

Enhanced FSS Architecture for Shift Power Optimization using Machine Learning

Shishira Shetty K S¹, Dr R. Jayagowri²

^{1,2}Dept. of Electronics and Communication, B.M.S College of Engineering, Bangalore, India.

Abstract—Scan shift power consumption is one of the major concerns in low power circuits. While there are multiple design for testability (DFT) techniques proposed in the literature for addressing both peak and average shift power optimization, most of the solutions impose additional design overhead which may impact functional performance of the device. In this work a novel frequency scaled segmented scan architecture which adopts a custom test method based on scan shift frequency scaling and shifting of scan data using multi-phase clock latching mechanism. Structured scan based testing is one of the widely used method to screen devices for various manufacturing defects. In addition, critical issue is the selection of tests covering the scenarios, which can lead to power issues during test application. After the accurate analysis of some tests covering worst-case scenarios, it might be still possible that other tests exist, which will cause power-related issues. Also the corrective iteration for some patterns may not be applicable to other test patterns. This work targets the identification of worst-case test power related issues by means of prediction. We propose the use of a Machine Learning (ML)-based prediction mechanism to characterize the test power behavior of all tests, thereby avoiding the detailed resource-consuming analysis of all test patterns.

1. INTRODUCTION

The testing power is one of the major issues of current generation IC testing. The extensive use of IP cores in SoC has further exaggerated the testing problem. Because of the hidden structure of IP cores, the SoCs containing large IP cores can use only those power reduction technique which do not require any modification or insertion in architecture of IP core and also do not demand the use of any test development tools like ATPG or scan insertion. These methods must be capable of using ready made test data. The dynamic dissipation is the dominant term of power dissipation. The dynamic power dissipation can be minimized by test vector set generated to minimize the frequency of switching at circuit lines during test application. From a large pool of diverse available techniques for switching activity reduction during the external testing suitable to hidden structure of IP core is difficult. The test vector reordering can address this issue. While many techniques have evolved to address power reduction in application use modes of chip operation since it defines the product specification, power consumption in the test mode of operation had often been overlooked until various issues such as burnt sockets due to thermal runaway during burn-in tests and spurious yield issues due to elevated test mode IR drop especially during scan tests have occurred. As the test application process is dominated by shift operations, average power mostly depends on scan shift power, and thus, the impact of capture power on average power is negligible.

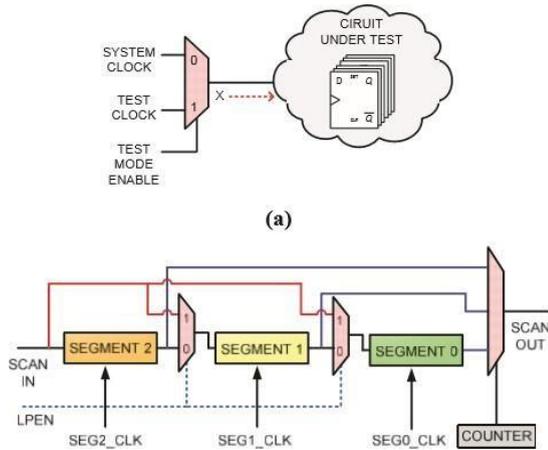
2. RELATED WORK

Many techniques have been proposed in the literature for scan shift power reduction recently. On ATPG front, test vector reordering [8, 9] and adjacent fill techniques [10, 11,] (X-filling) are the two most adopted methods to reduce scan shift power. Test vector reordering primarily focuses on reducing the number of scan-shift transitions between two consecutive vectors by changing the order of the test vectors. The adjacent fill technique takes the unmodified ATPG vector set and fills don't care bits in test vectors by replacing them with the most recent care bit value. The downside with this approach is it results in large number of test vectors as the test vector compaction algorithms don't work well when the number of X's is few. In practice though, fill techniques have only been moderately effective in industrial designs and in the presence of test compression technique. Additionally, several DFT based solutions such as scan-cell reordering, gate insertion, modified Scan cell designs and usage of reconfigurable inverters on scan paths have also been proposed to reduce scan shift power. But the need for additional hardware and routing which degrades the functional performance of the CUT has majorly limited the practical adaptability of these methods.

The other most used DFT method to solve the shift power problem is scan partitioning or segmentation via clock gating. This basically enables suppression of scan chain rippling effect by allowing only a few scan elements to shift at any given instant of scan operation. The chains are split into multiple segments or partitions with each segment separately clock gated. During scan operation the segments are sequentially activated by controlling the clock gates corresponding to each segment.

The other most used DFT method to solve the shift power problem is scan partitioning or segmentation via clock gating[3,4,5]. This basically enables suppression of scan chain rippling effect by allowing only a few scan elements to shift at any given instant of scan operation. The chains are split into multiple segments or partitions with each segment separately clock gated. Since this is baseline to the method proposed in this paper, let us analyze the operation of this method in the context of the circuit shown in Figure 1(a). The design under consideration has a single system (functional) clock which has override of test clock to be used in test mode. Figure 1(b) shows the implementation of a three segment native segmented scan on the same circuit where

each segment is controlled by separate clocks and the low power mode operation is enabled by LPEN (low power enable) control.



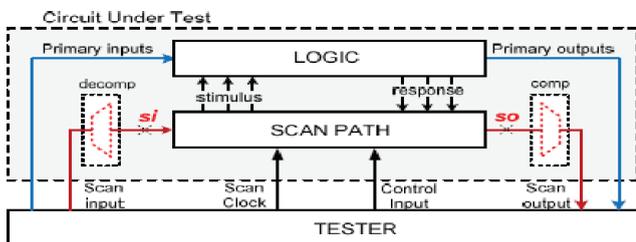
1) Native scan segment method

3. PROPOSED METHODOLOGY.

The proposed technique in this work makes use of an existing architecture [1] as a baseline on which further modifications are done to obtain even more successful power reduction. The existing architecture is known as FSS architecture where the shift frequency is reduced for decreasing the shift power consumption. In the normal scan architectures reducing of the frequency causes timing inflation proportional to the frequency reduction which is not desirable. The FSS architecture overcomes this disadvantage by using reduced frequency by considering N number of segments which are used in parallel with phase shifted clocking technique fed to each segment.

The FSS architecture has the following features:

- It does not add any design intrusive hardware which will affect the normal functional path.
- It avoids any clock tree separation on the normal functional path
- It can be scaled down to N number of segments and can hence be used for obtaining reduction in shift power or the time scaling by factor of N.
- It does not impose any additional restrictions on ATPG which requires certain modification to be done on the vector before application onto the architecture.



2) Environment for testing for a sample CUT

In the above diagram figure-2[1], the test vectors stimulus are loaded onto the logic from scan when the scan clock is triggered. The obtained response is captured and compared against the expected response in the tester to determine whether it was a pass/fail. The FSS architecture is implemented in the path SI

to SO which will be the part on which the FSS architecture and the proposed techniques are applied.

To have a better understanding let us consider of a test design which has a scan length of $N=6$ between the SI to SO. For examining the architecture, the chain is split into two segments each of length $N/2 = 3$ as shown in figure-3[1]. The segments are joined using muxes which is being controlled by the LPEN signal and SEGSW signals which aids in configuring into low power mode and to monitor the switching between the segments. The LPEN signal allows the FSS architecture to operate the scan shift-in and shift-out in two modes: 1) Default mode : where the $LPEN=0$, in which the scan chain gets configured as a single chain to during normal operation which does not use the FSS architecture.

2) Low power operation mode: where the $LPEN=1$, in which the scan chain acts as multiple segments which work in parallel for the operation.

The above diagram indicates the how the segments are configured to form FSS architecture with the addition of the MUX1 and MUX0 in-between.

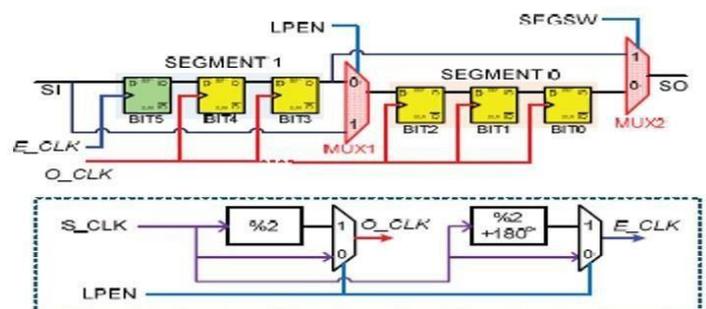
If the $LPEN=0$, then the FSS architecture is configured in the form of a single chain which is the default mode of operation. Here the scan elements are given a frequency which is the original rated scan clock (S_{clk}).

If the $LPEN=1$, then the FSS architecture gets configured in multi-segment mode of operation where the test vector from the source gets broadcasted in parallel to all the segments through SI with the aid of the phase shifted clock. Internally the architecture gets clocked by two versions of modified clock derived from the scan clock which runs at a lower frequency that is S_{clk} / N and the clocks are offset by a value of 180-degree phase shift.

The head FF is clocked by E_{clk} while the rest of the segments are clocked by O_{clk} . The first FF on the input side is referred to as head FF.

On the input side the data is fed at the rate of (S_{clk} / N) , this data gets latched alternatively using the o_{clk} and E_{clk} which efficiently de-serializes the incoming data into parallel format. The E_{clk} and O_{clk} works at half the frequency of the S_{clk} and E_{clk} is phase shifted by a value of 180-degree. Hence the shift-in cycle is reduced from 6 cycles to 3 cycles thus reducing the peak power activity to 3 cycles.

At the output side the reverse technology is used where the parallel data coming out of the segments are effectively converted into serialized data which is shifted out.

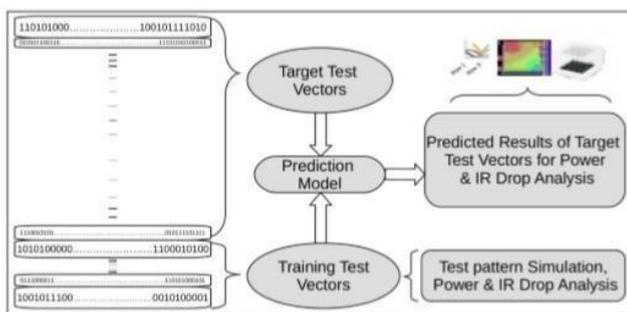


3) Example FSS circuit with the clocking module

In order to facilitate this type of parallelization of data to segments, the vectors which are to be fed in are scrambled before the application. This scrambling of the data is done post the ATPG generation and pre-application into the architecture.

To this above mentioned FSS architecture[1] a scan vector reordering technique can be applied to further reduce power consumption. To achieve this we have applied a machine learning based prediction of test power technique proposed in [2].

A critical issue is the selection of tests covering the scenarios, which can lead to power issues during test application. After the accurate analysis of some tests covering worst-case scenarios, it might be still possible that other tests exist, which will cause power-related issues. Also the corrective iteration for some patterns may not be applicable to other test patterns. This work targets the identification of worst-case test power related issues by means of prediction. We propose the use of a Machine Learning (ML)-based prediction mechanism to characterize the test power behavior of all tests, thereby avoiding the detailed resource-consuming analysis of all test patterns.



4) Illustration of the proposed method

Figure-04[2] illustrates the general idea. First, a few pre-selected tests are accurately simulated as in the regular flow. These tests are referred to as training test vectors (T_t). The training test vectors as well as the corresponding analysis results, i.e. the simulation data, are then used to train an ML model f. Then, this trained model is used to predict the power analysis result of the set of prediction target test vectors (T_p) without an explicit simulation of T_p. This methodology allows us to predict the overall power profile of a test. The application is done in two different ways:

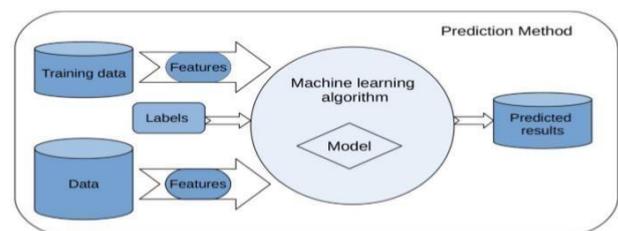
- 1) The overall (global) power consumption of a test t is targeted and predicted to identify critical tests.
- 2) Since a low global power consumption does not guarantee the absence of hot spots, the power consumption is related to the layout of the chip. In this way, local hot spots can also be predicted.

The necessary data required are as follows:

- Design related information – In order to read, store and check the extracted and learned data, design information is necessary, e.g., hierarchy and port information. ATPG generated test patterns – An essential part of the test is made up by the test patterns or test vectors, which are generated by ATPG tools.

heavily used. A scan test is first scanned in using shift cycles and then applied during the capture cycle(s). As a disadvantage in the context of test power, scan tests induce nonfunctional behavior, since the scanned state of the flip-flops is not guaranteed to be reachable in the functional mode. In our proposed methodology, the required data, i.e., the content of the scan cells during shift and capture operation for each scan chain and test pattern, is extracted from the test-pattern files and stored for further processing.

- Simulation and analysis data – Shift and during the capture operations of the applied tests cause activity in the logic parts of the design. The activity can be obtained by simulating the test patterns. After simulation, this information is extracted and can be stored, e.g., in a VCD file (or in another format), which describe the change in logic values over discrete time. Note that logic simulation usually does not process technology information and, therefore, it does not provide information regarding power consumption.
- Technology and physical layout information – For only the global power consumption is targeted, the physical location of the gates and other layout data is not required. However, this data has to be processed and stored in order to identify local hot spots. Therefore, the location and sizes of the cells are extracted from the layout information, e.g. from the .def and .lef files. This data can be fed into commercial EDA tools or other approaches in order to identify power-critical areas, e.g., by using heatmaps.



5) Prediction model

The above mentioned inputs are depicted in figure-5[2] During training, the learning algorithms, which are used as a black box, learn a function f_g to predict the value of P_t depending on the logic values at the input nodes of the circuit, i.e., the scan cells:

$$f_g(P_t) = f(v_1; v_2; \dots; v_k); [2]$$

where v₁; v₂; :::; v_k represent the value of the input nodes. This is done by continuous refinement of f_g(P_t) during training. For the i-th training example and the j-th node, the error e is internally represented as e_{j(i)} = t_{j(i)} - p_{j(i)}, where t represents the actual training value and p is the predicted value of this node. Depending on the learning algorithm used, a mathematical optimization function is applied to f_g(P_t) in order to minimize the error between the predicted value and the original value, e.g. by adjusting weights. After the training phase, the function f_g(P_t) is applied to the other test vectors T_p to predict their power profile based on the trained model f_g(P_t). The second application is the prediction of the power distribution over

the layout (local power). In order to apply the learning algorithms, the coordinates of the nodes $x; y$ in the layout have to be included in the learning process. $f_l(Pt(x; y)) = f(w_1(x_1; y_1); w_2(x_2; y_2); \dots; w_n(x_n; y_n))$ Therefore, a node n is replaced with its location $(x; y)$. By doing this, the layout of the circuit is considered for the prediction of the activity of each node. The training and the prediction of the test vectors is done in a similar way as described above. The outcome of the prediction process is then the switching activity related to the location.

4. RESULTS.

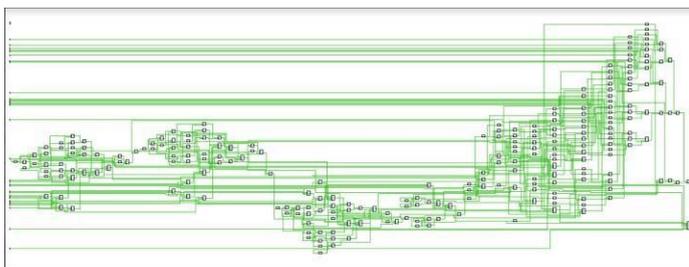
The proposed method was implemented on various benchmark circuits. For patterns generated with the help of commercial ATPG tools, pattern scrambling can be done with ease using simple user created Perl/shell scripts as an incremental step. However, given that vector scrambling process is simple and straightforward, this could effectively be integrated into the ATPG procedure, to directly provide modified patterns applicable for LPMODE.

The number of segments used is restricted to 4 here. The phase shifted clock values obtained is as indicated in the diagram.



6) Phase differed clocking

Example of the design on a S420 benchmark circuit



The above technique was tried out several benchmark circuits and the results are as tabulated below.

SL.NO	POWER MEASUREMENTS		
	CIRCUIT	FSS	MODIFIED FSS
01	S27	0.35uW	0.30uW
02	S420	1.43uW	1.41uW
03	S13207	39.3uW	36.2uW

CONCLUSION

The above-mentioned work shows that there is further reduction in the dynamic power when the vector modification is done on the existing FSS architecture. The modified vectors which are applied when similar valued vectors are fed into the segments which will in turn reduce a lot of switching activities caused by it. Hence the work provides a better reduction in power for an existing

architecture with the scaled frequency. Further, more robust and intelligent algorithms can be derived out.

REFERENCES

[1]W. Pradeep, P. Narayanan, R. Mittal and N. Maheshwari, N. Naresh “Frequency Scaled Segmented (FSS) Scan Architecture for Shift power Optimized Scan-Shift Power and Faster Test Application Time”, International Test Conference,2017.

[2] Harshad Dhotre, Stephan Eggersglu, Krishnendu Chakrabarty and Rolf Drechsler “Machine Learning-based Prediction of Test Power”, 24th IEEE European Test Symposium (ETS),2019.

2011.

[3] L. Whetsel, “Adapting Scan Architectures for Low Power Operation”, Intl. Test Conf., 2000.

[4] P. Rosinger, B.M. Al-Hashimi, and N. Nicolici, “Scan architecture with mutually exclusive scan segment activation for shift and capture-power reduction”, IEEE Trans. on Computer-Aided Design of Integrated Circuits, vol. 23, pp. 1142–1153,2004.

[5]H. S. Kim, C. G. Kim, and S. Kang, “A New Scan Partition Scheme for Low-Power Embedded Systems”, ETRI Journal, vol. 30, no. 3, pp. 412-420, 2008.

[6] P. Girard, N. Nicolici, and X. Wen, “Power-Aware Testing and Test Strategies for Low Power Devices”, Springer, 2010.

[7]S. Ravi, “Power-aware test: Challenges and solutions”, Intl. Test Conf., 2007.

[8]S. Ravi, R. Parekhji, and J. Saxena, “Low Power Test for Nanometer System-on-Chips (SoCs)”, JOLPE 4(1), pp. 81100, 2008.

[9] P. Girard, L. Guiller, C. Landrault, and S. Pravossoudovitch, “A Test Vector Ordering Technique for Switching Activity Reduction During Test Application”, IEEE Great Lakes Symp. on VLSI, pp. 24-27, 1999.

[10]J. Tudu, E. Larsson, V. Singh, and V. Agrawal, “On Minimization of Peak Power for Scan Circuit during Test”, IEEE European Test Symp., pp. 25-30, 2009.

[11] K.M. Butler, J. Saxena, A. Jain, T. Fryars, J. Lewis, and G. Hetherington, “Minimizing Power Consumption in Scan Testing: Pattern Generation and DFT Techniques”, Intl. Test Conf., 2004.