

Power Issues in SoCs: Power Aware DFT Architecture and Power Estimation

A THESIS
SUBMITTED FOR THE DEGREE OF
Doctor of Philosophy
IN THE COMPUTER SCIENCE AND ENGINEERING

by

Jaynarayan Thakurdas Tudu



Department of Computer Science & Automation
Division of Electrical Sciences
Indian Institute of Science
BANGALORE – 560 012

July 2016

©Jaynarayan Thakurdas Tudu

July 2016

All rights reserved

Dedicated to

Jagannatha

Declaration

I, **Jaynarayan Thakurdas Tudu**, with SR No. **4714-120-101-07951** hereby declare that the material presented in the thesis titled

Power Issues in SoCs: Power Aware DFT Architecture and Power Estimation

represents original work carried out by me in the **Department of Computer Science and Automation** at **Indian Institute of Science** during the years **2010 - 2016**.

With my signature, I certify that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property. I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I have understood that any false claim will result in severe disciplinary action.
- I have understood that the work may be screened for any form of academic misconduct.

Date:

Student Signature

Certificate

In my capacity as supervisor of the above-mentioned work, I certify that the above statements are true to the best of my knowledge, and I have carried out due diligence to ensure the originality of the report.

Prof. Matthew Jacob Thazhuthaveetil

Signature

Acknowledgements

It gives me immense pleasure to express my sincere thanks and gratitude to all those who have kindly helped me in carrying out this research work. I am very much thankful to Prof. Virendra Singh, IIT-Bombay and Prof Matthew Jacob for their guidance through out the research. From the very beginning of my PhD days Prof. Matthew and Prof. Singh have been guiding me in all the aspects of my research. I am particularly thankful to Prof Virendra Singh for he has kindly welcomed me to IIT Bombay and ensured my stay and research in every visit to CADS laboratory. Prof Matthew has been constantly monitoring my progress which helped me pursue the research even when I used to feel moral weakness. Prof. Matthew has provided me a several technical inputs though my research area is slightly different than his usual area of interest, the inputs has latter on resulted in the publications. Prof. Matthew has been very much accommodative in various situation particularly for my visit to IIT-Bombay. I am thankful to MHRD, Govt. of India and IISc for supporting my entire research by institute scholarship.

It was my pleasure to share and discuss some our ideas with Prof Shankar Balachandran, IIT Chennai. I respect his quick ideas and suggestions, many of our work got good shape partly because of his suggestions.

Prof Kewal Saluja of University of Wisconsin for all of our works has given many valuable suggestions and reviewed our ideas on Joint-scan architecture. In the starting of this work Prof. Saluja has given us a good motivation to look into the architectural aspect of Joint-scan. Prof Jank Patel, University of Illinois, UC, used be our frequent guest at CADSL, IIT Bombay. Many of our ideas which sprung from our daily meeting are strengthen by Prof. Janak Patel. He often boosts our moral in pursuing research with engineering attitude. Prof Janak has taught us how by taking care of minor aspects of research work, the larger problem get better understanding.

During my stay at IISc I have met with some of the fine professors whose words are filled with inspirations and encouragements. I thank Prof Y Narahari for his support in each and every difficulty of my IISc research life. Prof Narahari is the person who understand the student's difficulty from the first person perspective. Prof M Narasimha Murthy has always helped me and ensured my timely research and completion of the thesis. I am very much thankful to the current chairman of our department Prof.

Jayant Haritsa for he has always pushed me for early completion of the thesis, Professor has officially ensured all of my visits to IIT-Bombay for research to happen in fast track.

I would like to acknowledge the help I got from the technical staff member of ALTERA laboratory at Department of Electronic System Engineering, IISc in using Synopsys and other CAD tools. I also would like thank my colleagues at VLSI Design laboratory, EE, IIT-Bombay for providing me access to Synopsys and Cadence tools. I acknowledge the collaborative work with LSI Logic where we performed experiments on peak power issues.

Many of my friends have contributed to our research work. I am extremely thankful to my co-authors who have actually experimented upon many of our ideas. My colleague Deepak Malani has contributed in several aspect of our work on worst power estimation. Thanks to Alok Malakar for conducting a lot of experiments on flip-flop design, building libraries, doing place and routing of the proposed architecture. Thanks to Binod and Nehru for exploring and experimenting on dynamic and bombay scan. Our friend Satdev has done a work for which I am really thankful to him. We have designed a set of flip flops for high performance and low power scan test. Thanks to Rohini and Nihar for extending the work on worst power estimation. I am thankful to Suryakantji and Toral madam for their constant support in several technical difficulties. Our research atmosphere has always been kept vibrant by several of my friends at CADSL, I thank to Prashant, Sandipan, Nitant, Karthik, Virat, and Vijay for always keeping me encouraged in doing research of some or other kind.

IISc campus has always been a place of quest. Though I am ordinary researcher but my friends have always kept my quest alive for knowing nature or something. I thank Hrishikesh, Ramesh, Ashish, Thulasi Raman, Govind and other friends for helping me in several occasion for several reasons. We have always cherished the discussions on topic which great scientists sometime feel reluctant to discuss. Specially I thank to Ramesh and Ashish for sparing time for me to make me understand the CAD tools and the detail concept of VLSI technology.

Staying in Bangalore has always been a pleasant for me. The due goes to my senior brothers of Bhaktivedanta Institute, Bangalore who have always provided me an opportunity to explore the different horizon of life.

The love and care that I have received from my parents can not be acknowledged formally, those are priceless. I pray to lord that may I ever serve them and keep them in a happy life.

Saying is that without the mercy of supreme lord even a blade of grass does not move. Lord, you are great, may your greatness be always echoed in my mind.

Vita

I am a bottom line PhD research scholar in the Department of Computer Science and Automation, IISc, Bangalore. I have completed MSc(Engg) from the same department in the year of 2010. My master research work has been supervised by Prof. Virendra Singh and Prof. Matthew Jacob T. During my master I have worked on peak power issues in SoC testing. As part of the master work I have co-authored some of the papers in proceedings of leading test conferences. During my PhD I have worked on power issues in SoC during test and worst power estimation problem. I have co-authored a few number of papers in proceedings of some of the test conferences and symposiums. I have volunteered as a peer reviewer for ATS, VLSID, VDAT and few other conferences. Also, served as an organising member for some of the conferences and workshop on vlsi test.

I have done my BE (Bachelor of Engineering) in Computer Science and Engineering from Institute of Technical Education and Research (then affiliated to Biju Patnaik University of Technology), Bhubaneswar, Odisha in the year of 2005. I have spent a year in self learning in pursuance of higher study. Thereafter, briefly for one year I taught a BTech course at Silicon Institute of Technology, Bhubaneswar, Odisha from 2006 to 2007.

I am born and brought up in the state of Odisha. Most of my early life I have spent in village and now I am in middle of my life, spending time in moving around in different cities. I hope, one day I could return back to my village and spend a time in peace and silence with that scenic nature. I take interest in philosophy, often try to find out whether the technology have any link with the principle of life. I write blogs, visit: www.consciousnessasitis.blogspot.com for more articles.

Publications

Included in Thesis

1. Jaynarayan Tudu, Alok Malakar, and Virendra Singh, “An Architectural Approach to Efficient Hybridization of Serial and Random Access Scan”. [Manuscript under preparation for *IEEE Transaction in VLSI*].
2. Jaynarayan Tudu and Virendra Singh, “Efficient Test Control Mechanism for Joint-scan DFT Architecture”, *IEEE 25th Asian Test Symposium (ATS-2016)*, Hiroshima, Japan. [Under review].
3. Satyadev Ahlawat, Jaynarayan Tudu, Anzhela Matrosova, and Virendra Singh, “A High Performance Scan Flip-Flop Design for Serial and Mixed mode Scan Test”, *Proceedings in IEEE International Symposium on On-Line Testing and Robust System Design (IOLTS-2016)*, Catalunya, Spain, July, 2016.
4. Jaynarayan Tudu, “JSCAN: A Joint-scan DFT Architecture to Minimize Test Time, Data Volume, and Test Power”, *Proceedings in 20th International Symposium on VLSI Design and Test (VDAT-2016)*, Guwahati, India, May, 2016.
5. Jaynarayan Tudu and Satyadev Ahlawat, “Guided Shifting of Test Pattern to Minimize Test Time in Serial Scan”, *Proceedings in 20th International Symposium on VLSI Design and Test (VDAT-2016)*, Guwahati, India, May, 2016.
6. Jaynarayan Tudu, Deepak Malani, Virendra Singh, “Level-Accurate Peak Activity Estimation in Combinational Circuit Using BILP”, *Communication in Computer and Information Science, Springer Berlin Heidelberg, Proceeding in 17th VLSI Design and Test(VDAT)-2013*, Jaipur, India, pp. 345 - 352.
7. Jaynarayan Tudu, Deepak Malani, and Virendra Singh, “ILP Based Approach for Input Vector Controlled Toggle Maximization in Combinational Circuits”, *Springer Berlin Heidelberg, 16th International Symposium on VLSI Design and Test (VDAT) 2012*, Kolkata, India, pp. 172 - 179.

8. Jaynarayan Tudu, Erik Larsson, Virendra Singh, and Hideo Fujiwara, "Graph Theoretic Approach for Scan cell Reordering to Minimize Peak Shift Power", *Proc. in 20th IEEE/ACM Great Lakes Symposium on VLSI (GLSVLSI-2010)*, Rhode Island, USA.

Other Contributions

1. Rohin Gulve, Nihar Hage, and Jaynarayan Tudu, "On Determination of Instantaneous Peak and Cycle Peak Switching using ILP", *Proceedings in 20th International Symposium on VLSI Design and Test (VDAT-2016)*, Guwahati, India, May, 2016.
2. Binod Kumar, Boda Nehru, Brajesh Pandey, and Jaynarayan Tudu, "Bombay Scan: A Low Power Reconfigurable Scan Architecture", *IEEE-TTTC Reliability Aware System Design and Test (RASDAT-2016)*, Kolkata, India.
3. Binod Kumar, Boda Nehru, Brajesh Pandey, and Jaynarayan Tudu, "Skip-scan: A Methodology for Test Time Reduction", *Proceedings in 20th International Symposium on VLSI Design and Test (VDAT-2016)*, Guwahati, India, May, 2016.
4. Satyadev Ahlawat and Jaynarayan Tudu, "On Minimization of Test Power through Modified Scan Flip-flop", *Proceedings 20th International Symposium on VLSI Design and Test (VDAT-2016)*, Guwahati, India, May, 2016.
5. Satdev Ahlawat, Jaynarayan Tudu, Anzhela Matrosava, Virendra Singh, "A New Scan Flip Flop Design to Eliminate Performance Penalty of Scan", *Proc. in IEEE 24th Asian Test Symposium (ATS-2015)*, Mumbai, India.

Abstract

Test power, data volume, and test time have been long-standing problems for sequential scan based testing of system-on-chip (SoC) design. The modern SoCs fabricated at lower technology nodes are complex in nature, the transistor count is as large as billions of gate for some of the microprocessors. The design complexity is further projected to increase in the coming years in accordance with Moore's law. The larger gate count and integration of multiple functionalities are the causes for higher test power dissipation, test time and data volume. The dynamic power dissipation during scan testing, i.e. during scan shift, launch, and response capture, are major concern for reliable as well as cost effective testing. The excessive average power dissipation leads to a thermal problem which causes burn-out of the chip during testing. The peak power on other hand causes test failure due to power induced additional delay. The test failure has direct impact on yield. The test power problem in modern 3D stacked based IC is even a more serious issue. Estimating the worst case functional power dissipation is yet another great challenge. The worst case functional power estimation is necessary because it gives an upper bound on the functional power dissipation which can further be used to determine the safe power zone for the test.

Several solutions in the past have been proposed to address these issues. In this thesis we have proposed two sets of design-for-test (DFT) solutions: 1) Sequential scan chain reordering, and 2) *JScan*-an alternative Joint-scan DFT architecture to address primarily the test power power along with test time and data volume, and an integer linear programming based methodology to address the power estimation problem. We have proposed a graph theoretic based formulation for scan chain reordering and for optimum scan shift operation. For each formulation a set of algorithms is proposed. The experimental results on ISCAS-89 benchmark circuit show a reduction of around 25% and 15% in peak power and scan shift time respectively.

In order to have a holistic DFT architecture which could solve test power, test time, and data volume problems, a new DFT architecture called Joint-scan (*JScan*) have been developed. In *JScan* we have integrated the serial and random access scan architectures in a systematic way by which the *JScan* could harness the respective advantages from each of the architectures. The serial scan architecture suffers from test power, test time, and data volume problems. However, the serial scan is simple in

terms of its functionality and is cost effective in terms of DFT circuitry. Whereas, the random access scan architecture is opposite to this; it is power efficient and it takes lesser time and data volume compared to serial scan. However, the random access scan occupies larger DFT area and introduces routing congestion. Therefore, here we have proposed a methodology to realize the *JScan* architecture as an efficient alternative for standard serial and random access scan. Further, the *JScan* architecture is optimized and it resulted into a 2-Mode *2M-Jscan* Joint-scan architecture. The proposed architectures are experimentally verified on larger benchmark circuits and compared with existing state of the art DFT architectures. The results show a reduction of 50% to 80% in test power and 30% to 50% in test time and data volume. The proposed architectures are also evaluated for routing area minimization and we obtained a saving of around 7% to 20% of chip area.

Estimating the worst case functional power being a challenging problem, we have proposed a binary integer linear programming (BILP) based methodology. Two different formulations have been proposed considering the different delay models namely zero-delay and unit-delay. The proposed methodology generates a pair input vectors which could toggle the circuit to dissipate worst power. The BILP problems are solved using CPLEX solver for ISCAS-85 combinational benchmark circuits. For some of the circuits, the proposed methodology provided the expected worst possible power dissipation i.e. 80 to 100% toggling in nets.

Keywords

VLSI Testing, Design-for-Test, Random Access Scan, Multiple Sequential Scan, Congestion Aware Scan Architecture, Power aware test, Vector reordering, Scan reordering, Graph algorithm, Power Estimation, and BILP Programming

Notation and Abbreviations

C_L	Load Capacitance	IC	Integrated Circuit
C_p	Capture Pulse	ILP	Integer Linear Programming
E_{CL}	Energy Stored in Load Capacitance	JScan	Joint-scan
E_{VDD}	Total Energy Dissipation	K	Krishna
$G_w(E, V, W)$	Weighted complete graph	LOC	Launch on Capture
$G_{wd}(E, V, W)$	Weighted directed complete graph	LOS	Launch on Shift
I_{stat}	Static current	LSSD	Level Sensitive Scan Design
L_p	Launch Pulse	MISR	Multiple Input Signature Register
P_{dyn}	Average Dynamic Power	MOS	Metal Oxide Semiconductor
P_{stat}	Average Static Power	MSDFF	Multiplexed Scan D Flip-flop
V_{DD}	Drain Supply Voltage	MSS	Multiple sequential scan
2M-JScan	2-Mode Joint-scan	MT-Fill	Minimum Transition Fill
ATE	Automatic Test Equipment	MaxST	Maximum Spanning Tree
ATPG	Automatic Test Pattern Generation	MinST	Minimum Spanning Tree
ATSP	Asymmetric TSP	NMOS	N-type MOS
BILP	Binary Integer Linear Programming	OSC -	Optimum shift cycle
BIST	Built-In Self-Test	P-random	Partial random access scan
C-X	Care - don't care bit pair	P-serial	Paritial serial scan
CBR	Care Bit Ratio	PBSAT	Pseudo Boolean Satisfiability
CMOS	Complementary MOS	PCB	Printed Circuit Board
CP	Clock Signal/Pulse	PDN	Power Distribution Network
CUT	Circuit Under Test	PI	Primary Input
DC	Direct Current	PMOS	P-type MOS
DFT	Design-for-Test	PO	Primary Output
DSM	Deep Sub Micron	PPI	Pseudo Primary Input
G(V, E, W)	Unidirected Graph	PPO	Pseudo Primry Output

PRAS	Progressive Random Access Scan
PRAS-FF	Progressive random access scan flip-flop
RAM	Random Access Memory
RAS	Random Access Scan
RTL	Register Transfer Language
RVW	Response Vector Weight
SAF	Stuck-at-Fault
SAT	Satisfiability Problem
SE/SEn	Scan Enable
SF _{<i>i</i>}	<i>i</i> th Scan flip-flop
SSD-FF	Serial scan D flip-flop
SoC	System On Chip
TCL	Test Control Logic
TCLK	Test Clock
TC1	Test Clock
TDF	Transition Delay Fault
TSP	Travelling Sales Person Problem
TVM	Test Vector Weight
VLSI	Very Large Scale Integration
WOR-BIST	Word oriented build in self test
WTM	Weighted Transition Metric
X-C	Don't care - care bit pair
X-Fill	Don't Care Fill

Contents

Declaration	i
Certificate	ii
Acknowledgements	iii
Vita	v
Publications	vi
Abstract	viii
Keywords	x
Notation and Abbreviations	xi
1 Introduction and Motivation	1
1.1 Introduction	1
1.2 Power Dissipation in VLSI Circuit	3
1.2.1 Dynamic Power Dissipation	4
1.2.2 Static Dissipation	5
1.2.3 Metrics to Measure Power	6
1.3 Testing of VLSI Circuit	8
1.3.1 VLSI Design and Test Flow	9
1.3.2 Scan Architecture and Scan Testing	11
1.3.3 Shift and Capture Power	15
1.4 Power Estimation Problem	17
1.5 Thesis Contributions and Organization	18
2 Previous Work	20
2.1 Scope of the Literature Review	20
2.2 Minimization of Average Power	21
2.2.1 Circuit Modification Techniques	22
2.2.2 ATPG Based and Test Vector Modification Techniques	22
2.2.3 Test Vector Ordering Techniques	23
2.2.4 Scan Chain Ordering Techniques	23
2.3 Minimization of Peak Power	24
2.3.1 Circuit Modification Techniques	24
2.3.2 ATPG Based and Test Vector Modification Techniques	25
2.3.3 Test Vector Ordering Techniques	28
2.3.4 Scan Chain Ordering Techniques	29
2.3.5 Test Scheduling Approaches	29

2.4	Random Access Scan Architecture (RAS)	30
2.5	Worst Case Power Estimation	31
2.6	Motivation	32
3	Peak Power Minimization in Serial Scan	34
3.1	Introduction	34
3.2	Previous Work	36
3.3	Background	38
3.3.1	Overview of Peak Power During Scan Operation	38
3.3.2	Basis of Power Reduction by Scan Cell Reordering	39
3.4	Problem Formulation	39
3.4.1	Construction of Graph	39
3.4.2	Graph Theoretic Problem Formulation	41
3.5	Proposed Algorithm	43
3.5.1	Construction of Hamiltonian cycle	43
3.5.2	Construction of Hamiltonian path	46
3.6	Time and Space Complexity	48
3.6.1	Time Complexity	48
3.6.2	Space Complexity	48
3.7	Experimental Results	48
3.8	Conclusion	50
4	Test Time Minimization in Serial Scan	51
4.1	Introduction	51
4.2	Previous Work	53
4.3	Proposed Scan Chain Ordering Technique	54
4.3.1	Basic Idea	55
4.3.2	Graph Construction	56
4.3.3	Hamiltonian Path Problem Formulation	59
4.3.4	Kruskal's MaxST and MinST Algorithm	61
4.4	Test Pattern Reordering	63
4.4.1	Graph Construction	63
4.4.2	TSP and MST Based Heuristic	65
4.5	Experimental Result and Discussion	66
4.6	Conclusion and Prospects	67
5	JScan: A Joint-scan DFT Architecture	68
5.1	Introduction	68
5.1.1	Chapter Organization	70
5.2	Previous Work	70
5.3	Proposed JScan Architecture	71
5.3.1	Architecture	72
5.3.2	Functionality	74
5.3.3	Scan Cell Segregation Methodology	77
5.3.4	Test Time, Data Volume, and Power	79
5.4	Experimental Result and Discussion	80
5.5	2M-JScan Architecture	85
5.5.1	The Architecture	86
5.5.2	Limitations of JScan	88
5.5.3	Test Pattern Alignment	89
5.5.4	Test Control Mechanism	91
5.5.5	Functionality	91
5.5.6	Clustering of Scan Flip-flops	92

5.5.7	Computation of Test Parameters	94
5.6	Experimental Result and Discussion	96
5.7	Variants of Joint-Scan Architecture and Future Work	98
5.8	Conclusion	99
6	Scan Flip-flop Design	101
6.1	Introduction	101
6.2	Preliminaries and Proposed Scan Flip-Flop	105
6.2.1	Proposed Scan Flip-Flop Design	106
6.3	Application of Test Vectors	109
6.3.1	Stuck-at-fault test	109
6.3.2	Launch-on-capture test	111
6.3.3	Launch-on-shift test	111
6.4	The Proposed Scan Flip-flop as a Joint-scan Cell	112
6.5	Experimental Results	114
6.6	Conclusion	115
7	Worst Case Power Estimation	117
7.1	Introduction	117
7.1.1	Power Estimation Requirement and Challenges	117
7.1.2	Contribution	121
7.1.3	Chapter Organization	122
7.1.4	Previous Work on Power Estimation	122
7.2	Background on Binary Integer Linear Program	124
7.3	Power Estimation Considering Zero-delay Model: Formulation of BILP Problem	125
7.3.1	Input Output Constraints	126
7.3.2	Linearization Constraint	126
7.3.3	Toggle Constraints	127
7.3.4	Objective Function	127
7.3.5	Example	128
7.4	Evaluation on ISCAS-85 Circuits	130
7.5	Summary of Zero-Delay ILP Methodology	133
7.6	Level-Accurate Power Estimation: Methodology	133
7.6.1	Leveling of the Circuit	133
7.6.2	Power Estimation Methodology	134
7.7	Formulation of Level-BILP	135
7.7.1	Computation of Maximum Peak Activity	137
7.8	Evaluation on ISCAS-85 Circuits for Level-accurate BILP	137
7.9	Conclusion and Future Work	139
8	Conclusion and Future Work	140
8.1	Contributions	141
8.2	Future Work	144
A	Experimental Setup and Tools	147

List of Tables

3.1	Scan patterns and Original scan order	41
3.2	Computed TVW and RVW	42
3.3	Specification of ITC-99 (b04 to b10) and ISCAS-89 Benchmark Circuits	49
3.4	Experimental results for peak power minimization	50
4.1	Demonstration of basic idea	55
4.2	C-X and X-C pair indicator bit array	57
4.3	Experimental Results on ISCAS-89 Ckts	66
5.1	Joint-scan Modes of operation	75
5.2	Circuit specification of S-ISCAS-89 benchmark	80
5.3	JScan: P-random and P-serial size in #SFF	81
5.4	Peak and Average Power of JScan in comparison with MSS[77]	81
5.5	Test time in contrast with PRAS[10] and MSS[77]	83
5.6	Pattern volume in contrast with PRAS[10] and MSS[77]	84
5.7	Routing wire length of JScan, PRAS, and MSS	85
5.8	Variation in loading/unloading time	88
5.9	Circuit specification of S-ISCAS89 benchmark	95
5.10	Peak and Average Power in comparison with MSS[77]	96
5.11	Test time of 2M-JScan in contrast with JScan[110], PRAS[10] and MSS[77]	97
5.12	Test data volume of 2M-JScan in contrast with JScan[110], PRAS[10] and MSS[77]	98
5.13	Possible Variants of Joint-scan	98
6.1	Post Layout Timing Simulation Results at 500MHz	114
7.1	Properties of ISCAS-85[47] Benchmark Circuits	131
7.2	% of toggling nets as computed by CPLEX solver	132
7.3	Summary of BILP as presented in Section-7.3	135
7.4	Comparison of estimated switching activity and simulation time of CPLEX solver	138

List of Figures

1.1	Dynamic and leakage power phenomena in CMOS inverter	3
1.2	Basic test procedure	9
1.3	VLSI design flow [119]	10
1.4	Block diagram of basic serial scan architecture	12
1.5	Random access scan architecture	13
1.6	The power dissipation phenomena during shift cycle of scan test, Shift- i indicates the i^{th} shift cycle. 1 - 0 or 0 - 1 changes in i^{th} and $i + 1^{st}$ cycle in corresponding sff indicates the switching.	15
1.7	Power dissipation during last shift cycle for stuck-at and LOS test	16
1.8	Capture power dissipation (two toggles), Hamming distance between stimuli and response	16
3.1	Timing digram for at-speed scan testing	38
3.2	Basic principle of scan reordering	39
3.3	A complete vector-weighted graph	41
3.4	Hamiltonian cycle from Algorithm Part-1	45
4.1	A complete weighted graph constructed from Table-4.1	59
4.2	A complete directed weighted graph constructed from Table-4.1	64
5.1	The Proposed Joint-scan Architecture (the JScan)	73
5.2	Serial address shift register for PRAS using only one address pin	73
5.3	State transition machine for mode control: where, Q_j is Joint-scan mode, Q_r is Random mode, Q_s is Serial mode, and Q_f is functional mode.	76
5.4	Two configuration of JScan with respect to column address application	82
5.5	Test time behaviour for s-s15850 circuit over number of pins	83
5.6	Patten volume variation for s-s15850 circuit over number of pins	84
5.7	Proposed scan Architecture (the 2M-JScan)	87
5.8	Two states controller: Q_t for load/unload and launch of test, Q_f for response capture and normal function	91
5.9	Graphical demonstration of 1D clustering algorithm	94
6.1	Scan design performance overhead[21]	103
6.2	A conventional scan flip-flop design	105
6.3	Proposed scan flip-flop design	107
6.4	Launch and capture in <i>stuck-at</i> test	110
6.5	Launch of $V1$, $V2$, and capture in <i>LOC</i> test	111
6.6	Progressive Random Access Scan (<i>PRAS</i>) Cell [10]	112
6.7	Proposed design as a base scan cell for Joint-scan architecture	113
6.8	Post layout timing waveform at 500MHz	115
7.1	Demonstration circuit for ILP formulation with zero-delay model	128

7.2 Group of gates leveled based on the circuit depth 134

Chapter 1

Introduction and Motivation

The thesis primarily concentrated on two problems: 1) power issues in System-on-Chips (SoCs) during test and 2) worst case power estimation problem. The severity of scan test power issue is discussed with respect to modern SoC designs. The need and challenges of estimating worst case power is also discussed. This chapter introduces the problem of interest and gives a back ground on power dissipation in VLSI circuit and scan architecture.

1.1 Introduction

Moore's prediction has triggered a lot of changes in VLSI technology[53, 74]. The down scaling of device size has brought up number of innovations and along with that a set of challenges. Designs are getting complex with multiple number of functions being integrated in a single chip. Realizing such complex design is itself a challenging task. Three parameters have been at forefront of the design consideration. Performance and area have been driving the design from last few decades or so. Recently the power has gained utmost important as one of the critical design parameters. Low power design has become now a mandatory practice for modern designs across all the application domains[20, 35]. Testing being a mandatory phase of VLSI design flow (at least for critical application areas), it has witnessed a set of challenges due to the power consumption during test

process, particularly in scan based testing[21, 43]. The power consumption in general is measured in terms of three power metrics: 1) Energy, 2) Average power and 3) Peak power

Among these three, the average and peak power causes critical issues particularly during test. The average power causes thermal related issues where as the peak power causes timing related issues. How are these problems so different that they need special attention during testing? Why is the power aware design not suffice to tackle these problems? How and why the power consumed during test operation is considerably large compared to the functional power? All these questions are natural to be asked. The test researchers have investigated these questions and pin pointed the problems which requires research effort to come up with appropriate solutions. Section-1.2 and 1.3 discusses the problems in detail. First part, Chapter-3, Chapter-4, Chapter-5, and Chapter-6 of this thesis address these problems. Second part, Chapter-7, addresses the worst case power estimation problem.

Estimation of worst case power plays an important role in fixing crucial design parameters related to power distribution network and power dissipation in general[124]. For test, the worst case power estimation render a help in understanding the upper limit on power which further can be used to determine the safe power zone for the design under test[114]. Therefore, the problem have attracted a good amount efforts to fix the challenges in power estimation. There are two related challenges associated with the power estimation: 1) maintaining accuracy, and 2) consuming minimum possible computation time. Two directions have been explored, one is dynamic power estimation and the other is static power estimation. We have explored the dynamic power estimation with an objective of maintaining accuracy with reasonable amount of computational time. Section-1.4 gives a detail understanding of the problem and related definitions.

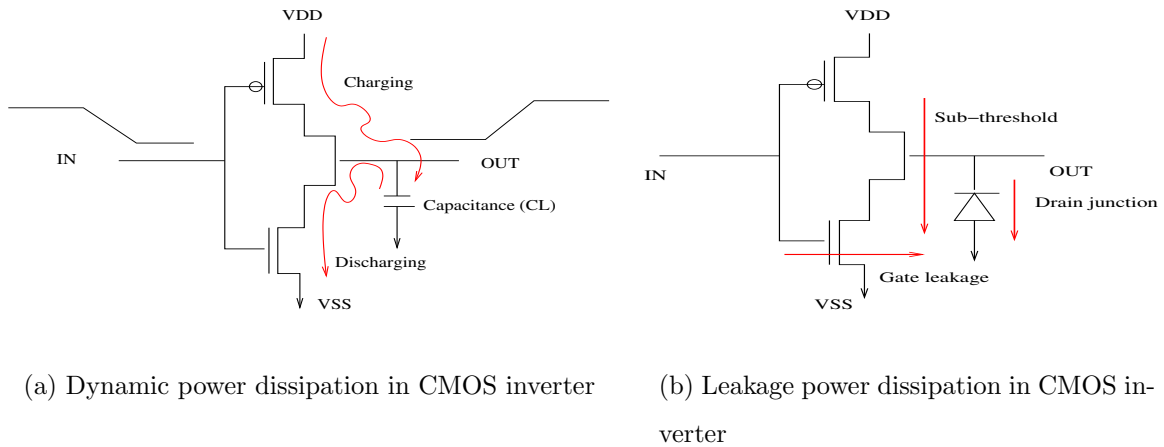


Figure 1.1: Dynamic and leakage power phenomena in CMOS inverter

1.2 Power Dissipation in VLSI Circuit

Power dissipation in digital CMOS circuit is caused by four sources [81]:

- The *leakage current*, which is primarily determined by the fabrication technology, consists of two components: 1) reverse bias current in the parasitic diodes formed between source and drain diffusions and the bulk region in a MOS transistor, and 2) the sub-threshold current that arises from the inversion charge that exists at the gate voltages below the threshold voltage.
- The *standby current*, which is the DC current drawn continuously from V_{DD} to ground.
- The *short-circuit* (rush-through) current, which is due to the DC path between the supply rails during output transitions
- The *capacitance current*, which flows to charge and discharge capacitive loads (C_L) during logic changes.

The above sources of power dissipation are explained briefly in the following sections.

1.2.1 Dynamic Power Dissipation

The *dynamic power* is defined as the sum of *capacitive* and *short-circuit* power dissipations.

Each time the capacitor C_L gets charged through the PMOS transistors as shown in Figure-1.1(a), it's voltage rises from 0 to V_{DD} , and a certain amount of energy is drawn from the power supply. Part of this energy is dissipated in PMOS device, while the remainder is stored on the load capacitor. During the high-to-low transition, this capacitor is discharged, and the stored energy is dissipated in the NMOS transistor.

The energy E_{VDD} , taken from the supply during the transition, as well as the energy E_C , stored on capacitor at the end of the transition, can be derived by integrating the instantaneous power over the period of interest. Equation 1.1 and 1.2 show the energy drawn from V_{DD} and the actual energy stored in C_L respectively. For these equations it is assumed that the NMOS and PMOS devices are never on simultaneously.

$$E_{VDD} = \int_0^\infty i_{VDD}(t)V_{DD}dt = V_{DD} \int_0^\infty C_L \frac{dv_{out}}{dt} dt = C_L V_{DD} \int_0^{V_{DD}} dv_{out} = C_L V_{DD}^2 \quad (1.1)$$

and

$$E_C = \int_0^\infty i_{VDD}(t)v_{out}dt = \int_0^\infty C_L \frac{dv_{out}}{dt} v_{out}dt = C_L \int_0^{V_{DD}} v_{out}dv_{out} = \frac{C_L V_{DD}^2}{2} \quad (1.2)$$

From above equations it can be observed that the energy stored at C_L , E_C is half the energy supplied at V_{DD} , E_{VDD} . This means that only half of the energy supplied by the power source is stored on C_L . The other half has been dissipated by the PMOS transistor. During the discharge phase the charge is removed from the capacitor, and its energy is dissipated in the NMOS transistor. In summary, each switching cycle (consisting of an low-to-high and high-to-low transition) takes a fixed amount of energy, equal to $C_L V_{DD}^2$. In order to compute the power consumption, we have to take into account how often the device is switched. If the gate is switched on and off $f_{0 \rightarrow 1}$ times per second, the power

(dynamic) consumption is given by

$$P_{dyn} = C_L V_{DD}^2 f_{0 \rightarrow 1} \quad (1.3)$$

For a complex circuit, the equation for dynamic power, Equation 1.3, can be rewritten as

$$P_{dyn} = C_L V_{DD}^2 P_{0 \rightarrow 1} f \quad (1.4)$$

where f is the maximum possible event rate of the inputs (which is often the clock rate) and $P_{0 \rightarrow 1}$ the probability that a clock event results in a $0 \rightarrow 1$ event at the output of the gate [86].

Power Dissipation due to Direct Path Current: The direct-path power is dissipated when both PMOS and NMOS transistor are conducting the current. This power is directly proportional to the switching activity similar to the capacitive power dissipation [86].

During testing the dynamic power is the primary concern. Reason being, the circuit has to be activated during the test operations. The activity that takes place during test are much more higher than the normal functional activity. Invariably all the experiment carried out till date reports that peak and average test power are around 30% and 6% more than the functional power [27, 57, 122]. However since the test process does not play any role distinctly in either increasing or decreasing the static power, it is not a critical issue for test.

1.2.2 Static Dissipation

The term *static power* (or steady-state power) dissipation refers to the sum of leakage and standby dissipations. The static power dissipation is expressed by the relation

$$P_{stat} = I_{stat} V_{DD} \quad (1.5)$$

where I_{stat} is the current that flows between the supply rails in the absence of switching activity.

Ideally, the static current of the CMOS inverter is equal to zero, as the PMOS and NMOS are never on simultaneously in steady-state operation. However, there is a leakage current flowing through the reverse-biased diode junctions of the transistor, located between the source and drain of the substrate. This contribution is, in general, very small and can be ignored.

An emerging source of leakage current is the sub-threshold current of the transistor, in Figure-1.1(b) this is shown as **sub-threshold** from PMOS to NMOS. The sub-threshold current increases as the threshold voltage approaches towards zero volt and the larger the static power consumption [86].

The emerging DSM (deep sub micron) technology has brought the concern for static power dissipation. The static power dissipation is now becoming a major concern for devices at 0.1 micron or below [58].

Though during test the leakage power is not a major concern the knowledge about the leakage power dissipation has to be considered while designing a design-for-test (DFT) circuitry. Nevertheless, there are few efforts been made to consider the leakage during test of VLSI circuitry[60].

1.2.3 Metrics to Measure Power

Energy: *Definition:* The total switching activities generated during the period of test application. Energy affects the battery lifetime during power up or periodic self-test of battery-operated devices[39].

Average power: *Definition:* Average power is the total dissipation of power over a given time period. The ratio of energy to test time gives the average power. Elevated average power increases the thermal load that must be vented away from the device under test to prevent structural damage (hot spots) to the silicon, bonding wires, or package[83, 86].

Instantaneous power: *Definition:* Instantaneous power is the amount of power consumed at any given instant. Usually, it is defined as the power consumed right after the application of a synchronizing clock signal. Elevated instantaneous power might overload the power distribution systems of the silicon or package, causing brown-out¹[83, 86].

Peak power: *Definition:* The highest power value at any given instant. Peak power determines the components thermal and electrical limits and system packaging requirements. If peak power exceeds a certain limit, designers can no longer guarantee that the entire circuit will function correctly. In fact, the time window for defining peak power is related to the chips thermal capacity, and forcing this window to one clock period is sometimes just a simplifying assumption. For example, consider a circuit that has a peak power consumption during only one cycle but consumes power within the chips thermal capacity for all other cycles. In this case, the circuit is not damaged, because the energy consumed will not be enough to elevate the temperature over the chips thermal capacity limit (unless the peak power consumption is far higher than normal). To damage the chip, high (not only highest) power consumption must last for several cycles [39, 81]. The peak power during test is a serious concern to avoid false test. Due to IR-drop effect the timing test fall in risk of false test[101].

Weighted Transition Metric (WTM): The weighted transition metric is an abstraction level power model to approximate the power dissipation with scan chain activity. It is a common practice to approximate the circuit power dissipation with the scan chain activity in test. It has been established experimentally that the switching activity that takes place in each of the flip-flop is linearly correlated with the whole circuit activities[102].

In this thesis the main focus has been on to minimize peak and average dynamic power dissipation during test and to estimate the worst case peak and average dynamic

¹A phenomena that arises due to excessive drop in power supply voltage, this leads to unexpected resetting or rebooting of the device

power dissipation. We have used the WTM method to compute the peak and average switching activity. The WTM method is time efficient and it gives an value which is proportional to the actual circuit power[102].

1.3 Testing of VLSI Circuit

Following the Moore's law, the number of transistors in an IC has doubled every 18 months. In the 1980s, the term "VLSI" was used for chips having more than 100,000 transistor and that has continued to be used over time to refer to chip with millions of transistors. In 1986, the first megabits random access memory(RAM) contained more than 1 million transistors was fabricated. Microprocessors produced in 1994 contained more than 3 million transistors. VLSI devices with many millions of transistor are commonly used in the modern computers and electronic devices. This is a direct result of the down scaling of transistor size, referred as **feature size**. The reduction in feature size has also resulted in increased operating frequencies; for example, in 1971, the first microprocessors runs a clock frequency of 108 KHz, while current commercially available microprocessors commonly run at several gigahertz, currently highest being the 5 GHz in IBM Power PC.

The reduction in feature size increases the probability that a manufacturing defect in the IC will result in a faulty chip. A very small defect can easily result in a faulty transistor or interconnecting wire when the feature size is less than 100 nm. Furthermore, it takes only one faulty transistor or wire to make the entire chip fail to function properly or at the required operating frequency. Yet, defects created during the manufacturing process are unavoidable, and, as a result, some number of ICs is expected to be faulty; therefore, testing is required to guarantee fault-free products, regardless of whether the product is VLSI device or an electronic system composed of many VLSI devices. It is also necessary to test components at various stages during the manufacturing process. There is general agreement with the **rule of ten**, which says that the cost of detecting a faulty IC increases by an order of magnitude as we move through each stage of manufacturing,

from device level to board level to system level and finally to system operation in the field.

Electronics testing includes IC testing, PCB testing, and system testing at the various manufacturing stages and, in some cases, during system operation. Testing is used not only to find the fault free devices, PCBs, and system but also to improve production yield at the various stages of manufacturing by analyzing the cause of defects when faults are encountered. In some systems, periodic testing is performed to ensure fault-free system operation and to initiate repair procedures when faults are detected. Hence, VLSI testing is important to designers, product engineers, test engineers, managers, manufacturers, and end-users.

1.3.1 VLSI Design and Test Flow

Testing of a VLSI circuit typically consist of applying a set of test vector at the input of the circuit while the response at the circuit output are analyzed. Circuit that gives the wrong response is considered to be the faulty chip where as the circuit with correct response is considered as the fault free circuit. The Figure 1.2 illustrates the basic test procedure. The input test patterns are generated either by automatic test pattern generator (ATPG) or by random process. The response at output are compared with the predetermined response, if the response are correct for all the test then the circuit under test is declared as fault-free circuit otherwise the circuit is faulty.

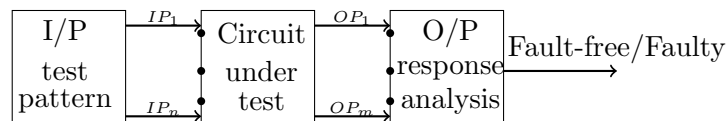


Figure 1.2: Basic test procedure

Testing plays an important role in VLSI design flow. The design has to be tested and verified after each design phase, hence the testing is performed at each phase of design flow. VLSI design flow consist of mainly following phases: 1. Design specification 2. Design 3. Fabrication 4. Packaging 5. Quality assurance. In Figure 1.3 the overall

design flow is illustrate with corresponding test requirement. The figure show that the testing in some form is needed in each phase of design flow. The first phase of design is to specify the design requirement. Once the specification is documented the next phase is to design and synthesize the circuit to generate net-list for further processing. At this stage of design the functionality of the circuit has to be verified (modern formal verification technique are used in general) against the given design specification.

Once the design is verified with zero design bugs the next phase is to perform the automatic test pattern generation (ATPG) for test. Once the pattern generation is completed the net-list will go for fabrication. In between these two steps are placement and routing. Once the fabrication is completed the ICs are tested to determine the defective devices. The chip that passes the wafer test is then extracted and packaged. To eliminate some of the chip with defective package a testing is performed again. Finally the chip which passes both wafer test and package test are tested for quality assurance. This final testing includes the important parameter like input/output timing, voltage, and current. In addition to above described testing, chip is also tested in high temperature which is known as burn-in test which accelerated the test by stressing the device.

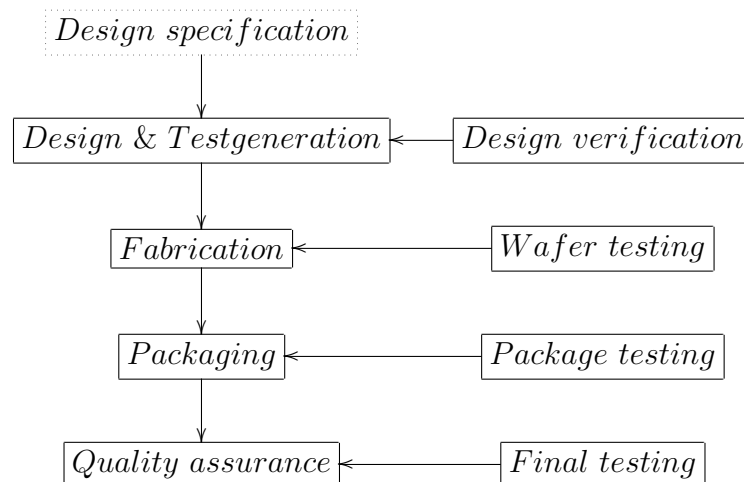


Figure 1.3: VLSI design flow [119]

1.3.2 Scan Architecture and Scan Testing

Scan based testing is one of the widely used design-for-test (DFT) technology. The motivation behind the scan design based approach is to improve the testability of a circuit by improving the controllability and observability. In scan design the sequential elements, the flip-flops, are converted in to scan flip-flop and these are connected as serial shift register to form a scan chain as shown in Figure-1.4. Now, the test pattern for the given sequential circuit can be generated using combinational automatic test pattern generation (ATPG). The standard serial scan design operates in two modes[119]:

Normal functional mode	Test mode
Functional operation	Shift operation
Response capture	Launch operation

In normal mode, all the test related signals are disabled and the circuit is operated in functional mode. In shift mode, scan test vector are shifted in to the scan chain and the test response are shifted out simultaneously. In launch mode, the shifted in test vector are applied to the combinational block. The response are captured while circuit is in normal functional mode.

Serial Scan Architecture:

Among the many varieties of scan architectures, the *full serial scan design* (loosely known as simply "full scan design"), *partial serial scan design* (or just partial scan design), and *random access scan design (RAS)* are the basic scan designs.

In full-scan architecture all the sequential cells are used as scan cells. The full-scan architecture can be implemented using *multiplexed scan D flip-flop (MSDFF)*, *clocked-scan cells*, *LSSD (Level sensitive scan design) scan cells* and the recently proposed mixed-mode scan flip-flop (detail is given in Chapter-6). Full-scan design provides the advantage of achieving highest possible fault coverage for a given ATPG algorithm compared to any other variation of its. The main disadvantage of full scan design is longer test application

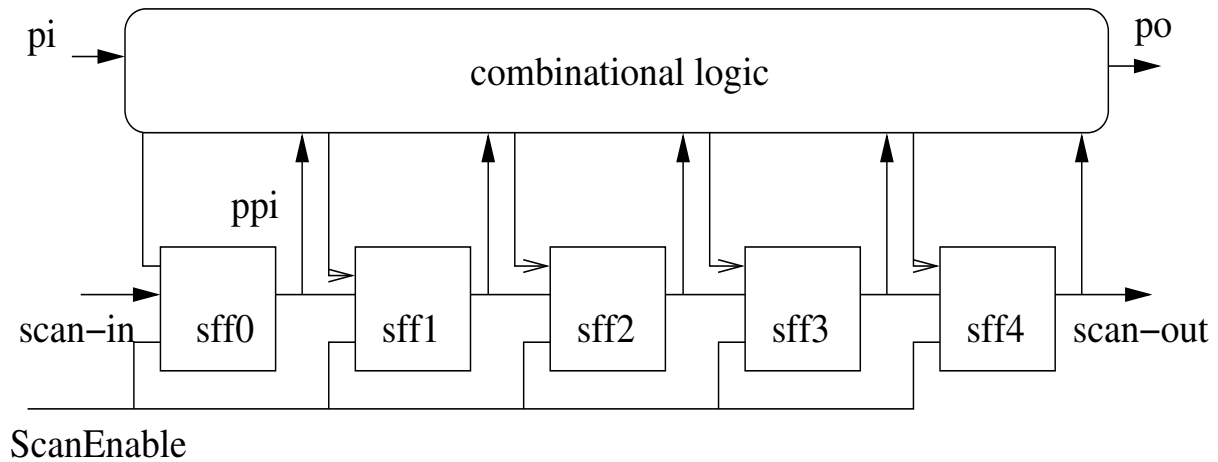


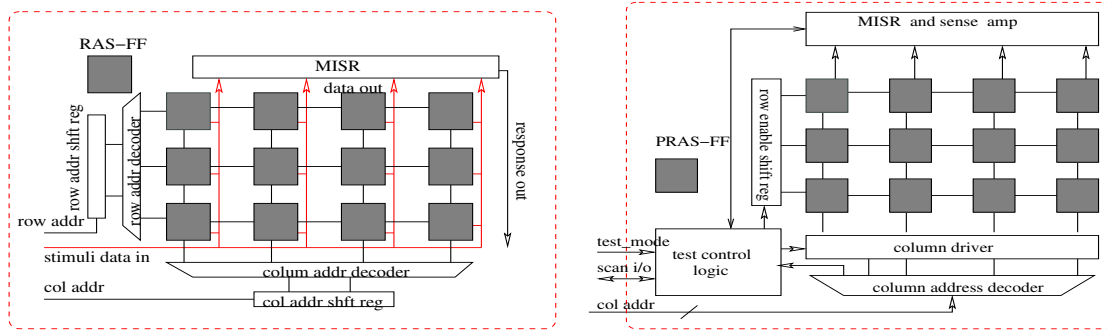
Figure 1.4: Block diagram of basic serial scan architecture

time and excessive toggling in scan cells as it requires serial shift operation. Some of the alternative design, some what better in test application time and power dissipation, are partial scan design and random access scan (RAS).

Unlike full-scan design where all the sequential cells are used as scan cells, partial-scan design requires only a subset of the sequential cells to be used as scan cells in scan chain. It can also be implemented using *multiplexed D flip-flop*, *clocked-scan cells*, and *LSSD (Level sensitive scan design) scan cells*. The test pattern for partial-scan design can either be generated by combinational ATPG or by sequential ATPG.

Random Access Scan Architecture:

Full-scan and partial-scan design are based on serial shift operation. The serial shift operation creates a lot of switching in scan cells which causes the excessive power dissipation. To avoid excessive power dissipation an alternative scan architecture was proposed by Ando in 1980 called random access scan (RAS). The basic idea in RAS is to access only one scan flip-flop at a time without disturbing or switching other scan flip-flops. Thus, RAS reduces the power drastically. The architecture of RAS is similar to the RAM (random access memory), to write a test data into a scan flip-flop, first the flip-flop has to be enabled by row and column address then the data has to be written. Similarly for



(a) Basic random access scan architecture (RAS)

(b) Progressive random access scan architecture (PRAS)

Figure 1.5: Random access scan architecture

reading the response. Two most commonly referred RAS architectures, the simplest random access scan and progressive random access scan, are shown in Figure-1.5. Chapter-5 elaborate the detail functions of RAS architecture and its advantages and disadvantages.

Scan Testing(Fault Models and Test Procedures)

Stuck-at fault Testing: The stuck-at fault testing is performed to test the stuck-at faults. This type of testing can be categorized into two types: internal testing and external testing. In case of internal testing the test related operation is performed using BIST (Built-In Self-Test) whereas in case of external testing ATPG and ATE (Automatic Test Equipment) are used.

The stuck-at testing for scan design (circuit with scan chain) consist of four basic phases: *Test generation* carried out by ATPG, *test application*, *response capture* and *response comparison*, all are performed by ATE. The test application step involves the shifting in and launching of test pattern. During shift operation the SE (Scan Enable) signal is kept high so that the circuit can operate at scan mode. The shift operation is performed at slow clock rate generally half the functional clock. This slow shifting is performed to guarantee the average power consumption below the specified limit. Once

the test patterns are shifted in, the primary input patterns are applied and the SE signal is disabled so that the circuit can now operate in normal mode. In the normal mode operation the responses from the combinational circuit are captured in scan flip flop. The capture operation of stuck-at fault testing is performed at slow clock rate. Once the responses are captured they are shifted out while shifting in another test pattern. The shifted out responses are compared with the previously generated golden responses by ATE for circuit to be declared as faulty or fault free.

Skewed-load Delay Testing: Unlike the stuck-at fault testing, the delay testing requires a test vector pair $\langle V_1, V_2 \rangle$ to be applied to test a set of delay faults. The first vector V_1 of the pair is used to set the circuit to an initial state and the second vector V_2 is used to launch the transition. The second vector V_2 of the delay test pair is one bit shift over the first vector V_1 in the pair. The capture cycle for this test application is now divided into two pulse: launch and capture pulse. In launch pulse the transitions are launched by the application of V_2 . The SE signal is now disabled and the capture operation is performed at functional clock speed [98]. Once the responses are captured they are shifted out while shifting in new test pattern. Finally the shifted out responses are compared with the golden responses. This method is also referred to as Launch-on-Shift (LOS) test.

One of the difficulty in skewed-load scan methodology is timing critical consideration of SE signal [68]. The power related problem that arises during the at-speed testing is discussed in Section 3.

Broad-side Delay Testing: A broad-side delay testing is a form of a scan-based delay testing, where the first vector V_1 of the pair $\langle V_1, V_2 \rangle$ is applied to the scan chain to initialize the chain and the test vector V_2 is applied to launch the transitions. The vector V_2 of the pair is the combinational circuits response to the first vector V_1 . This form of delay testing is called as broad-side since the second vector of the delay test pair is provided in a broad-side fashion, namely through the logic [99]. The broad-side test is

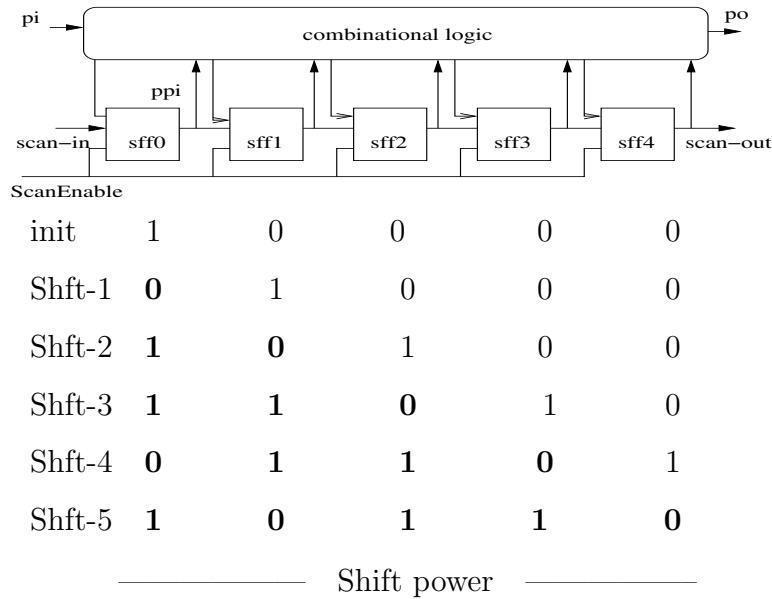


Figure 1.6: The power dissipation phenomena during shift cycle of scan test, Shft- i indicates the i^{th} shift cycle. 1 - 0 or 0 - 1 changes in i^{th} and $i + 1^{st}$ cycle in corresponding sff indicates the switching.

advantageous over the the skewed-load test because in this case the signal SE is no more a timing critical due to an additional dead cycle which allow SE to settle [68]. This method of applying the test is also referred to as Launch-on-Capture (LOC) test.

1.3.3 Shift and Capture Power

The scan testing involves shifts of test stimuli, a launch of stimuli, and a capture of response as discussed in Section 1.3.2. The shift operation shift-in and shift-out the test patterns and responses simultaneously. While shifting in and out the test patterns and responses the scan cells may changes their data in each shift clock cycles. The continuous changes of data in scan cells result in a large number of switching in combinational part of the design. This excessive switching causes the proportionate amount of dynamic power dissipation. The power dissipated during the scan operation can be called as “scan power”. The scan power can be classified into two classes based on the scan period: 1) *shift power* and 2) *capture power*. Further, the shift power can be classified in 1) the

scan chain →	sff0	sff1	sff2	sff3	sff4
second-last shift	0	1	1	0	1
last shift	1	0	1	1	0
Launch cycle (the last shift) power dissipation (four toggles) in stuck-at test					
V ₁	1	0	1	1	0
V ₂	0	1	0	1	1
V ₂ -launch cycle (the last shift) power dissipation (four toggles) in LOS test					

Figure 1.7: Power dissipation during last shift cycle for stuck-at and LOS test

scan chain →	sff0	sff1	sff2	sff3	sff4
current stimuli	1	0	1	1	0
captured response	1	1	1	0	0

Figure 1.8: Capture power dissipation (two toggles), Hamming distance between stimuli and response

power during only shift cycles and 2) power during the last shift launch cycle also known as the *test power*. The basic power dissipation phenomena is demonstrated in Figure-1.6 and 1.7.

Shift Power: The power dissipated during the shift operation is know as *shift power*. In general, the shift power includes the power dissipated over all the shift cycles including the last shift cycle. In literatures the shift power is studied in terms of peak and average power [39]. The peak shift power is computed as the maximum power dissipated over all the shift cycles for entire set of test vectors (both, stimuli and response) [36, 111]. The average power is the total energy consumed for all the vectors over the total shift periods [97]. There are two concerns for peak power during test. One is the thermal issue which could burn-out the circuit under test and the other is IR-drop induced delay problem.

In this thesis our primary objective is to provide a solution for the peak and average power problems in both the time periods: *during the shift* as well as *during the launch*.

The respective chapters, Chapter-3, Chapter-5 and Chapter-6 provides further details on the peak and average power problems.

Capture Power: The power dissipated during the capture cycle is known as *capture power*. For stuck-at fault testing and skewed-load (also known as launch-on-shift LOS) delay testing, there can be only one capture cycle which is followed by the last shift operation. Whereas for broad-side (which is also known as launch-on-capture LOC) testing there could be more than two capture cycles where the test vector pair $\langle V_1, V_2 \rangle$ could be generated in the consecutive cycles, the number of capture could be as large as 10 to 50 cycles.

The capture power dissipated due to application of a test vector in case of stuck-at and skewed-load testing is a Hamming distance between the test vector and its corresponding response. The capture power for broad-side test can be categorized into two: power dissipated during launch pulse and power dissipated during capture pulse. The power dissipated during launch pulse is due to the application of second test vector V_2 of test vector pair to launch the transition. The power dissipated during the capture pulse is due to the difference between test vector and its own response which is to be captured immediate next cycle at functional frequency.

The capture power has not been addressed in this thesis. The thesis concentrates on the peak and average power minimization during shift and launch cycle only.

1.4 Power Estimation Problem

The VLSI design flow need to know in advance about the worst power that a circuit may consume while operated functionally. The advance information on worst power consumption helps in determining supply voltage (V_{dd}), designing optimal power distribution network, deciding on packaging, and circuit optimization. The worst case power number also decides an upper bound on worst case power dissipation which can be used to define a power safe zone during testing. The Chapter-7 elaborates further detail on this topic.

1.5 Thesis Contributions and Organization

The primary goal of this thesis is to provide an innovative solutions for scan test power issues and worst case power estimation problem in SoC designs. The thesis has three major contributions: 1) scan chain reordering methodology for shift peak power and test time minimization, 2) a new scan design-for-test (DFT) architecture called *JScan* for simultaneous minimization of test power (peak as well as average), test time, and test data volume, and 3) integer linear programming methodology for worst case dynamic power estimation. Additionally, the thesis contributed a new scan flip-flop design for the *Jscan* architecture which also serve as replacement for standard multiplexer based scan flip-flop.

The scan chain reordering methodology contributed two graph theoretic formulations and corresponding algorithms to compute the optimum scan chain order for peak power and test time reduction. Effectiveness of the proposed methodologies are experimentally validated and detail discussion of results are carried. The *JScan*- a new joint-scan DFT architecture provides an alternative way of thinking for innovative scan architectures. It has minimized the peak and average test power along with test time and test data volume. The proposed scan flip-flop has eliminated the delay problem which is due to the conventional multiplexer based scan flip-flop; it is one step forward towards improving design performance. The last part of the thesis contributed a binary integer linear programming (BILP) approach for worst case design power estimation. The BILP approach is a different way of looking at the worst case power estimation problem which now can be solved using the powerful ILP solver for getting closer to accuracy.

Rest of the chapters in thesis are organised in following way. The Chapter-2 discusses the earlier work proposed in literatures till date. The Chapter-3 presents scan chain reordering methodology for peak power minimization. In similar line the Chapter-4 presents another scan chain reordering methodology for test time as well as peak and average power minimization. Next, the Chapter-5 presents two new DFT architectures: JScan and 2M-JScan. Exploring further the implementation issues in JScan, the

Chapter-6 proposes a new scan flip-flop design and gives detail functionality while operating as normal serial scan flip-flop and multi-mode Joint-scan flip-flop. The worst case power estimation methodology is presented in Chapter-7. Finally the thesis concludes in Chapter-8 with salient contributions of this thesis and possible problems for future work. The Appendix-A gives information on benchmark circuits, tools, and machines used for experiment. The references are listed in pages thereafter.

Chapter 2

Previous Work

2.1 Scope of the Literature Review

The importance of peak and average power reduction during test is highlighted in the Introduction chapter. The broader cause of excessive power dissipation is the scaling down of device size, the effect of Moore's law. The test power is largely influenced by the changes in design parameters like operating frequency, device density, and complex functionality [27, 57, 132]. Test time and test data volume which are also ramping up according to the Moore's law causes higher test power dissipation[53].

The excessive peak power during scan testing causes unnecessary yield loss and burn out of chips. The phenomena of higher peak power causes high rate of current in the power and ground rails, which decreases the supply voltage and causes ground bounce. This phenomenon is known as IR-drop. The larger IR-drop means the worse performance of the circuit. This degradation in performance leads to timing failure when the circuit is operated at high frequency, which is the case during at-speed test. This degradation in performance leads to incorrect capture of responses during the test process and this results in undesired yield loss [101]. Moreover, this excessive peak power during shift operation also puts limit in scan shift speed [131]. It is discussed by P. Girard [39] that the peak power also determines the thermal and electrical limits of circuit-under-test (CUT) [81]. If the peak power dissipation remains high for consecutively a sufficient

number of cycles then it may cause the thermal related issues. Hence to avoid all these problems it is required to investigate deeper into the possible solutions.

The problem of peak power consumption is also equally important in case of the modern multi-module system on chip (SoC) test. The multi-module SoCs are generally tested in parallel to reduce the over all test time. The parallel testing of SoCs are constraint by two parameters: the area and the power. Hence in this case also the reduction in excessive peak power can results into reduction in test application time.

It is also necessary to understand the upper bound on worst case power dissipation for a particular device to be tested. The worst case power number can be used to compute the safe test power limit for a device under test. Apart from that, the worst case power estimation is necessary for designing of an optimal/robust power distribution network, deciding V_{dd} value, and packaging.

This chapter provides a comprehensive review of literatures broadly in the domain of test power minimization and power estimation. We will provide a classification of the earlier works on peak and average power minimization. For each work we will discuss briefly the main idea, the results and compatibility with that of our proposed methodology viz. scan chain reordering, Joint-scan architecture, flip-flop design, and worst case power estimation. At the end of the chapter we will provide a motivation for our proposed work.

2.2 Minimization of Average Power

Several solutions have been proposed for average power reduction in various direction such as, scan masking [38], scan partitioning [78, 100, 122], circuit modification [59, 107], ATPG [22, 97], vector reordering [32, 41], and scan cells reordering [17, 40]. Along with the above mentioned techniques, the average power can also be controlled during the shift operation by shifting the scan patterns with low frequency test clock. However, this low frequency test clock affects the test application time. The following sub-sections explain the methodologies proposed in above mentioned work.

2.2.1 Circuit Modification Techniques

The work on scan masking approach to reduce the power consumption during scan shift operation is proposed by Gerstendorfer et al. [38]. This approach could able to reduce the scan shift power (both the average and peak) significantly. However, the scan masking methodology has its own limitation, it will not be able to control the power dissipated during *capture cycle* and it increases the delay in functional path because it inserts the masking circuit to check the toggling effect.

The approach proposed by Whetsel [122], Nicola et al. [78] and Saxena et al. [100] have explored the scan partitioning methodology for average power reduction. The scan partitioning technique partitions the scan chain into two or more segments. Each of the segment are operated individually while shifting in the scan vector. Since only a portion of the scan chain is active at a time the power dissipated during the shift operation is inverse proportionate to the number of partition. This technique however, require extra area for implementing the clock distribution circuit.

The methodology proposed in [107] modifies the scan cells by inserting some additional logic to control the transition that arises during the shift operation. This methodology does not degrade the circuit performance because the additional logic is not added in the functional path rather it is added in the scan path. This methodology is capable of reducing both the average and the peak power. The methodology proposed in [59] divides the circuit into multiple partition and test each of the partition independently. The independent testing of individual partition is the key in reducing the average power. The division of the circuit is performed at RTL level.

2.2.2 ATPG Based and Test Vector Modification Techniques

A pattern generation and DFT based technique is proposed by Butler et al. [22] to deal with excessive power dissipation during scan testing. The proposed methodology reduces both the average as well as peak power. In pattern generation technique, the X-filling heuristic known as “fill adjacent” is used to fill the unspecified X-bits in the order of

scan chain from scan-out to scan-in. And in DFT based techniques, the design partition methodology is used to reduce the switching in untested module.

Another work by Li et al. [64] explored the same idea of X-filling to reduce the average power and capture power. The idea in this proposal was that, the minimum possible number of Xs can be used to minimize the peak power only upto the required threshold limit while rest of the available Xs can be used to reduce the average power. The static test vector compaction technique is explored by Sankaralingam et al. [97]. This work is also based on the X-filling methodology.

Another work by V. Devanathan et al. [34] proposed an ATPG based low power methodology. They have proposed a new pattern generation methodology which consists of the technique of statistically profiling the scan patterns and rejecting the scan patterns those does not satisfy the power safety criterion.

2.2.3 Test Vector Ordering Techniques

Reordering of test vector is another technique that is suitable to reduce the power. The advantages of this approach is: it does not add any additional area and it can be done without interfering the design and test flow. P Girard et al. [41, 42] have formulated a graph theoretic problem for test vector reordering to reduce average power.

Dabholkar et al. [32] has achieved average power reduction using test vector and scan latch reordering methodology. The problem of test vector reordering and scan latch reordering was formulated as graph problem and solved by greedy ATSP heuristics, simulated annealing, and Christofied algorithm.

2.2.4 Scan Chain Ordering Techniques

Scan chain reordering is one of the efficient technique for reducing power; both, the average and the peak power. However, this technique comes up with power and routing area tradeoff. The reordering of scan chain also impacts on transition fault coverage, the impact can be either increase or decrease in fault coverage. Following are some of the

proposals based on scan chain reordering.

The proposal by Girard et al. [40] and Bonhomme et al. [17] formulated the scan reordering problem as a graph theoretic problem. The problem of finding the best ordering of scan chain is combinatorial problem which is NP-complete. As the problem is NP-complete they have solved it using a greedy based heuristic algorithm. Furthermore in [40] a methodology to fix the trade-off between power and routing area is proposed.

2.3 Minimization of Peak Power

Most of the proposals on average power also achieved reduction in peak power as a by-product. However, the emergence of deep sub-micron technology has brought a deeper concern for peak power during shift and capture cycle of scan operation. Particularly in case of at-speed testing the peak power becomes an important issue. Till date various direction have been explored for peak power reduction. Some of the earlier proposals are explained in following subsections.

2.3.1 Circuit Modification Techniques

Sankaralingam et al. [96] have proposed a technique based on circuit modification to minimize peak power during shift and capture cycle. The technique is based on the insertion of test point at the out put of scan flip-flops. The function of the test point is to block the transitions from propagating to the combinational block. The main idea of this work is to selectively choose the few numbers of flip-flops on which the test point can be inserted. The proposed methodology is designed to avoid inserting the test point in critical path. Further, the interesting contribution of this method is that it can also control the peak power during capture cycle. However our proposed methodologies (the test vector ordering, the scan chain ordering, and the test scheduling) can be complement to this kind of methodology because the test point that is inserted in the output of scan flip flops is independent of test patterns and scan chain architecture.

Rosinger et al. [88] have proposed a scan partition technique to deal with the peak

power during shift cycle and capture cycle. The idea behind this proposed methodology is similar to the original idea that was proposed by Whetsel et al. [122] and latter extended by J. Saxena et al. [100]. The main idea is to partition the scan chain in to different segments and operate each of the segment in non-overlapping fashion. In the earlier works of L. Whetsel and Saxena et al. the scan chain was split into length-balanced segments and enabling only one in each shift clock. Both of these earlier works performed well in reducing the average power and peak power during shift operation. However, in both of these approaches the capture clock is applied simultaneously to every scan segments which limits the technique from reducing the capture power. Based on this observation Rosinger et al. proposed an idea to overcome the limitation of simultaneous capture operations. The idea proposed by Rosinger et al. is to split the scan chain in to length-balanced segments, and only one segment is enabled during shift as well as capture cycles. The idea is realized by exploring the clock gating technique. This technique is shown effective in reducing both the power dissipated during shift operation and the capture power. Moreover, this technique does not depend on ATPG algorithm. It can reuse the patterns generated by ATPG without altering fault coverage. The minor drawbacks of this methodology is test time overhead and area overhead.

Few points to be observed on the relationship between this kind of methodology and the scan chain reordering, which we have proposed in this thesis, are: 1) the scan chain reordering can be applied to this segmented multiple scan chains, 2) these two methodology does not have conflicting requirements, 3) using both these technique together will results in better power minimization.

2.3.2 ATPG Based and Test Vector Modification Techniques

Corno et al. [30] have proposed an ATPG technique for sequential circuit to deal with peak power. The technique proposed is based on the identification of primary inputs which can be set to don't care bit without affecting the fault coverage and specifying these inputs to minimize the power consumption. Based on this idea, they have proposed a two steps, test sequence relaxation and test sequence reconstruction, methodology. In

the first step a set of primary inputs are identified which can be set to don't care bit and in the second step the identified unspecified inputs are assigned to appropriate bit. Although the authors proposed this technique for non-scan circuits, this can also be applied to circuit with scan dft. However, this kind of approach does not limit the test vector reordering and scan chain reordering technique. Both of these technique can still be applied without putting any side effect on this kind of techniques. Similarly the test scheduling also can be applied along with this kind of techniques.

A post ATPG methodology to reduce peak power during shift and capture operation is proposed by Sankaralingam et al. [36]. In this work, four type of power violation problems are identified; Scan capture, Order dependent, Scan-in, and Scan-out. Each of the problem are then solved by proposed bit-stripping, X-filling and test vector swapping heuristic. The scan capture power violation can occur if the peak power dissipated during the capture operation is higher than the threshold limit. The scan capture power as it is defined by author is approximated by hamming distance between the applied test vector and its response. The order dependent power problem is due to the particular ordering of test vectors. The scan-in power violation occurs during the scan in of test patterns. The scan-out power violation occurs during the scan out of test responses. To eliminate the power violation they have used three techniques. The bit-stripping technique is used to identify the particular bits in test vector which can be set to X value without affecting the fault coverage. Once the X bits are identified then a technique called as minimum transition fill [8] is applied to fill the Xs which minimizes the total number of transitions. This technique is applied to eliminate the scan-in, scan-out, and scan capture power violation. The order dependent power violation is eliminated by vector swapping method.

We have to note here that our proposed scan chain reordering methodology is well compatible with the test pattern modification based approach as similar to the work by Sankaralingam et al[36]. However, the Joint-scan architecture, presented in Chapter-5, has its own requirement for pattern modification because it is a new concept than purely serial scan chain.

Badereddine et al. [7, 8] have proposed a scan chain stitching and X-filling heuristic known as adjacent fill also known as MT-fill (Minimum Transition-fill) for peak power reduction. In this section we will discuss about the scan X-filling technique and the scan stitching is discussed in the Subsection Methodology for Scan Chain Ordering. The X-filling technique proposed here to reduce the peak power is one of the effective methodology to minimize the power. Here, the idea is to use a test generation process during which non-random filling is used to assign values to don't care bits (Xs) of each test pattern of the deterministic test sequence. For example, it is possible to apply the following non-random filling heuristics:

- Don't care '0': all don't care bits in a test pattern are set to '0'
- Don't care '1': all don't care bits in a test pattern are set to '1'
- Adjacent filling: all don't care bits in a test pattern are set to the value of the last encountered care bit (working from left to right). When applying adjacent filling, the most recent care bit value is used to replace each X value. When a new care bit is encountered, its value is used for the adjacent Xs

For example, consider a single test pattern that looks like the following: 0XXX1XX0XX0XX.

If we apply each one of the three non-random filling heuristics, the resulting pattern will be:

- 0000100000000 with '0' filling
- 0111111011011 with '1' filling
- 0000111000000 with adjacent filling.

In the direction of X-filling there are good amount of works has been done by Wen et al. [37, 126]. In [37] they have proposed a low capture power X-Filling (LCP X-Filling) technique. This technique mainly targets the capture power that is dissipated during the capture operation. They have classified the X-cases into four different categories. The Case 1 is when no Xs are present in pseudo primary input (PPI) and in pseudo primary output (PPO) and atleast one X is present in primary input (PI). The Case 2 is when PPI has atleast one X and PPO has no X. The Case 3 is PPI has no Xs and PPO has

atleast one X. The Case 4 is when both PPI and PPO has atleast one X. For each of the X-cases they have proposed corresponding X-Filling algorithm. Similarly in [126] they have proposed a double capture X-Filling (DC X-Filling) algorithm to reduce the peak power during capture cycle for launch-off capture scheme.

The proposed scan chain reordering methodology can be applied to either specified pattern set or to unspecified pattern set. The experiment we have performed for our proposed scan reordering methodology is based on specified pattern set. However, the second methodology presented in Chapter-4 to minimize the number of shift operation uses unspecified pattern set, in this methodology the X-fill is carried out post scan reordering. In our Joint-scan architecture the P-serial is free to use any of the X-fill methodology to further minimize the power.

2.3.3 Test Vector Ordering Techniques

A proposal by Girard et al. [42] was primarily formulated for average power reduction, however this could also able to reduce the peak power as by-product. A vector swapping approach is proposed by Sankaralingam et al. [36] (this technique is discussed in detail in Section 2.3.2) to minimize peak power during scan shift operation. Although the test vector ordering technique proposed by Girard et al. is able to reduce the average power, the peak power reduction is not much effective. From their experimental results it is observed that for some of the benchmark circuit the result of peak power reduction is negative. Other proposed technique discussed in Section 2.2.3 are also having similar kind of effect on peak power reduction.

Tudu et al.[111] have given a graph-theoretic formulation for test vector reordering to minimize shift cycle peak power. The results reported is show the effectiveness of test vector reordering. The advantage of test vector reordering is it can be applied on top of any of the architecture based methodology. Hence, test vector ordering always further improves the reduction.

2.3.4 Scan Chain Ordering Techniques

The scan chain reordering is another domain to explore for peak power minimization. However this domain was left unexplored till now, very few proposals have been made in peak power reduction during *test-cycle*. Badereddine et al. have explored this domain in [7, 9]. They have formulated the scan reordering problem as a constraint global optimization problem. The objective of formulation is to reduce the peak power during *test-cycle*. To solve the problem they have proposed a heuristic based on simulated annealing approach. Although the simulated annealing can provides near optimal solution - if it is allowed to run for sufficient number of iteration - the graph theoretic formulation will wider the solution space for scan reordering problem. The scan reordering methodology may alter the delay fault coverage for skewed-load testing, however it is reported in [9] that the change is very nominal.

Wu et al. [125] proposed a graph theoretic formulation of scan cells reordering for average power reduction. In this work, the Xs are preserved during the time of scan cells reordering for further optimization using MT-fill. This work also reduces peak power consumption along with average power as by-product.

Recent work on scan chain reordering methodology to minimize the peak power during test cycle (the cycle when the test stimuli is being launched) is proposed by Tudu et al. [112]. However this methodology does not considers the power dissipation during entire shift cycles. We have extended this idea and proposed a new formulation for scan shift power minimization using scan chain reordering, the proposed work is presented in Chapter-3. In similar line of scan chain reordering we have proposed yet another methodology to reduce scan shift time, Chapter-4 elaborates the proposed methodology.

2.3.5 Test Scheduling Approaches

Chou et al. proposed the usage of a single power value for each block of logic [28]. In the test scheduling process, the objective is to schedule the testing of blocks such that the total test application time is minimized while the power consumption at any time is below

a given power constraint. Zorian proposed for BIST cores also a scheduling technique under the peak power model [132]. However, the single power model is pessimistic and makes the schedules unnecessarily longer in test time. Rosinger et al. therefore have proposed a double-value test power model (one value representing a constant low power consumption, and the other value representing a constant high power consumption for a test) [89]. The power model was made further accurate by the cycle-accurate power model proposed by Sami et al. [94][95]. Instead of one or two values per test, a power value per cycle is used in order to find tighter test schedules with lower test application time. Samii et al. also have proposed a test architecture and test scheduling technique that defined the test access mechanism (TAM) and assigned tests to TAMs such that the total test time is minimized.

2.4 Random Access Scan Architecture (RAS)

Random Access Scan (RAS) architecture has been explored as a possible alternative to serial scan architecture for power aware testing. First it was proposed by Ando in 1980[5]. Wagner[116] and Ito[52] have implemented the RAS to evaluate its feasibility in 80s. Because of its excessive hardware overhead the RAS architecture was not given much attention that time.

Recently, Baik et al. [12] [10][11] have proposed a set of techniques on RAS to improve routing congestion, area overhead, and pattern loading time. A pattern reordering methodology is proposed in [12] to minimize the test time in RAS while shifting the column address bits. The routing wire length problem is further addressed by Mudlapura et al[75]. The proposed method eliminates the Scan-in and Scan-enable lines. However the technique introduced an additional gate delay in the clock path. To eliminate the clock gating, a modified T flip-flop based scan cell is proposed by Adiga et al.[2].

A cluster based grouping of scan cells to minimize the routing congestion is proposed by Hu et al.[51]. A word oriented random access scan, called as WOR-BIST, is reported by Yao et al.[29] to reduce area overhead in BIST environment.

Test time and pattern volume problems in random access scan architecture are addressed by Baik et al.[10] [13] and Adiga et al[1]. Adiga et al. used an additional data buffer for buffering test data (test stimuli) while shifting the address, therefore the test time as well as data volume are reduced. Baik et al. proposed a progressive random access scan (PRAS) which uses a row enable shift register to address each row one at a time and parallel reading of response and hence minimized the test time.

The random access architecture is highly power efficient. Almost all the work reported results show around 70% to 90% of reduction in peak and average power. Chapter-5 discusses further details on random access architecture and its use in the proposed *Joint-scan* architecture.

2.5 Worst Case Power Estimation

Accuracy and computational efficiency are the two motivations for research on power estimation methodology. Two directions has been explored: *Input vector based simulation* and *Vector less probabilistic method*[80].

A vector based approach proposed by S Devadas et al.[33], formulated as *weighted max-satisfiability(SAT)* problem, is one of the early attempts. The circuit was represented as *max-satisfiability(max-SAT)* problem and then was solved using exact and approximation SAT solver.

The idea of formulating the problem as *Pseudo-Boolean SAT* is explored recently in [69, 92, 106]. H Mangassarian et al.[69] and Sagahyoon et al. [91] have formulated the problem as Pseudo-SAT problem to estimate worst case power and worst case power-up current. The reported results show that the Pseudo-SAT based technique is efficient when the problem is solved using commercial SAT solver like CPLEX[31].

Some what a different idea, brought from ATPG(Automatic Test Pattern Generation) technique, is proposed by Chuan- Yu Wang et al. [117] [118] in 1996. Two methodologies were proposed in [117] for determining lower and upper bounds on maximum power dissipation. To calculate the lower bound, the authors have proposed an ATPG based

technique; while for upper bound they have proposed a *monte-carlo* based simulation technique. An improved version, based on D-Algorithm, of this technique is proposed in [118]. Genetic Algorithm based approach is examined by Hsiao et al[49]. In that work the estimation has been done for three different power matrices: instantaneous peak power, n-cycle peak power, and for sustainable peak power. The accuracy in this work has been accounted by employing four different type of delay models: zero delay model, unit delay model, variable delay model-1 (considers the fan-out with equal weights), and variable delay model-2 (considers the fan-out with variable driving gate size). Najeeb et al[54] and Shyamala et al[105] have proposed a controllability driven peak power estimation method in similar to the ATPG based methodology.

M Pedram and Q Wu et al. have explored probabilistic based ideas [82][123]. Q Wu et al.[123] have proposed a technique based on limiting distribution of extreme order statistics. The challenge in this approach is to show with higher confidence level that the estimated power is indeed the real power. As stated earlier such kind of techniques are time efficient. However difficulty is there in accuracy. More detailed survey of the earlier techniques are carried out by Farid N. Najm[76] and M Pedram[80].

Chapter-7 gives detail discussion on the pros and cons of some the earlier proposed methods mentioned in this chapter.

2.6 Motivation

The scan chain reordering methodology for average power minimization turns out to be an effective methodology. However, we did not find an efficient methodology for peak power minimization. Moreover the scan chain reordering for peak power minimization does not have any formal method. Therefore, we have explored the graph-theoretic formulation for said idea. Chapter-3 provides detail on this. Similarly in the same line, we have given in Chapter-4 a graph-theoretic methodology for reducing scan shift operation to minimize test time and scan activity.

Review of the advantages and the disadvantages of serial and random access scan

lead us to designing of a *Joint-scan* architecture. The Joint-scan architecture is a new approach to build a frame work where both, the serial and random scan, architecture can be implemented to derive the best from each of the scan architecture. The implementation challenge of *Joint-scan* lead us to designing of a new scan flip-flop with entirely new concept than the conventional scan flip-flop. Chapter-5 and 6 provides the elaboration on Joint-scan architecture and scan flip-flop design respectively.

The worst case power estimation is a challenging and a necessary problem for an efficient test and design of a chip. So far, the literatures have looked at this problem from either automatic test pattern generation (ATPG) perspective or pseudo SAT based approach. We gave a binary integer linear programming formulation for the worst dynamic power estimation which can than be solved using the available ILP solver. The proposed formulation initiates a new direction for further exploration to worst power estimation problem. Chapter-7 gives detail on worst case power estimation.

Chapter 3

Peak Power Minimization in Serial Scan

Scan circuit testing generally causes excessive switching activity compared to normal circuit operation. This excessive switching activity causes high peak and average power consumption. Higher peak power causes, supply voltage droop and excessive heat dissipation. This chapter explores idea of scan cell reordering methodology to minimize the peak power consumption during scan shift (which includes the last shift) operation. The proposed methodology first formulate the problem as graph theoretic problem then solve it by a linear time heuristic.

3.1 Introduction

A scan chain is a set of sequential elements, which during test mode are connected to each other in a serial fashion. In scan testing, the test patterns are loaded into the scan chain by serial shifting. During the time of shifting, a number of transitions occur in the scan chain. These transitions propagate and create excessive toggling in the combinational logic, which lead to high dynamic power consumption. The toggling activity that takes place during scan operation is in general much higher than the toggling during functional operation [132]. The dynamic power dissipated during scan shift is studied in terms of

peak and average power [39].

As most modern designs in the current era of deep sub-micron are functionally complex and operate at high speed, the power consumption becomes an important issue to address particularly during testing. Excessive average power can cause problems such as instant circuit damage, higher product cost, performance degradation and reduced battery life [39]; and excess peak power can result in IR drop and cross talk problems, leading to a good chip being falsely classified as defective. The IR drop and cross talk problems are of increasing concern for at-speed testing. During at-speed launch and capture operation, the excessive peak power causes high rate of current in the power and ground rails and hence leads to excessive noise at power and ground. This excessive noise can change the state of logic unexpectedly; hence, this may cause a good die to fail the test, which leads to yield loss [9]. Saxena et al. [101] discuss the IR drop and cross talk problems in detail.

Reduction of peak power during test becomes a necessary for three important reasons:

1. the high peak power causes yield loss due to power droop and cross talk,
2. to avoid chip being damaged during testing,
3. for an SoC with multiple module, a parallel testing can be performed to reduce the test time.

The main contributions of this chapter are two:

- The first one is to provide a graph theoretic problem formulation for scan cells reordering that takes power consumption during load/unload and shift & launch cycles into account.
- It also provides an approximation algorithm to solve the formulated problem. Here, it is to be noted that the graph problem is NP-complete[115].

The experimental results show that the peak power is significantly reduced compared to the default order provided by the industrial tool and the scan chain order computed using the earlier work by Bonhomme et al.[17].

Remaining sections of the Chapters are organised as follows. Section-3.2 gives an overview on previous works. In Section-3.3, a brief background to peak power problem and basic of scan reordering is presented. Section-3.4 and 3.5 presents the problem formulation and algorithm respectively. In Section-3.6, time and space complexity for the proposed algorithm is analyzed. Section-3.7 presents the experimental results and this Chapter concludes in Section-3.8.

3.2 Previous Work

The problem of test power reduction has been an active area of research for quite some-time. The most straight forward approach to reduce the dynamic power is to do test at reduced clock speed. However this solution is no longer a practical as modern designs are timing critical and sensitive to test application time. More over the peak power is independent of clock frequency, hence peak power needs different account.

Most of the previous approaches are aimed at average power reduction, however they have also achieved some reduction in peak power as a by-product [28, 32, 40, 97, 125]. The approach by Chou et al. [28] proposed the scheduling of test under power constraint, address power at module level, and not at circuit level. The work by Dabholkar et al. [32] has achieved average power reduction by test vector and scan latch reordering using some of the TSP heuristics. The work in the area of pattern compaction technique to control the average as well as peak power is done by Sankaralingam et al. [97]. The idea is to carefully merge the test patterns such that the power consumption in the resultant patterns is minimized. However, the scan reordering can still be applied on power aware merged patterns to further minimize the power consumption. Another work on average power reduction is done by Girard et al. [40] and Bonhomme et al. [17]. The basic idea behind these works is to reorder the scan cells. In both of these proposals, the problems are formulated as a graph theoretic problems. As the problems are NP-complete the solutions are heuristic based. Furthermore the proposal by Girard [40] also provides a power-area trade-off. Wu et al. [125] proposed a graph theoretic formulation of scan

cells reordering for average power reduction. In this work, the Xs are preserved during the time of scan cells reordering for further optimization using MT-fill. This work also reduces peak power consumption along with average power as by-product.

Butler et al. [22] and Li et al.[64] proposed X-filling technique, specifying the don't care bits in the test patterns, to minimize average power. Although the X-filling heuristic for average power also able to minimize peak power as by-product, the scan reordering can still be applied for further reduction.

As discussed in the previous section, peak power reduction is becoming increasingly important. A few works have been proposed in peak power reduction [7–9, 30, 36]. Sankaralingam et al. [36] first identify the peak power violation patterns, and then perform simulation to identify the Xs in those patterns. This kind of methodology generally involves the simulation process which is time consuming for large circuits. More over the scan reordering can still be used along with this kind of X-fill approach. The work by Badereddine et al. [7] provides two possible approaches: scan chain stitching and X-filling technique to deal with the peak power. Badereddine et al. [8] evaluate the peak power reduction using MT-fill. Corno et al. [30] also address the peak power problem. In this work X-bits are identified by simulation and specified for peak power reduction. A recent work reported by Tudu et al. [111] uses scan vector reordering to minimize peak power. However our approach of scan chain reordering can be applied along with such techniques. Hence further reduction can be achieved.

The work done by Badereddine et al. [9] is based on the scan reordering to minimize peak power during test cycle. In the work, the problem is formulated as a simulated annealing problem. As their approach is focused on test cycle, it considers only test patterns and not the responses which limits their approach to reduce peak power during load/unload cycles. However, they have also achieved nominal reduction in peak power during load/unload cycles as a by-product.

Given the above observation, in this Chapter we are proposing a methodology to explore the concept of scan cells reordering to minimize peak power during load/unload

cycle and test cycle. The proposed methodology does not affect the stuck-at fault coverage and test time. However it may incur nominal routing area overhead due to scan path modification and may affect the transition fault coverage due to launch-on-shift (LOS) test.

3.3 Background

3.3.1 Overview of Peak Power During Scan Operation

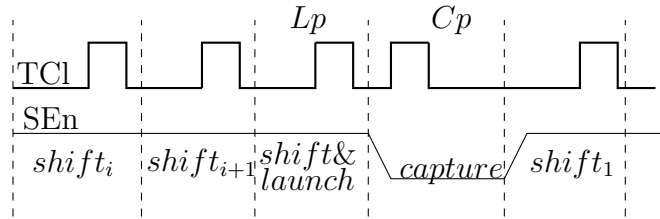


Figure 3.1: Timing diagram for at-speed scan testing

In the above figure, *TCl* is test clock, *SEn* is scan enable, *Lp* is launch pulse, *Cp* is capture pulse and *shift_i* is the i^{th} shift cycle.

The power consumed during entire scan operation at various clock cycle are, *shift power* during the shift cycles (which also include the launch pulse *Lp*) and *capture power* during the capture pulse *Cp*. The shift power at each cycle is known as instantaneous power and the maximum among all the instantaneous power (for all vectors) is known as peak power. The peak power during shift cycle causes excessive heat dissipation and the peak power during *Lp* causes the IR drop problem which is especially problematic for at-speed testing [39], [101]. The proposed methodology addresses both the problems: peak power during shift cycle and launch cycle.

3.3.2 Basis of Power Reduction by Scan Cell Reordering

The scan cell reordering mechanism used for peak and average power minimization basically rely on reduction of *intra pattern transitions* (the bit difference within a pattern). By looking at the differing bits in the patterns, each scan cell can be reconnected in a suitable order to reduce the *intra pattern transitions*. The reordering of scan cells brings all 1s close to each other as well as all 0s during the scan shift operation. Figure 3.2 demonstrates the basic principle of scan cell reordering. In this figure T is test and R is response.

	Initial ordering of scan chain		Reordered scan chain
	$SF_1 \rightarrow SF_2 \rightarrow SF_3 \rightarrow SF_4$		$SF_1 \rightarrow SF_3 \rightarrow SF_4 \rightarrow SF_2$
T	0 1 0 1		0 0 1 1
R	1 0 1 1		1 1 1 0

Figure 3.2: Basic principle of scan reordering

3.4 Problem Formulation

In this section a graph theoretic problem is formulated by first constructing a graph from given scan related information and then a problem resembling to scan cells ordering is defined on that graph. The following subsections explains the complete procedure.

3.4.1 Construction of Graph

The important informations that are needed to be reflected in the graph are scan cells, possible scan connection and peak power information for any pair of consecutive scan cells. Below is the procedure for graph construction:-

- For every scan cell SF_i there is a corresponding node N_i in the graph.

- The scan path between any two possible scan cells is represented by an undirected edge between respective nodes. *Note:* The edges are undirected because, the pairing of scan cells is symmetric with respect to the number of transition. For example, in Figure 3.2 a revers ordering of the reordered scan chain will also results into one transition.
- The instantaneous power consumed by any test vector and response vector due to any possible pair of scan cells is represented as a pair of weight named as *weight-pair* of the edge between respective nodes of the scan cells. The first element of the *weight-pair* is the weight corresponds to test vector and the second element corresponds to the response vector. Each element in the *weight-pair* are vector quantity, we name these element as *testvector-weight* (TVW) and *responsevector-weight* (RVW). The TVW and RVW are computed as follows:

Computation of TVW and RVW: Each TVW and RVW keeps the instantaneous power information for all tests and responses respectively. Instantaneous power information for an individual pattern (test or response) is computed by finding the bit difference between the original position of scan cell pair.

Table 3.2 shows the TVWs and RVWs for the tests and responses shown in Table 3.1 respectively. The rows in Table 3.1 are tests T_i and their corresponding response R_i and columns are scan flip flops SF_i . In Table 3.2 the scan cell pairs are given in column labeled with “scan cell pair” and corresponding TVW and RVW are given in columns t_1 to t_3 and r_1 to r_3 respectively. The columns t_1 to t_3 and r_1 to r_3 shows the peak power information for each test T_i and response R_i . The rectangular box in Table 3.1 shows the bit difference due to scan pair SF_1 and SF_2 and the corresponding transition information in Table 3.2 is shown inside a square box in t_1 minor column. Figure 3.3 shows the complete graph constructed from the data given in Table 3.1 & Table 3.2.

In the process of computation of *weight-pair*, we have considered tests as well as responses. Taking both tests and responses into account is necessary to capture the peak power information as tests and responses both are shifted during load/unload operation.

Table 3.1: Scan patterns and Original scan order

\longrightarrow	SF_1	\longrightarrow	SF_2	\longrightarrow	SF_3	\longrightarrow	SF_4	\longrightarrow
T_1	1		0		1		0	
T_2	0		1		0		1	
T_3	1		0		1		0	
R_1	1		0		1		1	
R_2	0		1		0		1	
R_3	1		0		0		0	

3.4.2 Graph Theoretic Problem Formulation

The basic objective of problem formulation is that the formulation should consider the ordering of scan cells and the minimization of peak power.

The instantaneous power consumed by scan patterns during scan shift operation due to scan cell pair $[SF_i, SF_j]$ is represented as the *weight-pair* of the $edge(N_i, N_j)$ in the graph. For example, the instantaneous power consumed by scan pair $[SF_4, SF_2]$ is represented as *weight-pair* $([0\ 0\ 0], [1\ 0\ 0])$ shown in Table 3.2. As the graph is complete graph, it is always possible to find an order of scan cells by constructing a *Hamiltonian path* (which visits each node of the graph exactly once). Each Hamiltonian path in the graph has a corresponding cost, $cost(N_s, N_e)$, of constructing the path $N_s \sim \dots \sim N_e$, where N_s and N_e are start and end nodes in the path. The computation of cost function $cost(N_s, N_e)$ for a Hamiltonian path (let say H_p) is carried out in following manner:

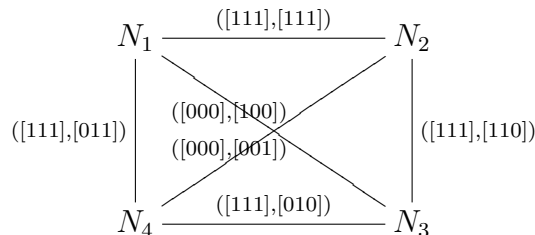


Figure 3.3: A complete vector-weighted graph

Table 3.2: Computed TVW and RVW

scan cell pair	TVW			RVW		
	t_1	t_2	t_3	r_1	r_2	r_3
[SF ₁ , SF ₂]	1	1	1	1	1	1
[SF ₂ , SF ₃]	1	1	1	1	1	0
[SF ₃ , SF ₄]	1	1	1	0	1	0
[SF ₄ , SF ₁]	1	1	1	0	1	1
[SF ₁ , SF ₃]	0	0	0	0	0	1
[SF ₄ , SF ₂]	0	0	0	1	0	0

- Sum up all the TVWs along the path H_p in vector addition manner and find out the maximum value in summation result.
- Similarly sum all the RVWs along path H_p and find out the maximum value in the summation result.
- The maximum of above computed maxima will be the $cost(N_s, N_e)$ for path $N_s \sim \dots \sim N_e$ which resembles the peak power value for corresponding scan order $S_s \sim \dots \sim S_e$.

In a more formal way,

$$PeakPower = cost(N_s, N_e) = Max(T_{max}, R_{max}) \quad (3.1)$$

where $T_{max} = Max\left(\sum_{i=1}^{l-1} TVW(n_i, n_{i+1})\right)$ and

$R_{max} = Max\left(\sum_{i=1}^{l-1} RVW(n_i, n_{i+1})\right)$, where n_i and n_{i+1} are the i^{th} and $(i+1)^{st}$ visited nodes in the path $N_s \sim \dots \sim N_e$, and l is the length of scan chain(or total number of node).

The following example illustrates the above described procedure.

Example 3.1. Let $N_1 \xrightarrow{([111],[011])} N_4 \xrightarrow{([000],[001])} N_2 \xrightarrow{([111],[110])} N_3$, be a Hamiltonian path from Figure 3.3. Then the T_{max} and R_{max} can be

computed as follow:

$$\begin{aligned} T_{max} &= \text{Max}(TVW(n_1, n_2) + TVW(n_2, n_3) + TVW(n_3, n_4)) \\ &= \text{Max}([1 \ 1 \ 1] + [0 \ 0 \ 0] + [1 \ 1 \ 1]) = \text{Max}([2 \ 2 \ 2]) = 2. \end{aligned}$$

$$\begin{aligned} R_{max} &= \text{Max}(RVW(n_1, n_2) + RVW(n_2, n_3) + RVW(n_3, n_4)) \\ &= \text{Max}([0 \ 1 \ 1] + [0 \ 0 \ 1] + [1 \ 1 \ 0]) = \text{Max}([1 \ 2 \ 2]) = 2. \end{aligned}$$

In this example $n_1 = N_1$, $n_2 = N_4$, $n_3 = N_2$, and $n_4 = N_3$. Finally the $\text{cost}(N_s, N_e)$
 $= \text{Max}(T_{max}, R_{max}) = 2$.

Hence from the above illustration the graph theoretic problem can formally be stated in following way:

Problem Statement: Given an undirected graph $G(V, E, W)$, find a Hamiltonian path which has minimum cost over all other possible paths. Where V is set of vertices V_i s, E is set of edges, and W is set of *weight-pair*.

Note: The hardness of problem is NP-complete as this can be reduced to TSP problem which is a known NP-complete problem.

3.5 Proposed Algorithm

As the problem is NP-complete we are proposing a greedy based polynomial time (with respect to $|V|$) algorithm. The complete algorithm has two subpart, Part-1 to find the Hamiltonian cycle and Part-2 to find the path from Hamiltonian cycle. Following subsequent subsection gives the details of algorithm for Part-1 and Part-2.

3.5.1 Construction of Hamiltonian cycle

Problem Statement: Given a graph $G(V, E, W)$ find a Hamiltonian cycle having minimum cost $\text{cost}(N_s, N_e)$, in this case $s = e$.

Algorithm Part-1(with reference to Figure 3.3):

Step 1: Chose an initial node randomly and assign it to *current-node* and initialize corresponding *cumulative-node-cost* as $([0 \ 0 \ 0], [0 \ 0 \ 0])$.

Note: *cumulative-node-cost* is an array of integer values computed dynamically for a node visited currently.

Step 2: In this step, the decision parameter *peak-power* and *average-power* are computed. These are used in step 3 to select the next node to visit. The *peak-power* and *average-power* are computed with respect to the set of nodes which are not yet visited and are neighbors of the current node. We call these node as *candidate-node* and any edge from current node to a *candidate-node* is called as *candidate-edge*. A vector addition operation is performed between the *cumulative-node-cost* and *weight-pair* of each *candidate-edge* separately. We call these sum *probe-path-weight*. The *peak-power* for each of the *candidate-edge* is computed by finding $Max(\text{probe-path-weight along the candidate-edge})$. And *average-power* is computed by finding $Sum(\text{probe-path-weight along the candidate-edge})$.

Step 3: A node N_i from the set of *candidate-node* is selected as the next node if the *peak-power* along the edge(*current-node*, N_i) is minimum among all other *peak-power*. If there are more than one node having equal *peak-power* then the corresponding *average-power* is used as a tie breaker. In case of equal average power, the next node will be chosen randomly. Following example helps understanding the above explanation.

Example 3.2. Considering Figure 3.3, the initial node will be N_1 with *cumulative-node-cost* $([0\ 0\ 0], [0\ 0\ 0])$. The neighboring node of N_1 are N_2 , N_3 , and N_4 which satisfies the property of being *candidate-node* because these are not yet visited. Based on the *peak-power* and *average-power*, a next node will be selected from these *candidate-nodes*. The corresponding *candidate-edges* are edge(N_1 , N_2), edge(N_1 , N_3), and edge(N_1 , N_4). The *weight-pair* of each of these *candidate-edge* are $([1\ 1\ 1], [1\ 1\ 1])$, $([0\ 0\ 0],[0\ 0\ 1])$, and $([1\ 1\ 1], [0\ 1\ 1])$ respectively. *Probe-path-weight* along each of the *candidate-edges* will be vector addition of *cumulative-node-cost* of current node with the *weight-pair* of each of these *candidate-edge* which are,

$$([0\ 0\ 0],[0\ 0\ 0]) + ([1\ 1\ 1], [1\ 1\ 1]) = ([1\ 1\ 1], [1\ 1\ 1]),$$

$$([0\ 0\ 0],[0\ 0\ 0]) + ([0\ 0\ 0], [0\ 0\ 1]) = ([0\ 0\ 0], [0\ 0\ 1]), \text{ and}$$

$$([0\ 0\ 0],[0\ 0\ 0]) + ([1\ 1\ 1], [0\ 1\ 1]) = ([1\ 1\ 1], [0\ 1\ 1])$$

respectively.

The *peak-power* and *average-power* along each *candidate-edge* can now be computed as follows,

peak-power:

$$\text{along edge}(N_1, N_2) = \text{Max}(\text{Max}([1\ 1\ 1]), \text{Max}([1\ 1\ 1])) = 1$$

$$\text{along edge}(N_1, N_3) = \text{Max}(\text{Max}([0\ 0\ 0]), \text{Max}([0\ 0\ 1])) = 1$$

$$\text{along edge}(N_1, N_4) = \text{Max}(\text{Max}([1\ 1\ 1]), \text{Max}([0\ 1\ 1])) = 1$$

average-power:

$$\text{along edge}(N_1, N_2) = \text{Sum}([1\ 1\ 1]) + \text{Sum}([1\ 1\ 1]) = 6$$

$$\text{along edge}(N_1, N_3) = \text{Sum}([0\ 0\ 0]) + \text{Sum}([0\ 0\ 1]) = 1$$

$$\text{along edge}(N_1, N_4) = \text{Sum}([1\ 1\ 1]) + \text{Sum}([0\ 1\ 1]) = 5$$

The next task is to select a next node based on the above computed *peak-power* and *average-power*. In this example *peak-power* along each *candidate edge* is equal *i.e.* 1, hence *average-power* is used as a tie breaker. *average-power* along edge(N_1, N_3 ,) is 1 which is less than the *average-power* of other *candidate edge*. Hence node N_3 will be selected as the next node to be visited.

Step 4: In this step *cumulative-node-cost* and *current-node* are updated. The *cumulative-node-cost* for next node to visit = *cumulative-node-cost* of *current-node* + *weight-pair* of edge(*current-node*, next-node-to-visit). For Example 2, *current-node* will be updated with node N_3 and *cumulative-node-cost* for $N_3 = ([0\ 0\ 0], [0\ 0\ 0]) + ([0\ 0\ 0], [0\ 0\ 1]) = ([0\ 0\ 0], [0\ 0\ 1])$.

Step 5: Repeat the algorithm from Step 2 until a Hamiltonian cycle is constructed.

At the end Algorithm part-1 will produce a Hamiltonian cycle. For Figure 3.3 the algorithm will give the Hamiltonian cycle shown in Figure 3.4 as output.

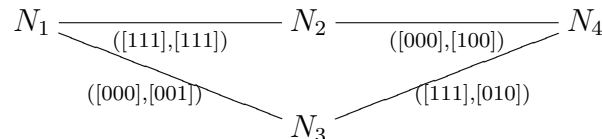


Figure 3.4: Hamiltonian cycle from Algorithm Part-1

3.5.2 Construction of Hamiltonian path

Problem Statement: Given a Hamiltonian cycle $H(V, E, W)$, find out a Hamiltonian path which have minimum cost $cost(N_s, N_e)$.

Algorithm Part-2(*Computation and Selection*):

Computation: In this step the peak and average power consumption will be computed for every possible Hamiltonian path. Peak power is the primary objective of power minimization where as average power is used a tie breaker which is the secondary objective. For a given H having total node $|V|$ there can be $|V|$ number of possible Hamiltonian path. For each path the peak and average power is computed. Computation of peak and average power for a Hamiltonian path H_p are done as follows:

Peak power is computed according to the Equation 1. Average power is computed according to the weighted transition metric method [97] as follows,

$$AveragePower = T_{sum} + R_{sum} \quad (3.2)$$

where $T_{sum} = Sum \left(\sum_{i=1}^{l-1} TVW(n_i, n_{i+1}) * i \right)$.

$R_{sum} = Sum \left(\sum_{i=1}^{l-1} RVW(n_i, n_{i+1}) * (n - i) \right)$, where $l = |V|$ is the total number of node in H, and n_i and n_{i+1} are the i^{th} and $(i + 1)^{st}$ node in H_p .

Example 3.3. Computation of peak and average power for all possible Hamiltonian path in Figure 3.4.

$$Path1: N_1 \xrightarrow{([111],[111])} N_2 \xrightarrow{([000],[100])} N_4 \xrightarrow{([111],[010])} N_3$$

$$PeakPower = Max(T_{max}, R_{max})$$

$$T_{max} = Max([1 1 1] + [0 0 0] + [1 1 1]) = 2$$

$$R_{max} = Max([1 1 1] + [1 0 0] + [0 1 0]) = 2$$

$$PeakPower = Max(2, 2) = 2$$

$$AveragePower = T_{sum} + R_{sum}$$

$$T_{sum} = Sum([1 1 1] + [0 0 0] * 2 + [1 1 1] * 3) = 12$$

$$R_{sum} = Sum([1 1 1] * 3 + [1 0 0] * 2 + [0 1 0]) = 12$$

$$\text{AveragePower} = 12 + 12 = 24$$

$$\text{Path2: } N_2 \xrightarrow{([000],[100])} N_4 \xrightarrow{([111],[010])} N_3 \xrightarrow{([000],[001])} N_1$$

$$\text{PeakPower} = \text{Max}(T_{max}, R_{max})$$

$$T_{max} = \text{Max}([0\ 0\ 0] + [1\ 1\ 1]) + [0\ 0\ 0] = 1$$

$$R_{max} = \text{Max}([1\ 0\ 0] + [0\ 1\ 0] + [0\ 0\ 1]) = 1$$

$$\text{PeakPower} = \text{Max}(1, 1) = 1$$

$$\text{AveragePower} = T_{sum} + R_{sum}$$

$$T_{sum} = \text{Sum}([0\ 0\ 0] + [1\ 1\ 1] * 2 + [0\ 0\ 0] * 3) = 6$$

$$R_{sum} = \text{Sum}([1\ 0\ 0]*3 + [0\ 1\ 0] * 2 + [0\ 0\ 1]) = 6$$

$$\text{AveragePower} = 6 + 6 = 12$$

$$\text{Path3: } N_4 \xrightarrow{([111],[010])} N_3 \xrightarrow{([000],[001])} N_1 \xrightarrow{([111],[111])} N_2$$

$$\text{PeakPower} = 2 \text{ and } \text{AveragePower} = 20$$

$$\text{Path4: } N_3 \xrightarrow{([000],[001])} N_1 \xrightarrow{([111],[111])} N_2 \xrightarrow{([000],[100])} N_4$$

$$\text{PeakPower} = 2 \text{ and } \text{AveragePower} = 16.$$

Selection: In this step the optimal Hamiltonian path will be selected based on the following criteria,

- The path which have lesser peak power than other will be the final Hamiltonian path
- In case of multiple path having equal peak power value, the minimum average power path will the final Hamiltonian path
- In case of failure of above criteria, the Hamiltonian path will be chosen with random choice.

For Example-3.3, the *Path2* will be chosen as the final Hamiltonian path as the peak power of this path is lesser than others. Hence the resultant scan chain which corresponds to Path2 will be $SF_2 \rightarrow SF_4 \rightarrow SF_3 \rightarrow SF_1$.

3.6 Time and Space Complexity

The proposed graph-theoretic problem being a NP-complete, the algorithm proposed here is a heuristic. Following Sections compute the time and space complexity.

3.6.1 Time Complexity

Algorithm Part-1(Hamiltonian cycle): The proposed heuristic has complexity of $O(n * l^2)$ where l is the length of scan chain and n is number of test vector. The proposed heuristic takes additional $2n$ time to compute peak and average power at *Step 2* in comparison to any other heuristic of same nature whose edge weight in graph are just a scalar quantity unlike in our case.

Algorithm Part-2 (Hamiltonian path): Time complexity for this part of algorithm is $O(n * l)$.

3.6.2 Space Complexity

Proposed heuristic does not use explicit space to represent the graph as the weight of edges are computed dynamically. An array of size l is used to keep the visited information for each node. To store the *cumulative-node-cost* requires $2n$ size of array. A temporary space to keep the *weight-pair* needs $2n$ size of array. Hence the overall complexity will be $O(l + n)$.

Hence the proposed heuristic does not suffer from any strain related to time and space.

3.7 Experimental Results

The experiments are carried out on ITC99 and ISCAS89 circuits. The algorithm is implemented in C++. Scan insertion, circuit synthesis, and test generation are performed using industrial tools. Test patterns are low power adjacent X'filled and compacted

Table 3.3: Specification of ITC-99 (b04 to b10) and ISCAS-89 Benchmark Circuits

Benchmark circuit	#Test pattern	Scan chain length	#PI	#PO	#Gates	Fault coverage
b04	49	66	11	8	628	92.85
b07	53	45	1	8	427	99.95
b08	40	21	9	4	171	100
b10	44	17	11	6	180	100
s420	53	16	18	1	218	99.32
s5378	107	179	35	49	2779	99.87
s9234	123	211	36	39	5597	99.93
s13207	115	638	62	152	7951	99.97
s15850	96	534	77	150	9772	100
s35932	13	1728	35	320	16065	100
s38417	69	1636	28	106	22179	100
s35854	111	1426	38	304	19523	100

stuck-at fault pattern. Specification of benchmark circuits and experimental results are provided in Table 3.3 and Table 3.4 respectively.

To show the effectiveness of the proposed work we have compared the experimental results with the scan order provided by industrial tool and with the average power scan order proposed by [17]. Table 3.4 show that the proposed work is able to reduce the peak power up to 48.29% in comparison to the industrial solution and 28.51% in comparison to [17] for circuit s13207. For circuit s35932 the proposed work and the work by [17] are able to reduce the power significantly due to the favourable combination of smaller number of pattern set and large scan chain. Also from Table 3.4 it can be observed that the proposed solution has achieved appreciable reduction for larger benchmark circuit in comparison to [17].

Table 3.4: Experimental results for peak power minimization

Benchmark circuit	#Peak trans. Ind. soln.	#Peak trans. [17]	%of reduct. by [17] wrt Ind. soln.	#Peak trans. Proposed soln.	%of reduct. wrt Ind. soln.	%of reductn. wrt [17]
b04	33	26	21.21	22	33.33	15.28
b07	27	19	29.63	17	37.04	10.53
b08	16	12	25	11	31.25	8.33
b10	12	11	8.33	10	16.67	9.09
s420	11	8	27.27	7	36.36	12.5
s5378	84	87	-3.57	68	20.24	21.82
s9234	118	99	16.1	89	24.58	11.11
s13207	176	148	15.9	91	48.29	28.51
s15850	108	117	-8.33	97	10.18	17.09
s35932	90	22	75.56	17	81.11	22.73
s38417	470	493	-4.9	364	22.55	26.16
s38584	440	421	4.32	353	19.77	16.15
Avg. Data	132.08	121.92	7.91	95.5	31.79	21.67

3.8 Conclusion

The chapter has proposed a graph theoretic approach for scan cell reordering to minimize peak power during scan shift operation. The results show that proposed methodology is capable of minimizing peak power in compared to industrial solution and [17]. The proposed work keeps stuck-at fault coverage and test time unaffected and also it does not incur any explicit area overhead. The scan reordering may affect the transition fault coverage and may incur small area overhead due to routing. In this work we have not taken these two parameter into account. However, the problem of routing can also be solved by integrating our work with [40], where the proposed methodology can be applied within the cluster to take care of power and cluster ordering can be used to minimize routing congestion. The transition fault coverage requires further work to co-optimize it with power and routing area.

Chapter 4

Test Time Minimization in Serial Scan

Scan test time has always been one of the priority issues for test researchers because it directly impact cost of the design. In this work we have addressed the issue through scan chain and test pattern reordering. The idea of limited scan shift is explored. We have proposed a graph theoretical framework for reordering of scan chain and test pattern. Graph theoretic problem is formulated for each, scan chain and test pattern, reordering. For each of the formulated problems corresponding approximation algorithms are proposed.

4.1 Introduction

Sequential scan DFT undoubtedly is a most widely used test methodology because of its simplicity in architecture and methodology. A scan architecture is designed by connecting the flip-flops in a back to back fashion to form a serial path called *scan-chain*. For clean observability and controllability it is often preferred to have full scan design over partial scan design[15]. Though serial scan DFT provides ease of testing several other problems are getting aggravated primarily due to increase in design size and technology scaling[93]. As the design size is moving towards billion of gates the test complexity is also increasing

proportionately[53][56].

As the length of scan chain increase the number of clock cycles required to load/unload the test/response patterns also increases in same proportion. Modern design with a gate count in order of 10^6 will be having the flip-flop count in the order of 10^5 [56]. Having a design with scan chain of length 10^5 is beyond the limit of affordable test time and cost. Each of the test cycle is associated with the test cost imposed by automatic test equipment. Larger is the number of test cycles larger is the test cost, and hence increase in over all product cost. For modern designs the test cost is dominating the overall production cost[53]. Therefor the test time has received special attention.

Several approaches have been proposed to solve test time and cost problems. One of the effective approach is to reduce the scan chain length by breaking it to n number of parallel scan chains. Hence, by this architecture a speed up of n is achievable. However this technique is limited by the available number of I/O pins. To mitigate the pin constraint a compression and output compaction based techniques are proposed[87][72]. Due to exponential growth in design size, and hence the test data volume, the compression and compaction based techniques are unable to meet the demand[16][120]. Hence further reduction by increasing more scan chains in parallel is getting limited.

In this work we have proposed a scan chain and test patterns reordering methodology to minimize the number of shift operations. The scan chain reordering is performed on the basis of test cube. The presence of abundant don't care bits (Xs) in the test cube facilitate the efficient scan chain ordering. Pattern analysis shows that around 80-90% of Xs are available in compacted test cubes[14]. The proposed technique explore the efficient use of Xs in getting the appropriate scan chain and test pattern ordering. We have formulated graph theoretic problems and proposed Kruskal's minimum and maximum spanning tree algorithm for scan chain reordering and TSP-approximation algorithm for pattern ordering. The proposed scan chain and test pattern order avoid the complete shifting of patterns with l shift cycle where l is length of scan chain. It provides an useful shift cycle, $c \leq l$, for each pattern. Additionally, since the number of shift operation are being reduced the proposed methodology also minimizes the average power dissipation.

The primary contributions of this chapter are following:

- graph theoretical formulation for scan chain and test pattern ordering,
- reduction in scan shift time,

Following observations are to be noted with respect to scan chain and test pattern ordering.

- The scan chain ordering technique can also be used in the multiple parallel scan chains to get an appropriate ordering of scan cells in each sub scan chain.
- The pattern compression may likely get benefited from the scan chain ordering[120].
- The test pattern reordering in this case is independent of the chain order.

Rest of the Chapter are organized as follows. Previous work on test time minimization and alternative scan chain architecture are discussed in Section-4.2. The proposed idea on scan chain reordering is presented in Section-4.3. In Section-4.4 we have explained the test pattern reordering methodology. Experimental methodology and direction for result analysis is explained in Section-4.5 The Chapter concludes in Section-4.6.

4.2 Previous Work

The increase in test time is primarily due to two reasons[21][53]:

1. increase in pattern volume,
2. increase in scan chain length.

Several realms have been explored in search for an efficient solution. Multiple parallel scan chains with compression, pattern compaction, partial scan design, broadcast based ideas like scan tree, scan forest, Illinois scan techniques, and limited scan shifting are proposed over the years.

Partial scan designs have been proposed with an objective to reduce the length of scan chain. Several researches have been carried out based on selecting the appropriate

flip-flops which are hard to control and observe for design of partial scan chain[63][108]. An efficient way of selecting scan flip-flops for partial scan chain is also proposed[127].

The test pattern compaction being a software based idea is most efficient and is aimed to reduce the pattern count. Several techniques have been proposed on static test compaction[85][67][84]. Simulation based dynamic test compaction methodology is proposed to reduce the test time[90]. The important observation to make here is the number of Xs in compacted test cube which is still around 70-80%. We have developed the scan chain ordering technique making the best use of Xs.

Another interesting technique for test time reduction is to partition the full scan chain into multiple parallel operated sub scan chains[77]. Since the number of I/O pins for test are limited so the number of parallel operated scan chains are limited[56]. Hence to overcome the problem of I/O pins constraint an input compression[87][120] and output compaction[72] based techniques are employed. Also the idea of scan chain partitioning are combined with parallel scan chains to isolate some of the scan cells which shifts the Xs bits[62]. There are other class of architecture based on pattern broadcasting idea are proposed to reduce test time[79][104][73][25]. The idea of limited scan shift has been proposed in [46, 48], the proposed methodology is proven to be effective in reducing test time.

Considering state of the art techniques in this work we have further explored the scan chain ordering methodology in the direction of limited scan shift to minimize test time. On top of the scan order we have also proposed test pattern reordering to further optimize the test time. Both, the scan and test ordering are formulated as a graph theoretic problem, and corresponding heuristic are proposed to solve the problems.

4.3 Proposed Scan Chain Ordering Technique

Scan chain ordering technique has been explored to solve wide variety of problems manifested out of scan DFT. Test power is minimised[17, 112], routing congestion[40] in scan path is reduced, pattern compression is enhanced[120] and further more, the scan chain

is sequenced to increase the fault coverage in LOS transition delay test. Here in we are proposing a mechanism to systematically arrange the scan chain in sequence to minimize test time by reducing the number of shift operation per pattern. The test pattern are also systematically ordered on top of the reordered scan chain. The mechanism involve a graph theoretic problem formulation, where a graph is being constructed out of scan cells and a maximum spanning tree algorithm is applied to find out a near optimal order of scan chain.

In following subsections the basic idea, graph construction, and MaxST algorithm are presented.

4.3.1 Basic Idea

The basic idea is to use the unspecified bits in test and response in such a way that the shifting of unspecified bits per pattern can be avoided as maximum as possible. If the scan cells can be ordered in a way that for most of the patterns only the specified bits could be loaded then the scan shift operation can be minimized. Hence the over all test time reduction can be achieved. Following example demonstrate the idea.

Table 4.1: Demonstration of basic idea

		Original Scan Chain							Reordered Scan Chain				
		F ₀	F ₁	F ₂	F ₃	F ₄			F ₀	F ₃	F ₁	F ₂	F ₄
T ₀	X	C	X	C	X	T ₀	X	C	C	X	X		
R ₀	X	X	C	X	C	R ₀	X	X	X	C	C		
T ₁	C	C	X	C	X	T ₁	C	C	C	X	X		
R ₁	X	C	X	X	C	R ₁	X	X	C	X	C		
T ₂	C	X	X	C	X	T ₂	C	C	X	X	X		
R ₂	X	C	C	X	C	R ₂	X	X	C	C	C		

Example 4.1. In the Table-4.1 the T_i and R_i denotes the test and response patterns respectively. C and X in the patterns denotes care and don't care bit respectively.

The original scan chain order is $F_0-F_1-F_2-F_3-F_4$ as shown in first major column and the reordered chain is $F_0-F_3-F_1-F_2-F_4$ shown in second major column. By looking at the reordered scan chain following can be observed; the care bits in T_0 are separated from don't care bits in to one side of scan chain. Similarly in R_0 the don't care bits are moved towards the head of scan chain and the care bits are moved towards the tail. Similarly in T_1 and R_1 . Now the test pattern T_0 can be loaded into the scan chain only by three shift cycle because the don't care bits in last two scan cells are not really needed and hence we can gain two cycles saving. Similarly the response R_0 needs only two clock cycles for shift out. Similarly for remaining test patterns respective number of shift cycles can be saved, in Table-4.1 the bits required to be shifted in or out are colored in gray. Hence an intelligent reordering of scan cells can benefit in saving of shift cycles. The idea is basically to explore the presence of don't care bits to obtain the scan chain order.

Hence the idea of scan chain reordering can formally be stated as follows.

Problem Statement: *Reorder the scan cells in such a way that maximum number of Xs in test patterns can be moved towards tail of scan chain and the maximum number of Xs in response can be moved towards head of the scan chain for every patterns.*

The scan chain ordering problem is further formulated as a graph theoretic problem and an algorithm is proposed to compute a shift cycle effective scan chain. We start with graph construction in following section.

4.3.2 Graph Construction

With reference to the demonstration in Table-4.1 following **two rules** has to be noted:

1. The scan cells having test bit C and response bit X (henceforth referred as C-X pair) needs to be moved towards head of scan chain.
2. The scan cells having test bit X and response bit C (henceforth referred as X-C pair) needs to be moved towards tail of scan chain.

The important information that are necessary to be reflected in the graph are scan cells, possible scan paths, and information built on Rule-1 and Rule-2 to make sure the

Table 4.2: C-X and X-C pair indicator bit array

	C-X pair			X-C pair			
	b_{p0}	b_{p1}	b_{p2}	b_{p0}	b_{p1}	b_{p2}	
I_{n0}^c	0	1	1	I_{n0}^x	0	0	0
I_{n1}^c	1	0	0	I_{n1}^x	0	0	1
I_{n2}^c	0	0	0	I_{n2}^x	1	0	1
I_{n3}^c	1	1	1	I_{n3}^x	0	0	0
I_{n4}^c	0	0	0	I_{n4}^x	1	1	1

proximity of scan cells having C-X and X-C pair.

Following **mapping procedure** is followed for construction of the graph.

- For every scan cell F_i there will be corresponding node N_i in the graph.
- A scan path P_{ij} between any two scan cells F_i and F_j is represented as an edge E_{ij} between node N_i and N_j .
- The proximity value on each edge E_{ij} due to C-X and X-C pair is represented as an edge weights W_i^j computed by formula given in Formula-1.

Formula-1: Computation of edge weight W_i^j

Weight W_i^j of an edge $E(N_i, N_j)$ is computed as the difference between $A_{xc}(I_{ni}^x, I_{nj}^x)$ and $A_{cx}(I_{ni}^c, I_{nj}^c)$. Where A_{xc} and A_{cx} are *ANDing* operation.

$A_{xc}(I_{ni}^x, I_{nj}^x) \Rightarrow$ Number of X-C proximity

$A_{cx}(I_{ni}^c, I_{nj}^c)^c \Rightarrow$ Number of C-x proximity

$$W_i^j = A_{xc}(I_{ni}^x, I_{nj}^x) - A_{cx}(I_{ni}^c, I_{nj}^c) \quad (4.1)$$

In the Formula-1 derived above the I_{ni}^c and I_{nj}^x represents the indicator bit array for C-X and X-C pair. Table-4.2 shows the indicator bit array for the pattern set and scan cells shown in Table-4.1. The bit 1 in array I_{ni}^c with position b_{pk} indicates the presence of C-X in the pattern $P_k(T_k, R_k)$ in the i th scan cell and 0 indicates the absence of C-X pair. Similarly in array I_{nj}^x the bit 1 in position b_{pk} indicates the presence of X-C pair in pattern $P_k(T_k, R_k)$ in j th scan cell and bit 0 indicates the absence of it. Following example elaborate the use of I_{ni}^c and I_{nj}^x in computing W_i^j .

Example 4.2. With reference to pattern set in Table-4.1 and the indicator bit array in Table-4.2 let I_{n0}^c , I_{n1}^c , I_{n0}^x , and I_{n1}^x be indicator array for C-X and X-C pair in F_0 and F_1 respectively. Then the *ANDing* can be computed as follows.

$A_{xc}(I_{n0}^x, I_{n3}^x)$	$A_{cx}(I_{n0}^c, I_{n3}^c)$
0 0 0	0 1 1
0 0 0	1 1 1
$0+0+0 = 0$	$0+1+1 = 2$

The corresponding edge weight will have following value,

$$W_0^1 = A_{xc}(I_{n0}^x, I_{n3}^x) - A_{cx}(I_{n0}^c, I_{n3}^c) = 0 - 2 = -2.$$

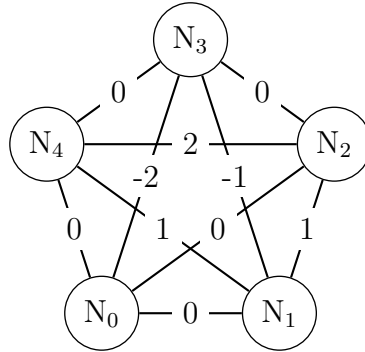


Figure 4.1: A complete weighted graph constructed from Table-4.1

Since any of the *ANDing* operation, A_{cx} or A_{xc} , can take arbitrary integer values the value of W_i^j can be -ve, 0, or +ve. The -ve weight indicates the scan cell pair are more suitable for to be placed towards the head of scan chain. The 0 indicates the placement of scan cell pair to be in the midway of scan chain. +ve weight indicates the scan cell pair to suitable for tail of the scan chain.

By using the equation Eqn.4.1 and the mapping procedure a graph shown in figure Fig.4.1 is constructed out of the pattern given in Table-4.1. The constructed graph is denoted by $G_w(E, V, W)$ where E , V , and W are set of edges, vertices, and edge weights respectively. Next, we will explain the problem formulation on the constructed graph. Henceforth the graph in Fig.4.1 will be used as the reference for scan chain sequence demonstration.

4.3.3 Hamiltonian Path Problem Formulation

Objective is to find a scan chain sequence such that C-X and X-C pair in test patterns should get segregated to two ends of the chain. The C-X pair to be moved towards entrance of chain and X-C pair towards exit of chain. The segregation of C-X and X-C pair will accumulate the consecutive Xs in response at the entrance of chain and Xs in test at the exit of chain.

Hence for the formulation of Hamiltonian path we have defined a cost function $F_c(P_h)$ to ensure the closeness of scan cells based on C-X and X-C pair. In the graph shown

in Fig.4.1 the closeness is taken care in the computation of edge weight W_i^j . The -ve edge weight implies the closeness in C-X pair, the +ve edge weight implies the closeness in X-C pair, and 0 implies the equivalent closeness in both. The complete Hamiltonian path is divided into two sub-paths for optimization. The first sub-path, named as P_h^c , is optimized for C-X pair and the second sub-path, named as P_h^x , is optimized for X-C pair. Each of the sub-paths contain equal number of edges for even number vertices and a difference of one edges for odd number of vertices. For example, considering following path,

$$P_h : N_0 \xrightarrow{0} N_1 \xrightarrow{1} N_2 \xrightarrow{0} N_3 \xrightarrow{0} N_4 \text{ the corresponding sub-paths will be}$$

$$P_h^c : N_0 \xrightarrow{0} N_1 \xrightarrow{1} N_2 \text{ and}$$

$$P_h^x : N_2 \xrightarrow{0} N_3 \xrightarrow{0} N_4$$

Now for each of these sub-paths we will define a cost function. For path P_h^c our objective is to minimize the path weight and for path P_h^x is to maximize the path weight. The cost function for paths P_h^c and P_h^x would be

$$F_c(P_h^c) = \sum_{i,j \text{ node index in } P_h^c} W_i^j \quad (4.2)$$

$$F_c(P_h^x) = \sum_{k,l \text{ node index in } P_h^x} W_k^l \quad (4.3)$$

Using Eqn.4.2 and Eqn.4.3 the cost function, $F_c(P_h)$ for the desired Hamiltonian path, P_h , is defined as follows.

$$F_c(P_h) = F_c(P_h^c) - F_c(P_h^x) \quad (4.4)$$

Our objective in defining the cost function F_c is to optimize it such that the cost $F_c(P_h^c)$ for C-X pair sub-path will be minimized and the cost $F_c(P_h^x)$ for X-C pair sub-path will be maximized. Hence the cost function F_c can be optimized by minimising its value. A formal problem statement to find the desired Hamiltonian path in graph $G_w(V, E, W)$ is stated below.

Problem Statement: Find a Hamiltonian path P_h in a given graph $G_w(V, E, W)$ in such a way that it minimizes the cost function $F_c(P_h)$.

To compute the P_h we have devised an approximation algorithm based on minimum spanning tree algorithm. Following section explains the algorithm and its efficiency.

4.3.4 Kruskal's MaxST and MinST Algorithm

To find out the path P_h in a given graph $G_w(V, E, W)$ we have proposed a modified Kruskal's algorithm. The algorithm in an iterative steps find out the set of edges to form a path which maximizes the cost function $F_c(P_h)$ defined in Eqn-4.4. The path resembles the ordered scan chain having the scan cells with C-X pair closer to head and the scan cells with X-C pair closer to the tail of scan chain.

The algorithm will build two spanning trees, a left spanning tree called *minimum spanning tree(MinST)* and a right spanning tree called *maximum spanning tree(MaxST)* over multiple iterations. The left spanning tree is computed to approximate the path P_h^c which resembles the sub scan chain with maximum C-X pair and the right spanning tree is computed to approximate the path P_h^x which resembles the sub scan chain with maximum X-C pair. Once both the sub-paths are computed they will be joined to form a Hamiltonian path P_h .

The algorithm is shown in Algorithm-1. The coloring of the nodes to white in initial then to red for MinST and to green for MaxST is performed to avoid the interference in edge selection. The MinST can not chose any edges whose adjacent edges are marked with green similarly the MaxST can not chose any edges which are marked with red. The edges are first sorted in non-decreasing order to ensure the formation of minimum and maximum spanning trees. Since the construction of T_{minst} and T_{maxst} is done by alternative edge selection the cardinality of edges in T_{minst} and T_{maxst} will be equal or difference of one edge. Hence the sub-paths formed using depth first tour[115] of T_{minst} and T_{maxst} will have equal or a difference of one edge. The sub-paths are approximation to Eqn-4.2 and Eqn-4.3. In Step-13 the joint operation is performed with an edge having weight $|W_m^n|$ as minimum as possible to ensure that neither C-X nor X-C pair shall reside

Algorithm 1 Computation of MinST and MaxST

- 1: Compute indicator bit array I_{ni}^c and I_{ni}^x for each node N_i .
 - 2: Construct a graph $G_w(E, V, W)$, where $E \leftarrow$ set of edges, $V \leftarrow$ set of vertices, $W \leftarrow$ set of edge weights.
 - 3: Let $T_{minst} = \phi$ and $T_{maxst} = \phi$ be minimum and maximum spanning tree respectively, initially be empty.
 - 4: Color all the nodes in set V to be white.
 - 5: Sort the all the edges in E in nondecreasing order and let Q be sorted edges.
 - 6: Select the next smallest edge e_s from Q and include it in set T_{minst} if and only if
1) it's adjacent vertices are colored *white-white* or *red-white* and 2) it does not form any cycle else select the next smallest edge. Increase the edge pointer *ptr-s* in Q to point next edge.
 - 7: Color the nodes adjacent to e_s with red.
 - 8: Select the next largest edge e_l from Q and include it in set T_{maxst} if and only if 1) it's adjacent vertices are colored with *white-white* or *green-white* and 2) it does not form any cycle else select the next largest edge. Decrease the pointer *ptr-l* in Q to point next edge.
 - 9: Color the nodes adjacent to e_l with green.
 - 10: Repeat the procedure from state:6 until all the nodes in V are colored to red or green.
 - 11: Compute a path P_h^c out of minimum spanning tree T_{minst} using depth fist tour.
 - 12: Compute a path P_h^x out of maximum spanning tree T_{maxst} using depth fist tour.
 - 13: Joint the two sub-paths P_h^c and P_h^x using the edge incident to the last node of P_h^c and the first node of P_h^x having edge weight near to zero. This will form the Hamiltonian path P_h with optimal cost.
 - 14: return P_h
-

in the mid of scan chain.

The time complexity of Algorithm-1 can be bounded by $O(|E| \log |E|)$. The dominating time consumer is sorting operation in Step-5. Other operations like coloring of nodes, MinST in Step-6 and MaxST in Step-8 are linear in terms of $|V|$. The space complexity can be bounded by size of indicator bit array i.e. $O(|T| * |S|)$ where $|T|$ is number of test patterns and $|S|$ is number of scan flip-flops.

On the top of the reordered scan chain a test vector ordering idea is formulated to further optimize the test time. The test pattern reordering is presented in Section-4.4.

4.4 Test Pattern Reordering

The test pattern reordering problem is formulated as a graph theoretic problem. The graph now is a complete directed weighted graph denoted as $G_{wd}(\vec{E}, V, W)$. The direction between any two nodes indicate that the order $T_i \rightarrow T_j$ is different from $T_j \rightarrow T_i$ where T_i and T_j indicates the i_{th} and j_{th} test pattern.

We have first constructed the graph $G_{wd}(\vec{E}, V, W)$ from a given pattern set and a given scan chain order. The necessary information for finding an optimal test pattern order are reflected in the graph. On the graph a travelling salesman problem (TSP) is formulated. The TSP is then solved using minimum spanning tree based approximation algorithm.

Following subsections describe the procedure for graph construction and presents the approximation algorithm.

4.4.1 Graph Construction

Following mapping procedure is followed for graph construction.

- Each test pattern T_i is represented as a vertex V_i in the graph G_{wd} .
- To take account of the initial scan chain content and the shifting of last response the graph includes two dummy nodes D_i and D_o respectively.

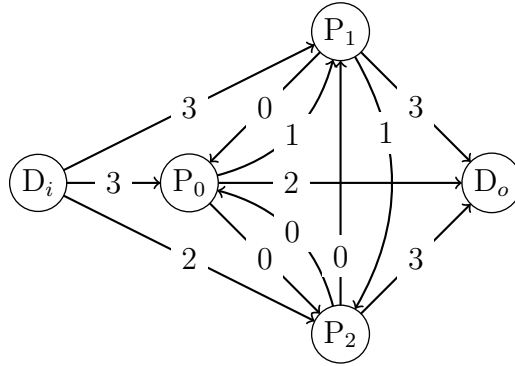


Figure 4.2: A complete directed weighted graph constructed from Table-4.1

- The order of applying the test pattern $T_i \rightarrow T_j$ is represented as an edge $\vec{E}(V_i, V_j)$. Since any test pattern can be applied in followed by remaining pattern the graph will be a complete graph.
- The edge weight W_i^j in the graph represents the number of *penalty-cycles*. The computation penalty cycle is done as follows.

Computation of *penalty-cycles*: The *penalty-cycles* is defined as the number of cycles wasted due to mismatch in the number of *optimum shift cycles*(*OSC*) needed to shift in the test pattern and optimum cycle needed to shift out the response. For a test pattern T_j being applied followed by T_i the penalty cycle is computed as per Eqn.4.5.

$$\text{penalty-cycles} = |\text{OSC}(R_i) - \text{OSC}(T_j)| \quad (4.5)$$

In the Eqn.4.5 $\text{OSC}(\text{pattern})$ is a function to compute *optimal shift cycle*.

For example, from Table-4.1, let test pattern T_2 be applied next to T_1 , then the edge weight W_1^2 of an edge $\vec{E}(V_1, V_2)$ is $|\text{OSC}(R_1) - \text{OSC}(T_2)| = |2 - 2| = 0$. The $\text{OSC}(R_1)$ is 2 because the response R_1 needs only 2 shift out cycles and $\text{OSC}(T_2)$ is 2 because the test T_2 needs only 2 shift in cycles. The graph constructed from the pattern set in Table-4.1 is shown in Fig. 4.2. For further reference we will be using the Fig. 4.2.

4.4.2 TSP and MST Based Heuristic

Algorithm 2 MST based Approximation Algorithm

```

1: procedure MSTAPPR(In:  $G_{wd}(\vec{E}, V, W)$ , Out:  $H_p$  )
2:   procedure FINDMST(In:  $G_{wd}(\vec{E}, V, W)$ , Out:  $T_{mst}$ )
3:     Sort all the edges of  $G_{wd}(\vec{E}, V, W)$  in non-decreasing order.
4:     Delete the next smallest edge and add it to  $T_{mst}$  if it does not form any cycle.
5:     Continue the process until  $|V| - 1$  edges are added to  $T_{mst}$ .
6:     DFS-TOUR(In:  $t_{mst}$ , Out:  $H_p$ )
7:   procedure DFS-TOUR(In:  $T_{mst}$ , Out:  $H_p$ )
8:     Mark the nodes in  $T_{mst}$  as per depth first traversal.
9:     Form a Hamiltonian path  $H_p$  according to the marking of nodes.
10:  Return  $H_p$ 

```

To find out the optimal order of test patterns we have formulated a travelling salesman problem(TSP) on the given graph $G_{wd}(\vec{E}, V, W)$. Since the TSP is known NP-complete problem we have applied a MST based approximation algorithm to TSP[115].

The mapping of pattern ordering problem to TSP is based on the following two consideration.

1. Since all the pattern has to be applied it is needed to visit all the node in $G_{wd}(\vec{E}, V, W)$ including dummy node D_i and D_o .
2. Since the total number of *penalty-cycles* as defined in eqn.4.5 is to be minimized the path chosen to visit all the nodes must satisfy this criteria.

The problem is formally defined as follows.

Problem Definition: Given a complete directed weighted graph $G_{wd}(\vec{E}, V, W)$ with edge set E and vertices V find a Hamiltonian path which will have minimum cost, $C_f(H_p)$. The cost function $C_f(H_p) = \sum(W_i^j)$ where i and j are indices of all the vertices belongs to H_p .

Table 4.3: Experimental Results on ISCAS-89 Ckts

Bench Ckt	SChain Length	#Test Pattern	ShiftCycle ToolOrder	ShiftCycle Proposed	%Reduction
s344	15	17	269	203	24.5
s713	19	29	489	398	18.60
s1423	74	39	2766	2104	23.9
s5378	179	98	17264	15892	7.95
s15850	534	109	52427	45628	12.97
s38417	1636	278	421590	357989	15.08
Average data			494805	422214	14.68

where Average data is computed as an average reduction over total shift cycles across all the bench mark circuit

In a high level pseudo code we have presented the MST based approximation algorithm in Algorithm-2.

4.5 Experimental Result and Discussion

The experiment has been performed on ISCAS-89 circuits. Single scan chain based architecture is considered for the experiment. The ATPG is performed using TetraMax tool. The unspecified pattern set is generated to retain the Xs in pattern cube. The shift cycle is computed for the given tool order of scan chain and test pattern, and for the order computed using proposed algorithm. The computation of shift cycle considers the *optimum shift cycle* for given tool order as well as for the proposed order. In this work we have considered variable shift cycle for each pattern. The experimental results are tabulated in Table-4.3. The results show up to 24% of reduction in shift time. For some of the circuits the default tool order also provided some amount of saving compared to full scan shift.

4.6 Conclusion and Prospects

We have presented a methodology to minimize scan shift cycle. Scan chain and test pattern reordering ideas are explored. Each of the reordering idea is translated into respective graph theoretic problem. For each of the graph a corresponding Hamiltonian path with desired cost function problem is solved using approximation algorithms. From the experimental results we can learn that the algorithms need improvement for further reduction in shift time. The proposed idea can seamlessly be incorporated with the parallel scan architecture which is of current industrial practice.

Over the years the reordering technique have been proposed for test power minimization, scan path routing optimization, and for efficient test compression. We believe that the proposed methodology has now opened a avenue for unified graph-theoretic approach for scan chain reordering where all the parameters of interest can be taken into consideration.

Chapter 5

JScan: A Joint-scan DFT Architecture

This chapter discusses the proposed Joint-scan design-for-test (DFT) architecture. Two architectures are presented: 1) **JScan** and 2) **2M-JScan**. The JScan architecture is the central concept whereas the 2M-JScan is enhancement on JScan with efficient test control mechanism. Both architectures are verified with through experiment on large bench mark circuits. The term Joint-scan is also used as a synonym to JScan unless mentioned specifically.

5.1 Introduction

Serial scan DFT methodology has been very successful for the last couple of decades, primarily because of its simplicity. The architecture is cost effective in terms of hardware area and the number of test pins. However, in recent years it has been observed that the serial scan architecture is no longer “the solution” for test[57][44]. Designs are getting increasingly more complex and dense with every new generation of technology. Due to rapid increase in complexity of designs newer challenges for testing are manifesting. Test power, test data volume, and test time are becoming aggressively high. These parameters also follow the trend of Moore’s law[93]. The increase in test time and test data volume

directly impact the test cost. The excessive test power, particularly the dynamic power, causes test failure and chip damage which indirectly adds to test cost[44].

Several solutions to mitigate the excessive power dissipation have been proposed in recent past[44]. Despite all these inventions the higher power dissipation during test still remain as one of the challenges in serial scan architecture. As an alternative to serial scan the random access scan (RAS) architecture has been proposed by Ando[5], back in 1980. The RAS was considered to be impractical at that time. However, the need for reliable test and uprisng test cost has provoked the test researchers to reconsider the RAS architecture in the current technology scenario. The scan architecture proposed by Ando is based on the idea of organizing the flip-flops as a random access memory, thereby enabling read and write to an individual flip-flop without creating unwanted activities in rest of the flip-flops.

Since the RAS provides a way to access each individual flip-flop independently the activity it creates during pattern loading is negligible. The activity per pattern load is proportional to the number of bits to be written (to be loaded) to the scan flip-flops. Hence the average and peak power consumption is proportional to only one flip-flop toggle[116][52][12]. Therefore power consumed during pattern loading in RAS is negligible. In our proposed work we have explored this ultra low power feature of RAS to reduce the test time and test data volume.

The primary idea in the proposed architecture is to integrate the serial and random access scan architecture as an unified single scan architecture. We segregate the given flip-flops into two groups to form a serial scan and a random access scan. The two separate scan architectures are operated concurrently with a common control unit, the functionality of architecture is elaborated in Section-5.3. The balanced segregation of flip-flops play an important role since it involve multiple parameters - test power, test time, and data volume. Thereby, an effort has been made here to work out a methodology for the new architecture. We call the proposed architecture as **JScan** for Joint-scan. The major contributions of this part of work are following:

- a methodology to segregate the given flip-flops

- controlling mechanism to operate the JScan
- minimization of test time, data volume, and test power.

5.1.1 Chapter Organization

The rest of the Chapter is organized as follows. Earlier works are discussed in Section-5.2. The proposed JScan architecture is presented in Section-5.3. Experimental results for JScan are discussed in Section-5.4. The 2M-JScan architecture is presented in Section-5.5. In Section-5.6 experimental results for 2M-JScan are discussed. We have presented a possible variants of Joint-scan architecture in Section-5.7. The chapter concludes in Section-5.8 with suggestions for future direction.

5.2 Previous Work

As mentioned in the Chapter-2 (Previous Work) and in Section-5.1 (previous section), several solutions have already been proposed in last couple of decades to solve test power, pattern volume, and test time problems. Scan cell gating, scan chain reordering, test vector reordering, low power X-fill, low-power ATPG, primary input control method, and power gating ideas has been explored for low power test[44]. Scan chain partition, partial scan, pattern compression, pattern broadcasting has been examined for test time and data volume reduction[109]. Despite these techniques, a reasonably good solution is still a goal[57][27][16].

Random Access Scan (RAS) architecture has been explored as a possible alternative. Wagner[116] and Ito[52] have implemented the RAS to evaluate its feasibility in 80s. Recently, Baik et al. [12] [10][11] have proposed a set of techniques on RAS to improve routing congestion, area, and pattern loading time. The single address decoder in the earlier architecture is now modified to have two, row and column address, decoders. The scan flip-flops are now connected by the row and column address lines in a grid arrangement. Baik et al. [11] have shown that grid of size $\sqrt{N} \times \sqrt{N}$ is optimal for routing congestion, where N is number of scan cells. A pattern reordering methodology

is proposed in [12] to minimize the test time. The routing wire length problem is further addressed by Mudlapura et al.[75]. The proposed method eliminates the Scan-in and Scan-enable lines. However the technique introduced an additional gate delay in the clock path. To eliminate the clock gating, a modified T flip-flop based scan cell is proposed by Adiga et al.[2].

A cluster based grouping of scan cells to minimize the routing congestion is proposed by Hu et al.[51]. A word oriented random access scan, called as WOR-BIST, is reported by Yao et al.[29] to reduce area overhead in BIST environment. The experimental results show that WOR-BIST needs comparatively less DFT area.

Test time and pattern volume problems in random access scan architecture are addressed by Baik et al.[10] [13] and Adiga et al[1]. Adiga et al. used an additional data buffer for buffering test data while shifting the address, therefore the test time as well as data volume are reduced. Baik et al. proposed a progressive random access scan (PRAS) which uses a row enable shift register to address each row one at a time and parallel reading of response and hence minimized the test time. In the proposed Joint-scan the PRAS is used along side the standard multiple sequential scan.

Other than the power issues, the multiple sequential scan also been challenged by prolonged test time and higher pattern volume. There have been several methodologies to minimize the volume and time. The most effective one is the pattern compression which serve both the purposes[87][104]. However due to growing design complexity the problem of pattern volume and test time seems to be a daunting task. The proposed architecture aim to overcome the limitations confronted by the earlier approaches and deliver a novel DFT methodology.

5.3 Proposed JScan Architecture

The JScan is a Joint-scan architecture, which integrates the two well known DFT architectures, the serial and random access scan, to harness the best benefit from of each architecture. Serial scan architecture is simple and requires minimal hardware, likewise

the random access scan is power efficient. On the other hand the serial scan dissipates undesirable power, and the RAS occupies additional chip area and introduces routing congestion due to wirings. Therefore the *Joint-scan* is designed to integrate these two architectures such that by leveraging certain amount of power (that should necessarily be equal to functionally maximum allowable power) the test time and pattern volume can be further reduced. Maintaining the minimum power during test is also necessary to avoid the test escape. Earlier experiment has shown that 40 - 50 % of test power reduction is sufficient to avoid power related issues and is necessary to maintain the minimum power requirement to avoid test escape[114].

We name the serial scan part of JScan as **partial-serial access scan (P-serial)** and the random scan part as **partial-random access scan (P-random)**. The part of the concept of JScan architecture is reported in [110]. In explaining the architecture we will be using the term Joint-scan and JScan interchangeably.

5.3.1 Architecture

We have chosen to implement the P-serial using multiple serial scan architecture (MSS)[77] and the P-random using progressive random scan architecture (PRAS)[10]. Possible variants of implementation are briefly outlined in Section-5.7. We identify the following three challenges in realizing the architecture:

- to integrate and implement the serial and random scan as one architecture
- devising an efficient controlling mechanism
- to segregate the flip-flops into two groups.

The scan flip-flops are segregated in to two groups (the algorithm to segregate the flip-flops is elaborated in Section-5.3.3). One group forms the MSS and the other forms PRAS architecture. Each flip-flop can either be in any one of the groups. In this architecture multiplexor based scan flip-flop is considered for MSS and the PRAS cell is considered for PRAS architecture.

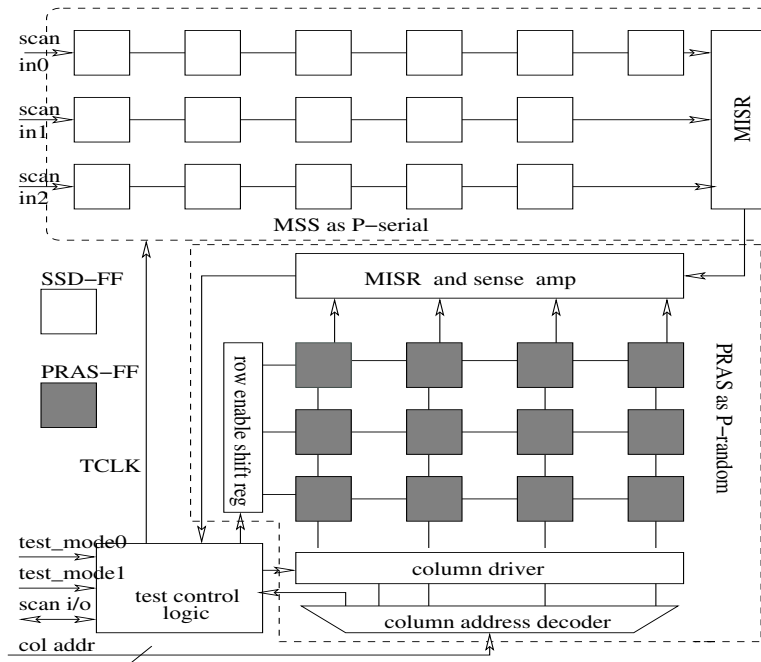


Figure 5.1: The Proposed Joint-scan Architecture (the JScan)

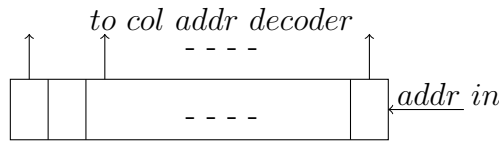


Figure 5.2: Serial address shift register for PRAS using only one address pin

The proposed architecture is shown in Figure 5.7. We have implemented two variants of PRAS architecture, one with parallel address pins for supplying column address and the other with serial shift register with only one address pin. The Figure 5.7 shows the implementation with parallel address pins. The Figure 5.2 shows the serial column address shift register- an alternative to parallel implementation.

The architecture primarily consists of three blocks: P-serial, P-random, and a Test Control Logic (TCL). In Figure 5.7 the P-serial is shown with small white square blocks and the P-random is shown with small grey square blocks. In P-serial chain (implemented as MSS), the *Scan-in0* to *Scan-in2* are the scan in lines used for shifting of test. The MISR connected at the output of MSS functions as response compacter. The *SE* signal

functions as a scan-enable signal for P-serial. In P-random (implemented as PRAS), the *Column address* lines supplies the column addresses, the *Col address decoder* which is connected with the column address lines decodes the addresses and generates address signal. The lines which are connected from *Column address* lines to *TCL* are used for generating some of the control signals. The *Column driver* control/drive the column address lines. In particular to PRAS architecture, the *column driver* drives two signal, bit and bit-bar, lines for writing test data and reading the response. The *Sense amplifier & MISR* block is a circuitry to read and compact the responses. The *MISR* in P-serial and *MISR* in P-random are connected serially for simplicity to shift out the responses via *Scan I/O* signal. The *TCL* circuitry generate control signals to oversee the functionality of Joint-scan. Two external signals *Test_mode0* and *Test_mode1* are connected to *TCL*, using these the *TCL* generate the necessary signals.

5.3.2 Functionality

The Joint-scan architecture operates both the scans, the P-serial as well as the P-random, concurrently. The complete functionality of Joint-scan involves pattern loading/unloading, launching of test stimuli, and response capture. While loading/unloading the patterns three scenario arises: 1) both the scans complete simultaneously, 2) P-random completes ahead of P-serial, and 3) P-serial completes ahead of P-random. These three test scenarios along with the normal circuit functionality are controlled using two test mode pins, *Test_mode0* and *Test_mode1*. The Joint-scan thereby operates in four different modes. The four modes of operation are listed in the Table 5.1 and elaborated here.

During test, the Joint-scan starts with *Joint mode* wherein the P-serial and P-random operate concurrently to load/unload the test/response pattern. On completion of loading/unloading of pattern in either P-serial or P-random the mode is changed either to *P-serial mode* (if the P-random is completed) or to *P-random mode* (if the P-serial is completed). Once the test pattern is loaded fully the mode is changed to *Functional mode* for test launching and subsequently response capture. The controlling mechanism

Table 5.1: Joint-scan Modes of operation

Test_mode0	Test_mode1	Mode of Operation
0	0	Functional mode
0	1	Joint mode
1	0	P-random mode
1	1	P-serial mode

of these scenario is depicted in Figure 5.3 as a state transition machine. In the state transition machine the state represents mode of operation and the transition captures the completion event.

The procedure repeats for loading/unloading of each subsequent test patterns. Detail control signals are explained here.

Functional mode:

The *Functional mode* controls three operations: 1) normal function, 2) launching of test stimuli, and 3) response capture. During functional mode the *SE* signal is disabled to ensure the scan flip-flops in P-serial as well as in P-random to operate in functional mode. On completion of capture operation the circuit would return back to test mode, i.e. Q_j in state transition machine, by enabling *SE* signal.

Joint-scan mode:

In this mode *SE* is enabled and both the scans, P-serial and P-random, starts to load/unload the test/response patterns. The pattern shift operation in P-serial is performed as usual. The reading/writing of response/test data in P-random is controlled by sequence of two steps. In the first step *row enable* signal is generated to enable the row and read response. The reading of response is done via sense amplifier. In second step the desired scan flip-flops, for which the previous response differ from the current test data and the current test data is care bit, are written with the current test data. The flip-flops are written one after another in column by column for which the column

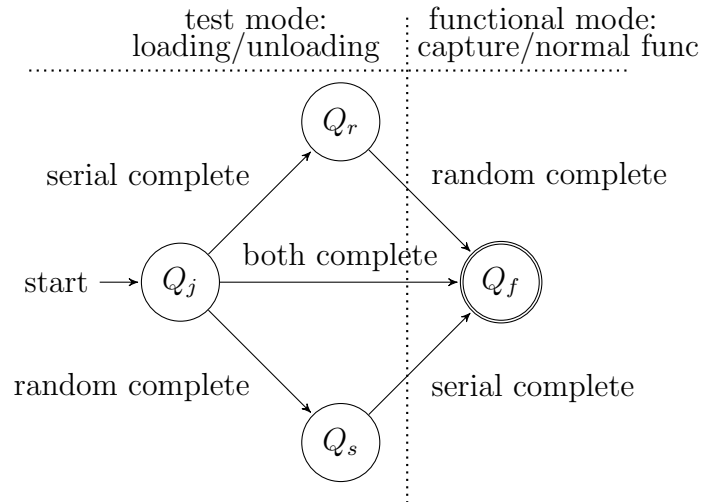


Figure 5.3: State transition machine for mode control: where, Q_j is Joint-scan mode, Q_r is Random mode, Q_s is Serial mode, and Q_f is functional mode.

address is applied through column address pins as shown in Figure 5.7. The same procedure repeats to read/write the cells in subsequent rows and columns. Once any one of the scans completes pattern loading/unloading operation the mode is changed accordingly to either *P-serial mode* or *P-random mode*. For the case of equality when both the scans, P-serial and P-random, complete the loading/unloading in same time the mode is changed to *Functional mode*. The transition of modes take place according to the state machine shown in Figure 5.3.

P-random mode:

While being at *Joint mode* whenever pattern loading in P-serial is completed ahead of P-random the *P-random mode* is activated. The P-serial scan now is paused, and the P-random continue to load/unload the pattern. One of ways to pause the P-serial is to hold the clock at high. On completion of load/unload the mode is changed to *Functional mode* by disabling *SE* signal to launch the test, and capture the responses. In state transition machine the transitions of states take place from $Q_j \rightarrow Q_r \rightarrow Q_f$.

P-serial mode:

Similar to *P-random mode*, this mode is activated to continue the pattern loading/unloading

in P-serial scan. During this mode only the P-serial is operated, the P-random is kept in hold state by disabling row enable and column driver. The test patterns are shifted in from *Scan-in_i* pins and the response are compacted in MISR. Once the load/unload is over the mode is changed to *Functional mode* for launch and capture. The state transition in this scenario would be $Q_j \rightarrow Q_s \rightarrow Q_f$.

Once all the patterns are applied and responses are collected in respective MISRs—for P-random and P-serial, a control signal is generated to shift out the response signature via *Scan-I/O*.

5.3.3 Scan Cell Segregation Methodology

The balanced segregation of scan cells plays a vital role in maintaining the test parameters at desired limits. The test power in Joint-scan is primarily dominated by the power dissipated due to P-serial scan chain. The test time however depends on the size of P-serial, P-random, and the test pattern set. Similarly, the pattern volume also depends on size of P-serial and P-random, and presence of don't care bits in pattern set. The logic area and routing wire length is a trade-off between P-serial and P-random. Larger the size of P-serial (implies lesser the size of P-random) lesser the wire length and routing area and vice-versa. Therefore to configure a balanced Joint-scan an appropriate criterion is needed. We propose a criterion based on the number of care bits present in the given pattern set. The proposed criterion is formulated using *care-bit ratio (CBR)*.

Criterion: *The P-serial scan chain will be formed with those scan cells wherein most of the patterns are specified with care bits. Whereas the P-random scan will be formed with those scan cells wherein most of the patterns are of don't care (X) bits.*

The *CBR* for each scan flip-flop is defined as follows.

$$CBR = \frac{\text{Total number of care bits}}{\text{Total number of patterns}} \quad (5.1)$$

Example 5.1 (Demonstration of CBR criterion).

Scan cells	Sff_0	Sff_1	Sff_2	Sff_3
Test Patt0	0	X	X	1
Test Patt1	X	0	X	1
Test Patt2	X	0	0	X
Test Patt3	1	1	X	X
Test Patt4	X	1	X	0
#Care-bits	2	4	1	3
CBR=#Care-bits/5	0.4	0.8	0.2	0.6

In the above example the bottom most line calculates the *care-bit ratio* for each scan cell-Sff_i. The ratio essentially indicates the density of care bits in a scan cell across the pattern set. As per the stated criterion, the scan cells with dense care bits are suitable to form P-serial whereas the scan cells with sparse care bits are suitable for P-random. The scan cells with larger *CBR* are suitable for P-serial because it minimizes the don't care bits and thereby increases the useful bits per pattern. The scan cells with smaller *CBR* are suitable for P-random because the don't care bits in P-random test patterns are not needed to be loaded into the scan cells. Therefore, pattern loading time as well as data volume can be minimized to the proportion of number of don't cares those can be avoided. Based on the *CBR* criterion, the **Algorithm 3** finds the optimum segregation.

Algorithm 3 Scan cell segregation methodology

- 1: Compute the CBR for each scan flip-flop
 - 2: Let Q \leftarrow sorted list of scan cells in non-increasing order of CBR value
 - 3: **for** i \leftarrow 1 to 9 **do**:
 - 4: pivot \leftarrow 10 * i
 - 5: P-serial \leftarrow First pivot% of flip-flops from Q
 - 6: P-random \leftarrow {Q - P-serial} flip-flops
 - 7: Joint-scan[i] \leftarrow {P-serial, P-random}
 - 8: Compute test time, pat volume, and test power for Joint-scan[i]
-

5.3.4 Test Time, Data Volume, and Power

Test time, T_t , test pattern volume, T_v , and power (Peak, P_p and Average, P_a) in Joint-scan are computed as follows. The following notations are used henceforth for the references.

N	Total number of patterns
n	Total scan cells ($= n_1 + n_2$)
n_1	Total scan cells in P-random
n_2	Total scan cells in P-serial
k	Number of scan-in pins for P-serial scan
l	Maximum length of scan chains in P-serial($= \lceil \frac{n_2}{k} \rceil$)
w	Number of write operation in P-random
c	Number of columns in P-random
r	Number of rows in P-random
a	Number of column address pins for P-random ($= \lceil \log_2 c \rceil$)
I	Total number of test pins in Joint-scan($= k + a + 3$)
P_{io}	Number of primary I/O pins
m	Size of MISR($=$ MISR in P-serial $+$ MISR in P-random)

Since the P-random and P-serial are operated concurrently the over all test time (#clock cycles) is computed as follows:

$$T_t = \sum_{i=1}^{N+1} (\text{Max}(T_{ts}^i, T_{tr}^i)) + m + N \quad (5.2)$$

where $T_{ts}^i = l$ and $T_{tr}^i = r + w$ are test times corresponding to P-serial and P-random for i_{th} pattern. The N is added to include the capture cycle.

The test data volume, T_v , is computed as follows.

$$T_v = T_{vs} + T_{vr} \quad (5.3)$$

where $T_{vs} = N * n_2 + k$ is data volume for P-serial and $T_{vr} = (N * c * r) + (w * c) + w + c$ is data volume for P-random. For P-serial the k is added to consider the MISR response

Table 5.2: Circuit specification of S-ISCAS-89 benchmark

Circuits	#PI/PO	#Flip-flop	#Faults	#Testpat	%FaultCov
S-S15850	77/150	9612	0.31E6	129	99.11
S-S5378	35/49	14857	0.51E6	1729	100
S-S9234	36/39	34815	0.89E6	608	99.23
S-S38584	38/304	39928	1.61E6	408	99.31
S-S38417	28/106	49080	1.38E6	418	100

signature. The computation of data volume considers only the scan patterns.

The test power dissipation is computed using the weighted transition metric (WTM) method. The *WTM* approximates the power by computing the toggles that arises in scan flip-flops during scan shift operation. For Joint-scan the peak and average power are computed as follows:

$$P_a = \frac{toggle_{ps} + toggle_{pr}}{T_t} \quad (5.4)$$

$$P_p = Max(Instantaneous\ toggle)$$

where $toggle_{ps}$ and $toggle_{pr}$ are total toggles in P-serial and P-random respectively. The *Instantaneous toggle* is the per-cycle toggles from P-serial and P-random.

5.4 Experimental Result and Discussion

The effectiveness of the proposed architecture is demonstrated through experiment on scaled ISCAS-89 benchmark circuits. The benchmark circuits are recreated from the ISCAS-89 circuit by replicating same module to multiple copy. The circuits are synthesized using Design Compiler[®] and ATPG is performed using TetraMax[®]. The architecture is evaluated for test time, test data volume, and test power. The results are compared with multiple sequential scan (MSS) and progressive random access scan

Table 5.3: JScan: P-random and P-serial size in #SFF

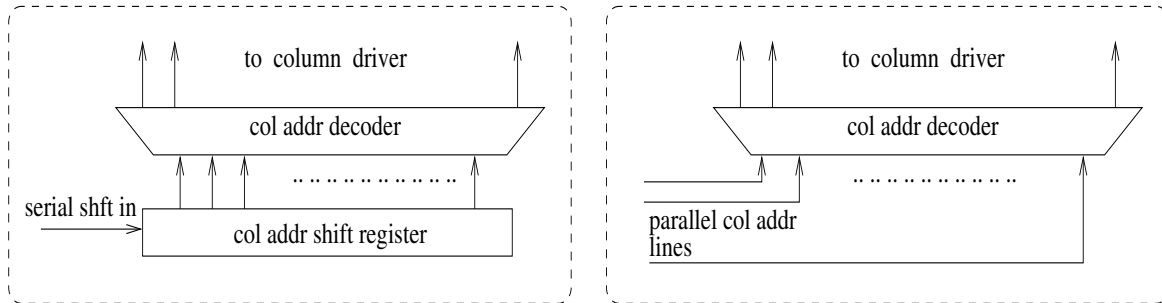
Bench Ckt→	S-S15850	S-S5378	S-S9234	S-S38584	S-S38417
#SFF in PRAN	7432	9539	25396	35191	43300
#SFF in PSER	2180	5318	9419	4737	5780

(PRAS) architecture. Test time, test data volume, and test power are computed using the formula stated in Section 5.3.4. For unbiased comparison the required number of test pins for all the three architectures are kept equal. For robust evaluation of Joint-scan the size of scaled ISCAS-89 circuits with respect to total number flip-flops are varied from 9,000 to 49,000. For all the circuits the complete scan design is considered. The ATPG for stuck-at fault is performed to generate test patterns with don't care bits, the patterns are statically compacted. The nature of benchmark circuits is provided in Table-5.9.

Table 5.4: Peak and Average Power of JScan in comparison with MSS[77]

Benchmark Ckts	Peak toggle			Average toggle		
	MSS	JScan	%of reduction	MSS	JScan	%of reduction
S-S15850	5091	1144	77.53	2774.4	661.37	76.16
S-S5378	7627	3386	55.6	4036.28	2007.5	50.26
S-S9234	17940	4994	72.15	9502.72	3292.69	65.35
S-S38584	20705	2448	88.17	10746.63	1572.18	85.37
S-S38417	24845	2998	87.92	11864.43	1543.43	86.99

It is naturally expected that in Joint-scan the power consumption is dominated by the activity generated from P-serial; the activity from P-random is negligible. Table-5.10 reports the peak and average toggle activity considering the same partition as shown in column-2 of Table-5.6. The computation of average and peak power follows the Eqn. 5.10 in Section-5.3.4. From the table it can be observed that the Joint-scan reduces the



(a) JScan Configuration C-1 with serial column address

(b) JScan Configuration C-2 with parallel column address

Figure 5.4: Two configurations of JScan with respect to column address application

power by large margin compared to MSS. Note that PRAS is the lower bound in power consumption. The peak and average power dissipation is computed using weighted transition metric (WTM) method. The JScan configuration considered for the computation is shown in Table-5.3.

Experiments for test time and pattern volume is performed for varied number of pins. The results are reported in Table-5.6. For Joint-scan, evaluation is carried out for all the possible partition obtained based on the CBR criterion as described in Section-5.3.3. The column-2 of Table-5.6 reports the best partition for the given number of pins. The experimental results show that the Joint-scan is effective to minimize the test time up to 48% and 56% in contrast to PRAS and MSS respectively for sufficient number of test pins. Pattern volume in PRAS primarily depends on the number of specified bits i.e. number of scan cells to be written. In MSS the pattern volume is essentially the function of number of scan flip-flops. The Joint-scan architecture basically take advantage of this trade-off and minimizes the overall pattern volume. The experimental results on pattern volume are reported in Table-5.6. Joint-scan effectively reduces the pattern volume by around 50% compared to PRAS and MSS. Figure 5.5 and Figure 5.6 plots the variation of test time and pattern volume over the pin count. It indicates that Joint-scan is scalable over increased number of pins. In the figures 5.5 and 5.6 the configuration C1 is a one type with serial shift of column addresses and C2 is with parallel column address

Table 5.5: Test time in contrast with PRAS[10] and MSS[77]

Ckts	#SFF in JScan: PRAN + PSER	#Test Pins	Pattern loading time in cycle				
			JScan	PRAS	MSS	%reduce wrt PRAS	%reduce wrt MSS
S-S15850	7432 + 2180	9	141847	152105	177253	6.7	20.7
		12	94706	141697	124108	33.2	24.7
S-S5378	9539 + 5318	9	3147710	4289836	3670674	26.6	14.2
		12	2117236	4107416	2569304	48.4	17.6
S-S9234	25396 + 9419	10	1912887	2018751	2646024	5.2	27.7
		13	1293867	1920956	1924939	32.6	32.8
S-S38584	35191 + 4737	10	969363	1085826	2036744	10.7	52.4
		13	646350	1015826	1481051	36.4	56.4
S-S38417	43300 + 5780	10	1335900	1413720	2564856	5.5	47.9
		13	1128854	1334182	1865127	15.4	39.5

application. The two different configurations are shown in Figure 5.4.

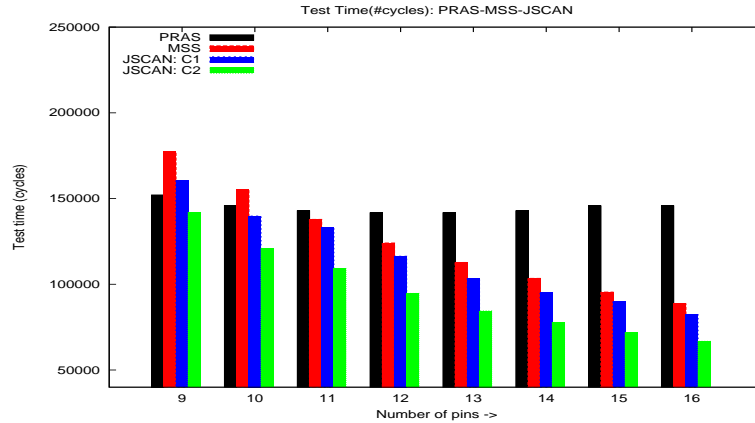


Figure 5.5: Test time behaviour for s-s15850 circuit over number of pins

We have performed a placement and routing for couple of benchmark circuits to evaluate the routing wire length (WL). The routing wire length is compared with the PRAS and MSS architecture. The TSMC 65nm technology node is considered for placement

Table 5.6: Pattern volume in contrast with PRAS[10] and MSS[77]

Ckts	#SFF in JScan: PRAN + PSER	#Test Pins	Pattern volume in bits				
			JScan	PRAS	MSS	%reduce wrt PRAS	%reduce wrt MSS
S-S15850	7432 + 2180	9	944427	1298562	1268851	27.3	25.6
		12	952732	1672092	1268725	43.1	24.9
S-S5378	9539 + 5318	9	19862981	34342177	25827809	42.3	23.4
		12	18887025	45417963	25820896	58.7	27.4
S-S9234	25396 + 9419	10	13683089	18372028	21208872	25.5	35.5
		13	13050185	23440197	21213131	44.3	38.5
S-S38584	35191 + 4737	10	7396300	10144322	16430168	27.1	54.9
		13	6974947	12734360	16426499	45.2	57.5
S-S38417	43300 + 5780	10	11288391	13073576	20571460	13.6	45.1
		13	11357339	16568044	20567701	31.4	44.7

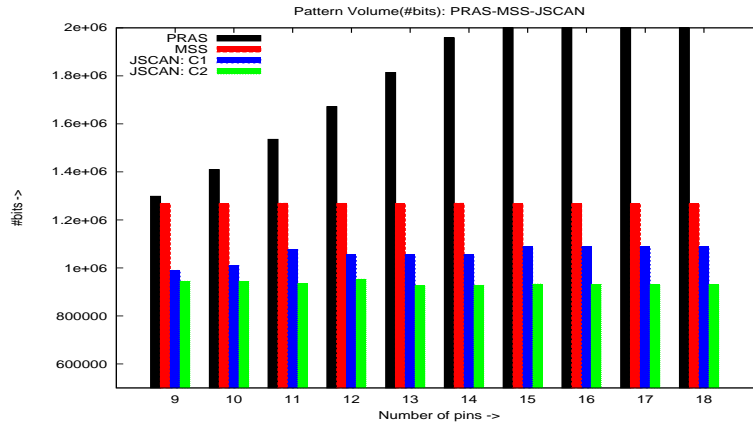


Figure 5.6: Patten volume variation for s-s15850 circuit over number of pins

and routing experiments. The placement and routing experiment is carried out with Cadence SoC Encounter. Table-5.7 shows the minimization in wire length by around 7 to 20% in comparison to the PRAS architecture. It is to be noted that the multiple sequential scan architecture is the lower bound in routing wire length as the architecture

Table 5.7: Routing wire length of JScan, PRAS, and MSS

Ckts	Layout area(μm^2)	Arch	Total WL(mm)	Avg WL/net (μm)	%red
S38417	65149.296	PRAS	279.539	45.02	0
		JOINT	221.278	35.29	20.84
		MSS	167.631	28.37	40.03
	75929.056	PRAS	281.586	45.32	0
		JScan	232.110	36.96	17.57
		MSS	179.171	30.71	46.37
S5378	7400.92	PRAS	29.521	28.72	0
		JScan	26.739	26.60	9.42
		MSS	23.997	26.87	18.72
	8616.944	PRAS	29.43	28.69	0
		JScan	27.30	27.16	7.4
		MSS	22.86	25.6	22.48

require fewer nets.

The following sections discuss the 2M-JScan architecture. This architecture additionally proposes a new test control mechanism and a clustering based algorithm. The 2M-JScan improves on test time minimization and reduction in hardware overhead.

5.5 2M-JScan Architecture

The proposed architecture integrates the serial and random access scan architecture to harness the best of each. Serial scan architecture requires minimal circuitry whereas the random access scan is power and data volume efficient. On the other hand the serial scan dissipates undesirable power and the random access scan occupies comparatively large layout area and introduces routing congestion. The proposed 2M-JScan architecture

simplify the test control mechanism and further reduces test time, keeping data volume and power same as JScan[110].

The 2M-JScan consist of two sub-scan architectures: Partial serial scan (P-serial) and Partial random scan (P-random). The available flip flops are segregated into two groups to form P-serial and P-random. The P-serial is a serial scan chain formed from the first group of flip flops and the P-random is a random access scan formed from the second group of flip flops. Following sub sections give detail on architecture.

5.5.1 The Architecture

The three primary components in architecture are P-serial, P-random, and test control logic (TCL). The P-serial and P-random are implemented with multiple serial scan chains (MSS) and progressive random access scan (PRAS)[10] respectively. To realise the proposed architecture we have identified three main challenges:

1. integrating and operating the P-serial and P-random
2. maintaining equilibrium in shift time across all patterns
3. grouping of flip flops to obtain the best results.

The proposed architecture is shown in Figure 5.7. The test control logic ensures correct operation of the architecture. The challenge in the part of test control logic is to operate both the architectures concurrently for load/unload of test/response, launch, and capture operations. The P-serial consists of scan-in lines/pins to shift in the test stimuli and MISR at the out put to compact the response and a scan-out line to shift out the final response. The test control logic controls the activity in P-serial using serial scan enable (SSE) and scan clock line. The detail control lines are implementation issues and may vary for different implementation, here we will primarily be focusing on architectural issues.

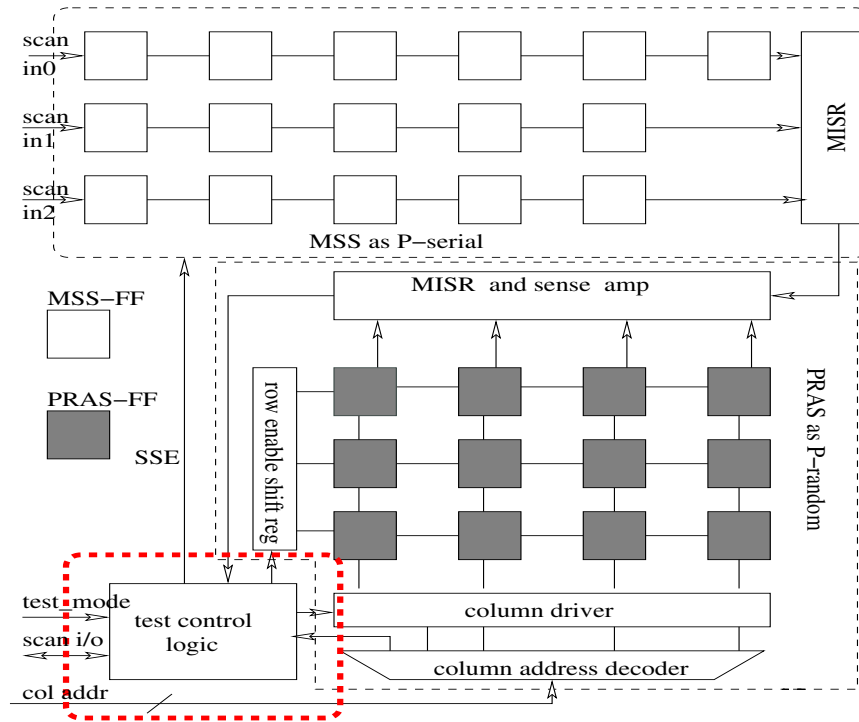


Figure 5.7: Proposed scan Architecture (the 2M-JScan)

The P-random, similar to PRAS, consists of three major units: row address register, column address decoder, and MISR. It also uses a sense amplifier to read responses. The scan flip-flops are arranged in grid fashion, and they are connected with row and column enable lines. The column enable lines are driven by column driver and the row enable lines are driven by row address shift register. The responses in P-random are compacted using MISR. The MISR is connected with test control logic through scan out line[10].

The test control logic in this architecture is simplified to operate with only one test mode line/pin unlike the JScan where two test mode pins were required. The serial scan enable signal is generated using test control logic. The test control logic also gets input from column address decoder. These inputs are generated from reserved address bits dedicated for generating control signals. Test control logic steers overall test and functional mode operations.

5.5.2 Limitations of JScan

In JScan, pattern loading/unloading time varies for P-serial and P-random. The variation arises due to the fact that in P-random number of write operations vary from a test stimuli to another test stimuli. The write of test stimuli bits in P-random is based on two criteria: 1) the stimuli must be specified bit (the don't care bits in stimuli are not considered for write), 2) the stimuli bit must differ from the don't bit of last response. This leads to variation in load/unload time. To control this variation the test control logic for JScan needs two test mode pins. Furthermore, additional circuitry is required to keep track of completion of stimuli/response loading/unloading in P-serial and P-random. Following example demonstrates the scenario.

Example 5.2. In Table-5.8, the P-serial for S_0 takes 2 cycle where as P-random takes 4 cycles which results into variation of 2 cycles. Similarly for S_1 it results in variation of 1 cycle. P-serial has two parallel scan chains, chain-1: sf_0, sf_1 and chain-2: sf_4, sf_6 and the P-random has two rows, row1: sf_2, sf_3 , and sf_5 , and row2: sf_7 and sf_8 , and three columns. For P-serial it is easy to see that it consumes two cycles to load/unload test/response. For P-random it varies for S_0 and S_1 , for S_0 it takes two cycles– one cycle to read response + one cycle to write bit 1 to sf_5 . Similarly P-random takes three cycles for S_1 . Therefore there is a variation in test time across the patterns. To keep track of this variation for each pattern a dedicated controller is employed.

Table 5.8: Variation in loading/unloading time

patterns	P-serial(2-chains)	P-random(2-row/3-col)	l/u time P-s P-r
	sf_0 sf_1 sf_4 sf_6	sf_2 sf_3 sf_5 sf_7 sf_8	
S_0	1 x 0 1	x x 1 0 x	2 4
R_0	x x 0 0	1 x 0 x x	
S_1	0 x 0 1	x x 1 x x	2 3
R_1	1 x x 1	0 0 1 0 x	

In the current proposed architecture we equalize the variation and thereby eliminate the need for additional controller. We have proposed a methodology to align the patterns for P-serial and P-random in such a way that the pattern loading time for each pattern turns out to be equal.

5.5.3 Test Pattern Alignment

In this section we will describe the procedure for padding the patterns such that each of the patterns will be aligned to have equal loading/unloading time.

First of all we have to find out the maximum load/unload cycle (this is synonymous to shift time in case of serial scan chain) by any individual pattern. Thereafter the candidate test patterns, which are to be appended, will be selected and padding bits will be computed.

Let $S_i(S_i^s, S_i^r)$ and $R_i(R_i^s, R_i^r)$ be i^{th} stimuli and response pattern respectively, where S_i^s and S_i^r are the stimuli pattern for P-serial and P-random respectively and R_i^s and R_i^r are response pattern for P-serial and P-random respectively. And, let $T_i = \{S_i, R_i\}$ be the i^{th} test set with stimuli and response pattern. The computation of maximum load/unload cycle can be done as follow:

$$maxcycle = Max_{i=0}^{N-1}(lutime(T_i)) \quad (5.5)$$

where $lutime(T_i)$ is load/unload time for test set T_i , and N be the total number of test set. Here we assume that stimuli loading time \geq response unloading time for P-serial as well as for P-random.

Now, the padding bit pattern for each test set, T_i , will be determined using $maxcycle$. We have to note that only the test stimuli needs to be padded with required number of bit patterns. The required number of bit patters are decided based on the number of additional loading/unloading cycles needed to equal the $maxcycle$. For example, if the stimuli S_1^s are to be padded then the number of padding bit pattern needed are in the order of $maxcycle - loadtime(S_1^s)$. Similarly for stimuli S_1^r the padding bit patterns are

in the order of $maxcycle - loadtime(S_1^r)$.

Appending for P-serial Pattern: For P-serial scan chain the padding bit patterns are appended at the beginning (right most places) of the test stimuli. The number of additional bits appended are equal to the $maxcycle - loadtime(S_j^s)$, where S_j^s is the j^{th} stimuli.

The bit pattern considered for appending would be all 0s (0 0 0 0 . . .) or all 1s (1 1 1 1 . . .). The appending pattern should be chosen such that it does not affect response compacter.

Example 5.3. Let's consider the scenario described in Table-5.8. Let P-serial be two parallel scan chains chain-1: sf_0 and sf_1 , and chain-2: sf_4 and sf_6 . Considering stimuli S_0 with chain-1: 1 x pattern and chain-2: 0 1 pattern, the new pattern with padding would be $\langle 1 x | 0 0 \rangle$ to be shifted to chain-1 and $\langle 0 1 | 0 0 \rangle$ to be shifted to chain-2.

Appending for P-random Pattern: For P-random the pattern appending procedure differs from that of P-serial. The basic procedure to write a scan flip flop in P-random differ. First a row is enabled to read the response, then a column address for the target flip flop is applied, thereafter the stimuli bit is applied for write. These three steps are essential. Accordingly, the pattern has to be formatted. One of the possible format is $\langle coladdr, bit \rangle$, where $coladdr$ is address of target flip flop in an enabled row and bit is stimuli to be written.

The pattern for appending has to be in the given format of $\langle coladdr, bit \rangle$. The last column to be written in last row of P-random scan grid is chosen as the desired pattern for appending. In case of no column to be written in last row, the first column is considered as the last column to be written and don't care bit is considered as stimuli bit to be written.

Example 5.4. Considering the scenario depicted in Table-5.8, let P-random be configured as two rows and three columns. The test set $T_0(S_0^r, R_0^r)$ does not need appending since it's $maxcycle - loadtime(S_0^r)$ is zero. However, $T_1(S_1^r, R_1^r)$ need one pattern appending. The appending pattern in this scenario would be $\langle 0 0 0, x \rangle$, where $[0 0 0]$ is the address of 0^{th} column and x is the stimuli bit.

In this way all the patterns are aligned by appropriate appending. The implication of alignment process is to enable load/unload the stimuli/response with equal time. As stated earlier the pattern alignment results in saving of hardware overhead due to test control logic.

5.5.4 Test Control Mechanism

Test control mechanism in this architecture is simplified compared to JScan. The mode control uses only one test mode pin whereas JScan uses two pins. This saving of one additional pin contribute in minimization of test time. The controller now does not need to keep track of completion of load/unload operation in either P-serial or P-random unlike in JScan. This reduces additional circuitry. The controller now is having only two states (shown in Fig-5.8): test and functional, whereas the JScan is having four: joint, random, serial, and functional as discussed in Section-5.3.1 of JScan architecture.

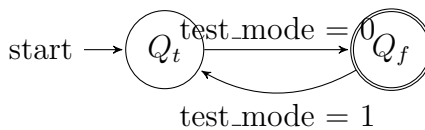


Figure 5.8: Two states controller: Q_t for load/unload and launch of test, Q_f for response capture and normal function .

5.5.5 Functionality

Functionality of the architecture is controlled in two modes: 1) functional mode, 2) test mode. Note that both the scans, P-serial and P-random, are operated concurrently.

Functional mode: Functional mode controls two primary operations: 1) normal function and 2) response capture. Normal function is when the circuit perform desired functional operation in normal mode. The test control logic keep serial scan enable (SSE) signal at low to operate P-serial scan flip-flops in normal mode. Similarly row address

shift register and column driver are disabled to operate P-random flip-flops in normal mode.

Response capture is part of test procedure performed in functional mode. Once the test stimuli is launched the mode of operation is changed to functional mode ($test_mode = 0$), and after a delay of one clock cycle (called dead cycle) the response is captured. Once completed the mode is changed to test mode ($test_mode = 1$).

Test Mode: Test mode controls three primary test operations: 1) loading/unloading of stimuli/response, 2) launching of stimuli, 3) shift out of response from MISR. The test mode is enabled by holding $test_mode = 1$. Loading/unloading operation in P-serial and P-random takes place simultaneously. The SSE signal is kept high during this mode. In P-serial the test stimuli are scanned in through *scan-in* lines and responses are compacted using MISR. The load/unload operation in P-random are performed row by row. First a row is enabled by signal generated from row address shift register, then the column driven activates the desired flip-flops in the enabled row for to be loaded with test stimuli bit one by one. Once a row is completed the same procedure is repeated for next row until the last row. Recall that the test patterns are aligned in such both, P-serial and P-random, will complete the load/unload operation simultaneously. Once load/unload is completed the test stimuli is launched and test mode is changed to functional mode ($test_mode = 0$) for capture.

Once all the test sets are applied and responses are compacted in MISRs the compacted response is shifted out for comparison. This completes entire test procedure.

5.5.6 Clustering of Scan Flip-flops

Grouping scan flip flops is another challenge because it affects test time, data volume, and test power. To minimize data volume the flip flops which for most of the test pattern contain don't care bits are to be grouped to P-random and flip flops which contain care bits for most of the test pattern are to be grouped with P-serial. Also the appending pattern has to be considered because it adds up to the data volume. Test time is largely affected by maximum number of write operation being performed for any test pattern

in P-random and the length of scan chain in P-serial. Therefore, size of both the scan chains play role. Test power in Joint-scan is primarily consumed and dominated by P-serial chain. Power consumed due to P-random is negligible compared to that of P-serial. Therefor maintaining length of scan chains is important. We propose here an *1D data clustering* algorithm for grouping. Following are problem statement and algorithm.

Problem:

Given a set of flip-flops and corresponding number of care and don't care bits, **find** two clusters **such that** the difference between two centroids, ie distance, will be optimum.

The optimum distance means the best possible minimization of test time, data volume, and power. Centroid and distance are compute as follow:

$$centroid = \frac{\sum_{i=0}^{k-1} x_i}{k} \quad (5.6)$$

and

$$distance = d_i - d_j \quad (5.7)$$

where x_i is don't care bit count, k is cluster size, and d_i is centroid of cluster i .

Algorithm:

First, the flip-flops will be sorted in increasing value of don't care bits. Starting with initial two clusters where cluster-1 will have first flip-flop and cluster-2 will have remaining $N - 1$ flip-flops from sorted list. We set a pivot pointer to progressively update the cluster population with new flip-flop(s) from sorted list, update of clustering can be done with varying number of flip-flops. Iteratively centroid and distance will be computed and compared with the previously computed cluster distance. At the last iteration the best clustering will be identified with optimum pivot point. Figure-5.9 demonstrates the working of algorithm. In the figure sf_i are scan flip-flops sorted in non-decreasing number of corresponding don't care bits. In each iteration the algorithm gives a two clusters: P-serial and P-random. Clusters are updated/populated in subsequent iterations and the final iteration provides the best possible clustering.

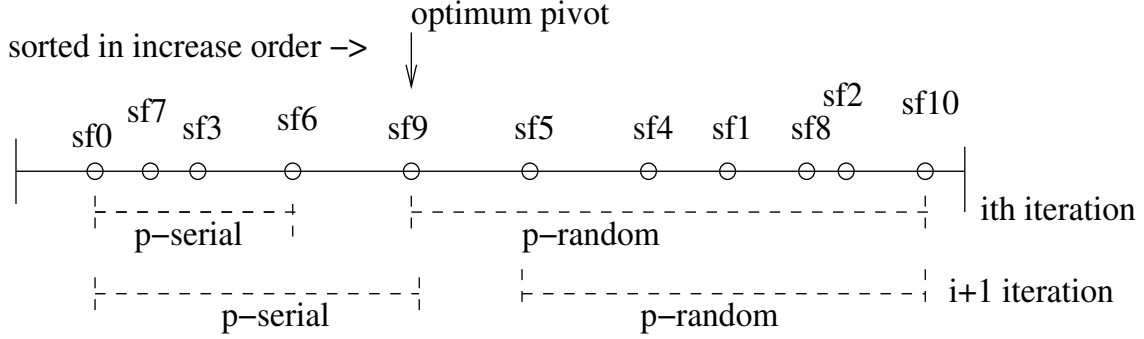


Figure 5.9: Graphical demonstration of 1D clustering algorithm

5.5.7 Computation of Test Parameters

Test time, t_{time} , test data volume, v_{data} , and power (peak, p_{peak} and average, p_{avg}) for the proposed architecture are computed as follow. Following notations are used henceforth.

- N Total number of test sets(T_0 to T_{N-1})
- n Total scan cells ($= n_1 + n_2$)
- n_1 Total scan cells in P-random
- n_2 Total scan cells in P-serial
- k Number of scan-in pins for P-serial scan
- l Maximum length of scan chains in P-serial($= \lceil \frac{n_2}{k} \rceil$)
- w Number of write operation in P-random
- c Number of columns in P-random
- r Number of rows in P-random
- a Number of column address pins for P-random ($= \lceil \log_2 c \rceil$)
- I Total number of test pins in 2m-jscan($= k + a + 2$)
- P_{io} Number of primary I/O pins
- m Size of MISR($=$ MISR in P-serial $+$ MISR in P-random)

Since the P-random and P-serial are operated concurrently the over all test time(in #clock cycle) is computed as follow.

$$t_{time} = (maxcycle * (N + 1)) + N + m \quad (5.8)$$

Table 5.9: Circuit specification of S-ISCAS89 benchmark

Circuits	#PI/PO	#Flip-flop	#Faults	#Testpat	%TestCov
S-S13207	62/152	9575	0.22E6	99	99.08
S-S15850	77/150	9612	0.31E6	129	99.11
S-S5378	35/49	14857	0.51E6	1729	100
S-S9234	36/39	34815	0.89E6	608	99.23
S-S38584	38/304	39928	1.61E6	408	99.31
S-S38417	28/106	49080	1.38E6	418	100

where $maxcycle$ is load/unload cycle for each test pattern.

The test data volume, v_{data} , is computed as follows.

$$v_{data} = v_{sdata} + v_{rdata} + (N * P_{io}) + m \quad (5.9)$$

where $v_{sdata} = N * n_2 + k + v_{s_{apnd}}$ is data volume in P-serial and $v_{rdata} = (N * c * r) + (w * c) + w + c + v_{r_{apnd}}$ is data volume in P-random. $v_{s_{apnd}}$ and $v_{r_{apnd}}$ are data volume due to pattern appending.

The test power dissipation is computed using the weighted transition metric method which approximate the power by computing the toggles arising in scan flip flops during shift operation. For 2M-JScan the peak and average power are computed as follows:

$$p_{avg} = \frac{toggle_{ps} + toggle_{pr}}{t_{time}} \quad (5.10)$$

$$p_{peak} = Max(instantaneous\ toggle)$$

where $toggle_{ps}$ and $toggle_{pr}$ are total toggle in P-serial and P-random respectively. The instantaneous toggle is the per-cycle toggles activated from P-serial and P-random.

Table 5.10: Peak and Average Power in comparison with MSS[77]

Ckts	Peak toggle			Average toggle		
	MSS	2M-JScan	%of reduce	MSS	2M-JScan	%of reduce
S-S13207	4915	1491	69.66	2729.4	879.84	67.7
S-S15850	5091	1438	71.75	2774.4	813.04	70.7
S-S5378	7627	3083	59.5	4036.28	1685.03	58.25
S-S9234	17940	5132	71.39	9502.72	2836.88	70.14
S-S38584	20705	2981	85.6	10746.63	1657.86	84.55
S-S38417	24845	3010	87.88	11864.43	1518.11	87.2

5.6 Experimental Result and Discussion

Experiment is performed on scaled ISCAS-89 benchmark circuits. The benchmark circuits are recreated from the existing ISCAS-89 circuit by replicating and interconnecting the top module to desired number of time. Design Compiler[®] is used for synthesis and TetraMax[®] is used for ATPG. Proposed architecture is evaluated for test time, test data volume, and test power. Results are compared with JScan[110], PRAS[10], and multiple sequential scan (MSS). Test time, test data volume, and test power are computed using the equations stated in Section 5.3.4. For unbiased comparison the required number of test pins for all the four architectures are kept equal considering PRAS as the baseline for pin count.

For robust evaluation the size of scaled ISCAS-89 circuits with respect to total flip-flops are varied from 9000 to 49,000. For all the circuits the full scan design is considered. The ATPG for stuck-at fault is performed to generate test patterns with don't care bits. The characteristics of benchmark circuits are listed in Table-5.9.

Experiments for test time and data volume is performed for varied number of pins. The results are reported in Table-5.11 and Table-5.12. The column-2 of Table-5.11 and Table-5.12 reports the optimum partition (PRAN and PSER) for given number of test

Table 5.11: Test time of 2M-JScan in contrast with JScan[110], PRAS[10] and MSS[77]

Ckts	#SFF in 2M-JScan PRAN + PSER	#Test Pins	Pattern loading time in cycle						
			proposed 2M-JScan	JScan	PRAS	MSS	%red wrt JScan	%red wrt PRAS	%red wrt MSS
S-S13207	6705 + 2865	9	95599	NA	109229	135439	NA	12.4	29.4
		12	57439	NA	101500	94852	NA	43.4	39.4
S-S15850	6915 + 2697	9	123893	141847	152105	177253	12.6	18.5	30.1
		12	87749	94706	141697	124108	7.3	38.1	29.2
S-S5378	8945 + 5912	9	2834564	3147710	4289836	3670674	9.9	33.9	22.7
		12	1873240	2117236	4107416	2569304	11.5	54.3	27.1
S-S9234	24722 + 10093	10	1697079	1912887	2018751	2646024	11.2	15.9	35.8
		13	1227703	1293867	1920956	1924939	5.1	36.1	36.2
S-S38584	34204 + 5724	10	921450	969363	1085826	2036744	4.9	15.1	54.7
		13	594642	646350	1015826	1481051	7.9	41.4	59.8
S-S38417	43300 + 5780	10	1270007	1335900	1413720	2564856	4.9	10.1	50.4
		13	1128723	1128854	1334182	1865127	0.0	15.4	39.5

pins. Table-5.11 show that 2M-JScan reduces test time compared to earlier architectures. Results show up to 50% reduction in test time compared to PRAS and MSS for sufficient number of test pins. Test data volume compared to PRAS and MSS is reduced by around 40 - 50%. The small quantity of negative reduction compared to JScan is expected because of pattern appending. The same for test time is not true because the 2M-JScan uses only one test mode pins unlike the JScan which uses two pins.

It is intuitive that in 2M-JScan the power consumption is dominated by the activity generated from P-serial, the activity from P-random is negligible. Table-5.10 reports the peak and average toggle activity considering the same partition as shown in column-2 of Table-5.11. The computation of average and peak power follows the Eqn. 5.10 in Section-5.3.4. From the table we can observe that the proposed architecture reduces the power by large margin compared to MSS, wherein it is to be noted that PRAS is the lower bound in power consumption.

Table 5.12: Test data volume of 2M-JScan in contrast with JScan[110], PRAS[10] and MSS[77]

Ckts	#SFF in 2M-JScan: PRAN + PSER	#Test Pins	Data volume in bits						
			proposed 2M-JScan	JScan	PRAS	MSS	%red wrt JScan	%red wrt PRAS	%red wrt MSS
S-S13207	6705 + 2865	9	741156	NA	950730	968524	NA	22.1	23.4
		12	628306	NA	1222386	968626	NA	48.6	35.1
S-S15850	6915 + 2697	9	946941	944427	1298562	1268851	-0.2	27.1	25.3
		12	976414	952732	1672092	1268725	-2.4	41.6	23.1
S-S5378	8945 + 5912	9	20051681	19862981	34342177	25827809	-0.9	41.6	22.3
		12	18929643	18887025	45417963	25820896	-0.2	58.3	26.6
S-S9234	24722 + 10093	10	13885425	13683089	18372028	21208872	-1.4	24.4	34.5
		13	13277928	13050185	23440197	21213131	-1.7	43.3	37.4
S-S38584	34204 + 5724	10	7698932	7396300	10144322	16430168	-4.0	24.1	53.1
		13	7121708	6974947	12734360	16426499	-2.1	44.1	56.6
S-S38417	43300 + 5780	10	11135560	11288391	13073576	20571460	-1.3	14.8	45.8
		13	11880188	11357339	16568044	20567701	-4.6	28.3	42.2

5.7 Variants of Joint-Scan Architecture and Future Work

The proposed Joint-scan architecture sets a framework on which the variants of serial and random access scan can be integrated. Some of the possible variants that could be fits in the proposed framework are identified and studied for possible implementation. Detail study with experiment and implementation are the subject of our future work.

Table 5.13: Possible Variants of Joint-scan

P-serial scan	P-random scan
Illinois Scan[24, 103]	PRAS[10]
Compression Based Scan[87]	Clustered RAS[51]
Scan Tree[18]	SIRAS[1]

The broadcast based Illinois scan architecture comes with a major advantage that it minimizes pattern volume and test time with fewer scan-in pins. Implementing this architecture as P-serial would result into saving of pins those could be used for P-random either with PRAS or clustered RAS. Similarly the compression based serial scan can be paired with PRAS as well clustered RAS. The SIRAS improves upon the test time on standard RAS architecture with very fewer number of test pins, less than 4. Hence SIRAS can be used along with compression based serial scan.

For each of these clubbing the crucial role is to be played by the flip-flop partitioning algorithm. It is important to identify based on set of characteristic the flip-flops for serial and random part of the Joint-scan chain. Once identified based on the characteristic the next important task is to determine the size of serial and random part (which we call P-serial and P-random respectively) of scan chain for optimal results. The strategy for grouping of flip-flops may differ for different configurations.

5.8 Conclusion

This chapter contributed two new scan DFT architectures: JScan and 2M-JScan. Both the architectures are aimed at minimizing the test time, data volume, and test power further from state of the art methodology. The experimental results demonstrate that both the architectures reduce test time, data volume, and test power with large margin compared to PRAS and MSS architectures. The JScan architecture could minimize test time up to 48% compared to PRAS and up to 56% compared to multiple sequential scan chain. Also the experimental results show that for increased number of test pins the architecture perform better.

The 2M-JScan architecture further minimizes the test time compared to JScan, PRAS, and MSS without much affecting test data volume and test power. New test control mechanism for 2M-JScan architecture is proposed which enables the architecture to function in similar way the standard DFT architecture does. Procedure for test

set alignment is described. An 1D clustering algorithm is proposed for efficient grouping/clustering of scan flip-flops.

The proposed architectures also provide a framework for further optimization of the test parameters by letting the designer to choose the best hybridization among the available DFT architectures for serial and random access scan. Some of the possible architectures are discussed in Section-5.7. The prediction is that the architecture would produce a good result when the grouping is performed optimally which could take advantage of available test pins.

We anticipate that both the architectures, JScan and 2M-JScan, could be well scaled for billion gates design by appropriately guiding the grouping algorithm to provide optimum size for P-serial and P-random. The proposed architecture is limited to stuck-at fault and transition fault testing with launch-on-capture strategy. Further research is needed.

Chapter 6

Scan Flip-flop Design

The scan flip-flop plays a pivotal role in designing an efficient scan chain. Traditionally, designing of scan flip-flop is motivated by two problems: performance and power. In the Joint-scan architecture also, the scan flip-flop plays an important role in designing efficient scan architecture. Particularly in the PRAS and MSS based implementation of Joint-scan, needs a new scan flip-flop for MSS chain. The traditional multiplexor based scan flip-flop is functionally incompatible with PRAS scan flip-flop. This motivates our work to design an efficient and compatible Joint-scan flip-flop.

6.1 Introduction

The testing of today's highly complex *SoC* designs is a big challenge faced by *VLSI* test community. Scan design is the only *DFT* approach that can effectively test a highly complex design with very high fault coverage. The objective of scan design is to achieve full controllability and observability of each and every flip-flop in the design. In a full scan design, each flip-flop is replaced by a scan flip-flop. A scan flip-flop is nothing but a muxed input master-slave based *D* type flip-flop. The scan multiplexer has two inputs: data input (*D*) and scan input (*SI*). The input selection is done using a control signal called *test_enable* (*TE*). In functional mode, data input is selected and the scan flip-flop

function as a regular flip-flop. In test mode, scan input is selected and all the scan flip-flops are connected into one or more serial shift register(s). The serial shift register(s) are popularly known as scan chain(s). All flip-flops of the scan chain are loaded with desired data by consecutive application of the clock signal. The full scan design reduces the sequential test problem to combinational test problem.

Serial scan is obviously not free from drawbacks. There are some inherent penalties associated with serial scan. These penalties include:

1. performance overhead,
2. test data volume,
3. test power consumption, and
4. test application time.

The performance overhead of serial scan is related to the scan multiplexer[4, 21]. The scan multiplexer falls into each clocked path and adds performance penalty of approximately two gate-delays. A circuit without scan design and with scan design are shown in Figure 6.1. As it is observable in Figure 6.1a, critical path of a sequential circuit without scan insertion is decided by the longest combinational path between two flip-flops. However, in a scan inserted sequential circuit (see Figure 6.1b) the same critical path is elongated by a scan multiplexer at the end of combinational path. Scan also adds an extra fanout at the output of flip-flop. Both of these factors increase the critical path delay and can reduce functional clock speed by 5 to 10% [21].

This makes it necessary to eliminate the performance overhead of scan multiplexer. Several solutions have been proposed to alleviate the performance penalty of scan design. One such solution that alleviates the performance overhead as well as other penalties associated with the serial scan design is the use of partial scan instead of full scan. In partial scan design, only a subset of all flip-flops in Circuit-Under-Test (*CUT*) are replaced by scan flip-flops to form a scan chain. This subset does not include flip-flops of the critical paths and hence reduces the performance penalty of scan. Additionally,

the partial scan design techniques also reduce test data volume and test application time which are directly related to testing cost. However, the partial scan design techniques may reduce fault coverage of the *CUT*. The selection of flip-flops to be included into partial scan design can be testability measures based [19, 55, 128, 129], structure based [6, 23, 26, 61], or *ATPG* based [3, 50, 66]. The structure based techniques select flip-flops to cut-off the feedback path. These techniques make use of heuristics based on network topology for selecting a minimum set of flip-flops and do not explicitly analyze the circuit behavior. The *ATPG* based or test generation based techniques select flip-flops that are useful for test generation. These techniques first use sequential *ATPG* to generate test vectors for all possible faults. For the faults which remain undetectable, the culprit flip-flops are found and are included into partial scan design. These techniques are computationally intensive and are un-affordable. Most of the partial scan techniques require computationally demanding sequential *ATPG*, and can not be afforded with ever increasing circuit complexity. Furthermore, partial scan design techniques do not provide as high fault coverage as provided by full scan design, and are difficult to integrate into the existing industry design flow.

Random access scan (*RAS*) is an alternative *DFT* technique that can alleviate the problems associated with serial scan. Literature shows that *RAS* can greatly reduce test application time and test data volume along with test power reduction up to 99% [1,

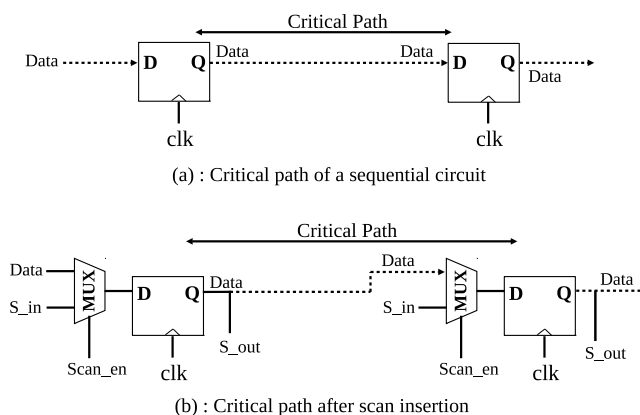


Figure 6.1: Scan design performance overhead[21]

10, 12]. However, the hardware overhead associated with *RAS* is prohibitively high. Furthermore, observability of storage cells and *RAS* architecture implementation are some other issues which need to be addressed properly [10]. Recently, we have started to explore a new architecture which intend to solve the problems of *RAS* and serial scan architecture. In Chapter-5, we have discussed in detail the principle of *Joint-scan* architecture. To recall, the Joint-scan architecture (we will also be referring the architecture as mixed mode DFT scan architecture in this chapter) integrates the serial and random access architecture to derive the best from each of the architectures. The experimental results reported in Chapter-5 show that the Joint-scan architecture is highly efficient compared to the PRAS and standard serial scan. *However, the Joint-scan also faces some unresolved issues at implementation level. One of such un-availability of a common scan cell that can be used for both, P-serial as well as P-random, DFT structures.*

In this Chapter we have proposed a new scan flip-flop design that can be used as both a serial scan cell as well as a random access scan cell. The major advantages of the proposed scan flip-flop design are as follows:

1. The proposed design eliminates the performance penalty of serial scan by removing scan multiplexer off the functional path.
2. The proposed scan flip-flop can be used as a serial scan cell as well as a *RAS* cell, in Joint-scan test. This can be very helpful in Joint-scan architecture implementation.
3. The proposed design does not use any extra control signal and has a minimal area overhead compared to conventional scan flip-flop.
4. The new scan flip-flop is capable of applying all sort of tests that can be applied with a conventional scan flip-flop. The proposed design fully comply with existing industry design and test flow.

The remainder of the Chapter is organized as follows: Section-6.2 describes the conventional scan flip-flop and the proposed scan flip-flop design. Section-6.2 further elaborates on the details of test application process and retainment of test quality. Section-6.4

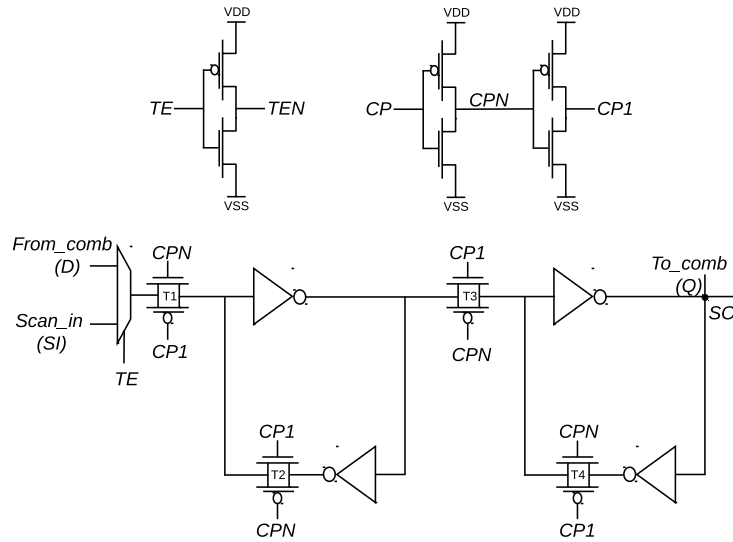


Figure 6.2: A conventional scan flip-flop design

explains the possible use of proposed design in Joint-scan architecture. Section 6.5 explains the post layout timing results and verifies the efficacy of proposed design in eliminating the performance overhead of scan design. Section 6.6 concludes the Chapter.

6.2 Preliminaries and Proposed Scan Flip-Flop

A large variety of scan flip-flop implementations are available in literature. A conventional scan flip-flop design is shown in Figure 6.2. This scan cell is a master slave latch based positive edge triggered muxed input D type flip-flop. The transmission gate $T1$, and the inverter pair connected back to back via transmission gate $T2$ forms the master latch. The slave latch comprises transmission gate $T3$ and the inverter pair connected back to back via transmission gate $T4$. The multiplexer at the input of master latch selects between functional input (D) or scan input (SI) depending upon the value of test control signal $test_enable$ (TE). In test mode, when TE is high (1), SI is selected and is connected to master latch's input. When the clock signal (CP) is low (0), the value of SI propagates to the master latch. In the mean time, slave latch retains the value from previous clock cycle. The value latched into the master propagates to slave

latch when CP turns to high (1), and to the output Q of scan flip-flop. Similarly, when $test_enable$ signal (TE) is set to 0, functional input D is selected and the circuit operates in functional mode.

6.2.1 Proposed Scan Flip-Flop Design

This section discusses the working of proposed scan flip-flop in different modes of operation. The proposed scan flip-flop schematic design is shown in Figure 6.3. Instead of a multiplexer at master latch's input, the proposed design uses a separate path for loading test vector values into the master latch. Furthermore, the proposed scan flip-flop uses a low cost dynamic slave latch for shifting of test vectors in test mode. In functional mode, functional slave latch's output Q drives the combinational circuit inputs. The master latch of proposed scan flip-flop is formed by transmission gate $T1$, and inverter pair ($i1, i2$) connected back to back via transmission gate $T2$. Similarly, the slave latch is formed by transmission gate $T3$, and inverter pair ($i3, i4$) connected back to back via transmission gate $T4$. The dynamic slave latch comprises transmission gate $T7$ and inverter $i7$. The test mode path is formed by adding transmission gate $T5$, $T6$, buffer $i5$, and inverter $i6$ to the master latch structure. It should be noted that the extra gates added to master stage to form the test mode input path are not on the functional path. This extra circuitry remains disabled during functional mode and the proposed scan flip-flop acts as a regular flip-flop. The master latch and the slave latch are controlled by functional clock signal CP . The test mode input path is disabled by $test_enable$ cum scan clock signal SCK . The details of working of the proposed design in different modes of operation is explained in the following subsections:

Functional mode

The proposed scan flip-flop works as a regular flip-flop in functional mode. In functional mode, scan clock signal SCK is kept at constant logic high (1) level. As long as SCK is at constant high (1) level the transmission gate $T5$, and $T6$ remain disabled. This

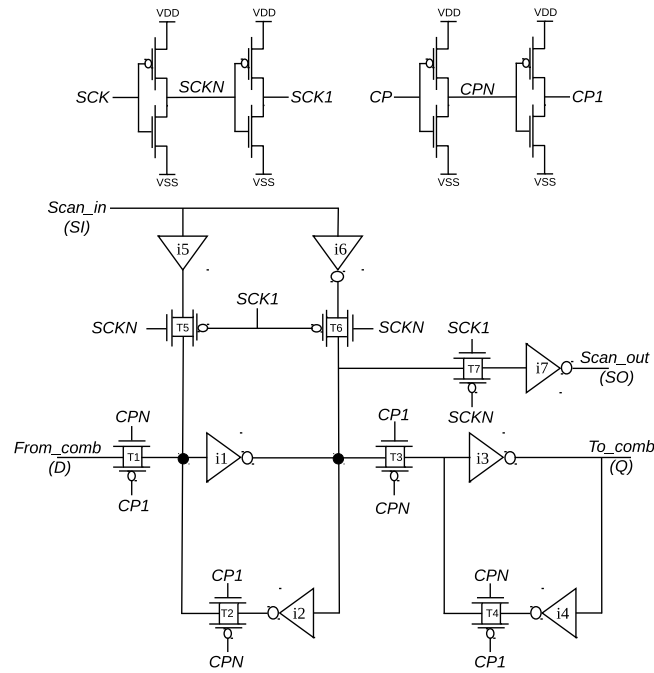


Figure 6.3: Proposed scan flip-flop design

disconnects the test mode input path from master structure and the proposed scan flip-flop functions as a regular flip-flop. The scan clock signal (SCK) held at constant high (1) level indicates functional mode operation. During the functional mode operation the transmission gate $T7$ always remains enabled. This keeps the dynamic slave latch always transparent during functional mode and makes the scan output (SO) toggle every time whenever there is a change in master latch's state. However, that is not of any concern as the scan output (SO) feeds the scan input path of successive scan flip-flop which remains disconnected from the master structure. In functional mode, the master latch gets its input from functional data input D . When clock CP is low, the value of functional input D propagates into the functional master latch. When CP turns to high, the value latched into the master propagates to functional slave latch, and to output Q of the scan cell.

Test mode

While keeping the functional clock CP held at constant high (1) level, consecutive application of scan clock SCK makes the proposed scan flip-flop to function in test mode.

As the functional clock CP is kept high (1), the transmission gate $T1$ always remains disabled in test mode. This disconnects the functional input D from the master latch. During test mode the master latch gets its input from *scan_input* SI . The consecutive application of scan clock SCK loads the test values into the scan flip-flops. As can be observed in Figure 6.3, when SCK gets to logic low (0), $T5$ and $T6$ get enabled and the value of SI is written into the master latch in a similar way to memory write operation. It should be noted that in test mode, since CP is always high(1), the feedback path transmission gate $T2$ always remains enabled. This makes the master latch always trying to retain its previous value. However, it can be observed in Figure 6.3, the test mode input path circuit force writes the SI value simultaneously at both input and output nodes of inverter $i1$ via buffer $i5$ and inverter $i6$ respectively. This makes the write process very fast as far as logical fighting is concerned. When the scan clock SCK gets high (1), the dynamic slave latch transmission gate $T7$ gets enabled, and the master latch starts driving both dynamic slave latch inverter $i7$, and functional slave latch inverter $i3$. This propagates the test value latched into the master during negative clock cycle, to dynamic slave latch, and to *scan_output* SO of the scan cell.

When scan clock SCK gets to logic low (0), $T7$ gets disabled, and the input parasitic capacitance of inverter $i7$ drives the successive scan cell's *scan_input* SI . Due to very high impedance of the inverter, the parasitic capacitance do not discharge immediately and takes a long time. The parasitic capacitance discharge time decides the minimum scan clock frequency at which scan shifting can be done. However, a lower shift frequency is undesirable as it increases the test time, which in turn increases the test cost. It should be noted that in test mode the transmission gate $T3$ always remains enabled. This keeps the functional slave latch always transparent during test mode and makes the output (Q) toggle every time whenever there is a change in master latch's state. Every master latch in scan chain gets its scan input from preceding scan flip-flop's SO output, except the very first master latch in the scan chain which gets its test input from a primary input pin. The scan output SO of the last flip-flop of the scan chain is connected to a primary output pin. The shifting of test vectors in scan chain is done using the dynamic

slave latch. Once the scan chain is loaded, test vector is launched via the functional slave latch. The test application process is elaborated in detail in the next section.

6.3 Application of Test Vectors

The proposed scan flip-flop allows to apply all kind of test vectors that can be applied using a conventional scan flip-flop. Before applying any kind of test vectors, scan chain integrity is verified by exercising scan flush test. Scan flush test is applied by propagating an all transition pattern, like 1100, through the scan chain without any response capture cycle in between. The scan clock SCK is always kept high during functional mode. When functional clock CP is high (1), falling edge on SCK switches the circuit from functional mode to test or scan mode. The functional clock CP is always kept high (1) during scan shift operation. On arrival of negative edge of SCK , the value of SI propagates into master latch via test input path. Next, the rising edge on SCK transfers the master latch value to dynamic slave latch and to the scan output node SO . By repetitive application of scan clock the flush pattern is propagated through the scan chain and is observed at primary output pin. The observation of correct input sequence at primary output pin verifies the integrity of scan chain. The scan integrity test covers all possible faults on the test mode input path and the input/output faults of dynamic slave latch as well as master latch. The scan integrity test does not cover input/output faults of functional slave latch. As we will see in the next subsection, these faults are covered during stuck-at fault test application.

6.3.1 Stuck-at-fault test

When clock CP is high (1), falling edge on scan clock SCK indicates the start of scan shifting. The *stuck-at-fault* test is applied by first loading the test vector via scan shifting path and then launching the test vector via functional slave latches. As explained earlier the functional slave latch always remains transparent during scan shifting process. So during the last *shift-cum-launch* cycle when negative edge at scan clock SCK comes, the

test vector is applied via output Q of functional slave latches. It should be noted that the test vector is launched in last shift cycle at negative edge of the scan clock SCK . After launch of test vector, the scan clock SCK is kept high (1) to disable the scan input path. Now, to capture the test response the functional clock CP is clocked once. When the functional clock CP gets low, functional response is latched into the master latch via functional input D . On arrival of positive edge on the functional clock CP , response is propagated into the functional slave latch as well as into the dynamic slave latch. Once the test response is captured, functional clock CK is kept at logic high (1) level. This disconnects the functional input D from the master latch. Now, falling edge on scan clock SCK switches the circuit operation from capture to shift mode. At the same negative edge on SCK , functional response stored into slave latch gets transferred to functional master latch of next scan cell via the scan input path. The unloading of test response is done with simultaneous loading of next test vector.

The timing diagram for *stuck-at-fault* test application is shown in Figure 6.4. As stated earlier, input/output faults of the functional slave latch are not covered by the scan chain integrity test. These faults are covered by *stuck-at-fault* test, as the test vectors are launched via functional slave latch. So, these faults if there exist any will manifest during the test vector launch cycle. The input/output faults of functional master latch will cause wrong test vector launch, which in turn will result into wrong test vector response. This way these faults will be detected at the end of test response unloading process. Since, the same set of test vectors are shifted and applied during *stuck-at-fault* test, the same set of faults are covered as in conventional scan design based test process.

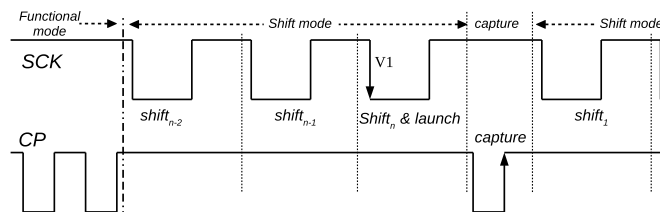


Figure 6.4: Launch and capture in *stuck-at* test

6.3.2 Launch-on-capture test

In *launch-on-capture (LOC)* testing, a test vector pair ($V1$, $V2$) is applied to the *CUT*. The first vector $V1$ initializes the circuit and the second vector $V2$ launches the transition. The launch of initialization vector $V1$ and first response capture is performed in a way similar to *stuck-at-fault* test. As we know that vector $V2$ is the functional response of vector $V1$ so, response capture of $V1$ also acts as launch of transition vector $V2$. The response of vector $V2$ is captured by applying an *at-speed* functional clock cycle. The loading/unloading of test/response is done in a way similar to *stuck-at-fault* testing. The timing diagram for the *Launch-on-capture* test application is shown in Figure 6.5.

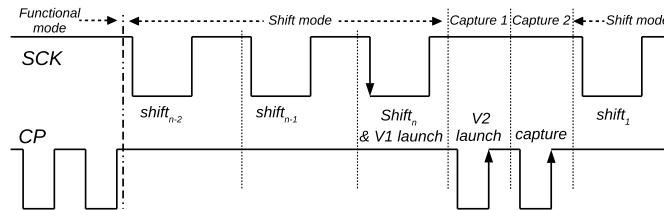


Figure 6.5: Launch of $V1$, $V2$, and capture in *LOC* test

6.3.3 Launch-on-shift test

In *launch-on-shift (LOS)* testing, transition vector $V2$ is one bit shift of initialization vector $V1$. In *LOS* test, response of $V1$ is not captured. As explained earlier $V1$ is launched at negative edge of scan clock SCK . As $V2$ is one bit shift of $V1$, $V2$ is launched at next negative edge of SCK . Now to capture the response of $V2$ *at-speed*, the scan clock SCK needs to be clocked at functional clock speed. To apply *LOS*, SCK must be timing closed. However it should be noted that SCK in the proposed design is an exact equivalent of *test_enable* signal TE in conventional scan flip-flop. In order to make SCK timing closed, that is also a global signal like the functional clock signal CK , it needs to be synthesized like a clock tree. That is a very costly task. Due to the high cost associated with *LOS* capable scan design, it is rarely exercised in industry. Application of *LOS* test even in conventional scan design is not possible without a timing

closed TE signal. The LOS based transition delay fault test, without a fast SCK (or timing closed TE in case of conventional scan design) can be exercised by using the $AND-OR-INVERT$ (AOI) circuitry proposed by Gefu et al. [71, 130]. Hence, neither conventional nor proposed design have the capability of exercising LOS based delay test. The proposed design does not use any extra control signal or testing constraint, and can be easily integrated into the existing DFT flow.

6.4 The Proposed Scan Flip-flop as a Joint-scan Cell

In a Joint-scan architecture, the serial and random access scan are integrated together to test the circuit. The serial part of Joint-scan and the random part both are operated simultaneously to load/unload the stimuli/response. The memory elements that are included into serial scan architecture are replaced by a serial scan flip-flop. The memory elements that are included into RAS test architecture are replaced by RAS cell. The proposed scan design eliminates the requirement of two separate scan cell libraries for serial scan cell and RAS cell. It provides a common scan cell that can be used for both

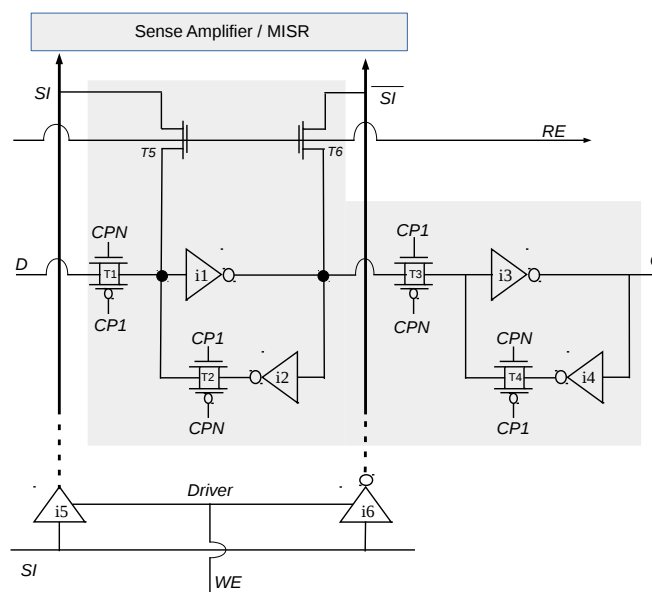


Figure 6.6: Progressive Random Access Scan ($PRAS$) Cell [10]

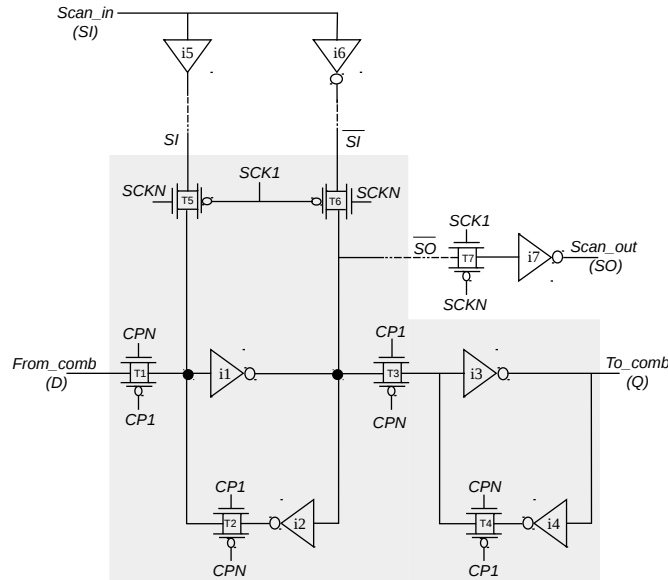


Figure 6.7: Proposed design as a base scan cell for Joint-scan architecture

serial scan and *RAS*. Schematic design of an area and performance efficient progressive random access scan cell proposed by Baik et al.[10] is shown in Figure 6.6. This *PRAS* cell is a modified design of a regular master slave based positive edge triggered *D* type flip-flop. The grey part in Figure 6.6 depicts the *PRAS* cell and the remaining circuit is part of *RAS* test architecture. Figure 6.7 represents the proposed flip-flop as a base scan cell. The grey box of the proposed flip-flop maps to the basic *PRAS* cell design except the access pass transistors are replaced by transmission gates (*T5*, *T6*). Both of these basic cells are functionally equivalent. Therefore, same basic cell of the proposed design can be used as a *PRAS* cell without modification. Note that the proposed scan cell has one extra output node *SO* which remains unconnected in *RAS* mode. The base scan cell can be synthesized as a *PRAS* cell by mapping the base scan cell in/out signals with corresponding *PRAS* in/out signals to use in *RAS* test architecture. Similarly, the base scan cell can be synthesized as serial scan cell with the full logic shown in Figure 6.7. It is worth to note that with a minor change in synthesis process the proposed scan flip-flop can be synthesized as both serial scan cell and *PRAS* cell. This can make the Joint-scan architecture implementation efficient and easy.

Table 6.1: Post Layout Timing Simulation Results at $500MHz$

Functional Mode			
Parameter ↓	Scan cell →	Conventional Scan Flip-Flop	Proposed Scan Flip-Flop
$Clock\ to\ Q\ (t_{cq})$		$0.332ns$	$0.378ns$
$Setup\ time\ (t_{setup})$		$0.400ns$	$0.290ns$
$Hold\ time\ (t_{hold})$		$0.0ns$	$0.0ns$
$t_{cq} + t_{setup} = t_{pd}$		$0.732ns$	$0.668ns$
$Time\ gain\ w.r.t\ Conventional$		$+64ps$	
Test Mode			
$Clock\ to\ Q\ (t_{cq})$		$0.332ns$	$0.284ns$
$Setup\ time\ (t_{setup})$		$0.400ns$	$0.545ns$
$Hold\ time\ (t_{hold})$		$0.0ns$	$0.0ns$
$t_{cq} + t_{setup} = t_{pd}$		$0.732ns$	$0.829ns$
$Time\ gain\ w.r.t\ Conventional$		$-103ps$	

6.5 Experimental Results

The post layout timing simulation of the proposed scan flip-flop has been carried out using *UMC's* $65nm$ technology at operating voltage of $1.2V$ for frequencies ranging from $500MHz$ to $1GHz$. The post layout timing diagram at a clock frequency of $500MHz$ has been put for reference in Figure 6.8. The timing diagram verifies the efficacy of the proposed design. The post layout timing simulation results are listed in Table 6.1. In functional mode, clock to Q (t_{cq}) delay, and setup time (t_{setup}) of the proposed design are found to be $378ps$, and $290ps$ respectively. It can be observed from Table 6.1, that t_{cq} of the proposed flip-flop is slightly higher than t_{cq} of the conventional flip-flop. This is due to the extra capacitive loading caused by dynamic slave latch. There is a considerable decrease in t_{setup} of the proposed scan flip-flop due to elimination of input multiplexer.

Overall the proposed design offers a time saving of $64ps$.

In test mode, propagation delay (t_{pd}) of the proposed scan flip-flop degrades by approximately $103ps$. However, test mode performance degradation is not of any concern as the test shifting is done at a frequency much lower than the functional frequency. Overall the time saving offered by proposed design is approximately equal to 3-4 inverters delay, where typical inverter delay in $65nm$ technology is approximately $15-18ps$.

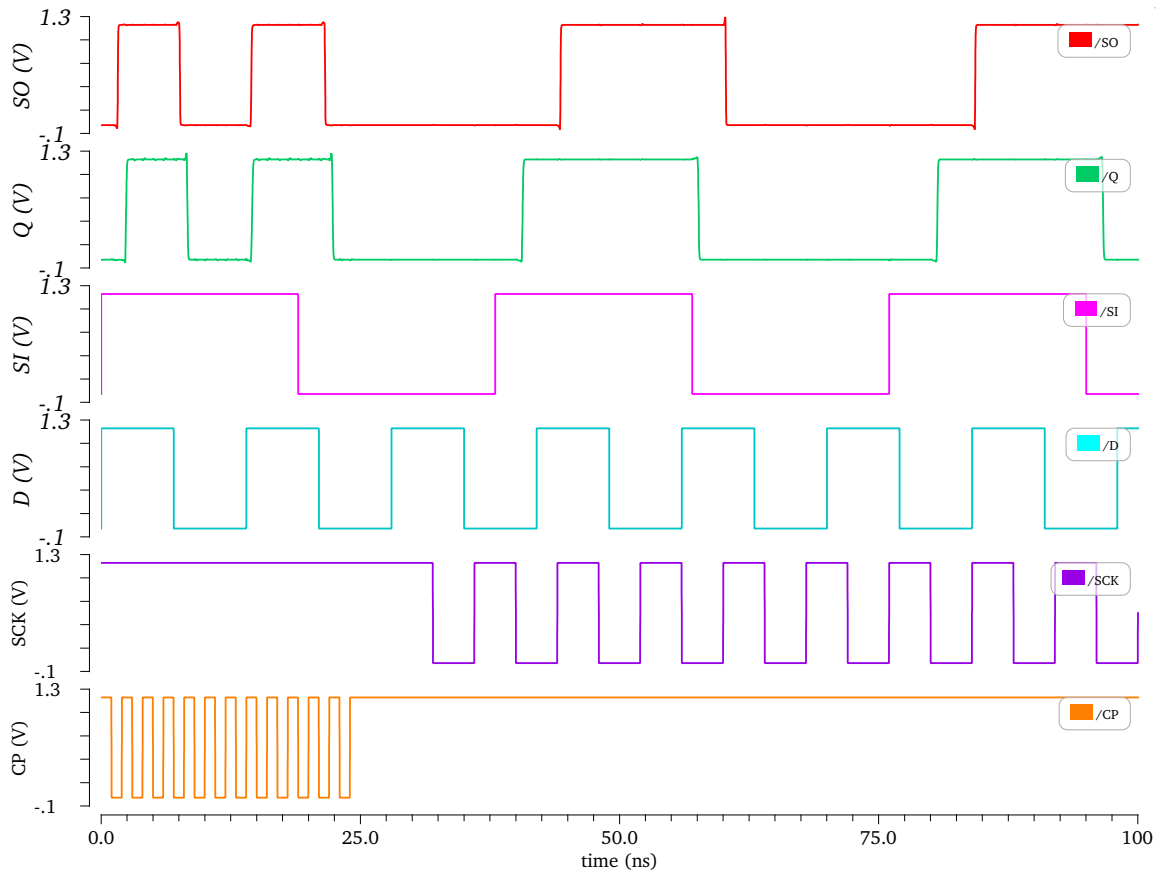


Figure 6.8: Post layout timing waveform at 500MHz

6.6 Conclusion

We have proposed a scan flip-flop design which eliminates the performance penalty of serial scan by removing scan multiplexer off the functional path. The proposed design uses a separate test mode input path and a low cost dynamic slave latch to scan shift

the test vectors. The new scan flip-flop is capable of applying all sort of conventional tests and fully comply with the conventional industry design and test flow.

Furthermore, the proposed scan flip-flop can be used as a serial scan cell as well as a random access scan cell, in Joint-scan test of *JScan* and *2M-JScan* architectures. This scan cell has now ease the process of scan insertion and synthesis because there will no requirement for two separate scan flip-flops for synthesis and design rule checking for flip-flops. The proposed scan flip-flop does not use any extra control signal other than the standard *Scan Enable* signal and has a minimal area overhead compared to conventional scan flip-flop. The Chapter has contributed a new scan flip-flop design which has dual advantages: solved the long-standing mux-delay problem on conventional scan flip-flop and rendered a ease for Joint-scan synthesis using this mixed-mode scan flip-flop.

— * — * —

Chapter 7

Worst Case Power Estimation

7.1 Introduction

Worst case power estimation is crucial for efficient design decision making. Determining the appropriate design for power distribution network (PDN) and the supply voltage (V_{dd}) value plays an important role for an efficient (high performance, low power, and reliable) circuit design. Determining these values during RTL level of design flow helps in design optimization. Furthermore, the worst case power estimation provides an upper bound on circuit activity which can be used as a threshold for test power dissipation.

7.1.1 Power Estimation Requirement and Challenges

With shrinking in size of a transistor (reduction in feature size) a large number of gates are packed in any digital logic circuits of current era, some of the system-on-chip (SoCs) like processor cores comes with billions of transistors at the latest 32nm or 22nm technology[53]. In other words, the transistor density per chip area are gradually increasing with this shrink phenomena. Coupled with higher operating frequencies, the dynamic power dissipation in these circuits has increased over the time. Also, the static power due to various leakages is increasing at exponential rate[58] (in this work we will be dealing with the dynamic power estimation). The increased power dissipation has developed concerns towards set of issues manifested due to excessive power consumption:

like formation of local hot-spots, burn-out of chip due to excessive thermal situation, long term reliability, and IR drop problem in modern day's integrated circuits. The variation in power dissipation also causes variation in delay which in turn has impact on efficient clock tree design. Therefore to deal with all these issues there is a need of effective power management in circuit. Though several solutions have been proposed to deal with these problems, the manifestation of these problems are largely due to two design parameters:

- Power distribution network (PDN)
- Supply voltage (V_{dd})

The design of power distribution network and selection of supply voltage are two important parameters for an effective circuit design. Hence for designing of optimized power distribution network and to determine supply voltage it is necessary to have prior knowledge of the maximum power that a circuit can dissipate. Therefore comes the need of prior estimation of worst case power. Apart from this the power estimation also helps in evaluating circuit reliability, and in decision-making for packaging and cooling requirement [76] [80]. In the design flow it is typically at the RTL level, where designers have access to RTL or gate level description of design, the power estimation is performed and the resultant parameters, worst case average and peak power, hot-spot, and IR drop, are supplied to PDN designer.

The power estimation basically relies on the following derivation.

$$\begin{aligned}
 P_{total} &= P_{stat} + P_{dyn} \\
 P_{stat} &= I_{stat}V_{dd} \\
 P_{dyn} &= \alpha CV_{dd}^2 f
 \end{aligned}
 \tag{7.1}$$

In this work our objective is dynamic power, P_{dyn} , estimation. The dynamic power is directly proportional to circuit activity (the charging and discharging of load capacitance C). This charging and discharging activity is accounted in the equation using the activity factor α . The power dissipation takes place only if the gate toggles from $0 \rightarrow 1$ (charging) and $1 \rightarrow 0$ (discharging). Often the activity factor is assumed to be $1/2$ for

the simplicity[121]. We have resorted on this fact that the worst case power is similarly proportional to the worst number of switching activity that takes place in the circuit. Therefor in this work our proposal is to estimate the worst case activity in the circuit.

Occurrence of activity in the circuit depends upon the signal transmission delay over circuit paths. In a circuit all the activity does not occur in same time, they occur over the span of time period depending upon the arrival of signal at the nodes/gates. Therefor the estimation must consider the timing delay for accurate estimation of worst case dynamic power. This timing consideration imposes a trade off between accuracy and computational complexity. Better the timing delay model accurate is the estimation and complex is the computation[49]. It is well researched subject that timing and power are interdependent. Irregular power dissipation often results into variation in timing, and for estimation of accurate power it is necessary to have timing delay model at first place. Therefor it is essential to consider both of these parameters simultaneously. Four different time delay models have been studied in the literatures[49]:

- Zero delay model:
Does not take gate delay or wire delay into account.
- Unit delay model:
Gate delay for all the gates are considered to be equal unit.
- Variable delay model-1:
Considers fan-out load and equal unit of output gate strength.
- Variable delay model-2:
Considers variable output gate strength along with variable delay model-1.

Power estimation poses two major challenges: 1) accuracy and 2) expensive computation. Estimation of worst case dynamic power requires activation of circuit to its maximum possible activity. The maximum possible activity includes functional toggle and glitch toggle. For accurate estimation each of the toggle has to be accounted. Here the challenge is how to figure out the input stimuli which can activate the circuit to its

maximum possible activity. Theoretically, the problem of finding such input stimuli is NP-Complete. In order to alleviate the computational problem the alternative way is to find out the toggle probability of each node/gate and use that to compute the power. In the later method however it is not guaranteed to be the maximum possible power, thereby the accuracy gets affected. Most of the tools currently resort to some variation of probabilistic method because it is time efficient. However the ideal goal is to estimate with certain acceptable accuracy.

The power estimation technique can primarily be categorized into two:

1. Vector less power estimation,
2. Vector based power estimation.

In vector less method generally the estimation is performed based on the analysis of physical properties of circuit parameter. These kind of technique generally uses probability to model the expected switching activity. Such kind of technique are very prompt in computing the power number however the accuracies are some time far from the reality. The second type of technique tries to bring the accuracy more closer to reality. Therefore it uses input vector pair or set of vectors for accurate estimation of power. This kind of technique although are more realistic in estimating the power they suffer from longer simulation time. Hence a motivation to conduct research in this direction is to find out an efficient technique to generate an optimal input vectors which can stress the circuit to maximum. Recently, some of the methodology have been proposed to enhance the efficiency of vector based methodology [118] [49] [92].

The power estimation generally are carried out at different level of design abstraction. The power estimation at RTL level is computationally easy compared to gate level estimation and gate level estimation is computationally easy compared to estimation at device level. As the abstraction level goes down towards fine granularity the accuracy increases at the cost of computational time [76]. In this work the power estimation is restricted at the gate level abstraction. The method we have proposed in this work is

vector-based approach. The methodology provides an efficient way of generating input stimuli which triggers worst possible activity in the circuit.

7.1.2 Contribution

The proposed power estimation methodology explores the *vector-based approach*. The proposed methodology is limited to the estimation for combinational circuit alone. The objective of proposed work is twofold:

1. to improve the efficiency of generating input stimuli,
2. to maintaining the accuracy by using appropriate delay model.

The problem of finding a pair of input stimuli has been modeled as *Binary Integer Linear Programming (BILP) Model*. The BILP formulation is based on the gate level behaviour of circuit. Two individual BILP problems are defined. The first problem is defined to estimate the power considering zero delay model, this essentially estimates the worst case toggles in circuit assuming that all the nodes toggle simultaneously. The estimated power can be termed as *instantaneous peak power*. The second problem considers the unit delay timing model and thereby it estimate the *level-accurate instantaneous peak power*. The second methodology distinctly account for the toggles that take place at certain delay interval. For both the problems respective objective functions and set of constraints are defined. Both the problems are solved with IBM CPLEX [31] ILP solver with necessary modification wherever needed. Experiment is performed on ISCAS-85 combinational circuits. Experimental results show that for most of the circuits the ILP solver could solve within stipulated time limit thereby confirms the worst case input stimuli. However for certain circuit the solver could not find the optimized solution thereby it solved for nearest optimum solution. Following summarises the contribution of this chapter:

1. Instantaneous peak power estimation
 - BILP formulation with zero delay model

- Improvement over accuracy
2. Level-accurate instantaneous peak power estimation
 - BILP formulation with unit delay model
 - Improvement over accuracy

7.1.3 Chapter Organization

The remaining sections are organised as follows. Section 7.1.4 gives an overview of previous works. In Section 7.3 level-accurate power estimation methodology is elaborated. Section 7.2 discusses a background on integer linear programming. The experimental results for first methodology is tabulated and discussed in Section 7.4. A brief summary of the first methodology is discussed in Section 7.5. The second methodology is presented Section 7.6. Level-accurate BILP problem formulation is elaborated in Section 7.7. In Section 7.8 the experimental results on ISCAS-85 benchmark circuits and analysis of results is presented. The chapter on power estimation concludes in Section 7.9 with vision for future work.

7.1.4 Previous Work on Power Estimation

Accuracy and computational efficiency are the two motivations for research on power estimation methodology. Two directions has been explored: *Input vector based simulation* and *Vector less probabilistic method*. The simulation based approach generally look for a suitable input vector pair or a set of pairs to trigger the circuit activity to maximum possible value. This kind of approach is motivated by accuracy of the estimation, however such technique suffer from solving an NP-Complete SAT problem to generate input patterns. On the other hand probabilistic approach perform static analysis of circuit based on toggle probability at each node. Such approach are time efficient since static analysis of circuit is a liner time problem. However the accuracy is more pessimistic, which leads to power expensive design[80].

A vector based approach proposed by S Devadas et al.[33], formulated as *weighted max-satisfiability(SAT)* problem, is one of the early attempts. The circuit was represented as *max-satisfiability(max-SAT)* problem and then was solved using exact and approximation SAT solver. The SAT formulation was modified as weighted SAT problem to consider load capacitance. One difficulty with SAT problem is computation time.

The idea of formulating the problem as *Pseudo-Boolean SAT* is explored recently in [69, 92, 106]. H Mangassarian et al.[69] and Sagahyroon et al. [91] have formulated the problem as Pseudo-SAT problem to estimate worst case power and worst case power-up current. The reported results show that the Pseudo-SAT based technique is efficient when the problem is solved using commercial SAT solver like CPLEX[31]. Motivated by the efficiency of CPLEX solver and trade-off between ILP and Pseudo-SAT [65], we have in our previous work [113] explored an alternative idea of formulating the problem as ILP and solving it using CPLEX.

Some what different idea, brought from ATPG(Automatic Test Pattern Generation) technique, is proposed by Chuan- Yu Wang et al. [117] [118] in 1996. Two methodologies were proposed in [117] for determining lower and upper bounds on maximum power dissipation. To calculate the lower bound, the authors have proposed an ATPG based technique; while for upper bound they have proposed a *monte-carlo* based simulation technique. An improved version, based on D-Algorithm, of this technique is proposed in [118]. Recently the similar ideas have been explored by Najeeb et al for peak dynamic power estimation[54]. A controllability driven input vector generation is explored by Shyamala et al[105].

Genetic Algorithm based approach is examined by Hsiao et al[49]. In that work the estimation has been done for three different power matrices: instantaneous peak power, n-cycle peak power, and for sustainable peak power. The accuracy in this work has been accounted by employing four different type of delay models: zero delay model, unit delay model, variable delay model-1 (considers the fan-out with equal weights), and variable delay model-2 (considers the fan-out with variable driving gate size).

M Pedram and Q Wu et al. have explored probabilistic based ideas [82][123]. Q Wu

et al.[123] have proposed a technique based on limiting distribution of extreme order statistics. The challenge in this approach is to show with higher confidence level that the estimated power is indeed the real power. As stated earlier such kind of techniques are time efficient. However difficulty is there in accuracy. More detailed survey of the earlier techniques are carried out by Farid N. Najm[76] and M Pedram[80].

Keeping accuracy and efficiency in forefront in this work we have proposed an alternative approach to formulate the problem as an ILP problem and solved it using the standard solver. The proposed method show that for some of the circuit the solver is able to find the optimal solution. Whereas it is also been observed, for some other circuit the solution is near optimal within the range of 80 - 90% in a given time limit.

7.2 Background on Binary Integer Linear Program

Binary integer linear programming (BILP) is a class of optimization method called linear programming (ILP) method. In BILP the variables are constraint to take only 0 or 1 value—a binary digit. The BILP problem is to optimize a given function called *objective function* subject to a given set of constraints. The general form of BILP can be stated as follows:

$$\text{Minimize : } f(x), \quad \text{Subject to : } Ax = 0, \quad x \in [0, 1]$$

where $f(x)$ is objective function, $Ax = 0$ is a constraint. This a particular type of linear programming where variables are constraint to take the integer values 0 or 1.

7.3 Power Estimation Considering Zero-delay Model: Formulation of BILP Problem

In this work we have proposed a input vector based methodology to estimate the worst case power. We call this methodology as input vector control method (IVC). The challenge in this methodology is to identify the input vectors which results into the worst power dissipation on given delay model. For any combinational circuit these two input vectors can be applied consecutively to trigger activity in the circuit. These toggle activity thereafter are recorded to estimate the power dissipation. In this work we have considered these toggle activity as the final power number. The consideration here is that since the dynamic power is directly proportional to toggle activity, the toggle count would suffice for comparison.

The formulation of Integer Linear Programming (ILP) problem is based upon the representation of combinational circuit in linear algebraic form. The linear algebraic representation resembles the boolean algebraic form with few additional constraints and the consideration of value 0 and 1 as integer value alike the binary in case of boolean algebra. The formulated ILP problem can be called as 0-1 integer linear programming (0-1 ILP) problem. There are three sets of constraints defined in order to represent the circuit functionality. The constraints are as follows:

- Input Output constraints (I/O constraints)
- Linearization constraints
- Toggle constraints

The I/O constraints express the input-output relation between the gates. The linearization constraints convert a nonlinear product terms into linear sum of variables, and thereby it transform the nonlinear integer program into linear integer program. The toggle constraints are used to reduce the computation time taken by the solver to find near optimal solution.

7.3.1 Input Output Constraints

The relation between gate inputs and outputs is encoded using a set of constraints. We have used Binary Integer Programming (BIP) to represent NOT, NAND, and NOR gates. Let the inputs of a n-input gate be x_1, x_2, \dots, x_n and its output be y . For NOT, NAND and NOR gates, the relations can be expressed as follows:

$$\begin{aligned}
 NOT : y &= 1 - x_1 \\
 NAND : y &= 1 - \prod_i x_i \\
 NOR : y &= 1 - \sum_i x_i + \sum_{i,j} x_i \cdot x_j + \dots + (-1)^n \prod_i x_i
 \end{aligned} \tag{7.2}$$

Similar expression can be derived for other type of gates, namely AND, OR, and XOR.

For simplicity we have synthesised the RTL with two input NAND, two input NOR, and Inverter gates only. The constraint for two input gates can be expressed as follows:

$$\begin{aligned}
 NOT : y &= 1 - x_1 \\
 NAND : y &= 1 - x_1 \cdot x_2 \\
 NOR : y &= 1 - (x_1 + x_2) + x_1 \cdot x_2
 \end{aligned} \tag{7.3}$$

7.3.2 Linearization Constraint

The constraints in the section above consist of non-linear functions where output variables are product of other variables. Since an ILP solver requires linear programming model, we need to linearize these equations. We rewrite the product constraints as linear functions by introducing a product variable and three additional constraints for each

product variable. For example, consider the following constraint:

$$\begin{aligned}
 y &= 1 - x_1 \cdot x_2 \\
 p &= x_1 \cdot x_2 \\
 y &= 1 - p \\
 \text{s.t. } p &\leq x_1 \\
 \text{s.t. } p &\leq x_2 \\
 \text{s.t. } x_1 + x_2 p &\leq 1
 \end{aligned} \tag{7.4}$$

The above three equations will give the product p when $x_1, x_2 \in [0, 1]$.

7.3.3 Toggle Constraints

Toggle constraints are defined for each net to measure the toggling event. The two inputs XOR function is used to measure the toggling of a net. The two inputs in this case are nothing but the values triggered by two input vectors. For example for any net x we define the following five variables to capture the toggle activity.

$$\begin{aligned}
 x_a &= x \\
 x_{abar} &= 1 - x_a \\
 x_b &= x \\
 x_{bbar} &= 1 - x_b \\
 x_{toggle} &= x_a \cdot x_{bbar} \oplus x_{abar} \cdot x_b
 \end{aligned} \tag{7.5}$$

The non-linear terms, $x_a \cdot x_{bbar}$ and $x_{abar} \cdot x_b$, in x_{toggle} variable are again linearized using linearization constraints.

7.3.4 Objective Function

The primary objective is to compute the worst case dynamic power i.e. the circuit must toggle to maximum. Therefore, the formulation is done in such that the sum of toggle

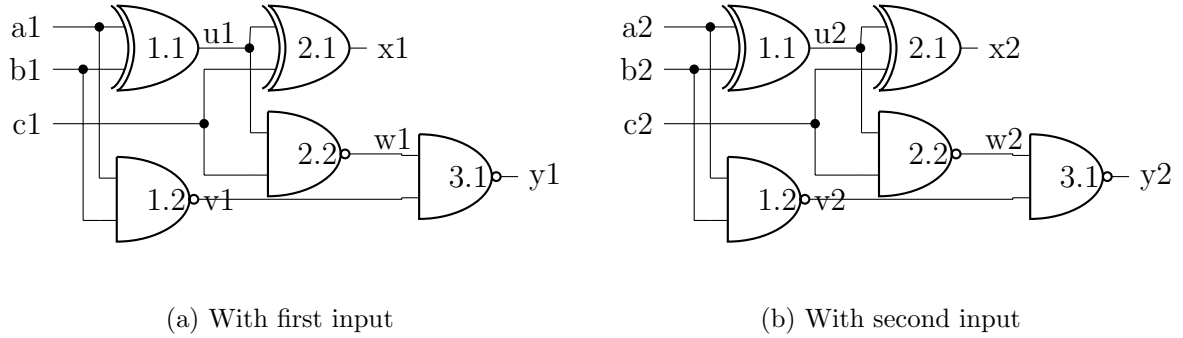


Figure 7.1: Demonstration circuit for ILP formulation with zero-delay model

activity across each node (boolean gate) output would be maximum. The objective function can hence be defined as follows:

$$f_t(x) = \sum_{i=0}^{n-1} x_{toggle}^i \quad (7.6)$$

where x_{toggle}^i is a toggle variable for i_{th} node of the circuit having total of n nodes indexed from 0 to $n - 1$.

With the above three sets of constraint and the objective function following binary integer linear program is formulated.

$$\text{maximize : } f_t(x),$$

Subject to :

$$1) \text{ I/O constraints} \quad (7.7)$$

2) *Linearization constraints*

3) *Toggle constraints*

7.3.5 Example

In this example the BILP problem formulation considering the zero-delay model is demonstrated. With reference to Figure-7.1, the Figure (a) and (b) shows a circuit with two inputs. The objective of this formulation is to compute $\langle a1, b1, c1 \rangle$ and $\langle a2, b2, c2 \rangle$

pair of input vector such that the circuit will display it's highest activity.

First of all the input-output constraint set will be formulated such that the circuit functionality will be ensured to be correct. Next, the linearization constraints will be formulated. Thereafter the switching constraints (or toggle constraints) will be formulated to account for the switching activity. Finally the objective function will be formulated. Following are the constraints and objective function for circuit shown in Figure-7.1.

I/O constraints): For XOR gate 1.1 the boolean output expression will be:

$$u1 = a1.\overline{b1} + \overline{a1}.b1$$

To transform it to equivalent algebraic expression, first the XOR will be expressed in terms of NAND-INV circuit. Hence the resultant algebraic expression would be:

$$\begin{aligned} u1 &= (1 - (a1.(1 - b1)).(1 - (1 - a1).b1) \\ &= (1 - m1).(1 - n1) \\ &= k1 \end{aligned}$$

In above expression the $m1n1$ and $k1$ are non-linear component, therefore a linearization constraint to be introduced for such variable.

Linearization constraints:

$$\begin{aligned} m1 &= a1.(1 - b1) \\ n1 &= (1 - a1).b1 \\ m1 &\leq a1 \\ m1 &\leq 1 - b1 \\ a1 + 1 - b1 &\leq 1 \\ n1 &\leq 1 - a1 \\ n1 &\leq b1 \\ 1 - a1 + b1 &\leq 1 \end{aligned}$$

Similarly for $k1 = (1 - m1).(1 - n1)$ following are the linearization constraint:

$$\begin{aligned} k1 &\leq 1 - m1 \\ k1 &\leq 1 - n1 \\ 1 - m1 + 1 - n1 &\leq 1 \end{aligned}$$

Finally the constraint set for XOR-1.1 can be expressed as below:

$$\begin{aligned}
 u1 &= k1 \\
 s.t. \quad m1 &\leq a1 \\
 s.t. \quad m1 &\leq 1 - b1 \\
 s.t. \quad a1 + 1 - b1 &\leq 1 \\
 s.t. \quad n1 &\leq 1 - a1 \\
 s.t. \quad n1 &\leq b1 \\
 s.t. \quad 1 - a1 + b1 &\leq 1 \\
 s.t. \quad k1 &\leq 1 - m1 \\
 s.t. \quad k1 &\leq 1 - n1 \\
 s.t. \quad 1 - m1 + 1 - n1 &\leq 1
 \end{aligned}$$

Similarly all the gates will be expressed as constraint set. Now, to compute the switching activity five toggle variables and corresponding constrain set are to be defined.

Following are the toggle variables: **Toggle variable and constraints:**

$$\begin{aligned}
 \text{let } t1 &= u1 \oplus u2 \\
 t2 &= v1 \oplus v2 \\
 t3 &= w1 \oplus w2 \\
 t4 &= x1 \oplus x2 \\
 t5 &= y1 \oplus y2
 \end{aligned}$$

All the above variable will be further expressed as ILP constraints in similar way to I/O constraints. Since this is a digital logic circuit all the variables has to be ensured to take only either zero or one value. Finally the objective function will be expressed as:

$$Max(t1 + t2 + t3 + t4 + t5)$$

Now, this binary integer linear programming problem (BILP) can be solved using any of the standard ILP solver.

7.4 Evaluation on ISCAS-85 Circuits

For examining our proposition, we have used benchmark circuits from ISCAS-85. These benchmark circuits are simplified into a combination of two input gates as could be used

Table 7.1: Properties of ISCAS-85[47] Benchmark Circuits

Circuits	Functionality	#PI	#PO	#Gates	#Nets	#Fanout
c04	NAND-INV tree	1	4	4	100	0.09
c05	3-lvl NAND-NOR	3	8	7	87.5	0.05
c11	NAND-NOR re-convergent	3	6	5	83.3	0.02
c14	NAND re-convergent	2	4	3	75	0.03
c17	NAND re-convergent	5	6	6	100	0.02
c432	interrupt controller	36	296	243	82	8.86
c499	SEC	41	626	360	57.50	3406.00
c880	8bit ALU	60	592	491	82.93	21.55
c1355	32bit SEC	41	690	415	60.14	3616.48
c1908	16bit SEC/DED	33	1291	832	64.44	2371.53
c2670	12bit ALU and control	233	1925	1362	70.75	559.65
c5315	9bit ALU and selector	178	4897	2635	53.80	3399.05

SEC : single error correction, DED : double error detection

in standard library. We have used CPLEX solver[31] for finding optimal input vector patterns. The TABLE I summarizes the percentage of nets that toggle with the resulting input vector pairs. In the table the MaxToggle column represents the maximum possible toggle the circuit can have i.e. equivalent to the total number of internal nets including output nets. The estimated toggle resulted from the CPLEX solver is tabulated in #Toggle column. The %Toggle column show the % of nets toggle with respect to the maximum number of nets. The number of toggle count does not consider the fan-out nets.

We observe that for small circuits the solution results into around 100% to 80% toggles. However for larger circuits the said toggle percentage varies from 80% down to about 50% depending on the circuit topology and complexity. This is quite expected because as the circuit complexity increases the dependency among the nets also increase.

The computations are performed on Intel Xeon CPU operating at 2.0GHz. In the TABLE I the solver run time is reported in Time(Sec.) column. It can be observed that some of the circuits have executed faster where as other have taken huge amount of time. The longer time is due to the nature of ILP problem being a NP-Complete type.

The experimental results show that the proposed method could able to solve the large circuit in reasonably lesser time. For some bench mark circuit like c432 the proposed method is performing better compared to SAT-based technique[91].

The CPLEX solver for some of the large circuit took more than 7-10 hrs of cpu time. Such circuits are a particular kind of circuits where the re-convergent fan-out are present in large quantity. This is one of the suspected causes for larger computation time. For such circuit we have chosen the nearest optimal solution for a given time bound.

Table 7.2: % of toggling nets as computed by CPLEX solver

Circuits	#PI	MaxToggle	#Toggle	%Toggle	Time(Sec.)
c04	1	4	4	100	0.09
c05	3	8	7	87.5	0.05
c11	3	6	5	83.3	0.02
c14	2	4	3	75	0.03
c17	5	6	6	100	0.02
c432	36	296	243	82	8.86
c499	41	626	360	57.50	3406.00
c880	60	592	491	82.93	21.55
c1355	41	690	415	60.14	3616.48
c1908	33	1291	832	64.44	2371.53
c2670	233	1925	1362	70.75	559.65
c5315	178	4897	2635	53.80	3399.05

7.5 Summary of Zero-Delay ILP Methodology

Any large combinational circuit can be represented as a binary integer program. An ILP solver is demonstrated to be useful in generating input vector pairs for toggle maximization. The cuts method speeds up optimal polyhedron computation over simplex methods for large circuits. We comprehend that SAT solvers have shown promise in solving of binary satisfiability problems. For more realistic consideration of practical circuits, we plan to incorporate glitches due to contamination delay and propagation delay. Also the estimation can improve when we increase weightage for high fan-out gates in our formulation. The model can be extended to consider pipelining in sequential circuits using local maximization and more number of states.

7.6 Level-Accurate Power Estimation: Methodology

As it is discussed in Section 7.1.4, all the power estimation techniques, particularly the peak power estimation techniques, proposed so far perform estimation of the peak power considering the nodes from entire circuit. Such approach leads to pessimistic results because all the circuit nodes does not switch simultaneously. Even though different delay models, zero delay, unit delay, type-1 variable delay, and type-2 variable delay models[49], are taken in to consideration to capture the spurious activity like glitches and hazards this leads to pessimistic value since the toggles are accounted from entire circuit. Our observation is that, since technology is shrinking towards nano meter geometry it is sufficient to estimate the worst case peak activity in a particular level to capture the information on IR-drop, ground bounce, and associated delays.

The proposed *level-accurate* method is demonstrated in following sections.

7.6.1 Leveling of the Circuit

The different levels of the circuit is determined by assigning the level number at each net. The assignment of level number is performed on per gate basis. For each gate the

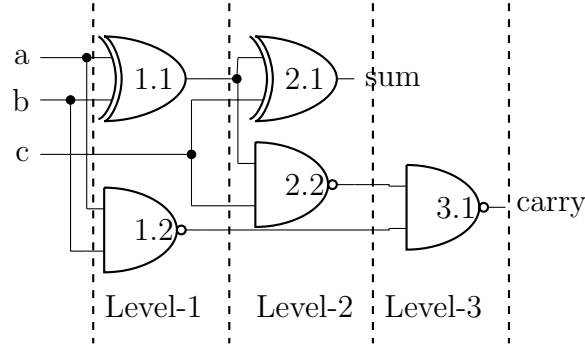


Figure 7.2: Group of gates leveled based on the circuit depth

level number of output net is determined based on the current levels of input nets by following simple formula.

$$L_{out}^i = \text{Max}(L_{in1}^i, L_{in2}^i, \dots) + 1 \quad (7.8)$$

where L_{out}^i is the level number of output net of i_{th} gate, L_{in1}^i is the level number of input net-1 of i_{th} gate.

In figure Fig-7.2 the leveling of each gates along with output net is shown as Level-1, Level-2, and Level-3. For example, for the circuit in Fig-7.2 the level number of output nets of gate-1.1(XOR gate) and gate-1.2(NAND gate) is 1, in Fig-7.2 it is shown as Level-1. This is computed using the formula in Eqn-7.8 by assigning $L_{in1}^{1.1} = 0$ and $L_{in2}^{1.1} = 0$ since they are primary input nets. Hence $L_{out}^{1.1} = 1$. The same formula can be applied progressively to determine rest of the levels.

7.6.2 Power Estimation Methodology

The proposed methodology estimates the maximum instantaneous power by determining the maximum activity among the activity at each levels of the circuit. Mathematically this can be expressed as follows.

$$P_{peak} = \text{Max}(A_l) \quad (7.9)$$

where P_{peak} is the estimated worst peak power and A_l is the maximal activity at l^{th} level of the circuit. Detail is explained in the following section, Section 7.7.

7.7 Formulation of Level-BILP

Table 7.3: Summary of BILP as presented in Section-7.3

Function Type	Examples
I/O Constraints	$NOT1 : y = 1 - x_1$ $NAND2 : y = 1 - x_1 \cdot x_2$ $NOR2 : y = 1 - (x_1 + x_2) + x_1 \cdot x_2$ where $y, x_i \in [0, 1]$
Linearization Constraints	$y = 1 - x_1 \cdot x_2$ $p = x_1 \cdot x_2$ $y = 1 - p$ $s.t. p \leq x_1, p \leq x_2, \text{ and } x_1 + x_2 - p \leq 1$
Toggle Constraints	x_n x_{n+1} $\bar{x}_n = 1 - x_n$ $\bar{x}_{n+1} = 1 - x_{n+1}$ $x_{toggle} = x_n \cdot \bar{x}_{n+1} + \bar{x}_n \cdot x_{n+1}$
Objective Function	$Maximize \sum_{i=1}^N x_{toggle}^i$

Formulation of BILP for each level is needed to determine the level-max peak activity and the corresponding input vectors in a modified way. For example with respect to Fig-7.2, to determine the peak activity at Level-1 the toggle variables at out-put net of XOR and NAND gate only has to be included in the objective function. Hence for Level-1 the

objective function as given in Table 7.3 can be rewritten in following way.

$$\text{Maximize } \sum_{i=1}^2 x_{toggle}^{1,i} \quad (7.10)$$

where $x_{toggle}^{1.1}$ and $x_{toggle}^{1.2}$ are toggle variable corresponding to XOR and NAND gate at Level-1.

Since here the objective is to determine the peak activity at Level-1 the rest of the circuit i.e. Level-2 to Level-3 can be ignored. Therefor the constraint set now will only consider the gates from Level-1.

Similarly when we consider the Level-2 the objective function will optimize the toggle variables corresponding to Level-2. And the constraint set will include the gates from Level-2 back to Level-1. Rest of the gates from Level-3 onwards can be ignored.

Hence from the above demonstration we observed that during the maximization at Level-j(j is the current level at which maximization is being performed) the circuit from Level-[j+1] to Level-L(L is the maximum number of levels) is in don't care state and circuit from Level-1 to Level-j is in active state or care state. Hence the BILP can now be applied only to the part of circuit that is in active state.

Hence the modified objective function can be expressed mathematically as follows.

Formulation of Level-objective Function:

$$\text{Maximize } \sum_{j,i} x_{toggle}^{j,i} \quad (7.11)$$

where j varies from 1 to L where L is the total number of levels and i varies from 1 to n where n is total number of nets in a particular level.

Constraint set:

All the three constraints, I/O constraints, linearization constraints, and toggle constraints (as given summarised in Table-7.3), are defined only for that part of the circuit which is currently in active state, which means that gates from Level-1 to Level-j are

considered and rest are ignored.

7.7.1 Computation of Maximum Peak Activity

Now we are all set to compute the maximum peak activity using *level-BILP* formula. The peak power is expressed as Eqn-7.9. More precisely, the final peak activity is the maximum among the peak activity at each level. Hence by iterative execution of *level-BILP* for every level the level peak activity, A_l in Eqn-7.9, can be computed. In programing structure this can be written in following way.

for $l = 1$ to L

$$A[l] = \text{Maximize} \sum_{l,i} x_{toggle}^{l,i}$$

Now the peak power will be computed using Eqn-7.9 as $P_{peak} = \text{Max}(A[l])$.

7.8 Evaluation on ISCAS-85 Circuits for Level-accurate BILP

To evaluate the accuracy and efficiency of the proposed estimation methodology, we have experimented on ISACS-85 circuits set. We have solved the binary integer linear program(BILP) using CPLEX ILP solver[31]. CPLEX solver is found to be faster for BILP problem[92]. The benchmark circuits are standardised to one-input and two-input gates.

The experimental results are shown in Table 7.4. Complexity of the circuit is reported in 2nd and 3rd columns, column #PI indicates the number of primary inputs and #Ckt Nets indicates the total number of nets in circuit. The 4th, 5th, and 6th columns reports the activity count. The column #Max Level nets indicates the maximum number of nets across the levels, column *Est Swit act* indicates the estimated switching activity computed by the proposed level accurate method, and columns % Swit act reports the percentage of estimated switching activity over maximum number of level nets provided

Table 7.4: Comparison of estimated switching activity and simulation time of CPLEX solver

Bench Ckts	Level accurate (Unit-delay)						Zero-delay[113]			
	# PIs	# Tot nets	#Max Level nets	Est Swit act	% Swit act	Time in sec.	Est Swit Act	# Tot level	Est Avg Switact	Time in sec.
c01	4	9	4	4	100	0.04	9	4	2.25	0.01
c04	1	4	3	3	100	0.02	4	2	2	0.09
c05	3	8	3	3	100	0.03	7	3	2.33	0.05
c11	4	6	4	4	100	0.03	5	3	1.66	0.02
c14	2	4	2	2	100	0.03	3	3	1	0.03
c17	5	6	2	2	100	0.04	6	3	2	0.02
c432	36	296	27	27	100	9.25	243	42	5.78	8.86
c499	41	626	194	76	39.17	34.45	360	25	14.4	3406.00
c880	60	592	58	58	100	22.34	491	39	12.58	21.55
c1355	41	690	194	76	39.17	44.92	415	27	15.37	3616.48
c1908	33	1291	92	92	100	269.2	832	52	16	2371.53
c2670	233	1925	147	144	97.95	545.7	1362	56	24.42	559.65

in *# Max level nets*. The last four columns are results from the previous work reported in [113] and included in Section-7.3 in this thesis. The comparison indicates that the zero-delay model is highly inefficient in computing the worst case peak power when compared with that of the level-accurate model.

For all the circuits the CPLEX Solver could solve in no more than 546 second which is around 7 time faster than the worst time in [113]. For circuit c499 and c1355 the simulation time is improved. Since in the case of earlier work the ILP problem is formulated for entire circuit which causes around 690 variables and 11K of constraints to be dealt with where as for the level based approach the ILP is formulated for each level and are solved independently in which case only 194 variables and 9K constraints are involved. The estimated toggle activity by proposed method is less than that of the earlier method. For circuits from c432 to c2670 the estimated toggle activity is significantly better.

The c499 and c1355 are both error correction and translation(ECAT) circuits. The % of toggle activity for these circuits are low compared to rest of the circuits because the ECAT circuit consists of xor tree with re-convergent fanout. The % of toggle indicates that i) if it is 100% then the solver has found the solution, ii) if it is less than 100% and solver run time reaches it's time out limit than the solution is approximate, iii)if the solver finishes before time out than the solution reached the maximum achievable toggle activity.

7.9 Conclusion and Future Work

In the proposed method peak power estimation in terms of toggle activity count using level based approach is explored. The proposed method generates input vector pair which maximizes the toggle activity for a level. Through experimental results and theoretical analysis we have demonstrated that the level based approach is better than whole circuit approach. The CPLEX solver runs 10 to 50 time faster for this approach than the whole circuit approach.

Consideration of variable delay model along with fanout load could make the estimation realistic. Recently we have worked on this problem further where we have considered the variable delay and fan-out load[45]¹. Extending the work to sequential circuit is a practical necessary and challenge. Hence we set forth the above as next step of this work.

— * — * —

¹The work in not included in Thesis

Chapter 8

Conclusion and Future Work

Power issues in the modern complex SoCs are unavoidable. As the design density and complexity is growing at exponential rate the associated problems demand innovative solutions. With this understanding the thesis studied the test power and power estimation issues and proposed a set of innovative solutions. The test power has been a standing-problem since last decade or so. We have studied/reviewed the methodology proposed since 1980 to till date. Based on the assessment and examination on limitations of the available methodology the thesis has explored two classes of methodology to minimize the test power. The thesis has proposed a graph theoretic formulation for scan chain reordering and minimum shift in scan chain and new DFT architecture. Another problem that we studied is estimation of worst case dynamic power. Dynamic power being primarily determined by the switching activity in circuit, it poses a greater challenge to attain accuracy at the cost of computation time. For this reason most of the modern tools go for approximate approach and use static power estimation methodology. The thesis attempted to resolve this limitation and explored a input vector pair based dynamic power estimation while considering the computational complexity. Following sections outline the contributions of this thesis.

8.1 Contributions

Chapter-3 and Chapter-4 presented a graph theoretic methodologies to minimize peak power and test time in serial scan chain. The specific issues due to excessive dissipation of peak power are discussed in detail. The graph theoretic formulation for scan chain reordering to minimize peak power is presented. Algorithm to compute the optimum scan path is presented and the experiment on ISCAS-89 and ITC-99 circuits is carried out. Similarly, the idea of limited scan shifting is formulated as a graph theoretic problem and an approximate algorithm is proposed to find the desired scan chain ordering. Experiment is performed on ISCAS-89 circuits validate the effectiveness of proposed algorithm.

From these two study following outcomes are to be noted:

- graph theoretic formulation provided a formal structure to the scan chain reorder methodology particularly for peak power and scan shift minimization;
- experimental results reveals reduction of around 30% in peak power and around 15% in test time;
- the methodology does not involve any additional DFT circuitry;
- the scan chain reordering methodology does not change the placement of scan flip-flops rather it only reconnects the scan paths;
- it does not influence the design timing closure however it may affect the scan path timing closure;
- the reordering of scan chain also affects the transition delay fault coverage either positively or negatively;
- the reordering of scan chain needs unspecified test patterns (test patterns with don't care bits);
- as a side effect of peak power minimization the average power also get affected;

- the scan chain reordering for reducing shift operation also reduces the total number of toggle activity, thereby the power also is minimized;
- the test pattern reordering for reducing test time is complimentary to scan chain ordering;

Chapter-5 contributed two new DFT architectures. In this work we have systematical explored the methodology to combine random and serial scan architecture, an attempt is made to derive the benefits of each scan architecture in *Joint-scan*. The research is centered around the following problems:

- organisation of random and serial scan as single architecture,
- defining test functionalities,
- designing of test control mechanism,
- grouping of scan flip-flops into random and serial scan.

The research reveals the following facts:

- systematic integration of serial and random scan results in minimization of test time and data volume further from state of the art scan methodology;
- minimization of power dissipation to around 50% to 60% is sufficient to ensure no test escape and no damage due to peak or average power;
- further, the power slack can be explored to reduce DFT area overhead and routing congestion;
- dimension- row/column size of random scan and length of serial scan chain impacts the test time and data volume;
- column address application mechanism influences the test time and data volume, experiment on serial vs parallel address application indicates that fully parallel column address produces better results (at least this is true for the implemented *Joint-scan* architecture, this observation may change for different *Joint-scan*);

- test pin count has greater influence on test power, test time, and data volume, as the pin count increases the Joint-scan efficiently take advantages to further optimize these parameters;
- experimental results reveals that the proposed Joint-scan architecture produces better results for larger design;
- efficient grouping, the criteria and algorithm, of scan flip-flops play a crucial role in optimizing all the parameters: test time, data volume, test power, DFT area overhead, and routing congestion;

Chapter-6 proposed a new scan flip-flop to overcome the critical path delay problem in traditional multiplexer based scan flip-flop. The proposed scan flip-flop also served as a mixed mode scan flip-flop for Joint-scan architecture. The proposed scan flip-flop is verified by post layout timing simulation at 65nm technology node from *UMC* library. Following observations and contributions are to be noted:

- the proposed scan flip-flop solved the incompatibility issue between P-serial and P-random of Joint-scan¹;
- the area overhead in terms of number of additional transistors (pmos or nmos) is 12, 11, and 10, compared to normal D flip-flop, PRAS flip-flop, and mux-based scan D flip-flop respectively;
- the proposed scan flip-flop can also be used as a standalone scan flip-flop for serial scan chain;
- it gains 64ps of time saving during normal functional mode of operation;

Chapter-7 addresses the power estimation issue in SoC design. Modern SoC being a complex design the estimation of worst case power becomes a challenge mainly for two reasons: accuracy and computational efficiency. The chapter proposed a binary

¹refer Chapter-5 and Chapter-6 for the incompatibility problem.

integer linear programming (ILP) based approach with the intention to improve upon the efficiency by using smart ILP solver. All the earlier proposed approaches to best of our knowledge resorted to either ATPG based heuristic or pseudo boolean satisfiability based approach. The proposed work contributed to following:

- explored the binary integer linear programming approach to power estimation;
- two problems are addressed considering zero and unit circuit delay models;
- the experimental results show timely generation of input vector pair;

While the thesis studied the test power and power estimation problem it has encountered a set of challenges and new problems which are beyond the scope of this dissertation. Some of the problems with possible solution space are posed here as future work.

8.2 Future Work

- **On Joint-scan architecture:**
 - The proposed frame work for integrating serial scan with random scan could be explored for various possible combination of serial and random access scan architecture. Various architectures may employ unique kind of strategy for grouping of scan flip-flops based on the requirements to optimize either test time, data volume, test power, and DFT area overhead and routing congestion. One of the potential solution could be to integrate the broadcast based serial scan architecture with progressive random access scan.
 - Next important work that must be carried out in order for Joint-scan to be deployed as a viable replacement of state of the art scan techniques is test mechanism for LOS and LOC based transition fault testing. We speculate that transition delay fault coverage can be improved using Joint-scan with nominal DFT area overhead.

- The proposed architecture is expected to ease the scan chain integrity testing where particularly the hold time violation for scan flip-flop is tested. In Joint-scan architecture it is not necessary to test the P-random scan cells for hold violation because those flip-flops are not interconnected back to back unlike in the case of serial scan chain. However the P-serial scan cells needs to be tested against hold violation. Therefor, for chain integrity test of Joint-scan the P-serial scan can possibly be tested as the normal scan chains where as the P-random scan chain could be tested using *march*² test.
- The proposed architecture can be used efficiently than the current serial scan to test the thermal sensitive 3D stacked IC. The proposal could be to test the middle dies which are suffocated with the temperature with P-random access scan where as the other dies, the bottom and top, could be tested with P-serial scan chain.

- **On flip-flop design**

- The proposed flip-flop could be extended to minimize power as well while being used in serial scan chain or P-serial scan chain by extending in the direction as proposed by Amit et al[70]
- The proposed scan flop-flop could be extended to reliability aware flip-flop to tolerate transient errors.

- **On scan chain reordering:**

- By now, the scan chain reordering methodology has been explored exhaustively for various optimization. We propose here to explore the possibilities towards formulating a unified methodology considering test time, peak and average power, routing area and congestion, hold violation, transition fault coverage, and pattern volume.

²The march test is one of the standard test methodologies for memory test[21]

- Algorithm plays a very very important role in getting the nice results. Therefor experimenting on other possible algorithms on formulated graph problem may reveal interesting results.
- Most of the scan chain reordering methodologies are based on test pattern set, this has one drawback with the design flow. It requires late re-connection of scan flip-flops. Therefor if some smart techniques can be thought of which could re-stich the scan flip-flops based on structure of circuit.

- **On power estimation:**

- The current work we proposed is limited to combinational circuit which obviously is not the practice. Therefore the technique must be extended to sequential logic. We propose here to explore the possibility of using 1) the time frame expansion based technique or 2) the scan chain based technique where the combinational and sequential logic are isolated from each other.
- Another limitations where further work is needed is to experiment with the larger real time SoC. This is a challenging problem because the power estimation problem is NP-complete³ in nature. We propose here to explore the modular based divide and conquer approach to compute the vector pair.
- It is quite expected that for some of the circuit the solver would not be able to meet the time limit therefor it is necessary to look at whether the power estimated at this time is near optimal or not.
- Yet another important task to be carried out is to generate the exact functional power in worst case scenario from the input vector pair.

– * – * –

³3-SAT problem is reducible to BILP

Appendix A

Experimental Setup and Tools

The standard benchmark circuits from ISCAS-89, ISCAS-85, ITC-99, and Scaled ISCAS-89 are used for most of the experiments carried out for each of the proposed methodologies. The scaled ISCAS-89 circuits are the larger SoCs built from the existing ISCAS-89 designs. Each of the circuits is synthesized using the Design Compiler® which does the scan chain insertion. The automatic test pattern generation (ATPG) is performed using TetraMax®. For *JScan* architecture the placement and route are carried out using SoC Encounter®. For synthesis of circuit with JScan, we have developed an in-house synthesis tool to insert the JScan architecture. All these experiments are performed on a system with following configurations: clock speed at 2GHz, with 16 GB of RAM and 500 GB of file storage. For power estimation, we have used the industry standard CPLEX solver from IBM. Computation of test power, test time, and test data volume and simulation of scan chains are coded with C++ and python script.

Bibliography

- [1] R. Adiga, G. Arpit, V. Singh, K. K. Saluja, H. Fujiwara, and A. D. Singh. On minimization of test application time for RAS. In *Proc. of the 23rd VLSI Design*, pages 393–398, 2010.
- [2] R. Adiga, G. Arpit, V. Singh, K. K. Saluja, and A. D. Singh. Modified t-flip-flop based scan cell for RAS. In *Proc. of the European Test Symposium, ETS '10*. IEEE Computer Society, 2010.
- [3] V. D. Agrawal, K. T. Cheng, D. D. Johnson, and T. Sheng Lin. Designing circuits with partial scan. *IEEE Des. Test*, 5:8–15, March 1988. ISSN 0740-7475. doi: 10.1109/54.2032.
- [4] S. Ahlawat, J. T. Tudu, A. Matrosova, and V. Singh. A new scan flip flop design to eliminate performance penalty of scan. In *24th IEEE Asian Test Symposium, ATS 2015, Mumbai, India, November 22-25*, pages 25–30, 2015. doi: 10.1109/ATS.2015.12.
- [5] H. Ando. Testing vlsi with random acecss scan. In *Digest of Computer Society International Conference*, pages 50–52, 1980.
- [6] P. Ashar and S. Malik. Implicit computation of minimum-cost feedback-vertex sets for partial scan and other applications. In *Proceedings of the 31st annual Design Automation Conference, DAC '94*, pages 77–80, New York, NY, USA, 1994. ACM. ISBN 0-89791-653-0. doi: <http://doi.acm.org/10.1145/196244.196283>.

- [7] N. Badereddine, P. Girard, A. Virazel, S. Pravossoudovitch, and C. Landrault. *Controlling Peak Power During Scan Testing: Power-Aware Dft and Test Set Perspective*. Springer Berlin / Heidelberg, August 2005.
- [8] N. Badereddine, P. Girard, S. Pravassoudovitch, C. Landrault, AVirazel, and H.-J. Wunderlich. Minimizing peak power consumption during scan testing: Test pattern modification with x-filling heuristics. In *Proc. of Int'l Conf. on Design and Test of Integrated System in Nano Technology*, pages 359 – 364, - 2006.
- [9] N. Badereddine, P.Girard, S. Pravossoudovitch, A.Virazel, and C. Landrault. Scan cell reordering for peak power reduction during scan test cycles. *IFIP International Federation for Information Processing*, 240:267 – 281, 2007.
- [10] D. Baik and K. Saluja. Progressive random access scan: A simultaneous solution to test power, test data volume and test time. In *Proc. of the International Test Conference*, pages 1–10, 2005.
- [11] D. Baik and K. K. Saluja. Test cost reduction using partitioned grid random access scan. In *Proc of the 19th VLSI Design*, 2006.
- [12] D. Baik, S. Kajihara, and K. Saluja. Random access scan: A solution to test power, test data volume and test time. In *Proc of the VSLI Design*, pages 883–888, 2004.
- [13] D. H. Baik and K. K. Saluja. State-reuse test generation for progressive random access scan: Solution to test power, application time and data size. In *Proc. 14th Asian Test Symposium*, pages 272–277, Dec 2005.
- [14] J. Balcrek, P. Fier, and J. Schmidt. On dont cares in test compression. *Microprocessors and Microsystems*, (0):–, 2014. doi: <http://dx.doi.org/10.1016/j.micpro.2014.07.006>.
- [15] R. G. Bennetts and F. P. M. Beenker. Partial scan: what problem does it solve? In *European Test Conference, Proc. of ETC 93., Third*, pages 99–106, Apr 1993.

- [16] S. Bhatia. Will test compression run out of gas? In *Test Conference, 2008. International*, pages 1–1, Oct 2008.
- [17] Y. Bonhomme, P. Girard, C. Landrault, and S. Pravossooudovitch. Power driven chaining of flip-flops in scan architectures. In *Proc. of International Test Conference*, pages 796 – 802, 2002.
- [18] Y. Bonhomme, T. Yoneda, H. Fujiwara, and P. Girard. An efficient scan tree design for test time reduction. *18th IEEE European Test Symposium (ETS)*, 0:174–179, 2004. doi: <http://doi.ieeecomputersociety.org/10.1109/ETSYM.2004.1347657>.
- [19] V. Boppana and W. K. Fuchs. Partial scan design based on state transition modeling. In *Proc. Int. Test Conf*, pages 538–547, 1996.
- [20] S. Borkar. Design challenges of technology scaling. *IEEE Micro*, 19(4):23–29, July 1999. ISSN 0272-1732. doi: 10.1109/40.782564. URL <http://dx.doi.org/10.1109/40.782564>.
- [21] M. Bushnell and V. D. Agrawal. *Essentials of Electronic Testing for Digital, Memory, and Mixed-Signal VLSI Circuits*. Springer, Heidelberger Platz 3, 14197 Berlin, Germany, first edition, 2005.
- [22] K. Butler, J. Saxena, T. Frayer, G. Herthenigton, A. Jain, and J. Lewis. Minimizing power consumption in scan testing: Pattern generation and DFT techniques. In *Proc. IEEE International Test Conference*, pages 335 – 364, 2004.
- [23] S. T. Chakradhar, A. Balakrishnan, and V. D. Agrawal. An exact algorithm for selecting partial scan flip-flops. In *Proceedings of the 31st annual Design Automation Conference, DAC '94*, pages 81–86, New York, NY, USA, 1994. ACM. ISBN 0-89791-653-0. doi: <http://doi.acm.org/10.1145/196244.196285>.
- [24] A. Chandra, H. Yan, and R. Kapur. Multimode illinois scan architecture for test application time and test data volume reduction. In *VLSI Test Symposium, 2007. 25th IEEE*, pages 84–92, May 2007. doi: 10.1109/VTS.2007.39.

- [25] Z. Chen, S. Seth, D. Xiang, and B. Bhattacharya. A unified solution to scan test volume, time, and power minimization. In *VLSI Design. 23rd International Conference*, pages 9–14. IEEE CS, 2010.
- [26] K. T. Cheng and V. D. Agrawal. A partial scan method for sequential circuits with feedback. *IEEE Trans. Comput.*, 39:544–548, April 1990. ISSN 0018-9340. doi: <http://dx.doi.org/10.1109/12.54847>.
- [27] S. Cheng and R. Kapur. How power-aware test improve reliability and yield. *EE Times EDA news online*, www.eetimes.com, 15th, Sept 2004.
- [28] R. Chou, K. Saluja, and V. Agrawal. Scheduling test for VLSI systems under power under power constraints. *IEEE Transactions on VLSI Systems*, 5(2):175 – 185, 1997.
- [29] Y. Chunhua and K. K. Saluja. A study of word oriented random-access-scan based BIST for low area overhead and low power. In *Proc. of Workshop on Impact of Low-Power design on Test and Reliability*, 2008.
- [30] F. Corno, M. Rebaudengo, M. S. Reorda, and M. Violante. On reducing the peak power consumption of test sequences. In *Proc. European Conf. on Circuit Theory and Design*, pages 247 – 250, 1999.
- [31] Corporation. IBM ILOG CPLEX] optimization studio. In *IBM Coporation*. <http://www-01.ibm.com/software/integration/optimization/cplex-optimization-studio/>, 2010.
- [32] V. Dabholkar, S. Chakravarty, I. Pomeranz, and S. Reddy. Techniques for minimizing power dissipation in scan and combinational circuits during test application. *IEEE Trans. on Comp.-Aided Design*, 17(12):1325 – 1333, December 1998.
- [33] S. Devadas, K. Keutzer, and J. White. Estimation of power dissipation in cmos combinational circuits. In *Custom Integrated Circuits Conference, 1990., Proceedings of the IEEE 1990*, pages 19.7/1 –19.7/6, may 1990.

- [34] V. Devanathan, C. Ravikumar, R. Mehrotra, and V. Kamakoti. Pmscan : A power-managed scan for simultaneous reduction of dynamic and leakage power during scan test. In *Test Conference, 2007. ITC 2007. IEEE International*, pages 1–9, Oct 2007. doi: 10.1109/TEST.2007.4437598.
- [35] H. Esmailzadeh, E. Blem, R. S. Amant, K. Sankaralingam, and D. Burger. Power challenges may end the multicore era. *Commun. ACM*, 56(2):93–102, Feb. 2013. ISSN 0001-0782. doi: 10.1145/2408776.2408797. URL <http://doi.acm.org/10.1145/2408776.2408797>.
- [36] R. S. et al. Controlling peak power during scan testing. In *Proc. IEEE VLSI Test Symposium*, pages 153 – 159, 2002.
- [37] X. W. et al. On low capture power test generation for scan testing. In *Proc. IEEE VLSI Test Symposium*, pages 265 – 270, 2005.
- [38] S. Gerstendorfer and H.-J. Wunderlich. Minimized power consumption for scan-based BIST. In *Proc. of International Test Conference*, pages 77 – 84, 1999.
- [39] P. Girard. Survey of low-power testing of VLSI circuits. *IEEE Design & Test of Computers*, pages 82 – 92, May 2002.
- [40] P. Girard and Y. Bonhomme. Low power scan chain design: A solution for an efficient trade-off between test power and scan routing. *Journal of Low Power Electronics*, 1:85 – 95, 2005.
- [41] P. Girard, C. Landrault, S. Pravossoudovitch, and D. Severac. Reduction of power consumption during test application by test vector ordering. *IEEE Electronics Letters*, 33(21):1752 – 1754, October 1997.
- [42] P. Girard, L. Guiller, C. Landrault, and S. Pravassoudovitch. A test vector ordering technique for switching activity reduction during test operation. In *Proceedings of Ninth Great Lakes Symposium on VLSI*, pages 24–30, 1999.

- [43] P. Girard, X. Wen, and N. Toubia. *System-on-Chip Test Architecture: Nanometer Design for Testability*. Morgan Kaufmann, 500 Sansome street, Sanfrancisco, CA 94111, 2007.
- [44] P. Girard, N. Nicolici, and X. Wen, editors. *Power-Aware Testing and Test Strategies for Low Power Devices*. Springer US, first edition, 2010.
- [45] R. Gulve, N. Hage, and J. Tudu. On determination of instantaneous peak and cycle peak switching using ilp. In *Proc. in 20th Symposium on VLSI Test and Design (VDATE-16), 2016*, May 2016.
- [46] R. Gupta and M. Breuer. Ordering storage elements in a single scan chain. In *Computer-Aided Design, Digest of Technical Papers, IEEE International Conference on*, pages 408–411, Nov 1991. doi: 10.1109/ICCAD.1991.185289.
- [47] M. C. Hansen, H. Yalcin, and J. P. Hayes. Unveiling the iscas-85 benchmarks: a case study in reverse engineering. *IEEE Design Test of Computers*, 16(3):72–80, 1999. ISSN 0740-7475. doi: 10.1109/54.785838.
- [48] Y. Higami, S. Kajihara, and K. Kinoshita. Reduced scan shift: a new testing method for sequential circuits. In *Test Conference, 1994. Proceedings., International*, pages 624–630, Oct 1994.
- [49] M. S. Hsiao, E. M. Rudnick, and J. H. Patel. Effects of Delay Models on Peak Power Estimation of VLSI Sequential Circuits. In *International Conference on Computer Aided Design*, pages 45 – 51. IEEE, 1997.
- [50] M. S. Hsiao, G. S. Saund, E. M. Rudnick, and J. H. Patel. Partial scan selection based on dynamic reachability and observability information. In *Proceedings of the Eleventh International Conference on VLSI Design, VLSID '98*, pages 174–, Washington, DC, USA, 1998. ISBN 0-8186-8224-8.
- [51] Y. Hu, X. Fu, X. Fan, and H. Fujiwara. Localized random access scan: Towards

- low area and routing overhead. In *Proc. Asia and South Pacific Design Automation Conference*, ASP-DAC '08, pages 565–570. IEEE Computer Society Press, 2008.
- [52] N. Ito. Automatic incorporation of on-chip testability circuits. In *Proc. Design Automation Conference*, pages 529–534, 1991.
- [53] ITRS-2013. *Test and Test Equipment*. International Technology Roadmap for Semiconductor, 2013 edition edition, 2013.
- [54] k Najeeb, G. Karthik, V. Kamakoti, and V. M. Vedula. Controllability-driven peak dynamic power estimation for vlsi circuits. In *Journal of Low Power Electronics*, volume 3, pages 280 –292, 2007.
- [55] P. Kalla and M. J. Ciesielski. A comprehensive approach to the partial scan problem using implicit state enumeration. In *Proceedings of the IEEE International Test Conference*, ITC '98, pages 651–657, Washington, DC, USA, 1998. ISBN 0-7803-5093-6.
- [56] R. Kapur and R. Ruiz. Maximum test cost reduction. *DFTMAX Compression Backgrounder*, Synopsys, 2009.
- [57] R. Kapur and T. Williams. Tough challenges as design and test go nanometer. *Computer*, 32(11):42 – 45, Nov 1999.
- [58] N. S. Kim, T. Austin, D. Baaruw, T. Mudge, K. Flautner, J. S. Hu, M. J. Irwin, M. Kandemir, and V. Narayan. Leakage current: Moore’s law meets static power. *IEEE Computer*, 36(12):68 – 75, December 2003.
- [59] H. F. Ko and N. Nicolici. Rtl scan design for skewed-load at-speed test under power constraints. In *Proceedings of IEEE International Conference on Computer Design*, pages 237 – 242, 2006.
- [60] S. Krishna Kumar, S. Kundu, and S. Chattopadhyay. Customizing completely specified pattern set targeting dynamic and leakage power reduction during testing.

- Integr. VLSI J.*, 45(2):211–221, Mar. 2012. ISSN 0167-9260. doi: 10.1016/j.vlsi.2011.08.001. URL <http://dx.doi.org/10.1016/j.vlsi.2011.08.001>.
- [61] A. Kunzmann and J. H. Wunderlich. An analytical approach to the partial scan problem. *Journal of Electronic Testing: Theory and Applications*, 1:163–174, 1990.
- [62] I.-S. Lee, Y. M. Hur, and T. Ambler. The efficient multiple scan chain architecture reducing power dissipation and test time. In *Test Symposium. 13th Asian*, pages 94–97, Nov 2004.
- [63] S. Y. Lee and K. Saluja. Test application time reduction for sequential circuits with scan. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 14(9):1128–1140, Sep 1995.
- [64] J. LI, Q. XU, Y. HU, and X. LI. On reducing both shift and capture power for scan-based testing. In *Proc. Design Automation Conference*, pages 653 – 658, 2008.
- [65] R. Li, D. Zhou, and D. Du. Satisfiability and integer programming as complementary tools. In *Design Automation Conference, 2004. Proceedings of the ASP-DAC 2004. Asia and South Pacific*, pages 880–883, 2004.
- [66] X. Lin, I. Pomeranz, and S. M. Reddy. Full scan fault coverage with partial scan. In *Proceedings of the conference on Design, automation and test in Europe, DATE '99*, New York, NY, USA, 1999. ACM. ISBN 1-58113-121-6. doi: <http://doi.acm.org/10.1145/307418.307545>.
- [67] X. Lin, J. Rajski, I. Pomeranz, and S. Reddy. On static test compaction and test pattern ordering for scan designs. In *Test Conference, 2001. Proceedings. International*, pages 1088–1097, 2001.
- [68] X. Lin, R. Press, J. Rajski, P. Reuter, T. Rinderknecht, B. Swanson, and N. Tamarapalli. High-frequency, at-speed scan testing. *IEEE Design and Test of Computer*, 20(5):17 – 25, September-October 2003.

- [69] H. Mangassarian, A. Veneris, S. Safarpour, F. Najm, and M. Abadir. Maximum circuit activity estimation using pseudo-boolean satisfiability. In *Design, Automation Test in Europe Conference Exhibition, 2007. DATE '07*, pages 1–6, april 2007. doi: 10.1109/DATE.2007.364519.
- [70] A. Mishra, N. Sinha, Satdev, V. Singh, S. Chakravarty, and A. D. Singh. Modified scan flip-flop for low power testing. *2012 IEEE 21st Asian Test Symposium*, 0: 367–370, 2010. ISSN 1081-7735. doi: <http://doi.ieeecomputersociety.org/10.1109/ATS.2010.69>.
- [71] A. Mishra, N. Sinha, Satdev, V. Singh, S. Chakravarty, and A. D. Singh. Modified scan flip-flop for low power testing. In *Proceedings of the 19th IEEE Asian Test Symposium, ATS 2010, 1-4 December 2010, Shanghai, China*, pages 367–370, 2010. doi: 10.1109/ATS.2010.69.
- [72] S. Mitra and K. S. Kim. X-compact: an efficient response compaction technique. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 23(3):421–432, March 2004.
- [73] K. Miyase and S. Kajihara. Optimal scan tree construction with test vector modification for test compression. In *Test Symposium, 12th IEEE Asian*, pages 136–141, Nov 2003.
- [74] G. E. Moore. Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85, Jan 1998. ISSN 0018-9219. doi: 10.1109/JPROC.1998.658762.
- [75] A. Mudlapur, V. Agrawal, and A. Singh. A random scan architecture to reduce hardware overhead. In *Proc of the International Test Conference*, number 15.1 in ITC '05, pages 1–9, 2005.
- [76] F. Najm. A survey of power estimation techniques in vlsi circuits. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 2(4):446–455, dec. 1994. ISSN 1063-8210. doi: 10.1109/92.335013.

- [77] S. Narayanan, R. Gupta, and M. Breuer. Configuring multiple scan chains for minimum test time. In *Computer-Aided Design, Digest of Technical Papers., IEEE/ACM International Conference on*, pages 4–8, Nov 1992.
- [78] N. Nicolici and B. M. Al-Hashimi. Scan latch partitioning into multiple scan chain for power minimization in full scan sequential circuits. In *Proc. IEEE/ACM Design Automation and Test in Europe*, pages 715 – 722, 2000.
- [79] A. Pandey and J. Patel. Reconfiguration technique for reducing test time and test data volume in illinois scan architecture based designs. In *VLSI Test Symposium, 2002. (VTS 2002). Proceedings 20th IEEE*, pages 9–15, 2002.
- [80] M. Pedram. Power minimization in IC Design: Principles and Applications. *ACM Transactions on Design Automation of Electronics System*, 1(1):3 – 56, January 1996.
- [81] M. Pedram. Power minimization in ic design: Principles and applications. *ACM Transactions on Design Automation of Electronics System*, 1(1):3 – 56, January 1996.
- [82] M. Pedram. Advanced power estimation technique. In J. Mermet and W. Nebel, editors, *Low Power Design in Deep Submicron Technology*. Kluwer Academic Publishers, 1997.
- [83] M. Pedram and S. Nazarian. Thermal modeling, analysis, and management in vlsi circuits: Principles and methods. *Proceedings of The IEEE*, 94(8):1487 – 1501, August 2006.
- [84] I. Pomeranz. Reducing the input test data volume under transparent scan. *Computers Digital Techniques, IET*, 8(1):1–10, January 2014.
- [85] I. Pomeranz and S. Reddy. Reducing test application time for full scan circuits by the addition of transfer sequences. In *Test Symposium, Proceedings of the Ninth Asian*, pages 317–322, 2000.

- [86] J. M. Rabaey, A. Chandrakasan, and B. Nikolic. *Digital Integrated Circuits A Design Perspective*. Pearson Prentice Hall, 482, F.I.E, Pataparganj, Delhi 110092, India, second edition, 2007.
- [87] J. Rajski, J. Tyszer, M. Kassab, and N. Mukherjee. Embedded deterministic test. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 23(5):776–792, May 2004.
- [88] 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 CADICS*, 23(7):1142 – 1153, July 2004.
- [89] P. M. Rosinger, B. Al-Hashimi, and N. Nicolici. Scan architecture for shift and capture cycle power reduction. In *Proc. International Symposium on Defect and Fault Tolerance in VLSI System*, pages 129 – 137, 2002.
- [90] E. Rudnick and J. Patel. Efficient techniques for dynamic test sequence compaction. *Computers, IEEE Transactions on*, 48(3):323–330, Mar 1999.
- [91] A. Sagahyroon and F. A. Aloul. Using SAT-based techniques in power estimation. *Microelectronics Journal, Elsevier*, 38:706–715, jun 2007.
- [92] A. Sagahyroon and F. A. Aloul. Using SAT based technique in low power state assignment. *Journal of Circuits, Systems, and Computers, World Scientific Publishing Company*, 20(8):1605–1618, 2011.
- [93] K. K. Saluja. Outstanding challenges in testing nano technology based integrated circuits. In *Proc. Asian Test Symposium*, 2003.
- [94] S. Samii, M. Selkala, E. Larsson, K. Chakrabarty, and Z. Peng. Cycle-accurate test power modeling and its application to SoC test scheduling. In *Proc. Intl. Test Conf.*, pages 1 – 10, 2006.

- [95] S. Samii, M. Selkala, E. Larsson, K. Chakrabarty, and Z. Peng. Cycle-accurate test power modeling and its application to soc test architecture design and scheduling. *IEEE Transaction on Computer-Aided Design*, 27(5):973 – 984, May 2008.
- [96] R. Sankaralingam and N. A. Touba. Inserting test points to control peak power during scan testing. In *Proc. IEEE Symp. Defect and Fault Tolerance*, pages 138 – 146, 2002.
- [97] R. Sankaralingam, R. R. Oruganti, and N. A. Touba. Static compaction technique to control scan vector power dissipation. In *Proc. of VLSI Test Symposium*, pages 35 – 40, - 2000.
- [98] J. Savir. Skewed-load transition test: Part i, calculus. In *Proceedings of International Test Conference*, pages 705–713, 1992.
- [99] J. Savir and S. Patil. Broad-side delay test. *IEEE Transanction on Computer-Aided Design of Integrated Circuit and System*, 13(8):1057 – 1065, August 1994.
- [100] J. Saxena, K. M. Butler, and L. Whetsel. An analysis of power reduction techniques in scan testing. In *Proc. IEEE Int. Test Conference*, pages 670 – 677, 2001.
- [101] J. Saxena, K. M. Butler, V. B. Jayaram, S. Kundu, N. V. Arvind, P. Sreeprakash, and M. Hachinger. A case study of ir-drop in structured in at-speed testing. In *Proc. IEEE Int. Test Conference*, pages 1098 – 1104, September 2003.
- [102] M. Selkälä. volume Master Thesis. Linköping University, Linköping, Sweden, June 2006.
- [103] M. Shah and J. Patel. Enhancement of the illinois scan architecture for use with multiple scan inputs. In *VLSI, 2004. Proceedings. IEEE Computer society Annual Symposium on*, pages 167–172, Feb 2004. doi: 10.1109/ISVLSI.2004.1339525.
- [104] M. Sharma, J. Patel, and J. Rearick. Test data compression and test time reduction of longest-path-per-gate tests based on illinois scan architecture. In *Proc. 21st VLSI Test Symposium*, pages 15–21, April 2003.

- [105] K. Shyamala, M. Shoaib, and V. Kamakoti. Peak dynamic power estimation of fpga-mapped digital designs. In *13th VLSI Design and Test Symposium(VDAT-09)*, pages 155 – 166, 2009.
- [106] K. Shyamala, J. Vimalkumar, and V. Kamakoti. Novel sat-based peak dynamic power estimation for digital circuits. In *Journal of Low Power Electronics*, volume 5, pages 429 – 438, December 2009.
- [107] O. Sinanoglu, I. Bayraktaroglu, and A. Orailoglu. Reducing average and peak test power through scan chain modification. *Journal of Electronics Testing: Theory and Application*, 19:457 – 467, 2003.
- [108] T. Takasaki, T. Inoue, and H. Fujiwara. Partial scan design methods based on internally balanced structure. In *Design Automation Conference, Proc. of the Asia and South Pacific*, pages 211–216, Feb 1998.
- [109] N. Toubia. Survey of test vector compression techniques. *Design Test of Computers, IEEE*, 23(4):294–303, April 2006.
- [110] J. Tudu. Jscan: A joint scan dft architecture to minimize test time, pattern volume, and power. In *Proc. in Symposium on VLSI Design and Test*, Guwahati, India, May 2016.
- [111] J. T. Tudu, E. Larsson, V. Singh, and V. D. Agrawal. On minimization of peak power for scan circuit during test. In *14th IEEE European Test Symposium, ETS 2009, Sevilla, Spain, May 25-29, 2009*, pages 25–30, 2009. doi: 10.1109/ETS.2009.36.
- [112] J. T. Tudu, E. Larsson, V. Singh, and H. Fujiwara. Scan cell reordering to minimize peak power during test cycle: A graph theoretic approach. In *15th European Test Symposium, ETS 2010, Prague, Czech Republic, May 24-28, 2010*, page 259, Nov 2010. doi: 10.1109/TEST.2005.1583994.

- [113] J. T. Tudu, D. Malani, and V. Singh. Ilp based approach for input vector controlled (ivc) toggle maximization in combinational circuits. In *Progress in VLSI Design and Test*, volume 7373 of *Lecture Notes in Computer Science*, pages 172–179. Springer Berlin Heidelberg, 2012.
- [114] M. Valka, A. Bosio, L. Dilillo, P. Girard, S. Pravossoudovitch, A. Virazel, E. Sanchez, M. D. Carvalho, and M. S. Reorda. A functional power evaluation flow for defining test power limits during at-speed delay testing. In *2011 Sixteenth IEEE European Test Symposium*, pages 153–158, May 2011. doi: 10.1109/ETS.2011.21.
- [115] V. V. Vazirani. *Approximation Algorithm*. Springer, India, first edition, 2001.
- [116] K. Wagner. Design for testability in the amdahl 580. In *Digest of Computer Society International Conference*, pages 384–288, 1983.
- [117] C.-Y. Wang and K. Roy. Maximum power estimation for CMOS circuits using deterministic and statistic approaches. In *VLSI Design, 1996. Proceedings., Ninth International Conference on*, pages 364 –369, jan 1996. doi: 10.1109/ICVD.1996.489636.
- [118] C.-Y. Wang, K. Roy, and T.-L. Chou. Maximum power estimation for sequential circuits using a test generation based technique. In *Custom Integrated Circuits Conference, 1996., Proceedings of the IEEE 1996*, pages 229 –232, may 1996. doi: 10.1109/CICC.1996.510549.
- [119] L.-T. Wang, C.-W. Wu, and X. Wen. *VLSI Test Principles and Architectures Design for Testability*. Morgan Kaufmann Publishers Elsevier, 500 Sansome street, Sanfrancisco, CA 94111, 2006.
- [120] S.-J. Wang, K. Shu-Min Li, S.-C. Chen, H.-Y. Shiu, and Y.-L. Chu. Scan-chain partition for high test-data compressibility and low shift power under routing constraint. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 28(5):716–727, May 2009.

- [121] N. Weste, D. Harris, and A. Banerjee. *CMOS VLSI Design A Circuit and System Perspective*. Dorling Kindersley (India) Pvt. Ltd., Pearson Education, 482, F.I.E., Patparganj, Delhi 110 092, India, 2009.
- [122] L. Whetsel. Adapting scan architecture for low power scan operation. In *Proceedings International Test Conference*, pages 863–873, 2000.
- [123] Q. Wu, Q. Qiu, and M. Pedram. Estimation of peak power dissipation in VLSI circuits using the limiting distributions of extreme order statistics. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 20(8):942–956, aug 2001. ISSN 0278-0070.
- [124] Q. Wu, Q. Qiu, and M. Pedram. Estimation of peak power dissipation in vlsi circuits using the limiting distributions of extreme order statistics. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 20(8): 942–956, aug 2001.
- [125] Y. Wu and M. C.-T. Chao. Scan-chain reordering for minimizing scan shift power based on non-specified test cubes. In *Proc. IEEE VLSI Test Symposium*, pages 147–174, April-May 2008.
- [126] e. a. X. Wen. Low capture switching activity test generation for reducing IR-drop in at-speed scan testing. *Journal of Electronics Tsting: Theory and Application*, 24:379–391, 2008.
- [127] D. Xiang and J. Patel. Partial scan design based on circuit state information and functional analysis. *Computers, IEEE Transactions on*, 53(3):276–287, March 2004.
- [128] D. Xiang and J. H. Patel. Partial scan design based on circuit state information and functional analysis. *IEEE Trans. Comput.*, 53:276–287, March 2004. ISSN 0018-9340. doi: <http://dx.doi.org/10.1109/TC.2004.1261835>.

- [129] D. Xiang, S. Venkataraman, W. K. Fuchs, and J. H. Patel. Partial scan design based on circuit state information. In *In Proc. of the Design Automation Conf., Las Vegas, NV*, pages 807–812, 1996.
- [130] G. Xu and A. D. Singh. Achieving high transition delay fault coverage with partial DTSFF scan chains. In *IEEE International Test Conference, ITC 2007, Santa Clara, California, USA, October 21-26*, pages 1–9, 2007. doi: 10.1109/TEST.2007.4437608.
- [131] T. Yoshida and M. Watati. A new approach for low power scan testing. In *Proc. International Test Conference (ITC-2003)*, pages 480 – 487, 2003.
- [132] Y. Zorian. A distributed BIST control scheme for complex VLSI devices. In *Proceedings of the 11th IEEE VLSI Test Symposium*, pages 4 – 9, 1993.