Transient Faults in DRAMs: Concept, Analysis and Impact on Tests

Zaid Al-Ars Ad J. van de Goor
Section 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, vdGoor@cardit.et.tudelft.nl

Abstract: Memory fault models have always been considered not to change with time. Therefore, tests constructed to detect sensitized faults need not take into consideration the time period between sensitizing and detecting the fault. In this paper, a new class of memory fault models is presented, where the time between sensitizing and detection should be considered. The paper also presents fault analysis results, based on defect injection and simulation, where transient faults have been observed. The impact of transient faults on testing is discussed and new detection conditions, in combination with a test, are given.

Key words: transient faults, functional fault models, DRAMs, defect simulation, detection conditions, memory testing.

1 Introduction

One of the basic assumptions about functional fault models (FFMs) in memories is that they are constant in time. That is, once an FFM is sensitized, the fault effect remains present, unless a subsequent operation is performed that results in eliminating the sensitized fault effect (either by directly overwriting the fault effect, or as a result of sensitizing a different FFM that is linked to this sensitized FFM). As a result of this assumption, tests constructed to detect sensitized FFMs need not take into consideration the time period between sensitizing and detecting the FFM, as long as precautions are taken so that the sensitized FFM is not eliminated [vdGoor98].

FFMs are represented as sets of fault primitives (FPs) [vdGoor90]. An FP is a simple, compact and unambiguous way to associate the faulty behavior exhibited by a given memory cell as a result of a given sensitizing operation sequence (SOS). FPps are defined as $<S/F/R>$, where $S$ stands for the SOS, $F$ for the state of the faulty cell and $R$ for the output on a read operation. In this paper, $S \in \{0, 1, 0u1, 1u0, 0u1, 1u0, 0r1, 1r1\}$, where, for example, the 0 in 0u1 represents the initial value of the cell before the u1 is applied. $F \in \{0, 1, ?\}$, where ? denotes the unknown or random state. $R \in \{0, 1, ?, -\}$, where - means an unpredictable read result, and - denotes the fact that the SOS does not end with a read operation to the victim. For example, the FP $<0u1/0/->$ represents the up-transition fault (TF^+). $S = 0u1$ denotes that the cell contains the value 0, after which a u1 operation is applied to perform an up-transition in the cell. $F = 0$ means that the cell remains in state 0, and $R = -$ means that $S$ does not apply a read operation to the victim.

Previous work has shown that a special variant of an FP exists: $<S/F/R>$, whereby the faulty behavior, sensitized by the SOS $S_F$, occurs only after a given time period $T$. An example of such an FP is a Data Retention Fault (DRF) $<1_F/0/->$ [Dekker90]. This fault causes the cell to loose its value (the value changes from 1 to 0) after a time period $T$. However, the other two components of an FP have, until now, been thought to have a permanent nature that is not time dependent.

Here, we show that some FFMs may, in fact, change in time without the need to perform any external memory operation at all. This change results sometimes in hiding the sensitized faulty behavior. In terms of FPs, we show that under some circumstances the faulty state of the cell ($F$) may have a limited lifetime $L$. Such a time dependent faulty state is denoted by $F_L$. For example, the defective DRAM cell shown in Figure 1, where a resistive element ($R_{def}$) is inserted between the pass transistor and the cell capacitor. If $R_{def}$ has a high enough resistance, the cell would be partially disconnected from the bit line (BL). As a result, a 0u1 operation would not be able to charge the cell up to 1, resulting in a TF^+ $=<0u1/0/->$. Yet, the 0u1 only partially charges the cell, thereby resulting in a weak stored 0. Since the cell has a PMOS pass transistor, leakage currents charge the cell up to 1. Therefore, the stored weak 0 is lost before it is possible to refresh the cell back to its incorrect 0 value. Since TF^+ is restored after some period of time $L$, this FFM is called transient and is
denoted as \( TP^+ = \langle 0w1/0L/- \rangle \). The actual value of the lifetime \( L \) depends on two basic factors:

- the amount of leakage current into the cell, and
- the initial voltage of the weak logic value stored in the cell, which in turn depends on the value of the defect resistance.

![Figure 1. DRAM memory cell with an open defect.](image)

This paper is organized as follows. First, a quantitative analysis of transient faults is made in Section 2. Then, Section 3 explores the space of transient faults. Next, Section 4 describes the consequences transient faults have on testing. Finally, Section 5 ends with the conclusions.

## 2 Quantifying transient faults

This section starts with enumerating the causes for leakage currents in Section 2.1. Then, Section 2.2 analyzes the timing of transient faults.

### 2.1 Leakage current mechanisms

Leakage currents are important to DRAMs because of the cell retention time requirements, which have increased from 16 ms in early 4 Mb DRAMs to as much as 256 ms for some types of low-power DRAM products [Adler95]. For current types of DRAMs, a number of leakage mechanisms exist that result in depleting the charge stored in a DRAM cell.

1. leakage through the pass transistor to the substrate
2. subthreshold leakage through pass transistor to bit line [Hanafi98]
3. leakage between cell capacitor plates through insulator [Romainenko98]
4. leakage between adjacent cell capacitors [Dietl90]
5. other technology specific leakage mechanisms [Adler95]

In a well constructed CMOS DRAM cell, the most important leakage mechanism of these is leakage through pass transistor to substrate (Mechanism 1). However, Hanafi[98] reports that subthreshold leakage current is the main leakage mechanism for silicon-on-insulator (SOI) DRAMs.

### 2.2 Timing of transient faults

In this section, the simulated memory is analyzed to approximate the value of the lifetime \( (L) \) of the observed transient FPs. Figure 2(a) shows a scale of the voltage within a DRAM cell, where a number of important cell voltages are indicated. If a memory is defect free, write operations set the stored voltage within the cell to either \( V_{dd} \) or GND, depending on the stored logic value. \( V_{mp} \) is the mid-point voltage. In a perfectly balanced sense amplifier, a stored voltage above \( V_{mp} \) is read as logic H, while a stored voltage below \( V_{mp} \) is read as logic L. However, parametric mismatches in the sense amplifier can amplify the bit line voltage differential in the wrong direction. Sarapeshkar[91] identifies four important parameters that cause sense amplifier imbalance:

1. transconductance of the sense amplifier transistors
2. gate-source parasitic capacitance of transistors
3. threshold voltage of transistors
4. bit line capacitance

As a result of sense amplifier imbalance, three voltage ranges within the cell can be identified: logic H, logic L and sense amplifier sensitivity (see Figure 2(a)). If a cell contains a voltage in the logic H range \( (V_{dd} \rightarrow V_{th}) \), then a read operation always succeeds in setting the sense amplifier to 1 (cell on BT) or to 0 (cell on BC). If a cell contains a voltage in the logic L range \( (V_{lo} \rightarrow GND) \), then a read operation always succeeds in setting the sense amplifier to 0 (cell on BT) or to 1 (cell on BC). If the cell contains a voltage in the sense amplifier sensitivity range \( (V_{th} \rightarrow V_{lo}) \), then the effect of a read operation is unpredictable; it depends on the parametric mismatches in the sense amplifier.

After a write operation has set the cell voltage to GND, leakage currents through the PMOS pass transistor gradually charge the cell up. The cell should be refreshed before the cell voltage reaches \( V_{lo} \) in order to maintain the stored logic value. Therefore, the refresh time \( (t_{ref}) \) is less than the time needed for leakage currents to charge the cell up from GND to \( V_{lo} \). If the memory is defective, write operations may not succeed in setting the cell voltage to GND, which means that it would take leakage currents less time to charge the cell up to \( V_{lo} \).

In order for transient FPs to eliminate the fault effect of a cell, leakage currents should be able to charge the cell up from \( V_{lo} \) to \( V_{th} \) in a time period \( L \leq t_{ref} \). This \( L \) is considered to be the minimum transition time for a transient FP. Next, we will attempt to calculate the value of \( L \) for the DRAM under consideration.

Consider the memory structure shown in Figure 2(b), where the following parameters are given [Al-Ars99].
• Storage cell capacitance: \( C_c = 50 \text{ fF} \)
• Logic 1 on BT is \( V_c = 2.4 \text{ V} \), and 0 is \( V_c = 0 \text{ V} \)
• Reference cell capacitance: \( C_r = 50 \text{ fF} \)
• Mid-point voltage in reference cell is \( V_r = 1.2 \text{ V} \)
• Bit line capacitance: \( C_b = 300 \text{ fF} \)
• Bit line precharge voltage is \( V_b = 2.4 \text{ V} \)

When a read operation is performed, the reference cell is connected to the complementary bit line. The voltage on the bit line stabilizes at

\[
V_b' = \frac{Q_r + Q_b}{C_r + C_b} \cdot \frac{1.2(50) + 2.4(300)}{50 + 300} = 2.23 \text{ V}
\]

At the same time, the memory cell to be read is connected to the true bit line. Depending on the content of the cell, the bit line voltage stabilizes at

Logic 1: \( V_b' = \frac{Q_c + Q_b}{C_c + C_b} \cdot \frac{2.4(50) + 2.4(300)}{50 + 300} = 2.4 \text{ V} \)

Logic 0: \( V_b' = \frac{0(50) + 2.4(300)}{50 + 300} = 2.06 \text{ V} \)

As shown in Figure 2, there is a linear relation between the stored cell voltage and the sense amplifier voltage. The slope of this linear relation is given by \( \frac{2.4 - 2.06}{2.23 - 2.06} = 7.06 \).

Assuming that the sensitivity of the sense amplifier is given to be \( \pm 50 \text{ mV} \) centered around 2.23 V, this gives a sensitivity range of 2.28V \( \rightarrow \) 2.18V for the sense amplifier.

This means that the voltage sensitivity limits within the cell are \( V_{hi} = 7.06(2.28 - 2.06) = 1.55 \text{ V} \) and \( V_{lo} = 7.06(2.18 - 2.06) = 0.85 \text{ V} \). Therefore, leakage currents have to charge up the cell from 0 V to 0.85 V in order to destroy a stored logic 0. Now, assume that it takes \( t_{ref} = 60 \text{ ms} \) for this charging up to take place, and that we approximate the charge up process as linear. Then, in order for leakage currents to pull the stored voltage from \( V_{lo} \) to \( V_{hi} \) (a voltage of 0.7 V), it takes

\[
L_{min} = \frac{V_{hi} - V_{lo}}{V_{lo} - V_{hi}} \cdot t_{ref} = 0.7 \cdot 0.85 \cdot 60 \text{ ms} = 50 \text{ ms}
\]

Therefore, a transient fault has a lifetime of about 50 ms or more (i.e., \( L \geq 50 \text{ ms} \)). In order to detect a transient fault before it is corrected, it is important to perform a detecting read operation within 50 ms of sensitizing the fault. This condition, however, does not ensure detecting transient faults since within the sensitivity region, the sense amplifier may give a correct or an incorrect readout of the stored cell voltage. The following condition ensures that all transient faults are detected.

To ensure that a transient fault is detected, a detecting read operation should directly be performed after sensitizing each fault.
3 Space of transient faults

This section starts with Section 3.1 where two types of SOS’s are identified, detecting and non-detecting. Section 3.2 analyzes the space of all possible single-cell transient FPs. Then, the proposed transient FPs are validated in Section 3.3.

3.1 Types of SOS’s

From a testing point of view, transient faults may affect the way detection conditions are constructed, which in turn modifies the resulting memory test. In order to analyze the impact of transient faults on the FP space, we divide SOS’s into two types: detecting and non-detecting. A detecting SOS is a sequence that sensitizes a fault and makes it detectable on the output. A non-detecting SOS is a sequence that only sensitizes the fault but does not forward the fault to the output. This definition can be given in a more formal way, as follows. Let $S$ be the SOS in $<S/F/R>$, where $S$ ends with an operation or state having $d$ as the expected data in the victim.

Definition: If $R = d$ or $R = d$ then $S$ is referred to as a non-detecting SOS. Otherwise, if $R 
eq d$ and $R 
eq d$ then $S$ is referred to as a detecting SOS.

As an example, the deceptive read destructive fault DRDF$_1$ ($<1r1/0/1>$) describes a fault that is only sensitized within the cell, but not forwarded to the output ($R = d = 1$) [Adams96]. Therefore, the SOS of DRDF$_1$ is a non-detecting SOS, because $d = 1$, which is the expected data of the $r1$ operation of $S$. Based on this definition, the impact of transient faults on testing can be described as follows:

- Non-detecting SOS: The detection of fault effects sensitized by a non-detecting SOS is done by a read operation, which may be part of the next march element, assuming that the fault effect is permanent. In case the fault effect is transient, the read operation has to be appended to the SOS, which sensitizes the fault effect. For example, to detect a permanent DRDF$_1$ = $<1r1/0/1>$, it is possible to perform either $\text{U}(\ldots, r1, r1, \ldots)$ or $\text{U}(\ldots, r1) \text{U}(r1, \ldots)$. To detect a transient DRDF$_1$ = $<1r1/0l/1>$ then it is necessary to perform $\text{U}(\ldots, r1, r1, \ldots)$.

- Detecting SOS: The detection of the fault effect sensitized by a detecting SOS does not require any extra read operation. Because of the fact that the fault is detected by the SOS sensitizing the fault effect, one does not have to distinguish between permanent and transient faults.

3.2 Space of transient FPs

The idea of transient FPs implies that after a fault is sensitized a mechanism results in correcting this faulty behavior before it is detected on the output. An FP has two components to describe a fault: $F$ (the value of the faulty cell) and $R$ (the output on a read operation). The value of $R$ may not be transient, since it is sensitized and detected on the output at the same time. However, as it has already been shown, it is possible for $F$ to change in such a way that a faulty state would be transformed into a proper state within the cell.

For example, the FP $<1lw1/0/0>$ represents a non-transition write 1 operation that sets a faulty 0 into the memory cell. Under the assumption that faults are constant in time this FP would remain sensitized until a read operation is performed to detect this fault on the output. Under the assumption of transient FPs, this FP would only remain sensitized for a limited period of time and then the state of the cell gets automatically corrected to its expected value (i.e., logic 1). As a second example, consider the FP $<1r1/1/0>$ that represents a faulty read operation, resulting on an incorrect value on the output. It cannot be transient since the cell is in a correct state and the output is sensitized and detected simultaneously.

Table 1 shows all single-cell static and some single-cell two-operation dynamic FPs [AI-Ars00]. Static FPs have an SOS with no performed operations ($S \in \{0, 1\}$, meaning that the state of the victim is specified), or contains one operation ($S \in \{0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1\}$). Dynamic FPs have an SOS with more than one performed operation. In our case, FPs with the following two-operation SOS have been analyzed: $S \in \{0w0r0, 1w0r0, 0w1r1, 1w1r1\}$. The FFM column in the table lists the name of the fault model, where an $L$ subscript indicates that the FFM may be transient. The FP column lists the FP of the fault, while the way these faults behave under the assumption of transient faults is listed in 'After time $L$'. $L$ represents the time needed for the described behavioral change to take place. The 'Comment' column lists whether the fault is transient or not; in case the fault is transient, it indicates whether this transient behavior has been observed in the analysis performed in Section 3.3. Note that the transient FPs in the table that describe a read destructive behavior (e.g., $<1r1/0/0>$) do not result in hiding the whole faulty behavior, but only the internal component of it. This does not result in a proper behavior, but results in modifying the fault into a different fault.

3.3 Practical validation

In order for transient faults to take place in practice, there should be a memory mechanism that works against the fail-
Table 1. Transient faults from single-cell static and 2-operation dynamic FPs.

<table>
<thead>
<tr>
<th>FFM</th>
<th>FP</th>
<th>After time $L$</th>
<th>Comment</th>
<th>FFM</th>
<th>FP</th>
<th>After time $L$</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF0L</td>
<td>$&lt;0/1L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Not observed</td>
<td>SF1L</td>
<td>$&lt;1/0L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Not observed</td>
</tr>
<tr>
<td>WDF0L</td>
<td>$&lt;0w0/1L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>WDF1L</td>
<td>$&lt;1w1/0L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
</tr>
<tr>
<td>TFI0L</td>
<td>$&lt;1w0/1L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>TFI1L</td>
<td>$&lt;0w1/0L/-/-&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
</tr>
<tr>
<td>IRF0</td>
<td>$&lt;0r0/0/1/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
<td>IRF1</td>
<td>$&lt;1r1/1/0/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
</tr>
<tr>
<td>DRDF0L</td>
<td>$&lt;0r0/1L/0/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>DRDF1L</td>
<td>$&lt;1r1/0L/1/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
</tr>
<tr>
<td>RDF0</td>
<td>$&lt;0r0/1/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>RDF1</td>
<td>$&lt;1r1/0/&gt;$</td>
<td>Not transient</td>
<td>IRF1L</td>
</tr>
<tr>
<td>IRF00</td>
<td>$&lt;0w0r0/0/1/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
<td>IRF0L</td>
<td>$&lt;0w1r1/1L/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
</tr>
<tr>
<td>IRF10</td>
<td>$&lt;1w0r0/0/1/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
<td>IRF1L</td>
<td>$&lt;1w1r1L/&gt;$</td>
<td>No change</td>
<td>Not transient</td>
</tr>
<tr>
<td>RDF00</td>
<td>$&lt;0w0r0L/1/&gt;$</td>
<td>Not transient</td>
<td>RDF1</td>
<td>$&lt;1w1r1/0/&gt;$</td>
<td>Not transient</td>
<td>RDF1L</td>
<td>$&lt;1w1r1/1/&gt;$</td>
</tr>
<tr>
<td>RDF10</td>
<td>$&lt;1w0r0L/1/&gt;$</td>
<td>Not transient</td>
<td>RDF0</td>
<td>$&lt;0w1r1/0/&gt;$</td>
<td>Not transient</td>
<td>RDF0L</td>
<td>$&lt;0w1r1/1/&gt;$</td>
</tr>
<tr>
<td>DRDF00L</td>
<td>$&lt;0w0r0/1L/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>DRDF1L</td>
<td>$&lt;1w1r1L/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
</tr>
<tr>
<td>DRDF10L</td>
<td>$&lt;0w1r1/0L/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
<td>DRDF10L</td>
<td>$&lt;1w0r0/1L/&gt;$</td>
<td>Proper behavior</td>
<td>Observed</td>
</tr>
</tbody>
</table>

An analysis has been made by simulating the faulty behavior of injected open defects into a memory model to validate which of the proposed transient fault models actually take place. Table 2 gives the count (#) or the number of different defects that result in sensitizing the transient FP in this analysis on the true (BT) and the complementary (BC) bit lines. For example, SF0L has a 0 count (# = 0) on BT and a 0 count on BC, which means that the transient fault SF0L has not been observed in cells connected to BT nor to BC. WDF0L has not been observed in cells on BC (count is 0), while it has been observed in cells on BC as a result of 5 different injected defects (count is 5). Explicitly indicating BT and BC is important because of the bias leakage currents have towards restoring a given voltage level back into the cell. The table shows clearly that specific types of transient FPs take place in cells on BT (those with faulty 1) or on BC (those with faulty 0). In addition the count of both transient faults on BT and on BC is equal, since transient FP instances on BC are the complementary of those on BT (they are caused by complementary defects). Note that although it is theoretically possible for state faults (SF0 and SF1) to be transient, this behavior has not been observed.

Table 2. Transient FFM's observed by simulation.

<table>
<thead>
<tr>
<th>FFM</th>
<th># on BT # on BC</th>
<th>FFM</th>
<th># on BT # on BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF0L</td>
<td>0 0</td>
<td>SF1L</td>
<td>0 0</td>
</tr>
<tr>
<td>WDF0L</td>
<td>0 5</td>
<td>WDF1L</td>
<td>5 0</td>
</tr>
<tr>
<td>TFI0</td>
<td>9 0</td>
<td>TFI1</td>
<td>9 0</td>
</tr>
<tr>
<td>DRDF0L</td>
<td>0 5</td>
<td>DRDF1L</td>
<td>5 0</td>
</tr>
<tr>
<td>DRDF10L</td>
<td>0 3</td>
<td>DRDF10L</td>
<td>3 0</td>
</tr>
<tr>
<td>DRDF010L</td>
<td>0 6</td>
<td>DRDF010L</td>
<td>6 0</td>
</tr>
</tbody>
</table>

In addition, static and dynamic incorrect read faults (IRF1 and IRF2) and read destructive faults (RDF1 and RDF2) are not given in the table since they cannot be transient.

4 Impact on tests

Memory testing assumes that a sensitized FFM remains sensitized unless an action has been taken that eliminates the fault. In the specific case of DRAMs, where leakage currents play a significant role, testing assumes that FFM's are always sensitized as strong faults such that the fault remains sensitized by memory refresh until a read operation detects the sensitized FFM. This assumption does not hold for transient faults, which need specific detection conditions to ensure their detection.

Only those transient faults that have a non-detecting SOS need special action to detect them. Table 3 lists all analyzed single-cell transient and non-transient faults and gives the needed detection conditions that ensure their detection. Note that, in case of a transient fault, the detection...
March T has four march elements, each of which performs either a transition or a non-transition write operation as required in Table 3. Furthermore, each march element contains two read operations to sensitize and then detect deceptive read destructive faults. In addition to transient FPs, March T detects the non-transient forms of the targeted FPs, since transient FPs need more restrictive detection conditions for their detection than non-transient FPs.

## 5 Conclusions

In this paper, the new class of transient faults has been presented, where the time between sensitizing and detecting the fault plays an important role in testing. A quantitative analysis of transient faults has been made whereby the lifetime of the fault has been calculated. The paper also analyzed the space of transient faults and validated their existence in commercial DRAMS using simulation. The impact of transient faults on testing has been discussed and a test to detect them has been given.

#### References


