

VirtualScan: A New Compressed Scan Technology for Test Cost Reduction

¹Laung-Terng (L.-T.) Wang, ²Xiaoqing Wen*, ³Hiroshi Furukawa, ⁴Fei-Sheng Hsu, ⁴Shyh-Horng Lin, ⁴Sen-Wei Tsai, ¹Khader S. Abdel-Hafez, and ¹Shianling Wu

¹SynTest Technologies, Inc.
505 S. Pastoria Ave., Suite 101
Sunnyvale 94086, U.S.A.

²Department of CSE
Kyushu Institute of Technology
Iizuka 820-8502, Japan

³NEC Micro Systems, Ltd.
2081-24 Tabaru, Mashiki-Machi, Kamimashiki-Gun
Kumamoto 861-2202, Japan

⁴SynTest Technologies, Inc., Taiwan
2F, No. 27, Industry E. Rd. 9
Hsinchu, Taiwan

Abstract

This paper describes the VirtualScan technology for scan test cost reduction. Scan chains in a VirtualScan circuit are split into shorter ones and the gap between external scan ports and internal scan chains are bridged with a broadcaster and a compactor. Test patterns for a VirtualScan circuit are generated directly by one-pass VirtualScan ATPG, in which multi-capture clocking and maximum test compaction are supported. In addition, VirtualScan ATPG avoids unknown-value and aliasing effects algorithmically without adding any additional circuitry. The VirtualScan technology has achieved successful tape-outs of industrial chips and has been proven to be an efficient and easy-to-implement solution for scan test cost reduction.

1. Introduction

Integrated circuit testing based on the full-scan methodology and automatic test pattern generation (ATPG) is the most widely used test strategy that is well supported by test engineers, electronic design automation (EDA) vendors, and tester makers. In a full-scan circuit, all functional storage elements are replaced with scan cells, which are structured into scan chains that are assessable from a tester. As a result, a sequential circuit is reduced to a combinational circuit in test mode. Test patterns for a combinational circuit can be readily generated with an ATPG program and are stored in a tester. A tester applies test patterns to an integrated circuit, collects test responses, and makes pass/fail judgment [1].

Despite the usefulness of scan-based manufacturing test, its applicability is being severely threatened by its rapidly growing cost. The cost of manufacturing scan test consists of many parts, including costs of tester capital investment, handlers, probe-cards, tester utilization, test development,

etc. [2]. The most significant, recurring, and unpredictable one is the tester utilization cost determined by test data volume and test cycles.

Due to system-on-a-chip (SoC) complexity and tens of millions of gates in size, test data volume and test cycles increase dramatically even for single-stuck-at faults with single-detection. With the wide-spread of deep sub-micron (DSM) processes, the need for test patterns for single-stuck-at faults with multi-detection, transition delay faults, path delay faults, bridging faults, cross-talk faults is also growing in order to maintain the quality level of next-generation integrated chips [2]. This need will further increase test data volume and test cycles.

A large volume of test data results in costly tester re-load and a large number of test cycles results in long tester utilization time, both leading to higher test cost. Obviously, it is unsustainable to tackle the test cost problem by keeping buying bigger and faster testers. The ultimate solution is logic built-in self-test (BIST) [1]. However, logic BIST has significantly different characteristics from the current full-scan/ATPG based test flow, in terms of fault coverage, overhead, logic and physical design efforts, as well as fault diagnosis [3]. As a result, for the foreseeable future, full-scan/ATPG based testing will remain as the major test strategy, especially for manufacturing test and for circuits without in-system test requirements. Therefore, finding an efficient method for reducing test data volume and test cycles in a full-scan/ATPG test environment, with making full use of existing testers in mind, is a very important task.

For a full-scan circuit, both test data volume and test cycles are proportional to the number of test patterns (N) and the longest scan chain length (L). Basically, one can try to reduce N , L , or both in order to reduce test data volume and test cycles.

* Part of this work was done while the author was at SynTest Technologies, Inc.

Various methods [4-5] have been proposed for reducing the number of test patterns (N) through compaction. All of them assume a *1-to-1 scan configuration*, in which the number of internal scan chains equals the number of external scan input/output ports. In addition, it has been shown that, for a circuit with multiple clocks, using a multi-capture clocking scheme instead of a one-hot clocking scheme can significantly reduce the number of test patterns [6-7].

Recently, several scan test cost reduction methods [8-23] have been proposed based on the idea of reducing the longest scan chain length (L). These methods assume a *1-to- n scan configuration*, in which the number of internal scan chains is n times the number of external scan input/output ports. Such a scan configuration can be obtained by splitting an original scan chain into n shorter ones, where n is called *split ratio*. Obviously, the longest scan chain length in a 1-to- n scan configuration is $1/n$ of that in a 1-to-1 scan configuration. As a result, test data volume and test cycles in a 1-to- n scan configuration can theoretically be $1/n$ of that in a 1-to-1 scan configuration, although the actual reduction ratio is often less than n because of stronger constraints at the interface between external scan ports and internal scan chains.

Conceptually, reducing the longest scan chain length (L) with a 1-to- n scan configuration is a more efficient approach to scan test cost reduction than reducing the number of test patterns (N). The reason is that a split ratio is user-controllable so that a large split ratio can be selected if a greater scan test cost reduction effect is required. However, the biggest issue in scan test cost reduction with a 1-to- n scan configuration is how to bridge the gap between external scan ports and internal scan chains since the number of internal scan chains is n times the number of external scan input/output ports. Generally, test data to be fed into internal scan chains needs to be compressed in one way or another in order to be applied through external scan input ports. In this sense, the test cost reduction approach based on a 1-to- n scan configuration can be called *compressed scan test*.

This paper describes a new compressed scan technology, called the *VirtualScan technology*, for bridging the gap between external scan ports and internal scan chains. The VirtualScan technology consists of the *VirtualScan architecture* and the *VirtualScan ATPG technique*. Scan chains in the VirtualScan architecture are split into shorter ones and the gap between external scan ports and internal scan chains are bridged with a broadcaster and a compactor. VirtualScan ATPG generates test patterns directly in a one-pass process, in which multi-capture clocking and maximum test compaction are supported. In addition, VirtualScan ATPG avoids unknown-value and aliasing effects algorithmically without any extra circuitry.

The paper is organized as follows: Section 2 describes the research background. Section 3 and Section 4 present the VirtualScan architecture and the VirtualScan ATPG technique, respectively. Section 5 outlines the VirtualScan design flow. Section 6 shows application results and Section 7 concludes the paper.

2. Background

Previous scan test cost reduction methods [8-23] based on the idea of reducing the longest scan chain length (L) can be divided into two categories: input-side solutions and output-side solutions, as described below:

2.1 Previous Input-Side Solutions

There are three major approaches to bridging the gap between external scan ports and internal scan chains on the input side, i.e., providing test data to a large number of internal scan chain inputs through a small number of external scan input ports. The first one is to use a decompression/compression scheme, the second one is to use a deterministic BIST scheme, and the third one is to use a broadcasting scheme.

The decompression/compression scheme [8-15] is based on the fact that a test cube generated by ATPG for a circuit with a 1-to- n scan configuration often contains a significant number of unspecified or don't care bits. It is possible to encode such a test cube with a compressed test vector of a smaller number of bits and later decompress the compressed test vector during test with an on-chip decompressor. The encoding is conducted by solving a set of linear equations, and a decompressor is a sequential circuit, such as a linear feedback shift register (LFSR), a ring generator, etc. In this scheme, final test patterns to be applied from a tester are generated in a two-pass process, i.e., test cubes are generated first and compression is then conducted. This means that ATPG may not conduct dynamic and static compaction at the highest level since a significant number of unspecified bits need to be left in the test cubes. As a result, the number of test cubes may be larger than that of test vectors generated with maximum test compaction. In addition, a decompressor is a sequential circuit, which is generally more costly to design, especially for clock and timing.

The deterministic BIST scheme [16] is based on the fact that most faults in a circuit can be detected with random patterns and that only a small number of faults need test vectors generated deterministically by ATPG. It uses a pseudo-random pattern generator (PRPG) for on-chip test data stream generation. Random-pattern-resistant faults are identified and test cubes are generated for them with ATPG. The test cubes are then compressed as seeds for the PRPG. Periodically, the PRPG is re-seeded to decompress a seed loaded through a shadow register to its

corresponding test cube. This scheme reduces the number of compressed test vectors by making use of pseudo-random patterns. It also uses a sequential circuit and the overhead may be higher because of a shadow register used for re-seeding.

The broadcasting scheme [17-19] uses an external scan input port to drive multiple internal scan chain inputs. This is a straightforward approach to bridging the gap between the number of external scan input ports and the number of internal scan chain inputs. In the previous broadcasting methods, an external scan input port is connected directly to multiple internal scan chain inputs without passing through any logic gates. As a result, the strong correlation among multiple internal scan chains driven by the same external scan input port may make it difficult to achieve high fault coverage in some cases.

2.2 Previous Output-Side Solutions

There are two major approaches to bridging the gap between internal scan chains and external scan ports on the output side, i.e., obtaining test responses from a large number of internal scan chain outputs through a small number of external scan output ports. The first one is to use a multiple-input signature register (MISR) and the second one is to use a compactor, as described below:

If a MISR is used, it is necessary to preserve the uniqueness of a signature by making sure no unknown values (X 's) propagating to the MISR. The propagation of X 's can be blocked with additional circuitry at the cost of significant impacts on design effort, timing, and overhead. Some methods [16, 20] use a mask network or a scan-out selector between internal scan chain outputs and a MISR to mask or avoid X 's without blocking them inside the circuit-under-test. These methods may need to handle issues such as increased overhead and sequential control complexity.

If a simple space compactor, usually composed of XOR gates, is used, it is necessary to deal with fault coverage loss due to unknown values (X -impact) and aliasing. Previous methods for solving this problem include the use of a selective compactor [21], a convolutional compactor [22], or an X -tolerant compactor [23]. These methods may need to handle issues such as increased overhead, sequential control complexity, and the number of X 's that can be tolerated.

2.3 The VirtualScan Technology

In order to solve the problems of the previous methods for bridging the gap between external scan ports and internal scan chains, this paper presents the VirtualScan technology, which consists of a new input-side solution as well as a new output-side solution.

The VirtualScan technology consists of the VirtualScan

architecture and the VirtualScan ATPG technique. The VirtualScan architecture is based on a 1-to- n scan configuration and the gap between external scan ports and internal scan chains are bridged with a broadcaster and a compactor. A broadcaster is a small and simple circuitry that is used to distribute test data from external scan input ports to internal scan chain inputs in minimally constrained manner in order to achieve higher fault coverage. A compactor is simply a set of XOR trees with minimal overhead for merging internal scan chain outputs to feed into external scan output ports. Test patterns for such a VirtualScan circuit are generated directly by VirtualScan ATPG, which is a one-pass process that supports multi-capture clocking and allows maximum test compaction during test pattern generation. In addition, VirtualScan ATPG is aware of the compactor structure and can assign proper values during test pattern generation to algorithmically avoid X -impact and aliasing without adding any additional hardware.

The VirtualScan technology is significantly different from previous solutions in that (1) the broadcaster is a small and simple circuitry, (2) ATPG is a one-pass process instead of a two-pass one in which test cubes must be generated and then compressed, and (3) X -impact and aliasing are avoided algorithmically instead of using any additional circuitry. These characteristics make the VirtualScan technology fit well into any full-scan/ATPG test environment, as an efficient, low-overhead, and easy-to-implement solution for scan test cost reduction.

3. VirtualScan Architecture

3.1 General Structure

The VirtualScan architecture consists of three major parts: a full-scan circuit with a 1-to- n scan configuration, a broadcaster located between external scan input ports and internal scan chain inputs, and a compactor located between internal scan chain outputs and external scan output ports [24].

Fig. 1 shows the general VirtualScan architecture for a split ratio of 4. The full-scan circuit has a 1-to-4 scan configuration. That is, one original scan chain is split into 4 shorter scan chains in a balanced way. The broadcaster is inserted between the external scan input ports (SI_1, \dots, SI_m) and the internal scan chain inputs ($s_{10}, s_{11}, s_{12}, s_{13}, \dots, s_{m0}, s_{m1}, s_{m2}, s_{m3}$). The compactor is inserted between the internal scan chain outputs ($t_{10}, t_{11}, t_{12}, t_{13}, \dots, t_{m0}, t_{m1}, t_{m2}, t_{m3}$) and the external scan output ports (SO_1, \dots, SO_m). The combination of the full-scan circuit, the broadcaster, and the compactor is a *VirtualScan circuit*.

Note that final test patterns, instead of test cubes, are generated with one-pass VirtualScan ATPG directly for the VirtualScan circuit, instead of the full-scan circuit.

This is different from other solutions [8-19], whose compressed test generation is mostly a two-pass process. This characteristic makes the VirtualScan technology fit well into any existing full-scan/ATPG test environment. Since the longest scan chain length is reduced by 4 times, theoretically test data volume and test cycles are also reduced by 4 times. Due to possibly stronger constraints induced by the broadcaster and the compactor, however, the actual reduction ratio may be lower than 4.

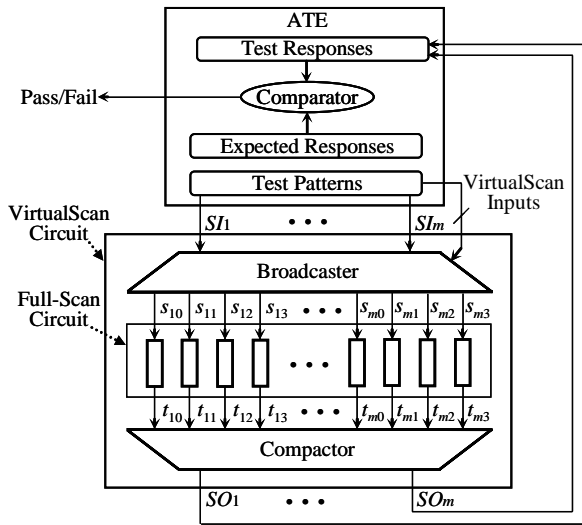


Fig. 1. VirtualScan Architecture

3.2 Broadcaster

A broadcaster is used to distribute test patterns from a small number of external scan input ports to a large number of internal scan chain inputs in a minimally constrained manner.

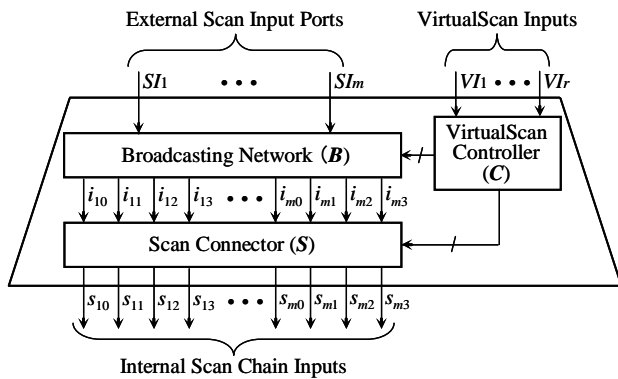


Fig. 2. General Broadcaster

Fig. 2 shows the general structure of a broadcaster for a split ratio of 4, which consists of a broadcasting network (B), a scan connector (S), and a VirtualScan controller (C). The broadcasting network is a combinational block composed of one or more logic gates, such as AND, OR, NAND, NOR, XOR, and XNOR gates as well as buffers

and inverters. It is used to distribute the values at m external scan input ports $\{SI_1, \dots, SI_m\}$ to $4 \cdot m$ internal signals $\{i_{10}, i_{11}, i_{12}, i_{13}, \dots, i_{m0}, i_{m1}, i_{m2}, i_{m3}\}$. The VirtualScan controller can be a combinational block (a random logic network, a decoder, etc.) or a sequential block (a shift register, a finite-state machine, etc.). It is used to provide control values to the broadcasting network for reducing value correlation at the internal signals. The scan connector consists of a number of multiplexers and optionally scan cells. The multiplexers can be controlled by one or more mode selection signals provided from the VirtualScan controller. When a mode selection signal is 1, each corresponding internal signal feeds its corresponding internal scan chain input so that test data reduction can be achieved. When a mode selection signal is 0, the corresponding split internal scan chains in a 1-to-4 scan configuration are connected back to the original scan chains in a 1-to-1 scan configuration so that fault coverage improvement by top-up ATPG, as well as fault diagnosis, can be conducted. Note that multiple mode selection signals can be used to provide different selection values for different multiplexers so that the test data reduction effect can be achieved without fault coverage loss.

Fig. 3 shows an example broadcaster. Here, the broadcasting network consists of only XOR gates. For example, SI_1 is distributed to i_{10} in a direct manner but to $i_{11}, i_{12},$ and i_{13} in a controlled manner. The control values are provided by a shift register in the VirtualScan controller. The shift register can either be loaded once for the whole test process or during each test session, or each time when new test data values are applied. A mode selection signal VI_2 is used to control all multiplexers in the scan connector. Note that the VirtualScan inputs VI_1 and VI_2 can be borrowed from other external scan input ports except SI_1 . This means that there can be no extra pin overhead.

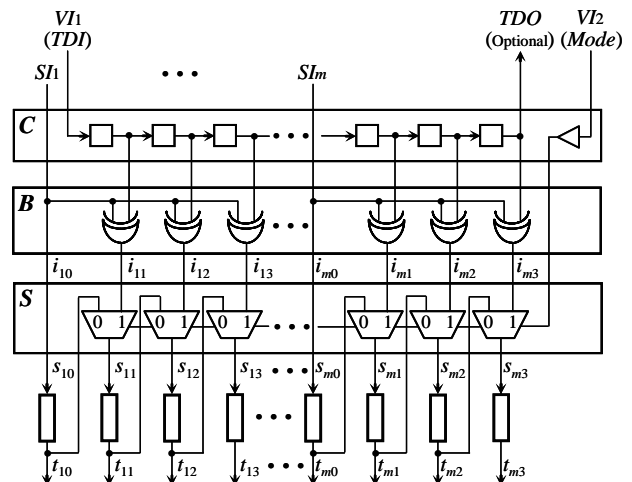


Fig. 3. Example Broadcaster

Generally, a broadcaster can use some external scan input ports to provide test data and uses others as VirtualScan inputs to provide control data in order to reduce value correlation at targeted internal scan chain inputs. The role of an external scan input port as a data source or as a control source can be switched dynamically based on test pattern generation requirements. This is determined automatically by the VirtualScan ATPG. As a result, fault coverage loss due to value correlation can be minimized at the cost of slightly more test patterns. Since the impact of splitting scan chains is much bigger than that of the increased number of test patterns, the VirtualScan technology can still achieve significant reduction in test data volume and test cycles.

Note that a broadcaster is a small and simple logic block. Generally, it is easier to implement than other sequential decompressors [8-16]. Being small and simple also make a broadcaster itself less vulnerable to physical defects. In addition, a broadcaster is local and its design only depends on the split ratio. As a result, a broadcaster can be easily incorporated at the register-transfer level (RTL) or in a hierarchical design.

3.3 Compactor

The VirtualScan architecture uses a compactor for space compaction of test responses from a large number of internal scan chain outputs to a small number of external scan output ports. A compactor is chosen over a MISR because of its simplicity (no clock involved) and low overhead (no *X*-blocking needed).

However, a simple compactor, usually composed of XOR trees, may suffer from two major issues: *X*-impact and aliasing. *X*-impact means that a fault cannot be detected if its effect feeds an XOR gate whose another input has an unknown value (*X*). Aliasing means that a fault cannot be detected if its fault effects appear on both inputs of an XOR gate, canceling each other. Instead of using a complex compactor that usually involves a sequential controller [21, 22], the VirtualScan technology uses an algorithmic technique to solve these issues during ATPG without adding any additional circuitry. This will be described in 4.3.

3.4 Support for Top-Up ATPG and Diagnosis

The VirtualScan architecture also provides a mechanism to switch between the 1-to-*n* scan configuration and the original 1-to-1 scan configuration in a full-scan circuit. This is achieved by using a scan connector composed of a number of multiplexors and one or more additional mode selection signals. As a result, top-up ATPG for the 1-to-1 scan configuration can be conducted after VirtualScan ATPG is done for the 1-to-*n* scan configuration in order to improve final fault coverage if necessary. In addition, switching back to the 1-to-1 scan configuration makes

fault diagnosis easy since an existing fault diagnosis flow can now be used without any modification.

4. VirtualScan ATPG

In the VirtualScan technology, final test patterns are generated directly for a VirtualScan circuit with VirtualScan ATPG. VirtualScan ATPG is unique and efficient because of three distinguishing characters: one-pass process, multi-capture clocking scheme, and algorithmic handling of *X*-impact and aliasing.

4.1 One-Pass Process

VirtualScan ATPG is significantly different from most previous solutions [8-16], in which intermediate test cubes (fault-detection assignments with a considerable number of unspecified bits) are first generated, and are then compressed into test patterns or seeds in order to be stored in a tester. Different from this two-pass approach, the VirtualScan ATPG is a one-pass process, in which final test patterns are generated directly for the entire VirtualScan circuit. As a result, maximum test compaction can be conducted dynamically and statically since there is no need to preserve a significant number of unspecified bits. Therefore, a smaller set of test patterns can usually be generated in comparison with test cubes.

4.2 Multi-Capture Clocking Scheme

It is very common that a circuit has multiple clocks, each controlling one clock domain, and that clock-tree design is performed on each individual clock domain. As a result, the clock skew in a clock domain can be minimized to the extent that all flip-flops in the clock domain operate correctly in both functional and test modes. However, the clock skew between two clock domains is usually large and unpredictable since clock trees for the two clock domains are designed separately. Because of this, it is not safe to activate the clocks in inter-related clock domains simultaneously to capture test response.

One widely used solution for this problem is the so-called one-hot clocking scheme. Suppose that a circuit has two clock domains CD1 and CD2, driven by clocks CLK1 and CLK2, respectively, and that CD1 transfers data to CD2. A test pattern is shifted into all flip-flops in both CD1 and CD2, and capture is first conducted for CD2 by only activating CLK2 but keeping CLK1 inactive. The captured test responses are shifted out while the next test pattern is shifted into all flip-flops in both CD1 and CD2. Capture is then conducted for CD1 by only activating CLK1 but keeping CLK2 inactive. Obviously, testing CD1 and CD2 once needs two test patterns. Generally, the one-hot clocking scheme results in a larger number of test patterns although it only needs a simple combinational ATPG program.

VirtualScan ATPG uses a complete multi-capture clocking scheme. Suppose again that a circuit has two clock domains CD1 and CD2, driven by clocks CLK1 and CLK2, respectively, and that CD1 transfers data to CD2. As shown in Fig. 4, a test pattern is shifted into all flip-flops in both CD1 and CD2, and capture is first conducted for CD1 by only activating CLK1 but keeping CLK2 inactive. After a delay larger than the clock skew between CD1 and CD2, capture is then conducted for CD2 by only activating CLK2 but keeping CLK1 inactive. The test responses obtained in two captures are then shifted out together. Note that there is no shift operation between the two capture operations. That is, testing CD1 and CD2 once only needs one test pattern. Different from other multi-capture ATPG solutions, VirtualScan ATPG employs a unique algorithm for handling sequential behaviors related to the multi-capture operation. The advantage is higher fault coverage with a smaller number of test vectors due to less test response information loss. This algorithm will be described in a separate paper.

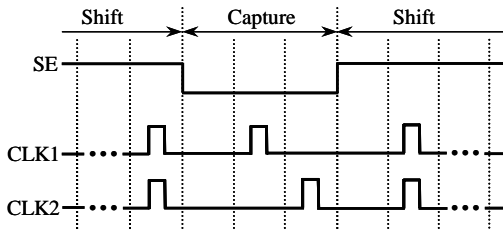


Fig. 4. Multi-Capture Clocking Scheme

4.3 Algorithmic Handling of X-Impact and Aliasing

The ATPG algorithms used in previous test cost reduction solutions [9-19] only target a full-scan circuit without taking into consideration the constraints induced by the circuits added for bridging the gap between external scan ports and internal scan chains. VirtualScan ATPG, on the other hand, is aware of the structures of the broadcaster and the compactor. The broadcaster structure information allows VirtualScan ATPG to generate final test patterns directly as described in 3.2. In addition, the compactor structure information enables VirtualScan ATPG to algorithmically handle X-impact and aliasing without adding any new circuitry.

An example of algorithmically handling X-impact in VirtualScan ATPG is shown in Fig. 5. Here, $SC1, SC2, \dots, SC4$ are scan cells connected to a compactor composed of XOR gates $G7$ and $G8$. a, b, \dots, h are internal signal lines, and f is assumed to be connected to an X-source (memory, non-scanned storage element, etc.). Now consider the detection of the stuck-at-0 fault f_1 . Obviously, logic 1 should be assigned to both d and e in order to activate f_1 . The fault effect will be captured by scan cell $SC3$. If the X on f propagates to $SC4$, the compactor output q will

become X and f_1 cannot be detected. To avoid this, VirtualScan ATPG will try to assign either logic 1 to g or logic 0 to h in order to block the X from reaching $SC4$. If it is impossible to achieve this assignment, VirtualScan ATPG will then try to assign logic 1 to c , logic 0 to b , and logic 0 to a in order to propagate the fault effect to $SC2$. As a result, fault f_1 can be detected. Thus, X-impact is avoided by algorithmic assignment without adding any new circuitry.

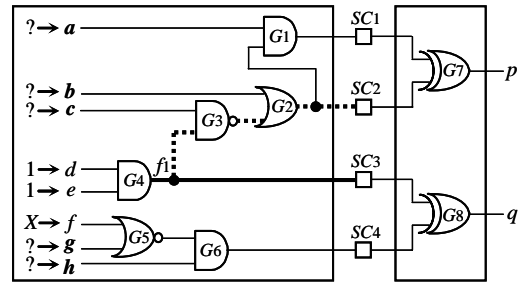


Fig. 5. Handling of X-Impact

An example of algorithmically handling aliasing in VirtualScan ATPG is shown in Fig. 6. Here, $SC1, SC2, \dots, SC4$ are scan cells connected to a compactor composed of XOR gates $G7$ and $G8$. a, b, \dots, h are internal signal lines. Now consider the detection of the stuck-at-1 fault f_2 . Obviously, logic 1 should be assigned to c, d , and e in order to activate f_2 , and logic 0 should be assigned to b in order to propagate the fault effect to $SC2$. If a has logic 1, the fault effect will also propagate to $SC1$. In this case, aliasing will cause the compactor output p to have a fault-free value, resulting in an undetected f_2 . To avoid this, VirtualScan ATPG will try to assign logic 0 to a in order to block the fault effect from reaching $SC1$. As a result, fault f_2 can be detected. This way, aliasing can be avoided by algorithmic assignment without any extra circuitry.

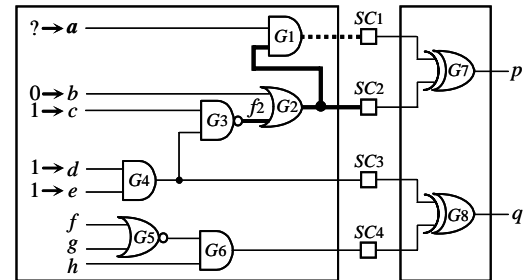


Fig. 6. Handling of Aliasing

5. Design Flow

The VirtualScan technology is both efficient and flexible. Its architecture simply requires a broadcaster and a compactor, which are small and simple. And its ATPG is a one-pass process. In addition, the broadcaster design, as

well as the compactor design, only depends on the split ratio, and has no relation with the structure and size of the circuit-under-test. All these make it easy to apply the VirtualScan technology in a gate-level design flow, a RTL design flow, or a hierarchical design flow.

5.1 Gate-Level VirtualScan Flow

Fig. 7 shows a gate-level VirtualScan design flow. In the gate-level VirtualScan flow, VirtualScan circuit generation is conducted at the gate level. A normal scan netlist, which only has a 1-to-1 scan configuration, is generated by conducting logic/scan synthesis on the functional RTL code. The original scan chains are then split into shorter ones based on a given split ratio, and a broadcaster and a compactor are added to form a VirtualScan circuit, on which layout is conducted and a final netlist is obtained. Then, VirtualScan ATPG is conducted on the final netlist. If fault coverage is not enough, top-up ATPG is conducted using a conventional ATPG engine.

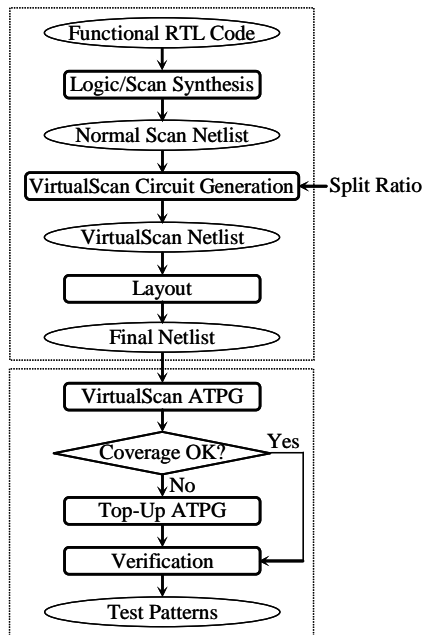


Fig. 7. Gate-Level VirtualScan Flow

Notably, this gate-level VirtualScan flow is very similar to any existing full-scan/ATPG test flow. This makes it easy to adopt the VirtualScan technology in any current DFT environment.

5.2 RTL VirtualScan Flow

Fig. 8 shows a RTL VirtualScan design flow. In the RTL VirtualScan flow, VirtualScan RTL blocks, including a broadcaster and a compactor, are created based on the number of scan chains and a given split ratio before logic/scan synthesis is conducted. This is possible because

the structures of both broadcaster and compactor are independent of the circuit under test. The logic/scan synthesis program then synthesizes both functional RTL blocks and VirtualScan RTL blocks, creates short scan chains, and connects them with the broadcaster and the compactor. The result is a VirtualScan netlist ready for layout. Then, VirtualScan ATPG is conducted on the netlist. If fault coverage is not enough, top-up ATPG is conducted. After verification, final test patterns are obtained, ready for manufacturing test.

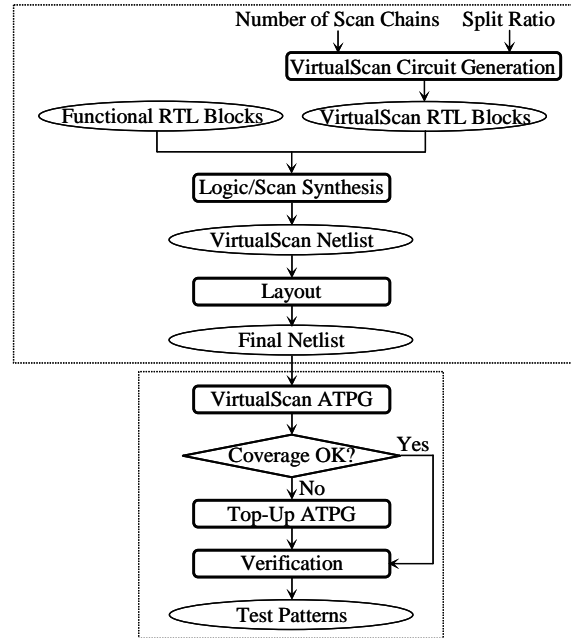


Fig. 8. RTL VirtualScan Flow

Understandably, this RTL VirtualScan flow can achieve higher performance since the logic/scan synthesis program has more flexibility to optimize a design including the circuitry added for the VirtualScan technology. In addition, moving such a design-for-testability (DFT) task to a higher level of abstraction greatly reduces the risk of design iterations caused by improper DFT insertion, thus improves the over-all design turn-around time.

5.3 Hierarchical VirtualScan

A complex SoC design usually adopts a hierarchical design style, in which individual modules are designed separately and are then integrated at the top level with some glue logic. At the module level, it is not only necessary to complete functional, logic, and even physical design, it is also preferable to complete DFT design including test pattern generation. The VirtualScan technology is suitable for this purpose.

An example is shown in Fig. 9. This design has two modules M1 and M2 as well as some glue logic at the top

level. One pair of broadcaster and compactor is inserted for each module and the top-level logic also has its own pair of broadcaster and compactor. Test patterns are also generated separately for the two modules and the top-level logic, and later combined together to form final test patterns.

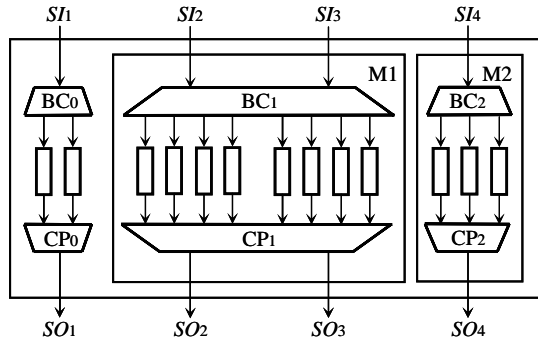


Fig. 9. Hierarchical VirtualScan

6. Application Results

The proposed VirtualScan technology has been applied in various experimental and practical settings. In the following, the results of applying the VirtualScan technology to three industrial designs are shown: one for evaluation and two for successful tape-outs.

6.1 Design Statistics

Table 1 summarizes the statistics of the three industrial designs. A clock group consists of clocks that are not inter-related with each other. That is, all clocks in a clock group can be activated simultaneously for test response capture without suffering any clock skew issue. This will greatly reduce test response capture time needed in the multi-capture clocking scheme. A tool has been developed to identify all independent clock groups. In addition, the number of scan chains is for a 1-to-1 scan configuration. That is, one external scan input/output port directly corresponds to one internal scan chain. In the applications, these scan chains were split according to a given split ratio.

Table 1. Design statistics

	Design A	Design B	Design C
Circuit Size	1.2M	4.2M	4.5M
Clocks	31	36	52
Clock Groups	7	9	12
Flip-Flops	102,647	203,578	250,364
Scan Chains	28	30	32
Max. Scan Length	3,666	7,005	7,824

Note that Design A and Design C are described by hierarchical netlists; while Design B is a flattened netlist. The VirtualScan technology is flexible enough to support both scenarios.

6.2 Results

The VirtualScan technology was applied to the three industrial designs listed in Table 1. The computer used was a SUN Blade 2000/2900 (900MHz). Tables 2 through 4 summarize the application results. The fault model used was the single stuck-at fault model.

Table 2 shows the result of applying the VirtualScan technology on Design A. In this application, the original test patterns were generated by an ATPG program (P-1) with the multi-capture feature as described in 4.2 on the 1-to-1 scan configuration of Design A. The fault coverage achieved by 2207 original test patterns was 92.14% with a turn-around-time (TAT) of 15 hours. Then, the VirtualScan technology was applied in two scenarios with two different split ratios of 10 and 20, respectively.

Table 2. Application Result for Design A

Design A		ATPG P-1	VirtualScan	
			Split = 10	Split = 20
<i>Quality</i>	Fault Coverage	92.14%	92.14%	92.03%
	Test Patterns	2,207	3,065	3,128
<i>Cost</i>	Test Data Volume (MW)	225.42	31.50	16.12
	Reduction Ratio	-	7.16	13.98
	Test Cycles	8,090,862	1,124,855	575,552
	Reduction Ratio	-	7.19	14.06
<i>Impact</i>	Design TAT-1 (hrs)	0	3	3
	Design TAT-2 (hrs)	15	12	28
	Overhead	0	0.2%	0.4%

As shown in Table 2, when a split ratio of 10 was used, the test data volume and test cycles were reduced by roughly 7 times, with no fault coverage degradation. When a split ratio of 20 was used, the test data volume and test cycles were reduced by roughly 14 times, with a slight fault coverage degradation of 0.11%. In many cases, such a slight fault coverage loss is tolerable, especially given the fact that many test patterns are currently thrown away as they cannot fit into one tester load in practice, which usually results in more significant fault coverage loss. If fault coverage loss is not allowed, top-up ATPG can be further applied. In the case of Design A under the split ratio of 20, 200 additional test patterns are generated by top-up ATPG to achieve the original fault coverage of 92.14%. For both split ratios, the time (TAT-1) needed for generating and integrating the VirtualScan circuit was 3 hours; while the VirtualScan ATPG time (TAT-2) was 12 hours for the split ratio of 10 and 28 hours for the split ratio of 20. The VirtualScan circuit overhead is 0.2% for the split ratio of 10 and 0.4% for the split ratio of 20.

Since both ATPG P-1 and VirtualScan have the same multi-capture feature, the test cost reduction effects shown in Table 1 were entirely due to the contribution of the VirtualScan architecture itself. Based on the experimental

results, it indicates that using the VirtualScan architecture alone could achieve about 70% of the theoretical maximum test cost reduction ratio, which is the split ratio. That is to say, if the split ratio is s , the test cost reduction effect from VirtualScan architecture only would be about $0.7 * s$.

In practice, ATPG programs based on the one-hot clocking scheme are still in wide use due to its simplicity but are one of the major causes of rapidly-rising test costs. Although using the multi-capture clocking scheme alone can alleviate this problem to some extent, the following two experiments showed that, by using the VirtualScan architecture together with a multi-capture based ATPG such as VirtualScan ATPG, test costs can be more significantly reduced even with a small split ratio.

Table 3 shows the result of applying the VirtualScan technology on Design B. This design was an entirely flattened netlist. A one-hot based ATPG program (P-2) was used to generate original test patterns for comparison. The split ratio was selected as 2 in this application. The time (TAT-1) for VirtualScan circuit insertion was 3 hours, and the VirtualScan circuit overhead is only 0.01%. The VirtualScan ATPG was then used to generate final test patterns. The final fault coverage is 92.61%. The test data volume and test cycles are reduced by 14 times compared with the original patterns. The 14 times reduction achieved by a split ratio of only 2 was due to the performance of the VirtualScan ATPG which has full multi-capture support. In this case, the VirtualScan ATPG generated a test set which is more than 7 times smaller than the original test set generated by the ATPG P-2 at a cost of longer CPU time (TAT-2) of 182 hours versus 89 hours for the ATPG P-2.

Table 3. Application Result for Design B

Design B		ATPG P-2	VirtualScan Split = 2
<i>Quality</i>	Fault Coverage	92.68%	92.61%
<i>Cost</i>	Test Patterns	19,094	2,546
	Test Data Volume (MW) Reduction Ratio	4379.71 -	312.85 14.00
	Test Cycles Reduction Ratio	133,829,846 -	9,529,678 14.04
<i>Impact</i>	Design TAT-1 (hrs)	0	3
	Design TAT-2 (hrs)	89	182
	Overhead	0	0.01%

Table 4 shows the result of applying the VirtualScan technology on Design C. A one-hot based ATPG program (P-2) was used to generate original test patterns for comparison. The split ratio was selected as 4 in this application. The time (TAT-1) for VirtualScan circuit insertion was 1 hour, and the VirtualScan circuit overhead is only 0.02%. The VirtualScan ATPG was then used to

generate final test patterns. The final fault coverage is 97.15%. The test data volume and test cycles were reduced by 18 times compared with the original patterns. The 18 times reduction achieved by a split ratio of only 4 was due to the performance of the VirtualScan ATPG which has full multi-capture support. In this case, the VirtualScan ATPG generated a test set which is more than 4.5 times smaller than the original test set generated by the ATPG P-2 at a cost of longer CPU time (TAT-2) of 273 hours versus 101 hours for the ATPG P-2.

Table 4. Application Result for Design C

Design C		ATPG P-2	VirtualScan Split = 4
<i>Quality</i>	Fault Coverage	97.49%	97.15%
<i>Cost</i>	Test Patterns	21,041	4,472
	Test Data Volume (MW) Reduction Ratio	5343.40 -	293.08 18.23
	Test Cycles Reduction Ratio	164,708,948 -	8,809,840 18.70
	Design TAT-1 (hrs)	0	1
<i>Impact</i>	Design TAT-2 (hrs)	101	273
	Overhead	0	0.02%

Experimental results on Design B and Design C indicate that, if the high test cost is caused by the one-hot clocking scheme, the VirtualScan technology can be used to reduce the test cost with a rather small split ratio. Note that the smaller the split ratio, the less the physical design impacts due to scan chain splitting, and the smaller the area overhead.

6.3 Discussions

From these and other experimental/application results, it can be observed that the test cost reduction effect achieved by applying the VirtualScan technology can be roughly estimated by $(s * d1)$ if the original test patterns are generated with the multi-capture clocking scheme and by $(s * d1) * (M * d2)$ if the original test patterns are generated with the one-hot clocking scheme. Here, s is the split ratio and M is the number of independent clock groups. $d1$ and $d2$ are two adjustment parameters. From previous results, it can be observed that $d1$ is often between 0.7 and 0.9 while $d2$ is often between 0.8 and 0.9. These parameters can help in estimating the split ratio that is needed to achieve a given test cost reduction goal.

In addition, the overhead of a VirtualScan circuit only depends on the split ratio and the number of internal scan chains. That is, the overhead of a VirtualScan circuit does not increase with the circuit size, making the VirtualScan technology applicable even for very large circuits.

Furthermore, there is no theoretical limit on the value of a split ratio. In practice, however, too many scan chains resulting from a large split ratio may cause difficulty in

layout design and this is true for any 1-to- n scan configuration. In addition, a large split ratio may result in higher coverage loss. However, as shown in experimental results, a reasonable split ratio can achieve a significant test cost reduction effect with no or negligible fault coverage loss.

7. Conclusions

This paper describes the VirtualScan technology for scan test cost reduction based on the idea of reducing the longest scan chain length in a full-scan circuit. The VirtualScan architecture uses a simple broadcaster for broadcasting test patterns from external scan input ports to internal scan chain inputs, and a simple compactor for compacting test responses from internal scan chain outputs to external scan output ports. Test patterns are generated directly with the one-pass VirtualScan ATPG, which fully supports the multi-capture clocking scheme and algorithmically avoids X -impact and aliasing without any extra circuitry. As a result, the VirtualScan technology can achieve significant scan test cost reduction with very low overhead on design and implementation, and can be readily adopted into any existing full-scan/ATPG test flow at either the gate level or the RT level.

More importantly, since the VirtualScan ATPG is a one-pass process, it can be readily extended to other complex fault models, such as delay faults. Related experiments are being conducted.

Acknowledgments

The authors would like to thank Mr. Tomotaka Odajima at Marubeni Solutions, Japan, for his technical support. The authors also thank reviewers for their helpful comments.

References

- [1] M. Abramovici, M. Breuer, and A. Friedman, *Digital Systems Testing and Testable Design*, Computer Science Press, 1990.
- [2] *ITRS Roadmap: Test & Test Equipment*, 2003 Edition, <http://public.itrs.net/>, 2003.
- [3] G. Hetherington, T. Fryars, N. Tamarapalli, M. Kassab, A. Hassan, and J. Rajski, "Logic BIST for large industrial designs: real issues and case studies," *Proc. Int'l Test Conf.*, pp. 358-367, 1999.
- [4] S. Kajihara, A. Murakami, and T. Kaneko, "On Compact Test Sets for Multiple Stuck-at Faults for Large Circuits," *Proc. Asian Test Symp.*, pp. 20-24, 1999.
- [5] X. Lin, J. Rajski, I. Pomeranz, S. M. Reddy, "On Static Test Compaction and Test Pattern Ordering for Scan Designs," *Proc. Int'l Test Conf.*, pp. 1088-1097, 2001.
- [6] V. Jain and J. Waicukauski, "Scan test data volume reduction in multi-clocked designs with safe capture technique," *Proc. Int'l Test Conf.*, pp. 148-153, 2002.
- [7] X. Lin and R. Thompson, "Test generation for designs with multiple clocks," *Proc. Design Automation Conf.*, pp. 662-667, 2003.
- [8] B. Koenemann, "LFSR-Coded Test Patterns for Scan Designs," *Proc. European Test Conf.*, pp. 237-242, 1991.
- [9] S. Hellebrand, J. Rajski, S. Tarnick, S. Venkataraman, and B. Courtois, "Built-in Test for Circuits with Scan Based Reseeding of Multiple Polynomial Linear Feedback Shift Registers," *IEEE Trans. on Computers*, vol. C-44, No. 2, pp. 223-233, Feb. 1995.
- [10] J. Rajski, J. Tyszer, and N. Zacharia, "Test Data Decompression for Multiple Scan Designs with Boundary Scan," *IEEE Trans. on Computers*, vol. 47, No. 11, pp. 1188-1200, Nov. 1998.
- [11] A. Jas, B. Pouya, and N. Touba, "Virtual Scan Chains: A Means for Reducing Scan Length in Cores," *Proc. VLSI Test Symp.*, pp. 73-78, 2000.
- [12] I. Bayraktaroglu and A. Orailoglu, "Test Volume and Application Time Reduction through Scan Chain Concealment," *Proc. Design Automation Conf.*, pp. 151-155, 2001.
- [13] R. Dorsch and H. Wunderlich, "Tailoring ATPG for Embedded Testing," *Proc. Int'l Test Conf.*, pp. 530-537, 2001.
- [14] C. Barnhart, V. Brunkhorst, F. Distler, O. Farnsworth, B. Keller, and B. Koenemann, "OPMISR: The Foundation for Compressed ATPG Vectors," *Proc. Int'l Test Conf.*, pp. 748-757, 2001.
- [15] J. Rajski, J. Tyszer, M. Kassab, N. Mukherjee, R. Thompson, K. Tsai, A. Hertwig, N. Tamarapalli, G. Mrugalski, G. Eide, and J. Qian, "Embedded Deterministic Test for Low Cost Manufacturing Test," *Proc. Int'l Test Conf.*, pp. 301-310, 2002.
- [16] P. Wohl, J. Waicukauski, S. Patel, and M. Amin, "X-Tolerant Compression and Application of Scan-ATPG Patterns in a BIST Architecture," *Proc. Int'l Test Conf.*, pp. 727-736, 2003.
- [17] K.-J. Lee, J. Chen, and C. Huang, "Broadcasting Test Patterns to Multiple Circuits," *IEEE Trans. on Computer-Aided Design*, vol. 18, No. 12, pp. 1793-1802, Dec. 1999.
- [18] I. Hamzaoglu, and J.H. Patel, "Reducing Test Application Time for Full Scan Embedded Cores," *Proc. Fault Tolerant Computing Symp.*, pp. 260-267, 1999.
- [19] F. Hsu, K. Butler, and J. Patel, "A Case Study on the Implementation of the Illinois Scan Architecture," *Proc. Int'l Test Conf.*, pp. 538-547, 2001.
- [20] R. Tekumalla, "On Reducing Aliasing Effects and Improving Diagnosis of Logic BIST Failures," *Proc. Int'l Test Conf.*, pp. 737-744, 2003.
- [21] J. Rajski, J. Tyszer, M. Kassab, and N. Mukherjee, "Selective Linear Compactor for Test Responses with Unknown Values," *U.S. Pending Patent Application*, 2000.
- [22] J. Rajski, J. Tyszer, C. Wang, and S. Reddy, "Convolutional Compaction of Test Responses," *Proc. Int'l Test Conf.*, pp. 745-754, 2003.
- [23] S. Mitra and K. Kim, "X-Compact: An Efficient Response Compaction Technique for Test Cost Reduction," *Proc. Int'l Test Conf.*, pp. 311-320, 2002.
- [24] L.-T. Wang, H.-P. Wang, X. Wen, M.-C. Lin, S.-H. Lin, D.-C. Yeh, S.-W. Tsai, and K.S. Abdel-Hafez, "Method and Apparatus for Broadcasting Scan Patterns in a Scan-Based Integrated Circuit," *United States Patent Application, 20030154433*, August 14, 2003.