Inconsistent Fails due to Limited Tester Timing Accuracy

Intaik Park[1], Donghwi Lee[1], Erik Chmelar[1,2], and Edward J. McCluskey[1]

[1]*Center for Reliable Computing*
Stanford University
http://crc.stanford.edu

[2]*Advanced Development*
LSI Corporation
http://www.lsi.com

{intaik | dhlee | echmelar | ejm}@crc.stanford.edu

Abstract

Delay testing is a technique to determine if a chip will function correctly at a specified frequency. If a chip passes delay tests, it will presumably function at the specified frequency in the field. This paper presents experimental results that show how chips can pass very thorough delay tests and still fail in the field. It is shown that some chips sometimes pass and sometimes fail when the same delay test is applied multiple times under the same test conditions. These chips are called inconsistent fails. This paper shows how tester timing edge placement accuracy can cause inconsistent fails and suggests the minimum requirements for guardbands that avoid the inconsistent test results.

1. Introduction

Delay testing, such as transition test [Waicukauski 87] and path-delay test [Smith 85], is a widely used technique for testing integrated circuits (ICs). A *delay test* is a two-pattern test; One to launch a transition at a fault site, and another to propagate and capture the response of the circuit at a flip-flop. The time period between the launch clock and the response transition at the capture flip-flop determines what logic value is captured. In principle, the same value should be captured every time a chip is tested in this manner. However, data from both test chips and commercial chips show that some flip-flops sometimes capture expected logic values and sometimes capture erroneous logic values when the chip is tested multiple times with the same delay tests and under the same test conditions (frequency, supply voltage, and temperature). The value captured by one of these flip-flops is called an *inconsistent bit* (also known as a *flakey bit*).

Inconsistent bits in an IC delay test are harmless provided there is at least one flip-flop that consistently captures an erroneous value[†], called a *consistent failing bit*. Given at least one consistent failing bit, the chip will always fail the test. However, if there are no consistent failing bits and there is at least one inconsistent bit, the chip may sometimes pass and sometimes fail the test. This is called an *inconsistent fail*.

Unlike intermittent fails, which appear in weak chips under changes in the operating environments [Hao 93], inconsistent fails are reproducible when tested using Automatic Test Equipment (ATE) and appear in both good chips and bad chips.

The limited tester *timing edge placement accuracy* (EPA) is suspected to be the main source of test inconsistency. Inconsistency due to limited tester timing EPA in memory test and high-speed I/O test were studied in [Dalal 99, Mohanram 03]. This paper discusses how limited tester timing EPA affects scan-based delay test of logic.

A delay test applied without considering inconsistency may not be sufficient to guarantee that a passing chip will function correctly in the field because it may be an inconsistent fail. For example, if an inconsistent failing chip passes delay test at the operating frequency, this chip would be considered to be defect-free and sold to a customer. However, this chip may fail at the same frequency in the field: a test escape.

Although chips are sometimes tested at multiple test conditions such as various frequencies, voltages, and temperatures (as in speed binning or corner testing) [Turakhia 06], an inconsistent fail can still affect an individual test result. For example, in speed binning, chips are tested at multiple frequencies and the fastest passing test determines the operating speed. However, the result of the fastest passing test may be inconsistent,

[†] Or, at least one bit from a group of inconsistent bits fails all the time

in which case the quality of speed binning can be compromised if inconsistent fails are not considered.

This paper presents a cause of inconsistent bits as well as techniques to avoid inconsistent fails. Section 2 describes the test chips used and the experimental setup. Section 3 presents the cause of inconsistent bits. Section 4 describes the effect of inconsistent bits on test results and future trends. Section 5 suggests techniques to avoid inconsistent fails. Finally, Sec. 6 concludes the paper.

2. Test Chips and Experimental Setup

The experiments for this paper utilize both ELF18 and ELF13 chips. The ELF18 chips are test chips manufactured in the Philips 0.18µm Corelib technology, containing 6 DSP cores that are tested individually. ELF13 is a commercial graphics processor manufactured by NVIDIA using 0.13µm technology. The following table summarizes characteristics of ELF18 and ELF13 chips.

Table 1. Characteristics of ELF18 and ELF13

Chip	Technology	Clock domain	Fastest clk freq.	Gate Count
ELF18	0.18µm	1	100MHz	53k
ELF13	0.13µm	10+	370MHz	7.2M

The experiments in this paper apply delay tests to each chip multiple times. To avoid the variations of test conditions such as contact resistance, tests are applied multiple times without removing the chip from the socket. This may cause temperature increase. However, chips behaved the same in terms of inconsistency at both room temperature and high temperature.

3. Inconsistent Bits

In production IC tests, ATEs, or testers, are used to drive signals to the chips being tested. However, these signals are not perfect. They contain uncertainties caused by limited tester timing accuracy. One of the uncertainties that affect the test result is the cycle-to-cycle jitter, which is the variation of the period from one cycle to another [Hodges 04]. When the tester-generated signal drives the test clock that governs delay tests, the test clock signal is inherently under influences of the tester timing jitter, or limited timing EPA. In other words, there would be variations of the test clock periods between launch pulses and capture pulses from cycle to cycle. This variation of the clock period lengths is suspected to produce the inconsistent bits.

When the propagation delay of a path being tested is very close to the time interval between launch and capture edges (*test clock period*), or the path slack is small, the logic value being captured depends on the uncertainties of the clock signal as shown in Fig. 1.

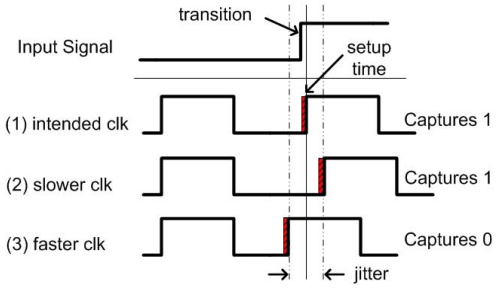

Figure 1. Timing uncertainty and captured value

In Fig. 1, the input signal to a flip-flop transitions from logic-0 to logic-1 and the flip-flop is supposed to capture logic-1, the value after transition. Case (1) is the ideal case when the applied clock period is exactly as intended. The flip-flop captures the value after transition. Case (2) depicts the rising edge of clock arriving late due to jitter, but it does not affect the captured value. However, in case (3), the clock edge is earlier than intended and the flip-flop captures the value before transition, which is logic-0. Hence, due to the variations of test clock periods, the flip-flop sometimes captures logic-0 and sometimes logic-1: an inconsistent bit.

In case (3) of Fig. 1, it is possible that the input signal transitions during the flip-flop setup time or hold time, which is the interval during which an input signal must be stable [Miczo 03]. When this happens, the state of flip-flop can be logic-0, logic-1 or metastable [Kim 90]. In all three cases, the captured logic value is unpredictable, hence, inconsistent.

Experimental results showing the existence of inconsistent bits in delay testing are presented in the following sub-sections.

3.1. Flip-flop States over Different Frequencies

To see how the states of flip-flops (captured bits) vary over different test clock periods, the clock period was gradually decreased, and the same delay test was applied 10 times at each frequency. At each frequency, the flip-flop states were examined.

Figure 2 shows results of an ELF18 core at four different test frequencies. Each square entry represents a scan flip-flop. Only flip-flops that failed within the specified frequency range are marked with letters. Each passing bit at low frequency (denoted as ‘P’) becomes an inconsistent bit as the frequency increases (denoted as ‘I’), and then becomes a consistent failing bit at a higher frequency (denoted as ‘F’). At a frequency where the captured value is inconsistent, the clock period is so close to the propagation delay of the path (small slack) that the pass/fail result is inconsistent due to the tester timing uncertainties.

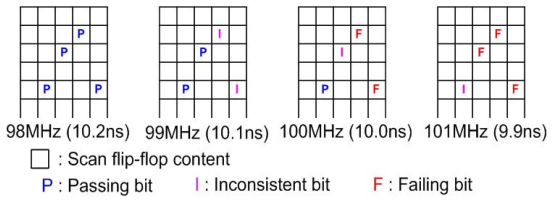

Figure 2. FF contents over different test speeds

3.2. Frequency Sweep

A second experiment was conducted in which delay test sets were applied 10 times at various test clock frequencies and the number of total failing bits was recorded for each test application. If the total number of failing bits varied during 10 test applications at a given frequency, the variance was denoted as the number of inconsistent bits at this frequency. The test clock frequency was stepped by 1MHz and the number of consistent failing bits and inconsistent bits was recorded.

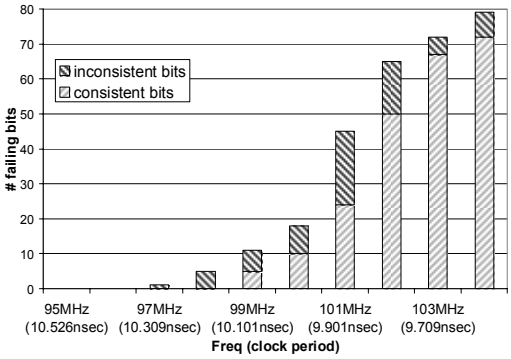

Figure 3. Frequency sweep: number of consistent and inconsistent bits (an ELF18 core)

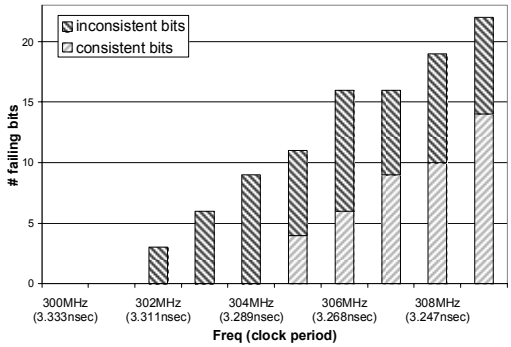

Figure 4. Frequency sweep: number of consistent and inconsistent bits (an ELF13 chip)

Figure 3 shows the experimental result of an ELF18 core and Fig. 4 shows the same for an ELF13 chip (all other cores showed similar results). These cores shown in the figures failed to operate at their specified clock frequencies. But, they operate as designed in lower frequencies. They were chosen to show inconsistent bits near the operational clock frequencies of the designs, at which chips are most likely tested.

In both graphs, as the test clock frequency increases, inconsistent bits start to appear, and the number of consistent failing bits increases. The inconsistent bits exist over a wide range of clock periods. That is because the delay test set tests many different paths with various timing margins. Hence, at a certain test clock period, some paths have enough slack to produce passing bits while other paths have no slack and produce consistent failing bits. There are also paths with very small slack which produce inconsistent bits. Therefore, an inconsistent bit in lower frequency is not necessarily the same inconsistent bit in high frequency. They are from different sensitized paths caught by different capture flip-flops.

3.3. Test Clock Period Accuracy and the Duration of Inconsistency

In scan-based delay test, the test clock period is determined by the distance between the rising edges of two consecutive clock pulses. The location of each edge is determined by a Gaussian distribution (Fig. 5).

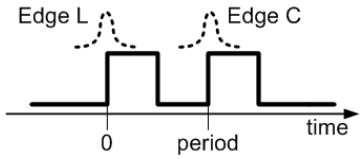

Figure 5. Test clock period and two timing edges

Hence, the uncertainty of the test clock period is derived from the timing edge uncertainties of the tester as described below. (1) and (2) describe the distributions of edge L and edge C locations. (3) is the probability of the size of test clock period being between time A and time B.

$$(1)\ P(\text{edge}\,L@0) = \exp(\frac{-x^2}{2\sigma^2})$$

$$(2)\ P(\text{edgeC@period}) = \exp(\frac{-(y-\text{period})^2}{2\sigma^2})$$

$$(3)\ P(A < \textit{test clock period} < B)$$

$$= \int_{-\infty}^{\infty}\left[\int_{x+A}^{x+B} \exp(\frac{-(y-period)^2}{2\sigma^2})dy\right]\exp(\frac{-x^2}{2\sigma^2})dx$$

The resulting *test clock period accuracy*, which represents the error of test clock period from the specified value, is shown in Fig. 6 along with a single edge timing EPA. In the figure, the graph suggests that the clock period accuracy is nearly ±150ps when the tester timing EPA is ±100ps. However, the measurement of test chips on tester with EPA of ±100ps showed that the uncertainty is smaller than expected.

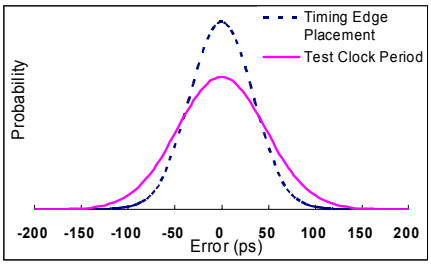

Figure 6. Tester timing edge and test clock period accuracy (EPA: 3 =±100ps)

To find the range of clock periods in which a path produces an inconsistent bit, called *the duration of inconsistency*, sweep experiments with highest resolution (10ps) of the tester were performed. In this experiment, a modified delay test pattern is used. To measure the inconsistency of a single flip-flop, all other scan flip-flops are masked and only one delay test pattern is used so that only one capture flip-flop is active and the path that ends in the flip-flop is tested. This test pattern is applied 10,000 times at each frequency and the number of failing tests was recorded. The test results of an ELF18 core on two different testers are shown in Fig. 7. (EPA: 3σ=±100ps) and Fig. 8. (EPA: 3σ=±170ps).

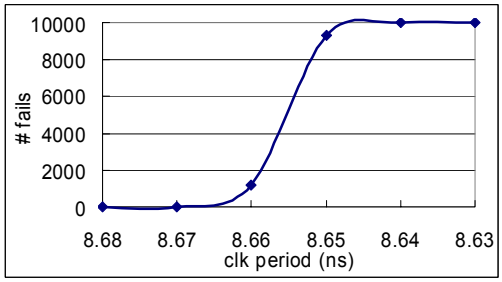

Figure 7. The duration of inconsistency of an ELF18 core (EPA: 3 =±100ps)

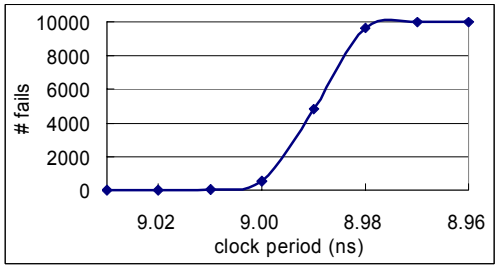

Figure 8. The duration of inconsistency of an ELF18 core (EPA: 3 =±170ps)

The shapes of the graphs in Fig. 7 and 8 are smoothed step functions. It can be explained as followings. The operation of a flip-flop can be modeled as a step function that produces 0 if clock signal arrives after input transition and 1 if clock arrives before transition (signal 1 in Fig. 9). The jitter of clock signal is derived from the uncertainty of tester timing edge placement that is a Gaussian distribution [SEMI 00] (signal 2). Finally, the frequency sweep experiment is a convolution operation of these two signals, which is a smeared step function (signal 3).

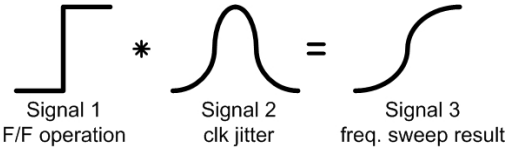

Figure 9. Freq. sweep as convolution operation

Table 2 shows the duration of inconsistency of 9 arbitrarily-picked ELF18 samples measured on two testers with different EPAs (tests are applied 10,000 times at each frequency). It is noticeable that the measured durations of inconsistency were shorter than the calculated clock period uncertainties in both cases.

Table 2. The duration of inconsistency on two testers with different EPAs (ELF18 cores)

EPA	Min	Max	Avg
=±100ps	40 ps	50 ps	43 ps
=±170ps	50 ps	80 ps	66 ps

4. Inconsistent Fails

The existences of inconsistent bits are harmless provided they co-exist with consistent failing bits and the purpose of the test is to reject defective chips (numbers of failing bits do not matter). However, there are certain situations, where inconsistent bits can affect the outcome of delay tests.

4.1. The Cause and Effect of Inconsistent Fails

When the test output contains only inconsistent bits and no consistent failing bits, a chip is sometimes accepted and sometimes rejected by the same delay test set when the test is applied multiple times under the same conditions. These inconsistent fails are observed in Fig. 3 (at 97MHz and 98MHz) and Fig. 4 (at 302MHz, 303MHz and 304MHz) in Sec. 3.

A delay test set may be applied under multiple test conditions in industry. However, if the test is applied only once under each condition, it may not be a sufficient screen for bad chips. An inconsistent failing chip could either pass or fail depending on the value of the inconsistent bit. If the chip happened to pass the test, this would be a test escape because the chip may fail in the field. An example of test escapes due to the inconsistent fail is presented in Table 3.

Table 3. ELF18 on multiple test applications

Test set	1st appl.	2nd appl.	3rd appl.	Final
Delay test1	171	0	+1	172
Delay test2	372	+4	+3	379
Delay test3	377	+8	+3	388
Delay test4	480	+8	+5	493

This table shows the number of cores that failed 4 different delay test sets applied multiple times at the same frequency. Column 2 shows the number of cores

failing delay tests when applied once. Column 3 and 4 shows additional cores failed the 2nd and the 3rd applications of the tests. The last column shows the total number of cores that failed any of 3 applications of delay tests. It is clear that for each test set, one-time application results in test escapes.

Delay test 4 in Table 3 has the highest defect coverage and delay test 1 has the lowest coverage. However, tests with higher defect coverage show more inconsistent fails. This is due the differences of the number and the length of paths being tested by each delay test set. Delay test 4 has the most number of patterns and tests the largest set of paths that contains more paths with small slacks.

Each defective ELF18 chip has been classified according to its defect behavior [Ferhani 06]. Chips that behave the same regardless of test frequency are classified as *timing-independent* (e.g. stuck-at defect) while chips that behave differently at varying test frequencies are classified as *timing-dependent*. All the chips in the latter class showed inconsistent failures at certain frequencies. In other words, all the chips that are suspected to have delay defects exhibited inconsistency at some frequencies.

The inconsistent failing frequencies of these chips were found by frequency sweep experiments as in Sec. 3 or from shmoo plots. In a shmoo plot, the frequency on the border line of the pass and fail region (first failing frequency) was noted and the same delay test was applied multiple times at that frequency. All the timing-dependent defective chips had inconsistent fails at these frequencies.

4.2. The Occurrence of Inconsistent Fails

The occurrences of inconsistent fails may increase as clock periods and slacks decrease given the tester timing accuracy does not improve. Table 4 shows clock periods and average slacks of three different generations of Stanford ELF test chips. It shows that both the clock period and the average slack of each ELF test chip decrease. With this trend, in more recent ELF chips, there would be more paths with small slack and hence small margin of error. Therefore, if the same tester is used to test these chips, ELF13 chips are the most susceptible to inconsistent bits.

Table 4. Clock period and slack of ELF chips

Test chip	Clk period (ns)	Avg. slack (ns)
ELF35	33.60	18.23
ELF18	10.00	4.53
ELF13	2.70	0.29

If a chip is tested by two different testers; one with higher timing accuracy and one with lower timing accuracy, the result from the latter would contain more inconsistent bits. This can be an obstacle of using legacy testers for high speed chips.

The implementation of design also affects the occurrences of inconsistent fails. If the developments of synthesis tools force paths to have similar lengths (more paths become critical paths), there would be more paths with small slacks [Williams 91], which increases the likelihood of inconsistent fails.

Another factor to consider is the test patterns. A significant amount of research has focused on methods to generate delay tests that sensitize long paths to catch fine delay defects [Gupta 04, Qiu 04]. These test patterns sensitize paths with small slacks. Consequently, these tests are more susceptible to inconsistent fails.

5. Avoiding Inconsistent Fails

The simplest way to avoid test escapes due to inconsistent fail is to apply the delay test multiple times and reject chips that fail the test at least once. However, this method may not be practical in production testing where many chips have to be tested and tester time is an expensive resource. This technique can be used in characterization or diagnosis where thoroughness matters more than efficiency.

Another way to prevent test escapes due to inconsistent fail is to apply the delay test at a faster speed. In the test industry, *guardbanding*, or testing beyond the values defined in the device specification, is employed to reduce the possibility of shipping marginal chips or to account for the differences between the test environment and the actual operating condition [Kim 03].

Deciding the size of guardband is a difficult task. If the guardband is too small, the test results can still be inconsistent. On the other hand, an excessively large guardband leads to over-testing that reduces yield.

A typical guardband size is the *Overall Timing Accuracy* (OTA) value of the tester [Dalal 99], which comprises of input, output edge placement accuracy and input to output timing accuracy [SEMI 00]. However, in scan-based delay testing, OTA value is not appropriate for guardband because scan flip-flops triggered by clock edges capture the responses and output timing edge is not involved during launch and capture. Instead, the test clock period accuracy discussed in Sec.3.3 should be considered for guardband because this value represents the duration of inconsistency in scan-based delay test. If a tighter guardband is required to increase yield as much as possible, the frequency sweep with single delay test pattern (Sec. 3.3) should be performed to measure the actual test clock period accuracy of the testing environment, which consists of tester, load board and the chips being tested.

Table 5 shows the number of failing ELF18 cores (the number of inconsistent fails in parenthesis) when

the size of the guardband is zero (0ps, column2), measured clock period accuracy (60ps, column3), calculated clock period accuracy (150ps, column4) and OTA value (200ps, column5). Tests were applied 5 times with each guardband. It is clear that the tighter guardband saves many possible yield losses in ELF18 experiments.

Table 5. Guardband and ELF18 Fails

Guardband	0ps	60ps	150ps	200ps
Delay test 1	168 (4)	175 (1)	203 (6)	222 (8)
Delay test 2	164 (1)	176 (4)	207 (7)	226 (7)

Another way to avoid inconsistency is to test using on-chip clock generators. To test high speed designs, techniques to generate test clocks using on-chip phase-locked loops (PLL) have been developed [Tendolkar 02, Iyengar 06]. Although they require area overhead and complicate the design process and test procedure, these techniques can alleviate inconsistency since PLLs reduce the jitter in on-chip high frequency clock signals derived from tester generated signals. A PLL typically functions as a low-pass filter, which passes signals in lower frequency range (clock signal) and suppresses signals in higher frequency range (jitter) [Li 2007]. Therefore, the limited tester timing accuracy has little effect on clock signals generated by PLLs.

6. Conclusion

In this paper, we show that the limited tester timing accuracy can make delay test results inconsistent, which may cause test escapes. As clock frequency increases while the tester accuracy does not improve, the likelihood of inconsistency tends to increase. To avoid test escapes due to inconsistency, a guardband is employed but it also causes yield loss. Experimental results in this paper suggest that guardband is required to avoid inconsistency but a tighter guardband can be used to save yield. However, to avoid both test escapes from inconsistency and yield loss from guardbanding, testing chips using on-chip PLL clock generator is more appropriate.

Acknowledgments

This work was supported by LSI corp. We thank Philips and NVidia for the chips and for tester time. We thank Verigy for the tester time. We thank Bill Price, Stefan Eichenberger, Fred Bowen, Long Nguyen, and Patrick Debaumarche from Philips; Bruce Cory from Nvidia and Erik Volkerink from Verigy for their support on test setup and the helpful discussions. We thank Subhasish Mitra, Nirmal Saxena, Samy Makar, Siyad Ma, and all the members of CRC for their help.

7. References

[Dalal 99] W. Dalal, S. Miao, "The Value of Tester Accuracy," *Proc. IEEE intl. Test. Conf*., pp. 518-523, 1999.

[Ferhani 06] F.-F. Ferhani and E. J. McCluskey, "Classifying Bad Chips and Ordering Test Sets," *Proc. IEEE Intl. Test. Conf.*, 2006.

[Gupta 04] P. Gupta, M. Hisao, "ALAPTF: A New Transition Fault Model and the ATPG Algorithm," *Proc. IEEE Intl. Test. Conf*., 2004.

[Hao 93] H. Hao, E. J. McCluskey, "Very-Low-Voltage Testing for Weak CMOS Logic ICs," *Proc. IEEE Intl. Test. Conf*., 1993.

[Hodges 04] D. Hodges, H. G. Jackson, R. A. Saleh, "*Analysis and Design of Digital Integrated Cicuits*," 3rd Edition, McGraw-Hiil, New York, 2004.

[Iyengar 06] V. Iyengar, et. al., "At-Speed Structural Test For High-Performance ASICs," *Proc. IEEE Intl. Test. Conf.*, 2006.

[Kim 90] L.-S. Kim and R. W. Dutton, "Metastability of CMOS latch/flip-flop," *IEEE J. Solid-State Circuits*, vol. 25, no. 4, pp.942-951, Aug, 1990.

[Kim 03] K. S. Kim, S. Mitra and P. G. Ryan, "Delay Defect Characteristics and Testing Strategies," *IEEE. Design & Test on Comp.*, vol. 20, pp. 8-16, 2003.

[Li 2007] M. P. Li, "*Jitter, Noise, and Signal Integrity at High-Speed*," First Edition, Prantice Hall, Boston, 2007.

[Miczo 03] A. Miczo, "*Digital Logic Testing and Simulation*," 2nd ed., A John Wiley & Sons, Hoboken, NJ, 2003

[Mohanram 03] K. Mohanram, N. Touba, "Eliminating Non-Determinism During Test of High-Speed Source Synchronous Differential Buses," *Proc. IEEE, VLSI Test Symp*., 2003.

[Qiu 04] W. Qiu, et. al., "K Longest Paths Per Gate (KLPG) Test Generation for Scan-Based Sequential Circuits," *Proc. IEEE Intl. Test Conf*., 2004.

[SEMI 00] Semiconductor Equipment and Materials International, "SEMI G80-0200: *Test Method for the Analysis of Overall Digital Timing Accuracy for Automated Test Equipment*," 2000.

[Smith 85] G. Smith, "Model for Delay Faults Based Upon Paths," *Proc. Intl. Test Conf*., pp. 342-349, 1985.

[Tendolkar 02] N. Tendolkar et. al., "Novel Techniques for Achieving High At-Speed Transition Fault Coverage for Motorola's Microprocessors Based on PowerPC Instruction Set Architecture," *Proc. VLSI test Symp*., pp. 3-8, 2002.

[Turakhia 06] R. P. Turakhia, et. al., "Changing Test and Data Modeling Requirements for Screening Latent Defects as Statistical Outliers," Proc. *IEEE Design & Test of Comp*., vol. 23, pp. 100-109, 2006.

[Waicukauski 87] J. Waicukauski, et. al., "Transition Fault Simultaion," *IEEE Design and Test*, pp. 32-38, April 1987.

[Williams 91] T.W. Williams, B. Underwood, M.R. Mercer, "The Interdependence Between Delay-Optimization of Synthesized Networks and Testing," 28th *IEEE Design Automation Conf*., pp. 87-92,1991