A 60nm NOR Flash Memory Cell Technology Utilizing Back Bias Assisted Band-to-Band Tunneling Induced Hot-Electron Injection (B4-Flash)

Shoji Shukuri, Natsuo Ajika, Masaaki Mihara, Kazuo Kobayashi, Tetsuo Endoh* and Moriyoshi Nakashima
GENUSION, Inc., 7-1-3 Douicho, Amagasaki, Hyogo, Japan
*Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Japan
Phone:+81-6-6416-6133,Fax:+81-6-6416-6134,e-mail:shukuri.shoji@genusion.co.jp

Abstract

A p-channel SONOS flash memory cell technology, which provides excellent scalability and high programming efficiency for NOR architecture, has been developed. The cells named B4-Flash utilizing novel Back Bias assisted Band-to-Band tunneling induced hot-electron (B4-HE) injection is proposed. By applying a moderate back bias to the cell during programming, the bit-line voltage can be reduced below the supply voltage, 1.8V. Resulting that the B4-Flash, applicable to NOR architecture, achieves the gate length of 60nm, for the first time. Basic operation of the 50nm B4-Flash cell is also confirmed. Proposed B4-HE injection scheme realizes not only extreme scalability but also high programming efficiency for NOR type cell.

Introduction

The most crucial limitation in scaling of NOR flash memory cells utilizing CHE injection programming is gate length shortening. This is mainly attributed to lowering of drain to source punch-through immunity below the drain voltage required for CHE. So that the physical limit of the gate length is said to be around 130nm(ITRS 2004, [1]).

Another issue is its low programming efficiency in comparison with those of data storage memory such as NAND and AND, which restricts the program throughput of NOR flash comparatively low.

In this paper, we propose a novel programming scheme using p-channel memory device, in which BTBT-HE injection with an adequate back bias is applied for programming. BTBT-HE programming for p-channel device was reported previously[2]. However, the programming drain voltage was still high around -6V to obtain required HE injection efficiency, and this may limit the gate length scalability, same as conventional NOR.

B4-HE programming scheme provides not only low drain voltage programmability below 1.8V but also high programming efficiency comparable to NAND, which enables to breakthrough the limit of the gate length scaling and program throughput of NOR flash.

Cell Structure and Programming Mechanism

Fig.1 shows the B4-Flash cell structure and its TEM micrograph. Gate dielectrics consist of 6nm top oxide, 8nm nitride and 3nm bottom oxide. Minimum gate length of the evaluated cell is 50nm. The 60nm B4-Flash cells have been extensively investigated.

Gate length(Lg) dependencies of the drain-to-source breakdown voltage(BVds) and the initial Vth are shown in Fig.2. When a back bias of 4V is applied, BVds decreases in 4V from the junction breakdown voltage(-6V) in longer Lg region(>200nm). In the region of Lg<200nm, however, BVds lowering by punch-through is significantly improved even at 50nm because of the back bias application.

Table 1 shows operation voltage conditions for B4-Flash cells. At the standby state, 4 terminals of the cells, those are common source(Vs), bit-lines(Vd), word-lines(Vg) and n-well(Vb), are pre-charged to 1.8V(Vcc). At the first stage of the program operation, back bias is increased and then the selected word-line voltage is raised, finally only the selected bit-line is grounded during programming.

The back bias assisted BTBT-HE generation model is shown in Fig.3. The generation of B4-HEs consists of two steps, (1)BTBT generation is controlled by the vertical electric field(Vg – Vd), and (2)the generated BTBT-electrons are accelerated in the depletion layer controlled by the junction electric field(Vd – Vb), as shown in Fig.3(a) and (b). At the source side, both the vertical and junction electric field is relaxed by applying the source voltage of 1.8V to inhibit the programming, as shown in Fig.3(c). In this programming operation, the voltage difference between Vd and Vs can be minimized and the gate length scalability is extremely improved resultingly. In the case of Vb=0V, drain voltage of about -4V is necessary to generate HE having enough energy to overcome the oxide barrier height. Therefore, the minimum gate length for the BTBT-HE without back bias is >100nm, as shown in Fig.2.

Erasure is carried out by Fowler-Nordheim(FN) tunneling hole injection.

Programming, Erase and Disturb Characteristics

Fig.5 shows the programming and disturb characteristics under the voltage conditions indicated in Fig.4. Typical programming time is about 1μs, and both disturb margins of drain and gate are about 3 orders of magnitudes, which shows the applicability of B4-HE programming to NOR architecture

Programming drain current is shown in Fig.6. After word-line charging(step-1), the drain leakage current increases with decreasing drain voltage(step-2). The maximum drain current, however, is below 1nA/cell. Assuming 1kB word-line simultaneous programming (page program) of 128kB B4-Flash cell array (conventional erase block size of NOR), the estimated total programming current is about 30μA. As the actual programming time is estimated to be less than 10μs including verify operation, the programming throughput of 100MB/s can be achieved by utilizing 1kB page programming.

For embedded applications, the high programming efficiency of B4-Flash may help to achieve high area efficiency of the memory allay due to the reduction of on-chip power supply capability.

Estimated read disturb lifetime is shown in Fig.7. High immunity to Vth lowering due to soft-writing can be seen even at 1.8V reading operation. Erase time of about 100ms is obtained by FN tunneling hole injection, as shown in Fig.8.

Cycling Endurance and Data Retention

Fig.9 shows cycling endurance characteristics, where Vth window narrowing due to fixed charge trapping is well controlled. Fig.10 predicts data retention lifetime of 10k cycled cells and shows that the maximum allowable temperature for 10 years retention is around 125C.

Conclusions

B4-Flash cell technology utilizing B4-HE injection for programming, is proposed and successfully demonstrated by using 60nm SONOS cells. B4-Flash realizes the high speed programming comparable to conventional NOR, high programming efficiency and excellent scalability comparable to NAND flash at the same time. Moreover, this programming technique, which can be easily evolved to a floating gate cell, has a good potential for achieving high programming throughput of 100MB/s and pushing forward NOR cell scaling beyond 50nm era.

Acknowledgments

The authors would like to thank Prof. K.Taniguchi, Osaka University, for his support.

References

[1] G.Servalli, et al., IEDM Tech. Dig., 35_1, 2005.
[2] T.Ohnakado, et al., IEEE Trans. EL, Vol.46, No.9, 1999, pp.1866-1870.

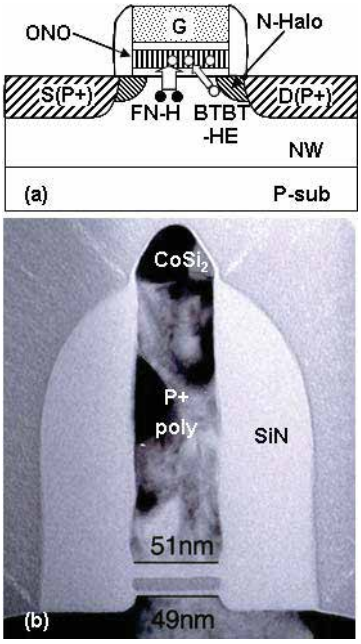

Fig.1 (a)Schematic B4-Flash cell structure and (b)its TEM micrograph

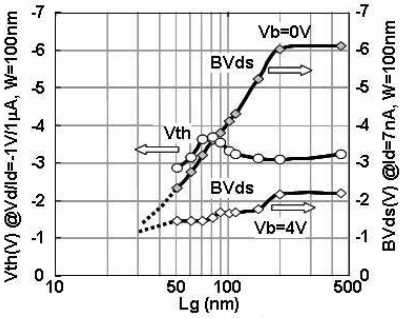

Fig.2 Lg dependencies of initial Vth and BVds measured at Vb=0 and 4V

Mode	Vg	Vd	Vs	Vb
Program	12	1.8⇒0	1.8	4.2
Erase	-12	1.8	1.8	1.8
Read	-2	1.8⇒0	1.8	1.8
Stand-by	1.8	1.8	1.8	1.8

Table 1 Voltage conditions of the selected cell (unit:V)

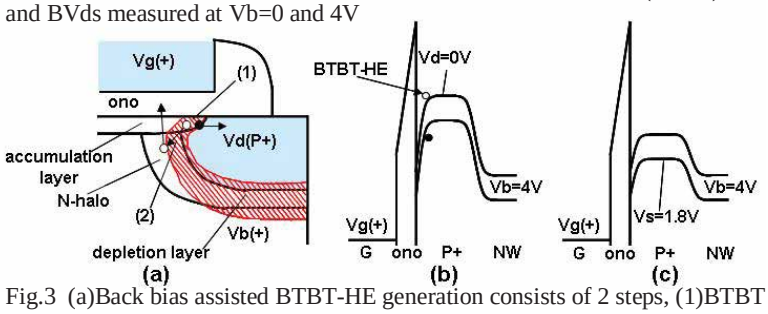

Fig.3 (a)Back bias assisted BTBT-HE generation consists of 2 steps, (1)BTBT generation, (2)electron acceleration, (b)band diagram at drain edge, BTBT-HE is generated, and (c)band diagram at source edge, BTBT is inhibited by Vs=1.8V

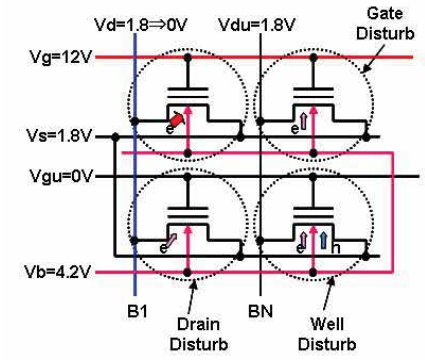

Fig.4 Programming and disturb conditions in NOR array

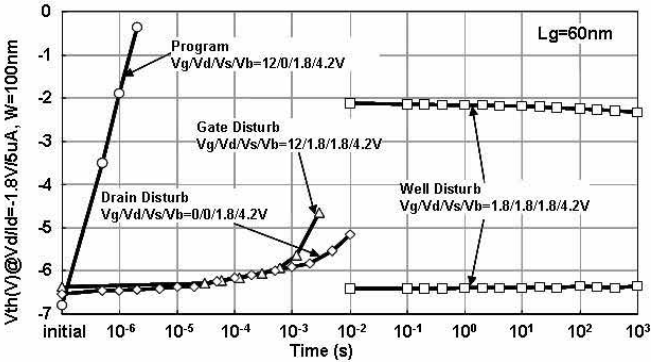

Fig.5 Programming, drain, gate and well disturb characteristics

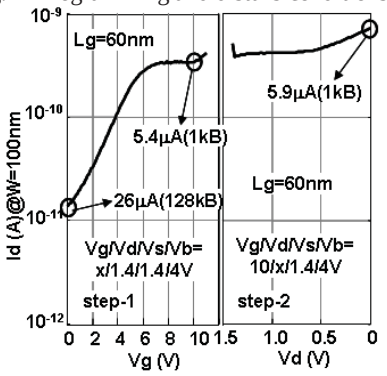

Fig.6 Programming drain current

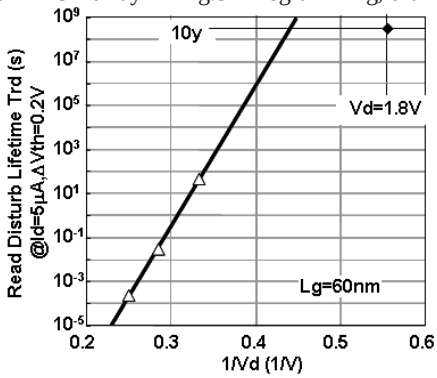

Fig.7 Read disturb lifetime prediction

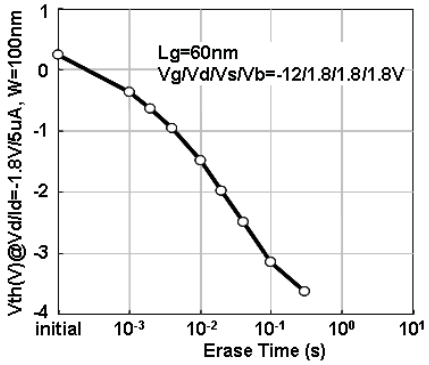

Fig.8 FN hole injection erase characteristics

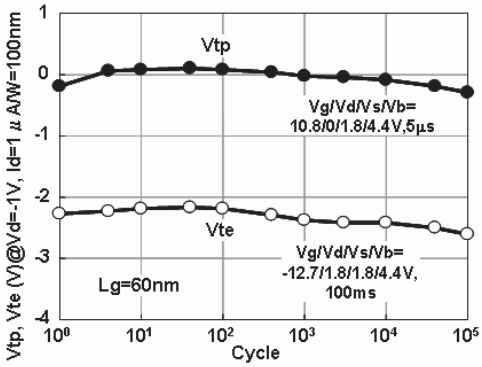

Fig.9 Cycling endurance characteristics

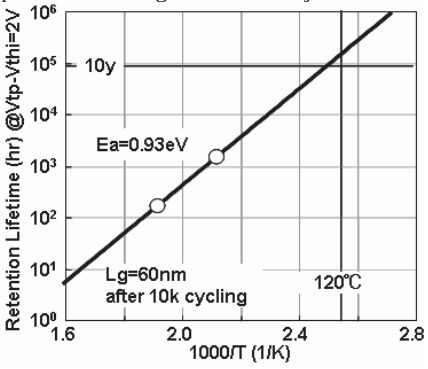

Fig.10 Retention lifetime estimation