

Segmented Addressable Scan Architecture

Ahmad Al-Yamani, Erik Chmelar, and Mikhail Grinchuck

TNT, LSI Logic Corporation

Milpitas, CA 95035

Abstract

This paper presents a test architecture that addresses multiple problems faced in digital IC testing. These problems are test data volume, test application time, test power consumption, and tester channel requirements. With minimal hardware overhead, the architecture provides at least an order of magnitude reduction to each of the above problems. The architecture relies on scan chain segmentation and multiple-hot decoders.

1. Introduction

The scan-path method is a widely used design for testability (DFT) technique. It is based on serialization of test data [McCluskey 86]. In scan-based testing, the flip-flops in the circuit under test are connected together to form one or multiple scan chains. Through this scan chain, arbitrary test patterns can be shifted into the flip-flops and applied to different parts of the circuit. The main advantage of scan testing is improving the controllability and observability of the circuit under test by having direct access to the states of the flip-flops.

In scan-path methods, the circuit is designed so that it has two modes of operation: normal functional mode and scan mode. In scan mode, the flip-flops in the circuit are connected as a shift register so that it is possible to shift an arbitrary test pattern into the flip-flops. By returning the circuit to normal mode for one clock cycle, the outputs of the combinational circuitry are stored in the flip-flops. The circuit can then be placed in the scan mode to shift out the contents of the flip-flops and compare them with the correct response.

Shifting a test pattern into the scan chains is a time consuming process. If the circuit has a single scan chain, shifting a test pattern takes as many clock cycles as there are flip-flops in the circuit. The stumps architecture, shown in Figure 1, groups the flip-flops into multiple scan chains to enable loading a test pattern into the

scan chains in parallel. On the downside, it requires a tester channel per scan chain input to load the test pattern and a tester channel per scan chain output to shift out the circuit response.

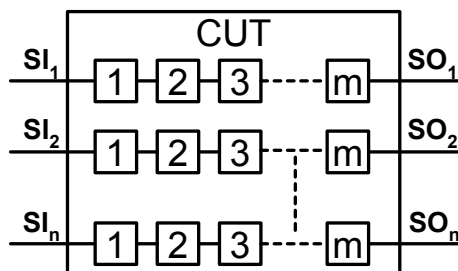


Figure 1 Multiple scan chains architecture.

Several compression techniques were presented in the literature to reduce the data volume and the tester channel requirements. A generic architecture for such techniques is shown in Figure 2. In such techniques, a compressed vector is loaded from the tester into the decompression circuitry, which expands the vector into a test pattern in the scan chains. The test response is also compressed into a smaller vector using the output compression circuitry. To name a few, [Koenemann 91], [McCluskey 03], [Alyamani 03] and [Rajski 04] discuss such compression techniques.

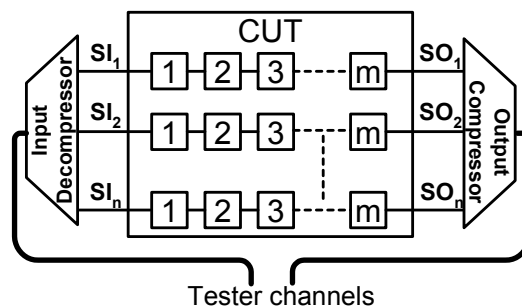


Figure 2 A generic architecture for input and output compression.

Illinois Scan Architecture (ISA) was introduced in [Hamzaoglu 99] to reduce data volume and test application time by splitting the scan chain into multiple segments and

broadcasting the data to all of them as long as the segments data are compatible.

This paper presents a new architecture and circuitry for significantly reducing test data volume and test application time with segmented scan architectures by enabling many different compatibility configurations among multiple segments. The architecture also reduces test power consumption and tester channel requirements.

Section 2 of the paper reviews some related literature. Section 3 discusses the new architecture. Section 4 presents the implementation of the decoder for the SAS architecture. Section 5 shows some experiments results and Sec. 6 concludes the paper.

2. Related work

Illinois Scan Architecture (ISA) was introduced to reduce data volume and test application time. The basic architecture for Illinois scan is shown in Figure 3 [Hamzaoglu 99]. A given scan chain is split into multiple segments. Since a majority of the bits in ATPG patterns are don't care bits, there are chances that these segments will have compatible vectors (not having opposite care bits in one location). In this case, all segments of a given chain are configured in broadcast mode to read the same vector. This speeds up the test vector loading time and reduces the data volume by a factor equivalent to the number of segments. In case if the segments within a given scan chain are incompatible, the test vector needs to be loaded serially by reconfiguring the segments into a single long scan chain. The fact that a majority of the ATPG bits (95-99% [Hiraide 03]) are don't care bits makes ISA an attractive solution for data volume and test time.

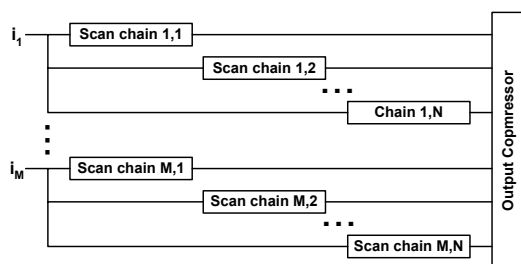


Figure 3 Illinois scan architecture.

Several enhancements to the scan architecture have been proposed and discussed in the literature for multiple reasons.

Lee et. al. presented a broadcasting scheme where ATPG patterns are broadcasted to multiple scan chains within a core or across multiple cores. The broadcast mode is used when the vectors going into multiple chains are compatible [Lee 99].

[Huang 01] introduced a token scan architecture to gate the clock to different scan segments while taking advantage of the regularity and periodicity of scan chains. Another scheme for selective triggering of scan segments was proposed in [Sharifi 03].

A novel scheme was presented in [Sinanoglu 02] to reduce test power consumption by freezing scan segments that don't have care bits in the next test stimulus. By only loading the segments that have care bits, data volume, application time, and test power consumption are all reduced at once. Only one segment of the scan chain is controlled and observed at a time.

A reconfigurable scheme was introduced in [Samaranayake 03] to use mapping logic to control the connection of multiple scan chains. This increases the chances of compatibility between multiple chains and hence makes room for additional compaction. The scheme we present here achieves a similar effect while eliminating the need for the mapping logic and reducing power consumption significantly.

A new scan architecture was proposed in [Xiang 03] to order the scan cells and connect them based on their functional interaction.

A circular scan scheme was presented in [Arslan 04] to reduce test data volume. The basic concept is to use a decoder to address different scan chains at different times. This increases the number of possible scan chains (2^{N-1} for an N -input decoder). Also, the output of each scan chains is reconnected to its input. This enables reusing the contents of the response captured in the chain as a new test stimulus if they are compatible.

The previous schemes are either limited in how much they can benefit from compatibility between some of the segments or don't address the issue of power consumption during scan or both.

Another attempt for using decoder-based segmentation is available in [Rosinger 04]. In this scheme the authors control the clocks to various segments through a regular decoder. The main advantage of the scheme is power reduction during scan and capture. The solution doesn't address data volume, or test application time.

The architecture presented in this paper enhances the benefit from all scan segmentation schemes by avoiding the limitation of having to have all segments compatible to benefit from the segmentation. In other words, any combination of segments can be compatible to lead to reduction in the test stimuli loaded. This is done with minimal overhead due to the multiple-hot decoder explained in Sec. 4. The scheme simultaneously addresses data volume, test time, power, and tester channel requirement.

3. Segmented Addressable Scan

The segmented addressable scan (SAS) architecture incorporates some of the basic concepts from Illinois scan [Hamzaoglu 99] and from scan segment decoding [Arslan 04] [Rosinger 04]. It has the following distinguishing features: (1) Multiple segments can be loaded simultaneously while maintaining the freedom of changing which segments are grouped together within a give test pattern. Unlike the technique in [Samaranayake 03], this is done without additional mapping logic. Such reconfiguration of compatibility allows for significant additional compaction leading to gains in data volume reduction. It also enables aggressive parallelism of the scan chains leading to reduction in test application time. (2) The compatible segments are loaded in parallel using a multiple-hot decoder (explained in Sec. 4) instead of loading them serially and this further reduces the time required to load the scan chains. For previous schemes, if only a subset of the segments satisfies compatibility, that subset cannot be loaded in parallel. (3) The segments that are not loaded within a given round are not clocked leading to power savings which in turn allows for faster clocking of the test patterns within the same power budget.

The basic blocks of the SAS architecture are shown in Figure 4.

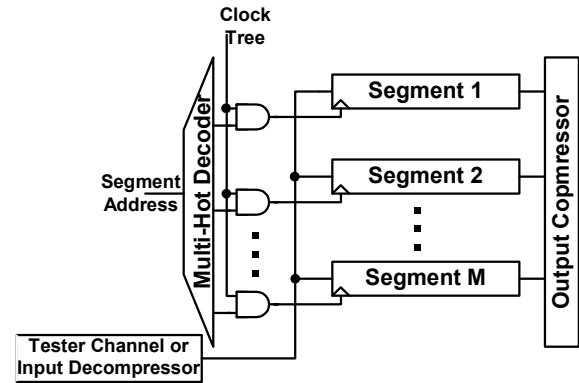


Figure 4 Segmented Addressable Scan (SAS)

The architecture shown can be implemented with a single decoder for the entire design or multiple decoders for multiple scan chains. The multiple-hot decoder (MHD) is used to activate all compatible segments within a given compatibility class. A given address is loaded into the MHD to refer to a single or multiple segments. The multiple activation of compatible segments allows the technique to benefit from the compatibility between the segments even when not all segments are compatible. This gives a significant advantage over previous segmentation techniques. Take for example the following 30-bit test pattern:

```

Pattern :
0X11XXX0110X0XX101X1X11X0X011X
Segment 1: 0X11X
Segment 2: XX011
Segment 3: 0X0XX
Segment 4: 101X1
Segment 5: X11X0
Segment 6: X011X

```

Assume that the scan cells are segmented into 6 5-bit segments as shown above. If a regular segmentation scheme (like ISA) is applied to the pattern, we will not be able to take advantage of partial compatibilities between segments. In other words, we will have to store and scan in the entire 30 bits. However, because of the segmentation scheme presented here, we can split the above pattern into 3 compatibility classes as follows:

```

Class 1= {Seg 1, Seg 5}= 01110
Class 2= {Seg 2, Seg 3}= 0X011
Class 3= {Seg 4, Seg 6}= 10111

```

A regular decoder scheme like the ones in [Arslan 04] and [Rosinger 04] would take advantage of the above compatibility for data volume reduction only. Because of the MHD used in the SAS architecture, test time can also be optimized based on this compatibility since the compatible classes will be loaded in parallel.

4. SAS multiple-hot decoder

The multiple-hot decoder is supposed to take the address of the segment(s) to be enabled and activate the clocks to those segments through the clock gating AND gates shown in Figure 4.

For regular one-hot decoders, the input is an address of the selected output. For the SAS multiple-hot decoder, the address can include don't care bits (d's) allowing multiple outputs to be activated. For the example shown above, compatibility class 1 can be loaded in parallel by combining the addresses of Segment 1 and Segment 5 {001, 101}, which means that the address for this class is d01. The address for class 2 is 01d, and the address for class 3 is 1d0.

For the multiple-hot decoder, don't care bits in the address need to be encoded too. We use positional cube notation [De Micheli 94] to encode 0s, 1s and don't cares as follows:

Code	Value
10	0
01	1
11	d
00	Unused

Positional cube encoding scheme results in an implementation for the multiple-hot decoder that requires the same hardware as a regular one-hot address decoder.

Here is an example for the implementation of a 2-input 4-output multiple-hot decoder:

Inputs				Outputs			
I ₁		I ₀		O ₃	O ₂	O ₁	O ₀
a	b	c	d				
0	0	0	0	X	X	X	X
0	0	0	1	X	X	X	X
0	0	1	0	X	X	X	X
0	0	1	1	X	X	X	X
0	1	0	0	X	X	X	X
0	1	0	1	0	0	0	1
0	1	1	0	0	0	1	0
0	1	1	1	0	0	1	1
1	0	0	0	X	X	X	X
1	0	0	1	0	1	0	0

1	0	1	0	1	0	0	0
1	0	1	1	1	1	0	0
1	1	0	0	X	X	X	X
1	1	0	1	0	1	0	1
1	1	1	0	1	0	1	0
1	1	1	1	1	1	1	1

The above function can be implemented with 4 2-input AND gates as shown in Figure 5. This is exactly the same hardware needed for a 2-input one-hot decoder. Similarly, for 8 segments, we need 8 3-input AND gates. In general, if we have S segments, we need S AND gates each with $\lceil \log_2 S \rceil$ inputs for the multiple-hot decoder. For clock gating, we need additional S 2-input AND gates.

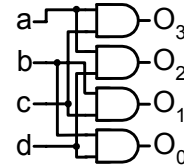


Figure 5 2-to-4 multiple-hot decoder.

The address for a given compatibility class can be loaded in parallel from the tester in which case the number of tester channels needed is $2\lceil \log_2 S \rceil + 1$ (the multiplication by 2 is a result of encoding 3 values 0, 1, and d). It can also be loaded serially. While shifting in a test pattern for one compatibility class, the address for the next compatibility class can be loaded into a shadow register. The total number of flip-flops for the shadow register is $2\lceil \log_2 S \rceil$.

The block diagram for SAS architecture with 4 segments and serial loading of address register is shown in Figure 6.

The decoder and the register hardware overhead is very minimal even with very aggressive segmentation. We calculated the size of the multiple-hot decoder, the clock gating circuitry, the address register and the shadow registers for different segmentation options. The resulting worst-case number of transistors is shown in Figure 7. The figure shows the number of transistors needed for the additional hardware. For 128 segments, we need less than 3000 transistors i.e., less than 1000 gates.

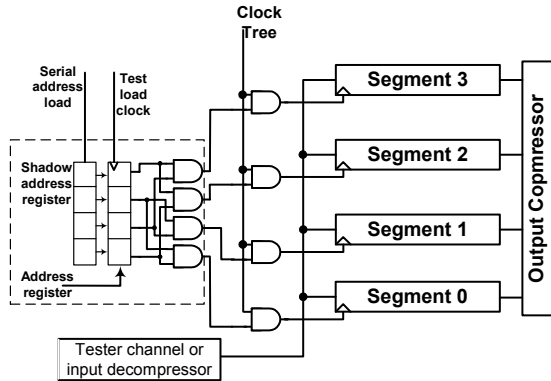


Figure 6 4-segments SAS block diagram.

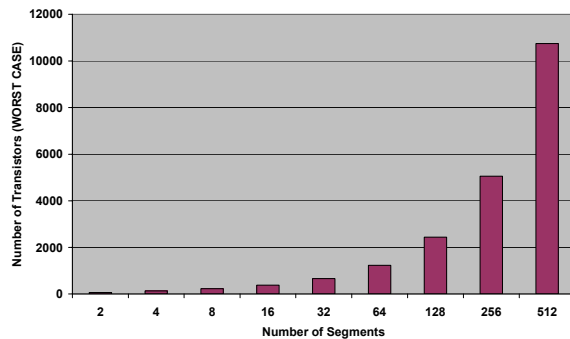


Figure 7 Number of transistors needed for SAS architecture with different segmentation options (worst case estimate).

5. Experiments and Results

SAS simulations were applied to three different circuits whose characteristics are shown in Table 1. The test patterns were generated and maximally compacted using a commercial ATPG tool.

Table 1 Circuit Characteristics

	flip-flops	Gate count	PIs	POs	Test Patterns
Ckt1	544	6,732	61	64	132
Ckt2	4,570	56,515	25	85	106
Ckt3	1,359	19,188	44	59	1,383

For the experiments, we applied simple ISA architecture and our SAS architecture with different segmentation options. The data volume requirements for both techniques and for the regular scan based testing are shown in Figure 8, Figure 9 and Figure 10 for the three circuits. We refer to regular scan data volume as the total data volume since it represents an upper bound.

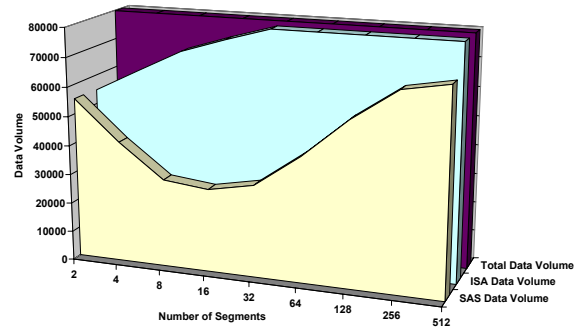


Figure 8 Data volume requirements (Ckt1).

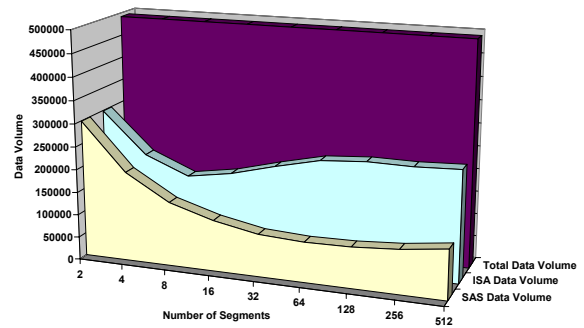


Figure 9 Data volume requirements (Ckt2).

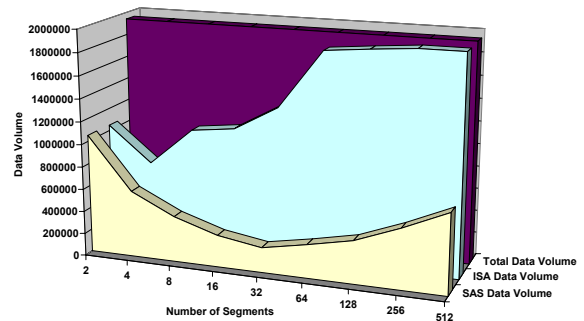


Figure 10 Data volume requirements (Ckt3).

We also compared the three options in terms of test time and the results are shown in Figure 11, Figure 12 and Figure 13. The total test time in these figures represents loading all scan cells and primary inputs serially since that represents an upper bound on shifting the test patterns.

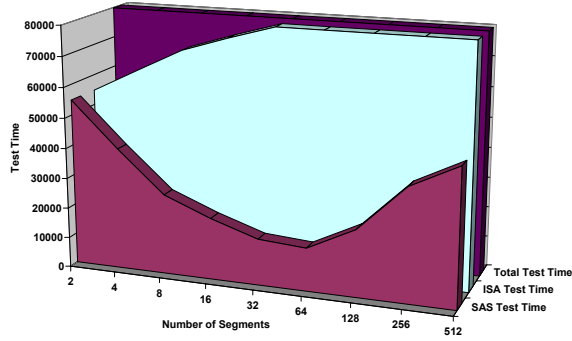


Figure 11 Test time (Ckt1).

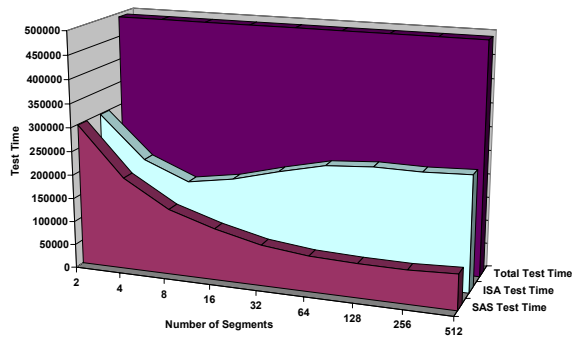


Figure 12 Test time (Ckt2).

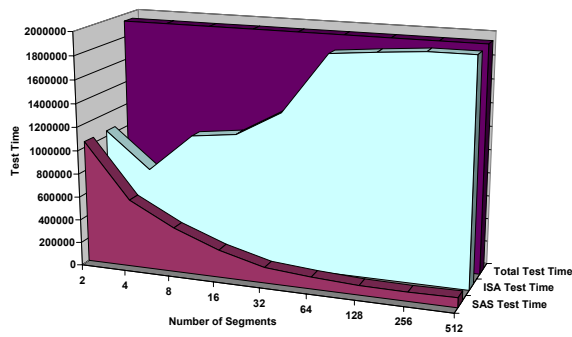


Figure 13 Test time (Ckt3).

The user of the scheme needs to select the number of segments. For selecting a certain segmentation over the others, a cost function has to be evaluated in terms of all variables (data volume, test time, and hardware overhead). The following is a general form for such a cost function:

$$Cost = C_t \cdot time + C_d \cdot data + C_h \cdot hardware$$

where C_t , C_d , and C_h are the weights for test time, data volume and hardware overhead. The

user controls these weights based on the importance of one factor over the others.

As an example we used the same weight for a clock cycle, a storage bit and a transistor to evaluate the cost function given test time, data volume and hardware overhead. The normalized cost function given different segmentations for the three circuits is shown in Figure 14. The normalization was based on the maximum cost for each circuit. A quick observation from the cost function curves show that as the circuit gets larger the optimal number of segments increases which allows for better optimization of test time, data volume and tester channels.

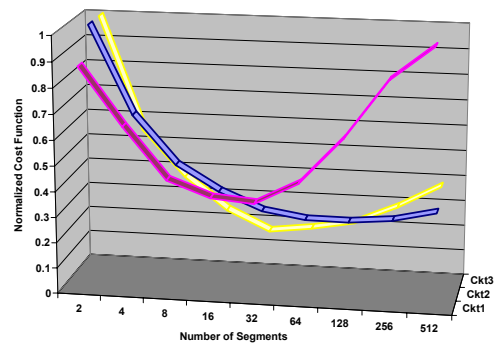


Figure 14 Example cost function.

6. Conclusions

The segmented addressable scan architecture presented in this paper provides a very attractive solution for several test problems that the industry is facing today. (1) As expected and shown by the experiments, it reduces the data volume significantly and can be combined with other compression or reseeding techniques for further reduction. (2) It also solves the test time problem by allowing very aggressive segmentation without requiring additional test channels and also by allowing parallel loading of compatibility classes. (3) It also solves the pin count problem by reducing the number of pins required to logarithm the number of scan segments. Further reduction is possible through serial loading. (4) Finally, it reduces the power consumption during test application by selectively activating the segments that need to be loaded and not clocking the other segments.

SAS provides a solution for all the above problems with very minimal hardware overhead and no routing overhead.

References

- [Alyamani 03] A. Al-Yamani and E.J. McCluskey, "Seed Encoding for LFSRs and Cellular Automata," *40th Design Automation Conference (DAC'03)*, June 2003.
- [Arslan 04] A. Arslan and A. Orailoglu, "CircularScan: A Scan Architecture for Test Cost Reduction," *Design, Automation and Test in Europe Conference and Exhibition (DATE'04)*, Vol. 2, pp. 1290-1295, Feb. 2004.
- [De Micheli 94] G. De Micheli, *Synthesis and Optimization of Digital Circuits*, McGraw-Hill, 1994.
- [Koenemann 91] B. Koenemann "LFSR-Coded Test Patterns for Scan Designs," *European Test Conference (ETC'91)*, pp. 237-242, 1991.
- [Koenemann 00] B. Koenemann "System for Test Data Storage Reduction," *US Patent 6,041,429*, Mar. 2000.
- [Lee 99] K-J. Lee, J-J. Chen and C-H. Huang, "Broadcasting Test Patterns to Multiple Circuits," *IEEE Transactions on Computer-Aided Design (TCAD)*, Vol. 18, No. 12, pp. 1793-1802, Dec. 1999.
- [Hamzaoglu 99] I. Hamzaoglu and J. Patel, "Reducing Test Application Time for Full Scan Embedded Cores" *IEEE International Symposium on Fault Tolerant Computing (FTC'99)*, pp. 260-267, 1999.
- [Hiraide 03] T. Hiraide, K.O. Boateng, H. Konishi, K. Itaya, M. Emori and H. Yamanaka, "BIST-Aided Scan Test – A New Method for Test Cost Reduction," *VLSI Test Symposium (VTS'03)*, pp. 359-364, Apr. 2003.
- [Huang 01] T-C. Huang and K-J. Lee, "A Token Scan Architecture for Low Power Testing," *International Test Conference (ITC'01)*, pp. 660-669, Oct. 2001.
- [McCluskey 86] E.J. McCluskey, *Logic Design Principles with Emphasis on Testable Semicustom Circuits*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1986.
- [McCluskey 03] E.J. McCluskey, D. Burek, B. Koenemann, S. Mitra, J. Patel, J. Rajski and J. Waicukauski, "Test Data Compression," *Design & Test of Computers*, Vol. 20, No. 2, pp. 76 – 87, March-April 2003.
- [Rajski 04] J. Rajski, J. Tyszer, M. Kassab and N. Mukherjee, "Embedded Deterministic Test," *IEEE Transactions on Computer-Aided Design (TCAD)*, Vol. 23 , No. 5 , pp. 776-792, May 2004.
- [Rosinger 04] P. Rosinger, B.M. Al-Hashimi, and N. Nicolici, "Scan Architecture With Mutually Exclusive Scan Segment Activation for Shift-and Capture-Power Reduction," *IEEE Transactions on Computer-Aided Design (TCAD)*, Vol. 23 , No. 7 , pp. 1142-1153, July 2004
- [Samaranayake 03] S. Samaranayake, E. Gizdarski, N. Sitchinava, F. Neuveux, R. Kapur and T. Williams, "A Reconfigurable Shared Scan-in Architecture" *VLSI Test Symposium (VTS'03)*, Apr. 2003.
- [Sharifi 03] S. Sharifi, M. Hosseinabadi, P. Riahi and Z. Navabi, "Reducing Test Power, Time and Data Volume in SoC Testing Using Selective Trigger Scan Architecture," *International Symposium on Defect and Fault Tolerance (DFT'03)*, 2003.
- [Sinanoglu 02] O. Sinanoglu and A. Orailoglu, "A Novel Scan Architecture for Power-Efficient, Rapid Test," *International Conference on Computer-Aided Design (ICCAD'02)*, pp. 299-303, Nov. 2002.
- [Xiang 03] D. Xiang, J. Sun, M. Chen and S Gu, "Cost-Effective Scan Architecture and a Test Application Scheme for Scan Testing with Non-scan Test Power and Test Application Cost" *US Patent Application 20040153978*, Aug. 2004.