large 10-year retention window of 3.1 V at 125° C. However, further downscaling the ENT is limited by the ion-implanted damage to tunnel oxide.

[0037] In this context, a high performance CTF memory at a record thinnest 2.5-nm ENT trapping layer is disclosed for the first time. This device has a large extrapolated 10-year retention memory window of 3.2 V at 125° C. and excellent endurance of 10^6 cycles, under fast 100 μs and low ±16 V P/E pulses. These were achieved using Ge reaction with HfON trapping layer to form the HfGeON, even at ultra-thin 2.5-nm ENT. Such excellent device integrity is unreachable for conventional Si_3N_4 CTF device due to nearly no trapping at 2-nm Si_3N_4.

EXPERIMENTS

[0038] The MoN—[SiO_2—$LaAlO_3$]—[Ge—HfON]—[$LaAlO_3$—SiO_2]—Si MONOS CTF devices were made on standard 6-in p-type Si wafers. The double tunnel oxide layers of 2.7-nm thick thermal SiO_2 was first grown on Si substrates and followed by depositing 2.5-nm thick $LaAlO_3$ by physical vapor deposition (PVD). Then the charge trapping layers of 8-nm thick HfON and 1.5-nm thick Ge were deposited by PVD. Next, a 6.5-nm thick $LaAlO_3$ was deposited by PVD and 6.5-nm thick SiO_2 was deposited by chemical vapor deposition (CVD) using Tetraethyl orthosilicate ($Si(C_2H_5O)_4$) to form the double blocking layers. Finally, a 200-nm thick MoN was deposited by PVD, followed by gate definition, reactive ion-etching (RIE), self-aligned 25 KeV P^+ implantation at 5×10^{15} cm^{-2} dose and 900° C. RTA to activate the dopant at source-drain region. The CTF devices were with 10-μm gate length, 100-μm width and isolated by field oxide. The fabricated devices were characterized by cross-sectional transmission electron microscopy (TEM), P/E, cycling and retention measurements to 125° C.

[0039] FIGS. **14**(*a*) and (*b*) show the schematic structure and energy band diagram of the MoN[SiO_2—$LaAlO_3$]-[Ge—HfON]—[$LaAlO_3$—SiO_2]—Si CTF device. An ultra-thin ENT of 2.5-nm was obtained from X-TEM. As shown in FIG. **14**(*a*), a flash memory **20** comprises a Si substrate **202**, a SiO_2 layer **204** formed on the Si substrate **202**, a $LaAlO_3$ layer **206** formed on the SiO_2 layer **204**, a HfON layer **208** formed on the $LaAlO_3$ layer **206**, a Ge/HfON layer **210** (that is formed by reacting Ge with HfON) formed on the HfON layer **208**, another $LaAlO_3$ layer **212** formed on the Ge/HfON layer **210**, another SiO_2 layer **214** formed on the another $LaAlO_3$ layer **212**, and a MoN layer **216** formed on the another SiO_2 layer **214**. In an embodiment, the flash memory **20** may further comprise another HfON layer formed on the Ge/HfON layer **210**, and another Ge/HfON layer formed on the another HfON layer. The energy band at top Ge/HfON may be narrowed due to the smaller bandgap of Ge (0.67 eV) and HfGeON formation. The small bandgap HfGeON lowers the E_C that improves the carrier retention. The band offset in $LaAlO_3/SiO_2$ double tunnel layers reduces the tunneling barrier for faster P/E speeds and better endurance. The high-κ blocking and trapping layers lower the P/E voltage, which in turn improves the erase saturation due to the higher electric field across the thin tunnel SiO_2.

[0040] FIGS. **15**(*a*) and (*b*) show the P/E characteristics of HfON and Ge/HfON CTF, respectively. The V_{th} increases with increasing P/E voltage and time. The larger program V_{th} of Ge/HfON CTF indicates the better trapping capability than that of control HfON device. Small erase saturation was obtained due to the using high-κ dielectrics to give a larger electric field in tunnel SiO_2 for better hole tunneling. Under the ±16 V and 100 μs P/E, the V_{th} difference (ΔV_{th}) between P and E are 2.7 V and 4.4 V for HfON and Ge/HfON CTF devices, respectively. The larger ΔV_{th} memory window of Ge/HfON CTF than control HfON device is due to the Ge reaction with HfON to form the HfGeON that has higher-κ for better erase and higher traps for more efficient charge-trapping.

[0041] FIG. **16**(*a*) shows the retention characteristics of HfON and Ge/HfON CTF devices. Under the ±16 V and 100 μs P/E, the 10-year extrapolated retention window at 125° C. for HfON CTF is 1.8 V that largely increases to 3.2 V for Ge/HfON device. Such large 10-year retention window allows multi-level cell (MLC) storage at 125° C. with a record thinnest 2.5-nm ENT. The good retention is due to the deep E_C of HfGeON for carrier storage, like the deep-E_C of poly-Si FG flash memory. In addition, the larger physical thickness of high-κ-$LaAlO_3/SiO_2$ double barriers improves the retention. Further, the large energy bandgap and small trap density of SiO_2, formed by CVD TEOS, in double barrier also helps the retention improvement. The fast 100 μs erase speed is due to bandgap-engineered lower hole barrier ΔE_V in the $LaAlO_3/SiO_2$ tunnel oxide. The better trapping capability, faster P/E speed, and lower hole tunnel barrier further lead to excellent endurance as shown in FIG. **16**(*b*), with a larger 10^6-cycled memory window of 3.6 V than the 2.2 V of control device, at the same ±16 V and 100 μs P/E. Such excellent cycling data are vitally important to allow further endurance improvement in highly scaled CTF device with fewer stored electrons.

[0042] The record thinnest 2.5-nm ENT for CTF, large 10-year retention window of 3.2 V at 125° C., and 10^6 cycled endurance were reached simultaneously under fast 100 μs and low ±16V P/E, which fit the requirements of ITRS shown in Table 2.

TABLE 2

lists a variety of data published by ITRS in 2009.

NAND Flash poly 1/2 Pitch (nm)	16	14	13	12	11	9
Cell size 0 area factor a in multiples of F^2SLC/MLC	4.0/1.0	4.0/1.0	4.0/1.0	4.0/1.0	4.0/1.0	4.0/1.0
Tunnel dielectric material	SiO_2 or ONO	SiO_2 or ONO	SiO_2 or ONO	SiO_2 or ONO	SiO_2 or ONO	SiO_2 or ONO
Tunnel dielectric thickness EOT (nm)	3-4	3-4	3-4	3-4	3-4	3-4
Blocking dielectric material	Al_2O_3	Al_2O_3	Al_2O_3	Al_2O_3	Al_2O_3	Al_2O_3
Blocking dielectric thickness EOT (nm)	6	5	5	5	5	5
Charge trapping layer material	SiN/High-K	SiN/High-K	SiN/High-K	SiN/High-K	SiN/High-K	SiN/High-K
Charge trapping layer thickness (nm)	4-6	4-6	4-6	4-6	3-4	3-4
Gate material	Metal	Metal	Metal	Metal	Metal	Metal
Highest W/E voltage (V)	15-17	15-17	15-17	15-17	15-17	15-17
Endurance (erase/write cycles)	1.00E + 04	1.00E + 04	1.00E + 04	1.00E + 04	1.00E + 04	1.00E + 04
Nonvolatile data retention (years)	5-10	5-10	5-10	5-10	5-10	5-10
Maximum number of bits per cell (MLC)	4	4	4	4	4	4

[0043] The foregoing descriptions of the detailed embodiments are only illustrated to disclose the features and functions of the present invention and not restrictive of the scope of the present invention. It should be understood to those in the art that all modifications and variations according to the spirit and principle in the disclosure of the present invention should fall within the scope of the appended claims.

1-4. (canceled)

5. A method, comprising:

providing a substrate;

forming a first SiO_2 layer on the substrate;

forming a first high-κ layer on the first SiO_2 layer;

forming a metal-implanted oxynitride layer on the first high-κ layer;

forming a second high-κ layer on the metal-implanted oxynitride layer;

forming a second SiO_2 layer on the second high-κ layer; and

forming a gate layer on the second SiO_2 layer.

6. The method of claim **5**, wherein the high-κ layer is Al_2O_3, La_2O_3, HfO_2, ZrO_2, TiO_2, SiN or ternary or quarternary combinations thereof, the oxynitride is AlON, LaON, HfON, ZrON, TiON, SiON or quarternary combinations thereof, and the implanted metal are heavy metals of As, Sb, Ga, or In.

7. The method of claim **5**, wherein the metal-implanted oxynitride is formed by implanting metal at tilted angle and <10 KeV.

8. The method of claim **5**, wherein the gate layer is made of metal or metal-nitride.

9. The method of claim **8**, wherein the metal-nitride is TiN, TaN or MoN.

* * * * *