Analyzing the Impact of Process Variations on DRAM Testing Using Border Resistance Traces

Zaid Al-Ars Ad J. van de Goor
Section of Computer Engineering, Faculty of Information Technology and Systems
Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
E-mail: z.e.al-ars@its.tudelft.nl

Abstract: *As a result of variations in the fabrication process, different memory components are produced with different operational characteristics, a situation that complicates the fault analysis process of manufactured memories. This paper discusses the issue of process variations, and shows how to deal with it in the context of fault analysis and test generation. The paper also introduces the concept of border resistance traces as a tool to optimize test stresses and inspect the impact of process variations on the optimization procedure. The concepts are discussed in the paper with the help of a practical example of a specific defect in the memory.*

Key words: *DRAMs, process variations, border resistance trace, defect simulation, memory testing.*

1 Introduction

The tolerance inherent in the fabrication process of memory devices results in memories fabricated with a distribution of device parameters which causes some memories to function at a faster or slower operation speed than others. This difference in operation speed makes it difficult to analyze the faulty behavior of the memory, difficult to generate a suitable test for the observed faulty behavior, and difficult to optimize different stresses for the generated memory test.

In practice, accounting for the impact of fast and slow components is done by statistically analyzing the test results on a large number of parts in order to identify a correlation between a given test and a specific speed range of manufactured components [Vollrath00]. Due to its statistical nature, this method provides limited insight into the specific relation between the applied test, the speed characterization of the tested components and the faulty behavior caused by the targeted defect.

This paper introduces a fault analysis method based on electrical simulation, designed to identify the impact of process variations on the faulty behavior caused by a specific defect. In addition, the concept of border resistance traces is introduced to optimize different test stresses, while taking process variations into consideration. The paper discusses a practical application of the introduced methods to analyze the problem of opens within the memory cell.

This paper is organized as follows. Section 2 discusses the issue of process variations and the way to model them. Section 3 introduces the fault analysis method used to analyze the faulty behavior. Section 4 shows how to use the results of the analysis to generate a test that is independent of process variations. This is followed by the introduction of the concept of border resistance traces to optimize stresses in Section 5. Finally, Section 6 ends with the conclusions.

2 Process variations

During the fabrication process of memory components, the characteristics of the produced components vary in a given range around their target specifications. These variations result in making some components operate slower or faster than the targeted speed of operation. It is important to understand the impact of process variations on the faulty behavior of the memory, since a test should be able to detect the faulty behavior for *any* produced memory component.

The fact that produced components operate with different speeds can be attributed to differences in speed of operation of the manufactured transistors on chip. Therefore, process variations can be modeled and simulated as variations in the transistor parameters used in the electrical simulation model [Foty97]. Since each type of transistor on a chip (there are a number of nFET and pFET transistors with different doping, oxide thickness, etc.) is produced with its own sequence of process steps, there should be a separate set of parameter variations for each type of transistor. This results, however, in a large number of possible variations, and therefore transistors are usually divided into

IEEE COMPUTER SOCIETY

sets that are supposed to have a correlated behavior.

Figure 1 shows how to model process variations considering nFET vs pFET transistor variations. The x-axis represents the measured saturation current of nFET transistors (In), while the y-axis represents the measured saturation current of pFET transistors (Ip). The ellipse represents the process spread around the center. Each point in the ellipse represents one possible combination of drain current for nFETs vs pFETs. Transistors with a drain current at the center of the ellipse (the crossing of the two principal axes) are modeled using the nominal transistor parameters. For each other point in the ellipse, a set of variations in the transistor parameters are added to the nominal parameter set.

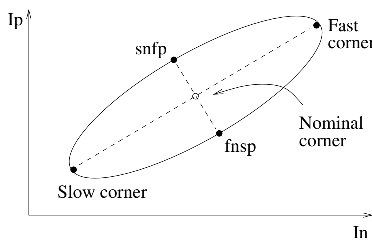

Figure 1. Modeling process variations using variations in transistor parameters for nFETs vs pFETs.

A *process corner* is a term that refers to a point in the ellipse of Figure 1. This means that each process corner has an associated set of transistor parameters describing transistor behavior at that point. In addition to the nominal corner, the figure shows four process corners:

- Corner "snsp" or slow corner—This refers to the process corner where nFETs and pFETs conduct the least amount of drain current. The symbol snsp stands for slow nFET and slow pFET.
- Corner "snfp"—This corner has slow nFETs and fast pFETs. It refers to the process corner with low nFET drain current and high pFET drain current.
- Corner "fnsp"—This corner has fast nFETs and slow pFETs. It refers to the process corner with high nFET drain current and low pFET drain current.
- Corner "fnfp" or fast corner—This refers to the process corner where both nFETs and pFETs conduct the most drain current.

Note that the fast and slow corners (fnfp and snsp) are the so-called 3σ-corners, since they have 3 times the standard deviation of the statistical spread of the fabrication process. Also, note that the distance between the snfp and the fnsp corners is less than the distance between the fnfp and snsp corners. This is caused by the correlation nFETs and pFETs have with each other because they are fabricated in close proximity on the same chip and share some of the fabrication steps [Foty97].

3 Fault analysis method

This section describes the fault analysis method to be used in this paper to analyze the faulty behavior of a defective DRAM. Only a summary is given here, since the method has already been described in the literature [Al-Ars02]. We will consider the defective DRAM cell shown in Figure 2 as an example, where a resistive open (R_{op}) between the access transistor and the cell capacitor limits the ability to control and observe the voltage across the capacitor (V_c). The open is injected into a memory cell and simulated as part of a simplified design-validation model of a real DRAM. The simplified model includes one folded cell array column (2×2 memory cells, load cells, precharge devices and a sense amplifier), one write driver and one data output buffer. The simulations are performed using an electrical Spice-based simulator called Titan, developed by Siemens/Infineon. The simulations performed in this section use the device parameters associated with the nominal corner.

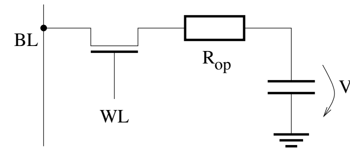

Figure 2. Open injected into memory cell.

The analysis is performed using the DRAM specific operations Act, Wr0, Wr1, Rd, Pre and Nop. First, a number of R_{op} values are selected, where the analysis is to be performed. Then the total faulty behavior of the DRAM is approximated by constructing two result planes that describe the impact of Wr0 and Wr1 operations on V_c, for a range of R_{op} values. Figure 3 shows an automatically generated result plane of Wr0, while Figure 4 shows the result plane of Wr1, for the cell open shown in Figure 2. The x-axes indicates the voltage within the cell using the ratio V_c/V_{dd} (to hide absolute voltage values). The y-axes represents the value of R_{op}, scaled using the scale factor r.

To generate these planes, V_c is initialized to V_{dd} and GND, then the operation sequence 'Wr0 Nop ... Nop' (for the Wr0 plane) or 'Wr1 Nop ... Nop' (for the Wr1 plane) is applied to the cell, which results in a gradual change in V_c toward GND or V_{dd}, respectively. The voltage level after each operation is recorded on each result plane, which results in a number of curves in the planes. We stop the sequence when ΔV_c becomes small enough (0.1 V in this example). V_{cs}, which distinguishes between reading a 0 and a 1, is also shown in the figures as a solid line.

It is possible to use the result planes to analyze a number of important aspects of the faulty behavior [Al-Ars02].

1081-7735/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY

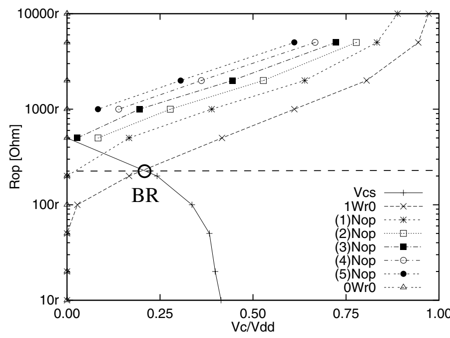

Figure 3. Result plane of Wr0 in the nominal corner (r is a scale factor).

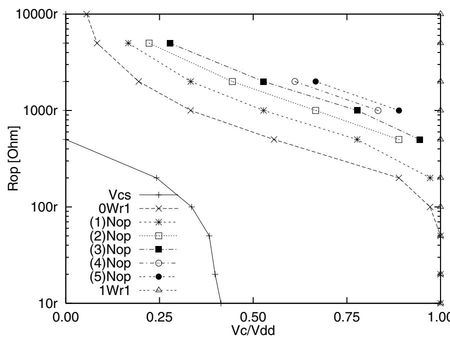

Figure 4. Result plane of Wr1 in the nominal corner (r is a scale factor).

One such aspect relevant to this paper is the *border resistance (BR)*, which is the R_{op} value where the cell *starts* to cause faults on the output. BR is derived on the result planes from the resistive value of the intersection point of the Wr operation curve and the V_{cs} curve. For the faulty behavior shown in Figures 3 and 4, only 1Wr0 intersects V_{cs} at about $210r\Omega$, which also means that BR $\approx 210r\Omega$. It is interesting to note that Wr1 curves do not intersect V_{cs}, which means that Wr1 operations never fail.

Another important aspect relevant to this paper is generating a test that detects the faulty behavior of the defect. Since BR has a value of $210r\Omega$, a test should be able to detect faults for defects with resistance above, but as close as possible to, BR. Inspecting the figure shows that with $R_{op} \geq 210r\Omega$, and with any voltage V_c, the sequence 'Act Wr1 Nop Nop Wr0 Pre Act Rd0' will detect a fault. For $R_{op} = 210r\Omega$, this can be validated by noting that performing 'Wr1 Nop Nop' charges V_c up from any voltage (GND or higher) to approximately V_{dd}. With $V_c = V_{dd}$, performing Wr0 sensitizes the fault which can then be detected by writing a 0 and trying to read it. Therefore, the detection condition $\Updownarrow$(..., Wr1, Nop, Nop, Wr0, Pre, Act, Rd0, ...) detects any faulty behavior for R_{op}.

4 Other process corners

All 5 process corners have been simulated and analyzed at nominal stress conditions (nominal timing and temperature) and for each process corner, the two result plains (Wr0 and Wr1) have been generated. For each process corner, Table 1 lists the two most important outcomes of the analysis: (1) the BR value at which the memory starts to fail, and (2) the detection condition needed to detect the faulty behavior.

Table 1. Summary of results in the 5 different process corners.

Corner	BR	Detection condition
Nominal	$210r\Omega$	$\Updownarrow$(..., Wr1, Nop, Nop, Wr0, Pre, Act, Rd0, ...)
fnfp	$220r\Omega$	$\Updownarrow$(..., Wr1, Nop, Nop, Wr0, Pre, Act, Rd0, ...)
snsp	$200r\Omega$	$\Updownarrow$(..., Wr1, Nop, Nop, Wr0, Pre, Act, Rd0, ...)
snfp	$210r\Omega$	$\Updownarrow$(..., Wr1, Nop, Nop, Nop, Wr0, Pre, Act, Rd0, ...)
fnsp	$205r\Omega$	$\Updownarrow$(..., Wr1, Nop, Nop, Wr0, Pre, Act, Rd0, ...)

The table shows that the value of BR varies between a minimum of $200r\Omega$ for the snsp corner and a maximum of $220r\Omega$ for the fnfp corner, with other process corners having BR values in between. This indicates that the fast corner is the least stressful corner for the faulty behavior, while the slow corner is the most stressful. Furthermore, the detection conditions needed to detect the faulty behavior caused by the defect have the same general structure in all process corners. The only difference is in the number of initializing Nops after the Wr1 operation needed to precharge V_c to V_{dd}.

To ensure detecting the faulty behavior in any process corner, the worst case detection condition (DC) should be used in testing DC = $\Updownarrow$(..., Wr1, Nop, Nop, Nop, Wr0, Pre, Act, Rd0, ...). From a practical point of view, we should consider that real manufactured silicon may deviate from the simulated behavior, such that the exact simulated voltages and currents may differ but the simulated tendencies of the behavior are still valid.

As a result, the exact number of operations needed to charge the memory to the desired voltage may differ. It is possible to increase the test time in an attempt to ensure coverage by increasing the number of Nops in the detection condition, as follows:

- The original DC = $\Updownarrow$(..., Wr1, Nop, Nop, Nop, Wr0, Pre, Act, Rd0, ...)
- Add 1 Nop DC^{+} = $\Updownarrow$(..., Wr1, Nop, Nop, Nop, **Nop,** Wr0, Pre, Act, Rd0, ...)
- Add 2 Nops DC^{++} = $\Updownarrow$(..., Wr1, Nop, Nop, Nop, **Nop, Nop,** Wr0, Pre, Act, Rd0, ...)
- etc.

Keep in mind that the more Nops are added, the less the impact of each Nop becomes.

1081-7735/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY

5 Optimizing stress conditions

In order to test a DRAM, it is not enough to generate a sequence of write and read operations to detect the faulty behavior, but it is also necessary to identify the proper way to optimize different *stress conditions (STs)* so that the memory is forced to fail under the applied test. The criterion for optimizing any ST can be stated as follows: A change in a given ST should modify the value of BR in that direction which maximizes the resistance range that results in a detectable functional fault.

The clock cycle time (t_{cyc}) can be optimized by inspecting the impact of a number of t_{cyc} values on the resulting BR. Inspecting the result planes in Figures 3 and 4, it is clear that BR is specified by the intersection point of the 1Wr0 curve and the V_{cs} curve. Further simulation shows that these two curves remain the same curves that define the value of BR across the desired range of t_{cyc}. Therefore, BR can be inspected by generating only the 1Wr0 and V_{cs} curves, and *tracing* their intersection point, instead of generating a whole range of new result planes for every t_{cyc} value.

Figure 5 shows 5 different BR traces for the different process corners, which indicate the impact of different process corners on the BR value as a function of t_{cyc}. The figure shows clearly that, for a given t_{cyc}, the fnfp corner results in the lowest coverage, while the snsp corner results in the highest coverage. This means that most of the test induced failed parts are located in the snsp corner, particularly when timing is used as a guard band. The fnsp, snfp and nominal corner traces stay close together and not only fall in between the BR traces of fnfp and snsp, but also keep a sizable margin away from them throughout the analyzed range of t_{cyc}. This can be explained by inspecting Figure 1 where it is clear that the fnsp and snfp corners are much closer to the nominal corner than fnfp and snsp are.

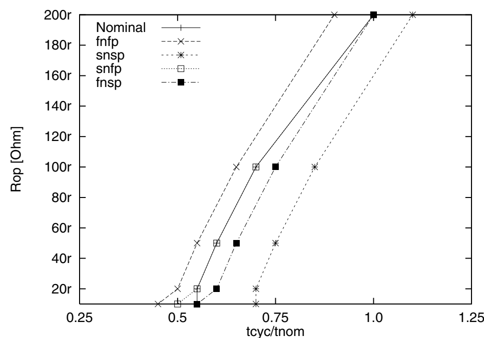

Figure 5. BR trace with t_{cyc} in different process corners.

In the same way, temperature can be optimized by inspecting the impact it has on BR, defined as the intersection of the 1Wr0 and V_{cs} curves. Until now in the paper, simulations have been performed using the room temperature of 27°. It is important to note here that, based on empirical measurements, the open defect analyzed in this paper (as shown in Figure 2) has a negative temperature coefficient (i.e., R_{op} decreases with increasing temperature).

Figure 6 shows the BR trace as a function of temperature in all process corners. The figure indicates that snsp results in the lowest BR values while fnfp results in the highest BR values. This means that most of the test induced failed parts are located in the snsp corner, particularly when temperature is used as a guard band. The nominal corner results in BR values that are between the snsp and fnfp corners. Between $-10°$ C and $27°$ C, BR has the same value for all process corners with the exception of the snsp corner. Note that as temperature increases, BR does not always increase but remains sometimes constant. This happens since the analysis has been performed on a grid in the (Temp, BR) plane.

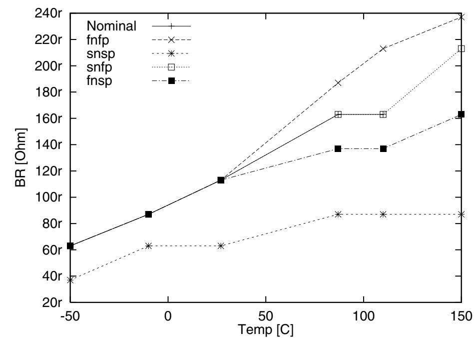

Figure 6. BR trace with temperature in different process corners.

6 Conclusions

This paper discussed the problem of fabrication process variations and the way it impacts fault analysis and testing. A method was presented to analyze the faulty behavior of the memory using defect injection and electrical simulation to generate a test that is independent of process variations. The paper also introduced the concept of border resistance traces, used to optimize different stress conditions for a specific defect, while taking process variations into consideration.

References

[Al-Ars02] Z. Al-Ars and A.J. van de Goor, "Approximating Infinite Dynamic Behavior for DRAM Cell Defects," *in Proc. IEEE VLSI Test Symp.*, 2002, pp. 401–406.

[Foty97] D. Foty, *MOSFET Modeling with Spice, Principles and Practice*, Prentice Hall Inc., New Jersey, 1997.

[Vollrath00] J. Vollrath, "Tutorial: Synchronous Dynamic Memory Test Construction, A Field Approach," *in Proc. IEEE Int'l Workshop Memory Technology, Design and Testing*, 2000.

1081-7735/03 $17.00 © 2003 **IEEE**

IEEE COMPUTER SOCIETY