

Series on Emerging Technologies in Circuits and Systems

Advances in **3D Integrated Circuits and Systems**

Hao Yu • Chuan Seng Tan

Series on Emerging Technologies in Circuits and Systems



Advances in **3D Integrated Circuits** and **Systems**

Hao Yu • Chuan-Seng Tan

NTU, Singapore

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Yu, Hao (Electrical engineer)

Advances in 3D integrated circuits and systems / by Hao Yu (NTU, Singapore), Chuan-Seng Tan (NTU, Singapore).

pages cm. -- (Series on emerging technologies in circuits and systems)

Includes bibliographical references and index.

ISBN 978-9814699006 (hardback : alk. paper) -- ISBN 978-9814699013 (pbk. : alk. paper)

I. Three-dimensional integrated circuits. I. Tan, Chuan Seng. II. Title.

TK7874.893.Y83 2015

621.3815--dc23

2015020144

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Copyright © 2016 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

In-house Editor: Amanda Yun

Typeset by Stallion Press

Email: enquiries@stallionpress.com

Printed in Singapore by B & Jo Enterprise Pte Ltd

Advances in
3D Integrated Circuits
and **Systems**

Series on Emerging Technologies in Circuits and Systems

Series Editor: Kiat-Seng Yeo (*Singapore University of Technology & Design,
Singapore*)

Published

Vol. 1 **Advances in 3D Integrated Circuits and Systems**
by Hao Yu and Chuan-Seng Tan

Preface

In the past few decades, the design of computers has been primarily driven by improving performance with faster clock frequency of single-core processor using transistor scaling. The transistor scaling towards high performance of fast clock frequency is, however, stuck recently due to the constraint of power density. By exploiting parallelism, multi-core processor based design of computers has emerged to scale up performance of throughput under power budget. As such, the scaling paradigm has changed to integrate as many processor cores as possible on one single chip. In the traditional 2D based memory-logic integration, the scalability of many-core integration is limited by the communication between the cores and memory by I/O interconnections, which have posed stringent requirements for high utilization efficiency of both bandwidth and power towards operating thousands of microprocessor cores on chip.

Stacking one layer (with core or memory blocks) above the other with short-distance RC interconnects of through-silicon vias (TSVs), 3D integration has become one promising alternative for many-core memory-logic integration with high bandwidth and low power. The other possible alternative for many-core memory-logic integration is 2.5D integration, to place core and memory blocks both on one common substrate, interconnected with the aid of middle-distance transmission lines (underneath the substrate) of through-silicon interposers (TSIs). Even though 3D and 2.5D integrations are potential candidates for many-core memory-logic integration, there are a few critical issues to be addressed as shown in this book.

From the device fabrication perspective, TSV involves a complex process due to its smaller dimension and demanded accuracy with existence of liner material around metal-fill. On the other hand, fabrication of TSI involves fabrication of TSV, microbumps and other metal layers, thus involving

additional complexity. The modeling of TSV/TSI can be quite different from the traditional RC interconnects. Existence of liner material around TSV metal-fill and also thermal-mechanical stress due to mismatch of coefficient of thermal expansion (CTE) between Si and TSV metal-fill can have critical effect on the active device parameters in the vicinity of TSV. In addition, as the fabricated TSVs may result in a poor yield, careful physical design and testing needs to be carried out. However, due to the existence of thousands of TSVs, a low complex physical design is required.

From the system management perspective, thermal management is one of the key challenges to be addressed in 3D integration. In addition to allocating dummy TSV for cooling, recent microfluidic channel based active cooling is also introduced. Microfluidic channels can be etched beneath the substrate of each layer in 3D and the microfluid flowing through the channel can help in cooling the substrate layer. However, an effective management of microfluid has to be performed such as flow-rate control to avoid unnecessary overhead. Similarly, dummy TSVs can be inserted to reduce the thermal and stress gradients across the chip. In addition to thermal management, power management needs to be performed in 3D many-core processors as well to avoid dark silicon dilemma. Many multi-core processor designs make use of off-chip power converters, which may not be scalable for the surge of the supply current demanded by many-core processors. As such, an effective power management of power I/Os needed to be carried with less number of power converters.

What is more, signal I/Os are utilized for the communication between memory and logic blocks. The available limited communication bandwidth is often the bottleneck. Bandwidth utilization of 2.5D TSI I/Os is less compared to that of 3D TSV I/Os. To improve the bandwidth utilization, memory controller needs to be explored with configurability. Previous works on memory controllers are either static or non-scalable for many-core processors. Further, with the increase in communication between memory and logic blocks, communication power also increases. Reduction in I/O voltage swing can help in reducing the communication power at the expense of bit-error-rate (BER). As some types of workloads can tolerate a certain amount of BER, I/O communication can thereby be performed with reduced voltage swing by the use of signal I/O management.

This book is intended to address all the above mentioned challenges. The need for many-core processors with according design challenges faced in the traditional 2D integration is presented in Chapter 1. The rest of the book is divided into five parts as follows.

Part 1 deals with the 3D-IC device modeling. We introduce fabrication methodology utilized for TSVs and TSIs and also evaluate their performance with different kinds of materials in Chapter 2. We further develop TSV and TSI device models in Chapter 3.

Part 2 presents the 3D-IC physical design and testing. We introduce macromodeling to reduce the physical design complexity of thousands of TSVs, which is followed by the corresponding TSV allocation methods for power grid and clock in Chapter 4 and Chapter 5 respectively. Moreover, to effectively evaluate the reliability of TSVs under test, we further propose the TSV testing with compressive sensing as discussed in Chapter 6.

Part 3 illustrates the thermal management of 3D-IC many-core processors. We introduce the system level thermal and power models in Chapter 7. We further discuss microfluid based thermal management in Chapter 8.

Part 4 deals with the power and signal I/O management of the 3D-IC many-core processor. To improve power-converter utilization, we present the power I/O management by space-time multiplexing in Chapter 9. To improve I/O channel utilization, we also present the signal I/O management by the space-time multiplexing and the adaptive voltage-swing in Chapter 10.

Part 5 presents 3D-IC design examples. We first show the 3D integration of MEMS sensor and CMOS readout circuit in Chapter 11. Next, we present the 2.5D I/O design such as pre-emphasis, adaptive driver, and clock-data recovery in Chapter 12. We further discuss an 8-core microprocessor with accelerator and memory blocks integrated by 2.5D I/Os. Lastly, we discuss the 3D integration of the non-volatile memory in Chapter 14.

Lastly, the authors would like to thank their students at Nanyang Technological University: Sai Manoj P. D., Lin Zhang, Hongyu Li, Hantao Huang, Kanwen Wang, Yuhao Wang, Dongjun Xu, Xiwei Huang, Hanhua Qian, Shunli Ma, Yang Shang, Jie Lin and Shikai Zhu, for their contribution to this book. Sai Manoj P. D. in particular has made the significant help for this book. The authors would also like to thank their colleagues: Prof. Paul Franzon, Prof. Sungkyu Lim, Prof. Dennis Sylvester, Prof. Sheldon Tan, Prof. Yuan Xie, Prof. Eby Friedman, Prof. Chiphong Chang, Prof. Zhiyi Yu and Prof. Wei Zhang, for their kind suggestion and collaboration. The relevant research is funded by MOE Tier-2, A*STAR PSF, DIRP and NRF CRP grants from Singapore.

Hao Yu
Chuan Seng Tan
Singapore

Contents

<i>Preface</i>	v
1. Introduction	1
1.1 Thousand-core On-chip	1
1.2 State-of-the-Art Many-core Microprocessors	5
1.3 Memory-logic Integration	7
1.3.1 2D Integration Challenges	7
1.3.1.1 Scalability	7
1.3.1.2 Channel Loss	8
1.3.1.3 I/O Circuit Design	9
1.3.1.4 Testing	10
1.3.1.5 Thermal Management	10
1.3.1.6 Power Management	11
1.3.1.7 I/O Management	11
1.3.2 3D Integration	12
1.3.3 2.5D Integration	14
1.4 Organization of the Book	15
 Part 1. Device Modeling	 17
2. Fabrication	19
2.1 Introduction	19
2.2 TSV Structure and Fabrication	21
2.2.1 Structure Design	21
2.2.1.1 Wafer Layout and Mask Design	21
2.2.1.2 Electrical Structure Design	22

2.2.1.3	Thermal Structure Design	24
2.2.1.4	Dummy TSV Blocks	25
2.2.2	Fabrication Process	25
2.2.2.1	Electrical Structure Fabrication Process	25
2.2.2.2	Thermal Structure Fabrication Process	28
2.2.3	Process Control and Optimization	30
2.2.3.1	DRIE Si Etch	30
2.2.3.2	Dielectric Liner Deposition	31
2.2.3.3	Ta Barrier/Cu Seed Layer Deposition and Cu ECP	33
2.2.3.4	Cu CMP	35
2.3	TSV Electrical Characterization	37
2.3.1	Measurement Setup	37
2.3.2	Conventional PETEOS Oxide Liner	38
2.3.2.1	Electrical <i>CV</i> Measurement	38
2.3.2.2	Electrical <i>IV</i> Measurement	42
2.3.3	Black Diamond Low-k Liner	42
2.3.3.1	Electrical <i>CV</i> Measurement	44
2.3.3.2	Electrical <i>IV</i> Measurement	46
2.3.4	Al ₂ O ₃ /Oxide Bi-layer Liner	48
2.3.4.1	Electrical <i>CV</i> Measurement	48
2.3.4.2	Electrical <i>IV</i> Measurement	53
2.4	TSV Thermal Characterization Results	54
2.4.1	Measurement Setup	54
2.4.2	Cu-TSV Thermal Modeling	55
2.4.3	Cu-TSV Induced Stress Modeling	56
2.4.4	Cu-TSV Induced Stress Measurement by Micro-Raman Analysis	61
2.5	TSI Structure and Fabrication	64
2.5.1	Structure Design	65
2.5.2	Fabrication Process	68
2.6	Summary	69
3.	Device Model	71
3.1	Introduction	71
3.2	Nonlinear MOSCAP Model	72
3.3	TSV Device Model	75
3.3.1	Electrical Model	75
3.3.2	Thermal Model	77

3.3.3	Mechanical Model	82
3.3.4	Delay Model	86
3.3.4.1	Electrical-Thermal Coupled Delay Model	86
3.3.4.2	Electrical-Mechanical Coupled Delay Model	87
3.3.4.3	Electrical-Thermal-Mechanical Coupled Delay Model	88
3.3.5	Power Model	90
3.4	TSI Device Model	91
3.4.1	Delay Model	92
3.4.1.1	T-line Model	92
3.4.1.2	Delay of T-line	93
3.4.2	Power Model	94
3.4.2.1	TSV and TSI Comparison	95
3.4.2.2	Energy-efficiency Analysis	95
3.5	Summary	97

Part 2. Physical Design 99

4.	Macromodel	101
4.1	Introduction	101
4.2	Power and Thermal Integrity	103
4.3	Macromodeling	105
4.3.1	Complexity Compression	106
4.3.1.1	Complexity Compression of States	106
4.3.1.2	Complexity Compression of I/Os	107
4.3.2	Parameterization	109
4.4	Summary	111
5.	TSV Allocation	113
5.1	Introduction	113
5.2	Power Ground Design	114
5.2.1	Problem Formulation	114
5.2.2	Sensitivity based TSV Allocation	116
5.3	Clock-tree Design	122
5.3.1	Problem Formulation	123
5.3.2	Sensitivity based TSV Allocation	123
5.3.2.1	Reduction of Thermal Gradient	124

5.3.2.2	Reduction of Stress Gradient	126
5.3.2.3	Clock-skew Reduction	128
5.4	Summary	137
6.	Testing	139
6.1	Introduction	139
6.2	3D IC Test	140
6.2.1	System Architecture	140
6.2.2	Problem Formulation	142
6.3	Compressive Sensing and Recovery of Testing Data	143
6.3.1	Sparsity of Test Data	143
6.3.2	Lossless Compression and Recovery	145
6.3.3	OMP Solver	146
6.4	TSV Testing	147
6.4.1	Testing Circuit	147
6.4.2	Pre-bond TSV Test	149
6.4.3	Post-bond TSV Test	152
6.5	Summary	153
Part 3. Thermal Management		155
7.	Power and Thermal System Model	157
7.1	Introduction	157
7.2	3D System Power Model	159
7.2.1	Core and DRAM Power Model	159
7.2.2	System Power Breakdown	160
7.2.2.1	Thermal Runaway Failure	162
7.3	3D System Thermal Model	164
7.3.1	Microfluidic Channel Thermal Model	166
7.3.2	Steady State Thermal Analysis	167
7.3.2.1	Software Implementation	168
7.3.2.2	Validation of Proposed Model	169
7.4	3D Cyber-physical System	171
7.4.1	System Architecture	171
7.4.2	Problem Formulation	172
7.5	Summary	173
8.	Microfluidic Based Cooling	175
8.1	Introduction	175

8.2	3D Cyber-physical Thermal Management	176
8.2.1	Real-time Temperature Demand Estimation . . .	179
8.2.2	Prediction and Correction	179
8.2.3	Clustering of Microchannels	182
8.2.4	Allocation of Flow Rates	189
8.2.5	Minimization of Cooling Effort	191
8.3	Summary	193

Part 4. I/O Management 195

9.	Power I/O Management	197
9.1	Introduction	197
9.2	3D Power I/O Management	199
9.2.1	System Architecture	200
9.2.2	Problem Formulation	204
9.3	ILP based Optimization	205
9.4	Space-time Multiplexing	206
9.4.1	Adaptive Clustering	207
9.4.1.1	Power-signature Extraction	207
9.4.1.2	Space Multiplexing: Grouping	210
9.4.1.3	Time Multiplexing: Subgrouping	211
9.4.1.4	Allocation of Power Converters	212
9.4.2	Scheduling of Workloads	221
9.4.2.1	Slack Calculation	222
9.4.2.2	Scheduling of Workloads	224
9.5	Summary	227
10.	Signal I/O Management	229
10.1	Introduction	229
10.2	3D Bandwidth and Voltage-swing Management	231
10.2.1	System Architecture	231
10.2.2	Problem Formulation	232
10.3	Signal I/O Bandwidth Management	235
10.3.1	Memory-access Data-pattern	235
10.3.1.1	LLC MPKI Pattern	235
10.3.1.2	Quality of Service	237
10.3.2	Reconfigurable Memory Controller	238
10.3.2.1	Space Multiplexing: Channel Allocation	238

10.3.2.2	Time Multiplexing: Time-slot Allocation	238
10.3.2.3	Space-Time Multiplexing based I/O Bandwidth Management	239
10.4	Signal I/O Voltage-swing Management	245
10.4.1	Reinforcement Q-learning	248
10.4.1.1	Q-learning Theory	248
10.4.1.2	System Model	250
10.4.1.3	Adaptive I/O Voltage-swing Tuning	251
10.4.1.4	Accelerated Q-learning	254
10.4.1.5	Comparison of Adaptive Tuning by Conventional and Accelerated Q-learning	259
10.5	Summary	261
11.	Sensor	263
11.1	Introduction	263
11.2	3D Sensor Design	265
11.2.1	SOI MEMS Accelerometer	265
11.2.2	CMOS Readout Circuit	268
11.3	Testing and Measurement	272
11.3.1	Hermetic Test	273
11.3.2	Bonding Reliability	277
11.3.3	Measurement Results	277
11.4	Summary	278
12.	I/O	281
12.1	Introduction	281
12.2	2.5D I/O Buffer Design	283
12.2.1	LVDS I/O Buffer	284
12.2.2	CML I/O Buffer	286
12.2.3	Simulation Results	287
12.2.3.1	EM Simulation of TSV and TSI I/O	290
12.3	2.5D Adaptive I/O	291
12.3.1	Transmitter and Receiver	293
12.3.1.1	Error Correcting Code	295
12.3.2	Simulation Results	296
12.4	2.5D I/O with Clock Data Recovery	298

12.4.1	CTLE Equalization	298
12.4.2	Current Sampling	299
12.4.3	Phase Interpolator	300
12.4.4	Measurement Results	301
	12.4.4.1 Data Recovery	301
	12.4.4.2 Clock Recovery	302
12.5	Summary	304
13.	Microprocessor	305
13.1	Introduction	305
13.2	Building Blocks	307
13.2.1	Core	307
13.2.2	Router	309
	13.2.2.1 Core-to-core Communication	310
	13.2.2.2 Core-to-memory Communication	311
13.2.3	Accelerator	312
13.2.4	I/O Link	314
	13.2.4.1 Transmitter	315
	13.2.4.2 Receiver	315
	13.2.4.3 Voltage-controlled Oscillator	317
13.3	2.5D Multi-core Microprocessor	318
	13.3.1 System Design	318
	13.3.2 Fabricated Chip	318
	13.3.3 Simulation Results	318
13.4	Summary	321
14.	Non-volatile Memory	323
14.1	Introduction	323
14.2	3D Hybrid Memory with ReRAM	324
14.3	ReRAM-crossbar Memory Design	326
	14.3.1 Performance Modeling	331
	14.3.1.1 Delay Model	331
	14.3.1.2 Power Model	334
	14.3.1.3 Area Model	336
	14.3.1.4 Stacked CBRAM-crossbar Memory	337
	14.3.2 Design Space Exploration	337
14.4	ReRAM Block-level Incremental Data Retention	340
	14.4.1 Dirty Bit Set-up	341

14.4.2	Incremental Write-back	342
14.4.3	Simulation Results	344
14.5	Summary	347
<i>Bibliography</i>		349
<i>Index</i>		373