Signature Analysis for Testing, Diagnosis, and Repair of Multi-Mode Power Switches*

Zhaobo Zhang[1], Xrysovalantis Kavousianos[1,2], Yan Luo[1], Yiorgos Tsiatouhas[2] and Krishnendu Chakrabarty[1]
[1]Dept. of Electrical & Computer Engineering
Duke University, 27708 Durham, NC, USA
[2]Dept. of Computer Science
University of Ioannina, 45110 Ioannina, Greece

Abstract- **Power-gating structures for intermediate power-off modes offer significant power saving benefits as they reduce the leakage power during short periods of inactivity. However, reliable operation of such devices must be ensured by using adequate test methods. We propose a signature analysis technique to efficiently test power-gating structures that provide intermediate power-off modes. In particular, the proposed technique can be used to test and diagnose an efficient multi-mode power-gating architecture that was proposed recently. In addition, we propose a methodology to repair catastrophic and parametric faults, and to tolerate process variations. Analysis and extensive simulations demonstrate the effectiveness of the proposed method.**

I. INTRODUCTION

As devices keep shrinking in deep sub-micron technologies, transistor threshold voltage decreases due to the drain-induced barrier-lowering effect as well as due to supply-voltage scaling, a technique that has served as the mainstream approach for reducing dynamic power dissipation. However, as threshold voltage decreases, leakage currents and the resulting static power increase considerably, thereby offsetting the gains derived from reduced dynamic power consumption.

A number of techniques have been proposed in the literature to reduce static power consumption. One such technique uses high-V_t cells in high-slack domains of the circuit [4]. Static-power reduction can also be achieved by partitioning the system into islands with separate supply rails and unique power characteristics [8], [11], [12]. These techniques allow separate power management policies to be applied to each island and thus they can best exploit the potential of each island to reduce power during both active and idle modes of operation. During active mode, power consumption is reduced by dynamically scaling the supply voltage whenever performance constraints are not violated [1], [5], [15]. During idle mode, the power supply is completely disconnected from the core, by using high-V_t power switches inserted between the core and the power supply or the ground rail. These switches are turned off, thereby suppressing the leakage current [4], [7].

Even though high-V_t power switches are very effective in reducing static power, they cannot be exploited during short periods of inactivity. This is attributed to the long *wake-up time*, that is, the time required to re-activate the core from the power-off mode. In [2], [6], [10] design techniques were presented to achieve one intermediate power-off mode that offers moderate suppression of leakage current and thus shorter wake-up times than the wake-up time required when the power supply is completely turned-off. The authors in [13] presented comprehensive studies to show that there are further power reduction benefits when multiple intermediate power-off modes are supported. They also proposed a power-switch scheme to support two intermediate power-off modes. However, this scheme consists of difficult-to-test analog components that are also highly sensitive to process variations. Thus it is unsuitable for realistic applications. An efficient and robust power-switch architecture was proposed in [16], [17]; this design is suitable for realistic applications and practical workloads as it supports more than two intermediate modes, requires minimal design effort, and is tolerant to process variations.

Even though power switches offer significant benefits in reducing static power, their adoption in practice depends on the availability of test, diagnosis methods that can guarantee their correct operation in the field. In prior work, test methods have been presented only for power switches that support the complete power-off mode [3], [14]. However, testing of power switches with intermediate power-off modes is more challenging due to their analog characteristics arising from intermediate voltage levels at the various power-down modes.

In this paper, we present a signature analysis method for the testing and diagnosis of power switches with multiple intermediate power-off modes. Even though the proposed technique is generic and can be applied to any such structure, we focus on the power switches proposed in [16], [17] since they have been shown to constitute a highly effective multi-mode power-gating architecture. The proposed method converts the analog characteristics of the power switches into a digital signature that can be transferred to the tester using simple built-in test structures [19]. We present a thorough fault analysis of the multi-mode architecture presented in [16] and we show that the proposed approach is effective in testing and diagnosing catastrophic and parametric faults affecting this architecture. Finally, we exploit the inherent properties of the power-switch design to develop a scheme to repair catastrophic and parametric faults, and to tolerate process variations at negligible cost.

II. BACKGROUND

The multi-mode power-switch architecture proposed in [16], [17] is shown in Fig. 1. It consists of a large footer high-V_t transistor M_p and a couple of small transistors M_0, M_1 that are connected between the core and the ground rail (note that we refer to M_p as a large transistor even though it is typically implemented using several transistors in parallel). When M_p is "on", the core operates in the normal functional mode. Note that M_p is large enough to constitute a strong driver and to thus

*This research was supported in part by the National Science Foundation under grant no. CCF-0903392, and by the Semiconductor Research Corporation under contract no. 1992.

1530-1877/11 $26.00 © 2011 IEEE
DOI 10.1109/ETS.2011.18

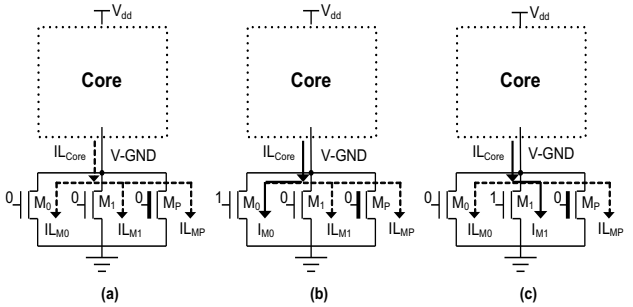

Figure 1. Multi-mode power-gating architecture

minimize the impact on circuit performance during normal operation. When M_p is "off" (i.e., during idle mode) and transistors M_0 and M_1 are also off, the virtual ground rail (*V-GND*) charges to a voltage level close to the power supply. Therefore, it suppresses the leakage currents of the transistors of the circuit due to the body effect. In this mode (which is called hereafter *snore mode*), the leakage current of the core, IL_{core}, is equal to the aggregate leakage current flowing through transistors M_0, M_1, M_P ($IL_{core} = IL_{M0}+IL_{M1}+IL_{MP}$), which is very small (note that M_0, M_1 are very small low-V_t transistors and M_P is a large high-V_t transistor). Thus the voltage level at virtual ground rail $V_{V\text{-}GND}$ approaches V_{dd} and the circuit consumes a negligible amount of energy.

In order to restore the voltage of the virtual ground rail to its nominal value when the circuit transitions from the power-off mode to the active mode, the parasitic capacitance at the *V-GND* node has to be completely discharged through the power switch M_p. However, this switch is not large enough (due to area constraints) to quickly restore the voltage level at the virtual ground node. Thus the wake-up time is usually long relative to circuit clock period. Transistors M_0 and M_1, which selectively remain on during the short idle periods, set the virtual ground node to an intermediate voltage level and thus offer two intermediate power-off modes (M_0 implements the *dream mode* and M_1 implements the *sleep mode*). Specifically, in the dream mode transistor M_0 is on and transistors M_P and M_1 are off as shown in Fig. 1(b). In this case, the current flowing through transistor M_0 increases compared to the snore mode because M_0 is on ($I_{M0} > IL_{M0}$). The exact value of I_{M0} depends on the size of transistor M_0, and it sets the virtual ground node at a lower voltage level, compared to that in the snore mode. Thus the static power consumed by the core increases compared to the snore mode, but the wake-up time drops.

In order to get into the sleep mode, transistor M_1 is turned on, and M_P, M_0 are turned off as shown in Fig. 1(c). Provided that M_1 has larger aspect ratio than M_0 ($W_{M1}/L_{M1} > W_{M0}/L_{M0}$), the aggregate current flowing through M_0, M_1, and M_P increases even more when M_1 is on (note that $I_{M1}>I_{M0}$). Consequently, the voltage level at the virtual ground node is further reduced compared to the dream mode and thus the wake-up time decreases at the expense of increased static power consumption.

Note that in both sleep and dream modes, static power consumption is much lower than that in the active mode. Thus, there are significant benefits if the core is put into an intermediate mode for a short period instead of remaining idle in the active mode [13]. Besides M_0 and M_1, additional transistors can be used to offer more power-off modes if they are properly sized, as shown in [16], [17].

TABLE 1. EFFECTS OF CATASTROPHIC FAULTS

	V-GND voltage (mV) / Wake-Up Cycles / Static Power (mW)				
	Fault-free	M_0 stuck-on	M_0 stuck-open	M_1 stuck-on	M_1 stuck-open
Snore mode	795 mV 8 cycles 0.73 mW	528 mV 5 cycles 1 mW	Same as fault-free case	410 mV 3 cycles 1.28 mW	Same as fault-free case
Dream mode	528 mV 5 cycles 1 mW	Same as fault-free case	795 mV 8 cycles 0.73 mW	343 mV 2 cycles 1.51 mW	Same as fault-free case
Sleep mode	410 mV 3 cycles 1.28 mW	343 mV 2 cycles 1.51 mW	Same as fault-free case	Same as fault-free case	795 mV 8 cycles 0.73 mW

III. FAULT ANALYSIS

We first describe the simulation framework used in this work. We used a logic core consisting of nine million transistors; such a core is representative of a realistic industrial circuit in terms of static power consumption during DC operation. The logic core and the power switches from [16] were implemented using the 45 nm predictive technology [18] for 1.1 volts power supply. As suggested in [16], the width of the main power switch M_P was set equal to 43.2×10^6 nm, i.e., 12% of the total width of the nMOS transistors in the logic core (the length of M_P is set equal to 45 nm). The sizes (W, L) of the power switches are (250 nm, 45 nm) for M_0 and (480 nm, 45 nm) for M_1 [16]. The static power consumption of the core in idle mode without the use of power switches was equal to 10 mW. The clock frequency was set to 1 Ghz.

Table 1 presents the effects of catastrophic faults on the operation of the logic core during the power-off modes. Each row of Table 1 corresponds to a different mode and each column to a different fault (except the first column, which reports the fault-free values). Each entry of the table reports the voltage at *V-GND*, the number of wake-up cycles, and the static power. Note that this data is reported only for those faults that cause a deviation from the fault-free case (the remaining entries are denoted as "*Same as fault-free case*"). Stuck-on faults reduce the wake-up time but increase static-power consumption; they do not affect the functionality of the core. On the other hand, stuck-open faults increase the wake-up time and they seriously affect the reliability of the core. Note that if the core does not fully wake-up in a predetermined number of cycles due to a fault then it will exhibit slower operation, which might corrupt the state of the core and cause logic errors. The bridging fault between the gates of transistors M_0 and M_1 has not been considered as it can be trivially avoided by not placing them close to each other in the layout.

Next we study the effects of parametric faults involving ±10%, ±20% and ±30% variation in the width, length, and threshold voltage of M_0 and M_1 (one fault is inserted at a time).

TABLE 2.
CHANGE IN VIRTUAL GROUND VOLTAGE FOR PARAMETRIC FAULTS (mV)

	Fault:	**Deviation Of Parameter**					
		−30%	**−20%**	**−10%**	**+10%**	**+20%**	**+30%**
M_0 Dream Mode	Width	+68	+43	+20	-18	-34	-49
	Length	-114	-66	-27	+19	+35	+49
	$V_{threshold}$	-33	-24	-13	+14	+30	+48
M_1 Sleep Mode	Width	+63	+39	+18	-16	-30	-43
	Length	-83	-47	-20	+16	+30	+42
	$V_{threshold}$	-25	-18	-10	+12	+26	+43

TABLE 3.
CHANGE IN NUMBER OF WAKE-UP CYCLES FOR PARAMETRIC FAULTS

	Fault:	Deviation Of Parameter					
		–30%	–20%	–10%	+10%	+20%	+30%
M_0 Dream Mode	Width	**+1 cycle (30/30)**	**+1 cycle (3/30)**	0 cycles	0 cycles	-1 cycle (30/30)	-1 cycle (30/30)
	Length	-2 cycles (30/30)	-1 cycle (30/30)	-1 cycle (30/30)	0 cycles	0 cycles	**+1 cycle (30/30)**
	$V_{threshold}$	-1 cycle (30/30)	-1 cycle (30/30)	0 cycles	0 cycles	0 cycles	**+1 cycle (30/30)**
M_1 Sleep Mode	Width	**+1 cycle (30/30)**	0 cycles	0 cycles	0 cycles	-1 cycle (30/30)	-1 cycle (30/30)
	Length	-1 cycle (30/30)	-1 cycle (30/30)	-1 cycle (1/30)	0 cycles	0 cycles	**+1 cycle (29/30)**
	$V_{threshold}$	-1 cycle (30/30)	0 cycles	0 cycles	0 cycles	0 cycles	**+1 cycle (29/30)**

Table 2 presents the effect of these faults on the virtual-ground voltage. For each fault, we report the deviation of the voltage level at *V-GND* from the nominal value, as an average of Monte Carlo simulations with 30 samples (30 samples are adequate according to [9]). For the fault-free case, the voltage at *V-GND* in the dream (sleep) mode is 520mV (410 mV).

In Table 3, we report the deviation in the number of cycles required for wake-up due to each parametric fault. Any positive (negative) value denotes the number of additional (fewer) cycles compared to the fault-free case required for re-activating the core in the existence of the fault. Below the number of cycles, we report the number of samples that were observed to cause the reported deviation. For example, when the width of M_0 is reduced by 20%, 3 out of 30 samples were found to increase the number of wake-up cycles by one. Bold-faced entries correspond to parametric faults that increase the number of cycles required for re-activating the core and which thus constitute a serious reliability issue for the operation of the core. The rest of the faults either do not affect the operation of the core or cause an increase in static power (the number of wake-up cycles decreases).

IV. PROPOSED TEST STRUCTURE

The testing of any structure offering intermediate power-off modes can be done by measuring the virtual ground voltage and by verifying that the wake-up time has not increased and that the goal of reducing static power at any power-off mode is achieved. One straightforward method is to add an analog pad and measure the voltage at *V-GND*. However, this probing method is costly and inaccurate, and moreover, one extra pin needs to be added to the original design. Another solution is to embed an analog-to-digital converter (ADC) which however is very expensive in terms of area and is also very sensitive to process variations.

In this work, we propose a simple and purely digital built-in test circuit, which can be embedded on-chip to accurately measure the virtual ground voltage. Note that even though we focus on the power-off switches proposed in [16] the proposed test method is generic and it can be used to test any similar structure. The basic idea of the proposed test structure is to convert the voltage of the virtual-ground node to a frequency reading. Specifically, we use a voltage control oscillator (VCO) to convert the voltage of *V-GND* into a signal toggling with a frequency that depends on this voltage. Then we use this signal to trigger a binary counter that provides a quantification of the frequency. By using this technique, the digital value that is stored in the counter at the end of the testing period is directly proportional to the voltage at *V-GND*.

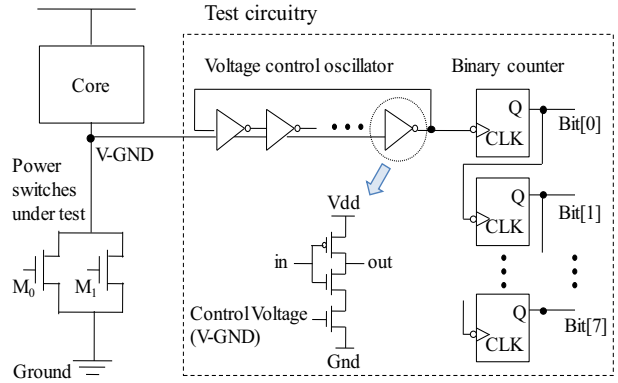

Figure 2. Proposed test circuitry for the multi-mode power gating architecture.

The proposed test circuit is illustrated in Fig. 2. It consists of a chain of inverters forming a ring oscillator and a number of D flip-flops forming a ripple counter. The VCO is constructed using a ring of current-starved inverters. We used the structure of inverters shown in Fig. 2 as they provide high resolution in the detection of voltage variations but any similar structure can be also used. The virtual ground is connected to the VCO as the control voltage. The propagation delay of each inverter depends on the control voltage. Thus, the frequency of the ring oscillator varies with the voltage at *V-GND*. When the voltage at *V-GND* is high, the discharging current through the tail NMOS transistor is large. Subsequently, the frequency of the ring oscillator is high. The opposite happens when the voltage at *V-GND* is low. The frequency of the ring oscillator increases monotonically as the control voltage increases before the tail NMOS transistor goes into the saturation region.

The frequency of the ring oscillator can be easily calculated using a binary counter. In our design, a ripple counter was implemented using D flip-flops (DFFs). The D input of each DFF is connected to its output Q′ (it is omitted in Fig. 2 for simplicity). The output Q of each DFF is connected to the clock input of the flip flop at the next stage of the counter. At each two successive toggles of each flip flop, one toggle of the next flip flop is triggered, thereby providing binary counting. Thus the value of the binary counter at the end of the test period is directly proportional to the frequency of the VCO.

The output of the counter comprises a test signature that represents the voltage level at the virtual ground node. The die signature can be easily transferred to the tester by using design-for-testability structures such as scan chains and test wrappers that are common in logic cores. In order to identify defective dies, the signature of each tested die must be compared to a pre-computed fault-free signature. However, a single fault free signature is not sufficient to determine whether a die is faulty or fault free. The reason is that the voltage at *V-GND* may deviate from the expected value and be still acceptable provided that the number of wake-up cycles or the static power limits of the core are not exceeded. Thus, it is more realistic to assume that the power switches are defect-free if the voltage at *V-GND* lies between an upper bound and a lower bound; these bounds depend on the power-off mode.

Voltage values above the upper bound increase the number of wake-up cycles while voltage values below the lower bound increase the static power. Thus, for each power-off mode, a range of acceptable signature values (consisting of a lower-bound and an upper-bound signature) must be computed.

A major concern in this test method is that process variations may affect the operation of VCO, which in turn can shift the signature of a fault-free die out of the nominal range of acceptable signature values. Thus, the acceptable signature bounds must be computed in such a way as to consider process variations inherent in the underlying technology node. To this end, we adjust both bounds as well as the measured signatures using the following approach. Initially we connect the input of the VCO to the power-supply voltage using a pull-up transistor (the virtual ground node is temporarily disconnected using a transistor switch from the input of the VCO in this case) and we get a golden reference signature using simulation assuming no process variations. Next we run Monte Carlo simulation and produce N circuit samples affected by process variations. For each sample i, we measure again the reference signature and we determine the difference D_i between this signature and the golden reference signature. This difference reflects the magnitude of the effect of process variations on the operation of the VCO, and in particular on the computation of the signature.

Note that the power-supply is not affected by process variations and thus it provides an excellent means for capturing the process variations affecting only the VCO. In addition, power droop and power-supply noise in the logic core does not affect the testing of the power switches. For each sample, we calculate the lower-bound and the upper-bound signature values for each power-off mode and we add D_i to both of them. By adding D_i to these bounds, the acceptable signature range for each intermediate power-off mode of each sample is shifted towards the process-variation-free case. Using this technique, we eliminate most of the effects of process variations on the computation of these ranges. The new ranges computed for each circuit sample and each power-off mode are called adjusted-acceptable ranges (AAR).

Despite the benefits offered by the AAR, there still remains some variation among the computed AARs for different Monte Carlo samples. In addition, it is difficult to ascertain the exact AAR for each die as it is challenging to accurately measure process-variation parameters for a die. However, we need to select one AAR (which is denoted hereafter as Test-AAR) that will be used to screen defective dies based on their adjusted signatures. Hence we pose the following question: How should we choose the appropriate Test-AAR out of the AARs computed using Monte Carlo simulations? One solution is to set the Test-AAR for each mode equal to the common sub-range of the AARs of all Monte Carlo samples for that mode. However, this selection will provide a rather strict and over-constrained Test-AAR. The Test-AAR can be relaxed if we exploit inherent properties of the power switches. Suppose for example that for a given die, the signature for an arbitrary power-off mode violates the upper bound of its own AAR due to a fault. Then the effect of this fault is that the number of cycles required for waking up the core increases and thus the fault must be detected as it affects the functional correctness of the core. Therefore, the upper bound of the Test-AAR must ensure that if the signature of any die violates the upper bound of its own AAR, it also violates the upper bound of Test-AAR (and thus the defect is detected). On the other hand, if the signature of the die violates the lower bound of its AAR due to a fault, then the effect is less serious as it does not cause functional errors; it may increase the static power and the number of wake-up cycles may reduce in this case.

Based on the above observations, we select the upper bound of the Test-AAR as the minimum of the upper bounds of all individual AARs for all samples generated using Monte Carlo simulations. In that way, no die signature that is above the upper bound of its own AAR can be interpreted as being from a fault-free die. On the other hand, there is no strict constraint for the lower bound and it is set equal to the average value of the lower bounds of the AARs of the individual samples produced during Monte Carlo simulations. Alternative lower bounds that are more strict or loose can be also selected depending on the increase of static power that can be tolerated.

Note that with this approach, there might be dies providing signatures that are higher than the upper bound of Test-ARR but which still work correctly. These correspond to the cases where the upper bound of their ARR is higher than the upper bound of the Test-ARR used. Due to the strict constraint enforced for the selection of upper bound of Test-AAR, these cases are misidentified as parametric faults or as process variation effects. Nevertheless, we can easily cope with this problem using the repair mechanism presented in the next section; hence no yield loss is introduced. On the other hand, suppose a signature violates the lower bound of its own AAR due to a fault but it does not violate the lower bound of Test-AAR. This situation can arise due to the loose nature of the lower bound. Then, the signature is accepted as being for a fault-free die and the parametric fault is not detected. However, in this case, the number of wake-up cycles is not increased (in fact it may decrease) and only the static power slightly increases. Thus the die will operate correctly with no logic error.

In practice, the tester first receives the reference signature from the chip and then it receives the signatures for each power-off mode. Subsequently the tester computes the D_i value of each die with respect to the golden reference signature, adjusts the signatures by adding D_i to each one of them, and compares the adjusted signatures with the Test-AARs computed using simulation. If any signature lies outside that range, a catastrophic or parametric fault is detected else the chip is defect-free with respect to the power-off switches.

V. Repair And Calibration Scheme & Methodology

In this section, we present the design and test of a power-switch architecture that repairs most faults affecting switches M_0 and M_1. Both stuck-open and parametric faults are easy to repair if transistors M_0 and M_1 are duplicated. If such redundancy is available, defective transistors can be replaced after the defect is diagnosed. Note that these transistors are small; therefore duplicates can be inserted at negligible cost. The selection of the non-defective transistors can be done post-manufacture using a programmable structure, e.g., fuses commonly used for built-in memory self-repair. This mechanism cannot repair stuck-on faults. However, as we have shown in

TABLE 4. TEST-ADJUSTED ACCEPTABLE RANGES

Golden Reference Signature = 202	Snore	Dream	Sleep
No of cycles for wake-up	8	5	3
Range of acceptable voltage values at *V-GND*	720~795 mV	510~569 mV	390~450 mV
Test-AAR	160~164	72~80	23~29

Section III, stuck-on faults do not affect the reliability of the cores as they do not increase the wake-up time. Their only effect is that they reduce the benefits in static power reduction offered by the power switches during the power-off modes. Thus, even in the presence of such defects, parts containing these switches can be shipped with no adverse impact on yield.

It is more difficult to be tolerant to process variations. In the presence of process variations either the number of wake-up cycles increases or the static power of the core at an intermediate power-off mode increases. Both behaviors result from small variations in the voltage level of the virtual ground node. For example, if the *V-GND* voltage in the dream mode is less than that in the nominal case (without process variations), the static power consumption increases. However, if the width of M_0 is slightly reduced then there will be an opposite effect on *V-GND,* which will counterbalance the effect of the process variations. Since we cannot alter the size of transistors after manufacturing, we provide embedded redundant structures. Specifically, each transistor M_0, M_1 is fabricated as a set of transistors $\{...M_j^{-2a}, M_j^{-a}, M_j, M_j^{+a}, M_j^{+2a}, ...\}$ with different aspect ratios connected in parallel ($j = 0$ for transistor M_0 and $j = 1$ for M_1). Let $\beta = \alpha, 2\alpha, \ldots$ etc. Then the aspect ratios of transistors $\left(M_j^{-\beta}, M_j^{+\beta}\right)$ are calculated as follows:

$$W_{M_j^{+\beta}} / L_{M_j^{+\beta}} = (1+\beta/100) W_{M_j} / L_{M_j}, W_{M_j^{-\beta}} / L_{M_j^{-\beta}} = (1-\beta/100) W_{M_j} / L_{M_j}$$

where $\beta = k \cdot a$, $\alpha \in (0\%, 100\%)$ and $k = 1, 2, \ldots$. The parameter a and the number of transistors of each set are selected in the following way. The smaller the value of α or the larger the number of transistors used, the more tolerant is the proposed scheme to process variations. Note that the use of transistors in each set with their aspect ratios shifted by a%, $2a$%, $3a$%, … above or below the nominal value increases the probability that one of the transistors of each set provides the required voltage at *V-GND* in the presence of process variations.

Both faults and process variations can be tolerated by duplicating the sets of transistors $\{...M_j^{-2a}, M_j^{-a}, M_j, M_j^{+a}, M_j^{+2a}, ...\}$. The area overhead increases, but considering the very small area overhead of the M_0, M_1 transistors, the overhead of the redundancies remains insignificant. Even if 10 redundant transistors are used for M_0 and another batch of 10 for M_1 the overhead still remains lower than the overhead of the method proposed in [13] (for the case at hand they occupy less than 0.0002% of the area occupied by the main power switch M_p). The redundant transistors have the same connectivity as the transistors forming M_p (they are both connected between ground rail and *V-GND* node) and thus they can be used to replace a few of the transistors forming the large M_p power switch (they remain on during active mode of operation). In this case, no additional area is required for the repair and calibration structure. However, it is necessary to include as few redundant transistors as possible, since the area of the control circuit (e.g., a decoder) for the selection of the redundant transistors needs to be taken into account as well.

TABLE 5. SIGNATURES FOR CATASTROPHIC FAULTS

Fault	Power-Off Mode	Test-AAR	Adjusted Signature
M_0 stuck-open	Dream	72~80	181
M_0 stuck-on	Snore	160~164	88
M_1 stuck-open	Sleep	23~29	181
M_1 stuck-on	Snore	160~164	30

VI. SIMULATION RESULTS

We implemented a VCO with 9 current-starved inverters driving an 8-bit binary counter. For the voltage range of interest (350 to 800 mV) at *V-GND* the frequency of the VCO is in the range 0.64 GHz to 9.1 GHz. These voltage ranges were selected based on the fault analysis presented in Section III. The VCO-Counter system counts for 20 ns each time. The clock frequency of the core was set equal to 1 GHz.

The golden reference signature and the Test-AARs are calculated using a 30-sample Monte Carlo experiment (they are reported in decimal form in Table 4). For example, if the core under test is put into the snore mode, any signature (adjusted to the reference signature) in the range 160-164 is acceptable and the power-off switches are defect-free for this range.

Catastrophic and parametric faults are inserted to both the M_0 and M_1 power switches and the adjusted signatures are computed using HSpice simulation. The adjusted signatures for catastrophic faults are listed in Table 5. For each fault, we report the Test-AAR as well as the adjusted fault signatures. It is evident that all catastrophic faults are detected, and the signatures in the presence of catastrophic faults fall far outside the Test-AARs. Table 6 presents the adjusted signature values as well as the number of wake-up cycles for each parametric fault studied in Section III. The boldfaced signatures correspond to faults that yield signatures outside the Test-AAR. The boldfaced data (number of cycles) correspond to faults that affect the number of wake-up cycles of the core. Based on this notation, there are four types of faults:

1. *Faults with both entries (no. of cycles and signature) bold*: these faults affect the wake-up time and fail the test.
2. *Faults with both entries non-bold*: these faults do not affect the wake-up time and pass the test.
3. *Faults with only the signature bold*: these faults do not affect the wake-up time but they fail the test (incorrect classification).
4. *Faults with only the number of cycles bold:* these faults affect the wake-up time but they pass the test (test escape).

TABLE 6. SIGNATURES FOR PARAMETRIC FAULTS

Deviation		Parameter: Width of M_0		Parameter: Length of M_0		Parameter: V_{th0} of M_0	
		cycles	signature	cycles	signature	cycles	signature
-30%	Dream Mode 5 cycles Test-AAR:72~80	**6**	**119**	**3**	**32**	**4**	**69**
-20%		**6**	**107**	**4**	**53**	**4**	72
-10%		5	**92**	**4**	72	5	79
+10%		5	76	5	**94**	5	**93**
+20%		**4**	**68**	5	**103**	5	**101**
+30%		**4**	**61**	**6**	**110**	**6**	**108**
Deviation		Parameter: Width of M_1		Parameter: Length of M_1		Parameter: V_{th0} of M_1	
-30%	Sleep Mode 3 cycles Test-AAR:23~29	4	**58**	**2**	**7**	**2**	**22**
-20%		3	**46**	**2**	**15**	3	24
-10%		3	**37**	3	23	3	27
+10%		3	24	3	**36**	3	**35**
+20%		**2**	**20**	3	**42**	3	**41**
+30%		**2**	**16**	**4**	**48**	**4**	**48**

It is obvious that in contrast to catastrophic faults, parametric faults give signatures outside the Test-ARRs but not far from these ranges. We can make the following observations:

a) All critical faults (i.e., those that increase the number of wake-up cycles) are detected by the test mechanism as they fail the test (these cases are highlighted in Table 6).
b) The vast majority of parametric faults that decrease the number of wake-up cycles (increase the static power) fail the test. Only 2 cases pass the test; their signatures are equal to the lower bound of the Test-AAR.
c) A few parametric faults fail the test even though they do not cause any increase in the number of wake-up cycles.

Faults in Category (a), which are of the highest concern, are all detected. Very few faults in Category (b) are missed, but these faults do not affect the reliability of the core's logic operation. Note that faults in category (b) that increase the static power significantly, i.e., they reduce the number of wake-up cycles by two or more, are easily detected as they give signatures outside the Test-AARs (e.g., the case of 30% reduction in the length of M_0). Finally, the faults in Category (c) do not affect the core operation, but are identified as problematic due to the strict way the upper bound is selected. These faults invoke the repair mechanism which selects the duplicate transistor to tolerate them. If the duplicate transistor suffers from the same problem (e.g. in the case that the malfunction is caused by process variations and not by parametric faults) then a redundant transistor with higher aspect ratio is selected as it provides lower signature value. Even though this choice causes a slight increase in static power, there is no yield loss.

It is obvious that catastrophic faults can be easily diagnosed by the proposed test mechanism as different signatures correspond to different catastrophic defects at each power-off mode. Moreover, the faulty switches in the case of parametric faults can be easily identified. Thus, all faults are easily diagnosed.

Finally, we implemented the repair and calibration structure for α = 10% and k = 3, and we ran simulations considering process variations, catastrophic and parametric defects. We verified that among all redundant transistors there is at least one transistor offering the required wake-up time in the presence of process variations. All catastrophic faults (except of the stuck-on faults) as well as the critical parametric faults of category (a) were corrected by replacing the faulty power switch by the appropriate redundant switch. A percentage of 5.5% of faults were found to be parametric faults that increased static power and were not detected by the test mechanism due to the loose lower bound used for computing Test-AARs. Another 30% of the faults fall in Category (c) and thus they are unnecessarily corrected. However, in the last two cases where the faults are either not detected or unnecessarily corrected, the only effect on the core operation is a slight increase in static power during the power-off mode. However, the static power reduction offered by the faulty power switches at the various power-off modes is hardly 1% lower than the reduction offered by the non-faulty switches. Therefore, even though they are faulty, these switches still offer an effective means for decreasing static power during power-off modes.

VII. Conclusion

We have presented a new method to test power switches offering one or more intermediate power-off modes. The proposed method efficiently detects all catastrophic and most parametric faults for a highly effective power-switch architecture and it also provides the capability to locate defects in these switches. Finally, a simple solution is proposed to repair almost all catastrophic and parametric faults and tolerate process variations affecting all parts of the power-switch circuit.

References

[1] ARM 1176JZ(F)-S documentation, http://www.arm.com/products/CPUs/ARM1176.html.
[2] M. Chowdhury, J. Gjanci and P. Khaled, "Innovative Power Gating for Leakage Reduction", in Proc. ISCAS, pp. 1568-1571, 2008.
[3] S. K. Goel, M. Meijer, and J. P. de Gyvz, "Testing and diagnosis of power switches in SOCs", Proc. of IEEE ETS, pp. 145-150, 2006.
[4] S. Ingunji, "Case Study of Low Power MTCMOS based ARM926 SoC: Design, Analysis and Test Challenges", in Proc. Int. Test Conf., 2007.
[5] Intel Corp, "Intel XScale Core Developer's Manual, 2003", http://developer.intel.com/design/intelxscale/
[6] S. Kim et al.,, "Experimental measurement of a novel power gating structure with intermediate power saving mode", in Proc. Int. SLPED., pp. 20-25, 2004.
[7] S. Kosonocky et al., "Enhanced Multithreshold (MTCMOS) Circuits with Variable Well Bias", in Proc. IEEE ISPLED, pp. 165-169, 2001.
[8] D. Lackey et al., "Managing power and performance for system-on-chip designs using voltage islands'. Proc. ICCAD, pp. 195–202, 2002.
[9] R.C. Lopez, et al., "Reuse-based methodologies and tools in the design of analog and mixed-signal integrated circuits", Springer, 2006.
[10] E. Pakbaznia and M. Pedram, "Design and Application of Multimodal Power Gating Structures", in Proc. ISQED, pp. 120-126, 2009.
[11] R. Puri et al., "Pushing ASIC performance in a power envelope", Proc. IEEE/ACM Design Automation Conference, pp. 788–793, 2003.
[12] R. Puri, D. Kung and L. Stok, "Minimizing power with flexible voltage islands" in Proc. IEEE ISCAS, pp. 21–24, 2005.
[13] H. Singh et al., "Enhanced Leakage Reduction Techniques Using Intermediate Strength Power Gating", IEEE Trans. VLSI, vol. 15, No 11, pp. 1215-1224, 2007.
[14] L. Souef C. Eychenne, E. Alie, "Architecture for Testing Multi-Voltage Domain SOC", Int. Test Conf., Paper 16.1, 2008.
[15] Transmeta Corporation. "Crusoe processor documentation 2002" http://www.transmeta.com.
[16] Z. Zhang, X. Kavousianos, K. Chakrabarty and Y. Tsiatouhas, "A Robust and Reconfigurable Multi-Mode Power Gating Architecture", in Proc. IEEE Int. Conf. on VLSI Design, pp. 280-285, 2011.
[17] K. Chakrabarty, X. Kavousianos and Z. Zhang, "Power Switch Design and Method For Reducing Leakage Power In Low-Power Integrated Circuits", US Patent Application no. 12/882,776, filed by Semiconductor Research Corporation, September 15, 2010.
[18] W. Zhao and Y. Cao, "New generation of Predictive Technology Model for sub-45nm early design exploration," IEEE Trans. Electron Devices, vol. 53, no. 11, pp. 2816-2823, 2006.
[19] L. Dermentzoglou, A. Arapoyammi and Y. Tsiatouhas, "A Built-In Self-Test Technique for RF Mixer", IEEE International Symposium on Design and Diagnostic of Electronic Circuits and Systems, pp. 88-92, 2010.