Highly Extendible Memory Cell Architecture for Reliable Data Retention Time for 0.10μm Technology Node and beyond

*Jaegoo Lee, Changhyun Cho, Juyong Lee, Sooho Shin, Jinwoo Lee, Donghwa Kwak,
Kyuhyun Lee, Byunghyug Roh, Taeyoung Chung, and Kinam Kim
Advanced Technology Development Team, Memory Division, Samsung Electronics Co.,
San #24, Nongseo-Lee, Kiheung-Eup, Yongin-City, Kyungki-Do, Korea
Tel) 82-31-209-4741, Fax)82-31-209-3274, E-mail)amusante@samsung.co.kr

Abstract

*In this paper, data retention time has been investigated for the high speed and low power 512Mb DRAM (**D**ynamic **R**andom **A**ccess **M**emory) with 0.10μm design rule. As the technology generation of DRAM has been developed into sub-quarter micron regime, the control of junction leakage current at storage node is much more important due to the increased channel doping concentration. In order to obtain high performance DRAM with design rule 0.10μm, novel SAC (**S**elf **A**ligned **C**ontact) process using SRP (**S**AC spacer **R**emoval after **P**lug implantation) is developed to improve data retention time characteristic and minimize short channel effect in cell transistor. We also tried to cure the surface defect and minimize junction leakage current using gate dual spacer and DS (**D**own **S**tream) surface cleaning process. The high capacitance is realized by DMO (**D**ual **M**olded **O**xide) capacitor process. This novel storage node structure gives much better mechanical stability of capacitor. With novel cell architecture, the dramatic increase of data retention time and device yield for 512Mb DRAM can be obtained. The developed cell architecture can be fairly extendible to the future high density DRAM beyond 0.10μm technology node.*

1. Introduction

As the DRAM density enters into the giga-bit era, one of the important characteristics of DRAM is the data retention time. It has been reported that the data retention time needs to be doubled as the memory cell density increases 4 times in every generation due to the requirement of high speed and low power consumption [1]. The data retention time is a period for weak cell to lose the stored data due to the leakage currents before refreshing it. It has been reported that the major failures in the refresh operation of DRAM cells are due to leakage currents of storage node [2]. The leakage currents at cell node which determine DRAM data retention time consist of cell transistor sub-threshold leakage current, cell capacitor dielectric leakage current and storage node junction leakage current from locally enhanced electric field and surface damage layer [3]. It is well known that the cell junction leakage current is caused by TFE (**T**hermionic **F**ield **E**mission) in depletion region and is strongly affected by defect density and its distribution. Therefore, in order to improve data retention time, it is indispensable to reduce not only junction defects but also surface damage and defect [4]. This paper describes novel cell architecture for 512Mb DRAM with 0.10μm design rule. The key technologies are gate dual spacer (SiO_2+Si_3N_4) instead of single Si_3N_4 spacer, down stream surface cleaning process, SRP process and DMO (BPSG+PE-TEOS) capacitor formation.

2. Highly Extendible Cell Architecture

The key processes of novel cell architecture are described in Fig.1 and Fig.2. The cell is fabricated with aforementioned key technologies such as gate dual spacer, down stream surface cleaning process, SRP process and COB (**C**apacitor-**O**ver-**B**it line) structure with DMO capacitor using 0.10μm technology.

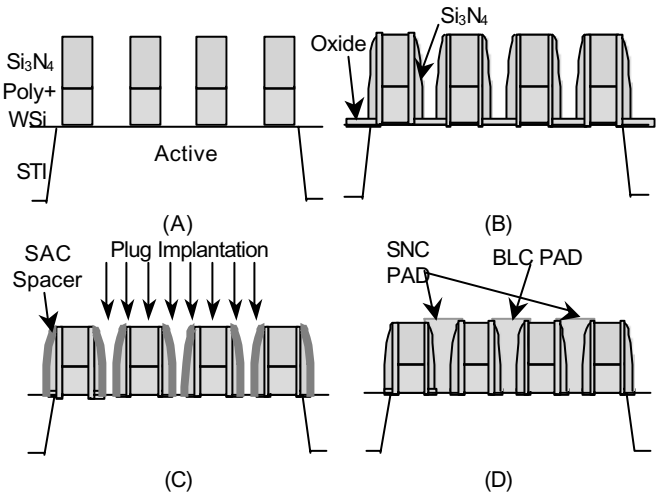

Figure 1 Gate dual spacer, SRP process and down stream process sequences

The gate dual spacer, SRP process and down stream process sequences are as follows: After the formation of STI (**S**hallow **T**rench **I**solation) with 0.25um depth is proceeded, the channel implantation, 55 Å thin oxide growth on silicon surface as gate oxide are executed and W (Tungsten)-polycide is deposited and etched for gate formation as shown in Fig.1-(A). Gate dual spacer is formed by oxide and Si_3N_4 layers which are deposited and etched as shown in Fig.1-(B). Cell contact holes are defined and etched. SAC spacer with 100 Å dielectric layer is formed at inside of contact hole followed by plug implantation in cell array as shown Fig.1-(C). SAC spacer is removed using down stream process and then cell contact pads which are BLC (**B**it **L**ine **C**ontact) pad and SNC (**S**torage **N**ode **C**ontact) pad are formed in cell array as shown in Fig.1-(D). We also used DMO capacitor for increasing storage capacitance. The DMO capacitor process sequences are as follows: Two oxide layers (BPSG+PE-TEOS) which have different wet etching rate are deposited on Si_3N_4 layer, which is etch stopping layer during storage node etch process as shown in Fig.2-(A). Two oxide layers are etched by anisotropic etching as shown in Fig.2-(B). The bottom area of storage node is enlarged during pre-cleaning process due to different wet etching rate as shown in Fig.2-(C). 400 Å thick poly-Si was deposited to form OCS (**O**ne **C**ylindrical **S**torage node) capacitor as shown in Fig.2-(D).

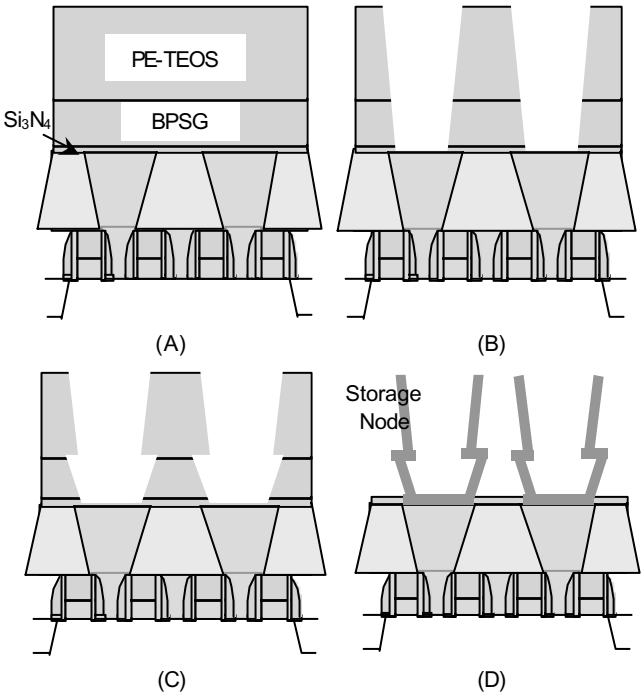

Figure 2 DMO capacitor process sequences

3. Results and Discussion

We developed gate dual spacer and down stream processes to reduce surface damage and leakage current after cell contact hole etch. Figure 3 shows cross-section TEM image of gate with gate dual spacer and cell contact node. Gate dual spacer, composed of 60 Å-SiO_2 and 300 Å-Si_3N_4, can minimize the plasma damage on Si surface during gate spacer etch by etch-stopping at oxide layer.

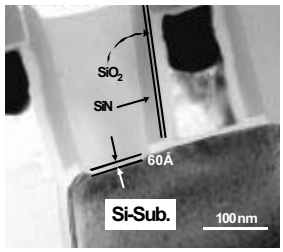

Figure 3 TEM image showing cell transistor using gate dual spacer process

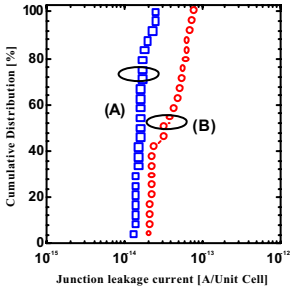

Figure 4 Cell junction leakage current for (A) gate dual spacer process and (B) single gate spacer process.

The merit of gate dual spacer process is less silicon consumption in the cell array. For the conventional process, severe silicon recess occurs because the silicon surface is exposed twice to etching environments during Si_3N_4 spacer formation and SAC etching processes.

However, for gate dual spacer process, the silicon surface is exposed only once to the etching environment of the SAC etching process. Therefore, gate dual spacer process decreases cell junction leakage due to less silicon surface damage by plasma etching. As a result, the cell junction leakage current is reduced to one third of that of the single nitride spacer process as shown in Fig.4.

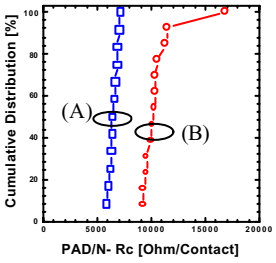

Figure 5 Cell contact resistance between the pad and active for (A) DS process and (B) no DS process.

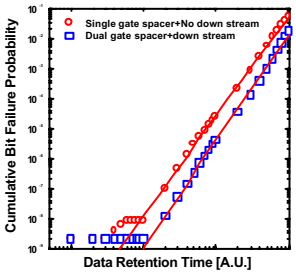

Figure 6 Cumulative fail bit probability curve for single and gate dual spacer processes.

As the technology generation of DRAM has been developed into 0.10μm regime, the storage node contact resistance of the interface between the pad and active region has been increased due to excessive polymer generation during high selective contact hole etch. In order to get rid of the excessive polymer which cannot be easily removed by using conventional Si-treatment method, we used the down stream surface cleaning process which is found to be very effective in remove polymer while minimizing plasma damage. The etching chemistry of down stream process is a mixture $C_4F_8/O_2/Ar$. As a result, we can reduce the contact resistance between the pad and active region more than 30% as shown in Fig.5. Since the junction leakage current and contact resistance at storage node are dominant factor for data retention time, data retention time characteristic of novel process using gate dual spacer and down stream process are much better than that of conventional process as shown in Fig.6.

As the technology generation of DRAM has been scaled down into 0.10μm regime, it is very difficult to keep the effective channel length of cell transistor to be free from short channel effect. It is mainly due to the difficulty to form shallow junction of memory cell transistor whose primary reason is high thermal budget of DRAM integration process. In order to suppress the short channel effect of cell transistor, we developed SRP process where heavily doped source and drain junctions can be effectively away from the gate as much as twice of SAC spacer thickness(10nm). As a result, we can increase effective channel length of memory cell transistor. The enlarged effective channel length of memory cell transistor can allow lower channel implantation dose while maintaining the same threshold voltage. By reducing cell channel implantation dose, junction electric field can be reduced. The decrease of electric field strength reduces cell junction leakage current and thus it improves DRAM data retention time at tail distribution region as shown in Fig.7.

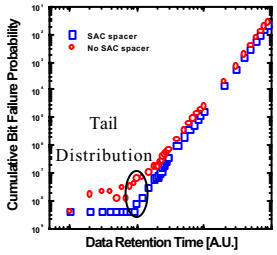

Figure 7 Cumulative fail bit probability curve for SRP and conventional process

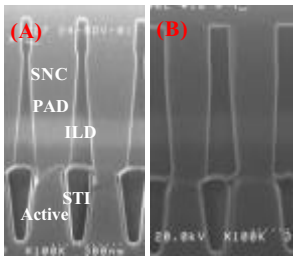

Figure 8 The vertical profile of cell landing pad after the formation SNC pads (A) Conventional process (B) SRP process

The other problem of conventional SAC process is rapid wet etching rate of ILD (**I**nter-**L**ayer **D**ielectric) whose material is BPSG with high concentration of P and B. The loss of ILD due to excessive side consumption during the pre-cleaning process before the deposition of poly-silicon may result in pad bridges between contact pads as shown in Fig.8-(A). However, SRP process can overcome this problem using SAC spacer as a blocking layer around ILD layer during the pre-cleaning process as shown in Fig.8-(B).

As the COB stack DRAM cell is scaled down below sub-quarter micron technology, mechanical stability of COB stack cell becomes important and it limits the stack height of the memory cell capacitor [5]. Therefore, the height of storage node structure with good stability is the key factor to increase the cell capacitance. However, the mechanical stability of storage node rapidly decreases as cell size shrinks. Therefore, we developed

DMO capacitor process to overcome this problem. DMO capacitor has better mechanical stability compared to the conventional capacitor, because it has wider area at the bottom region of storage node. The DMO capacitor process used two different oxide layers that have different wet etching rates. Therefore, the bottom area of DMO storage node is more enlarged than that of the conventional capacitor. Therefore, the structure of DMO capacitor can achieve the enlarged bottom area without losing its mechanical stability. The bottom area and stack height of DMO capacitor increase more than 20% compared with those of conventional cell capacitor as shown in Fig.9. The structure of DMO capacitor is expected to be better structure to prevent the storage node leaning bridges because its enlarged bottom area is contributed to overcome the deformation stress between storage node and lower layer.

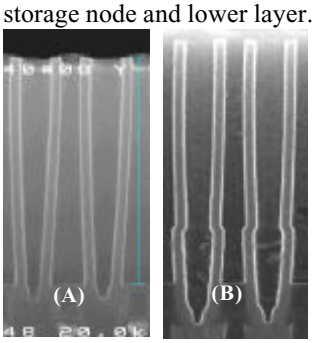

Figure 9 Vertical profiles of cell capacitor (a) conventional capacitor (b) dual molded oxide capacitor

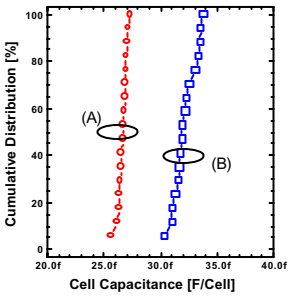

Figure 10 Electrical characteristic of cell capacitor (a) conventional capacitor (b) DMO capacitor

As a result, more than 32fF capacitance per cell can be obtained as shown in Fig.10, where the capacitor dielectric, Ta_2O_5 (tantalum oxide), has the equivalent oxide thickness of 35Å.

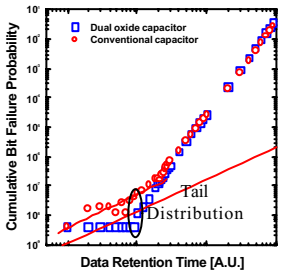

Figure 11 Cumulative fail bit probability curve for DMO capacitor and conventional capacitor.

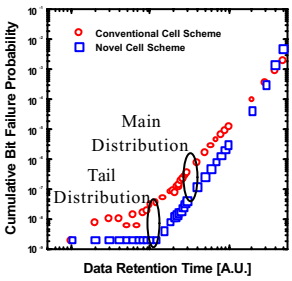

Figure 12 Cumulative fail bit probability curve for novel and conventional cell architecture.

The data retention time characteristic of DMO capacitor is superior to that of conventional one, especially in the tail distribution region shown in Fig.11. With the help of previous key technologies, we can achieve data retention time at least 3 times longer than the conventional technology as shown in Fig.12.

4. Conclusion

In this study, novel 0.10 μm DRAM technology is developed with several key technologies such as gate dual spacer, down stream, SRP processes and DMO capacitor technology. Since it was found that the main components of leakage current were originated from cell junction region and silicon surface, data retention time is significantly improved by the reducing junction leakage current and the surface defect using gate dual spacer, down stream and SRP processes. In addition, a novel DMO capacitor scheme can provide more higher stack height than conventional one, and thereby giving higher capacitance which greatly increases the data retention time. These technologies are expected to be easily extendible beyond 0.10μm technology node and beyond.

Reference

[1] T. Hamamoto et al., "Well concentration: A novel scaling limitation factor derived from DRAM retention time and its modeling.", *IEDM Tech. Dig.,* pp.915-918 (1995)

[2] Hyung Soo Uh et al., "A Strategy for Long Data Retention Time of 512Mb DRAM with 0.12um Design Rule:", *Symp. On VLSI Tech*., pp. 27–28 (2001)

[3] Jooyoung Lee et al, "Novel Cell Transistor Using Retracted Si_3N_4-Liner STI for the Improvement of Data Retention Time in Gigabit Density DRAM and Beyond.", IEEE Trans. Electron Devices, vol.48, pp1152-1158, (2001)

[4] Jaegoo Lee et al., "A Novel DRAM Technology using Dual Spacer and mechanically Robust Capacitor for 0.12μm DRAM and beyond.", *European Solid State Device Research Conference* pp.127-130 (2001)

[5] Yongjik Park et al., "COB Stack Cell Technology beyond 100nm Technology Node:", *IEDM Tech. Dig.,* pp.391-394 (2001)