

# A Circuit Failure Prediction Mechanism (*DART*) for High Field Reliability

Yasuo SATO\*, Seiji KAJIHARA, Yukiya MIURA, Tomokazu YONEDA,  
Satoshi OHTAKE, Michiko INOUE and Hideo FUJIWARA

**Abstract** — *This paper presents a novel circuit failure prediction mechanism for high field reliability. On-line testing at a power-on/off time of a system detects the circuits' delay degradation that is caused by aging. Dedicated test vectors are applied using BIST architecture. Embedded ring oscillators are utilized to compensate the measured delay values for temperature or voltage shift. The concept and necessary conditions for the mechanism are introduced and some preliminary experimental results show the possible effectiveness of the approach.*

**Index Terms** — **failure prediction, on-line testing, aging, delay degradation, power-on test, power-off test, ring oscillator.**

## I. INTRODUCTION

Ubiquitous VLSIs in the electronic world play important roles for system reliability. Many applications, such as automobile, aircraft, satellite, medical treatment, network equipment, or power plant, require high field reliability. Any malfunction of a chip may cause catastrophic situations, and it should be strictly avoided.

Transistor aging has been known as troublesome phenomenon in deep sub-micron process [1-4]. Channel Hot Carrier (CHC) increases a threshold voltage of an NMOS transistor under a source-drain voltage stress. Negative Bias Temperature Instability (NBTI) increases a threshold voltage of a PMOS transistor under a gate-source voltage stress. Time Dependent Dielectric Breakdown (TDDB) degrades a gate oxide film. Electro Migration has been a serious problem for a long time, and Stress Migration has also paid special attention in Cu-damascene process.

These aging mechanisms have been studied intensively, and their mechanisms are explained in detail [2-3]. The studies suggests that designers need to estimate the amount of degradation over the life time of the production chips (e.g. 10 years), and the chips need to have some margin for their performance (e.g. 10%) [4]. Moreover, designers also have to take process variation margin into consideration. Even if a chip is designed with large amount of margin, it might not be enough because there are many unknown parameters in their

<sup>1</sup> Yasuo SATO and Seiji KAJIHARA are with the Kyusyu Institute of Technology, Kawazu, Iizuka 820-8502, Japan (e-mail: {sato, kajihara}@aries30.cse.kyutech.ac.jp).

Yukiya MIURA is with the Tokyo Metropolitan University, Hino, Tokyo 191-0065, Japan.

Tomokazu YONEDA, Satoshi OHTAKE, Michiko INOUE and Hideo FUJIWARA are with the Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan.

All of the authors are with the Japan Science and Technology Agency, CREST, Chiyoda-ku, Tokyo 102-0075 Japan.

estimation. For instance, it is well known that temperature on the field greatly affects the transistor aging speed. It is also well known that the time length in which a PMOS transistor is on-state significantly affects the aging speed.

Some studies on aging-tolerant circuit design have been started recently. One approach is using a special circuit to detect unstable signal transition at flip-flops [5-7]. An input signal to a flip-flop is bypassed and its delay is shifted small amount, and it is compared to the original signal when a clock is triggered. If the two signals don't match each other, then it is concluded that the input signal is unstable for some reason such as a soft error, noise, or delay caused by degradation. A technical advantage of the approach is that the test can be done during a user operation mode. However, the size of the flip-flop becomes approximately three times larger than the normal one, and this will be unbearable for cost-oriented designs.

Another approach [8-9] is an on-line testing during test mode, which is conducted by an operating system (OS). Li, et al. [8] applied their method for a multi-core processor. They test a core when it is in an idle time. Testing is done based on the conventional transition delay test. However, their method is not applicable to SoCs or ASICs, and it cannot measure the amount of degradation. Khan, et al. [9] proposed a method that is more tightly connected to the OS. Functional test that measures  $F_{max}$  or  $V_{min}$  is applied at checkpoints that OS controls. When an illegal value is measured, system performance is tuned with a frequency or voltage control. This approach is not applicable to SoCs or ASICs, and it can neither measure the amount of degradation.

This paper presents a novel approach that is available for SoC with low test cost. The idea is a kind of on-line testing of a chip during test mode utilizing power-on or power-off time of a system. The scan structure is used as well as a production testing. The delay increase by degradation is measured with high accuracy in order to exclude the influence of the variation of test environment such as temperature or voltage. We try to correct the measured delay using embedded ring oscillators.

The remainder of the paper is constructed as follows. Section 2 introduces the concept of the proposed method. Section 3 shows the proposed DFT structure and testing strategy. Section 4 discusses the importance of thermal controlling technology for accurate delay testing with some preliminary experiments. Section 5 concludes the paper and discusses the future study.

## II. CONCEPT AND REQUIRED FEATURES

This section discusses the concept of the proposed method and some required feature for the mechanism.

### A. Degradation Factors

CHC is accelerated when current flows through an NMOS transistor (i.e.  $V_{ds} > 0$ ). It is unrecoverable. The degrading amount of threshold voltage  $V_{th}$  is estimated as follows [2]:

$$V_{th} \propto T^n \quad (1)$$

where  $T$  is the accelerated time, and  $n$  is a constant that depends on the fabrication process (around 0.45). It should be noted that CHC is accelerated only during current transition time in CMOS technology. NBTI is different from CHC. Fig. 1 illustrates the mechanism of NBTI. It is accelerated when the transistor is on (i.e.  $V_{gs} < 0$ ,  $V_{gb} < 0$ ). It is recoverable when  $V_{gs}$  is greater than zero. Therefore, the degrading amount of threshold voltage  $V_{th}$  is estimated as follows [2]:

$$V_{th} \propto (\alpha T)^n \quad (2)$$

where  $T$  is the accelerated time,  $\alpha$  is the time ratio that the PMOS transistor is ON, and  $n$  is a constant that depends on the fabrication process (around 0.16/0.25). It should be noted that NBTI is accelerated during the gate voltage is biased even if it is stable.

Fig. 2 illustrates the purpose of burn-in test for NBTI [4]. By the temperature and voltage stress at burn-in test, aging is accelerated. From the above expressions (1) and (2), we can find that degradation is very fast at the early stage. After some time, the degradation becomes very slow. Designers should estimate this increase of delay during the life time.

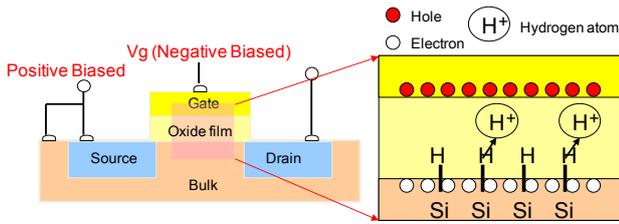


Fig. 1. Mechanism of NBTI

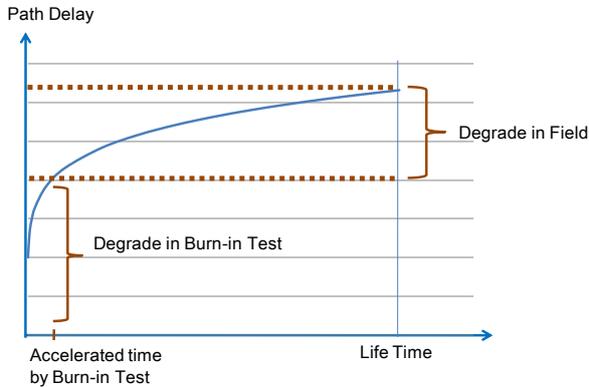


Fig. 2. Burn-in test and life time degrade

### B. Required Features of On-line Testing

The purpose of the proposed mechanism is to measure the performance of the chip repeatedly, and to report it to the system before the degradation becomes catastrophic. Fig. 3

shows this concept. We have set the following features for our study.

#### (1) Short test time

Allowed testing time during power-on time is usually very short compared to production tests. Although it depends on application systems, testing in a few dozen of milliseconds is required for some systems.

#### (2) Low circuit overhead

Any special circuits for user logic are not welcomed. Therefore, we decided to utilize the scan structure that exists as a resource for production test.

#### (3) Accurate measurement

As discussed in Section 1, the design margin for degradation would be no more than several hundred of picoseconds. This is not a large number because delay value varies depending on temperature or voltage very much. In the proposed method, the measured values are corrected by taking test environment such as temperature or voltage into consideration. The correction would be made using embedded ring oscillators [10-11]. The reason why we chose ring oscillators instead of special temperature sensors or voltage sensors is that temperature varies on chip greatly and we have to put sensors at many places.

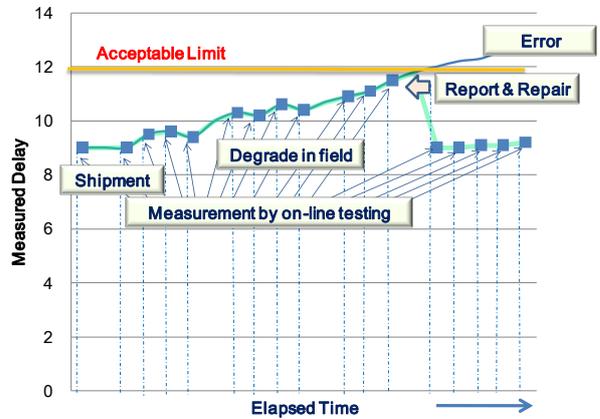


Fig. 3. The concept of measuring degradation

### C. Indexes of Improvement (DART)

We have set the following four indexes to evaluate the effect of our approach. We call these four indexes "DART", and try to quantify them as follows:

#### 1. $D$ (Degrade Factor)

The number of degrading factors should be targeted as many as possible.

#### 2. $A$ (Accuracy)

Delay values excluding test environmental factors of temperature and supply voltage should be accurate as much as possible.

#### 3. $R$ (Report & Repair)

Enough information for warning or repairing should be reported to the system.

#### 4. $T$ (Test Coverage)

An adequate fault model should be assumed and high test coverage should be achieved.

### III. DFT STRUCTURE AND TESTING STRATEGY

#### A. Proposed DFT Structure

Fig. 4 shows the proposed DFT structure. It consists of four parts: “Main Control IP”, “Test Control IP”, “Test Clock Generator IP”, and “Delay Monitor IP”. “Main Control IP” controls the whole DFT structure and analyzes the measured values, which are stored in a non-volatile memory (e.g. flash memory). “Test Control IP” applies test vectors for each core under test. It also controls test timings to measure the fastest functional speed (i.e.  $F_{max}$ ). The measured  $F_{max}$  values are stored in the memory through “Main Control IP”. “Test Clock Generator IP” 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 using the ODCS (On-Die Clock-Shrink) method [12-13] without changing PLL frequency. “Delay Monitor IP” includes embedded ring oscillators (RO), which are customized for DART. One or more IPs may be located in each core under test. The measured delay values of each oscillator are compared to pre-computed characteristics with respect to temperature and voltage. Then, we can estimate the temperature and voltage in the test environment of the system and the measured  $F_{max}$  values are corrected using these information.

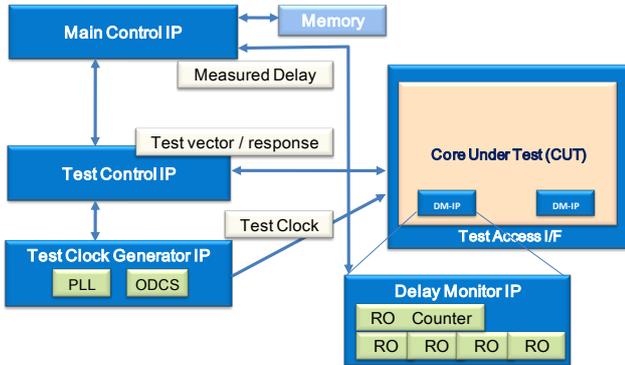


Fig. 4. The Proposed DFT Structure

#### B. Testing Strategy

We showed that the CHC or NBTI degradation occurs very slowly and gradual delay increase will be observed. On the other hand, the delay increase due to Electro/Stress Migration might occur suddenly as shown in Fig. 5. The progress of migration might be very slow, but remained metal will conduct current. Especially, barrier metal [1] in Cu-damascene process will remain until the last stage of migration, and it will conduct current. Considering these aspects, we take the following testing strategy that will make testing be faster and more effective.

1. Dedicated vectors that target CHC or NBTI-prone gates will be applied to measure  $F_{max}$ . These are accurate delay test.
2. Exhaustive transition test vectors, which detect sudden delay increase, will be applied at the system speed or at a speed a little higher than the system speed.

Table 1 shows the distributions of the ratio of stressed (i.e. on-state) PMOS transistors in ISCAS’89 and ITC’99 benchmark circuits. The ratio was calculated using a logic simulator for 10,000 random vectors. From the table, we see that the distributions differ depending on circuits. However, there are some data whose distribution is quite interesting, that is, a large number of PMOS transistors are almost always on (e.g. b15s, b17s, s38417, s38584). They are quite prone to degrade in NBTI mode. This implies that on-line test targeting these transistors would be necessary for NBTI detection.

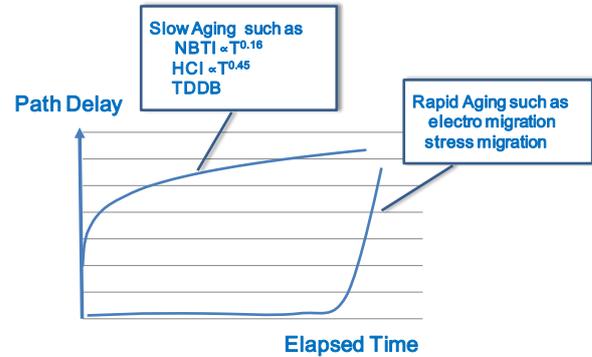


Fig. 5. Two types of delay aging

TABLE I  
THE ON RATIO OF PMOS TRANSISTOR

Data	#PMOS	The Ratio of PMOS transistors' on-state (%)						
		0%	0-20%	20-40%	40-60%	60-80%	80-100%	100%
b15s	18,367	17.8	19.94	3.17	7.25	7.79	25.70	18.36
b17s	51,489	30.65	10.76	2.62	3.71	2.70	14.81	34.74
b20s	19,259	0.99	20.43	19.92	35.57	10.67	10.78	1.64
b21s	20,435	0.89	20.01	22.38	35.37	9.94	10.21	1.20
b22s	31,061	0.67	18.88	21.59	37.06	9.75	10.54	1.52
s35932	33,453	1.13	15.77	3.17	39.23	11.80	27.23	1.67
s38417	35,411	17.47	12.69	4.72	26.44	5.31	15.57	17.79
s38584	39,520	7.96	9.19	10.41	22.07	18.51	17.47	14.39

### IV. THERMAL CONTROL FOR ACCURATE DELAY TEST

Thermal control for accurate delay testing is discussed with some preliminary experiments. Fig. 6 shows a temperature variation of chip when scan test vectors are applied to. An ITC’99 benchmark circuit was used to evaluate the temperature variation of a chip. Thermal simulation was performed for 100 scan vectors. The figure shows an extrapolated result so as to match to the power density ( $2.14\text{W}/\text{mm}^2$ ) in ITRS2007 [1]. The left part of the figure shows temperature variation over a chip, and the right part of the figure shows temperature variation of each block over test time. From this figure, we see the followings.

1. Temperature increases very rapidly in the first several milliseconds.
2. Temperature increases gradually as time passes.
3. Temperature of each block differs very much.

It is well-known that temperature affects circuit delay

significantly. Fig. 7 shows the temperature sensitivity of a ring oscillator (27 stages, 180nm process technology). 300ps delay increase is observed for each ten degrees. This means that accurate temperature control is mandatory for accurate delay measurement. Therefore, we set the following target features for thermal uniformity aware testing.

- Temperature increase should be very fast at the first term. Test would be done after this term.
- Temperature variation should be small in testing time span.
- Temperature variation should be small in space (over a chip).

Fig. 8 shows a preliminary experiment of the testing, which reduces space variation of temperature. A virtual chip of 256 mm<sup>2</sup> with 64 blocks was evaluated after applying 1 million scan cycles. This is only the first trial of the testing, but, it indicates the effectiveness of *DART*.

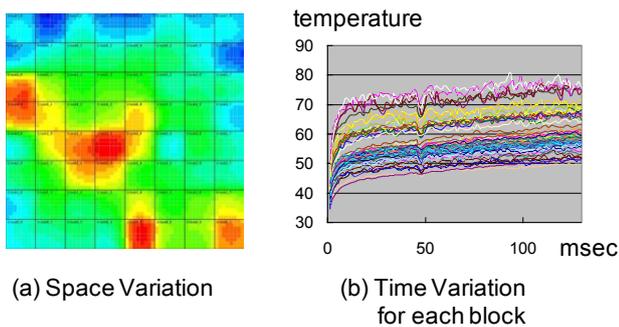


Fig. 6. Temperature variation during test

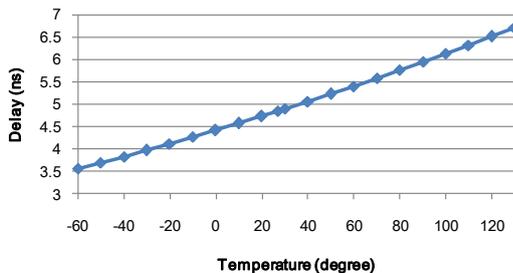


Fig. 7. Temperature sensitivity

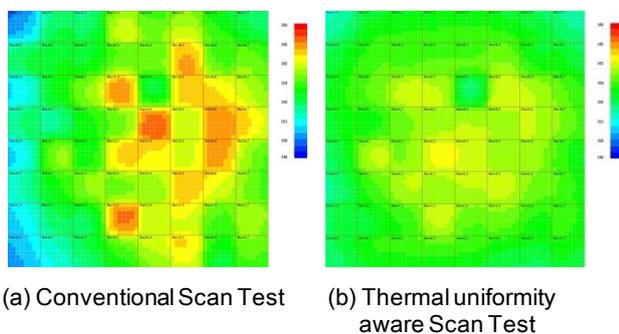


Fig. 8. Thermal uniformity aware testing

## IX. CONCLUSION

The current test technology focuses mainly on reducing power during test. However, as we have shown, it is not enough for field on-line testing. In this paper we introduced a

novel circuit failure prediction mechanism for high field reliability. A new concept of on-line testing includes accurate delay measurement mechanism and accurate delay testing methodology under strict temperature control. Our preliminary results indicated the possible effectiveness of the approach. Although the research has just started, it would lead to a new application of testing technology with new values and the conventional testing concept would be reviewed from this point.

## ACKNOWLEDGMENT

We would like to thank many industrial researchers and engineers who discussed on our study and gave us many suggestive advices. We also thank Prof. Xiaoqing Wen, KIT, for his educational advice, and Mr. Mitsumasa Noda and Mr. Toru Inoue, master students of KIT, for their help on this work.

## REFERENCES

- [1] International Technology Roadmap for Semiconductors, 2007 Edition.
- [2] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan, and Y. Cao, "Compact Modeling and Simulation of Circuit Reliability for 65-nm CMOS technology", IEEE Transaction on Device and Material Reliability, VOL. 7, NO. 4, December 2007.
- [3] T. W. Chen, K. Kim, Y. M. Kim, and S. Mitra, "Gate-Oxide Early Failure Prediction", Proc. IEEE VLSI Test Symp., 4A-1, 2008.
- [4] V. Reddy, J. Carulli, A. Krishnan, W. Bosch, and B. Burgess, "IMPACT OF NEGATIVE BIAS TEMPERATURE INSTABILITY ON PRODUCT PARAMETRIC DRIFT", Proc. Intl. Test Conf., pp. 148-155, 2004.
- [5] T. Nakura, K. Nose, and M. Mizuno, "Fine Grain Redundant Logic Using Defect-Prediction Flip-Flops", IEEE international Solid-State Circuits Conference, pp. 402-403, 2007.
- [6] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, "Circuit Failure Prediction and Its Application to Transistor Aging", Proc. IEEE VLSI Test Symp., pp. 277-284, 2007.
- [7] M. Agarwal, V. Balakrishnan, A. Bhuyan, K. Kim, B. C. Paul, W. Wang, B. Yang, Y. Cao, and S. Mitra, "Optimized Circuit Failure Prediction for Aging: Practically and Promise", Proc. Intl. Test Conf., paper 26.1, 2008.
- [8] 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. 885-890, 2008.
- [9] 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.
- [10] M. Nourani and A. Radhakrishnan, "Testing On-Die Process Variation in Nanometer VLSI", IEEE Design & Test of Computers, pp. 438-451, November-December, 2006.
- [11] Z. Abuhamdeh, V. D'Alessandro, R. Pico, D. Montrone, A. Crouch, and A. Tracy, "Separating Temperature Effects from Ring-Oscillator Readings to Measure True IR-Drop on a Chip", Proc. Intl. Test Conf., pp. 11.2.1-11.2.10, 2007.
- [12] D. D. Josephson, S. Poehlman, and V. Govan, "Debug Methodology for the McKinley Processor", Proc. Intl. Test Conf., pp. 451-460, 2001.
- [13] S. Naffziger, B. Stackhouse, T. Grutkowski, D. Josephson, J. Desai, E. Alon, and M. Horowitz, "The Implementation of a 2-Core, Multi-Threaded Itanium Family Processor", IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 41, NO. 1, JANUARY 2006.