Extended Abstracts of the 1997 International Conference on Solid State Devices and Materials, Hamamatsu, 1997, pp. 8-11

Invited

Root Cause Analysis of Thin Gate Oxide Degradation during Fabrication of Advanced CMOS ULSI Circuits

Sang U. Kim*
Sematech, 2706 Montopolis, Austin, Texas 78741. Phone/Fax: (512)-356-7468-3437,
E-mail: sang.kim@sematech.org.*On assignment form Intel

1. Introduction

This study is focused on thin gate oxide degradation of an advanced CMOS technology in high volume manufacturing today. The technology uses four metal levels, 6nm gate oxide, .35um feature size, and dual poly-gates. For such a fabrication line, more than 30 plasma processes and approximately 10 CMP planarization steps are employed. Therefore, the gate oxide degrades due to the following main causes: 1) plasma induced charging damage since so many plasma steps are used and plasma damage is accumulative; 2) process induced defect since wafers are constantly exposed to numerous process induced particles and contaminants; and 1) + 2). The root cause analysis showed that the gate oxide deteriorated predominantly due to 1) + 2), that is, a combination of plasma process induced charging and process induced defect. Such a combination makes gate oxides to be more susceptible to plasma damage, and therefore, leads to an enhanced gate oxide degradation. The root cause analysis was performed through both electrical and physical analysis which identified the nature of the defect and its formation. The defect formation was explained by a proposed model. The enhanced gate oxide degradation during high volume manufacturing of advanced CMOS technology may limit gate oxide scaling.

2. Experiment

Plasma Process Induced Charging

The gate oxide degradation caused by a combination of plasma charging damage and process defect was verified by using an unconventional plasma charging test structure. In this structure, the antenna is fixed, and the transistor size is varied, in contrast to the conventional structure where the transistor is fixed and the antenna size is varied. In this way, the plasma damage can be readily distinguished from the process defect by making the transistor size large. The details on the test structure design, its salient features and test results can be found in reference[1]. Ig was solely used as a plasma charging damage monitor since Vt, gm , Dit and Qbd were found to be not sensitive enough for Tox < 7nm. The test result indicated a much higher number of p+ poly gate oxide fallout than the n+ poly gate oxides[1,2,3]. In the case of severe plasma damage, both n+ and p+ oxides were damaged to the same extent, but still a higher number of p+ oxide fallout in some cases. In the case of mild plasma charging, only the p+ gate oxide damage occurred, with minimum n+ gate oxide damage. This suggests that the plasma charging alone can not be responsible for the high susceptibility of the p+ gate oxides. There must be some sort of process induced defect associated with the p+ gate oxides. Therefore, the efforts were focused on electrical and physical analysis of the n+ and p+ oxides with high gate oxide leakage.

Electrical Analysis

Fig.1 shows a typical Ig -Vg characteristics of the p-channel transistor gate oxide leakage caused by plasma damage. It exhibits an asymmetric behavior, significantly higher gate oxide leakage with a positive bias voltage. It is similar to a p-n diode characteristic, a large current with Vg positive (forward bias) and a small current with Vg negative (reverse bias). Next, it was electrically determined where the gate oxide leakage originates in the transistor. With the measurement conditions as shown in Fig. 2, if Ig = Inw, it signifies presence of a defect in the channel region, if Ig = Is, a defect in the source region, and if Ig = Id, a defect in the drain region. This technique was applied to both n+ and p+ gate oxides damaged by plasma charging. A surprise result was obtained, especially for P+ gate oxides. It was found that n+ gate oxide leakage occurred predominantly at the source or drain edge of the gate oxide, even for the largest (100um x 100um) transistor. On the other hand, p+ gate oxide leakage occurred predominantly in the channel region of the transistor, even for the minimum channel (20um x .32um) transistor. The implication of an area defect in such a small or edge intensive transistor was not expected since the plasma damage occurs at the highest field point which normally exists at the source and drain edge.

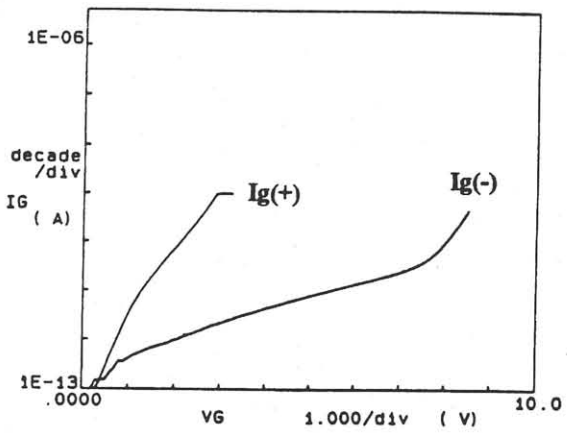

Fig. 1 : A p-n diode characteristic, a large current with Vg+ (forward bias) and a small current with Vg- (reverse bias).

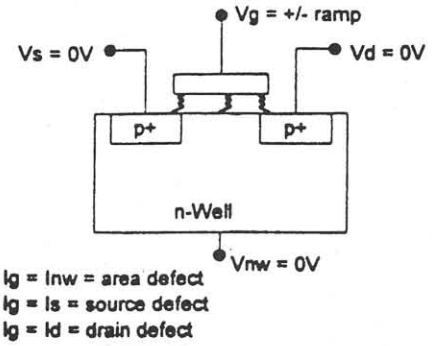

Fig. 2: Electrical measurement set-up to determine where the gate oxide leakage occurs.

EBIC Analysis

In order to confirm the electrical result, EBIC analysis was performed. EBIC can show not only where the gate oxide leakage originates, but more importantly, it can also localize or pinpoint the leakage site as shown in Fig. 3. The EBIC results were in complete agreement with the electrical results. For n+ gate oxides, the EBIC emission sites indicated by hot spots occurred predominantly at the edge, not at channel area even for the largest area (100um x 100um) transistor as shown in Fig. 3a. For p+ gate oxides, the emission sites occurred predominantly at the channel area, not at the edges even for the minimum channel length (.32um) transistor, as shown in Fig. 3b. Notice a very small localized nature of the emission site. This clearly indicates that some sort of process defects are formed in the p-channel gate oxides, contributing to the high susceptibility to plasma charging damage.

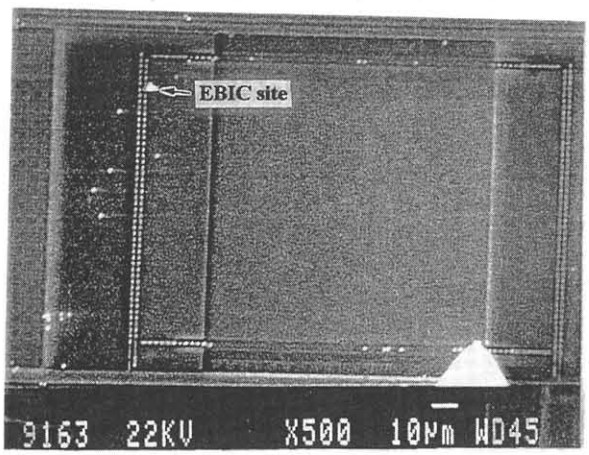

Fig. 3a: For the n+ gates, the EBIC emission spot occurs predominantly at the source and drain edge of the transistor, even for the largest area (100um x 100um) transistor as shown.

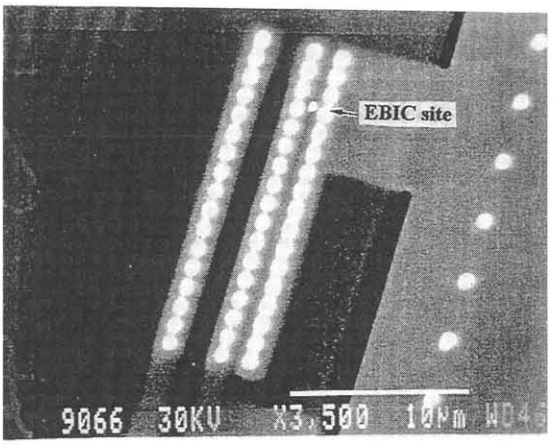

Fig. 3b: For the p+ gates, the emission spot occurs predominantly at the channel area, not at the edge even for the minimum channel (..32um) transistor as shown. It occurs at the center of the minimum transistor.

FIB and TEM Cross Section

The next step was to perform FIB cross section through the EBIC emission sites to determine the nature of the defect. The FIB cross section was found to be extremely difficult to do because the emission site where the gate oxide leakage originates was very small, ~500A or less in diameter. The cross section was performed in several increments. For each increment, the sample was taken out of the FIB machine, and was observed under a high resolution SEM. This procedure was repeated for each increminal ion milling. Only two out of eight samples were successful. One of these is shown in Fig. 4. Notice a missing gate oxide or a pin hole (~500A) created by plasma damage. A slight decorating etching was done to highlight the image. TEM cross section was not successful because of the difficulties involved in the sample preparation for such a small cross section area.

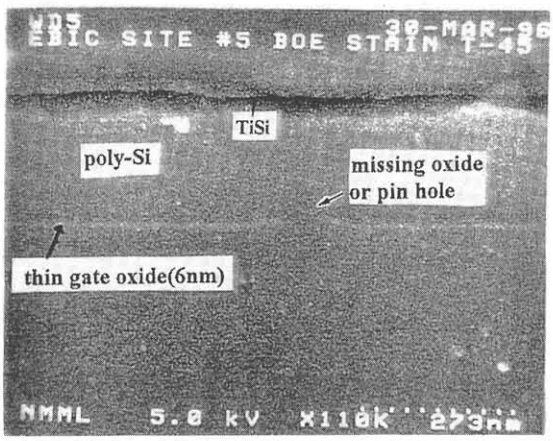

Fig. 4: FIB cross section through the EBIC emission site. A missing oxide or a pin hole created during plasma charging is clearly seen. Boron diffusion through the pin hole can form a p-n diode.

Continued on the 574 page

Nature of the Defect and Its Formation

It was conclusive from the EBIC and FIB cross section results that the enhanced p+ gate oxide degradation was due to a combination of process defect and plasma process charging. The defect, however, was not a random process induced defect. It is very plausible for some boron atoms in the p+-poly gate to diffuse through the pin hole into the n-well during the rupture process, and form a very small localized p-n junction which gives rise to an asymmetric electrical behavior as shown in Fig. 1. Therefore, the defect results from a localized p-n diode formation. The question is, then, how does such a defect form at the center of .32um minimum channel area? The defect formation is based on boron incorporation into the p+ gate oxide. It has been reported that the boron diffusion into the gate oxide of PMOS transistor is non-uniform across the channel, significantly higher at the center than at the source and drain edges[4]. The boron atoms diffused into the gate oxide are known to generate charge trapping sites[5], and thus the peak trap density must also occur at the center of the channel. These trapping centers may give rise to a localized weak spot in the gate oxide. During plasma charging, a large gate oxide leakage current may develop at the very small, localized weak spot where the highest boron concentration (or trap centers) exists. As a result, the thin gate oxide may be ruptured or melted due to excessive, localized heating, creating a pin hole as shown by the FIB cross section in Fig. 4. Some B atoms can readily diffuse through the pin hole into n-well below during the localized heating, forming a point contact (~500A) p-n diode. The p-n diode model is schematically illustrated in Fig.5. The p-n diode so formed is non-ideal, and its current varies,

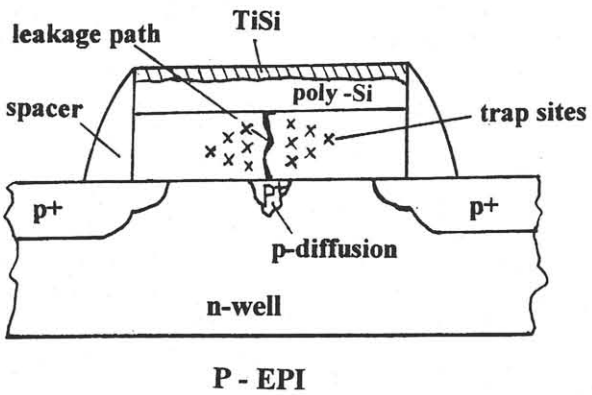

Fig. 5: A schematic illustration of boron induced defect formation. A point contact p-n diode is formed by boron diffusion through the very small, localized pin hole into n-well. The pin hole or a missing gate oxide results from excessive, localized heating as a result of a high gate current flow during plasma charging. The pin hole occurs near the center of the channel where highest concentration of boron induced trap sites occur. See the text for details.

depending on the amount of boron diffusion into the n-well and the pin hole area. The asymmetric Ig-Vg characteristic and other bias temperature stress results reported in the literature can be easily explained by the model. The positive gate bias is always the worst condition since with positive Vg on the gate, the p-n diode is forward biased, while a much smaller current flows with a negative Vg since the diode is reverse biased. The Ig-Vg characteristic, especially Vg positive as shown in Fig. 1 is often misinterpreted as SILC. It is not. It is a forward biased p-n diode current as shown here. As for the n+ gate oxides, the As or P induced defect formation in such oxides is not known. As or P atoms still can diffuse into the source and drain edge in the case of a severe plasma charging, creating a leakage path. However, the boron induced defects were found to be much more susceptible to plasma charging, causing an enhanced p+ gate oxide degradation.

3. Summary and Conclusion

The p+ thin gate oxides of the advanced CMOS technology in manufacturing today were found to be inherently more susceptible to gate oxide degradation than n+ gate oxides. This is because the p+ oxide degradation is caused by a combination of the process induced defect and plasma charging. The defect ,however, is not a random defect induced during process. The nature of the defect and its formation were determined by electrical and physical analysis such as EBIC and FIB cross section. The defect formation was modeled. The defects were found to result from boron incorporation into the p+ gate oxides during PMOS process. The p-channel gate oxide degradation will be worse with gate oxide scaling and may limit the device scaling. Therefore, future device scaling should consider a realistic gate oxide scaling limit which is based on p+ gate oxide degradation during high volume manufacturing of the advanced CMOS technology. The boron induced defect formation must be suppressed during high volume manufacturing.

Acknowledgments

The author would like to acknowledge Greg Neri for device measurements and data collection, Chris Schwappach for excellent FIB cross section, Tom Tribble and Das Chowdhry for TEM support. Also, many thanks to Analytical Solutions for EBIC analysis services and coorporation. Finally, thanks to Sematech and Intel for providing atmospheres and support to complete this work.

References

1) S. Kim: 1996 International Integrated Reliability Workshop, P57
2) W. Abadeer et al: 1996 1st P2ID Conference, P160
3) S. Ma et al: 1997 2nd P2ID Conference, P25
4) C. Subramanian et al: IEDM-95, P423
5) E. Hasegawa et al: IEDM-95, P327

Continued from the 9 page