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

Advanced Technology Development Team, Memory Division, Samsung Electronics Co., San #24, Nongseo-Lee, Kihueung-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 (Dynamic Random Access Memory) 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 (Self Aligned Contact) process using SRP (SAC spacer Removal after Plug 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 (Down Stream) surface cleaning process. The high capacitance is realized by DMO (Dual Molded Oxide) 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 (Thermionic Field Emission) 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 (SiO2+Si3N4) instead of single Si3N4 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 (Capacitor-Over-Bit line) structure with DMO capacitor using 0.10µm technology.
The gate dual spacer, SRP process and down stream process sequences are as follows: After the formation of STI (Shallow Trench Isolation) with 0.25μm 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 Si₃N₄ 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 (Bit Line Contact) pad and SNC (Storage Node Contact) pad are formed in cell array as shown in Fig.1-(D). 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 Si₃N₄ spacer formation and SAC etching processes.
transistor whose primary reason is high thermal budget difficulty to form shallow junction of memory cell free from short channel effect. It is mainly due to the keep the effective channel length of cell transistor to be scaled down into 0.10\(\mu\)m regime, it is very difficult to 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.

As the technology generation of DRAM has been developed into 0.10\(\mu\)m 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 \(\text{C}_2\text{F}_6/\text{O}_2/\text{Ar}\). 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 0.10\(\mu\)m 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.

As the technology generation of DRAM has been developed into 0.10\(\mu\)m 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 \(\text{C}_2\text{F}_6/\text{O}_2/\text{Ar}\). 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 developed into 0.10\(\mu\)m 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 \(\text{C}_2\text{F}_6/\text{O}_2/\text{Ar}\). 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.
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.

As a result, more than 32ff capacitance per cell can be obtained as shown in Fig.10, where the capacitor dielectric, Ta$_2$O$_5$ (tantalum oxide), has the equivalent oxide thickness of 35Å.

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