Evidence for a Correct SiO₂ Voltage Acceleration Model

W. M. Miller*, C. Messick**, N. F. Smith* & J. A. Shideler***

* Sandia National Laboratories, M/S 1081, PO Box 5800, Albuquerque, NM 87185-1081
(505) 844-7161, FAX: (505) 844-2991, e-mail millerwm@smtplink.mdl.sandia.gov

** National Semiconductor Corporation, 3333 West 9000 South, West Jordan, UT 84088-8838
(801) 562-7546, FAX: (801) 562-7500

*** National Semiconductor Corporation, 2900 Semiconductor Drive, M/S A-2513, PO Box 58090,
Santa Clara, CA 95052-8090, (408) 721-7558, FAX: (408) 721-7266

Abstract

We present experimental evidence that SiO₂ breakdown is best described by the E-model of Baglee and McPherson as opposed to the 1/E model of Lee, Chen and Hu. The experiment was performed on 1765 fully processed capacitors over a range of electric fields of 7.25 MV/cm to 11 MV/cm and temperatures of 25C to 200C. In addition, we also present a curious data point at low electric fields that we ascribe to a change in breakdown mechanism.

Introduction

I'm going to begin now and talk about evidence for a correct SiO₂ voltage acceleration model. This is some work that was done in close collaboration between Sandia Labs and National Semiconductor. You'll see the contributions that National made very shortly. Let me begin by suggesting why one might want to have a voltage acceleration model. If you ever hope to make an extrapolation of data taken at wafer level (large electric field) to the lower electric fields that are more typical at use conditions, you're going to have to have a correct physical model of oxide breakdown.

Why do we need an SiO₂ Voltage Acceleration Model?

- For extrapolations of breakdown data from one electric field stress to another
- To provide insight into the physical mechanism of breakdown
- To estimate the time needed to complete a breakdown test

I'm not claiming that you can presently make such an extrapolation; I'm not claiming that you will ever be able to, but as a prerequisite for it, you do need to have the correct physical model. Also, if you do have the correct model it may be able to provide you some insight into the physical mechanism of breakdown. This can help you understand your oxides better. Finally, and this is a more practical thing, if you

do have such a model of oxide breakdown it will help you to optimize wafer-level testing by helping you to predict when your test structures will break down under test.

These Topics Will be Covered in the Talk

- What Models Are Used to Describe Oxide Breakdown?
- Oxide Parameters & Test Structure Design
- Experiment Design and Description
- Test System Description and Definition of Breakdown
- Results of the Experiment
- Some Conclusions

Now, the first topic that I am going to cover in this talk concerns the models that are currently used to describe oxide breakdown. I'm then going to go on and describe the oxide parameters of the samples that we have looked at. Next, I'll deal with the test structure design. I'll describe the design of the experiment. Then I'll describe the test system itself and what we define as breakdown. The definition of breakdown is very very important so that you can interpret the results. Finally, I'll talk about the results of the experiment itself and then give you some of our conclusions.

Models of Oxide Breakdown

Those of you who were at the tutorial have already seen the following slide that lists the models that are used to describe oxide breakdown. There are several models. However, there are two major competing models that exist today and that no one can seem to choose between.

The first is the E-model, also known as the McPherson-Baglee model that was based mainly on thermodynamic arguments¹. It predicts a number of things. First, it predicts that the activation energy of the oxide breakdown varies with electric field. It also predicts that the electric field acceleration factor varies with temperature. If you make extrapolation with this models it tends to be more conservative than the 1/E model.

The 1/E model, on the other hand, was proposed by Berkeley group of Lee, Chen and Hu².

What Models are Used to Describe Oxide Breakdown?

• E Model Physics -Based on Thermodynamics • Predictions -E Varies with Field -Electric Field Acceleration Factor Varies with Temperature • Extrapolations -More Conservative than 1/E Model	• 1/E Model Physics -Based on Conduction • Predictions -E Varies with Field -Electric Field Acceleration Factor Varies with Temperature • Extrapolations -Predicts Longer Lifetimes

It is based on conduction mechanisms, namely on Fowler-Nordheim tunneling. It also predicts, if you look into it carefully, that you can expect the activation energy to vary with electric field and also the electric field acceleration factor to vary with temperature. It predicts similar effects when compared with the E-model. Where it differs is in how it extrapolates time-to-breakdown as a function of electric field. When you extrapolate, the 1/E model tends to predict longer lifetimes. So, the two models predict qualitatively similar things, but there are some quantitative differences.

Test Structure Design

I will now describe the test structure design. This test structure is National Semiconductors FLGOI test structure. It

Test Structure Design and Oxide Parameters

- We used National's FLGOI Test Structures (see inset photo)
- All samples taken from ONE wafer
- Area: $1.0695 \times 10^{-3} \text{ cm}^2$
- Thickness: 225Å
- Doping: p-type
- Process: Full
- Gate: Silicide
- Contacts: All topside

was designed by one of the co-authors, Cleston Messick. As those of you who came to the tutorial will probably see, the FLGOI structures obeys most of the oxide test structure design "rules" that I talked about there. Cleston didn't design the structure based on these rules, but he did it based on his own sound engineering judgment. It resulted in a very nice test structure. We had absolutely no problems with this structure during test. The oxides failed cleanly. We did not have any

extraneous failures. In many ways it was an ideal structure for the experiment which I'm about to describe. Another important feature of this study, was that all of the samples that I will talk about today were taken from a single wafer. We processed a wafer lot with nothing but these test structures on it. We were able to get several thousand test structures out of one wafer that we were then able to use for what follows. This is very important from the standpoint of wafer-to-wafer variation. There was none, because there was a single wafer. In fact there was very little die-to-die variation when we did some initial wafer-level testing on this single wafer. In all ways we believe this to be a uniform set of samples. The area was slightly in excess of 10^{-3} cm^2 , so that we were dealing with a fairly good size test structure. These are 225Å oxides with p-type doping in the underlying silicon. These structures were fully processed, so that there was no question about the results coming from only a short-loop process. We were therefore studying what an oxide would look like after the full process was complete. The gate is a polycide. All contacts were to the top side of the structure. This removes the potential problems with backside contacts.

Design and Description of The Experiment

Let me now talk a bit about the experiment design and exactly what it is we wanted to accomplish. What I've drawn on the graph below are the two models. These are not based on any data, rather it just shows the functional dependence of

Experiment Design and Description

- To distinguish between the two models you need:
 - A wide range of electric field data
 - A large sample size to minimize the error bars
- Our experimental parameters were:
 - Electric Field: 7.25 MV/cm to 11 MV/cm
 - Temperature: 25C to 200C
 - Sample Size: 108 to 120 parts per condition
 - Total Sample Size: 1765 parts
 - Breakdown Time Resolution: 10µsec

the two models. One of the interesting things that comes out of this is that the two models look very similar in the regime that you would look at in wafer-level testing. Say, between 8 and 11 MV/cm. You can't easily distinguish between the two models in this regime. You notice that they are very close together. As you drop in electric field, however, the two models tend to diverge. As they diverge, you'll notice that by the time that you get down to typical operating conditions you have predicted times to breakdown, given by the two models, that are quite a bit different. They can be many orders of magnitude different. That just again points out the fact that what you would really like to know which model is correct. Now then, there are two ways that you can go about dealing with determining which model is correct. In fact, you'll be

hearing a paper here in a little while from John Suehle that describes the other way of doing it. One way to determine which model is correct is to push your experiment down to very low electric fields. Unfortunately, you cannot afford to wait long enough for the breakdowns to occur. John got around that problem by going to higher temperatures. We decided to do a somewhat different type of experiment. That was, that at the modest fields that we could easily do in about a nine month time period (which was the longest that any of our experiments ran), we used a sample size that was relatively large. By doing that we were able to drive the standard deviation (in fact, the 95% confidence limits) small enough that we believe that we could tell the difference between the E-model and the 1/E model.

I'll now discuss our experimental conditions for this work. We studied oxide breakdown at electric fields between 7.25 MV/cm and 11 MV/cm. We have another data point which I will also present at 6.5 MV/cm, or should I say the lack of a data point at 6.5 MV/cm. You'll see that is just a little bit. The temperatures that we stressed the test structures at ranged from room temperature (25C) up to 200C. In order to get the amount of data that we needed at low electric fields the majority of the data that we took was at 200C. The sample size varied between 108 and 120 parts. The sample size was chosen based on the requirement that we be able to distinguish between the E-model and the 1/E model under the electric field acceleration conditions that we were using. The total sample size for this experiment was 1765 parts. The breakdown time resolution on my equipment was 10μsec. So that allowed us to measure the time-to-breakdown of the very early failures.

Test Procedure

Our equipment was specifically designed to run this experiment. We built the apparatus at Sandia specifically to answer E versus 1/E question. The system capacity was almost 3500 capacitors. The temperature was controlled with a uniformity of less than 5°C within the temperature control chamber shown in the figure below. We provided a backup power source, because any experiment that runs for nine months is bound to suffer some power glitches if you are not careful. We did not have the computers go down catastrophically in that time (they had backup power).

The test procedure is given on the slide below. To begin with, each capacitor was tested individually, at voltage, for 300 seconds. Our reasoning was that if we had a lot of early failures and had put all of the capacitors on in parallel initially, we would get so many failures that would overlap that the system would not be able to pick those failures up. So, we started this experiment out by testing each individual capacitor for 300 seconds at voltage while we were monitoring the current so that we could determine whether or not the capacitor broke down. At that point, once the parts had passed through this test, that we called a short-term test, we then hooked them all in parallel. The electric field stress was then

put on all of them in parallel. The total current drawn by the capacitors was then monitored, so that if

Our Equipment Was Specifically Designed to Run This Experiment

- System Capacity: 3456 capacitors
- Temperature Uniformity: <5C
- Backup power source provided
- Test Procedure:
 - Each capacitor tested individually for 300 sec.
 - All capacitors then hooked in parallel
 - Total current draw monitored
 - Increase in current draw triggers failure routine
 - Time recorded
 - Failed capacitor found and electrically removed
 - Fail time for that capacitor recorded
 - Test resumed

at any time any one part failed, it could be detected by the test unit. It would then tell us that there had been a failure. If there was an increase in current drawn during this time, it triggered the failure routine within the computer running the experiment. The moment that the trigger pulse was received the computer recorded the time and removed the bias from the parts. It turns out that with this kind of a setup, working in parallel, we were able to achieve about a 10μsec resolution for breakdown. This gets us around some of the problems that had been seen earlier in trying to scan through a large number of capacitors on test and find breakdowns that way. After the trigger time was recorded, the computer went into a routine which went through and tested each capacitor at low voltage. Low voltage being 5V in this particular case. This is far below the voltage that corresponded to the lowest electric field that we applied during the experiment. We determined that this voltage was insignificant in terms of total stress that was accumulated during the experiment. When the failed capacitor was found, it was noted in the computer and the time tag that was associated with the failure was added to the file that was associated with that capacitor. The capacitor was then electrically removed from the system. The test then automatically started back up again and continued. In all of these cases, we tested all samples in the group until 100% of them failed, with the exception of the samples tested at 25C and 8MV/cm. We did not test to 10% failures and then extrapolate to 50% (median) breakdown.

Data Analysis Procedure

Now, for a short description of how we collected, sorted, analyzed and plotted our data. You will see an example shortly. Occasionally, although it did not always happen, we would find bimodal distributions. In order to understand what was going on with the "intrinsic" portions of the curves, what we wanted to do was separate out the effect of the extrinsic (or defect) population (which was relatively small) from that. We separated that out using a technique that is described by Stu Peck in the Accelerated Testing Handbook of Technology Associates. The result of all of this is two best-fit lines. One

for the defect population and the other for the main ("intrinsic") population. We were only interested in the line for the main population, primarily because we had

Description of Data Analysis Procedure

- Data collected, sorted and plotted
- If bimodal, both modes separated using technique of Stu Peck (see "Accelerated Testing Handbook", Tech. Assoc.)
 - Results in two best fit lines:
 - One for the defect population
 - One for the main population
- Median Tbd plotted for main population only, as defect Tbd data was limited

a limited data set for the defect population. What you see below is probably the best example that we have of both populations.

Bimodal Distributions Required Deconvolution

This has many more defects that we typically saw. This shows a bimodal distribution and how we deconvolved it. The data points are shown as black squares. Using the technique that you find in the Accelerated Testing Handbook, we found both a main distribution and a defect distribution curve. There is a program that I have written using the software known as Mathcad that does this for me. All you need to do is pick a defect fraction and it will split these two curves apart for you. What we can then do is take the two curves and see how well it fits the data. We reconvolved the two curves to create a composite distribution. As you can see here it very accurately fits the data. We have a great deal of confidence, then, that the main and defect distributions that we have extracted from the data, are in fact the correct ones. As you can see towards the bottom of the plot, we did occasionally have some initial failures. From this point on we will be working with the main distribution, and in fact the median time to failure of that main distribution.

Discussion of the Data

This is an example of the data that we got out of the experiment. One of the things, as you recall, that both models predict is that the electric field acceleration factor

Field Acceleration Factor Varies with Temperature

should vary with temperature. That is what both models say. We, in fact, saw something that looks strikingly like that. These are all best-fit lines. Of course the best-fit through two points (on the upper data at 25C) is a straight line. You can't exactly get much better than that, can you. As you see, as you "walk up" in temperature, as temperature increases from 25C to 68C to 125C to 200C, the slopes of these line decrease. This is something that is predicted by both the E and the 1/E models. This is then no big surprise. Again, I have plotted median time to breakdown in seconds on the logarithmic y-axis and electric field on the x-axis. So, based on this result, there is nothing too surprising. Both models seem to make sense.

In the plot of activation energy we can also see something else. Both the E and the 1/E model predict that the activation energy should vary with electric field. And in fact, surprisingly and strikingly enough, we saw exactly that. Now, there is something that I would like to point out here. I'm plotting median time to breakdown (in seconds) on a logarithmic scale on the y-axis. I'm also plotting one over temperature (1/T) on the x-axis. This is a standard Arrhenius plot. There is an important point to be noted here. That is, that at relatively high electric fields, in this case 11MV/cm, you are getting activation energies on the order of 0.27eV, or 0.3eV, or something on that order. If you look in the literature, this is what most people give as the activation energy of oxides. The interesting thing is, that as you drop down to lower and lower electric fields the activation energy tends to increase. The slope is increasing. So at 9MV/cm we are dealing with 0.51eV and then finally when we get down to 8MV/cm we have 0.94eV. Both models predict that there is going to be a change. In other words, the activation energy is predicted to increase as the electric field gets lower. The point that I am bringing out here is that again, we have no distinction between the two models. It is very important to notice this particular effect.

Activation Energy Varies with Electric Field

And now we get to what I consider to be the most important data. That is, the one plot that may be able to distinguish between the two models. What I have done below is plot the median time to breakdown in seconds on the logarithmic y-axis, versus the electric field in MV/cm on the x-axis.

E versus 1/E Experiment Comparison of E-model to 1/E Model

Consequently, if the data behaves as the E-model, we should have a straight line. If the data behaves as the 1/E model on this kind of a plot, we should have an upwardly curving line. The upwardly curving line is the black line marked as 1/E curve in the preceding figure that is predicted from the 1/E model. It was based on the particular data set that we have. What we have done is chose enough samples, such that the 95% confidence limits for median time to breakdown are enclosed within approximately one of the diamonds representing each data point. There are no error bars shown as I could not put on error bars using the graphing package that I have. But the size of one of these data points is approximately equal to the 95% confidence limit. What you see, at least down to the 8MV/cm data point is something that is fitted very cleanly by the E-model. It is in no way described by the 1/E model, although the two models are both very close to the data points. We are trying to do something that is very difficult here. That is, distinguish between the two models. Really one would like more evidence at lower electric fields. When we went to try and get that evidence, however, something odd happened. At 7.25 MV/cm we got the point shown, which

caused us to be somewhat concerned. We are not certain about whether the data point is associated with a change in breakdown mechanism or if something else is happening. Clearly, we believed our data point, but we were very concerned about it. The next thing that we did was to take a data point at 6.5 MV/cm. If, as we had suspected, we were seeing a change in mechanism, we would have expected a median time to failure in a reasonable number of seconds. The unfortunate thing was that as of last Thursday (just before the workshop), the parts had accumulated a little over 3×10^6 seconds. Out of the 100 parts that we had on test not one of them had failed! We are not exactly sure what is happening out there at 6.5 MV/cm. That data point tends to call into question what we observed at 7.25 MV/cm. Clearly, if the 1/E model was correct, we should never see the failures at 6.5 MV/cm. Quite frankly, I doubt we ever will see failures if the E-model is correct either.

To me the most convincing part of the entire analysis can be seen by picking up your workshop notebooks. Hold them up sideways and look down the breakdown curve. It will give you a much better view of what I mean by data points that we have at relatively high electric fields following more closely to an E-model than a 1/E model. I do not necessarily consider this the "last word" in figuring out whether the E or the 1/E model is correct. It certainly is another piece of data that indicates that you just can't automatically throw away the E model as a possibility.

In addition, we also went into the literature and replotted some data by Shiono and Manabu³ that was presented at this last years (1993) International Reliability Physics Symposium (IRPS). In that paper the two authors suggested that their data

Replot of Shiono & Manabu Data Also Supports Use of E-model

very clearly is described by the 1/E model. In fact when the data is replotted, it causes one to wonder whether or not either model might not fit well. Although, if you look at the data, then except for one data point, which may be significant, (they had a small sample size so the error bars are much larger than in our case) the data appears to follow the E-model. If you include that one data point then it looks more like a "wash". This is yet another piece of evidence that you can't

throw away the E-model just yet and go with the 1/E model, no matter how physically beautiful it may in fact seem.

Conclusions

So then, I believe that I have presented some data from both our experiment and a little bit from the literature that does not rule out the E-model. In fact, it somewhat supports the use of the E-model. The definitive experiment,

Data From Our Experiment and From the Literature Supports Use of the E-model

- Temperature variation of field acceleration factor consistent with E and 1/E oxide breakdown models
- Field variation of thermal activation energy also consistent with E and 1/E models
- Median times to breakdown BEST described by E model for both our data and one set of data from the literature
- **HOWEVER:**
 - 1) Correct model may be process or technology dependent
 - 2) Our data at 7.25 MV/cm may indicate a change in mechanism
 - 3) Data below 7 MV/cm seems critical to determining proper model unless very large sample sizes are used and testing is continued over a long period of time

I would say, almost hasn't been done. Then again, you are going to be hearing a paper from John Suehle in a little bit, dealing with these exact same samples in which he was able to make measurements down to much lower electric fields, albeit at higher temperatures. As you can see from the data that I have presented, temperature variation of the electric field acceleration factor, namely the fact that it does vary, is consistent with both the E and 1/E models. So far, so good. The electric field variation of the thermal activation energy also is consistent with both the E and the 1/E models. However, the median breakdown times of our data are best described by the E-model. Also, the data from the literature shows that that is not an unreasonable conclusion.

Now, I'm going to give you a couple of cautions here. The one thing that I have learned about oxides is that every time that you think that you have them "pinned down", you'll get another data point that shows that you don't really understand their behavior. The "correct" model may in fact be technology dependent. There is no evidence to show that this cannot be the case. From a physics standpoint, it doesn't make me feel very good that technology may have some say in all of this. The best strategy may be to accept neither model for right now. Check your own oxides out, this is very important.

Our data below 7 MV/cm might indicate a change in mechanism. Then again, something else may be going on. If nothing else, that data also suggests caution. Do not just blindly take data at very high electric field and try to extrapolate them to much lower electric fields. In fact, it is probably wise to take data at electric fields as low as you can conveniently go, before you try to do anything else with it. Data below 7 MV/cm does seem to be critical in determining

the proper electric field acceleration model unless you can afford to use extremely large sample sizes to pin things down. Unfortunately, when you do use those large sample sizes below 7 MV/cm it will require continuous testing over a very long period of time.

My final conclusion is that we have presented some evidence to show that the E-model may in fact work. As I have said, I do not believe that the definitive experiment has been done yet. You'll get a little better flavor for the evidence we have here after a couple more papers. Thank you.

Questions and Answers

Warren Gladden (AMD): Based on the data that you have would you recommend that the activation energy that we chose for gate oxide failures be changed from 0.3 eV?

Bill Miller: I would recommend that you test your oxides to make certain of what you have. The real problem is that activation energy measurements are made at the wafer level and consequently at very high electric fields. You get 0.3 eV and everyone feels good about it. The problem with that is that as you go to lower electric fields, that activation energy, at least in my experience, increases. That has a major effect on how you interpret your data. If that holds true for other oxides also, the implication is that as you go to more modest electric fields (getting closer and closer to typical use conditions) thermal acceleration factors become more and more important. That is not something that you would ordinarily think of. Most people believe that thermal acceleration is only a small effect compared with voltage acceleration. But in fact, at more modest electric fields these data indicate that that is not true.

Warren Gladden (AMD): The other question that I had was that looking at the data that you showed, have you tried to compare it to what Boyko and Gerlach⁴ presented at IRPS a couple of years ago? They actually showed that the electric field dependence bent over at lower electric fields.

Bill Miller: Originally, when I saw that data point at 7.25 MV/cm I thought that I was seeing an effect similar to what they had been seeing. Right now I can't draw a final conclusion as to what is going on at that data point. Boyko's data was taken, I believe at 125C. We were at 200C. We thought that he was seeing an effect of a change in mechanism between 5 and 6 MV/cm and that perhaps at higher temperatures that "kink" occurs at higher electric fields. I'm not so sure now that that is the case.

Peter Kowalczyk (Delco Electronics): I was noticing here that you were at some high electric fields. Then you went down in electric field. You were saying that the parts were tested at 5V. Did all of the parts fail? If they had failed at the high electric field, did you now say that they had failed?

Bill Miller: Yes, absolutely. What would happen is that we would test until we got an increase in current. We would always, in fact, be able to find which part had failed. These parts had this nice characteristic, partly because of the good test structure design. When they failed, they failed as dead shorts. We could immediately pick them up. So we had no problem picking them up.

Ken Boyko (AT&T): I want to get a clarification on your test procedure. You said that you tested the capacitors individually for 300 seconds.

Bill Miller: Correct.

Ken Boyko (AT&T): Then they were hooked in parallel. The question is: Was the electrical bias removed at that point?

Bill Miller: The electrical bias was removed between the short and the long term tests. As it was, also, each time that a part failed. So, if a part failed, we would remove the bias. We would then go through at a much lower bias, look for the failure site, find it, remove it and then start the bias up again. In answer to the implied question, we were a little bit concerned about it causing other kinds of effects, because we were constantly removing and adding the bias. And in fact, in experiments in which we removed and added the bias quite frequently we saw no effect on the type of breakdown times that we had. Admittedly, during one of these tests, it would only happen on the order of 100 times.

Hareesh Bhat (Philips): Do you have any comments on the effect of defect distribution being affected by the electric field? In other words, have you done a similar study on the group of parts that have a higher defect density than what you showed?

Bill Miller: The answer is, unfortunately, no. The samples that we got were very clean; they had a very low defect density. In order to do this kind of experiment on defective parts... To begin with you have to do this kind of experiment on parts with lots of defects. As you have seen, in order to drive the standard deviations to something reasonable, you need to have quite a bit of data. I would need to have on the order of a hundred or so "defective" oxides. This would mean that in most cases I would have to start with thousands, tens of thousands or even more samples in order to hope to do the experiment this way. The answer again is no. I wish that I could, but what can I say, National just made good oxides for me. (Applause) I hear National clapping!

Paul Marcoux (Hewlett Packard): Your changing activation energy implies some change in the mechanism. Do you have any comments on how that mechanism might be changing?

Bill Miller: That is an interesting question. I would not necessarily say that it implies a changing mechanism. That is all. Both of these models assume a single mechanism. The

nature of the models indicate that you can see such changes in activation energy.

Bud Trapp (Technology Associates): You said that this was a silicide gate? Is this not polysilicon and silicide together?

Bill Miller: I'll ask Cleston, if you are if, if you could field the processing question. Cleston?

From the Audience: Polycide.

Bud Trapp (Technology Associates): Polycide. Thank you.

References

- ¹ McPherson, J. W. & Baglee, D. A., "Acceleration Factors for Thin Gate Oxide Stressing," Proc. Rel. Phys. Symp., 23,p.1,(1985).
- ² Lee, J., Chen, I-C & Hu, C., "Statistical Modeling of Silicon Dioxide Reliability," Proc. Rel. Phys. Symp., 26, p.131, (1988).
- ³ Shiono, N. & Manabu, I., "A Lifetime Projection Method Using Series Model and Acceleration Factors for TDDB Failures of Thin Gate Oxides," Proc. Rel. Phys. Symp., 31, p.1, (1993).
- ⁴ Boyko, K. C. et. al., "Time Dependent Dielectric Breakdown of 210A Oxides," Proc. Rel. Phys. Symp., 27, p.27, (1989).

The work performed at Sandia National Laboratories is supported by the U.S. Department of Energy under contract DE-AC04-94AL85000.

William M. Miller, Don Pierce & Eric S. Snyder
Reliability Physics Department

Sandia National Laboratories
Run by Martin Marietta for the
U. S. Department of Energy

Electronic Qualification Industry Center

References

- What is Wafer Level Reliability
- Dielectric Breakdown - Bill Miller
- Electromigration - Don Pierce
- Hot Carrier Degradation - Eric Snyder

- WLR is done with WAFERS, although some systems will also handle die (with difficulty) or packaged parts
- It is FAST compared to standard package part tests
- But tests are usually done **SERIALLY**, stepping from die to die, rather than being done in parallel!

- The effect of process changes on reliability may be rapidly determined
- Product reliability may be routinely monitored
- Large sample sizes may be tested quickly to allow low defect levels to be measured
- Testing at wafer level does not require expensive and time consuming packaging

- **It cannot predict product reliability under typical use conditions**
 - As tempting as it may be, any extrapolations may be off by orders of magnitude and should not be trusted as definitive numbers.
 - Most wafer level tests do not run at operating frequency.
- **It cannot catch all reliability problems**
 - Package issues still dominate overall reliability.
 - Cannot catch wafer sawing induced defects.
 - Reliability while being used in final circuit cannot necessarily be predicted.

- **Manual**
 - Uses micromanipulators or card
 - Practically limited to 6-8 pads with micromanipulators
 - Requires operator to move between die
- **Semiautomated**
 - Uses micromanipulators or card
 - Requires operator to move between wafers
- **Fully Automated**
 - Only uses card
 - Can test multiple boats of wafers without intervention

Electronic Test Equipment is Also Needed

- Benchtop equipment may be used
 - e.g. HP4145
- Automated Test System may also be used (e.g.)
 - Hewlett Packard
 - Keithley
 - Reedholm
 - QualITau
 - Micro Instruments
 - etc.

■ Sandia National Laboratories

Wafer Level Testing Strategy

■ Sandia National Laboratories

The Correct Physical Model is Important for Reliability Prediction

- Must Extrapolate From Test to Use
- Conditions are Very Different
- Small Variations in Models Lead to Very Large Differences in Predictions

■ Sandia National Laboratories

E and 1/E Models Differ

• E Model Physics – Based on Thermodynamics • Predictions – E_a Varies with Field – Acceleration Factor Varies with Temperature • Extrapolations – More Conservative than 1/E Model	• 1/E Model Physics – Based on Conduction • Predictions – E_a Varies with Field – Acceleration Factor Varies with Temperature • Extrapolations – Predicts Longer Lifetimes

■ Sandia National Laboratories

Correct Statistical Model Needed For Practical Reliability Prediction

- "Parts fail at the tail"
- Practical tests make use of reasonable sample sizes
 - The larger the sample size the greater the confidence
 - But, the larger the sample size the longer the test
- Extrapolations to the tail of the distribution require the correct statistical model

■ Sandia National Laboratories

Four Common Statistical Models Are Used for Oxide Reliability

- Normal Distribution (Gaussian)
- Lognormal Distribution
- Weibull Distribution
 - Many Adjustable Parameters
- Extreme Value Distribution
 - Is NOT the Weibull Distribution
 - Is a special case of the Weibull Distribution

■ Sandia National Laboratories

Oxides Follow Area Dependence Predicted by Extreme Value Distribution

Is there really an "intrinsic" Ebd?

■ Sandia National Laboratories

The Extreme Value Distribution Describes Ebd Data Best

- Predicts observed area effect
- Is physically reasonable
 - The larger the area the larger the likelihood of finding a lower Ebd
- Calls into question concept of "intrinsic" breakdown field
 - Caution should be exercised when sizing oxide test structures

■ Sandia National Laboratories

Proper Test Structure Design Needed for Proper Reliability Interpretation

- Assumptions:
 - Test Structure Failures Mimic Production Part Failures
 - Test Structure Design Simple Enough to Understand Failures
 - Test Structure Processing Has Important "Full Process" Steps
 - Test Structure Area Sufficient to Find Defects
- Limitations:
 - Capacitors Consume Large and Valuable "Real Estate"
 - Standard Bond Pad Layouts Force Non-ideal Designs
 - Need for Rapid Process Characterization Forces Short Loop Processing

■ Sandia National Laboratories

The EQRC's "Six Rules of Oxide Structure Design"

- Minimize Conductor Length
 - Maximize Conductor Width
 - Maximize Gate Contact Area
 - Maximize well/substrate/tub Contact Area
 - Provide a Source of Minority Carriers
 - Choose Appropriate Structure Area
- Other Issues to Consider:
- Edge Effect Structures Where Edge Defects are a Concern
 - Parallel Gate Structures to Mimic Product Transistor Effects
 - Topography to Mimic Product

■ Sandia National Laboratories

Minimize Conductor Length

- Short Conductors Lead To:
 - Low Voltage Drops and Correct Field Calculations
 - No Joule Heating Problems
 - No Fusing or Electromigration of Metal Lines
 - Lower Chances for Interlevel Dielectric Breakdown
- Short Conductor Runs Ensure That the Failure Detected Is That for Which the Structure Was Designed

■ Sandia National Laboratories

Maximize Conductor Width

- Wide Conductors Lead To:
 - Lower Current Densities and Less Chance of EM Failures
 - Less Current Crowding Effects if Corners are Needed

■ Sandia National Laboratories

Maximize Gate Contact Area & Maximize Well/Sub/Tub Contact Area

• Large Contact Area Leads to:

- Better Charge Distribution
- Lower Likelihood of Contact Fusing Problems
- Lower Contact Resistance
- More Accurate Electric Field Determination
- Lower Likelihood of Adjacent Structure Damage Due to Metal Splatter
- Less Gate Oxide Heating Problems

■ Sandia National Laboratories

Provide a Minority Carrier Source

• Minority Carrier Source Leads to:

- Understandable CV Curves
- No Need to Provide Light to Structure
- Ability to Easily Test in Inversion (if wanted)
- A Structure that Mimics a Gate Transistor More Closely

■ Sandia National Laboratories

Choose Appropriate Structure Area

• Appropriate Structure Area Leads to:

- Correct Defect Calculation
- Believable "Intrinsic" Breakdown Field Strengths
- Better Product-to-Structure Correlations

• When Possible, Use of Several Widely Varying Areas is a Good Idea

- Allows Area Effect to be Quantified

■ Sandia National Laboratories

Appropriate Complexity is Needed

• Vary Area

- Small Area
- Large Area

• Vary Topography

- Flat
- Rugged

• Vary Edge Length

- Minimize
- Maximize

• Vary Level of Processing

- Short Loop
- Full Process

■ Sandia National Laboratories

WLR Oxide Test Routines Must Identify Oxide Failures

• We Use JESD-35

- 1 MV/cm-sec
- 5V, 1uA Initial Test
- 5V Post test
- Failure Classification Performed Electrically

• Tests Performed on HP4062UX with EG3001, EG2001, KLA and PW Probers

■ Sandia National Laboratories

Failure Classification Important to Identify Potential Problems

• Initial Failure:

- Due to defect

• Catastrophic Failure (during ramp):

- Typical failure type

• Masked Catastrophic Failure (Vbd undetected):

• Non-catastrophic Failure (Fails but is good on retest)

• Other (Never breaks down)

TYPICAL RESULTS

■ Sandia National Laboratories

Masked Catastrophic Failures Indicate Testing Problems

- Can occur when failure is into a resistive state
 - Most failures are into a highly conductive state
- Can occur when failure current limit is too high
 - Especially true when failure current is defined to be above compliance current

Sandia National Laboratories

Non-catastrophic Failures Usually Indicate Structure Design Problems

- Can occur with metal gates
 - Self-healing failures
 - Failure causes stored energy to eject gate over bad area
- Can occur with long conductor runs
 - Causes line to fuse
- Can occur with small contact areas
 - Causes contacts to open up

Sandia National Laboratories

"Other" Failures Usually Indicate Prober Problems

- Can occur with non-planar probes
- Can occur with insufficient overdrive
- Can occur with unopened passivation
- Can occur with large contact resistance
 - Backside contacts
 - Insufficient contact implant
 - Open contacts

Sandia National Laboratories

Example of Raw TDDDB Data

Sandia National Laboratories

A Few Conclusions

- Correct physical model needed for reliability extrapolation
 - So far E-model seems to be correct
- Correct statistical model needed to correlate test data with product response
 - Tail of the distribution is what limits reliability
 - Extrapolation with EV Model best for Ebd data
- Test structures must be designed to fail due to oxide rupture only
- Test routines must detect failure type if validity of data is to be confirmed

Sandia National Laboratories

END

**DATE
FILMED**

31 9 194