DART: Dependable VLSI Test Architecture and Its Implementation

DART: Dependable VLSI Test Architecture and Its Implementation

Yasuo Sato[†]*, Seiji Kajihara[†]*, Tomokazu Yoneda[‡]*, Kazumi Hatayama[‡]*, Michiko Inoue[‡]*
Yukiya Miura[††*], Satoshi Ohtake[‡‡]*, Takumi Hasegawa[†††], Motoyuki Sato[†††] and Kotaro Shimamura[‡‡‡]
† Kyushu Institute of Technology, Fukuoka, Japan
‡ Nara Institute of Science and Technology, Nara, Japan
†† Tokyo Metropolitan University, Tokyo, Japan
‡‡ Oita University, Oita, Japan
††† Hitachi Ltd., Information&Telecommunication Systems Company
Hardware MONOZUKURI Division, Tokyo, Japan
‡‡‡ Hitachi Ltd., Hitachi Research Laboratory, Ibaraki, Japan
* Japan Science and Technology Agency, CREST, Tokyo, Japan

Abstract

Although many electronic safety-related systems require very high reliability, it is becoming harder and harder to achieve it because of transient faults such as soft errors or delay-related failures, which are caused by decreased noise margin. This paper describes a technology named DART and its implementation. The DART repeatedly measures the delay of a circuit and the amount of degradation in field, in consequence, confirms the marginality of the circuit. The system employing the DART will be informed the significant reduction of delay margin in advance and be able to repair it in an appropriate time. The DART also equips a technique to improve the test coverage using the rotating test and a technique to consider the test environment such as temperature or voltage using novel ring-oscillator-based monitors. The authors applied the proposed technology to an industrial design and confirmed its effectiveness and availability with reasonable resources.

1. Introduction

Many electronic safety-related systems such as network equipment, servers, automobiles, aircrafts, satellites, medical treatments or social infrastructure systems require very high reliability. However, it is becoming harder and harder to achieve it because of increasing system complexity and geometrically reducing semiconductor fabrication process. In a case of automobiles, a failure late less than 10 FIT is typically required in field for a single chip. To tackle this problem, the international standards such as IEC61508 [1] or ISO26262 [2] have been established. They propose systematic approaches which improve functional safety of the systems.

Trends of random hardware failures show that transient faults, such as soft errors or delay-related faults which are caused by decreased noise margin due to process variations or device aging, are increasing despite that permanent faults are decreasing [3]. Process variation is caused by the successive reduction in feature sizes well below optical wavelengths, which are difficult to control on silicon in high accuracy. In consequence, on-chip-variation, die-to-die-variation or parameter-shift on individual chip would significantly reduce delay margins. The device aging phenomena due to fine device structures such as hot carrier injection (HCI), negative bias temperature instability (NBTI), time dependent dielectric breakdown (TDDB), electro migration (EM) or stress migration (SM) are well-known [4-5]. These phenomena degrade device performance in a long term and reduce tolerance to noisy variation in field, which includes a soft-error, a voltage-shift or a thermal-induced additional delay.

Various approaches are investigated to overcome these issues. The first approach is process-variation-aware design and aging-aware design [6-9]. This approach requires performance degradation in exchange for design margin. Another tough problem is the difficulty of predicting reliability in life time because the amount of aging depends not only on the process variation of each chip but also on its functions and temperature environment. The second approach is on-line testing [3, 10-13]. Double/triple voting logics, self-checking circuits, fail-safe circuits, soft-error hardened circuits or monitors such as reliability indicators, current monitors or thermal monitors, are widely used for detecting soft-errors and transient noise problems. They are usually application-specific and require recovery or mask functions. Their coverage may not be enough because of their area penalty. As the last approach, some techniques using built-in self-test (BIST) during test mode in field (i.e. at a power-on/off time of a system or at some system's idle time) are proposed [14-21], and they have potentials to detect faults with high coverage. However, the assigned short time makes difficult to improve the coverage, and the fault detection after its occurrence may not be enough to reduce mean time to repair (MTTR) of the system. For instance, a long repair time of power plants or rail-way systems will cause severe confusion in our life.

Our proposed technology is categorized in the third approach mentioned above. It measures the delays of circuits and the amount of degradation in field, in consequence, confirms the marginality of the circuits. The

system will be informed the significant reduction of delay margin in advance and be able to repair it safely in an appropriate time (i.e. reduce MTTR of critical failures [1, 3]). The technology also utilizes some techniques to improve the test coverage using the rotating test [18, 22-24] and so on [25-27], and some techniques to consider the test environment such as temperature or voltage using novel ring-oscillator-based monitors [28] and so on [29, 30]. The authors applied the proposed technology that is called DART (Dependable Architecture with Reliability Testing) [31, 32] to an industrial design and confirmed its effectiveness and availability with reasonable resources.

This paper is organized as follows. Section 2 describes our concept and the applied design model. Section 3 describes the proposed design mostly from DFT perspective. Section 4 explains DART's test design. Section 5 discusses the experimental results using an industrial design. Section 6 concludes the paper.

2. Concept and Design Model

2.1 Background

The performance of a chip degrades as time passes. The aging speed of HCI or NBTI is reported as follows [5]:

$$\Delta D \propto (\alpha t)^n \qquad (1)$$

where t is the elapsed time, α is the time ratio that the transistor is accelerated depending on system functions, and n is a constant that depends on the fabrication process. HCI is accelerated when current flows through an NMOS transistor (i.e. $V_{ds} > 0$) and n is around 0.45. NBTI is accelerated when the transistor is on (i.e. $V_{gs} < 0$, $V_{gb} < 0$) and n is around 0.16 or 0.25. It is also affected by temperature. Usually, 3-5 percent delay degradation is taken into the design margin [8-9]. This shows a very slow speed of degradation ΔD. Then, it suggests that ΔD can be known accurately by repeatedly measuring with time and care in field. It also suggests that a set of partitioned tests does not lose its fault coverage for this kind of faults (See the rotating test in Section 4).

2.2 Concept

Fig. 1 shows the concept of the proposed technology. The largest delay D will be repeatedly measured in field until it reaches the threshold delay D_{th}. Once it reaches D_{th}, the system will be informed the risk with an alarm (the significant reduction of delay margin) before it causes a critical failure, and be repaired safely in an appropriate time. Therefore, a potential failure will be prevented and it means the MTTR reduction of critical failures. Let λ_{orig} be the original critical failure rate and DC be the diagnosis coverage (test coverage of critical faults by the given test), λ_{orig} will be reduced to $\lambda_{orig} \times (1\text{-}DC)$. D_{th} is assigned so as $D_{th} + M_{err} \le D_{crit}$. Here, M_{err} is the error of delay measuring and D_{crit} is the minimum allowable delay of the chip (e.g. the machine cycle or less than the machine cycle). Assigning D_{th} in this way, there is no risk of missing detectable faults. However, larger M_{err} would make the interval between the alarm and the actual failure longer. Measuring error due to random noise (e.g. noise due to hazard, etc.) can be reduced evaluating the moving average of successive measured values (e.g. $\frac{1}{N}\sum_{i=0}^{N-1} D_{M-i}$, D_P is the measured delay at time P). It is known that the moving average of N values makes the random error $N^{-1/2}$-times smaller.

Table 1 clarifies the failures that the proposed technology focuses on. Delay-related failures such as HCI, NBTI and TDDB-induced failure, which will be detected before actual failures, are mainly targeted. These failures can be common cause failures in a multiple voting logics [1, 3]. Noise-related failures are not directly targeted on. However, as the delay margin is repeatedly measured and confirmed, possibility of these failures will be also reduced. For permanent faults, the technology has almost the same capability as the conventional field BIST-based technologies. Although on-line testing techniques referred in Section 1 cover most of the failures, their coverage may not be so high because of area penalty. And it is not possible to predict the failures. The proposed technology can be used not only for field test, but also for system debugging or tuning to evaluate delay margin reduction due to process variation or aging.

The detail of the technology is described in Section 3 and 4. We named this technology *DART* and four factors are considered in the *DART*: *D* (Degrade Factor), *A* (Accuracy), *R* (Report) and *T* (Test Coverage).

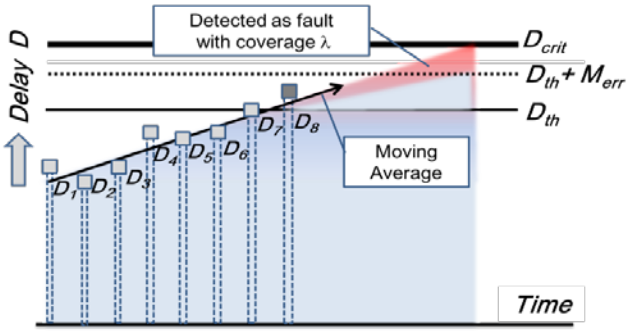

Fig. 1 Concept of Proposed Technology

Table 1 Focused Fault

Fault	Characteristic	Test Method			
		Multiple Voting	On-line Test	BIST	Proposed (DART)
Stuck	Permanent	Full	Partial	Full	Full
EM/SM					
HCI/NBTI	Delay-related	Partial	Partial	Full	Full (Advance)
TDDB				Full	
Noise	Noise-related	Full	Partial	NA	Partial (Advance)
Soft Error				NA	

2.3 Design Model

The authors applied the DART technology to a design of an industrial chip that is related to a safety-related system [3, 33]. The design consists of double voting logics (two micro- processers (μP), comparison logic and other control logic). Although it has high dependability, the reliability technique should be revised to tolerate delay-related failures that would be more severe in future deep-submicron process. Avoiding sudden system failure by increasing MTTR and preventing social confusion was the first priority.

Table 2 shows the overview of the design in 90 nm technology. It has 7.2M gates of 12 clock domains, in which the fastest clock frequency is 300 MHz, 234 static RAMs and 30 register file memories. The system has chances to diagnose each chip once a day. However, as most of the time is assigned to system operations, backup/recovery of system data and system reset, the assigned time to field test was less than 200 *ms*.

One important issue is what to measure in field. The system will be used in wide range of temperature and its operations will be different according to regions. Therefore, it would be difficult to measure delay margin correctly in real operation. DART focuses not only in measuring the delay, but also in measuring the delay increase monitoring the log data of field test. To exclude temperature or voltage variation, ring-oscillator-based special monitors (TVM: temperature and voltage monitor) are embedded in a chip and the measured delay of user circuit is normalized to nominal conditions.

3. Design Structure of DART

3.1 Design Structure

Fig. 2 shows the design structure of DART for the target design. It reuses DFT for production test as much as possible. The target design is equipped with IEEE 1149.1 (JTAG)-based BIST architecture, where both logic BIST (LBIST) and memory BIST (MBIST) are applied through the test access ports (TAP). The DART controller manipulates LBIST and MBIST through TAP. Five JTAG pins (TCK, TRST, TMS, TDI and TDO) are switched to the signals to/from the DART controller during the field test. One mode pin (DRTPIN) is added for field test. Input user pins are isolated using the boundary scan cells, and output user pins are isolated on board during field test. Some memories, which keep test information about DART or some parts of user data, are protected from writing with dis-enabling write-enable signals of RAMs.

In addition to the original BIST architecture, DART uses a variable test timing generator. The test timing generator shares users' PLL for testing. PLL clocks are modified to generate test clocks, which consist of scan clocks and release-capture clocks. The release-capture clock timings are shortened on-the-fly using the on-die clock-shrink (ODCS) method [34] without changing PLL frequency. This time, it is designed using a buffer-based chain, which means a unit timing is the delay of a buffer (e.g. 40 *ps*).

Five TVMs are embedded in the chip to estimate the temperature and the voltage of a domain under test (DUT) during test using measured frequency of ring oscillators (RO). The measured delay values of the ROs are compared to pre-computed characteristics with respect to temperature and voltage. Then, temperature and voltage values in the field test environment are estimated and the measured delays normalized using these values. The detail will be described in Section 3.3.

Table 2 Overview of the Design

Items	Values
Gates	7,198,340
Flip Flop	355,961
LSI pin	451
Memory(SRAM)	234
Memory(RGF)	30
Clock Domain	13
Frequency (MHz)	300/150/75/25

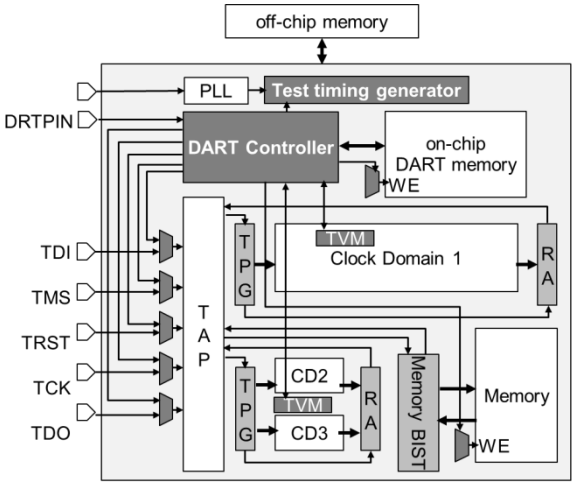

Fig. 2 Design Structure of DART

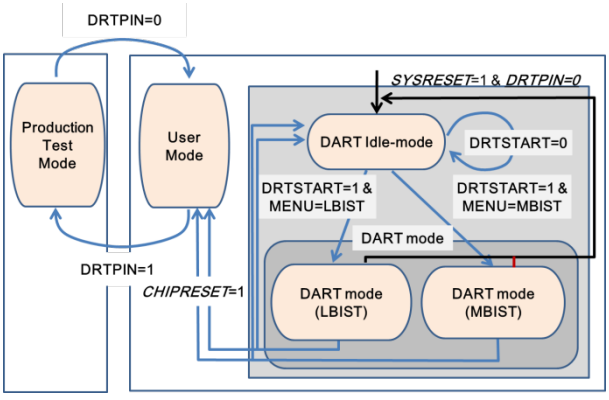

Fig. 3 State Diagram

3.2 Interface

Fig. 3 shows the state diagram of the design. It has the following three major modes.

Production test mode: a mode where the production test is executed. DUTs and the DART controller are tested in this mode. Most parts of the DART controller are tested by BIST and a functional test is needed for the remaining part.

User mode: a mode where the user functions are executed. The μP is used in the user mode. The DART controller is idle until it is invoked by the μP user program.

DART mode: a mode where the DART test is executed. This mode is a special test mode where the DART controller and the on-chip DART memory are still in the functional mode, whereas other DUTs are in the test mode. The DART mode consists of LBIST mode and MBIST mode.

The major difference between the production test mode and the DART mode is that the DART controller is tested in the production test mode and is not tested in the DART mode. For differentiating these two modes DRTPIN is used. The DART controller is idle under the user mode and waits for the invoking signal from the μP. The user program's procedures to invoke DART are as follows:

(a) Backup user register data to memories (protected SRAMs),

(b) Read log data and select a test session (See the detail in Section 4) and set the test specification corresponding to the selected test session to a memory (a protected SRAM for DART),

(c) Invoke the DART test (Send the invoking signals),

(d) When the DART mode is over, the chip-reset signal is sent from the DART controller and the μP restarts (As restart period is very unstable, special care is needed),

(e) Recover user register data from memories (See (a)),

(f) Store the log data to a non-volatile memory,

(g) Analyze the log data and check whether the delay margin is critical or not.

The procedures (b), (f) and (g) can be implemented as hardware logic. However, as this is the first experience and there is no know-how based on the real data, the authors have assigned these functions to the user software program.

The memory mapping strategy is as follows. An off-chip non-volatile memory (flash memory) is used for storing several test sessions and the log data. Although its capacity is rather large, it has the limitation of rewritable number (e.g. less than 10k times) and the unit of writing is 256kB. The on-chip memory used for DART is an 8kB SRAM. Therefore, when a test session for the current test chance is selected, the specification data of the test session is copied from the off-chip memory to the on-chip memory (DART memory). The test specification data mainly consists of clock domain and its timing information, BIST structure information (e.g. pattern generator, scan chain, etc.) and BIST test information (e.g. seeds, signatures, options). The log data is stored to the DART memory and copied to the off-chip memory, after the DART test mode is over.

3.3 TVM design [28]

A TVM consists of three types of ROs, which have different frequency characteristics regarding temperature and voltage (Fig. 4). By measuring three frequencies, temperature and voltage are derived using the following linear equations.

$$\Delta T = a_1\Delta F_1 + a_2\Delta F_2 + a_3\Delta F_3 \quad (2)$$

$$\Delta V = b_1\Delta F_1 + b_2\Delta F_2 + b_3\Delta F_3 \quad (3)$$

Here, ΔT, ΔV, ΔF_1, ΔF_2 and ΔF_3 are differences from the initial measurement in field. Therefore, the initial measurement is a kind of characterization to reduce process variation impact. The coefficients are defined with SPICE simulation. Using three frequencies, the 2-dimensional factors are cancelled and linear equations are derived. To improve the accuracy, the ranges of T and V are divided into three intervals, respectively, which overlap considering process variation. The TVM has the following features.

- It has a special structure to tolerate NBTI.
- It consists of logic gates in a standard cell library.
- It monitors temperature and voltage concurrently.
- No reference voltage or current is required.
- No ADC or DAC is required.
- It can be reused for other process monitoring purposes.

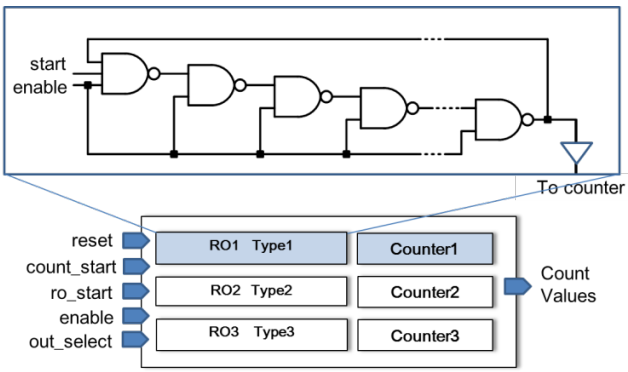

Fig. 4 TVM Design [28]

4. Test Structure of DART

4.1 Test Hierarchy

Fig. 5 shows the test hierarchy of DART test. Rotating test consists of some test sessions that are executed in an assigned time during user operation. Rotating test tests a part of the design and is executed with only one clock at a time. Test menu defines the test specification of the

focused clock domain (intra-domain/inter-domain). It includes several seeds for LBIST test pattern generators (TPG) and options that are used for test coverage improvement or low power test. One test group is a test that targets a part of memories in a design with different test algorithms for test quality-related reason or test power reason. For instance, a large SRAM and a small register file are tested with different algorithms.

4.2 Test Flow

LBIST [35-38] has 10 user-defined instructions, while MBIST has 14 user-defined instructions. All the instructions or data are provided through the TAP interface using the TMS and TDI ports. LBIST is repeated with different test timing until it finds the minimum test timing that the test passes. To reduce the number of iterations, the last test timing used for the same rotating test is read from the log data and used as the first test timing in the current test.

To execute some designated intra/inter-domain test, TPGs and response analyzers (RA) related to the DUTs are initialized first, where the related TPGs and RAs mean TPGs and RAs connected directly with the DUTs and also TPGs connected indirectly with DUTs through some clock domains. The outline of the test flow in the LBIST mode is shown in Fig. 6, where D is a set of DUTs, M_i is a set of test menus for DUT d_i, and SS_{ij} is a set of pairs of seed and signature for test menu m_{ij} for DUT d_i.

DART has two test modes of LBIST and MBIST. In the LBIST mode, the DART controller repeatedly applies delay test using LBIST with variable test timings. According to each test result, it varies test timing and finds the minimum timing with which the delay test passed. During test execution, the DART controller also measures temperature and voltage of the circuit by using the TVMs, and the measured temperature and voltage are used to analyze the delays of paths more accurately. In the MBIST mode, some memory test algorithms are applied with a fixed test clock, and the DART controller finds the applied tests are passed or failed.

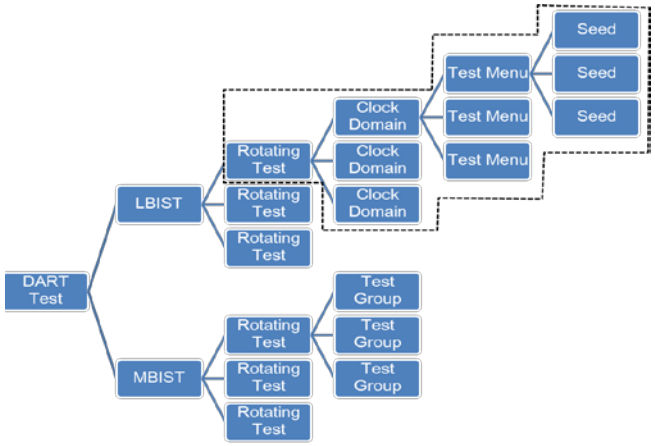

Fig. 5 Test Hierarchy

LBIST test flow

initialize all of TPGs and RAs;
for each d_i in D **do** /* clock domain */
  set up clock configuration of d_i for clock generator;
  set up initial test timing for d_i;
  /* read from the last log */
  repeat
  for each m_j in M_i **do** /* test menu */
    set up option parameters to LBIST controller;
    for each ($seed_k$, $signature_k$) in SS_{ij} **do**
      shift-in $seed_k$;
      run test;
      shift-out response and compare it with $signature_k$:
      /* the next seed is shifted in at the same time */
  initialize TPGs and RAs related to d_i ;
  /* prevent X-propagation to other domain */
  if test is passed for d_i
    then decrease test timing;
    else increase test timing; /* test failed */
  until the minimum passed test timing is found
  store the timing into the latest log data;

Fig. 6 LBIST Test Flow

5. Experimental Results

5.1 Test Data Volume and Test Time

The target design has strict constraints on both test data volume and test time as shown in Section 2.3. The on-chip memory size available for DART is less than 8*kB*, and test time acceptable for one test chance is 200*ms*.

We firstly estimated the test data volume needed for the LBIST to test 12 clock domains in one test chance, where the DUTs include 12 intra-domains and 52 inter-domains. As the complexity of inter-domains is simpler than those of intra-domains, we tried to apply more seeds to intra-domains than inter-domains. Table 3 shows the relation between the number of seeds and the required test data volume. The required test data includes not only seeds and signatures themselves but also the information about the test specifications and the log data for the domains.

Table 3 Test Data Volume Estimation (N_{TC} = 1)

# of seeds		test data volume (*kB*)
intra-domain test	inter-domain test	
1	1	10.3
2	1	11.3
2	2	16.6
3	1	12.3

It is noted that an inter-domain test uses more memory than an intra-domain test since it requires the seeds for both domains. In general, test data volume, *TDV*, is calculated as follows,

$$TDV = c_1 NS_{intra} + c_2 NS_{inter} + c_3 \quad (4)$$

where, *c1*, *c2* and *c3* are constants depending on LBIST structure, and NS_{intra} and NS_{inter} are the number of seeds for intra-domain test and inter-domain test, respectively.

In Table 3 we assume all tests are executed in one test chance, that is, the number of test chances, N_{TC}, is 1. Because of the limitation of the on-chip DART memory size and test time, we use several test chances to execute all tests. Table 4 shows the case for N_{TC} = 6.

Table 4 Test Data Volume Estimation (N_{TC} = 6)

# of seeds		test data volume
intra-domain test	inter-domain test	(*kB*)
1	1	4.8
2	2	5.6
5	5	8.0
11	4	7.9
15	3	8.0
22	1	7.9

From these results, we can see that multiple test chances for all tests are effective to keep the on-chip DART memory size low while increasing the number of seeds to enhance test quality.

We next estimated the test time needed for the LBIST. The test time consists of test setup time, test application time and test result collection time. The test setup time includes test circuit configuration for specified clock domain(s), parameter setting for specified options, seed setting, and so on. Assume that the maximum scan chain length is 300, the number of patterns for each seed is 64, test run is repeated 4 times to decide the minimum passed test timing of the DUT, and DART circuits operate at 25MHz, and so on. Then we estimated the relation between test time and the number of seeds as shown in Table 5 and Table 6 for the cases of N_{TC} = 1and N_{TC} = 6, respectively.

Table 5 Test Time Estimation (N_{TC} = 1)

# of seeds		test time
intra-domain test	inter-domain test	(ms)
1	1	229.4
2	1	271.0
2	2	458.7
3	1	312.6

Table 6 Test Time Estimation (N_{TC} = 6)

# of seeds		test time
intra-domain test	inter-domain test	(ms)
1	1	28.6
2	2	57.2
5	5	143.0
11	4	163.0
15	3	169.1
22	1	174.3

From these results, we can see that multiple test chances for all tests are effective to satisfy test time constraint. It should be noted that the number of seeds and the number of test patterns for each seed are manageable to meet both of the DART memory size and test time constraints.

5.2 Seed Selection

The selection of effective seeds is also crucial to realize high test quality under strict test constraints. We selected seeds using a method based on seed-ordering [25, 26] with statistical delay quality level (SDQL) as the metric of test quality [39, 40]. Though our proposed method considers SDQL, we simply consider transition fault coverage FC in this evaluation. The outline of the method is summarized in Fig. 7.

Seed ordering

initialize *S* to empty; /* *S*: ordered seed set */
for each *s* in *Sbase* **do** /* *Sbase*: base seed set */
run fault simulation for the expanded patterns from *s* to obtain *a list of detected transition faults*;
for *s* with the maximum FC **do**
move *s* from *Sbase* to *S*;
repeat
for each *s* in *Sbase* **do**
calculate FC for $S \cup s$
for *s* with the maximum FC **do**
move *s* from *Sbase* to *S*;
until *Sbase* becomes empty;

Fig. 7 Outline of Seed Selection Method

We evaluated the effectiveness of the method using three of clock domains of the target design. The profile of each domain is shown in Table 7.

Table 7 Profile of Target Clock Domain

domain name	clock frequency	# flip-flops
T0	300MHz	78k
T2	150MHz	39k
T7	75MHz	27k

The fault coverage curve using selected seeds is compared with that for random pattern test using one randomly selected initial seed. The results for clock domains *T0*, *T2* and *T7* are shown in Fig. 8, Fig. 9 and Fig. 10, respectively. For domains *T0* and *T2*, the cases where basic test pattern count (64 patterns) are expanded from each seed are evaluated, while for *T7* we added two cases, case2 and case3, where double and triple test patterns were expanded from each seed. For all cases, we add 50k random patterns at the end to show that the method can achieve higher fault coverage in less pattern count.

From these results, we can see the following two facts. First, the proposed seed-selection derives enough reduction of test pattern count without sacrificing the target fault coverage. For example, in case 1 of Fig. 10, fault coverage of 82% is achieved with 11k test patterns,

which is just before additional 50k patterns, while the conventional random pattern test requires nearly 20k patterns to reach the same fault coverage. Moreover, higher fault coverage is achieved with the similar number of test patterns. For example, in case 3 of Fig. 10, fault coverage of 86.9% is achieved with 82k patterns while 82k random patterns reach only 85.2%. In Fig. 9, the fault coverage for domain *T2* is low. It is because this domain needs a lot of inter-domain test to increase the fault coverage. Nevertheless, we can see that the approach improve test pattern quality even for such a low fault coverage case.

These results show the effectiveness of the seed-selection method for DART test, which should be executed under severe test time and memory constraints.

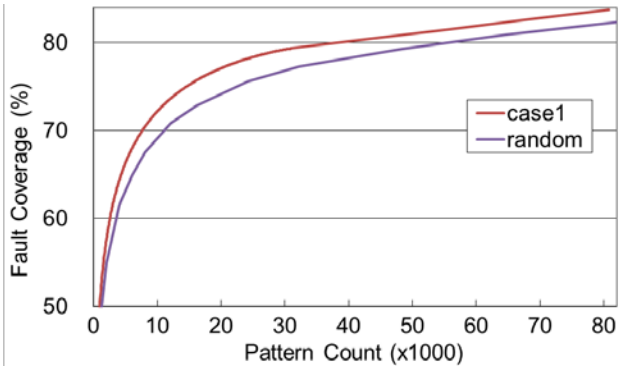

Fig. 8 Fault Coverage Curve for *T0*

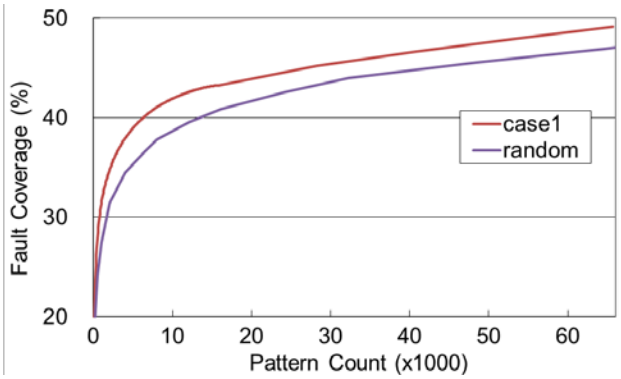

Fig. 9 Fault Coverage Curve for *T2*

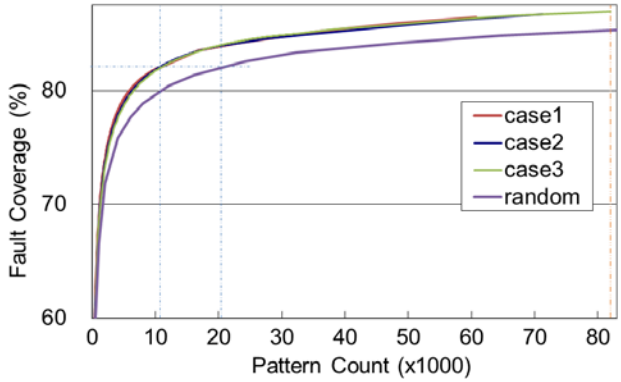

Fig. 10 Fault Coverage Curve for *T7*

5.3 Area Penalty/Investment

The DART controller and other related-circuits were coded in Verilog-HDL. It took 5,042 lines for the LBIST, 5,985 lines for the MBIST, 1,128 lines for control logic, 768 lines for the TVM and totally 12,923 lines for all the circuits (comments included). It took nearly 4 months by 2 people from design to debugging. Regarding the TVM circuit, an 180nm technology TEG has been designed and fabricated in parallel to confirm its functionality.

Table 8 shows the area penalty/investment of DART. Nearly 14k gates are used and it is only 0.2% of the original DUT circuit. The DART circuit operates with the dedicated slow clock of 25MHz and there is not any tight timing requirement. An exception is the TVM, in which ROs operate at nearly 300MHz. However, the counter values of the ROs are collected after the ROs have stopped and waited for a long time (in μ*sec* order). Therefore, there is no tight timing problem.

5.4 Measuring Error

Table 9 and 10 show the estimation accuracy of temperature and voltage using the TVM. It is evaluated using SPICE simulation [28]. Firstly, temperature and voltage are roughly estimated to assign the proper range of estimation. Then, precise estimation is performed in the assigned temperature and/or voltage ranges. As described in Section 3.3, the error caused by process variation is reduced by taking difference approach and these ranges also include process variation margins. Moreover, the TVM has a NBTI tolerant special structure. However, the tolerance to process variation should be studied more.

The tables show that temperature with the error of -3.3°C to 3.0°C and voltage with the error of -15mV to 8mV are able to be estimated using the TVM. Even, a sophisticated analog sensor [41] has an error of 3°C for 3σ. Then, the result seems to be in a proper level.

Table 11 shows the path delay variation due to temperature or voltage variations. Three types of paths were simulated using SPICE. Path *A* consists of 107 NAND gates with no fan-outs. Path *B* consists of 33 OR-NAND gates with 4 fan-outs for each gate. Path *C* consists of 55 NAND gates with 7 fan-outs for each gate. It shows 0.12%/°C maximum delay increase and 0.13%/mV maximum delay decrease. From these, TVM error (M_{TV}) will result in 2.31% (=0.12%/°C×3°C+0.13%/mV×15mV). As only the delay difference from the first measurement is evaluated (Section 3.3), M_{TV} for a 0.3*ns* increase of 3*ns* path (e.g. a 300MHz path with 10% delay margin) will be 7*ps* (0.3*ns*×2.31%). The total measuring error (M_{err}) will be 27*ps* as follows.

$$M_{err} = M_{tim} + M_{TV} + M_{rand} \quad (5)$$

where, M_{tim} is an error of the test timing generator (Section 3.1), which will be 20*ps* (a half of a unit timing), and M_{rand} is a random error (e.g. noise due to hazard, etc.), which

will be reduced to a negligible level evaluating the moving average of successive measured values (Section 2.2).

Table 8 Area Penalty/Investment of DART

Items	# of Cells	Comb. Gates	Non-comb. Gates	Total
LBIST	1,102	1,999	187	2,186
MBIST	1,063	1,607	218	1,825
Control	491	157	2,601	2,758
TVM	2,538	6,771	393	7,164
Total	5,194	10,534	3,399	13,933

Table 9 Temperature Estimation with TVM

			Temperature Range (°C) (Defined by 1st Rough Estimation)		
			-40 ~ 20	20 ~ 80	80 ~ 110
Voltage Range (V) (Defined by 1st Rough Estimation)	1.00 – 1.10	Max.	*2.4°C*	*2.4°C*	*2.0°C*
		Min.	*- 2.6°C*	*- 2.6°C*	*- 1.3°C*
		Std. Dev.	*1.2°C*	*1.1°C*	*0.8°C*
	1.10 – 1.20	Max.	*3.0°C*	*2.6°C*	*2.3°C*
		Min.	*- 3.3°C*	*- 2.7°C*	*- 2.1°C*
		Std. Dev.	*1.4°C*	*1.3°C*	*1.1°C*
	1.20 – 1.30	Max.	*1.9°C*	*1.9°C*	*1.9°C*
		Min.	*- 2.8°C*	*- 1.8°C*	*- 1.3°C*
		Std. Dev.	*1.2°C*	*0.9°C*	*0.7°C*

Table 10 Voltage Estimation with TVM

Estimation Error	Temperature Range (°C) (Defined by 1st Rough Estimation)		
	-40 ~ 20	20 ~ 80	80 ~ 110
Max.	8 mV	8 mV	5 mV
Min.	- 13 mV	- 15 mV	- 8 mV
Std. Dev.	5 mV	5 mV	3 mV

Table 11 Delay Variation due to T/V Variations

Path	Delay(Typ.) (*ns*)	Temperature Increase (*ps* / °C)		Voltage Increase (*ps* / mV)	
A	2.05	2.46	0.12%	-2.68	0.13%
B	2.19	2.02	0.09%	-2.68	0.12%
C	3.42	3.54	0.10%	-3.62	0.11%

5.5 Threshold Delay

Let us define the threshold delay D_{th} as 32*ps* (=M_{err}+5*ps*) lower than the minimum allowable delay *3.3ns* (D_{crit}). Here, 5*ps* is a margin. Then, the measured delay that looks greater than D_{th} has the possibility of 3.26*ns* (=D_{th} - M_{TV}) ~ *3.3ns* (=D_{crit}) because of M_{err}. This means that the detected possible fault might be 5 ~ 39*ps* (M_{TV}+M_{err}+5*ps*) shorter than D_{crit}.

Fig. 11 shows an example of delay increase caused by NBTI, which is modeled with the equation (1) in Section 2.1. The original delay is 3*ns* and 5% delay increase in 10 years is assumed. The early increase is very rapid and it becomes very slow as time passes. From this figure, it is known that 39*ps* delay increase takes a year to several years.

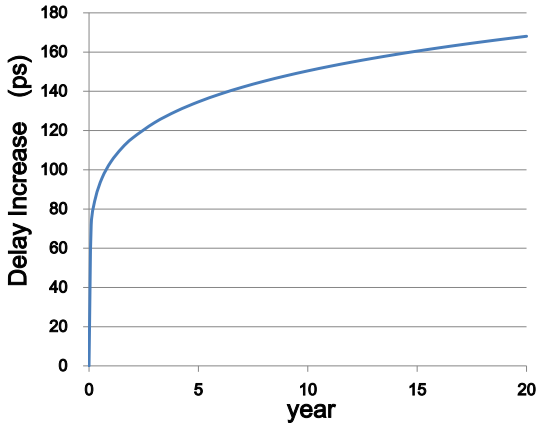

Fig. 11 Delay Increase of NBTI Aging

5.6 Improvements

To improve the effectiveness of DART, following themes are investigated by the authors. These technologies should be more enhanced and implemented for future applications.

As discussed above, the TVM greatly improves delay measurement error. The test timing generator is designed with standard cells and some amount of unit timing errors. However, it is known that more sophisticated design may improve the accuracy of the test timing generator to a few picoseconds [34].

The authors have also developed the test pattern optimization technology that reduces spatial and temporal temperature variation during testing [29, 30]. It is currently available to scan test and is enhancing for BIST. It will enable more stable measurement regarding temperature variation with the combination of the TVM approach.

To detect the dangerous delay increase efficiently, an adaptive rotating test methodology is proposed [32], which schedules test order dynamically and tests more often the critical paths than others.

Delay increase caused by aging-induced faults such as TDDB does not necessarily occur on the longest path. Even a delay increase on a short path may cause catastrophic failure. Therefore, measuring delay increase on short paths will also be needed in future. The authors propose a faster-than-at-speed testing technology [27], which enables to detect small delays on short paths.

For the EM/SM-induced sudden failures, the rotating test approach has possibility not to be able to detect such failures in time. There is a trade-off between the test coverage and the test interval. The proper pattern partitioning method to improve reliability is also investigated [23, 24].

6. Conclusions

Many electronic safety-related systems require very high reliability and variety of approaches are needed to achieve it. The authors have focused on the technology to measure system's delay margin repeatedly which will decrease gradually in a long term. The system will be informed the significant reduction of the delay margin in advance and be able to repair it in an appropriate time. In this paper, DART (dependable architecture with reliability testing) technology was implemented in an actual industrial design.

Applying the 6-time rotating test and the seed-selection method, test coverage of more than 80% was achieved with 8kB memory and 200*ms* time resources. The novel monitor measures temperature with the error of -3.3°C to 3.0°C and voltage with the error of -15mV to 8mV, which will reduce delay measurement error in various environments in field. Production test mode, user mode and DART mode are properly determined controlling the LBIST and the MBIST through the TAP. The area penalty/investment is only 0.2% of the original DUT. Consequently, the effectiveness and availability of DART were confirmed.

Acknowledgements

Many other people helped with our project, and the authors would particularly like to thank Prof. Fujiwara of Osaka Gakuin University, Dr. Uchiyama, Dr. Kanekawa and Mr. Yamamoto of Hitachi Ltd. for their invaluable discussion and support for the implementation. We also thank Dr. Miyase of Kyushu Institute of Technology, Dr. Yamato of Nara Institute of Technology and Prof. Yi of Hanbat National University for their contribution in research development.

References

[1] International Electrotechnical Commission, IEC61508, "Functional safety of electrical / electronic / programmable electronic safety-related systems," Ed.2.0, 2010-4, http://www.iec.ch/functionalsafety/.

[2] ISO26262 "Road vehicles -Functional safety- ," First Edition, 2011-11.

[3] N. Kanekawa, et al., "Dependability in Electronic Systems," Springer, ISBN 978-1-4419-6714-5, 2010.

[4] International Technology Roadmap for Semiconductors, 2011 Edition, http://www.itrs.net/, 2011.

[5] W. Wang, et al., "Compact Modeling and Simulation of Circuit Reliability for 65-nm CMOS Technology," *IEEE Trans. on Device and Material Reliability*, Vol. 7, No. 4, pp. 509-517, 2007.

[6] Y. Li, et al., "Overcoming Early-Life Failure and Aging for Robust Systems," *IEEE Design & Test of Computers*, Vol. 26, No. 6, pp. 28-39, 2009.

[7] T. W. Chen, et al., "Gate-Oxide Early Failure Prediction," *Proc. VLSI Test Symp.*, pp. 111-118, 2008.

[8] V. Reddy, et al., "Impact of Negative Bias Temperature Instability on Product Parametric Drift," *Proc. Int'l Test Conf.*, pp. 148-155, 2004.

[9] S. Bhardwaj, et al., "Predictive Modeling of the NBTI Effect for Reliable Design," *Proc. Custom Integrated Circuits Conf.*, pp. 189-192, 2006.

[10] M. Nicolaidis and Y. Zorian, "On-Line Testing for VLSI-A Compendium of Approaches," *Journal of Electronic Testing: Theory and Applications*, Vol. 12, No. 1-2, pp. 7-20, 1998.

[11] D. Ernst, et al., "Razor: A Low-Power Pipeline Based on Circuit-Level Timing Speculation," *Proc. Int'l Symp. on Microprocessor*, pp. 7-18, 2003.

[12] S. Mitra, et al., "Combinational Logic Soft Error Correction," *Proc. Int'l Test Conf.*, paper 29.2, 2006.

[13] T. Sakata, et al., "A Cost-Effective Dependable Microcontroller Architecture with Instruction-Level Rollback for Soft Error Recovery," *Proc. Int'l Conf. on Dependable Systems and Networks*, pp. 256-265, 2007.

[14] H. Al-Asaad, et al., "Online BIST for Embedded Systems," *IEEE Design & Test of Computers*, Vol. 15 No. 4, pp. 17-24, 1998.

[15] S. Dikic, et al., "BIST and Fault Insertion Re-use in Telecom Systems," *Proc. Int'l Test Conf.*, pp. 1011-1016, 2001.

[16] J. Braden, et al., "Use of BIST in FIRETM Servers," *Proc. Int'l Test Conf.*, pp. 1017-1022, 2001.

[17] J. Qian et al., "Logic BIST Architecture for System-Level Test and Diagnosis," *Proc. Asian Test Symp.*, pp. 21-28, 2009.

[18] Y. Li, S. Makar, and S. Mitra, "CASP: Concurrent Autonomous Chip Self-Test Using Stored Test Patterns," *Proc. Design Automation and Test in Europe*, pp. 885-89, 2008.

[19] H. Inoue, et al., "VAST: Virtualization-Assisted Concurrent Autonomous Self-Test," *Proc. Int'l Test Conf.*, paper 12.3, 2008.

[20] O. Khan and S. Kundu, "A Self-Adaptive System Architecture to Address Transistor Aging," *Proc. Design Automation and Test in Europe*, pp. 81-86, 2009.

[21] Y. Li, et al., "Operating System Scheduling for Efficient Online Self-Test in Robust Systems," *Proc. Int'l Conf. on Computer-Aided Design*, pp. 201-208, 2009.

[22] J. P. Robinson, "Segmented Testing," *IEEE Trans. Comput.*, Vol. C-34, No. 5, pp. 467-471, 1985.

[23] X. Fan, et al., "Genetic Algorithm Based Approach for Segmented Testing," *Int'l Conf. on DSNW*, pp. 85-90, 2011.

[24] S. Wang, et al., "A Pattern Partitioning Algorithm for Field Test," *Proc. 2nd Int'l Workshop on Reliability Aware System Design and Test*, pp. 31-36, 2011.

[25] T. Yoneda, M. Inoue, A. Taketani and H. Fujiwara, "Seed Ordering and Selection for High Quality Delay Test," *Proc. Asian Test Symp.*, pp. 313-318, 2010.

[26] M. Inoue, et al., "Test pattern selection to optimize delay test quality with a limited size of test set," *Proc. European Test Symp.*, pp. 260, 2010.
[27] T. Yoneda, K. Hori, M. Inoue and H. Fujiwara, "Faster-Than At Speed Test for Increased Test Quality and In-Field Reliability," *Proc. Int'l Test Conf.*, paper 2.2, 2011.
[28] Y. Miura, et al., "On-chip Temperature and Voltage measurement for Field Testing," *Proc. European Test Symp.*, to appear, 2012.
[29] T. Yoneda, M. Inoue, Y. Sato and H. Fujiwara, "Thermal-Uniformity Aware X-Filling to Reduce Temperature-Induced Delay Variation for Accurate At-Speed Testing," *Proc. VLSI Test Symp.*, pp. 188-193, 2010.
[30] T. Yoneda, et al., "Temperature-Variation Aware Test Pattern Optimization," *Proc. European Test Symp.*, pp. 214, 2011.
[31] Y. Sato, et al., "A Circuit Failure Prediction Mechanism (DART) for High Field Reliability," *Proc. Int'l Conf. on ASIC*, pp. 581-584, 2009.
[32] H. Yi, et al., "A Failure Prediction Strategy for Transistor Aging," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, to appear, 2012.
[33] K. Shimamura, et al., "A Single-Chip Fail-Safe Microprocessor with Memory Data Comparison Feature," *Proc. Pacific Rim Int'l Symp. on Dependable Computing*, pp. 359-368, 2006.
[34] D. D. Josephson, et al., "Debug Methodology for the McKinley Processor," *Proc. Int'l Test Conf.*, pp. 451-460, 2001.
[35] M. Nakao, et al., "Low Overhead Test Point Insertion for Scan-Based BIST," *Proc. Int'l Test Conf.*, pp. 348-357, 1999.
[36] Y. Sato, et al., "A BIST Approach for Very Deep Sub-Micron (VDSM) Defects," *Proc. Int'l Test Conf.*, pp. 283-291, 2000.
[37] K. Hatayama, et al., "Application of High-Quality Built-in Test to Industrial Designs," *Proc. Int'l Test Conf.*, pp. 1003-1012, 2002.
[38] K. Hatayama, et al., "At-Speed Built-in Test for Logic Circuits with Multiple Clocks," *Proc. Asian Test Symp.*, pp. 292-297 , 2002.
[39] Y. Sato, et al. "Evaluation of the Statistical Quality Model," Proc. Asia and South Pacific Design Automation Conf. (ASP-DAC'05), pp. 305-310, 2005.
[40] Y. Sato, et al., "Invisible delay quality – SDQM model lights up what could not be seen," *Int'l Test Conf.*, paper 47.1, 2005.
[41] S. Remarsu and S. Kundu, "On Process Variation Tolerant Low Cost Thermal Sensor Design in 32nm CMOS Technology," *Proc. Great Lakes Symp. on VLSI*, pp. 487-492, 2009.