A Physical-Location-Aware X-filling Method for **IR-Drop Reduction in At-Speed Scan Test**

Wen-Wen Hsieh, I-Sheng Lin and TingTing Hwang Department of Computer Science, National Tsing Hua University HsinChu, Taiwan 300 {wwhsieh, islin, tingting}@cs.nthu.edu.tw

Abstract—IR-drop problem during test mode exacerbates delay defects and results in false failures. In this paper, we take the Xfilling approach to reduce IR-drop effect during at-speed test. The main difference between our approach and the previous X-filling methods [7]-[9] lies in two aspects. The first one is that we take the spatial information into consideration in our approach. The second one is how X-filling is performed. We propose a backward-propagation approach instead of a forwardpropagation approach taken in previous work. The experimental results show that we have 42.81% reduction for the worst IRdrop and 45.71% reduction in the average IR-drop as compared to random fill method.

I. INTRODUCTION

With the progress of fabrication processes and the growing complexity of chip design, delay-fault testing has become more and more important. The main reason comes from the fact that the defect characteristics of processes at 0.13-micron and below are timing-related.

In order to ensure that a circuit meets timing requirements, at-speed scan test is widely used to detect delay defects. For example, [1] reports that the defects per million (DPM) rates are reduced by 30 to 70 percent when at-speed testing is added to the traditional stuck-at tests. Moreover, [2] shows that the escape rate goes up nearly 3 percent if at-speed scan tests are removed from the test program with 0.18-micron feature size.

However, at-speed test suffers the test-induced yield loss. Because the switching activity of whole circuit during test mode is much higher than that during normal mode, large portion of gates simultaneously switching contribute to serious IR-drop delay. Thus, propagation delay does not meet the timing constraint only at test mode. This IR-drop problem during test mode exacerbates delay defects and results in false failures. The analysis from [3] states that IR-drop effect increases up to 16 percent during at-speed test as compared to normal mode. In addition, both [4] and [5] address IR-drop issue and emphasize the importance of avoiding false delay test failures caused by IR-drop.

Recently many researchers have worked on generating test patterns to reduce IR-drop effect during at-speed testing so as to avoid false failures of delay test. The preferred fill proposed in [6] focuses on reducing the Hamming distance between flipflops outputs. Wen et al. propose a series of work [7]-[9] by using X-filling technique to reduce the IR-drop effect. In [7], only flip-flops transition activity is considered while in [8], both flip-flops and gates transition activity are studied. In [9], the authors focus on the critical paths related activities so that the delay malfunction is reduced, and *ifill* proposed in [10] is to to reduce both shift power and capture power during 978-3-9810801-5-5/DATE09 © 2009 EDAA

at-speed testing by improving the X-filling technique. Besides X-filling technique, in [11], a framework for test pattern generation is proposed where process variation information, powergrid topology and regional constraints on switching activity are taken into consideration to generate power-safe scan test patterns. The main idea of the above-mentioned approaches except [11] utilize switching probability of gates to predict the behavior of circuit and assign values to the unspecified bits accordingly in a test cube so that the fully-specified test vectors is expected to reduce IR-drop when applied in test mode. However, an assigned value may not propagate through a long path to target gate. Therefore, the assignment may result in no effect on reducing IR-drop. In addition, except [11], the location information such as power-grid topology and the adjacency of gates switching at the same time are not taken into consideration. Although the local switching activities are considered in [11], the extra test patterns have to be inserted and therefore increase test time.

The method of filling unspecified bits in test cube has the advantages of requiring only minimal changes of the existing ATPGS, and not affecting the test data volume and the test time. Therefore, in this paper, we take the X-filling approach to reduce the IR-drop effect during at-speed test. The main difference between our approach and the previous X-filling methods lies in two aspects. The first one is that we take the spatial information into consideration because the dynamic IR-drop occurs while a large number of gates switch simultaneously in a small region. Therefore, physical location should be taken into consideration to efficiently reduce the dynamic IR-drop. The second one is how X-filling is performed. In previous work [7]-[9], a forward-propagation approach was taken. Scan flip-flop is assigned with a value first. Then, the value is propagated to a target gate. In many cases, side-inputs along a path will block the propagation and thus target gate will not be assigned an expected value. Instead, we propose a backward-propagation approach. For a target gate, a target value is assigned to reduce IR-drop in the target region. Then the value is backward-propagated to scan flip-flop by solving *Pseudo Boolean* (PB) constraints [12]. Our method ensures that a target gate is assigned its expected value.

The rest of this paper is organized as follows. Section II describes the at-speed testing model. Section III presents our main algorithm to reduce IR-drop by taking physical location of gates into consideration and utilizing PB constraints to achieve our goal . Section IV shows the experimental results. Section V concludes this paper.

II. AT-SPEED TESTING MODEL

Before describing the motivation of our work, at-speed testing model is introduced first. In our work, we adopt the *launch-off-capture* scheme for at-speed scan-based delay test. Figure 1 shows an example waveform.



Fig. 1. The waveform of launch-off-capture scheme.

Since the false delay path occurs between the launch pulse and the capture pulse, our goal is to minimize the impact of IR-drop effect during this test cycle. In other words, our goal is to minimize the switching activity of gates between the lastshift cycle (before launch cycle) and the launch cycle.

III. THE PROPOSED METHOD

A. The Power Grid Architecture

The power/ground distribution network proposed in [3] is adopted. Figure 2 shows the power grid architecture. The power rings are created to carry power around the core chip area. Four VDD and VSS pads are inserted to the respective rings. The stripes directly connected to the power/ground ring and the rails draw power/ground from the nearest power stripe. By using stripes and rails, the power and ground is routed to the standard cells.



Fig. 2. The power/ground distribution network.

B. Overview of Our Proposed Method

Most of previous work [7]–[9] was performed at logic level. In order to take physical locations into consideration, our algorithm is performed after physical synthesis. Figure 3 illustrates the overall algorithm flow. First, a given circuit netlist is placed and routed. After placement and routing, physical information of the circuit is collected to be utilized in the following steps. Next, *region_partition* is performed. Since the power grid topology is taken into consideration, the chip is first divided into several regions according to the physical layout. Each cell in the design is then assigned to a region by using its physical coordinate. Based on the characteristic of the power grid architecture described in Section IV.A, each region is bounded by stripes and limited number of rails. If a large number of cells switch simultaneously within the same region, the IR-drop effect in this region is regarded as very serious and may increase the propagation delay. Our main idea is to minimize switch activities of regions. In order to achieve this goal, appropriate values (0 or 1) are assigned to unspecified bits within each region.

The next step is *target_region_selection*. With a set of given critical paths and test cubes with unspecified bits, the switching activity of each region is first calculated. The region with the highest switching activities is selected as target region for assigning values to unspecified bits. Then, in the *target region*, target bit is selected and its value is set one by one. This step is called value_set_up_for_target_bit. In the selection of target bit, sub-circuit in target region is modeled as a Psuedo Boolean constraint problem (PB-problem). The solution to the PB-problem is used to assist us to select *target bit* and its value. After each selection, the value of target bit is propagated to primary input and primary output. The selection continues till the switching activity is smaller than a threshold value. Once a target region is processed completely, the switch activities of all regions are re-calculated. The procedure repeats till no more region to be processed. If there are unspecified bits left in the test vector after all steps mentioned-above are processed, the greedy fill is applied. The greedy fill step is similar to the value_set_up_for_target_bit expect that the target bit is selected in the whole circuit.



Fig. 3. The overall flow of the proposed method.

The details of *target_region_selection* and the *value_set_up_for_target_bit* steps are described in the following sections.

C. Target Region Selection

In this subsection, we propose a cost function to guide the selection of a region with serious IR-drop effect. We first apply test vectors to the circuit to observe the switching activities of each gate during at-speed test cycle. Our cost function, weighted switching activities (WSA), is defined to represent the IR-drop impact of each region caused by the test vectors. Give a region, region j, its $WSA_{regionj}$ is defined as

$$WSA_{region j} = \sum_{gate i}^{vgate i \in region j} ((1 + \alpha \cdot Cri_i) \times$$

switching weight(type_i) ×
$$\sum_{k \in fanout \ of \ i} capacitance_k)$$
 (1)

where Cri_i represents if *gate i* is on critical path and defined as

$$Cri_i = \begin{cases} 1, & \text{if } gate \ i \text{ is on critical paths} \\ 0, & \text{otherwise.} \end{cases}$$

while α is the weight used to emphasize the importance of gate i if gate i is on the critical path. Since gates on critical paths switching simultaneously would cause serious IR-drop delay and result in delay defects, the region with more gates on critical paths will has higher IR-drop impact. In our experiment, α is set to 0.5 to balance the importance of IR-drop and criticality. $Type_i$ in equation(1) is the type of toggles of gate i. Here, the toggle type includes all kinds of transitions, i.e. $1 \rightarrow 0, 0 \rightarrow 1, X \rightarrow 0, X \rightarrow 1, 1 \rightarrow X, 0 \rightarrow X \text{ and } X \rightarrow X$. For each type of toggle, we give its switching weight, and switching weight($type_i$) represents the Switching weight of $type_i$, where type 1 has the highest weight and type 4 has zero weight. Switching weight represents the preference and flexibility to assign a specific transition. The higher the value, the less the preference (flexibility) is. The last term of WSAis to calculate the fanout capacitance of gate i.

WSA is defined to estimate the IR-drop of each region because it represents the supply current demand and the power consumption. Based on the cost function, the region with the highest WSA is selected as *target region* and is the next region to be processed.

D. Value Set Up for Target Bit

After the *target region* is selected, the *X-filling* technique will be applied to reduce IR-drop in this *target region*. Since values assigned in the region are outputs of internal gates, it is required to justify this assignment in the primary input. In addition, the correlations of internal gates between the last shift cycle and the launch cycle has to be satisfied. Therefore, we model this assignment problem as a *Pseudo Boolean* (PB) constraint problem and solve it by a PB-solver.

Since a circuit can be represented easily by CNF formulas, in our method, we use PB-constraint to express the correlation between internal gates at two timing frames. In addition, an objective function is defined to minimize IR-drop effect within a *target region*.

Before we describe the modeling, some variables are defined first. Given a sub-circuit with n gates and flip-flops, where $\{GF_i \mid i = 1, 2, ..., n\}$ represents a gate or flip-flop in the sub-circuit. The variables are:

 S_gfi is the value of GF_i at the last shift cycle.

 L_gfi is the value of GF_i at the launch cycle.

 T_GF_i is the toggle of GF_i and defined as

$$T_GF_i = \begin{cases} 1, & \text{if } S_gfi \neq L_gfi \\ 0, & \text{if } S_gfi = L_gfi \end{cases}$$

Since our goal is to reduce the total WSA within the *target* region, the objective function to be minimized is defined as following:

$$objective \ function = \sum_{GF_{i=1}}^{n} \left(\sum_{K \in fanout \ of \ GF_{i}} capacitance_{K} \cdot T_GF_{i} \right)$$
(2)

Next, three constraints are set to model the switching of circuit. They are *output constraint*, *consistency constraint* and *value constraint*.

- Output constraint: This constraint is used to represent the relationship between the switching activity and the value between the last shift cycle and the launch cycle. Therefore, it is modeled as XOR S_gfi and L_gfi .
- *Consistency constraint*: This constraint is used to maintain the function consistency at both last shift cycle and launch cycle.
- *Value constraint*: This constraint is used to indicate the circuit state after the test cube is applied to the circuit in the *target region*.

Finally, these constraints mentioned-above are translated to the corresponding PB-constraint expression. With the objective function, a PB-solver [12], [13] is called to solve the problem.

After we solve the above modeled PB problem, we obtained a set of value assignment of the gates in the *target region*. This assignment is to minimize the object function and, in terms, to minimize IR-drop in the region.

IV. EXPERIMENTAL RESULT

We perform the experiments to demonstrate the effectiveness and efficiency of our method. Figure 4 illustrates our experimental flow.



Fig. 4. The overall flow of the experiment.

The experiments are performed on ITC'99 benchmark [14]. The circuits are described in VHDL and synthesized by Synopsys Design Compiler with TSMC 90nm cell library. Six kinds of gates including *BUFFER*, *INVERTER*, *NAND*, *NOR*, *AND* and *OR* are used for synthesis. The gate-level netlist and the test protocol in STIL format are generated. The input to SoC Encounter (SOCE) is the netlist and the output is a layout design with detailed physical location and the timing information after floorplaning, placement, and routing. Then the netlist and the test protocol are fed into TetraMax to generate test patterns with unspecified bits. Meanwhile, the timing information is used as input file for PrimeTime to produce critical paths information. Our algorithm takes the netlist of layout design, the test pattern with unspecified bits and the critical paths information as input files, and generates IR-dropaware fully specified test patterns by utilizing MINISAT+ [13] as the PB-solver.

To produce precise IR-drop analysis by using RedHawk, VCD (value change dump) format file which records the accurate transition is generated by simulating the fully specified test patterns. Then, RedHawk reports IR-drop information according to the VCD file and layout netlist. Next, taking IRdrop effect into consideration, PrimeTime generates IR-dropaware critical paths and path delay.

Table I shows the information of benchmark circuits and test vectors. The columns labeled #PI, #FF, and #Gate represent the number of primary input, the number of scan-chain flip-flop and the total number of gate count, respectively. The number of vertical stripes and the number of rows used to divide a layout into regions are listed in columns 5 and 6. Column 7 labeled as #Region is the number of regions in each design. As to the test cube information, the column labeled #Vec and Fault *Coverage* (%) represent the number of test vectors and the fault coverage by running these test patterns. The last column labeled as X-bits (%) is the X-bits ratio.

IHE INFORMATION OF CIRCUITS AND TEST CUBE									
	Circuit Information						Test Cube Information		
circuit	#PI	#FF	#Cells	#Strap	#Row	#Region	#Vec	#Fault	X-bits
								coverage.	(%)
b11	11	31	824	3	16	35	161	98.21	62
b12	9	121	935	3	16	41	160	98.19	79
b14	36	215	20328	3	99	202	200	85.78	54
b15	40	417	10289	3	115	147	200	89.17	77
b20	36	430	25089	5	293	228	200	85.72	61
b21	36	430	25619	5	59	280	200	79.71	72

 TABLE I

 The information of circuits and test cube

Table II shows the voltage drop reported by RedHawk. We compare our PB-based X-filling method to random fill, and CPA-based X-filling [9]. The results are labeled as *r-fill*, *CPA-based*, and *PB-based*, respectively. The results of *r-fill* is our baseline result reference, and the columns labeled (%) are the reduction ratio as compared to *r-fill* method. The columns labeled *MAX* and *Avg*. list the maximum voltage drop and the average voltage drop of each benchmark circuit.

From this table, we can see that our *PB-based* method can significantly reduce both maximum and average voltage drop as compared to *CPA-based* method [9]. The improvement ratio of our *PB-based* method can achieve 42.18% for the maximum voltage drop and 44.48% in the average voltage drop while the improvement ratio of *CPA-based* method is 26.45% for the maximum voltage drop and 30.08% in the average voltage drop.

TABLE II COMPARISONS OF MAXIMUM AND AVERAGE VOLTAGE DROP r-fill CPA-based [9] PB-based circuit MAX Avg. MAX Avg MAX Avg (mV) (%) (mV) (%) (mV) (mV)|(mV)| (%) (mV) (%) 10.5 31.82 41.18 67.53 0.9 47.06 b11 15.4 1.7 1 5 b12 17.1 1.8 9 47.37 0.9 50 9 47.37 0.6 66.67 b14 280.5115.5 276.7 1.35 96.2 16.71 273.1 2.64 74.9 35.15 110.9 25.77 17.4 32.82 59.6 60.11 25.9 9.9 b15 149.4 61.78 b20 160.9 30.9 116.1 27.84 25.4 17.8 107.9 32.94 21 32.04 b21 114.6 27.3 86.5 24.52 21.3 21.98 65.9 42.5 20.7 24.18 26.45 30.08 42.18 44.48 Avg.

V. CONCLUSION

In this paper, we have proposed a backward-propagation *X-filling* method taking spatial information into consideration to reduce the IR-drop effect during at-speed scan test. First, the circuit was divided into several regions according to the physical layout. Then, we estimated the IR-drop effect for each region by our WSA cost function. Next, each region is modeled as a PB-problem to reduce its switching activity. The experimental results show that we have 42.18% reduction for the worst voltage drop and 44.48% reduction in the average voltage drop while the reduction of previous work [9] is 26.45% for the worst voltage drop and 30.08% in the average voltage drop.

REFERENCES

- [1] B. Swanson and M. Lange, "At-speed testing made easy," *EE Times*, 3, June 2004.
- [2] J. Gatej, L. Song, C. Pyron, R. Raina, and T. Munns, "Evaluating ATE features in terms of test escape rates and other cost of test culprits," in Proceedings of the International Test Conference, pp. 1040-1048, October 2002.
- [3] N. Ahmed, M. Tehranipoor, V. Jayaram, "A novel framework for fasterthan-at-speed delay test considering IR-drop effects," in *Proceedings of the IEEE/ACM International Conference on Computer Aided Design*, pp.198-203, November 2006.
- [4] J. Saxena, K. M. Butler, V. B. Jayaram, et.al., "A Case Study of IR-drop in Structured At-Speed Testing", in Proceedings of the International Test Conference, pp. 1098-1104, September-Octover 2003.
- [5] S. Ravi, "Power-aware Test: Challenges and Solutions," in Proceedings of the International Test Conference, pp.1-10, October 2007.
- [6] S. Remersaro, X. Lin, Z. Zhang, S. M. Reddy, I. Pomeranz and J. Rajski, "Preferred Fill: A Scalable Method to Reduce Capture Power for Scan Based Designs," in Proceedings of the International Test Conference, pp. 1-10, October 2006.
- [7] X. Wen, Y. Yamashita, S. Morishima, S. Kajihara, L. T. Wang, K. K. Saluja, and K. Kinoshita, "Low-Capture-Power Test Generation for Scan-Based At- Speed Testing," in *Proceedings of the International Test Conference*, pp., November 2005.
- [8] X. Wen, K. Miyase, T. Suzuki, Y. Yamato, S. Kajihara, L.-T. Wang, and K. K. Saluja, "A Highly-Guided X-Filling Method for Effective Low-Capture-Power Scan Test Generation," in Proceedings of the International Conference on Computer Design, pp. 251-258, October 2006.
- [9] X. Wen, K. Miyase, T. Suzuki, S. Kajihara, Y. Ohsumi, and K. K. Saluja, "Critical-Path-Aware X-Filling for Effective IR-Drop Reduction in At-Speed Scan Testing," *in Proceedings of the Design Automation Conference*, pp. 527-532, June 2007.
- [10] J. Li, Q. XU, Y. XU and X. Li, "iFill: an Impact-Oriented X-filling Method for shift-and capture-Power Reduction in At-Speed Scan-Based Testing," in Proceedings of the Design, Automation, and Test in Europe Conference and Exhibition, pp. 1184-1189, March 2008.
- [11] V. R. Devanathan, C. P. Ravikumar and V. Kamakoti, "A Stochastic Pattern Generation and Optimization Framework for Variation-Tolerant, Power-Safe Scan Test", in *Proceedings of the International Test Conference*, pp.1-10, October 2007.
- [12] F. A. Aloul, A. Ramani, I. Markov, K. Sakallah. "Generic ILP versus Specialized 0-1 ILP: an Update", in Proceedings of the International Conference on Computer Design, pp. 450-457, November 2002.
- [13] N. Een and N. Sorensson, "Translating Pseudo-Boolean Constraints into SAT," in Journal on Satisfiability, Boolean Modeling and Computation, Vol. 2, pp. 1-26, March 2006.
- [14] "http://www.cerc.utexas.edu/itc99-benchmarks/bench.html"