COMPARATIVE ANALYSIS OF SIMULATION TECHNIQUES: SCAN COMPRESSION AND INTERNAL SCAN

Parul Patel¹, Arvind Rajawat¹ and Pooja Jain²

¹Department of Electronics and Communication, Maulana Azad National Institute of Technology Bhopal, India
²Automotive and Discrete group STMicroelectronics Private Limited
Greater Noida, India

ABSTRACT

With advancement in technology, the feature size of transistors is shrinking and the transistor count in a circuit design is exponentially increasing. As a result, it is hard to control and observe internal nodes leading to complexity in locating and debugging faults specially for sequential circuits. Design for Testability (DFT) provides a way for fault detection of the circuit under test in less simulation duration with little increase in area. Many techniques are proposed under DFT for pattern simulation. In this paper, we have compared two such pattern simulation techniques namely scan compression and internal scan. The experiment is performed on different benchmark circuits, it is observed the simulation time is significantly reduced with increased coverage and a little area overhead.

KEYWORDS

DFT, Scan Compression, Internal Scan

1. INTRODUCTION

CMOS technology has attained very high advancement in scaling i.e., feature size, the designs have become more and more compact and much improved in terms of multi functioning, speed, reliability, and robustness. With these advancements there is high demand of these computing devices now a days in many industries and sectors whether it be medical sector, military sector, telecom department, domestic sectors – homes and shops etc. are all somehow related to this VLSI industry and so the demand of electronic products have boomed to a greater extent. The credit of this expansion can be allotted to the increase in memory density, low computing power and intellectual property modules. With the current trends in VLSI because of the compactness, low power consumption, higher frequency of operation there has been a huge necessity of testing of those components before delivering it to the consumers. The improvement in these areas has led to increased complexity and so the occurrence of manufacturing defects in ICs increases with decrease in transistor sizing.

Design for Testability (DFT)[1] based on scan and ATPG has been adopted as a reliable and broadly acceptable methodology that provides very high test coverage. Though for large circuits, the growing test data volume causes a significant increase in test cost because of much longer test time and elevated tester memory requirements. Scan compression techniques[2] are introduced to reduce test time. The compressed scan architecture is similar to the scan at the chip-level interface, but it contains combinational compression logic and uses many more scan chains of
shorter lengths called compressed scan chains, within the chip core. Figure 1 shows the DFT insertion flow. The first step of the DFT flow is the RTL design. In RTL design the DFT related blocks are inserted into the design. Flip flop are preceded by a multiplexer called a scan cell. Adding Multiplexer to flip flop results in Unrouted netlist. Stitching all such scan cells in chains and making requisite connection forms routed netlist. Routed netlist undergo optimization and design rule check leading to final netlist termed as DFT inserted netlist.

![DFT insertion flow](image)

2. **LITERATURE REVIEW**

The first step of the DFT flow[3] is the RTL design. In RTL design the DFT related blocks are inserted into the design and then the design is verified whether the desired response is obtained. The synthesis of the design generating netlist for further verification is carried out. Scan insertion is performed on synthesis netlist. The SDF file and library files obtained after routing and CTS are used for scan insertion. The scan synthesis process reads the RTL design, synthesizes it, tests it again, performs scan insertion and analyses the post DFT- Design. In the next step ATPG is done in which random test patterns are generated. These test patterns are passed as scan chains and simulation is carried out to test the results.

2.1. **Scan Cell**

Figure 2 shows a basic scan cell[4]. Scan cell is the building block of DFT architecture. It has two inputs namely Data in, Scan in and one out Data out. A scan cell can operate in two modes depending on the control signal scan Enable. One mode is functional mode in which the flip flop does its normal function for which it is placed in design with data in as input and another mode is test mode in which it acts as part of the scan chain with Scan in as input. A scan cell always contains at least one memory element.
2.2. Scan Chain

When all the scan cells are stitched together that is scan out of preceding scan cell is connected to scan in of succeeding scan cell then the resulting chain is termed as Scan chain[4].

Figure 3 shows a design without scan(left), in which flip flop is represented by square and with scan(right) in which scan cell is shown in square where the rectangle inside square represents multiplexer. With full scan design(right), the red line shows connection forming scan chains and all scan cells have the same control signal (scan enable) shown in blue. A practical circuit may contain more than one scan chain.

When scan enable is ‘1’, the scan chain is active i.e., the scan cell is operating in scan mode. When scan enable is ‘0’, the scan chain is inactive i.e., the scan cell is operating in normal mode.

2.3. Scan Operation

Initially, the scan enable is set to ‘1’, activating scan mode[5]. In this mode, Scan data is loaded in scan flops, Shift-In Phase (refer to Figure 4). The clock applied in this period to loading data is termed as shift clock. Once all the scan cells are loaded with test vector then the scan enable is set to ‘0’, activating functional mode. Test vector is forced at primary input and after action of combinational logic, output is available at certain scan cells input. In figure 3, clouds represent combinational logic. One more clock pulse is required to observe the output at the scan out port.
This pulse is termed as capture pulse. Once the output is captured, scan enable is again set to ‘1’, enabling next pattern load.

**Figure 4. Scan Operation**

---

### 2.4. Fault

Defect arises when the implemented hardware behaviour varies from the desired/required behaviour (design specification). Manifestation of defect at abstracted functional level is termed as fault[6]. Source of fault can be improper masking during fabrication, missing contact window, surface impurities, parasitic formation. Faults are classified as stuck at fault, bridging fault, transition fault, path delay fault, transistor fault. Stuck at fault occurs when a line is stuck at logic ‘0’ or logic ‘1’ implying if a line is permanently connected to ground or permanently connected to power supply. There can be single or multiple stuck at fault. Methodology of representing fault abstractly is called fault model. While designing a fault model to reduce complexity, it is assumed only one fault occurs at a time.

### 2.5. Coverage

All possible defects cannot be detected by single fault model. Effectiveness of any fault model is measured in terms of coverage[5]. Coverage is a quantitative quality formulated as

\[
\text{Fault coverage} = \frac{\text{Total number of faults detected}}{\text{Total number of faults}}
\]

### 3. **Internal Scan**

In this technique[7], a long chain of scan cells is formed such that the design is divided into partially or fully isolated combinational blocks. A design can have more than one scan chain comprising thousands of scan cells. All chains scan enabled are connected, thus all scan chains are in scan mode or normal mode at the same time. This method has simplified the pattern generation problem and enabled serial shift property in sequential elements other than their normal function. This test mode is created along with compressed scan insertion automatically. Refer to figure 5, left side arrangement containing two scan chains refers to internal scan chains.
4. SCAN COMPRESSION

In scan compression[7], the long scan chain is divided into smaller scan chains without altering the number of scan in and scan out ports. To do so, a compressor and decompressor are used. Compressor and Decompressor[8] can have different logic to perform compression and decompression action like by applying XOR[9] on smaller chain scan to get final output from compressor. As more elements are added to circuit (codec), there is a little variation in the number of faults of internal scan and scan compression mode. Figure 5 shows that the 2-scan chain arrangement(left) is internally broken into smaller chains(right) but at the chip level the scan-in pins will be same as in internal scan. This methodology reduces the simulation time significantly as the number of scan cells per chain decreases, number of clocks required in shift cycle reduces and all scan chains are acting in parallel. Amount of compression is calculated as

\[
\text{Compression ratio} = \frac{\text{Number of internal scan chains}}{\text{Number of scan channels}}
\]

Refer to figure 6, compression ratio is 5:2 as the number of internal scan chains (chain within decompressor and compressor) is 5 and 2 scan channels that are feeding the decompressor.

5. EXPERIMENTAL SETUP AND RESULT

This experiment is performed on 5 benchmark circuits consisting of three ISCAS’89 circuit, one ITC’99 and one FARADAY circuit namely s38584, s38417, s35932, B19, RISC respectively. Refer table 1 for basic information about benchmark circuits used. The Netlist/RTL of mentioned circuits are first converted to DFT inserted netlist using Synopsys –Design Compiler. Synopsys –TetraMAX (Version - p-2019.03-sp5)[10] tool is used to generate test patterns. This tool takes a test protocol file and DFT inserted netlist as input and produces test patterns in stil format. Using
Tetramax utility stil2verilog the test pattern in Verilog format is generated which is used for simulation. Only Static faults are targeted in this experiment. Cadence -Simvision tool is employed (Version Xcelium 19.09.009) for simulation. Table 2 shows the comparison between the methodology in terms of number of scan cells per chain, number of faults detected, simulation time and coverage percentage. As the number of scan cells are less in scan compression, the number of scan chains are more. Number of flip flops in design (column 2 of Table 1) is greater than the number of scan cells in the internal chain (column 4 of Table 2) as during optimization certain flip are depicted as memory elements and those flops are not part of scan insertion. To test memory flops, usually built-in self-test circuits are embedded in the design.

Table 1. Description of Benchmark circuit

<table>
<thead>
<tr>
<th>Benchmark Circuit</th>
<th>Number of Flip Flop</th>
<th>Number of inverters</th>
<th>Number of gates</th>
</tr>
</thead>
<tbody>
<tr>
<td>s38584</td>
<td>1452</td>
<td>7805</td>
<td>11448</td>
</tr>
<tr>
<td>s38417</td>
<td>1636</td>
<td>13470</td>
<td>8709</td>
</tr>
<tr>
<td>s35932</td>
<td>1728</td>
<td>3861</td>
<td>12204</td>
</tr>
<tr>
<td>B19</td>
<td>6592</td>
<td>20866</td>
<td>98726</td>
</tr>
<tr>
<td>RISC</td>
<td>7599</td>
<td>7370</td>
<td>44872</td>
</tr>
</tbody>
</table>

Table 2. Comparison Table

<table>
<thead>
<tr>
<th>Benchmark Circuit</th>
<th>Simulation Technique</th>
<th>Scan cells per chain</th>
<th>Number of chains</th>
<th>Fault detected</th>
<th>Coverage (%)</th>
<th>Simulation time(ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>s38584</td>
<td>Internal Scan</td>
<td>915</td>
<td>1</td>
<td>49529</td>
<td>87.48</td>
<td>2223752</td>
</tr>
<tr>
<td></td>
<td>Scan Compression</td>
<td>14 or 15</td>
<td>79</td>
<td>56810</td>
<td>94.39</td>
<td>84025</td>
</tr>
<tr>
<td>s38417</td>
<td>Internal Scan</td>
<td>1425</td>
<td>1</td>
<td>52205</td>
<td>85.21</td>
<td>3210775</td>
</tr>
<tr>
<td></td>
<td>Scan Compression</td>
<td>14 or 15</td>
<td>117</td>
<td>58591</td>
<td>95.64</td>
<td>148025</td>
</tr>
<tr>
<td>s35932</td>
<td>Internal Scan</td>
<td>1380</td>
<td>1</td>
<td>53128</td>
<td>80.07</td>
<td>1105625</td>
</tr>
<tr>
<td></td>
<td>Scan Compression</td>
<td>14 or 15</td>
<td>107</td>
<td>60141</td>
<td>90.64</td>
<td>16525</td>
</tr>
<tr>
<td>B19</td>
<td>Internal Scan</td>
<td>6142</td>
<td>1</td>
<td>600743</td>
<td>98.91</td>
<td>86476825</td>
</tr>
<tr>
<td></td>
<td>Scan Compression</td>
<td>79 or 80</td>
<td>79</td>
<td>604380</td>
<td>99.51</td>
<td>187025</td>
</tr>
<tr>
<td>RISC</td>
<td>Internal Scan</td>
<td>7341</td>
<td>1</td>
<td>328315</td>
<td>97.01</td>
<td>96009750</td>
</tr>
<tr>
<td></td>
<td>Scan Compression</td>
<td>109 or 110</td>
<td>72</td>
<td>330960</td>
<td>97.80</td>
<td>3249025</td>
</tr>
</tbody>
</table>

6. CONCLUSIONS

As the technology is shrinking, testability is getting complex. By this comparative analysis in terms of coverage, time, and fault, it is evident that scan compression has the upper hand over internal scan. Scan compression not only reduces test time significantly but also yields better coverage. For s38584, simulation time for internal scan is reduced to 96.2% in scan compression simulation technique. Similarly, reduction in simulation time for s38417, s35932, B19, RISC is 95.4%, 98%, 99.7%, 97% respectively. Though the coverage increase is a little around 9% for
small circuits s38584, s38417, s35932 and around 1% for large circuit B19 and RISC. Above results clearly show the difference in two simulation techniques. When simulation performed on different circuits having different numbers of sequential elements, it is evident that as the number of flip flops increases the simulation time required for internal scan technique is significantly longer than scan compression. Nowadays, even the smallest industrial design contains more than 20,000 flip flops, scan compression gives better coverage in less time and as chains are smaller, debugging is easier.

Figure 7 shows area variation with respect to number of components like AND gate, flip flop, inverter in design. Blue bar is for normal scan insertion and orange bar reflects cells after scan compression architecture is embedded. Using Scan compression is very time efficient. Though area overhead increases nearly by 10-20% after scan compression, 12.2% increase in RISC area and 13.6% increase in B19 area.

![Figure 7. Graph showing area variation.](image)

### 7. Future Scope

From the experiment it is clear that scan compression yields better coverage in lesser time, but the area overhead is considerable. In future, work can be done in reducing the area of decoder logic and reducing pattern count keeping into account achieved coverage and time.

### REFERENCES


AUTHORS

Parul Patel is an Intern in STMicroelectronics, India and pursuing M. Tech in VLSI Design and Embedded Systems from MANIT Bhopal, India. She is part of DFT team and deals with scan insertion and boundary scan in design.