HAL
open science

Critical sensitivity of flash gate dimension spread on electrical performances for advanced embedded memory

El Amine Agharben, A. Roussy, Marc Bocquet, M. Bileci, S Bégouin, A. Marchadier

To cite this version:

El Amine Agharben, A. Roussy, Marc Bocquet, M. Bileci, S Bégouin, et al.. Critical sensitivity of flash gate dimension spread on electrical performances for advanced embedded memory. 2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), May 2015, Saratoga Springs, United States. 10.1109/ASMC.2015.7164428 . hal-01737953

HAL Id: hal-01737953

https://hal.science/hal-01737953v1

Submitted on 20 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Critical sensitivity of Flash gate dimension spread on electrical performances for advanced embedded memory

El Amine Agharben, A.Roussy
EMSE-CMP, 880 Avenue de Mimet 13541
Gardanne, France

El Amine Agharben, M.Bocquet
IM2NP, 38 rue Juliot Curie, F-13451
Marseille, France

El Amine Agharben, M.Bileci, S.Bégouin, A.Marchadier
STMicroelectronics, 190 Avenue Célestin Coq, Zone Industrielle 13106, Rousset, France
elamine.agharben@st.com

***Abstract*—In this paper a correlation between specific inline and electrical parameters has been deeply investigated. This study shows the impact of the flash gate intra-die and intra-wafer spread on the memory performances. A process recipe optimization has been evaluated to recover the effects of these variations. This optimization allows us to reduce the written threshold voltage intra-wafer spread by fifty percent.**

Keywords: Non Volatile Memory; Embedded Memory; Semiconductor; Intra-wafer Variability, Intra-Die spread

I. Introduction

Flash memory represents the most important media to store all types of data. Depending on application characteristics, different architectures and devices have been developed over the last few years in order to satisfy all the customers' needs. Size scaling, faster access time and lower energy consumption have been the three pillars of scientific research in this area.

In a production environment, processes for transistor manufacturing suffer from permanent fluctuation caused by several factors. Some examples are changes in raw material properties (photo resist, substrate) and the evolution of equipment over time (aging, temperature change ...).

Because of high depreciation costs of microelectronics equipment, the semiconductor manufacturer is encouraged to adapt older generations of equipment to some advanced manufacturing processes. This strategy is not without effect on the spread of physical (geometric dimension) and electrical (current, voltage) characteristics spread. The yield may be degraded resulting in a reduction of profits.

In this context the variability of a parameter, whether physical or electrical, is twofold: spatial and temporal. The spatial variation is the combination of intra-wafer variation, for example due to an etching process to radius-dependent speed or an irregular spin coating, and intra-field changes, that may be linked to the previous process of photolithography. The variation within fields is due to the lens aberrations, the circuit architecture or intra-reticle errors. Finally the temporal variability, lot-to-lot or wafer-to-wafer is mainly coming from raw material variation and process drifts [1].

In this paper, we are going to analyze the impact of the Flash gate dimension on the electrical performances. We will essentially investigate the effects of intra-die (gate profile and word line uniformity) and then the intra-wafer variations.

II. Experimental details

A. Process description

Flash memories are organized in arrays of rows (Word Line or WL) and columns (Bit Line or BL). The type of connection determines the array architecture (Fig.1.a). In this paper a NOR architecture is used (Fig.1.b). The cells are connected in parallel and in particular the gates are connected together through the Word Line, while the drain is shared along the Bit Line.

In Flash process technology, the flash gate stack etch is one of the most challenging process steps. In our case the process must etch through an anti-reflective layer (ARL), a layer of Ashable Hard Mask (AHM), a thick layer of polysilicon (Control Gate), an Oxide/Nitride/Oxide (ONO) dielectric and a second layer of polysilicon (Floating Gate), while stopping on thin gate oxide (Fig.2).

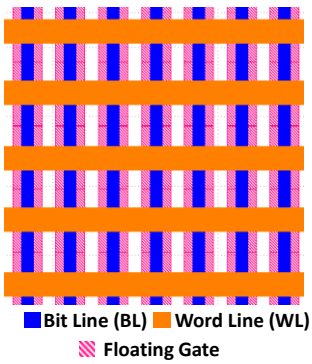

Fig.1.a : NOR Layout

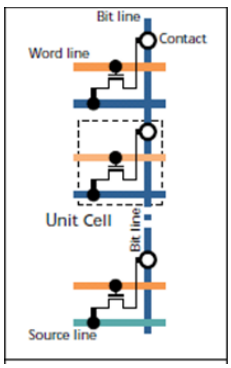

Fig.1.b: NOR circuit description

Fig.2 : Flash gate etch process definition

B. Analysis Methodology

Technology scaling has led to a significant increase in process variability due to effects such as random dopant fluctuation and imperfections in lithographic patterning of small devices. These variations cause significant unpredictability in the power and performance characteristics of integrated circuits. Process variation is typically observed at several different scales such as lot-to-lot, wafer-to-wafer, intra-wafer and intra-field variation. Lot-to-lot and wafer-to-wafer variations are caused by long term drifts in tools and wafer processing conditions.

Intra-wafer variation primarily occurs due to wafer-level non-uniformities such as Post Exposure Bake (PEB), to temperature gradient [2] and to resist thickness variation [3]. Intra-field variation, on the other hand, stems from optical sources such as focus variations across the field, dose variation [4] and mask errors [5]. In addition to the above sources, intra-die variation can also be caused by layout dependent systematic effects such as pitch and density dependent line-width variability [6] and microscopic etch loading [7].

The electrical test reflects a high spread of some critical parameters of Flash memory. After inline analysis it has been observed that these variations were due to fluctuation of the flash gate etch process. Since the product used in this experiment is under development we will analyze the spatial variation of the flash gate.

III. Intra-Die Variation

Intra-die variations can be classified into two categories systematic and random. Systematic variations have [8] a degree of correlation between different devices that changes with distance and are related to manufacturing imperfections or can be proximity dependent. The degree, distance, and type of correlation differ based on the source of variation.

Random variations are completely independent of layout and come mostly from the natural limits of scaling. These variations can cause a difference in the electrical characteristics of two otherwise identical devices (with the same geometry, layout, and neighborhood). Largely, the impact is on threshold voltage and mobility. At the current scale, factors such as surface roughness and the location of dopant atoms inside the bulk region affect also the electrical properties. There are three principal sources of intrinsic variations- Random Dopant Fluctuations (RDF), Line Edge Roughness (LER) and Oxide Thickness Variations (OTV).

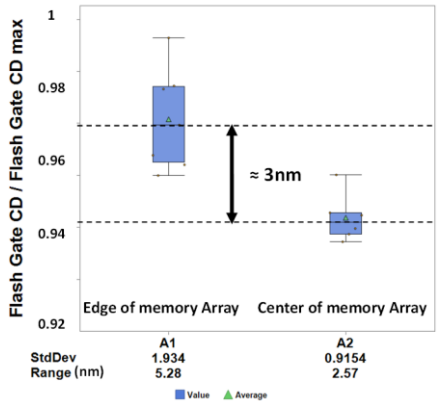

Fig.3 : Flash gate CD spread between the edge and the center of the memory array

In this part we tried to understand the impact of the intra-die variation on the cycling performances of the flash memory. The yield of the device used in this experiment is found to be very sensitive to the local flash gate Critical Dimension (CD). A reliable 3nm mean spread in the flash gate CD was found between the edge and the center of the memory array (Fig.3), by using the SEM as a metrology tool.

With a deep analysis, it has been found that basically it is the cells from the first and second bit lines of the memory array that are related to the intra-die variation discovered in Fig.3. This spread results in a degradation in cycling performance. With the physical analysis a morphological deformation has been high-lighted (Fig.4). The edge of the memory array environment can completely justify this distortion.

At that point the challenge was to find a way to measure and quantify this deformation. Several different metrology "Line Width Roughness" methods are under analysis to have a measurable evaluation of the flash gate CD at the edge of the memory array.

This distortion has a strong impact on the performances of the flash memory. An electrical wafer sorting test was performed on two different blocks (each one of 64 bit lines), to analyze the effect of the morphological uniformity on the cycling performance at 105 °C. The following chart (Fig.5) shows the flash memory threshold voltage distribution after that cycling test. As we can observe, tests with the entire block have an extrinsic limited slow erase, while the tests that exclude both first and last bit lines have a standard distribution.

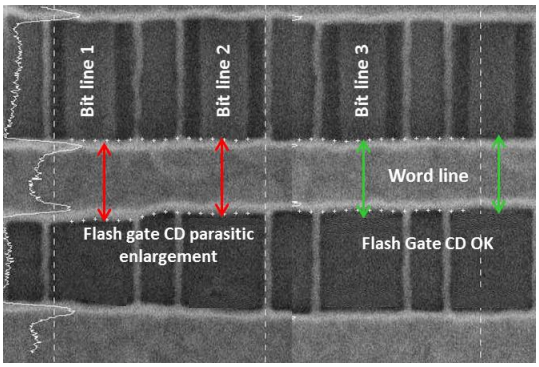

Fig.4: SEM top view for Intra-line CD spread measurements

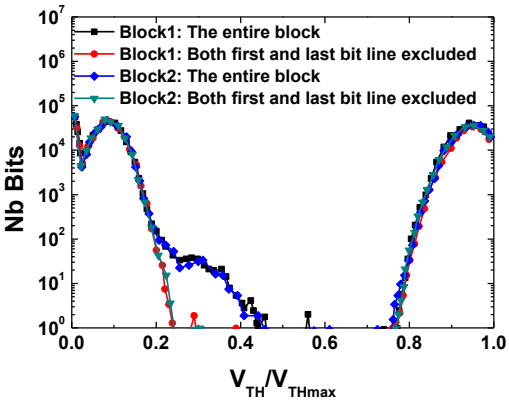

Fig.5 : Cycling test: V_{th} distribution at 105°C after 500K cycle

IV. Intra-wafer spread

In this part the impact of the flash gate CD intra-wafer variation on the electrical parameters will be analyzed. An etch process change will be proposed to improve the flash gate CD intra-wafer spread, and an analysis of the electrical parameters for each process change will show an improvement of in the spread of some electrical parameters.

A. Inline analysis

To analyze the flash gate CD intra-wafer spread we perform several full mapping measurements and this allows us to quantify the impact on some electrical parameters of the flash memory. These measurements show a 14 nm range that may be due to the uniformity of the resist deposition or the hard mask roughness.

To recover the marginalities here described, an alternative integration and process recipe optimization have been proposed to minimize the flash gate CD intra-wafer spread. In this article we show five different process solutions (Fig.6). We essentially try to add a cure treatment at different steps or modify the layer sequence. To summarize the results:

- Process 1 shows a big slope and some deformation.

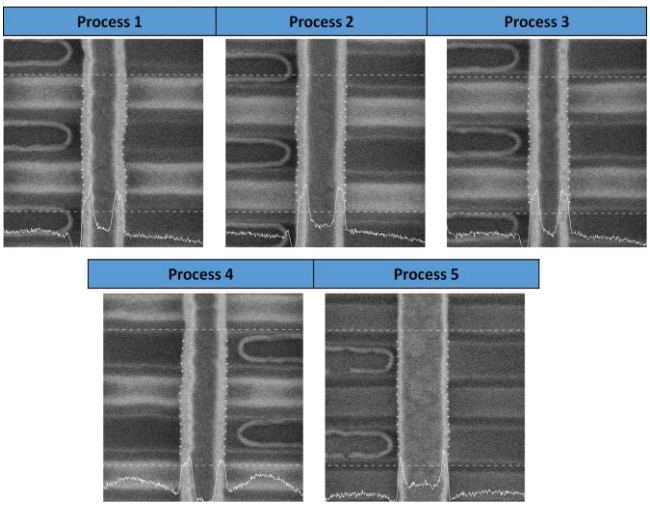

Fig.6: Flash gate SEM top view for each process improvement

- Process 2: the slope profile is improved compared to Process 1.
- Process 3 doesn't show any improvement of the Flash gate profile.
- Process 4: shrinks the Flash gate CD without any improvement of the profile.
- Process 5: good gate profile with negligible deformation and a straight profile.

For each process used to improve the flash gate profile, several full mapping measurements have been performed to analyze the flash gate CD uniformity. Fig.7 shows the CD intra-wafer variation for each process used in this experiment, Process 2 and Process 5 show around fifty percent spread compared to the others. The better gate profile is directly responsible for this improvement. In particular with Process 5 a more straight profile with essentially no roughness has been found.

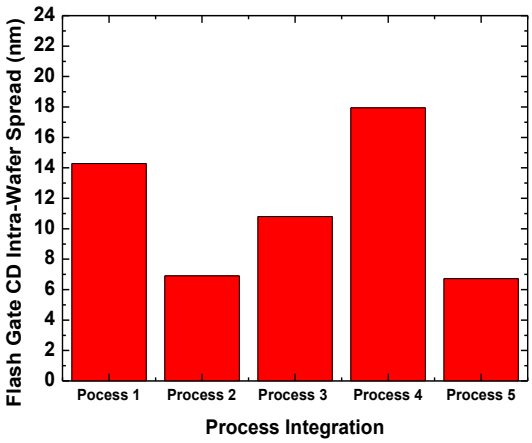

Fig.7 : Flash gate CD Intra-wafer spread

B. Electrical parameters analysis

Since the beginning of these experiments, the goal was to improve the flash gate CD intra-wafer spread to reduce the variation in electrical parameters. The electrical analysis shows a relationship between the flash gate CD and the threshold voltage write. This parameter depends on the flash gate profile definition and the intra-wafer spread. Fig.8 shows an improvement of the threshold voltage write spread. Both processes 2 and 5 give a better threshold distribution because of less line deformation and better CD spread uniformity.

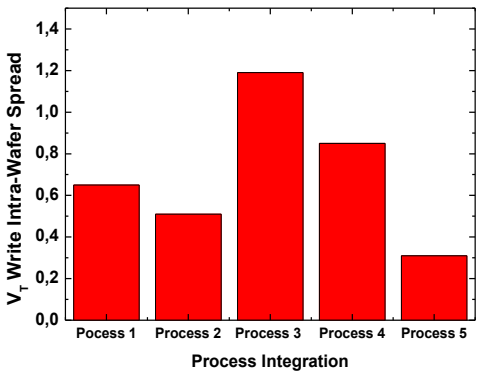

Fig.8 : V_T write Intra-wafer spread

V. Results and discussion

Several samples have been characterized in order to evaluate the variation in inline parameters on the electrical performances of the flash memory. The goal of this study is to analyze the effect of the spatial variation.

The intra-die analysis shows the impact of a slight deformation on the cycling performances of this device. The line width roughness is one of the best ways to inspect and quantify the flash gate deformation in order to evaluate the process drift that may be due to the temporal variation.

The impact of the intra-die variation led us to exclude the first and second bit line from the electrical test during device development. For an industrialization of this device two solutions have been proposed. The first one requires the use of Optical Proximity Correction (OPC) to redefine the size of the flash gate dimension on the edge of the memory array. The second one involves an addition of bit line dummies but it would affect the size of the embedded device.

However an analysis of the flash gate CD intra-wafer variation reveals a correlation with some electrical parameters like the threshold voltage write. The variability of these parameters directly impacts the reliability of the flash memory. In order to enhance the reliability of this device, an etch process improvement was done. The etch process change allows us to reduce the flash gate CD intra-wafer variation by fifty percent and to improve the spread in the threshold voltage write.

The choice of Process 5 rather than Process 2 as the new etch process was due to the flash gate profile.

Further investigations are running to improve the intra-wafer spread of the Flash gate CD, including even the intra-die variation. Only once the spatial variation is fixed will deeper analysis be performed to evaluate the temporal variation and to improve the process window of Process 5. A deployment of a Run-to-Run loop regulation between the flash gate lithography and the etch process is under development.

VI. Conclusion

In the work presented here we succeeded in decorrelating and quantifing the impact of intra-die and intra-wafer CD gate variation on electrical cell performances in a NOR memory array. We have found and improved process conditions to reduce these CD variations by fifty percent. Besides, it is clear that a wider process window is mandatory before entering the full production phase, therefore additional investigation and process improvement will be studied to reach the best possible CD spread for the most stable electrical response.

Acknowledgements

This research is part of the project MAGE, financed by the “Programme d’investissement d’Avenir” of the French Government.

References

[1] Brian E Stine, “Analysis and decomposition of spatial variation in integrated circuit processes and devices” IEEE Transactions on Semiconductor Manufacturing, 10:24-41, 1997.

[2] Y. Lee et al., “Temperature Rising Effect of 193nm Chemically Amplified Resist during Post-Exposure Bake”, SPIE, vol. 3999, pp. 1000-1008, 2000.

[3] C. Berger et al., “Critical Dimension Variations of I-line Processes due to Swing Effects”, SPIE, vol. 6153, pp. 61523T, 2006.

[4] Y. Borodovsky, “Impact of Local Partial Coherence Variations on Exposure Tool Performance”, SPIE, vol. 2440, pp. 750-770, 1995.

[5] A.K. Wong, R. Ferguson, S. Mansfield, “Mask Error Factor in Optical Lithography”, IEEE Trans. on Semi. Manuf., vol. 13, pp. 235–242, May 2000.

[6] D.G. Chesebro et al., “Overview of Gate Linewidth Control in the Manufacture of CMOS Logic Chips”, IBM J. of Res. and Dev., vol. 39, pp. 189-200, Jul 1995

[7] C. Hedlund, H. Blom, S. Berg, “Microloading Effect in Reactive Ion Etching”, J. of Vacuum Sci. and Tech., vol. 12, pp. 1962–1965, 1994.

[8] L.T. Pang, and B. Nikolic, “Measurement and analysis of variability in 45nm strained-Si CMOS technology”, CICC’08, p.129-132, 2008.