Testing for Tunneling Opens

James C.M. Li and Edward J. McCluskey

Center for Reliable Computing
Stanford University

Abstract

A tunneling-open failure mode is proposed and carefully studied. A circuit with a tunneling open could pass at-speed Boolean tests but fail VLV testing or I_{DDQ} testing. Theoretical calculations as well as Boolean and I_{DDQ} experiments confirm the existence of tunneling opens. The Murphy experimental data show that seven out of nine VLV-only failure circuits can be explained by this failure mode. All these seven circuits survived 366 hours temperature burn-in. Finally, a cost effective screening strategy is proposed.

1. Introduction

A *tunneling open* is a very thin open that allows electrons and holes to tunnel through. Circuits with tunneling opens were first reported by [Henderson 91]. In that experiment, a large current was forced to flow through a metal wire and produced a narrow (about 100Å) opening at the input of an inverter. The logarithm of the tunneling current was found to be proportional to the voltage across the opening. The circuits only operated at very slow speeds (110KHz maximum) because the tunneling current was small.

Here we present a careful analysis of the tunneling-open failure mode. Theoretical calculations and experimental results show that a circuit with a tunneling open can pass at-speed tests at nominal voltage but slow down significantly at very low voltage. A burn-in experiment shows that the circuits with tunneling opens may not be detected by temperature burn-in. However, the existence of tunneling opens might indicate process problems. At the end of the paper, a cost effective screening strategy is proposed.

Very-low-voltage (VLV) testing is defined as Boolean testing performed at a very low supply voltage (around 2 to 2.5 times V_t) [Hao 93][Chang 96b]. VLV testing has been shown theoretically to be able to detect weak chips that contain *flaws* which do not cause functional failure at nominal operating condition but may cause intermittent or early-life failure. Table 1 compares the effectiveness of VLV, I_{DDQ} testing, temperature and voltage burn-in. Five failure modes are discussed as follows. (1) *Threshold voltage shift* is caused by hot carrier effects or process variation. This failure mode causes the circuit delay to increase significantly at very low voltage and therefore can be detected by VLV testing [Chang 96a]. This failure mode could also cause the background leakage current to increase and hence be detected by I_{DDQ} testing. Voltage burn-in is effective in accelerating the hot-carrier injection effect because the electrons are accelerated to a higher energy as the V_{DD} increases [Hnatek 95][Leblebici 93]. (2) *Gate oxide shorts* are caused by defective gate oxide. One simple model of this failure mode is a resistive short (between transistor gate and source or drain) which can cause excessive I_{DDQ} current. It has been shown that VLV testing is effective in detecting gate oxide shorts [Chang 96a]. Burn-in is also effective in screening this failure mode because high temperature and high voltage accelerate the oxide degradation. (3) *Metal shorts* are unexpected shorts between metal wires. This failure mode may cause high I_{DDQ} and it may also cause circuit delay to increase significantly at very low voltage. It has been reported that some circuits with metal shorts failed after burn-in [Righter 98].

Table 1. Effectiveness of VLV, I_{DDQ}, temperature and voltage burn-in

Failure Mode	VLV	I_{DDQ}	Temp. Burn-in	Volt. Burn-in
(1) V_t shift	Y	Y	*	Y
(2) Gate oxide shorts	Y	Y	Y	Y
(3) Metal shorts	Y	Y	Y	Y
(4) High resistance interconnect	N	N	**	**
(5) Tunneling opens	Y	***	N	?

* low temperature is effective [Hnatek 95]
** depends on the cause ** depends on defect thickness

(4) *High resistance interconnect* can be caused by various mechanisms such as electronmigration, stress

voids, or defective vias. This failure mode introduces extra RC delay which can be detected by delay testing. On the other hand, VLV testing does not detect this failure mode because RC delay does not scale when the supply voltage is reduced. Neither can I_{DDQ} detect this failure mode. Burn-in is effective in some cases (such as electronmigration) but ineffective in the others (such as silicide open [Tseng 00]). (5) In this paper, it will be shown that tunneling opens can pass at-speed tests at nominal voltage but fail VLV testing. Whether I_{DDQ} testing is effective or not depends on the defect thickness. It will be shown experimentally that temperature burn-in is not effective in screening tunneling opens. Whether voltage burn-in is effective is unknown so far.

The data shown in this paper are collected from the Murphy experiment [Franco 95] [Chang 98 ab] [McCluskey 2000]. A digital CMOS test chip was built in 0.7µ technology with gate oxide thickness = 200Å. It has five combinational *circuits under test (CUTs)* and the total gate count is 25K. In this experiment, nine chips (out of 5,500 tested) were found to have *VLV-only failures*. These chips passed millions of test patterns (including single stuck-at fault, transition fault, path or gate delay fault and exhaustive patterns) at nominal voltage (5V) but failed some tests at VLV (1.7V).

Now the question arises: can the failure modes listed in Table 1 explain all the VLV-only failures? To answer this question, $I_{DDQ}(t)$ testing was performed on those VLV-only failure CUTs. *$I_{DDQ}(t)$ testing* is defined as making multiple continuous I_{DDQ} measurements (for every test pattern) to observe the change in I_{DDQ} over time. Figure 1 shows the experimental results from one of the VLV-only failure CUTs. The chip was powered up and its I_{DDQ} current was allowed to settle before time "zero." At time zero, a test pattern was applied and 20 continuous I_{DDQ} measurements were taken (without changing the test pattern). For a good circuit, the $I_{DDQ}(t)$ values should be constantly low (less than 1µA). But for this VLV-only failure CUT, some test patterns caused an *$I_{DDQ}(t)$ drift over time* phenomenon, which is present if the I_{DDQ} values change significantly during the interval of observation. How fast the I_{DDQ} drifts down to zero is a function of supply voltage V_{DD} (section 4.2).

It has been reported that a circuit with an open defect may cause such an $I_{DDQ}(t)$ drift over time [Maly 88] [Soden 89]. In both cases, the opens, caused by missing metal or polysilicon wires, are too large to permit tunneling. In Maly's experiment, an open defect was artificially injected in the source of a transistor. They explained this $I_{DDQ}(t)$ drift phenomenon by the slow charging process of the reverse biased source-well junction. It took 15 seconds for the circuit to switch. In Soden's experiment, the circuit contained an open which caused the circuit to fail some tests. However, neither of these two cases mentioned above can explain the VLV-only failures in the Murphy experiment because a circuit with a large open could not pass so many thorough at-speed tests at nominal voltage. Some via or contact defects have been reported in [Needham 98] [Campbell 91] where the defects are modeled as high resistance interconnects (resistive opens). However, resistive opens cannot explain the observed $I_{DDQ}(t)$ drift over time phenomenon.

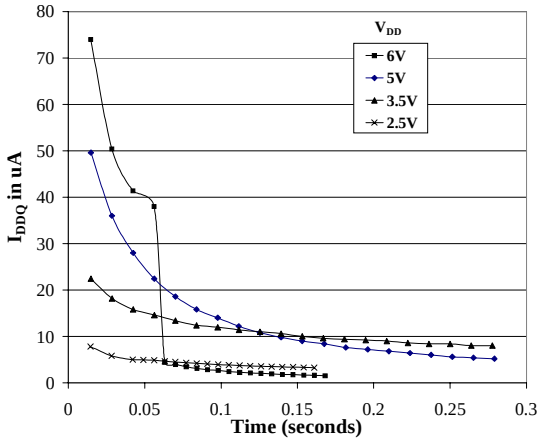

Figure 1. $I_{DDQ}(t)$ drift over time

The other explanation for the $I_{DDQ}(t)$ drift phenomenon of the Murphy chips could be a tunneling open. The tunneling open proposed here is much thinner (10-20Å) than that in Henderson's experiment. Such a narrow opening could be present at a via or a contact due to the incomplete oxide etching. The tunneling current through such a narrow opening is large enough to make the circuit operate at speed at nominal voltage. However, due to the exponential dependence of the tunneling current on the electric field across the opening, the circuit fails at very low voltage.

Table 2 shows the test results from nine VLV-only failure CUTs. They are classified into three categories according to their faulty behavior. The first two CUTs fail VLV testing (see section 4.1). These two CUTs have high and constant I_{DDQ} (section 4.2), and they failed after 6 hours of temperature burn-in (section 4.3). Their faulty behavior cannot be explained by tunneling opens. However, for the other seven CUTs, they operate at very slow speed at very low voltage. Their $I_{DDQ}(t)$ values either stay at a constant low value or drift down over time. Their behavior can be explained well by tunneling opens. All these seven CUT were unchanged after 366 hours of burn-in.

Table 2. Test results of nine VLV-only failure CUTs

category	CUT #	Speed @1.7V (sec. 4.1)	high I_{DDQ}? (sec. 4.2)	$I_{DDQ}(t)$ drift? (sec. 4.2)	fail after burn-in? (sec. 4.3)	explained by tunneling open?
I	1	hard failure	Yes	No	Yes	No
	2	hard failure	Yes	No	Yes	No
II	3	Very slow	No	No	No	Yes
	4	Very slow	No	No	No	Yes
III	5	Very slow	Yes	Yes	No	Yes
	6	Very slow	Yes	Yes	No	Yes
	7	Very slow	Yes	Yes	No	Yes
	8	Very slow	Yes	Yes	No	Yes
	9	Very slow	Yes	Yes	No	Yes

The organization of this paper is as follows. Section 2 introduces the basic physics of the tunneling effect. Section 3 first qualitatively explains the VLV-only failure by a tunneling open and then performs calculations to prove it. Section 4 provides a prediction of the behavior of the faulty circuits and then verifies with experimental results. Section 5 discusses some questions associated with tunneling opens. Finally, section 6 summarizes the paper.

2. Physics of the Tunneling Effect

When an insulator is thin enough, it is possible for electrons and holes to tunnel through it. The most commonly seen tunneling effects in CMOS technology can be classified into three categories according to their physical mechanisms. They are described in the following sections.

2.1 Trap-Assisted Tunneling

This tunneling effect is assisted by the traps which are generated by the impurities in the oxide. This effect happens even at a fairly low electric field. The magnitude of this tunneling current depends on the quality of the oxide. The typical values lie in the range of 10^{-4} to 10μA/cm^2 [Mozzami 92][Gupta 97].

2.2 Fowler-Nordheim Tunneling

If the *electric field across the oxide (E_{ox})* is strong enough, it causes band bending which makes the barrier thinner. The tunneling current therefore gets higher as the E_{ox} gets stronger. This field dependent tunneling effect is called Fowler-Nordheim tunneling and can be quantitatively modeled by the following equation [Fowler 28].

$$J_{FN} = \alpha \cdot Eox^2 \cdot e^{(-\beta/Eox)} \qquad (1)$$

In the case of a metal/oxide/silicon structure, α = $9.92\text{x}10^{-7}$ AV^{-2} and β = $2.635\text{x}10^{8}$ Vcm^{-1} [Weinberg 82].

2.3 Direct Tunneling

Direct tunneling can occur in a very thin oxide at very low electric field. The direct tunneling current can be modeled by multiplying the Fowler-Nordheim tunneling current with a correction factor as in the following equation [Schuegraf 92],

$$J_{DT} = J_{FN} \cdot \left(1 - \sqrt{(\phi_b - V_{ox})/\phi_b}\right)^{-2} e^{\beta\,[(\phi_b - V_{ox})/\phi_b]^2 / Eox} \qquad (2)$$

where ϕ_b is the barrier height. Equation (2) is applicable when the *voltage drop across the oxide (V_{ox})* is less than the barrier height.

2.4 Calculations

Figure 2 shows the tunneling current density for different silicon dioxide thicknesses (i.e., the thicknesses of the tunneling open) as a function of the *voltage across oxide (V_{ox})*. The numbers were obtained from equations (1) and (2), assuming an Al/oxide/n$^+$-polysilicon structure. In reality, the actual numbers may vary with the material and the quality of oxide. However, the shape of the curves should be similar. This figure shows that different tunneling effects dominate different V_{ox} regions. Take the 20Å curve for example, if V_{ox} is higher than 3V, FN-tunneling dominates and the logarithm of the tunneling current increases with V_{ox}. For V_{ox} between 3V and 1V, direct tunneling current dominates over FN tunneling current. For V_{ox} lower than 1V, trap-assisted tunneling current becomes the most significant.

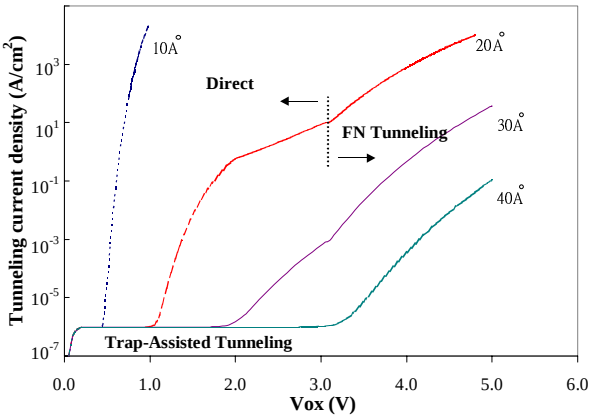

Figure 2. Tunneling current density for different oxide thickness (calculated)

3 Circuit Behavior with a Tunneling Open

3.1 Qualitative Description of a Tunneling Open

Figure 3 represents a tunneling open located at a contact hole. The metal and poly which should have been connected are separated by a very thin layer (10-20Å) of oxide. This kind of defect could be caused by incomplete etching or native (room temperature grown) oxidation after etching. Similar defects can also occur at a via where a thin air gap exists between the tungsten plug and metal wire. In the case of an air gap open, the tunneling current is smaller than the oxide case because the dielectric constant of air is about a third of that of oxide.

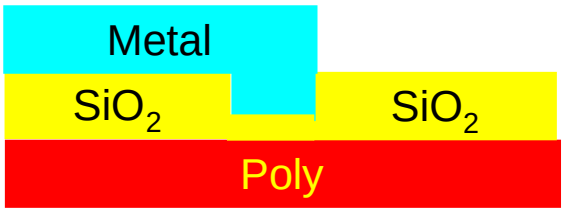

Figure 3. A tunneling open located at contact (not to scale)

Figure 4 shows an example circuit with a tunneling open located at the input of an inverter. As the voltage on the metal (V_m) ramps up, the electric field across the opening will cause the tunneling current (J_{tunnel}) to charge the gate capacitor (C_{gate}). The voltage on the poly (V_p) therefore rises. At nominal voltage, the field dependent tunneling current is large enough to switch the inverter in a very short time. However, at very low voltage, the tunneling current is so small that the inverter takes a very long time to switch. This explains why a circuit with a tunneling open can pass at-speed tests at nominal voltage but fail at very low voltage.

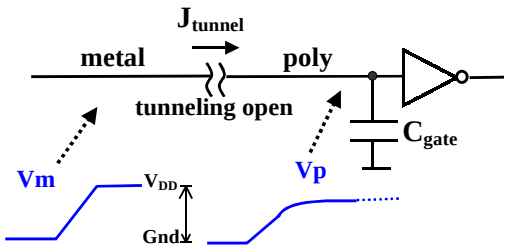

Figure 4. An example circuit with a tunneling open

When V_p stays at an intermediate voltage level, the PMOS and NMOS of the inverter are both turned on and high I_{DDQ} current is observed. As V_p continues to rise to V_{DD}, the electric field across the defect decreases and causes the tunneling current to diminish. Trap-assisted tunneling which happens at low electric field slowly charges the gate capacitor. The voltage on poly will gradually rise to V_{DD} and the I_{DDQ} will gradually drift down to zero. This slow charging process explains why a circuit with a tunneling open has very long $I_{DDQ}(t)$ drift over time phenomenon. Note that if the tunneling open is very thin, V_p may rise so fast that high I_{DDQ} is not measurable. Therefore the observation of this phenomenon depends on the defect thickness.

Although it is not shown in the figure, the same arguments also apply to the opposite case in which V_m ramps down. However, the fall time is different from the rise time due to the polarity of the tunneling effect [Shi 98].

3.2 Theoretical Analysis

Consider the circuit of Fig. 4. Its voltage behavior and $I_{DDQ}(t)$ drift behavior are calculated in sections 3.2.1 and 3.2.2 separately.

3.2.1 Voltage Behavior

Assume that the tunneling open has area A and thickness d. The *capacitance across the tunneling open* C_{defect} is $\varepsilon_{ox}A/d$. The *total capacitance of the gate capacitance* is C_{gate}. If there is no tunneling effect, the total charge on the poly is conserved. The voltage on the poly due to the coupling effect can be expressed as:

$$V_{p\_no\_tunneling}(t) = V_m(t) \cdot \frac{C_{defect}}{C_{defect} + C_{gate}} \qquad (3)$$

However, if the tunneling effect is taken into account, the total charge on the poly increases as the tunneling current flows through the open to charge the gate capacitors. In this case, the voltage on the poly can be expressed as:

$$V_{p\_tunneling}(t) = \left[V_m(t) + \frac{A\int J_{tunnel}(t)dt}{C_{defect}} \right] \cdot \frac{C_{defect}}{C_{defect} + C_{gate}} \qquad (4)$$

Compared with equation (3), equation (4) has one more term which corresponds to the total charge that tunnels through the open. The following calculations will demonstrate that this term cannot be ignored if the tunneling effect is significant.

Figure 5 illustrates the calculated waveforms for the voltage on poly (solid lines) given a ramp input voltage on the metal (dotted line) at nominal voltage. Based on the process technology of the Murphy experiment [LSI 93], the following numbers are assumed: nominal supply voltage V_{DD} = 5V, gate capacitance C_{gate} = 40fF, defect area A = 1μm^2 and defect thickness d = 12Å. If no tunneling effect is considered (equation 3), the $V_{p\_no\_tunneling}$ (thin solid line) eventually settles at 2V and the inverter is not switched (assuming logic threshold = $V_{DD}/2$ = 2.5V). However, if the tunneling effect is considered (equation 4), the $V_{p\_tunneling}$ (thick solid line) eventually goes higher than 2.5V and the inverter is switched. This calculation shows that the tunneling

effect cannot be ignored when the open is as thin as 12Å.

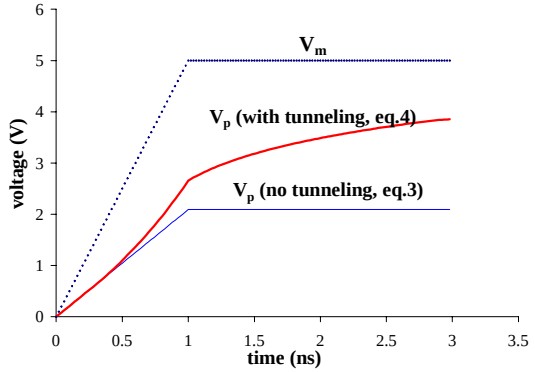

Figure 5. Voltage waveforms in circuit of Fig. 4 (V_{DD}=5)

Figure 5 illustrates two issues in testing tunneling opens. One is that tunneling and coupling effects take place so fast that no significant delay can be observed (only 0.5 ns delay from V_m = 2.5V to V_p = 2.5V). It is therefore very difficult to detect the tunneling open by transition or path/gate delay tests at nominal voltage. The other thing is that the V_p may not stay below (V_{DD}-Vt) for a long time. It depends on the thickness of the defect. Therefore it is not guaranteed that I_{DDQ} testing can detect a tunneling open.

In the Murphy experiment, the supply voltage of VLV testing is two times the transistor threshold voltage ($2V_t$ = 1.7V) [Chang 96b][Chang 98b]. Figure 6 shows the calculated waveforms at 1.7V given the same circuit as in Fig. 4. In this case, the voltages on poly calculated from eq.3 and 4 are almost the same (thin and thick lines overlap). The tunneling effect fails to boost the V_p above the logic threshold ($V_{DD}/2$ = 1.7V/2 = 0.85V) because the tunneling (FN or direct) current is small at low voltage. Although the circuit fails to operate at-speed, it eventually can operate correctly with the help of the small trap-assisted tunneling current given a very long wait time.

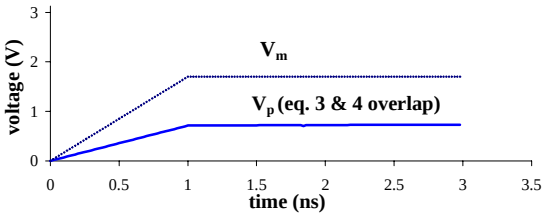

Figure 6. Voltage waveforms in circuit of Fig. 4 (V_{DD}=1.7)

3.2.2 I_{DDQ}(t) Drift Behavior

Consider the same circuit as shown in Fig. 4. Two assumptions are made: 1) the I_{DDQ}(t) drift over time phenomenon is dominated by the trap-assisted tunneling current (J_{TA}), and 2) the trap-assisted tunneling current remains constant. These assumptions, though not accurate, give a reasonably accurate approximation of J_{TA} as shown in the following equation,

$$J_{TA} = \frac{C_{gate} \cdot \Delta V_p}{A \cdot T_{drift}} \qquad (5)$$

where T_{drift} is defined as the time interval from the moment when trap-assisted tunneling current dominates to the moment when I_{DDQ}(t) drifts down to zero. ΔV_p is the voltage change of V_p during the time interval T_{drift}.

In the nominal voltage simulation (Fig. 5), V_p rises slowly after reaching 4V which represents the end of FN (direct) tunneling effect and the beginning of trap-assisted tunneling effect. Before V_p reaches (V_{DD}-Vt), both PMOS and NMOS are turned on. Consequently, high I_{DDQ} is observed at the inverter. As the Vp rises from 4V to (V_{DD}-Vt), the I_{DDQ} gets smaller and finally drifts down to zero. Assuming that V_t is 0.8V, the voltage change of V_p before the I_{DDQ}(t) drifts down to zero is therefore (V_{DD} - V_t) - 4.0V = 0.2V which is ΔV_p. According to the experimental data, T_{drift} is in the range of 0.1 and 1 second. Using the same values of C_{gate} and A as in the last section, the estimated J_{TA} from eq. 5 would be around 8 to 0.8 μA/cm^2 which falls in the range of the typical values of trap-assisted tunneling current. This calculation shows that a tunneling open can explain the I_{DDQ}(t) drift over time phenomena of the VLV-only failure CUTs.

4 Predictions and Experimental Results

This section first makes predictions for the faulty behavior of the circuits with tunneling opens. Experimental results collected from the Murphy chips are then shown to verify the predictions. Boolean and I_{DDQ} test results are shown in sections 4.1 and 4.2 respectively. Section 4.3 shows the burn-in results. Section 4.4 quotes some failure analysis results from the Sematech project [Nigh 97].

4.1 Boolean Tests

4.1.1 Predictions

Table 3 compares the speed ratio of four non-tunneling failure modes at nominal voltage and very low voltage [Chang 96a]. The *speed ratio* is defined as the speed of a defective-free circuit divided by the speed of a defective circuit. These numbers are obtained from simulations by injecting representative defects. This table shows that the difference between good and defective circuits becomes significant at very low voltage. A circuit with any of these non-tunneling failure modes could pass at-speed Boolean tests at nominal voltage but slow down 3 to 42 times at very

low voltage. In some cases, the defective circuit cannot operate at all (*hard failure*).

Table 3. Speed ratio of non-tunneling failure modes

Failure mode	Speed ratio	
	NV	VLV
Transmission gate open	3.3	Hard failure
V_t shift	1.3	3.0
Diminished drive gate	1.9	3.2
Degraded signal	1.7	42.0

Figure 7 compares the expected speeds of three circuits: a good circuit, a circuit with a tunneling-open failure and a circuit with a non-tunneling failure. At nominal voltage (NV), all three circuits have similar speeds. At very low voltage (VLV), the circuit with a non-tunneling failure mode (thin solid line) slows down 3 to 42 times (or fails) while the circuit with a tunneling open slows down much more than 42 times. Note that the curve of a tunneling open has a knee. Above the knee, the FN or direct tunneling effect is large enough to make the circuit operate at speed. Below this knee, the small trap-assisted tunneling current makes the speed slow down significantly. It may take as long as milliseconds for the small tunneling current to charge the gate capacitors. Also note that the circuit with a tunneling open can operate at the same lowest V_{DD} ($V_{DD,min}$) at which a good circuit can function. On the contrary, the circuit with a non-tunneling failure may fail completely above $V_{DD,min}$.

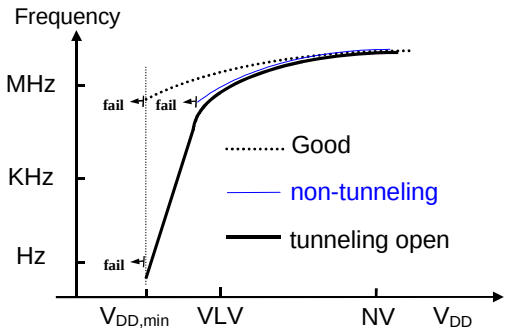

Figure 7. Predicted speeds of three circuits

4.1.2 Experimental Results

A speed measurement was performed to verify the predictions. The maximum speeds of the circuits were measured at different supply voltages. Table 4 lists the results of a good CUT and the nine VLV-only failure CUTs. The first row shows the speed of a good circuit. Its speed ranges from 37MHz to 4.17MHz as the supply voltage drops from 5V to 1.4V which is the lowest supply voltage ($V_{DD,min}$) at which a good circuit can function. At nominal voltage (5V), all CUTs have similar speeds. At very low voltage (1.7V), CUTs #1 and #2 failed to function (the slowest speed tested was 0.06Hz). CUT #3 operated 58 times slower than a good circuit. The other six CUTs were 20K times slower than a good circuit.

Table 4. Experimental circuit speeds (Hz)

CUT	5.0V	2.5V	1.7V	1.4V
good	37.0M	18.2M	7.69M	4.17M
1	35.7M	hard failure	hard failure	hard failure
2	33.3M	hard failure	hard failure	hard failure
3	37.0M	10.2M	132K	41.7K
4	35.7M	8.33M	1.75K	143
5	37.0M	3.33M	500	30.3
6	37.0M	6.25M	769	12.8
7	35.7M	5.88M	1.04K	16.1
8	35.7M	10.5M	3.57K	62.5
9	37.0M	12.7M	2.77K	50.0

Figure 8 shows the experimental circuit speeds of the good and the nine VLV-only failure CUTs based on the numbers in Table 4. The first two CUTs fail below V_{DD} = 4V and 3.5V respectively (their curves overlapped with the others). They do not match the expected behavior of tunneling opens. The last seven CUTs match the expected behavior of tunneling opens. The knee happens at 2.5V. They all operate at the same $V_{DD,min}$ as a good circuit.

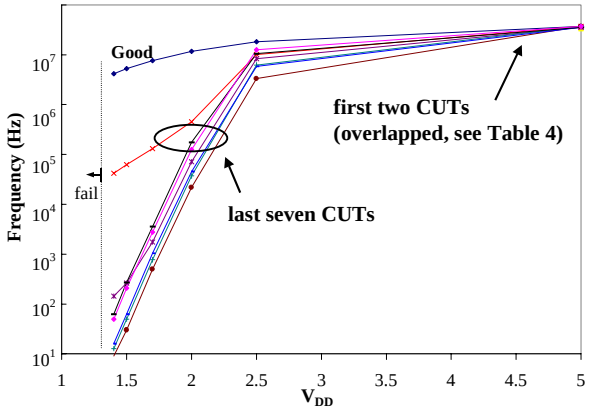

Figure 8. Experimental circuit speeds (see Table 4)

4.2 I_{DDQ} Tests

4.2.1 Predictions

Table 5 shows the predicted I_{DDQ} behavior of different failure modes that might cause VLV-only failure or $I_{DDQ}(t)$ drift over time. The first four failure modes have been discussed in connection with Table 1. The fifth failure mode, *defective PN junction*, could happen in the following cases: 1. PMOS gate to

source(drain) shorts, 2. NMOS gate to substrate shorts, and 3. source(drain) - substrate(well) junction leakage. These defective PN junctions have very small reverse bias current which may cause long $I_{DDQ}(t)$ drift as reported in [Maly 88].

Table 5. Predicted $I_{DDQ}(t)$ of different failure modes

failure mode	High I_{DDQ}?	time constant of drift, τ	V_{DD} dependency of τ
(1) High resistance interconnect	No	No drift	No drift
(2) V_t shift (3) Gate oxide shorts (4) Metal shorts	Yes	No drift	No drift
(5) Defective PN junction	Yes	> ms	V_{DD} decrease, τ decrease
(6) Tunneling opens	Y/N*	> ms	V_{DD} decrease, τ increase

*depends on defect thickness

The first column shows whether the failure modes cause high I_{DDQ} or not. Depending on the defect thickness, a tunneling open may or may not cause high I_{DDQ}. The second column shows the *time constant τ of the $I_{DDQ}(t)$ drift* which is defined in eq. 6. The $I_{DDQ}(t)$ drift (see Fig. 1) is modeled by the following equation,

$$I_{DDQ}(t) = I_{drift} \cdot e^{-t/\tau} + I_{final} \qquad (6)$$

where I_{drift} is the amplitude of the $I_{DDQ}(t)$ drift and the I_{final} is the final I_{DDQ} value. τ is referred to as the *time constant of $I_{DDQ}(t)$ drift*. For the first four failure modes, I_{DDQ} should not change with time. For a defective PN junction and a tunneling open, the time constant may be as long as milliseconds or seconds.

The third column indicates whether the time constant is a function of supply voltage. For the case of a defective PN junction, since the current through a reverse biased PN junction is independent of the voltage, the time constant of $I_{DDQ}(t)$ drift should go down linearly as the supply voltage goes down. On the contrary, for a tunneling open, the time constant should go up as the supply voltage goes down. Because FN and direct tunneling become insignificant at low voltage, the defective node stays at an intermediate voltage for a longer time.

4.2.2 Experimental Results

An $I_{DDQ}(t)$ testing experiment was performed to verify the previous predictions. In this experiment, twenty continuous I_{DDQ} measurements were taken after applying each pattern. Each I_{DDQ} measurement took about 7-21 milliseconds. Table 6 lists the results collected from a good CUT as well as the nine VLV-only failure CUTs. The first row shows that the good circuit did not have I_{DDQ} higher than 1μA. The first two VLV-only failure CUTs had high and constant I_{DDQ} over time. The third and fourth CUT had low and constant I_{DDQ} over time. They can be explained with very thin tunneling opens (see section 3.1). The last five CUTs had $I_{DDQ}(t)$ drift over time. None of the last five VLV-only failure CUTs has a time constant that decreases as the supply voltage drops. These results show that the last seven CUTs match the expected behavior of a tunneling open.

Table 6. $I_{DDQ}(t)$ experimental results

CUT	first I_{DDQ} (μA)	I_{DDQ} drift?	I_{drift} (μA)	$\tau(V_{DD})$ in ms		
				6V	5V	3.5V
good	0.9	No	-	-	-	-
1	1800	No	-	-	-	-
2	328	No	-	-	-	-
3	0.8	No	-	-	-	-
4	0.7	No	-	-	-	-
5	65	Yes	24	14	28	84
6	45	Yes	40	28	55	180
7	50	Yes	50	28	44	70
8	35	Yes	35	28	63	91
9	40	Yes	40	28	63	140

Based on both experimental results shown in sections 4.1.2 and 4.2.2, the behavior of the first two CUTs can be explained by non-tunneling failure modes. The behavior of the last seven CUTs can be explained well by the tunneling-open failure mode.

4.3 Burn-in Experiment

4.3.1 Predictions

If the tunneling open is so thin, will it break down easily? It has been shown that for a carrier tunneling through the oxide, the minimum distance between collisions is about 8-15Å [Fischetti 85]. It means that in a very thin oxide, it is very possible for the carriers to undergo direct tunneling (ballistic transport) without damaging the oxide. Quantitatively, the charge to breakdown (Q_{BD}) is $1\sim10^4$ Coul/cm^2 for a 30Å oxide and $10^3\sim10^7$ Coul/cm^2 for a 25Å oxide [Schuegraf 94].

To estimate the lifetime of a 12Å thin oxide (as in the case of section 3.2.1), the following assumptions were made. The Q_{BD} is 10^7 Coul/cm^2. The charge going through the tunneling open is $V_{DD}/2$ x C_{gate} (= 2.5V x 40fF = 100fCoul) per switch. Calculation shows that a 12Å thin oxide can survive 10^{12} switches before it breaks down. This number is very large compared with the number of test patterns (in the order of millions, 10^6). So the tunneling opens may survive a long period of usage without breaking down.

After the tunneling open breaks down, the defective contact (or via) may become a resistive open. This

does not mean the chip is "healed". Suppose the defective oxide thickness in Fig. 3 is not uniform. Only a very small area (A) is thinner so that most current flows through this small area. In this case, electron-migration could be serious due to localized high current density. Then the device may have reliability problems. Based on the above discussion, whether tunneling opens cause reliability problems is not known yet.

4.3.2 Experimental Results

Since high temperature is effective in speeding up the oxide degradation process [Schuegraf 94], a 366 hours temperature burn-in experiment was performed. Characterization tests (including speed measurement and I_{DDQ} testing) were performed before, during and after the burn-in process. The burn-in temperature was 130 degree C. The burn-in voltage was 5V, which is the highest available power supply voltage for the burn-in equipment used. Exhaustive test patterns were applied during burn-in (i.e., dynamic burn-in).

Table 7 shows the experimental results for good CUTs and the nine VLV-only failure CUTs. The changes of max I_{DDQ} values as well as the speeds at nominal voltage (5V) and very low voltage (1.7V) during the burn-in are listed. The first row shows the average values of a reference group of good CUTs. The first two VLV-only failure CUTs which are not suspected to have tunneling opens failed after the first six hours of burn-in. The other seven VLV-only failure CUTs which are suspected to have tunneling opens survived the 366 hours burn-in and their circuit behavior remained almost unchanged (except that CUT #7's I_{DDQ} increased after 6 hours of burn-in). Although all suspect CUTs survived the temperature burn-in, voltage burn-in will be performed in the future to further verify whether tunneling opens cause early-life failure or not.

4.4 Failure Analysis

Although failure analysis has not yet been done on these nine VLV-only failure CUTs, data from other experiments also support the existence of a tunneling open. In Sematech's experiment [Nigh 98], four out of seven "low voltage sensitive" dies which passed tests at nominal voltage but failed at low voltage were sent to failure analysis. Two of them were found to have gate-substrate or drain-substrate shorts. Nothing was found on the other two dies. This implies that some of the VLV-only defects are difficult to observe in failure analysis. A tunneling defect can be the culprit.

5 Discussion

5.1 Indication of Process Problems

Since the circuits with tunneling opens can operate at-speed at nominal voltage and they do not fail after burn-in, why do we want to test for tunneling opens? The answer is that the tunneling open can be an indication of process problems.

5.2 Screening Strategy

What is the best strategy to screen out the chips with tunneling opens? Traditionally, people use I_{DDQ} testing (single measurement per pattern, single pass/fail limit) to screen out weak chips. However, it has been shown in the previous sections that some tunneling opens may not fail traditional I_{DDQ} tests. Figure 9 shows a proposed screening strategy for tunneling opens. It should be performed after regular nominal voltage tests. This is a two-stage screening strategy which combines VLV and $I_{DDQ}(t)$ testing. Only if a circuit fails at VLV does it go to $I_{DDQ}(t)$ testing. If its $I_{DDQ}(t)$ values remain at a high and constant value (do not drift over time), the circuit is not a tunneling open. The circuit can cause early-life failure and must be rejected. If its $I_{DDQ}(t)$ values remain at a constant low level or its $I_{DDQ}(t)$ values drift over time, then the circuit might have a tunneling open. Compared with traditional I_{DDQ} testing, the proposed screening strategy is less costly because VLV testing requires shorter test time.

Table 7. Burn-in results (I_{DDQ} in μA, speed in Hz)

CUT #	before burn-in			after 6 hours burn-in			after 150 hrs. burn-in			after 366 hrs. burn-in		
	max I_{DDQ}	speed (Hz)		max I_{DDQ}	speed (Hz)		max I_{DDQ}	speed (Hz)		max I_{DDQ}	speed (Hz)	
		5V	1.7V		5V	1.7V		5V	1.7V		5V	1.7V
good	0.9	37.0M	7.69M	0.9	37.0M	7.69M	0.9	37.0M	7.69M	0.9	37.0M	7.69M
1	1080	35.7M	HF	1200	HF	HF	1420	HF	HF	1380	HF	HF
2	1480	33.3M	HF	2300	HF	HF	1300	HF	HF	1380	HF	HF
3	0.8	37.0M	132K	0.8	37.0M	125K	0.8	37.0M	133K	1.1	37.0M	145K
4	0.7	35.7M	1.75K	0.8	35.7M	1.51K	0.8	35.7M	1.58K	0.8	35.7M	1.54K
5	63.4	37.0M	500	63.8	37.0M	384	61.8	37.0M	370	62.8	37.0M	357
6	45.4	37.0M	769	45.8	37.0M	625	44.6	37.0M	714	45.2	37.0M	625
7	50	35.7M	1.04K	2980	35.7M	1.00K	2980	35.7M	0.91K	2980	35.7M	0.83K
8	32.4	35.7M	3.57K	32.4	35.7M	2.94K	32.4	35.7M	2.94K	32.4	35.7M	2.78K
9	36.0	37.0M	2.77K	36.4	37.0M	2.4K	35.8	37.0M	2.5K	30.8	37.0M	2.38K

HF: hard failure

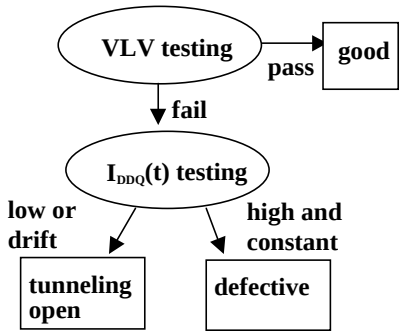

Figure 9. Proposed screening strategy for tunneling opens

The most effective test voltage to detect tunneling opens is the lowest possible supply voltage. As far as the test pattern is concerned, single-detect single stuck-at test patterns may not be effective because the tunneling effect has polarity dependence. Both the rise and fall transitions of a node have to be tested. Test patterns that have higher transition fault coverage such as transition fault test patterns or multiple-detect single stuck-at test patterns are more effective in detecting tunneling opens.

To verify the previous discussion, an experiment was performed [Chang 98a]. To change the number of transitions in a test pattern, the test pattern was modified into five different versions:

1. reordering the sequence of the vectors
2. padding an all-one vector before every vector
3. padding an all-zero vector before every vector
4. padding a bit-wise complemented vector before every vector
5. padding a one-bit-shifted vector before every vector

The testing was done at very low voltage (1.7V) and the test speed was the rated speed of a good circuit at 1.7V. The numbers of total failures were recorded. The first two non-tunneling open CUTs in Table 2 showed the same test results for the modified and original versions of test patterns. The last seven tunneling open CUTs showed different test results for some modified versions than for the original version. The experimental results support the previous arguments that the tunneling opens do not behave like stuck-at faults. The test results change with the number of transitions.

For I_{DDQ}(t) testing, test patterns that have high coverage of node transition are effective because the I_{DDQ}(t) drift over time can only be observed after the defective node has transition. To save test time, two I_{DDQ} measurements for each test pattern might be enough. According to this experiment, the wait time between two measurements can be tens or hundreds of milliseconds. If the second measurement is less than a certain percent (e.g., 80%) of the first measurement, then the circuit is suspected of having a tunneling open.

5.3 Defect Coverage

What is the defect coverage of VLV testing for tunneling defects? Will it still be valid in more advanced technology than 0.7μ? As is shown in the calculation example in section 3.2.1, the defect thickness (d) which is detectable by VLV testing is a function of gate capacitance (C_{gate}) and defect area (A). Therefore the detectable defect thickness of the VLV testing varies with locations in the same circuit. Table 8 lists the calculated minimum detectable defect thickness (d_{min}) that VLV testing can detect. Some typical numbers were estimated. The calculation is done in the same way as described in section 3.2.1. Note that in this calculation, it is assumed that no charge leaks through the gate oxide at the moment of gate switching. This is a valid assumption because the charges that tunnel through the tunneling open (dominated by direct or FN tunneling) is orders of magnitude more than the charges that leak through the gate oxide (dominated by trap-assisted tunneling).

Table 8 shows that VLV testing is capable of detecting tunneling opens with defect thickness d_{min} as thin as 7Å. A question that might arise in connection with Table 8 is that since the nominal supply voltage is already lower in the advanced technology, why is VLV testing still effective? The answer is that the defect area (A) and gate capacitance (C_{gate}) are also scaled down. Therefore the calculation shows that VLV testing is still needed to screen out tunneling opens in advanced technologies.

Table 8. Minimum detectable defect thickness

feature size	NV (V)	VLV (V)	A (μm^2)	C_{gate} (fF)	d_{min} (Å)
0.7μm	5	1.7	1*	40*	12
.35μm	3.3	1.2*	0.3*	10*	8
.18μm	1.8	0.9*	0.07*	3*	7

* estimated numbers

6 Summary

Nine VLV-only failure CUTs were found in the Murphy experiment. A tunneling open is proposed to explain seven of these VLV-only failures. A tunneling open can occur at a contact or a via where a very thin layer of oxide allows tunneling current to flow through it. Theoretical calculations show that field-dependent tunneling current allows the circuit to pass at-speed tests at nominal voltage but causes it to operate at very slow speed at very low voltage. It is also shown that

the trap-assisted tunneling current can cause $I_{DDQ}(t)$ to drift over an interval of hundreds of milliseconds.

The results of various experiments verified the theory. Table 2 summarizes the experimental results. These nine CUTs can be divided into three categories according to their faulty behavior. The first category failed at very low voltage. This category does not match the expected behavior of a tunneling open. This category could be caused by some non-tunneling failure mode, such as gate oxide shorts. The second category has very long delay at VLV but does not have high I_{DDQ}. The third category has very long delay at VLV and $I_{DDQ}(t)$ drift over time. The last two categories (seven CUTs) can be explained as tunneling-opens.

Although the tunneling open circuits have survived the 366 hours temperature burn-in, the existence of tunneling opens might indicate process problems. VLV testing followed by $I_{DDQ}(t)$ testing is proposed as a cost effective screening strategy.

Acknowledgements

The authors would like to thank the following people for their help and valuable comments. Jonathan Chang, Scott Keller (LSI Logic), Moon-Jung Kim, Siegrid Munda, Prakash (Pycon), Mike Purtell (Advantest), Don Sireci (Advantest), Mike Stover (LSI Logic), and Chaowen Tseng. This work was supported by LSI Logic Corporation.

References

[Campbell 91] Campbell, A., E. Cole, C. Henderson, and M. Taylor, " Case History: Failure Analysis of a CMOS SRAM with an Intermittent Open Contact," *Proceeding of 17th International Symposium for Testing and Failure Analysis*, pp. 261-269, 1991.

[Chang 98a] Chang, J., *et. al.*, "Analysis of Pattern-Dependent and Timing Dependent Failure in An Experimental Test Chip," *Proc. of International Test Conference*, pp.184-193, 1998.

[Chang 98b] Chang, J., *et. al.*, "Experimental Results for I_{DDQ} and VLV Testing," *IEEE VLSI Test Symposium*, pp.118-123, 1998.

[Chang 96a] Chang, J. and E.J. McCluskey, "Detecting Delay Flaws by Very-low-voltage Testing," *Proceeding of International Test Conference*, pp.367-376, 1996.

[Chang 96b] Chang, J. and E.J. McCluskey, "Quantitative Analysis of Very-low-voltage Testing," *IEEE VLSI Test Symposium*, pp.332-337, 1996.

[Fowler 28] Fowler, R.H. and L.W. Nordheim, "Electron Emission in Intense Electric Fields, " *Proc. R. Soc. Lond. A*, vol. 119, pp. 173-181, 1928.

[Franco 95] Franco, P. and *et. al.*, "An Experimental Chip to Evaluate Test Techniques Chip and Experiment Design," *Proc. International Test Conference*, pp.653-662, 1995.

[Fischetti 85] Fischetti, M.V., *et. al.*, "Theory of High Field Electron Transport in Silicon Dioxide," *Phys. Rev. B*, vol.31, pp.8124, 1985.

[Gupta 97] Gupta, A., *et. al*., "Accurate Determination of Ultrathin Gate Oxide Thickness and Effective Polysilicon Doping of CMOS Devices," *IEEE Electron Device Letters*, VOL. 18, No. 12, pp.580-582 December, 1997.

[Hao 93] H. Hao and E.J. McCluskey, "Very-Low Voltage Testing for Weak CMOS Logic IC's," *Proceeding of International Test Conference*, pp.275-284, Oct. 1993.

[Henderson 91] Henderson, C., Soden and Hawkins, "The Behavior and Testing Implications of CMOS IC Logic Gate Open Circuits," *Proceeding of International Test Conference*, pp. 302-310, 1991.

[Hnatek 95] Hhnatek, E.R., *Integrated Circuit Quality and Reliability*, Second Edition, Marcel Dekker, Inc. 1995.

[Leblebici 93] Leblebici, Y., and S.M. Kang, *Hot-carrier Reliability of MOS VLSI Circuits*, Kluwer Academic Publishers, 1993.

[LSI 93] Lsi Logic, "0.7-Micron Array-Based Product Databook," pp.1-20 ,1993.

[Maly 88] Maly, W., Nag and Nigh, "Testing Oriented Analysis of CMOS ICs with Opens," *ICCAD*, pp.344-346, 1988.

[McCluskey 2000] McCluskey, E. J., Tseng, C.W., "Stuck-Fault vs. Actual Defects," *Proceeding of International Test Conference*, 2000.

[Moazzami 92] Maozzami, R. and C.Hu, "Stress-Induced Current in Thin Silicon Dioxide Films," *IEEE International Electron Device Meeting*, pp. 6.3.1-6.3.4, 1992.

[Nigh 97] Nigh, P., W. Needham, K. Bulter, P. Maxwell, and R. Aitken, " An Experimental Study Comparing the Effectiveness of Functional, Scan, IDDQ and Delay-fault Testing, " *IEEEE VLSI Test Symposium*, pp.459-464, 1997.

[Nigh 98] Private communication, 1998.

[Needham 98] Needham, W., C. Prunty, and E. Yeoh, "High Volume Microprocessor Escapes, An Analysis of Defects Our Tests Are Missing," *Proceeding of International Test Conference*, pp.25-34, 1998.

[Soden 89] J. Soden and R. Treece, "CMOS IC Stuck-open Fault Electrical Effects and Design Considerations", *Proc. of International Test Conference*, pp.423-430, 1989.

[Schuegraf 94] Schuegraf, K and C. Hu, "Effect of Temperature and defects on Breakdown Lifetime of Thin Sio2 at Very Low Voltages," *IEEE Reliability Physics Symposium*, pp.126-135, 1994.

[Schuegraf 92] Schuegraf, K and C. Hu, "Ultra-thin Silicon Dioxide Leakage Current and Scaling Limit," *IEEE Sym. On VLSI Technology*, pp.18-19, 1992.

[Shi 98] Shi, *et.al.*, "Polarity Dependent Tunneling Current and Oxide Breakdown in Dual-Gate CMOSFET's," *IEEE Electron Device Letters*, vol. 19, No. 10, pp.2355-2360, Oct, 1998.

[Tseng 00] Tseng, C.W., X. Shao, D.M. Wu and E.J. McCluskey, " Cold Delay Defect Screening," *IEEE VLSI Test Symposium*, 2000.

[Weinberg 82] Weinberg, Z.A., "On Tunneling in Metal-oxide-silicon Structures," *J. Appl. Phys.*, vol. 53, pp.5052-5056, 1982.