Delay testing is one of key processes in production test to ensure high quality and high reliability for logic circuits. Test escape missing defective chips can be reduced by introducing delay testing. On the other hand, we need to concern yield loss caused by delay testing, i.e., over-testing. Many methods and techniques have been developed to solve problems on delay testing. In this paper, we introduce fundamental techniques of delay testing and survey recent problems and solutions. Especially we focus on techniques to enhance test quality, to avoid over-testing, and to make test design efficient by treating circuits described at register transfer level.

1. Introduction

With miniaturization of VLSI technology, defects that affect timing behavior of logic circuits are increasing\(^1\). For example, a resistive open defect or crosstalk noise may cause a delay fault. On the other hand timing margin of logic circuits is being reduced, and hence the circuits are more sensitive for defects with additional delay. Although the production test intended for stuck-at faults has been performed since before and it plays an important role even now, the traditional stuck-at testing is not effective for the detection of delay faults necessarily. These days, delay testing is essential in production test, and its role is becoming more and more important. Even for subtle additional delay due to not only defects but also process variations, which are referred to as small delay faults, it is easy for the circuit to cause malfunction. To ensure the quality of VLSIs, many researches on detection of small delay faults have been done.

In synchronous sequential circuits, a delay fault appears as a set-up time violation of a flip-flop, i.e., delay testing tries to detect the delay of signal transition at a flip-flop. Hence fault detection needs more than one clock cycle which is for launch and capture of the signal transition. In addition, the test clock must be applied for the circuit at the system clock frequency. Such test methodology is called at-speed testing.

In order to achieve high fault coverage for delay faults, circuits cannot avoid scan-based design. However scan testing makes delay testing complex because delay testing does not allow scan shift operations between launch and capture clocks. In addition, the circuit behavior during scan testing is often much different from that of the normal operation in terms of power consumption, noise, and/or temperature\(^2\)–\(^4\). Since these differences affect the circuit delay, they lead to over-testing\(^5\). Even if a chip failed in a test, the chip may not be defective, i.e., yield loss occurs. In order to avoid yield loss due to over-testing, many methods have been proposed. Research to control the factors that influence delay during testing is one of current hot topics on delay testing.

Research on test generation and DFT techniques at gate level circuits has a long history. However the gate level techniques sometimes face a problem arising from the huge number of elements and the high complexities of the circuit. For delay testing, several techniques for test generation and DFT at higher levels than the gate level such as register-transfer level and behavioral level have recently been proposed.

Thus many researchers have been challenging to provide solutions for the problems on delay testing. In this paper, we survey recent problems and solutions on delay testing for synchronous logic circuits. In Section 2, we explain fault models for delay testing and related issues including metrics to measure test quality of test patterns. In Section 3, we describe test methods to apply test patterns to scan circuits. Furthermore, techniques to improve test quality are given. In Section 4, we describe test methods which are developed to avoid over-testing. In Section 5, we mention test techniques using high level circuit description. In Section 6, we conclude the paper.
2. Fault Models and Test Evaluation

2.1 Basic Fault Models

There are several fault models for test generation or fault grading for delay testing. The fault models are classified based on the view of defective area. Any delay fault model deals with two possible transitions associated with each fault site, rising transition and falling transition.

Transition fault model assumes that a line connecting to a gate has an additional delay whose size is large enough to cause malfunction. A defect such as a resistive open or a resistive bridge results in a transition fault. Recently more detailed analyses based on defects are also reported. In order to detect a transition fault on line \( l \), a test pattern must be applied such that line \( l \) has a transition and the value at \( l \) after the transition can be verified by propagating the value to a flip-flop. The test pattern to verify the value after transition can detect a stuck-at fault on \( l \) too.

Gate delay fault model assumes that a switching delay of a logic gate has an additional delay, i.e., a signal transition of a fan-in line of a gate is propagated to an output of the gate with additional delay. Segment delay fault model assumes a fault on a subpath of the circuit where the length of the subpath is predetermined. When the subpath length is one, it corresponds to a transition fault. When the subpath length is two, it corresponds to a gate delay fault. Path delay fault model assumes that a path from a flip-flop to a flip-flop has a delay. Since it models localized as well as distributed excessive delays on the path, a test pattern generated for the path delay fault can detect most of other delay faults such as transition faults on the path.

When test patterns are generated for delay testing, either path delay faults or transition faults is typically considered. Transition faults can cover the circuit comprehensively, and ATPG (Automatic Test Pattern Generation) for transition faults can be performed easily by extending ATPG for stuck-at faults. One of the disadvantages of transition fault testing is that the delay size caused by a defect is not considered. Since the detectable delay size depends on generated test patterns, the detection of a small delay fault is not guaranteed. The discussion on small delay faults using the transition fault model is introduced in Section 2.2.

2.2 Transition Fault with Small Delay

The original transition fault model does not take the size of the additional delay of faulty lines into consideration. Therefore the detection of an actual transition fault is pattern-dependent. Even though a transition fault was judged as detectable for a test pattern by fault simulation, it might not be detected because of its too small delay size. Only when the delay size is enough large compared to the slack of paths used for fault detection, the fault is guaranteed to be detected. Since fault coverage for the transition faults is not enough to evaluate test quality of test patterns, Sato, et al. proposed a metrics that deals with small delay faults, which are called the statistical delay quality model (SDQM). Since recent research papers on delay testing uses the SDQM for evaluation of test patterns and it is implemented in commercial EDA tools too, we briefly explain the SDQM in this section.

The SDQM assumes a delay defect distribution based on the defect probability in a fabrication process, and then investigates the sensitized transition paths and calculates their delay. The detectable delay defect sizes are determined from the difference between the test clock timing and the tested path lengths. Therefore, the probability of missing small delay faults is calculated by multiplying the occurrence probability for each defect size. The value calculated as a defect level is called the statistical delay quality level (SDQL).

Although the path delay fault model is also popular with delay testing as well as the transition fault model, it has different features from the transition fault model. It is practically impossible to test all paths in a circuit because the number of paths is possible to increase exponentially to the circuit size (the number of gates). For example, the ISCAS-85 benchmark circuit c6288, which is a 16-bit multiplier, has more than \( 10^{19} \) paths. Hence in test generation for path delay faults, we need to select a subset of paths to be targeted. Beside, when a circuit has a large number of paths, it is known that most of paths are untestable. In Section 2.3, methods of path selection and untestable path identification are introduced.
patterns is 4 ns. Also suppose the system clock timing \( T_{MC} \) is 6 ns and the test clock timing \( T_{TC} \) is 7 ns. For the fault, minimum detectable delay size \( T_{det} \) is 3 ns (7 ns–4 ns). Define the difference between the delay size of longest sensitzable path and \( T_{MC} \) as \( T_{mgn} \). If delay size of a fault is less than \( T_{mgn} \), the fault is untestable for any test pattern. In this case, \( T_{mgn} \) is 1 ns (6 ns–5 ns). So if a delay size is greater than 1 and less than 3, the fault remains undetected. Depending on the delay size (defect size), its fault coverage varies significantly, i.e., the percentages of untestable faults, detected faults and undetected faults vary. Figure 2 shows an example of those fault coverage graph. Let \( T_{det}(f_k) \) and \( T_{mgn}(f_k) \) be \( T_{det} \) and \( T_{mgn} \) for fault \( f_k \). SDQL is calculated as follows:

\[
\sum_{k=1}^{2N} \int_{T_{mgn}(f_k)}^{T_{det}(f_k)} F(s) ds
\]

where \( N \) is the number of circuit lines, i.e., \( 2N \) is the number of assumed transition faults and \( F(s) \) is the probability of small delay defects of size \( s \).

The SDQL corresponds to the area indicating undetected faults when the distribution probability of delay size is uniform. While \( T_{MC} \) and \( T_{mgn} \) are test pattern independent, \( T_{det} \) is test pattern dependent. \( T_{det} \) is calculated from two types of delay values: (1) an arrival time of a transition at a faulty site from flip-flops, (2) arrival times of the faulty value at observable points from the faulty site. The sum of these two delay values is the sensitized path delay, and \( T_{det} \) means the difference between the sensitized path delay and test clock timing \( T_{TC} \).

By being aware of the length of sensitization paths in test generation for transition faults, high quality test patterns can be generated. Such test generation methods are called timing-aware ATPG, which are introduced in Section 3.2.

### 2.3 Path Delay Testing

In testing path delay faults, only a small subset of paths is tested because the number of paths is too large to test all of them. Since paths that are likely to be faulty should be tested, longer paths are usually selected according to a certain criterion. A simple approach of path selection is to select K longest paths in order of the path length such that the length of any selected path is longer than the length of any unselected path. However, the location of the selected paths may not be distributed all over the circuit and may be locally concentrated in a part of the circuit. An alternative approach of K longest path is to select a set of paths that contains at least one of the longest paths through each line (17)–20). These approaches are based on structural information of the circuit. In the DSM era, the structurally longest paths may not be actual longest paths in a manufactured circuit due to process variation and/or noise (1). Statistical or dynamic analysis based approaches for path selection have been proposed too (21)–23). However, it is difficult to know exact delay distribution of the manufactured circuit and its computational complexity is high. In addition, the longest paths may be different for each manufactured circuit. Hence, statistical approaches may be still insufficient.

Tekumalla and Menon proposed a concept of primitive faults (24),25). A primitive fault is a multiple path delay fault that satisfies the following two conditions: (1) The multiple path delay fault is static sensitizable, and (2) No proper subset of the multiple path delay fault is static sensitzable. If all primitive faults are tested, timing behavior of the circuit is guaranteed without testing other paths.
The disadvantage of primitive faults is scalability, i.e., the computation of a set of primitive faults is time-consuming for large circuits. Although Krstic, et al. proposed a method to identify primitive faults and to insert test points, it is still impossible for large circuits to identify all primitive faults.

In order to make up for the incompleteness of path selection, Pomeranz and Reddy proposed a test generation method that selects two subsets of paths. For paths in the primary set consisting of the longest paths, test patterns are guaranteed to be generated. For paths in the secondary set consisting of next-longest paths, fault detection is not guaranteed, but it is considered so as to maximize accidental detection by the test patterns for paths in the primary set.

During path selection, we need to be aware of the existence of untestable paths because it is known that there are many untestable paths in a circuit. If untestable faults are included in the target faults, the fault coverage would be so low that additional paths need to be selected until a sufficient number of selected paths are testable. This can be a time-consuming process because of the need to perform test generation for untestable paths. Therefore it is desirable that untestable paths are excluded from a fault list as much as possible.

Methods for identifying untestable paths were described earlier in Refs. 29–31. The first cost-effective method whose run time is independent of the number of paths is the one from Ref. 32. This method finds pairs of lines such that any path including both lines of a pair is untestable. The effectiveness of this method was further established by similar approaches. While these methods of untestable path identification are based on implication, Padmanaban and Tragoudas proposed a method based on decision diagrams.

3. Test Methods to Improve Test Quality

3.1 Test Application Methods for Scan Circuits

Delay testing requires application of at least two successive patterns at at-speed. Even for delay testing, scan designs are required to apply test patterns and to observe test responses because test patterns with high fault coverage can hardly be achieved without scan designs. There are three major methods of at-speed scan testing: LoC (Launch-on-Capture, or broad-side) method, LoS (Launch-on-Shift, or skewed-load) method, and enhanced flip-flop method. Both LoC and LoS are available for scan design with typical scan flip-flops.

The LoC sets a test pattern for initialization by the scan shift operation, and the launch clock for signal transition is applied using the system clock for the functional operation. Then logic values after signal transition are captured using the system clock near the functional operation between scan-in and scan-out operations. The LoC is widely used as a method of at-speed testing because of ease of implementation, i.e., it is available for standard scan design.

The LoS sets a test pattern for initialization by scan shift operation as well as the LoC, but the launch clock for signal transition is applied using the test clock for the scan shift operation too. Then logic values after signal transition are observed using the capture clock, and the captured values are observed at primary outputs through the scan shift operation. The capture clock is applied once, but the clock interval between the last scan shift clock and the capture clock must be short. Therefore if LoS is assumed, physical design of the circuit becomes more complex to meet the timing constraints.

The use of enhanced scan flip-flops allows the circuit to be tested at-speed. The enhanced scan flip-flop takes a special architecture including three latches, while the normal scan flip-flop with a multiplexer includes two latches. The value of the extra latch changes in the scan shift operation, but the output of the original flip-flop keeps its values during the scan shift operation. Since the circuit state does not change in the scan shift operation, a test pattern is applicable between the launch clock and the capture clock. Therefore, in test generation, a test pattern for the capture of the signal transition can be generated independently of a test pattern for the launch of the signal transition. Testability of the circuit is increased dramatically but area overhead due to the extra latches and the risk of over-testing are increased too.

3.2 Timing-aware Test

As mentioned in Section 2.2, detection of a transition fault is pattern-dependent because of undefined delay size. If a fault is detected through a long path, the detectable delay size becomes small. Recently timing-aware test generation methods to detect small delay faults are proposed that try to generate test
patterns so as to maximize delay test quality with the SDQM. Timing-aware test generation takes timing information into account to test a target fault through the longest path whose length is determined from the sum of the fault activation path length and the fault propagation path length.

The disadvantages of timing-aware test generation are the increase of test pattern count and runtime of ATPG. For the reduction of the test pattern count, Yilmaz, et al. and Inoue, et al. proposed methods of test pattern selection from a generated test set. Usually generated test patterns include some test patterns that do not contribute to the improvement of delay test quality. Since only a subset of the generated test patterns can achieve high delay test quality, it allows the reduction of the number of test patterns with a very slight loss of delay test quality of the given test set.

Target fault selection is also an effective way for the reduction of not only pattern count but also ATPG runtime. Lin, et al. proposed a method that apply timing-aware ATPG only for timing critical transition faults with minimal static slack. If a transition fault is not timing critical, normal transition ATPG is applied for the fault. Thus reducing the number of target faults for timing-aware ATPG allows to save test pattern count and the runtime of ATPG.

3.3 Faster-than-at-speed Testing

A defect resulting in a small delay fault may not cause an error for any input pattern if it is not on a long sensitizable path. However, the fault accelerates aging the circuit and causes a reliability problem after shipping. Therefore detection of the small delay fault is meaningful although it is undetectable at the system clock frequency. Faster-than-at-speed test is known as an effective technique of small delay fault detection using the faster test clock cycle than the system clock cycle.

An advantage of faster-than-at-speed test is that test generation does not have to be aware of timing during ATPG. Because the test clock cycle can be varied, a small delay fault can be detected even through a short, non-critical path. On the other hand, the faster-than-at-speed test has some problems to be solved on hazard, noise of IR-drop, and test clock application.

A flip-flop value just before the arrival of the system clock may not be stable due to a hazard because the faster test clock cycle may violates the set-up time of flip-flops. If a hazard appears around the system clock arrival, it can lead to both yield loss due to false identification of good chips and test escape due to fault masking. In addition, the faster clock speed increases the issue of IR-drop during testing as described in Section 4.2. It results in yield loss. Therefore hazard-free or hazard-aware test generation methods for faster-than-at-speed test have been proposed in Refs. 46–52).

Test clock application is another issue for Faster-than-at-speed test. When the fast test clock is applied from an external ATE (Automatic Test Equipment), there are concerns for accurate test clock frequencies by parasitic capacitance and resistance, and skew of the ATE. And the impact on the test cost is large because such an ATE is very expensive. Therefore on-chip clock generators/controllers have been used. McLaurin, et al. proposed a clock control circuit using on-chip PLL. Tayada, et al. and Pei, et al. proposed on-chip programmable clock generators using a logic circuit. These methods are available for faster-than-at-speed test.

4. Techniques to Avoid Over-testing

During testing VLSIs, unreachable states (i.e., illegal state) or state transitions which never appear in the system operation are used. Such states and state transitions may result in test-induced malfunction because power, noise such as IR-drop and crosstalk, and temperature can affect circuit delay. In order to avoid yield loss, test methods or test pattern generation have to take care of test-induced malfunction. In this Section, we introduce techniques to avoid over-testing.

4.1 Test Relaxation for Post-ATPG

In general, test patterns include don’t care bits (X bits) to which assignments of logic values do not affect fault coverage. Although ATPG can output test patterns with X bits, such ATPG increases both the test pattern count and it is reported that even when a highly compacted test pattern set which does not include X bits is generated initially, more than 50% of input bits are Xs implicitly. Test relaxation is a technique to identify X bits in test patterns without sacrificing fault coverage, and is useful for post-ATPG methods to avoid over-testing. This section explains test relaxation techniques.
While ATPG generates test patterns for modeled faults, not all bits of the test patterns are necessary to detect the faults. Kajihara and Miyase proposed a method to identify X bits in a generated test set for full scan circuits. The procedure of the method consists of two phases. At the first phase, the method decides faults which should be detected by each test pattern. At the second phase, the method specifies logic values of each test pattern that are needed to detect the faults. After the second phase, unspecified values of test patterns are identified as X bits. The first phase is based on fault simulation. Although the second phase employs the implication procedure used in ATPG algorithms, there is no need to search. Therefore the complexity of the proposed method is lower than that of ATPG.

The X bits can be filled appropriately for the purpose of controlling power and temperature described in Sections 4.2 and 4.3, as well as test compaction or test compression. However, these uses are better served if the Xs can be placed in desired/specific bit positions of the test patterns. Miyase, et al. proposed a method for maximally fixing X bits on specific bits of each test pattern.

From other aspects of test relaxation, Lin proposed a method of physically-aware N-detect test relaxation. Higami, et al. and El-Maleh, et al. proposed test relaxation methods for non-scan sequential circuits, respectively. This is important for delay test to take the thermal impacts into consideration.

### 4.2 Power-aware Test Method

At-speed scan testing often causes higher switching activity of a circuit than the normal operation. Excessive switching activity occurs during scan shift operation and launch of signal transition of delay testing, and causes excessive power dissipation and IR-drop resulting in increased delay. Many test generation methods to reduce power and/or noise by the first capture clock of the LoC have been proposed in the last decade. A signal transition produced by the first capture clock can have an additional propagation delay caused by excessive power and/or IR-drop. Even when a flip-flop captures an incorrect value by the second capture clock of the LoC, the circuit may not be defective. This situation is over-testing.

Most methods of power-aware test generation and noise-aware test generation are based on post-ATPG process such as X-filling that fills X bits in test patterns. The advantages of X-filling are that there are no impact on design, fault coverage, test pattern count, and ATPG tools. Although there are methods that employ a commercial ATPG tool, they increase test pattern count and runtime.

Recently there are works on noise-aware test generation focusing on controlling noise of paths to be affected, or circuit region in which logic gates may share the same power grid simultaneously. Because the longest paths is easily affected by noise, IR-drop around the paths should be avoided.


### 4.3 Thermal-aware Test Method

As one of the environmental parameters, temperature has a large impact on the circuit delay. For example, every 10°C rise in temperature causes approximately 5% delay increase. Similarly, some defects are sensitive to a certain temperature level. For example, metal interconnect defects may pass a delay test at nominal temperature but fail the same test at a high temperature. Therefore, it is important for delay test to take the thermal impacts into consideration.

In scan-based DFT architectures, each test pattern creates a significant amount of switching activities not only in the scan chains but also in logic cells during scan-shift operation. The continuous excessive switching activity increases overall circuit temperature. Peak temperature can easily exceed 100°C and causes overheating problem. Overheating can lead to problems such as abnormal delay increase and even permanent circuit damage, and result in yield loss. Even if the circuit could escape permanent damage, the exposure to high temperature during a test might reduce its lifetime or reliability. Besides, the scan test creates complex power profile during its application, and causes localized heating, so called hot-spot. Since a hot-spot occurs faster than circuit-wide heating, it can create temperature variation across the circuit. Such spatial temperature variation on a circuit can be as high as 40°C to 50°C, and typical time intervals for temperature changes are on the order of milliseconds. Consequently, it is difficult to screen temperature-sensitive delay defects since the entire circuit cannot be exposed to a particular temperature condition during test application.
tion. Moreover, test-induced spatial temperature variation can cause a clock skew problem and lead to an erroneous pass or fail in delay test \(^8^2\). Therefore, it is important to avoid overheating and temperature variation for accurate delay test.

Several approaches have been proposed to tackle the above-mentioned thermal-related problems \(^8^4\)–\(^8^8\), and the approaches can be classified into two well-known categories: (1) test pattern ordering \(^8^4\),\(^8^5\),\(^8^8\) that manipulates the switching activity by modifying the order in which testers apply test patterns to the CUT and (2) X-filling \(^8^6\),\(^8^7\) that manipulates the switching activity by assigning binary values to don’t-care bits. The basic principle of the thermal-aware approaches is to manage the power consumption by manipulating test patterns. However, unlike the power-aware ATPG approaches, they focus on the scan shift power since it is dominant part in scan test application, and utilize the layout information to take the localized heating events into consideration.

Cho, et al. proposed an algorithm for pattern ordering to minimize the peak temperature during scan testing \(^8^4\). They used the weighted transition count \(^8^9\) as a measure of the power consumption in each circuit block for each test pattern. They first predict a hot-spot (circuit block) in the CUT by taking neighbor circuit blocks into account, and then minimize the peak temperature of the hot-spot. The basic strategies of the pattern ordering are (1) to order high power consuming patterns earlier to maximize heat dissipation with larger thermal gradient and (2) to reduce the total power consumption to minimize heat generation.

Bahukudumbi, et al. also proposed a test pattern ordering framework for wafer-level test-during-burn-in where the junction temperature needs to be maintained at a constant value \(^8^5\). They assumed that, even though each block in the circuit has different activity, latch-up occurs in some part of the circuit and the local temperature increase rapidly spreads over the entire circuit. They used the cycle-by-cycle total number of scan-cell transitions as a measure of the power consumption during test for the CUT. The objective of the pattern ordering is to minimize the statistical temporal variance in test power consumption for the CUT. In their later work \(^8^6\), they integrated an X-filling technique into the pattern ordering framework to further minimize the variation in power consumption during test application.

Yoneda, et al. proposed a thermal-uniformity-aware X-filling technique to minimize the spatial temperature variation for high quality and accurate delay testing \(^8^7\). They also proposed a power estimation method that can accurately predict the total power (including dynamic and leakage) consumed in each circuit block for the response-pattern pair which are simultaneously shifted during test application. The objective of the X-filling is to minimize the statistical spatial variance in test power consumption for each response-pattern pair among the circuit blocks. In their later work \(^8^8\), they proposed a test pattern ordering technique that minimizes the temporal temperature variation while preserving the spatial temperature variation achieved by the thermal-uniformity-aware X-filling technique.

### 4.4 Pseudo Functional Testing

Over-testing is caused by the use of unreachable (or illegal) states or state transitions in scan testing. If a state scanned-in is a reachable state at the system operation, over-testing does not happen. Pseudo functional testing, which is a solution to minimize the risk of over-testing, is a scan test methodology \(^8^0\),\(^8^1\). Before ATPG, unreachable states are extracted and then test patterns are generated with constraints such that the extracted unreachable states are not included. As any state delivered by scan-in operation is close to a reachable state, state transition produced by the system clock such as LoC is more like functional. As a result, the test-induced over-testing is avoided.

Research on pseudo functional testing focuses on unreachable state identification. The more unreachable state is identified, the less the risk of over-testing. Lin, et al. employed a SAT-solver to extract the functional constraints \(^8^0\),\(^8^1\). While the SAT-solver allowed us to find almost all unreachable states, the computational complexity is high. Therefore there remains a scalability problem. The method proposed by Wu, et al. uses a mining technique to extract functional constraints for unreachable states \(^8^2\). They first extract sequential relations as the candidate of functional constraints by analysis with random pattern simulation for a multiple time-frames combinational circuit. Then they verify the actual unreachability of the candidates using a SAT-solver. Syal, et al. proposed implication-based unreachable state identification method \(^8^3\), and Yuan, et al. proposed a justification-based method \(^8^4\). Yuan, et al. then discussed on
compression-aware pseudo functional testing where X bits are used to express unreachable state and to compress test patterns with the Xs.

Pomeranz considers test generation for scan circuits using only reachable states\(^{96}\). While there is no risk of over-testing unlike unreachable state identification, the method has a scalability problem.

5. High Level ATPG and DFT

At register-transfer level (RTL), circuit descriptions can be classified into the following two categories: \textit{structural RTL description} which describes the interconnection of RTL primitives of the circuit and \textit{functional RTL description} which describes the circuit in a cycle-accurate algorithm kind of fashion. In structural RTL description, a circuit generally consists of two separate parts: a \textit{controller} and a \textit{data path}. The former is represented by a state transition graph (STG) and the latter by hardware elements (e.g., registers, multiplexers, and operational modules) and lines.

Several DFT techniques at RTL have been proposed. The techniques try to utilize available circuitry for test instead of using scan FFs. These are classified into scan-based methods and non-scan based ones.

5.1 Scan Based Methods

The most scan-based methods at RTL are designed for testing static faults or single pattern application per scan. For structural RTL circuits, Bhattacharya, et al. proposed \textit{H-SCAN} \(^{97}\). H-Scan uses functional paths, which only pass through multiplexers, between registers as scan paths. Since a functional path in a data path usually have bit width, parallel scan paths are naturally embedded in the data path. Norwood, et al. also developed a method called \textit{orthogonal scan} \(^{98}\). Orthogonal scan shares functional logic with test logic. The method is applied during allocation and binding of hardware resources in high level synthesis and orthogonal scan paths are embedded in the resultant structural RTL circuit.

Aktouf, et al. proposed a method of scan insertion for functional RTL circuits \(^{99}\). In the method, sequential elements are stitched together in a random order in a functional RTL description so that scan paths are embedded in the circuit as a function of the circuit. Huang, et al. provided an effective approach for RTL scan by arranging registers in scan chains through cost rules, which ensure the lowest possible area overhead for the circuit \(^{100}\).

For testing delay faults in a circuit with a DFT structure introduced above, the LoC test strategy can be applied, i.e., any pattern can be set to the FFs as the first pattern of a two pattern test and the second pattern can be prepared as the response of the first pattern.

Obien, et al. introduced functional scan called \textit{F-scan} and its corresponding test generation method \(^{101}\). F-scan utilizes available functional paths and logic as much as possible. Unlike the above described methods which allow application of any single pattern, F-scan only guarantee to propagate test patterns and its response in the range of values used in the original function. This has ability to reduce overtesting but the range is needed to be taken into consideration during test generation. The authors provided constrained ATPG method for generating test patterns for F-scan circuits.

To improve fault coverage of delay faults, a method for supporting LoS test strategy has been proposed by Fai Ko, et al. \(^{102}\) for structural RTL circuits. In LoS test, fault coverage depends on the order of FFs in a scan chain. The method analyzes functional dependency of registers using RTL structural information and orders the registers in functional scan paths to avoid making logical dependency between two consecutive registers on scan paths.

F-scan also provide LoC and LoS test strategies for delay fault testing \(^{103,104}\). As mentioned above, since an F-scan path may not propagate any test pattern, the second pattern depends not only on the order of registers on the F-scan paths but also on functions between registers. Therefore the authors proposed a new test generation method utilizing a commercial ATPG for supporting both LoS and LoC test strategy for F-scan circuits in Ref. 104).

5.2 Non-scan Based Methods

Non-scan DFT methods at RTL have also been proposed. A concept called \textit{hierarchical test generation} \(^{105}\) is usually used for the non-scan based methods. The hierarchical test generation consists of the following two steps: (1) generating test patterns for each individual module in the circuit at gate level and (2) generating justification sequences for applying patterns to respective modules and propagation sequences for the response from the respective modules.

Altuf-Ul-Amin, et al. proposed \textit{hierarchical two pattern testability (HTPT)} of
A data path with HTPT is guaranteed to have justifiability of any consecutive two pattern test to each module and observability of any response from the module. A DFT method to make a given data path hierarchically two pattern testable has also been proposed in the literature. The method achieves complete fault efficiency for every module by paying a hardware overhead. The authors extended their HTPT method for controller-data path circuits in Ref. 107). Yoshikawa, et al. proposed single port change two pattern testability to reduce the hardware overhead of the HTPT method at the sacrifice of not guaranteeing testability of functionally sensitizable path delay faults 108). Since the authors only focused on testing robust and non-robust path delay faults, two consecutive patterns are necessary to be justified at least one input port of a module but the other input ports are not at a time, i.e., for each module, application of single input change two pattern tests is guaranteed.

5.3 Methods for Over-testing Reduction
A DFT method makes untestable faults in an original circuit testable in the resultant circuit to ease test generation. Testing such faults causes over-testing as described in the previous section. A series of identification of untestable delay faults prior to test generation and generation of tests which do not test identified untestable faults is a solution to alleviate over-testing. False path identification using high level design information is an efficient strategy for alleviating over-testing of path delay faults. Yoshikawa, et al. and Ohtake, et al. proposed methods of false path identification utilizing RTL and high level synthesis information, respectively 109).110). In the literatures, a concept of false path at RTL is introduced and conditions for the RTL false path are shown. The proposed methods are designed based on the conditions and it is shown that these methods can efficiently identify RTL false paths. To propagate RTL false path information to gate level through logic synthesis, Ohtake, et al. also proposed a synthesis-based method of path mapping from RTL to gate level 111). In the method, a minimal number of RTL signals are selected to preserve the falseness of paths during logic synthesis. This method makes false path information derived at RTL available at gate level for alleviation of delay test-induced yield loss.

6. Conclusion
In this paper we introduced problems and solutions on delay testing for logic circuits. After explaining delay fault models and fault model-dependent test issues, we mentioned recent topics on delay testing. As methods to achieve high test quality, we introduced at-speed scan test, timing-aware test and faster-than-at-speed test. Then, we introduced methods to avoid over-testing that results in test-induced yield loss. Also, we introduced delay testing methods that allow delay testing to be efficient using register transfer level information. Because the importance of delay testing still has been increasing, techniques on delay testing hereafter will continue to grow.

References
11) Houarche, N., Comte, M., Renovell, M., Czutro, A., Engelke, P., Polian, I. and Becker, B.: An Electrical Model for the Fault Simulation of Small Delay Faults


(Received March 9, 2011)
(Released August 10, 2011)

(Invited by Editor-in-Chief: Hidetoshi Onodera)
Tomokazu Yoneda received his B.E. degree in information systems engineering from Osaka University, Osaka, Japan, in 1998, and M.E. and Ph.D. degrees in information science from Nara Institute of Science and Technology, Nara, Japan, in 2001 and 2002, respectively. Presently he is an Assistant Professor in the Graduate School of Information Science, Nara Institute of Science and Technology. His research interests include VLSI CAD, design for testability, and SoC test. He is a senior member of IEEE (Computer Society) and a member of IEICE.