On-Chip Resonance in Nanoscale Integrated Circuits

By

Jonathan Rosenfeld

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Eby G. Friedman

Department of Electrical and Computer Engineering

Arts, Sciences, and Engineering

School of Engineering and Applied Sciences

University of Rochester

Rochester, New York

2009

“In this short life
That only lasts an hour,
How much, how little,
Is within our power!”

Emily Dickinson

Curriculum Vitae

Jonathan Rosenfeld was born in Simferopol, Ukraine in 1973. He grew up in Israel, served in the Israeli Defense Force - IDF, and received a bachelor degree in mechanical engineering in 1999 from the Technion - Israel Institute of Technology, Department of Mechanical Engineering.

He worked as a process development engineer, conducting research on electro-chemical polishing methods at HAM-LET. He then joined CAMTEK, where he provided world-wide customer support services for the company's automated optical inspection (AOI) systems used by the semiconductor manufacturing and packaging, IC substrate, and printed circuit board (PCB) industries.

He obtained a second bachelor degree, this time in electrical engineering, from Ort Braude College, Department of Electrical Engineering in 2003, graduating with honors. In 2003, Jonathan Rosenfeld began his Ph.D. studies at the University of Rochester, Department of Electrical and Computer Engineering, under the supervision of Professor Eby G. Friedman. In 2005, Jonathan obtained his Masters degree in Electrical and Computer Engineering.

Jonathan Rosenfeld was an intern at Intrinsix Corporation, Fairport, NY, in 2005, where he designed Gm-C circuits for an ΣΔ-modulator ADC for an FM tuner. In 2007, he was an intern at the Eastman Kodak Company, Rochester, NY, where he developed a column-based multiple ramp integrated ADC for CMOS image sensors. His primary research interests are in the areas of interconnect design, resonant clock and data distribution networks, on-chip DC-DC converters, and the design of analog and mixed-signal integrated circuits.

Acknowledgments

It is quite difficult to express gratitude to all of the individuals who have affected my life since my day of birth until the moment these lines are written. I truly believe that all of these people have shaped my personality, contributing to my achievements. Naturally, my parents had the greatest impact on my early life and education and the pursuit for excellence. I would like to thank them dearly for their encouragement, patience, understanding, and support throughout these past years and the many years to come.

Undoubtedly, my dearest wife Merav and our lovely and joyful son, Blaze Barak, have impacted my life during the recent intense years of my PhD studies. The support, happiness, and enjoyment they provide me every day is priceless. The years I spent pursuing my PhD degree would not be the same without them. Thank you so much my little dearests.

Obtaining a PhD degree could not be realized without Professor Eby G. Friedman. I was extremely lucky to be embraced by one of the leading researchers in the country in the area of VLSI circuits. Professor Friedman gave me the unique opportunity to bring my creativity and scientific enthusiasm to higher levels. The exceptional way Professor Friedman manages his PhD program allowed me to thrive and brought out the best from me. His ability to foresee future and challenging areas of potential research resulted in this PhD thesis as many others before. I learned so much from his remarkable verbal and writing presentation skills, realizing how important it is to deliver your ideas in a clear, consistent, and accurate way. His attention to my academic and personal requests is highly appreciated. Thank you so much for being my mentor!

Spending my days in Professor Friedman's lab would not have been the same without his assistant, Ms. RuthAnn Williams. Ms. Williams was always there for me (as she was for the other students), providing me help and advice on many issues. Her bubbly personality and happiness of life is one of a kind and unmatched in any person I know. I would like to thank her for years of support and, of course, for her delicious lasagna. Many thanks Ruthie (and thank you for the cookbooks you gave me).

I would like to thank Professor Yaakov Bar-Shlomo and Dr. Pinchas Schechner from the Ort-Braude College in Israel, who chose me to pioneer an internship program that Professor Friedman established with the College. Their confidence in my abilities brought me to this moment.

I would also like to thank Dr. Mücahit Kozak, who supervised me during my internship in Professor Friedman's lab. His encouragement, positive spirit, and technical knowledge helped me to develop a basic understanding of analog integrated circuit design. This precious knowledge has helped me tremendously during my PhD studies. His positive approach to life inspired me throughout my PhD studies. Thank you Mujo.

I would like to thank Professors Jacob Jorne, Hui Wu, and Zeljko Ignjatovic for serving on my proposal and PhD committees. Their suggestions, comments, and valuable time are highly appreciated.

Last but certainly not least, I would like to thank my PhD colleges: Dr. Guoqing Chen, Dr. Mikhail Popovich, Dr. Vasilis Pavlidis, Emre Salman, Renatas Jakushokas, Ioannis Savidis, and Salçuk Köse. I learned new things from all of you. The technical and personal discussions I had with each of you definitely enriched my PhD experience. Dear gentlemen, thank you for your attention.

Abstract

Relentless scaling of integrated circuits has resulted in significant performance improvements. Although active devices mostly benefit from scaling, passive interconnect networks have degraded in performance with scaling. Interconnect parasitic effects therefore must be considered throughout the design process. Furthermore, novel and innovative design methodologies for interconnect networks are required to maintain high performance in these highly complex integrated circuits.

The focus of this thesis is on three important interconnect networks: clock, data, and power generation and distribution networks. Design and analysis methodologies to improve the performance of these networks have been developed. Specifically, the following three topics have been addressed in this thesis.

Exploiting resonance for distributing high frequency clock signals is a promising technology to reduce power dissipation, clock skew, and jitter. A comprehensive methodology for designing these resonant networks has been developed. A case study of a 5 GHz clock signal within a resonant H-tree network has been demonstrated in a 180 nm CMOS technology, resulting in a substantial 84% reduction in power consumption as compared to a traditional H-tree network.

On-chip resonance has also been used to design a novel data distribution network. By eliminating the need for traditional buffer insertion, a significant reduction in power and latency has been observed. A methodology for designing these networks has been developed. A case study of a 5 Gbps data signal distributed within a 5 mm long interconnect has been

demonstrated, exhibiting 90% and 40% improvements in power consumption and latency, respectively, as compared to repeater insertion and several different exotic techniques.

A distributed rectifier for a buck converter implemented in three-dimensional (3-D) technology has also been developed. The proposed rectifier eliminates the need for a traditional *LC* filter, enabling the on-chip integration of DC-DC converters. A test circuit of the distributed rectifier has been designed for manufacture in the MIT Lincoln Laboratories 150 nm CMOS technology. Additionally, an on-chip hybrid buck converter based on switching and linear DC-DC converters has been developed, demonstrating superior efficiency and conversion range as compared to conventional buck converters.

The development of these novel design methodologies will compensate for the detrimental effects of scaling on interconnect networks. High performance operation of highly complex integrated circuits has been demonstrated to be feasible. A combination of novel design methodologies, materials, and integration technologies, is required for future nanometer integrated circuits.

Contents

Curriculum Vitae ..... **iii**

Acknowledgments ..... **iv**

Abstract **vi**

Contents **viii**

List of Tables ..... **xiv**

List of Figures ..... **xvi**

Chapter 1 Introduction **1**

1.1 Where Are We Heading? Trends and Predictions 3

1.2 Thesis Outline 8

Chapter 2 Global Clock and Data Distribution Networks **11**

2.1 Clock Networks in Synchronous Digital Integrated Circuits 11

2.2 Design Methodologies of Data Distribution Networks 16

2.2.1 Low Swing Interconnects 17

2.2.2 Repeater Insertion 18

2.2.3 Shielding 19

2.2.4 Differential Signaling 21

2.2.5 Bus Swizzling 23

2.2.6 Tapered Interconnects 23

2.3 Criteria for Designing Interconnects and Clock Distribution Networks..... 25
2.3.1 Clock Skew 25
2.3.2 Jitter 28
2.3.3 Power Consumption 33
2.3.4 Delay 34
2.4 Summary 35

Chapter 3 Oscillatory Clock and Signal Modulated Data Distribution Networks 37
3.1 Introduction to Transmission Line Theory 37
3.2 Novel Global Clock Distribution Networks 40
3.2.1 Coupled Standing Wave Oscillators (SWO) 41
3.2.1.1 Theoretical Principles and Design Guidelines 41
3.2.1.2 Circuit Implementation 49
3.2.2 Rotary Traveling Wave Oscillators (RTWO) 50
3.2.3 Resonant Distributed Differential Oscillators (DDO) 56
3.2.3.1 Theoretical Principles and Design Guidelines 56
3.2.3.2 Circuit Implementation 59
3.2.4 Comparison of Clock Distribution Networks 60
3.3 Novel Transmission Line Networks 63
3.3.1 Inductance Dominated Interconnects 63
3.3.1.1 Theoretical Principles 64
3.3.1.2 Circuit Implementation 67
3.3.2 Pulse Signal Interconnects 70
3.3.2.1 Pulsed Current-Mode Signaling 72
3.3.2.2 Pulsed Voltage-Mode Signaling 74

3.3.3 Comparison of Data Distribution Networks 76
3.4 Summary 76

Chapter 4 On-Chip Inductors 78
4.1 Types of On-Chip Inductors 79
4.1.1 Spiral (Planar) On-Chip Inductors 79
4.1.2 Advanced On-Chip Inductors 83
4.1.3 Active Inductors 88
4.1.3.1 Principles of Active Inductors 88
4.1.3.2 Performance Comparison between Passive and Active Inductors 90
4.2 Modeling On-Chip Spiral Inductors 91
4.2.1 Loss Mechanisms in On-Chip Spiral Inductors 91
4.2.2 Models and Performance Metrics of Spiral Inductors 93
4.2.2.1 Basic Model, *Q*, and Self-Resonance of On-Chip Inductors 93
4.2.2.2 Advanced Models of On-Chip Inductors 96
4.3 Design of On-Chip Inductors 98
4.3.1 First Order Inductance Equations 98
4.3.2 Design Practices and Tradeoffs 101
4.4 Summary 104

Chapter 5 Design Methodology for Global Resonant H-Tree Clock Distribution Networks 106
5.1 Introduction 107
5.2 Background and Problem Formulation 108
5.3 Design Guidelines for H-Tree Sector 111

5.3.1 H-Tree Sector Model 111
5.3.2 Model of On-chip Inductor 114
5.3.3 On-Chip Inductor, Capacitor, and Output Resistance of the Driving Buffer 116
5.4 Case Study 121
5.5 Sensitivity of H-Tree Sector 130
5.6 Summary 140

Chapter 6 Quasi-Resonant Interconnects: A Low Power, Low Latency Design Methodology 142
6.1 Introduction 143
6.2 Principle of Quasi-Resonant Interconnect 145
6.3 Interconnect and Spiral Inductor Models 148
6.3.1 Interconnect Model 148
6.3.2 On-Chip Spiral Inductor Model 149
6.4 Design Methodology 150
6.4.1 Transmitter Design 150
6.4.2 Receiver Design 153
6.4.3 Network Input Impedance and Transfer Function 156
6.4.4 Power Consumption Model 158
6.4.5 Signal Delay Model 160
6.4.6 Design Guidelines 161
6.5 Case Study 163
6.6 Simulation Results and Comparison 173
6.7 Summary 177

Chapter 7 Power Supply Generation and Distribution for Three-Dimensional ICs . 178

7.1 Introduction 178

7.2 Background 180

7.3 Principle of a Distributed Rectifier 184

7.4 Design Methodology 188

7.4.1 Physical Structure and Current Load Properties of a 3-D Rectifier 188

7.4.2 Transfer Function of a 3-D Rectifier 192

7.4.3 Efficiency and Area of a 3-D Rectifier 196

7.4.4 Design Guidelines 197

7.5 Case Study 202

7.5.1 Current Load Characterization 202

7.5.2 Conventional Rectifier 203

7.5.3 Distributed 3-D Rectifier 204

7.6 Performance Analysis 207

7.7 Conclusions 211

Chapter 8 3-D Distributed Rectifier Test Circuit 212

8.1 Analysis and Design of a Distributed Rectifier 212

8.2 Design of Current Load, Control, and Noise Measurement Circuits 216

8.3 MIT Lincoln Laboratories 3-D Fabrication Process 220

8.4 Rectifier Layout Design 223

8.5 Summary 229

Chapter 9 A Hybrid Methodology for Wide Range DC-DC Conversion in 3-D Circuits 230

9.1 Introduction 230
9.2 Background 234
9.3 Design Methodology 238
9.3.1 Evaluation of Switching Converter Efficiency 238
9.3.2 A Linear Converter in 3-D Technology 241
9.3.3 A Switching Converter in 3-D Technology 245
9.4 Case Study 250
9.4.1 3.3 Volts to 2.5 Volts Conversion 251
9.4.2 3.3 Volts to 1.0 Volt Conversion 252
9.5 Performance Comparison 257
9.6 Conclusions 261

Chapter 10 Conclusions 263

Chapter 11 Future Research 267

11.1 Spectral Analysis of Jitter in Resonant Clock Distribution Networks 268
11.2 On-Chip Resonant Buck Converters 268
11.3 On-Chip Buck Converter Rectifier (Filter) 269
11.4 On-Chip Inductors in 3-D Circuits 270
11.5 Summary 273

Bibliography 274

Appendix A Modeling of a Resonant H-Tree Clock Network 295

Appendix B Modeling of a Quasi-Resonant Interconnect 304

List of Tables

Table 2.1 Comparison of signal distribution networks 25

Table 3.1 Comparison of clock distribution networks 62

Table 3.2 Comparison of data distribution networks 76

Table 4.1 Coefficients for the current sheet inductance [112] 100

Table 5.1 ESRs of spiral inductors at 5 GHz 116

Table 5.2 Resonant H-tree parameters 122

Table 5.3 Resonant H-tree extracted transmission line parameters 122

Table 5.4 Buffer sizes in the non-resonant case 127

Table 5.5 Sensitivity evaluation of resonant H-tree network 137

Table 5.6 Four cases of unbalanced inductor and capacitor variations 139

Table 6.1 Transistor width of the receiver and transmitter circuits, L = 0.18 μm channel length 165

Table 6.2 Simulated and analytic power consumption and delay of the receiver, transmitter, and QRN 171

Table 6.3 Comparison of QRN model with simulation: (a) power and delay using analytical model, (b) power and delay using simulation, (c) error between analytical model and simulation 174

Table 6.4 Comparison of power consumption and delay 175

Table 6.5 Performance comparison of quasi-resonant method with different approaches ... 176

Table 6.6 Area comparison of repeaters and on-chip spiral inductors for different technologies 177

Table 7.1 *RLC* Interconnect and 3-D via impedances 205

Table 8.1 Electrical parameters of interconnects and 3-D vias..... 214

Table 9.1 Width of the transistors within the 3-D linear converter 252

Table 9.2 Performance comparison of different on-chip DC-DC converters in the literature 261

Table 9.3 Performance summary of the proposed 3-D hybrid converter (this work) 261

List of Figures

Figure 1.1 Taxonomy of digital VLSI challenges..... 2

Figure 1.2 Predictions of skew and jitter [9]..... 5

Figure 1.3 The power budget breakdown for several microprocessors [20]..... 5

Figure 1.4 2005 ITRS predictions: (a) total interconnect length of Metal 1 and five intermediate levels, (b) interconnect *RC* delay for a 1 mm Cu global wire [12]... 7

Figure 2.1 Hierarchical topology of a clock distribution network..... 12

Figure 2.2 Global level distribution networks: (a) H-tree, (b) trunk..... 13

Figure 2.3 Global and semi-global H-tree network..... 14

Figure 2.4 Global H-tree and semi-global grid network..... 15

Figure 2.5 Global H-tree and semi-global length matched serpentine network..... 16

Figure 2.6 Principle of low swing interconnects..... 18

Figure 2.7 Repeater insertion in *RLC* interconnect..... 19

Figure 2.8 Shielded interconnect: (a) passive shielding, (b) active shielding..... 20

Figure 2.9 Differential transmission network..... 21

Figure 2.10 Differential low swing interconnect..... 22

Figure 2.11 Example of bus swizzling..... 23

Figure 2.12 Example of shielded clock lines: (a) tapered interconnect, (b) uniform interconnect [68]..... 24

Figure 2.13 Local data path..... 26

Figure 2.14 Clock timing diagrams..... 27

Figure 2.15 H-tree with eight leafs..... 28

Figure 2.16 Illustration of clock jitter 29

Figure 2.17 Taxonomy of different jitter forms 30

Figure 2.18 Cycle jitter 30

Figure 2.19 Cycle-to-cycle jitter 31

Figure 2.20 Typical section of a clock and interconnect network 33

Figure 3.1 Infinitesimal segment of a transmission line 38

Figure 3.2 Terminated transmission line with load Z_L 40

Figure 3.3 Transmission line terminated with a shorted load 42

Figure 3.4 Standing waves in a lossless transmission line 43

Figure 3.5 Transmission line with distributed negative transconductance 43

Figure 3.6 Equivalent infinitesimal segment 44

Figure 3.7 A standing wave oscillator 46

Figure 3.8 Two coupled SWOs 47

Figure 3.9 Grid of coupled SWOs [70] 48

Figure 3.10 Voltage waveforms of a SWOs grid [70] 48

Figure 3.11 A micrograph of a 10 GHz circuit composed of eight coupled SWOs [70] 49

Figure 3.12 Negative resistance circuit 49

Figure 3.13 Clock buffer 50

Figure 3.14 Operation of a traveling wave loop: (a) in a transmission line, (b) in a square shaped transmission line, (c) in a differential closed loop 51

Figure 3.15 Rotary clock ring 52

Figure 3.16 Section of a rotary clock ring 53

Figure 3.17 Rotary clock distribution network 55

Figure 3.18 A micrograph of a 965 MHz four coupled RTWOs [71] 55

Figure 3.19 Resonant differential oscillator 57

Figure 3.20 Circuit model of the DDO 58

Figure 3.21 Transconductance circuit 59

Figure 3.22 Clock buffer 60

Figure 3.23 Micrograph of a 1.1 GHz DDO with four differential on-chip inductors [74] 60

Figure 3.24 Example of *LC* dominated interconnects: (a) transfer function, (b) attenuation coefficient, (c) phase velocity 66

Figure 3.25 Block diagram of the near speed of light interconnect network 67

Figure 3.26 Modulation of the transmitted signal 68

Figure 3.27 Passive ring mixer 68

Figure 3.28 Active double-balanced mixer 69

Figure 3.29 Self biased differential amplifier 69

Figure 3.30 Micrograph of a 20 mm interconnect long transmission network [81] 70

Figure 3.31 Spectrum of a periodic pulse signal: (a) periodic pulse with width T_s and a period T_b, (b) spectral amplitudes of a pulse with T_s = 50 ps and T_b = 250 ps, and (c) spectral amplitudes of a pulse with T_s = 100 ps and T_b = 200 ps 71

Figure 3.32 Pulsed current interconnect network 73

Figure 3.33 Micrograph of a 3 mm pulsed current interconnect [84] 73

Figure 3.34 Pulsed wave interconnect: (a) generation and transmission of a pulse signal, and (b) network model 74

Figure 4.1 Planar spiral inductors: (a) square inductor [89], (b) octagonal inductor [90], (c) circular inductor [91] 80

Figure 4.2 Patterned ground shield [93] 81

Figure 4.3 Deep trench inductor [96]: (a) side view, (b) three dimensional view 82

Figure 4.4 Reduced loss inductors [90]: (a) tapered spiral, (b) symmetric spiral 82

Figure 4.5 Three layer stacked inductor [94] 83

Figure 4.6 Suspended inductor [95]: (a) single level, (b) double level 84

Figure 4.7 A toroidal inductor [97] 84

Figure 4.8 Vertical planar inductor [98]: (a) before the PDMA process, (b) after the PDMA process 85

Figure 4.9 A bond wire inductor [99] 86

Figure 4.10 Magnetic thin film inductors: (a) thin film ferrite [100], (b) magnetic thin film inductor [100], (c) cross section of metal and magnetic layers [101], (d) spiral inductor with magnetic film [101] 87

Figure 4.11 A typical configuration of an active inductor 89

Figure 4.12 Generic lumped model of a spiral inductor 93

Figure 4.13 On-chip inductor model considering high frequency effects 96

Figure 4.14 An on-chip spiral inductor model: (a) lumped model, (b) simplified T model ... 97

Figure 4.15 A square spiral inductor 99

Figure 5.1 H-tree sector with on-chip inductors and capacitors 109

Figure 5.2 A global clock distribution network, consisting of 16 resonant clock sectors and a total of 256 leafs 110

Figure 5.3 Distributed model of a transmission line 111

Figure 5.4 Distributed *RLC* network representation of an H-tree network 112

Figure 5.5 Resonant H-tree network simplified to a distributed *RLC* line 113

Figure 5.6 Simplified model of an on-chip spiral inductor 115

Figure 5.7 Effective series resistance of spiral inductors 115

Figure 5.8 One-port network driven by a voltage source 116

Figure 5.9 Structure of resonant H-tree sector 122

Figure 5.10 Design tradeoffs for an H-tree sector: (a) Output resistance as a function of the spiral inductor, (b) ρ as a function of the on-chip spiral inductor 124

Figure 5.11 Output resistance and ρ as a function of the on-chip capacitance with L_s = 2 nH 125

Figure 5.12 Output waveform at the leaf nodes 126

Figure 5.13 Magnitude of the transfer function and input impedance 126

Figure 5.14 H-tree sector with buffers 127

Figure 5.15 Comparison of power consumption between analytic model and SpectreS Spice simulation 128

Figure 5.16 Normalized voltage swing at the leaf node 129

Figure 5.17 Magnitude of the transfer function as a function of the inductance at a 5 GHz operating frequency 130

Figure 5.18 Voltage swing and power consumption as a function of variations in the output resistance 131

Figure 5.19 Voltage swing and power consumption as a function of variations in the on-chip inductor width 132

Figure 5.20 Voltage swing and power consumption as a function of variations of the on-chip capacitor 133

Figure 5.21 Voltage swing and power consumption as a function of variations in the signal line width 135

Figure 5.22 Voltage swing and power consumption as a function of variations in the shield line width 136

Figure 5.23 Voltage swing and power consumption as a function of spacing variations between the signal and shield lines 137

Figure 5.24 Three cases of voltage swing and power consumption as a function of four simultaneous variations: Case 1 - buffer output resistance, spiral inductors, capacitors, and signal line width are varied; Case 2 - buffer output resistance, spiral inductors, capacitors, and shield line width are varied; Case 3 - buffer output resistance, spiral inductors, capacitors, and spacing are varied 138

Figure 6.1 Quasi-resonant network 146

Figure 6.2 Model of an on-chip spiral inductor: (a) structure of an octagonal on-chip spiral inductor, (b) lumped model of the spiral inductor 149

Figure 6.3 Example of transmitting a “1011” bit stream 151

Figure 6.4 Transmitter circuit: (a) modulation scheme, (b) gate level circuit, (c) transistor level circuit, (d) signal waveforms 153

Figure 6.5 Receiver circuit: (a) sample and hold circuit, (b) transistor level circuit, (c) delay element circuit, (d) signal waveforms 155

Figure 6.6 Quasi-resonant network 157

Figure 6.7 One-port network driven by a voltage source 159

Figure 6.8 Flow diagram of QRN design process 163

Figure 6.9 Layout of a resonant transmission line network 164

Figure 6.10 A design example of a 5 mm long interconnect operating at a 5 Gbps transmission frequency: (a) minimum power-delay product as a function of inductance, (b) insertion location as a function of inductance, (c) driver resistance as a function of inductance 166

Figure 6.11 Power and delay: (a) power consumption, (b) signal delay 167

Figure 6.12 Frequency response of the transfer function: (a) magnitude, (b) phase..... 168

Figure 6.13 QRN input impedance: (a) real part, (b) imaginary part 168

Figure 6.14 Magnitude of the transfer function as a function of the inductance at a 5 GHz operating frequency..... 169

Figure 6.15 Ten bit data stream example: (a) data at input of transmitter, (b) data at output of transmitter, (c) data at input of receiver, (d) data at output of receiver..... 170

Figure 6.16 Noise analysis: (a) noisy power supply and ground signals, (b) eye diagram of an interconnect with repeaters, (c) eye diagram of quasi-resonant interconnect ... 172

Figure 7.1 Three-dimensional circuit with multiple power supplies 179

Figure 7.2 Conventional DC-DC converter: (a) buck converter circuit [139], (b) signal at the output of the power MOSFETs (node A)..... 181

Figure 7.3 *RLC* transmission line..... 185

Figure 7.4 Transfer function of the magnitude of an interconnect for different capacitive loads 185

Figure 7.5 Magnitude of transfer function: (a) single stage *RC* filter, (b) two stages *RC* filter, (c) *LC* rectifier, (d) distributed rectifier..... 187

Figure 7.6 A distributed rectifier 189

Figure 7.7 Current load profile in high speed digital circuits 190

Figure 7.8 Required magnitude of the transfer function..... 198

Figure 7.9 Design space of a distributed rectifier: (a) modified duty cycle as a function of R_d, (b) duty cycle ratio as a function of R_d, (c) capacitance as a function of R_d, (d) rectifier efficiency as a function of R_d..... 199

Figure 7.10 Effect of interconnect length on performance: (a) capacitance as a function of interconnect length, (b) rectifier efficiency as a function of interconnect length 201

Figure 7.11 Magnitude of the transfer function of an *LC* rectifier..... 204

Figure 7.12 Magnitude of the transfer function at ω_s for different lengths and capacitances206

Figure 7.13 Simulation and model of the magnitude of the transfer function of a 3-D rectifier and a conventional *LC* rectifier circuit 206

Figure 7.14 DC voltage at the output of a distributed rectifier 207

Figure 7.15 Conventional *LC* rectifier: (a) capacitance, (b) inductance 209

Figure 7.16 3-D distributed rectifier: (a) efficiency, (b) total capacitance per plane, (c) total area per plane 210

Figure 8.1 A distributed rectifier of a DC-DC power converter in a three-plane structure .. 213

Figure 8.2 Current load profile 214

Figure 8.3 Magnitude of the transfer function for different line lengths and capacitor magnitudes 215

Figure 8.4 Current load controlled by switches 216

Figure 8.5 Control logic for the current mirrors 217

Figure 8.6 Post-fabrication control circuit 217

Figure 8.7 Distributed rectifier: (a) current load, (b) output voltage supply..... 218

Figure 8.8 Power supply noise detector..... 219

Figure 8.9 Three independent tiers fabricated using a conventional 2-D technology [156]. 220

Figure 8.10 Tier 2 aligned and bonded to tier 1 [156] 221

Figure 8.11 Tier 2 electrically connected to tier 1 with Tungsten 3-D vias [156] 221

Figure 8.12 Formation of the back side vias and metal on tier 2 [156] 221

Figure 8.13 Tier 3 is aligned and bonded to the back side metal of tier 2 [156]..... 221

Figure 8.14 Formation of the back side metal on tier 3 [156]..... 222

Figure 8.15 Layer thicknesses within the three-tier structure [156] 222

Figure 8.16 Four blocks within the power delivery test circuit 223

Figure 8.17 Schematic illustration of the distributed rectifier circuit 224

Figure 8.18 Physical layout of the 3-D rectifier circuit (all three planes)..... 224

Figure 8.19 P-type capacitor: (a) model, (b) layout..... 226

Figure 8.20 Distributed rectifier structure: (a) plane A (bottom), (b) plane B (middle), (c) plane C (upper)..... 227

Figure 8.21 Circuitry on plane A: (a) post-fabrication control circuit, (b) power supply noise measurement circuit 228

Figure 8.22 Circuitry on plane B: (a) ring oscillators and buffers, (b) switched current loads 228

Figure 9.1 2-D and 3-D cells: (a) 2-D horizontal plane, (b) three stacked planes 231

Figure 9.2 3-D circuit: (a) multi-core processor combined from stacked planes, (b) cross section of a 3-D circuit..... 232

Figure 9.3 Different partitions of voltage domains within a 3-D circuit 233

Figure 9.4 Linear converter..... 235

Figure 9.5 Efficiency as a function of conversion ratios in different technologies 239

Figure 9.6 Efficiency of linear and switching converters with dual-V_{dd} input voltage (the two dashed lines represent the two input voltages)..... 241

Figure 9.7 Linear converter within a 3-D circuit 242

Figure 9.8 Current mirror bias circuit..... 243

Figure 9.9 A single-ended differential amplifier 243

Figure 9.10 A typical current load profile in a high performance processor 244

Figure 9.11 A switching converter with the 3-D distributed rectifier..... 246

Figure 9.12 Voltage signal at the input of the rectifier 246

Figure 9.13 Power breakdown of the power MOSFETs and rectifier filter, distributing a 200 mA peak current 250

Figure 9.14 Output voltage of the linear 3-D converter..... 253

Figure 9.15 Transient output voltage response: (a) 200 mA increase in current load, (b) 100 mA decrease in current load..... 254

Figure 9.16 Power components of the 3-D switching converter, distributing 2.5 Amperes of peak current 255

Figure 9.17 Power MOSFET widths as a function of $R_{MOS}/R_{MOS,opt}$..... 256

Figure 9.18 Output voltage of the switching converter 257

Figure 9.19 MOS capacitor density for different technologies..... 258

Figure 9.20 Capacitance per plane as a function of V_{out}/V_{in} 259

Figure 9.21 Power efficiency of different approaches as a function of V_{out}/V_{in} 260

Figure 11.1 Resonant buck converter 269

Figure 11.2 Proposed filter within the buck converter..... 270

Figure 11.3 On-chip inductor topologies in 3-D circuits: (a) planar 2-D inductor, (b) stacked planar 2-D and vertical 3-D inductors..... 272

Chapter 1

Introduction

> *"So I toyed around in my mind: How can I do this experiment -- I was still being lazy; of course I could have evacuated the thing, but that was the sample that had been fixed up to put down in this thermos and it wasn't exactly conducive to evacuation; and it would have taken some time to do this.*
>
> *I think I suggested, 'Why, John, we'll wax the point.' One of the problems was how do we do this, so we'd just coat the point with paraffin all over, and then we'd push it down on the crystal. The metal will penetrate the paraffin and make contact with the semi-conductor, but still we'd have it perfectly insulated from the liquid, and we'll put a drop of tap water around it. That day, we in principle, created an amplifier."* -- Walter Brattain, January 1964 [1]
>
> *"This circuit was actually spoken over and by switching the device in and out a distinct gain in speech level could be heard and seen on the scope presentation with no noticeable change in quality."* -- Walter Brattain, December 24, 1947, lab notebook [1]

Events that took place at Bell Laboratories between November 17 and December 23, 1947 ignited the microelectronics era. Inventing the transistor led to the most important economic development of the 20th century: the information revolution. The driving force at the heart of what has proven to be an exceptionally productive industry is microelectronic integrated circuit technology. Since the invention of the integrated circuit in 1959 by Fairchild Semiconductor, the productivity of silicon integrated circuit technology has increased more than a billion fold while performance has increased more than one hundred thousand fold [2]. The accomplishments of the microelectronic integrated circuit technology can be attributed to the continuous scaling of semiconductor devices and the use of digital signal processing paradigms. Inevitably, the leading semiconductor technology, *i.e.*,

complementary metal-oxide-semiconductor (CMOS), is approaching both physical and economic limits that will severely restrain the current rate of advance.

The taxonomy of the current challenges of very large scale integration (VLSI) systems and the relations among the sub-networks are shown in Figure 1.1. Noise, power, and speed are the three vertices of the triangle, indicating the primary performance criteria of digital VLSI circuits.

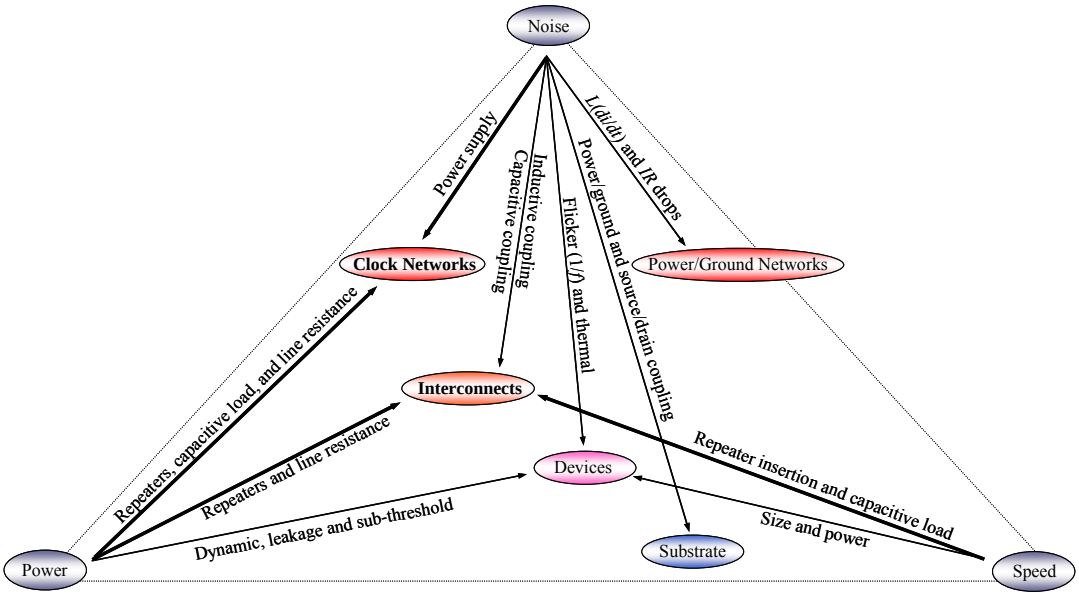


Figure 1.1 Taxonomy of digital VLSI challenges

The location inside the triangle of Figure 1.1 illustrating the primary design criteria in VLSI circuits, including clock networks, power/ground networks, interconnect, devices, and substrate, are analogous to the physical location within an integrated circuit. The arrows show the relationship among the different components and performance criteria, as well as the primary sources of performance degradation. Clock and interconnect distribution networks

(bold titles) and the related design issues (bold arrows) are the primary focus of this thesis with respect to noise, power, and speed.

1.1 Where Are We Heading? Trends and Predictions

The prediction of the international technology roadmap for semiconductors (ITRS) [3] ends in the year 2020 with more than a 100 billion transistors occupying a single IC with a typical transistor channel length of 10 nm. At this transistor size, semiconductor technologies have reached fundamental physical limitations, requiring the development of new technologies. In order to support a system with 10 nm length transistors, four requirements have been identified [4]:

1) Develop a cost efficient 10 nm fabrication technology;
2) Devise effective methods to handle low yield and defective devices;
3) Manage power and heat dissipation; and
4) Develop global interconnect technology to complement 10 nm transistors.

Interconnect networks included in the requirements for future technologies since interconnects do not scale well, consume a significant amount of power, and place severe limitations on processor clock speeds.

Clock signals in digital systems are simultaneously distributed to physically remote locations across an integrated circuit (IC). The clock signal provides a time reference that permits different parts of a circuit to operate in the correct order, thereby producing correct logical operation [7]. A clock signal is usually distributed from a common global source through metal interconnect networks and clock drivers, introducing delay. Unfortunately, the delay at every point on an IC cannot be precisely maintained, due to delay uncertainty [8].

Clock skew, which is the difference in the arrival time of the clock signal between sequentially-adjacent registers, can lead to catastrophic logic failure [7]. Another undesirable effect is clock jitter which occurs when the edges of the clock signal fluctuate in time. This behavior occurs due to imperfections in the clock generator and power supply noise [18]. Changes in the coupling capacitance and variations of the input capacitance of the registers also add random noise while increasing jitter.

Predictions regarding skew and jitter for different future technologies are depicted in Figure 1.2 [9]. Based on these trends, by the year 2013, at a critical node technology of 32 nm, skew and jitter will dominate synchronous performance, consuming 62% of the total clock period. Furthermore, synchronizing digital circuits at high frequencies has become more difficult since interconnect geometries do not scale as easily as transistors, producing longer wire delays.

Additionally, during each cycle, the entire clock capacitance is charged and discharged to ground, dissipating the stored energy as heat. In modern microprocessors, the largest portion of the power budget is consumed by the clock distribution network [20], as depicted in Figure 1.3. In this example, 40% of the power budget is consumed by the clock distribution network within the processor, such as the DEC Alpha 21164 [10] and the Motorola MCORE micro-RISC [11]. As technology scales, this trend will worsen, requiring the development of novel design methodologies and technologies for low power clock distribution networks.

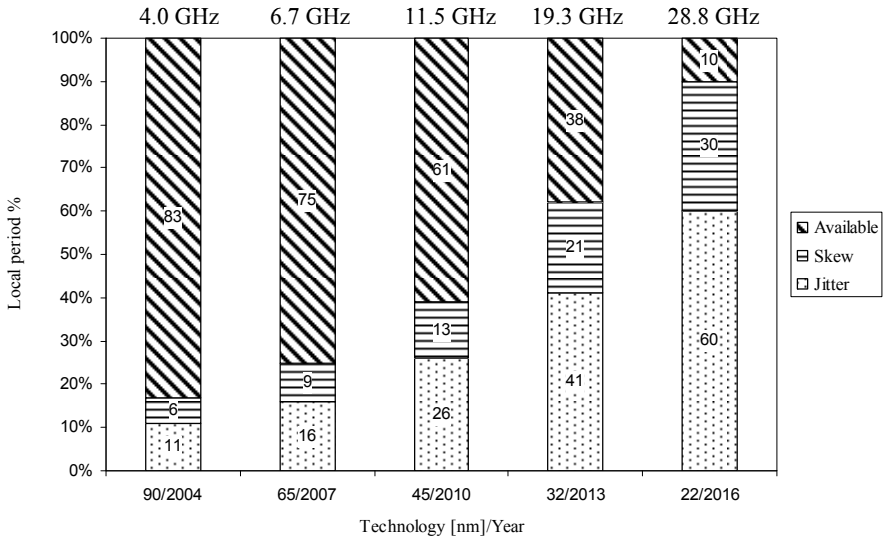


Figure 1.2 Predictions of skew and jitter [9]

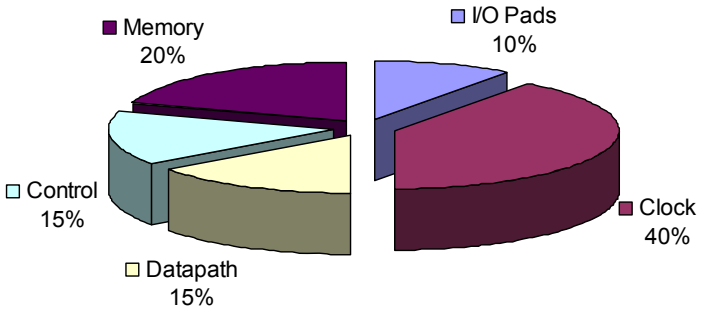


Figure 1.3 The power budget breakdown for several microprocessors [20]

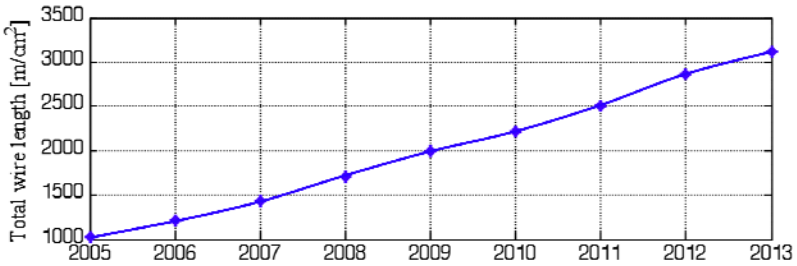
An alternative synchronization scheme has been proposed in the form of globally asynchronous, locally synchronous (GALS) systems [5]. In this scheme, individual digital modules are treated as synchronous networks, while the operation of the inter-module communication is not fully synchronous. Once the number of these multiprocessors increases, intra- and inter-module communication among the processors and the shared memory will require significant time [4]. An additional difficulty in GALS networks is the design of

reliable circuit interfaces to manage communication among the synchronized blocks while avoiding metastability between synchronous and asynchronous logic domains [6].

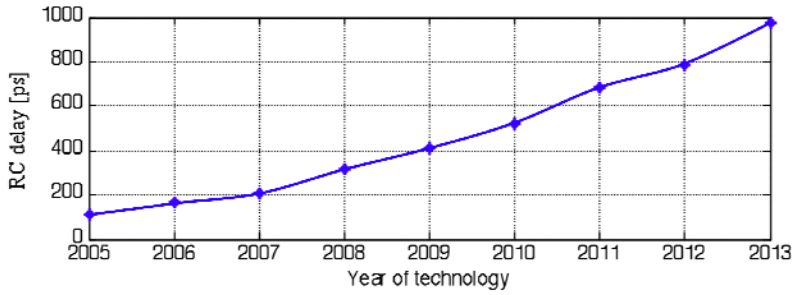
A related challenge in high performance, high complexity integrated circuits is the on-chip interconnect [12]. Transmitting clock and data signals over large die areas requires long and thicker interconnections among the various circuit modules. Consequently, as technology scales, the interconnect cross section decreases while operating frequencies increase. The impact of these trends on high performance systems is significant. Long interconnects with smaller cross sections exhibit increased capacitance and resistance, resulting in greater power consumption, latency, and signal attenuation. Furthermore, wire inductance can no longer be ignored due to high signal frequencies and long wire lengths. Two figures of merit that determine the length of interconnect in which inductance effects are significant have been provided in [13].

Some of these challenges are evident in the international technology roadmap for semiconductors (ITRS) [12], as shown in Figure 1.4. According to ITRS predictions, the total on-chip wire length will increase linearly with technology, reaching above 2000 m/cm^2 by the end of this decade (see Figure 1.4(a)). This trend is consistent with the assumption that long interconnects will occupy large die area in future technologies. The prediction for *RC* interconnect delay is depicted in Figure 1.4(b). Note that the *RC* delay increases quadratically with technology, reaching above 500 ps (for a 1 mm Cu global wire) by 2010. Since both *R* and *C* are proportional to the wire length, the *RC* time constant is proportional to the square of the line length, placing limitations on the bandwidth of the global interconnects. Additionally, as the width of the interconnects decreases, the *RC* time constant increases due to increased resistivity and a decreased cross-section. These trends demonstrate that long interconnects with small cross sections exhibit increasing latency and delay uncertainty. To

combat these phenomena, traditional repeater insertion methods have been widely developed [50]-[61]. Unfortunately, as interconnect lengths increase, a larger number of repeaters are required. This results in significant power dissipation, increased delay and delay uncertainty, and additional area.



(a)



(b)

Figure 1.4 2005 ITRS predictions: (a) total interconnect length of Metal 1 and five intermediate levels, (b) interconnect *RC* delay for a 1 mm Cu global wire [12]

1.2 Thesis Outline

This thesis is organized into ten chapters. Global clock and data distribution networks are described in Chapter 2. In this chapter, the characteristics and topologies of conventional clock distribution networks are discussed. Traditional design methodologies for high speed, low power, and low noise data transmission networks are also presented. Specifically, low swing interconnects, repeater insertion, shielding, differential interconnects, bus swizzling, and tapered interconnects are discussed and compared with respect to power consumption, noise, area, and transmission speed. Finally, to analyze the performance of these networks, design metrics such as clock skew, signal jitter, power consumption, and propagation delay are reviewed.

Advanced clock and data distribution networks in the form of oscillatory clock and signal modulation, respectively, are presented in Chapter 3. To support the discussion of these networks, a brief introduction to transmission line theory is provided. Three approaches for multi-gigahertz clock distribution networks are reviewed: coupled standing wave oscillators, rotary traveling wave oscillators, and resonant distributed differential oscillators. Advanced communication links in the form of inductance dominated interconnects and pulse signal interconnects are discussed. A comparison between the two interconnect families is also provided.

The importance of on-chip inductors for high speed analog and, recently, digital circuits require the development of high performance on-chip inductors. Therefore, insight into the characterization and analysis of on-chip inductors is provided in Chapter 4. Specifically, on-chip inductors are classified into common families. Models of spiral planar inductors, design guidelines, and tradeoffs are also discussed.

The following two chapters propose novel design methodologies for clock and data distribution networks, respectively. A design methodology for global H-tree clock distribution network is presented in Chapter 5. In this chapter, the background and problem formulation of resonant clock networks are described, providing guidelines for designing these structures. A case study is presented to demonstrate the design of these networks. The sensitivity of a resonant clock distribution network to certain design parameters is also examined.

A low power, low latency design methodology for quasi-resonant interconnects is presented in Chapter 6. Techniques for exploiting on-chip resonance in integrated circuits as well as the principle of resonance for data transmission are presented. A quasi-resonant interconnect design methodology is described, followed by a case study demonstrating a 5 Gbps data signal distribution along a 5 mm interconnect. Simulation results and a comparison to other communication schemes are provided.

A methodology for designing an on-chip distributed rectifier of a 3-D buck converter is described in Chapter 7, followed by a case study demonstrating a 3.3 to 1.2 volt conversion while providing a peak current of 700 mA to a designated plane. Performance analysis of the distributed rectifier as well as a comparison to conventional filters is also provided.

In order to evaluate and validate the operation of the proposed rectifier, a test circuit has been designed in the MIT Lincoln Laboratories 150 nm CMOS 3-D technology. Several design issues related to the physical layout have been addressed and are described in Chapter 8.

A novel hybrid on-chip converter for application to 3-D circuits based on linear and switching converters is described in Chapter 9. The proposed buck converter provides high efficiency for a wide range of conversion ratios while distributing 1 Ampere of DC current.

The distributed rectifier described in Chapter 7 is included in the design methodology of the proposed hybrid buck converter.

In Chapter 10, the research work described within this thesis is summarized. Future research directions are described in Chapter 11. Possible topics for future research include spectral analysis of jitter noise in resonant clock distribution networks. A resonant on-chip DC-DC converter circuit as a higher efficiency alternative to conventional converters as well as a novel approach for a buck converter filter is proposed. Finally, different topologies of 3-D on-chip inductors in 3-D circuits are explored.

Chapter 2

Global Clock and Data Distribution Networks

Most modern VLSI circuits are synchronized by a clock distribution network that spans across the physical die area. Synchronizing a large number of registers is an effective method to simultaneously manage the data flow process. Clock distribution networks provide a common framework for synchronizing data signals within high speed processors.

Transmission of data signals among the synchronized processing blocks results in complex interconnect networks. In modern processors, these interconnect networks have become a performance bottleneck, exhibiting high power consumption and long latency. To improve interconnect performance, novel methods for high speed, low power, and low noise interconnects have been proposed.

In this chapter, the characteristics and topologies of conventional clock distribution networks are discussed in section 2.1. Traditional design methodologies for high speed, low power, and low noise data transmission networks are presented in section 2.2. To analyze the performance of these networks, design metrics are reviewed in section 2.3. Finally, a summary is provided in section 2.4.

2.1 Clock Networks in Synchronous Digital Integrated Circuits

Conventional on-chip clock distribution networks generally consist of a series of buffers and interconnects that distribute a clock signal while minimizing delay uncertainty and power

consumption. Most microprocessors utilize multiple distributed buffer stages that drive a small portion of the network load. A typical hierarchy of a high performance clock distribution network is shown in Figure 2.1.

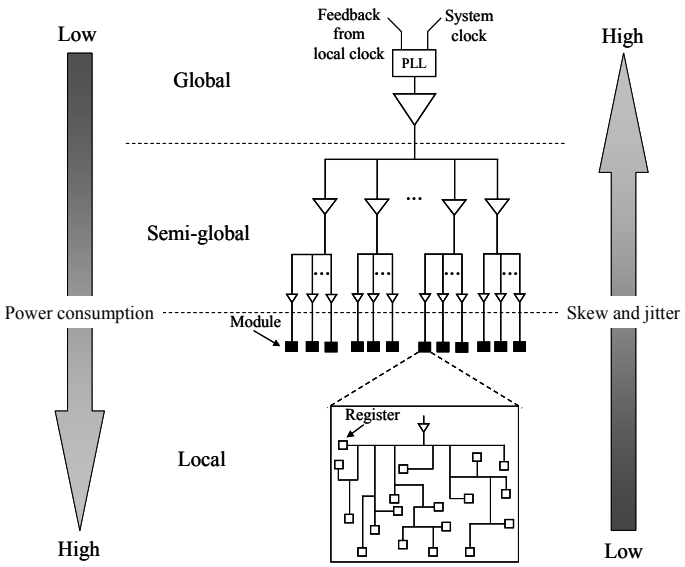


Figure 2.1 Hierarchical topology of a clock distribution network

A clock network is typically partitioned into three domains: global, semi-global, and local networks. The global clock distribution network spans the largest area and is driven by a global buffer at the root of the network. Design issues at the global level are mostly related to signal integrity and delay uncertainty due to the accumulation of clock skew and jitter over long signal paths. Global networks, therefore, introduce the highest skew and jitter to the clock signal. Global networks, however, exhibit smaller capacitance than local networks, relatively little power is dissipated at this level [23].

The semi-global distribution network may consist of several levels of buffering and is usually considered to be part of the global clock network. The semi-global level spans over

less area and drives less capacitance than the global network, resulting in relatively small power consumption. Skew and jitter are accumulated and propagated through the buffers due to power supply noise and process variations. Power dissipation, clock skew, and jitter are moderately affected by the semi-global network, as illustrated in Figure 2.1.

The final level of a clock distribution network is the local network. This network drives the final load of the clock network (*i.e.*, registers), dissipating a large portion of the power. The local clock network dissipates an order of magnitude more power than the global and semi-global networks combined [23]. Note that the clocked elements at the local level are typically connected by asymmetric interconnections if automated place and routing tools are used.

The design of the different hierarchical levels within a clock distribution network depends upon the architecture and performance characteristics. In modern processors, the global level is commonly implemented as a symmetric H-tree structure [32] or a trunk [57], as shown in Figure 2.2. These topologies are advantageous at the global level of the clock distribution network due to the simple symmetric pattern that covers a large area while minimizing the distance between the PLL and the leaf nodes.

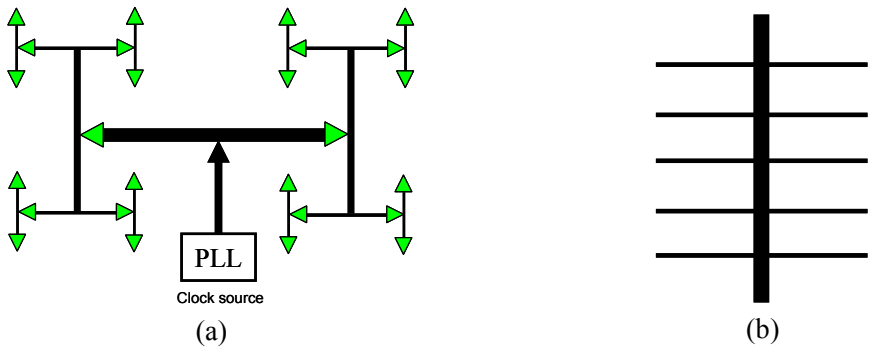


Figure 2.2 Global level distribution networks: (a) H-tree, (b) trunk

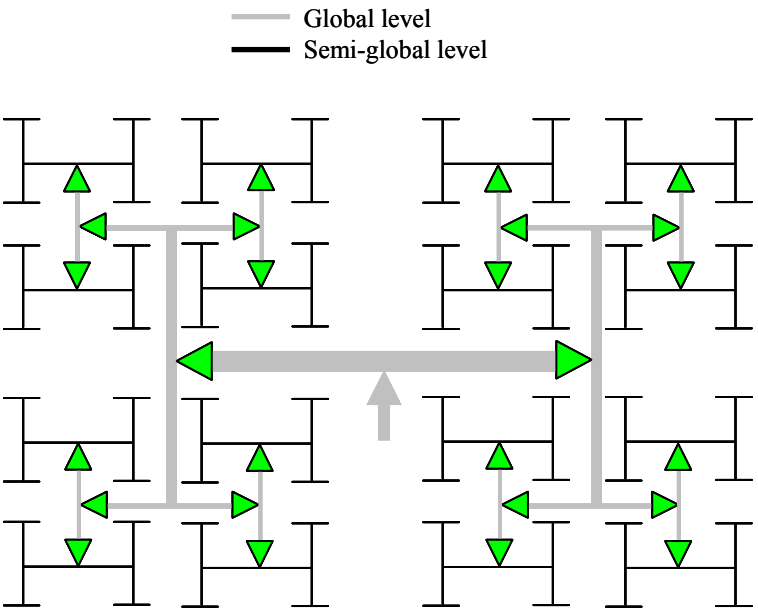


Figure 2.3 Global and semi-global H-tree network

In contrast to the global level, several alternatives are suitable for the semi-global portion of the clock distribution network. Numerous semi-global clock distribution networks are implemented in the form of an H-tree topology due to the simplicity of the structure, exploiting the inherent advantages of these structure [33], [34], [35], as illustrated in Figure 2.3. Alternatively, it is common to connect a large mesh to the global H-tree as the semi-global clock network [36], [37], [38], [39], as shown in Figure 2.4. The grid configuration provides a symmetric structure, reducing skew and jitter; however, a large capacitive load is introduced.

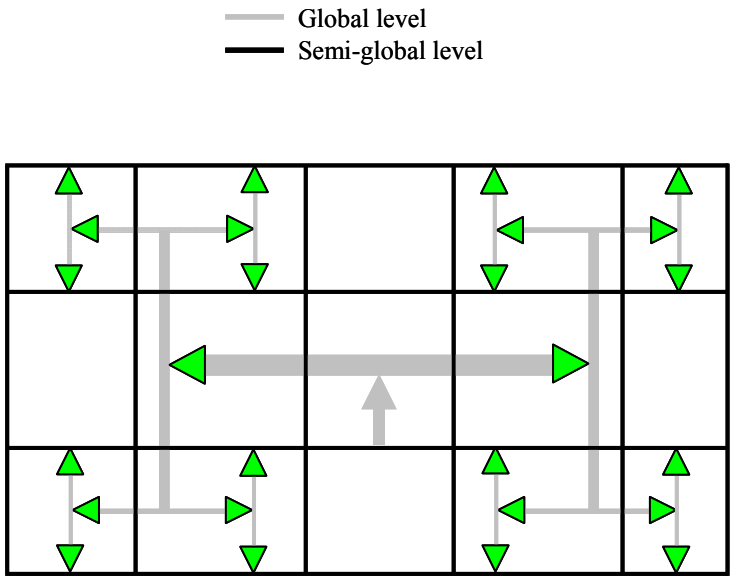


Figure 2.4 Global H-tree and semi-global grid network

A different approach for controlling clock skew is to utilize length-matched serpentines to distribute the clock at the semi-global level [40], as shown in Figure 2.5. Similar to the H-tree topology, the serpentine lines produce nominally zero skew; however, additional latency and load capacitance are introduced as compared to an H-tree.

The design of the final portion of the clock distribution network at the local level is usually achieved by manual (or semi-automated) design methods due to the relatively limited span of these local networks. Since most modules incorporate irregular interconnect structures, local networks typically exhibit an asymmetric tree structure that can be length matched to satisfy specific skew specifications.

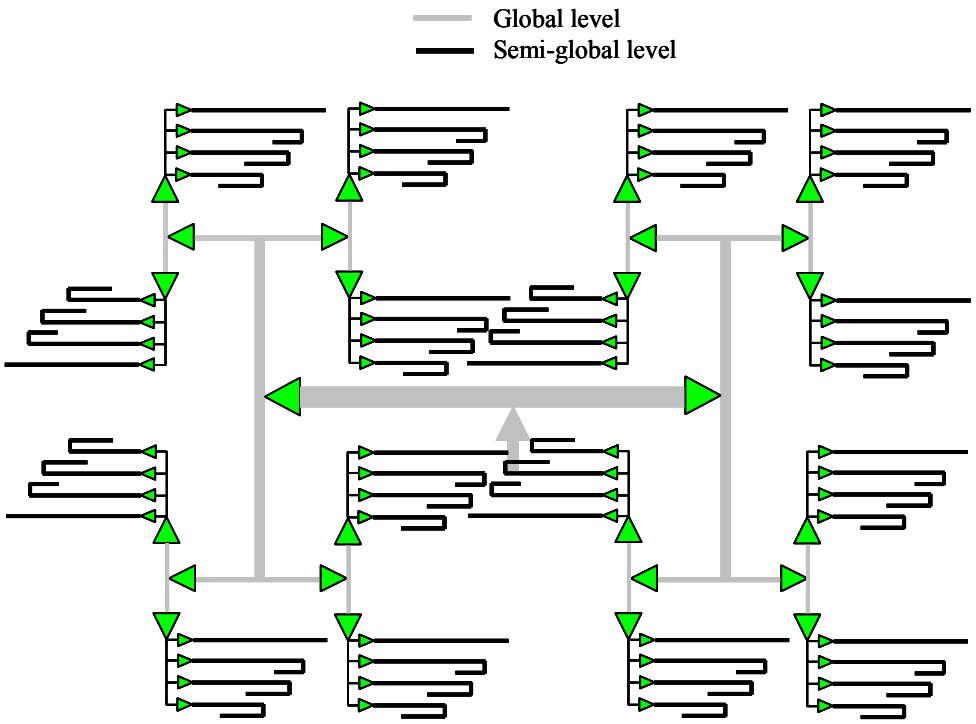


Figure 2.5 Global H-tree and semi-global length matched serpentine network

2.2 Design Methodologies of Data Distribution Networks

To satisfy high performance operation of on-chip interconnects in a high complexity VLSI environment, a variety of signaling techniques have been proposed. These techniques target a decrease in power consumption, coupling noise, and physical area while increasing signal transmission speed. Since these performance characteristics often exhibit opposing trends, optimizing for all performance requirements is difficult and sometimes not feasible. Tradeoffs in design to achieve different performance goals are therefore frequently encountered.

The most common design techniques for interconnects are briefly described in this section. Specifically, low swing interconnects, repeater insertion, shielding, differential

interconnects, bus swizzling, and tapered interconnects are discussed and compared with respect to power consumption, noise, area, and transmission speed in the following subsections.

2.2.1 Low Swing Interconnects

The primary power consumption component in typical CMOS circuits is the dynamic switching power. During voltage transitions, dynamic power is dissipated while charging and discharging the parasitic capacitances within a circuit. The average dynamic power consumed by a CMOS gate driving a capacitive node during a time period T_s is

$$P = \frac{1}{T_s} C_L V_{DD} V_{swing}, \tag{2.1}$$

where C_L is the equivalent lumped capacitance of the drain-to-body junction capacitance of the driver gate, the equivalent capacitance of the interconnect, and the gate oxide capacitance of the load gate. The voltages V_{dd} and V_{swing} are the power supply and voltage transition across C_L during a time period T_s, respectively. The power consumption can be reduced by increasing T_s or reducing C_L, V_{dd}, and V_{swing}, as evident from (2.1). For high performance circuits, however, the most practical approach to reduce power consumption is to decrease V_{dd} and V_{swing}, resulting in low swing operation.

The concept of low swing interconnects is illustrated in Figure 2.6. This scheme usually requires multiple power supply voltages to generate a low swing signal, as well as transmitter and receiver circuits at the near end and far end of the interconnect, respectively. The

transmitter and receiver are interface circuits that convert a full swing signal to a low swing signal and back to a full swing signal, respectively [41], [42], [43], [44], [45].

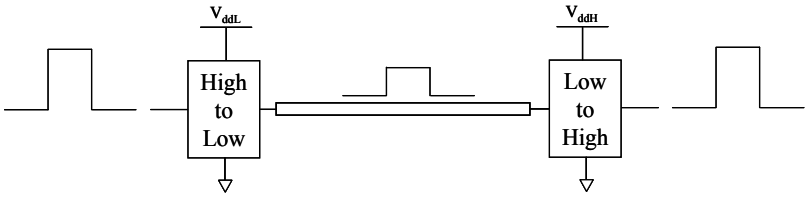


Figure 2.6 Principle of low swing interconnects

The primary inherent disadvantages of low swing interconnects are the lower speed and additional power consumption due to the interface circuits. Additionally, a low signal-to-noise ratio is exhibited by the low swing signals propagating across long resistive interconnects. Careful design of low swing interconnect networks, however, can result in significant power savings, achieving up to 60% improvement in power [43], [44], [45].

2.2.2 Repeater Insertion

Repeater insertion is a widely used technique for reducing propagation delay and coupling noise. As an example, consider the delay of an *RC* line, $0.38RCl^2$ [49], which is proportional to the square of the interconnect length *l*. By dividing the interconnect into *m* *s*horter segments, as shown in Figure 2.7, the delay is linearized with respect to the length $0.38RCl^2 / m$. Furthermore, the coupling noise between interconnects is also reduced due to the shorter length between neighboring lines. These repeaters, however, introduce additional delay, as well as increased power consumption, area, skew, and jitter.

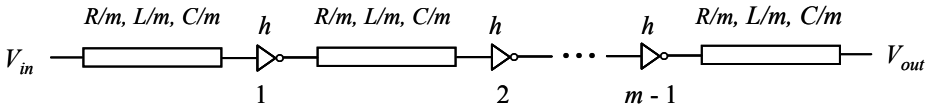
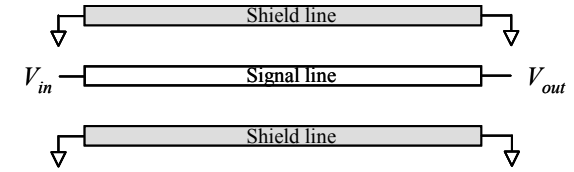


Figure 2.7 Repeater insertion in *RLC* interconnect

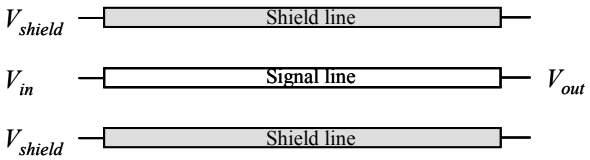
A variety of design methodologies have been proposed for efficiently inserting repeaters in *RC* [50] and *RLC* lines [51], as well as in interconnect trees [56]. The primary goal of these repeater insertion techniques is to determine the optimal number (m) and size of the repeaters (h), while achieving multiple design goals. For example, optimization for minimum delay is proposed in [52], [53], [53], [54]. Repeater insertion for minimum delay while considering power consumption is presented in [55], [58], [59]. Minimizing delay and crosstalk noise is described in [60]. Optimization of power consumption considering thermal effects is proposed in [61].

2.2.3 Shielding

Shielding techniques are widely used in integrated circuits to reduce capacitive and inductive coupling. By inserting a shield line between signals lines, changes in the effective interconnect capacitance is significantly reduced, resulting in less delay uncertainty. Inductive coupling and self inductance can also be reduced by inserting shield lines, since the shield line provides a nearby current return path [62].



(a)



(b)

Figure 2.8 Shielded interconnect: (a) passive shielding, (b) active shielding

Shielding techniques can be categorized into two types: passive shielding and active shielding, as shown in Figure 2.8. In passive shielding, the shield lines are typically connected to the power or ground networks. In active shielding, the signal switching activity on the shield lines is exploited to change the effective capacitance and loop inductance, thereby reducing delay and coupling noise [63]. Generally, switching the shield and signal lines in the same direction can reduce delay, improve slew rate, and increase the inductive coupling noise [62]. Note that clock distribution networks are usually shielded due to the reduced delay uncertainty and high speed and integrity requirements of the clock signals [64]. The primary disadvantage of shielding is the additional area occupied by the shield lines (and higher load capacitance in active shielding).

2.2.4 Differential Signaling

The use of differential data transmission has been adapted from off-chip and chip-to-chip communication to on-chip applications. The primary advantage of differential interconnects is the high signal-to-noise ratio due to the common mode noise rejection. Differential signaling, however, requires double the number of signal paths as compared to single-ended transmission networks. A typical topology is composed of a transmitter, differential transmission lines, and a receiver, as illustrated in Figure 2.9. The transmitter at the near end of the network converts the single-ended signal into an opposite polarity differential signal, while the receiver at the far end of the transmission lines converts the differential signal into a single ended form.

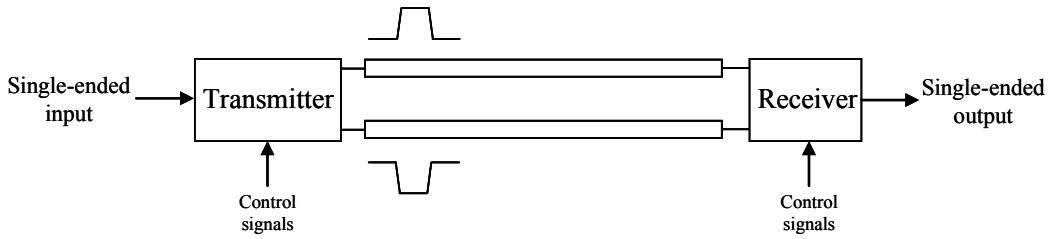


Figure 2.9 Differential transmission network

A differential current-mode signal transmission network can achieve 65% and 51% improvement in delay and energy per unit length, respectively, as compared to inserting repeaters [65]. In this scheme, a transmitter produces a differential current, charging the first transmission line while discharging the second line. Once the current difference is sufficiently large to be sensed, control signals enable and disable the receiver and transmitter, respectively, reducing static power dissipation. To address the problem of crosstalk between

neighboring differential lines in buses, a methodology for optimal positioning of twists in the differential interconnects is provided in [66].

Since differential interconnects exhibit high signal-to-noise ratio, low swing techniques are commonly combined with differential interconnects. A possible implementation of a low swing differential interconnects is shown in Figure 2.10 [46]. In this circuit, the transmitter uses NMOS transistors for both the pull-up and pull-down networks. Since differential interconnects are used, the transmitter provides two complementary signals (*i.e.*, V_{ddL} and 0). At the receiver end, a differential pair amplifies the signal (transistors M_1 and M_2). The cross connected transistors (M_3 and M_4) provide a positive feedback loop for the two N-type common source amplifiers (consisting of transistors M_3, M_1, and M_4, M_2), amplifying the signal until saturation occurs. The signals at the output of the differential pair are further amplified by the buffers, driving an SR flip-flop that stores the final digital state. During each clock cycle, the clocked transistors (M_5, M_6, and M_7) reset the output of the differential pair to zero voltage. When the clock signal is high, transistor M_7 turns off, saving power.

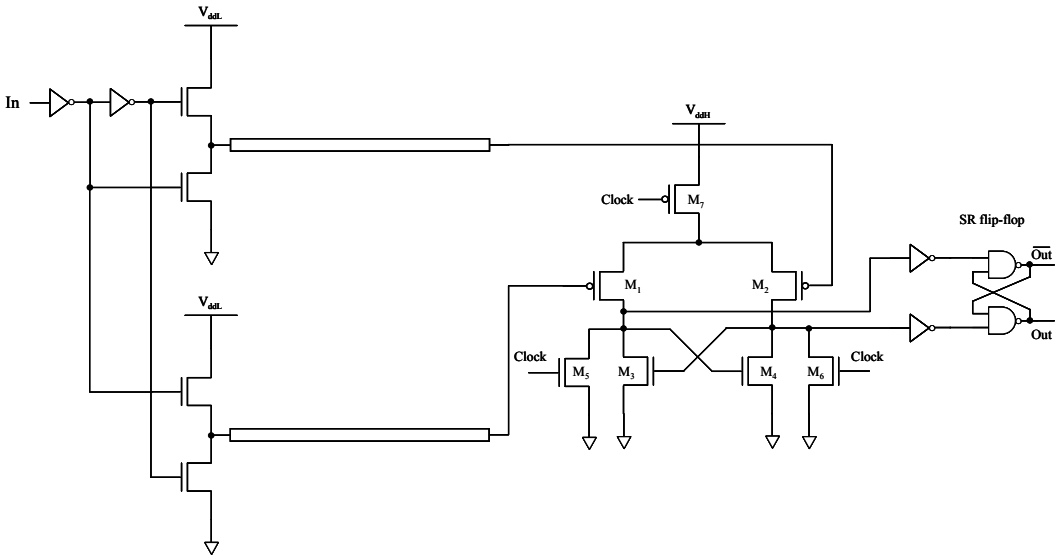


Figure 2.10 Differential low swing interconnect

2.2.5 Bus Swizzling

By intentionally permuting (or swizzling) the interconnects, capacitive coupling among the wires averages out, reducing both the worst case delay and delay uncertainty [47]. Intentional delay is introduced by the swizzled section of the interconnect, resulting in different arrival times of the signal, thereby improving the performance of the critical nets. This behavior occurs since capacitive coupling depends upon the relative arrival time and slew rate of the signals. An illustration of a four bit swizzled bus is shown in Figure 2.11. A maximum reduction of 25% in delay is achieved in a 65 nm technology for a four bit swizzled bus [47]. The bus swizzling technique can also reduce the mutual inductance, as described in [48].

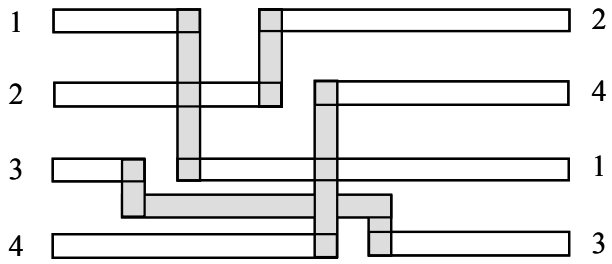
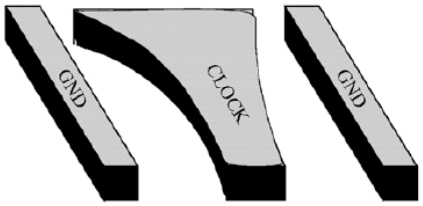


Figure 2.11 Example of bus swizzling

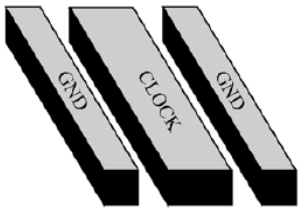
2.2.6 Tapered Interconnects

Non-rectangular geometric interconnect shapes can be used to reduce both power and delay. Specifically, exponential tapering has been shown to be the optimum shape function for minimum delay in *RLC* lines [67]. A tapered interconnect, shown in Figure 2.12(a), can

achieve the same delay with reduced power consumption as the uniform interconnect shown in Figure 2.12(b), due to the smaller capacitance. Furthermore, tapered interconnects reduce the coupling capacitance between neighbor lines, resulting in lower coupling noise. Tapered interconnects, however, exhibit greater resistance due to narrowing of the geometric shape, producing less inductive behavior [67]. The primary disadvantages of tapered lines are the increase in area and the added complexity of the CAD tools and fabrication processes that typically require strictly polygonal shapes. A comparison of trends among the different design methodologies with respect to power, noise, area, and speed is listed in Table 2.1.



(a)



(b)

Figure 2.12 Example of shielded clock lines: (a) tapered interconnect, (b) uniform interconnect [68]

Table 2.1 Comparison of signal distribution networks

Technique	Power	Noise	Area	Speed
Low swing interconnects	⇩	⇧	⇧	⇩
Repeater insertion	⇩	⇩	⇧	⇧
Shielding	⇧	⇩	⇧	
Differential signaling	⇧	⇩	⇧	⇩
Bus swizzling		⇩	⇧	⇧
Tapered interconnects	⇩	⇩	⇧	⇧

2.3 Criteria for Designing Interconnects and Clock Distribution Networks

Clock skew, jitter, power consumption, and delay are the primary metrics for the design and analysis of interconnect and clock networks. These design criteria are characterized and described in this section.

2.3.1 Clock Skew

In synchronous digital circuits, two types of skew, *i.e.*, local and global, are of significance. Local skew refers to the difference in the clock signal arrival time between two *sequentially-adjacent* registers [7],

$$T_{skew,local} = T_{C_i} - T_{C_j}, \tag{2.1}$$

where T_{Ci} and T_{Cj} are the clock delay from the clock source to the registers R_i and R_j, respectively, as shown in Figure 2.13.

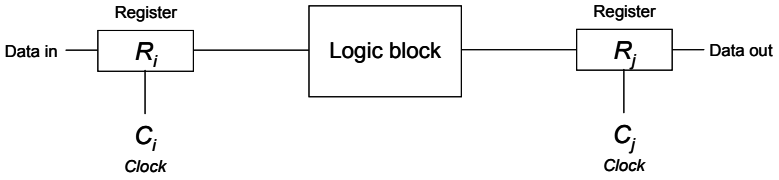


Figure 2.13 Local data path

The clock signals C_i and C_j synchronize the two registers R_i and R_j to ensure that the correct timing constraints are preserved for proper operation of the logic block. In an ideal clock distribution network, the local skew T_{skew} is zero. However, in certain cases, skew can be beneficial to the overall system performance. For instance, intentional skew can be introduced to decrease the delay in certain data paths. It is important to note that local skew is only relevant in local data paths, such as the path shown in Figure 2.13. Skew between nonsequentially connected registers, from an analysis viewpoint, has no effect on system performance [7].

Local skew can be positive or negative as illustrated in Figure 2.14. Positive skew occurs when the arrival of the clock signal at the final register T_{cj} leads the arrival of the clock signal at the initial register T_{ci} of the same sequential data path. Positive skew is, therefore, the additional amount of time which must be added to the minimum clock period to reliably apply the following clock signal at the final register T_{cj}. Thus, positive skew decreases the maximum operating frequency of the circuit, thereby degrading circuit performance. Negative skew occurs when the clock signal arrives at R_i before arriving at R_j, *i.e.*, the clock signal T_{cj} lags the arrival of T_{ci}. As opposed to positive skew, negative skew can improve the

system performance by decreasing the delay of a critical path. However, if not designed carefully, negative skew can create race conditions [7].

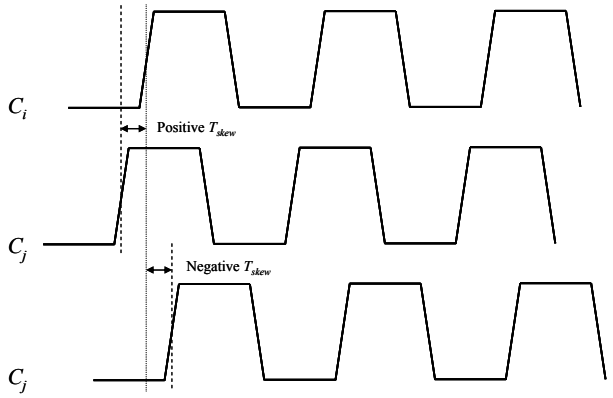


Figure 2.14 Clock timing diagrams

Global clock skew refers to the difference in clock signal arrival time between *any* two registers connected to a global clock distribution network, as illustrated in Figure 2.15. Note that from a design perspective, global clock skew places constraints on the permissible local skew [7]. The maximum difference of the clock arrival times at any two different registers during the same clock cycle is [14]

$$T_{skew,global} = \max\left(\max_{i,j}\left|T_{C_i} - T_{C_j}\right|\right), \tag{2.2}$$

where i and j are any pair of leaf nodes and T_{ci} and T_{cj} are the corresponding clock arrival times, respectively.

Local or global clock skew in synchronous digital systems occurs due to a variety of reasons [15]:

1) variations in line lengths from the clock source to the clocked register;
2) variations in passive interconnect parameters, such as the line resistivity, dielectric constant and thickness, via/contact resistance, line and fringing capacitance, and line dimensions; and
3) variations in active device parameters, such as the threshold voltage, channel mobility, and MOS geometric size, which affect the delay and power dissipation characteristics of the active buffers.

It should also be noted that for a well designed and balanced clock distribution network, the distributed clock buffers are the primary source of clock skew.

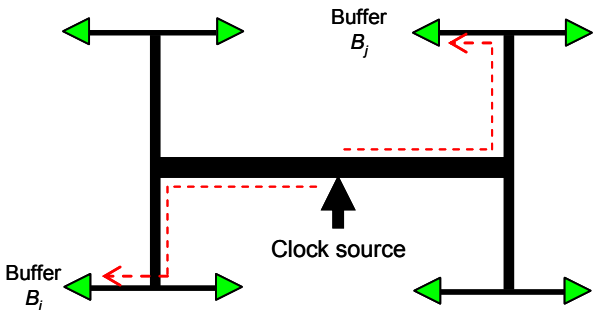


Figure 2.15 H-tree with eight leafs

2.3.2 Jitter

The term jitter (also called phase noise) refers to a deviation in the *time period* (or frequency) of a clock signal from an ideal time period, as illustrated in Figure 2.16. Difficulties in modeling, designing, measuring, and eliminating clock jitter can lead to circuit failure.

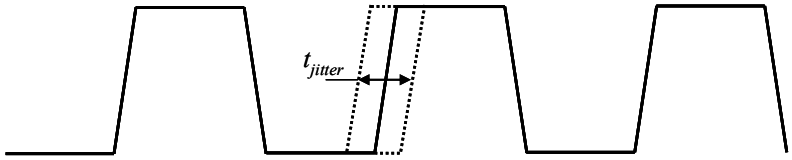


Figure 2.16 Illustration of clock jitter

Jitter can be classified into two primary categories: random (unbounded) and deterministic (bounded) jitter. Deterministic jitter is divided into three sub-categories as shown in Figure 2.17: periodic, data dependent (also known as inter-symbol interference), and cycle-to-cycle jitter.

A random process characterizes random jitter since this form of noise is unpredictable. Random jitter exhibits a Gaussian distribution since the primary source of random noise in many electrical circuits is thermal noise, represented by a Gaussian distribution [16]. Additionally, the combined effect of many uncorrelated noise sources with different individual distributions approaches a Gaussian distribution according to the central limit theorem. A peak value of a Gaussian variable can theoretically reach infinity. While most samples of this random variable is clustered around the mean value, any particular sample can differ from the mean by an arbitrarily large amount, resulting in an unbounded peak-to-peak magnitude of the overall distribution. Taking a large number of samples of such a distribution produces a larger peak-to-peak value.

Deterministic jitter is repeatable and predictable. The peak-to-peak value of this jitter is therefore bounded, and the bounds can be predicted based on a low number of observations. Jitter that repeats in a cyclic fashion is called periodic jitter. Since any periodic waveform can be decomposed into a Fourier series composed of sinusoidal functions, periodic jitter is also referred to as sinusoidal jitter. Periodic jitter is assumed to be uncorrelated with any other

periodically repeating pattern in the data stream, and is typically caused by an external deterministic noise source coupling into the system. For example, power supply noise, RF signals, or an unstable clock-recovery PLL are typical sources of periodic jitter.

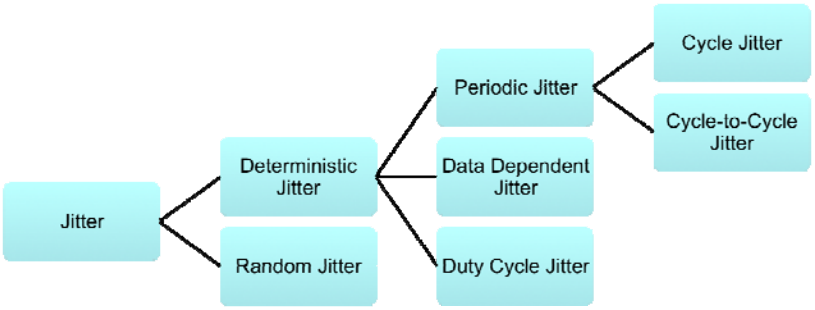


Figure 2.17 Taxonomy of different jitter forms

The maximum cycle jitter is the change in the time period of a clock signal as compared to the ideal time period. This form of jitter is shown in Figure 2.18. In this case, the actual time period of the signal can vary from T_{ideal} - t_{j1} to T_{ideal} + t_{j2}. Note that t_{j1} and t_{j2} are random variables and are not necessarily equal. The maximum period jitter is

$$T_{cycle,\max} = \max\left\{\left|t_k - T_{ideal}\right|\right\}. \tag{2.3}$$

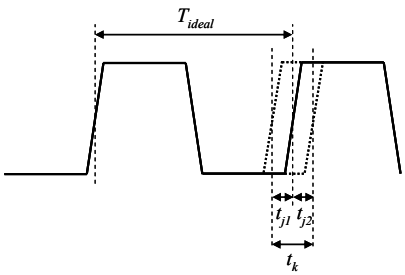


Figure 2.18 Cycle jitter

Contrary to the maximum cycle jitter, the average cycle jitter is the root-mean-square (RMS) over a large number of clock cycles,

$$T_{cycle,RMS} = \lim_{N\to\infty} \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(t_k - T_{ideal}\right)^2}. \tag{2.4}$$

This type of jitter is the most practical form of jitter for processor-based clock distribution networks. The RMS jitter provides meaningful information about the relative uncertainty in the clock arrival time and is therefore a useful metric for designing clock distribution networks.

Cycle-to-cycle jitter is the change in the time period of a clock signal relative to the previous clock cycle [17], as shown in Figure 2.19. In this form of jitter, the difference between t_2 and t_1, and t_3 and t_2 displays a random characteristic which is not equal to zero. This form of jitter is the most difficult to measure due to the nonidealities of the measuring equipment. The RMS cycle-to-cycle jitter is

$$T_{cycle-to-cycle,RMS} = \lim_{N\to\infty} \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(t_{k+1} - t_k\right)^2}, \tag{2.5}$$

where N is the number of clock cycles and t_k is the time period [73].

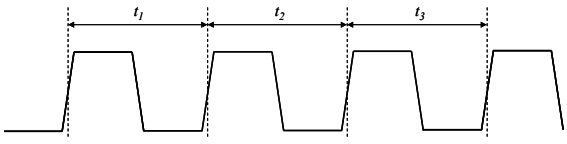


Figure 2.19 Cycle-to-cycle jitter

Note the difference between cycle jitter and cycle-to-cycle jitter. Cycle jitter compares the time period with the mean period, while cycle-to-cycle jitter compares the period with the preceding period. Cycle-to-cycle jitter therefore describes a short term jitter while cycle jitter described a long-term dynamic response.

Jitter correlated with the bit sequence in a data stream is called data dependent jitter. This type of jitter is often caused by the non-ideal frequency response of a medium that the signal propagates through. For example, consider a data sequence that propagates within an interconnect with a low pass frequency response. In this case, the data sequence at the destination exhibits less than a full swing voltage unless several bits in a row are of the same polarity. Data dependent jitter is also called inter-symbol interference (ISI).

Jitter that can be predicted based on the rising or falling edge of a bit stream is called duty-cycle jitter. This type of jitter typically originates from the difference between the slew rate of the rising and falling edges. Additionally, variations in the decision threshold can generate duty cycle jitter.

The dominant sources of clock jitter are [18]:

1) Power supply and temperature variations affecting the circuit delay from path to path.
2) Changes in capacitive load due to coupling between the clock lines and adjacent signal lines, and variations in the gate capacitance.
3) The generation of an on-chip clock signal usually includes a voltage controlled oscillator (VCO). A VCO, however, is sensitive to intrinsic device noise, power supply variations, and substrate noise coupling. Hence, the generation of the clock source contributes jitter to the signal.

In a well designed clock generation and distribution network, power supply variations are the dominant source of clock jitter.

2.3.3 Power Consumption

As described in Chapter 1, high power consumption is a fundamental issue in high complexity, high performance integrated circuits. Interestingly, local and global clock networks make up about 1% of the total interconnections, with about 4% of the total routing length [19]. However, a significant portion of the total power budget (varying between 40% to 70%) [20] is consumed by the clock distribution network.

A typical section of a clock or interconnect network is shown in Figure 2.20. A buffer drives an *RLC* interconnect, generally terminated with a time varying capacitive load. The total power consumption of this structure is

$$P_{total} = P_{buf} + P_{int}, \tag{2.6}$$

where P_{buf} and P_{int} are the power consumption of the driver and interconnect, respectively.

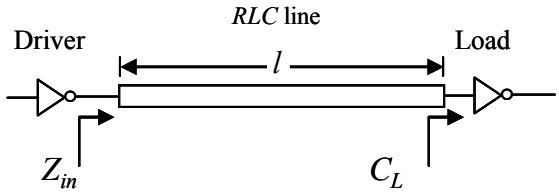


Figure 2.20 Typical section of a clock and interconnect network

The power consumption of the driver P_{buf} consists of the following components [21],

$$P_{buf} = P_{dynamic} + P_{leakage} + P_{short-circuit} + P_{DC}, \tag{2.7}$$

where $P_{dynamic}$ is the dynamic switching power dissipated while charging and discharging a capacitive load. $P_{leakage}$ is the sum of the subthreshold leakage power due to the non-ideal off state characteristics of the MOSFET transistor and the gate leakage caused by carrier tunneling through the thin gate oxide. The third component $P_{short\text{-}circuit}$ is the power dissipated during an input signal transition when both the pull-up and pull-down networks of a CMOS gate are simultaneously on. Finally, P_{dc} is the static DC power consumed when a CMOS circuit is driven by a low voltage signal, resulting in both the pull-up and pull-down networks turned on and therefore conducting current.

The second component of (2.6) is P_{int}, which is the power dissipated by an interconnect due to the resistive losses associated with the metal wires. The power dissipation of a lossy interconnect is [22]

$$P_{int} = \frac{1}{2}\Re\left\{V \cdot I^*\right\} = \frac{1}{2}V_{in,rms}^2 \cdot \Re\left\{\frac{1}{Z_{in}}\right\}, \tag{2.8}$$

where $V_{in,rms}$ is the root-mean square value of a periodic signal and Z_{in} is the input impedance of the interconnect, as shown in Figure 2.20.

2.3.4 Delay

Delay models for a typical interconnect structure such as shown in Figure 2.20 have been proposed in a multitude of research publications [24]-[31]. Based on a modified Bessel function, delay expressions of an *RLC* interconnect have been rigorously developed in [24] and [25]. This effort has resulted in highly complicated and cumbersome expressions. To

simplify these solutions, the transfer function of the interconnect is truncated and approximated with a few dominant poles [26], [27], and [28]. The analysis and characterization of more complicated interconnect structures, *e.g.*, *RLC* tree networks, is provided in [29], [30], and [31]. In long interconnects where the dominant component of the delay is the propagation of a traveling wave, the time of flight can be used as a first order approximation of the signal delay,

$$t_d \approx \frac{l}{\sqrt{2}} \cdot \sqrt{L \cdot C \cdot \left(\sqrt{1 + \left(R/\omega L \right)^2} + 1 \right)}, \tag{2.9}$$

where R, L, and C are the resistance, inductance, and capacitance per unit length, respectively, and ω and l are the radian frequency and length of the interconnect, respectively. Note that in clock distribution networks, the absolute delay from the central clock source to the registers is not significant (but can increase the skew and jitter). As described previously, the relative skew between two clocked registers, however, is important [7], [23].

2.4 Summary

Traditional global clock distribution networks and several signaling techniques are reviewed in this chapter. Different topologies and characteristics of global, semi-global, and local clock distribution networks are described. Data transmission techniques in the form of low swing interconnects, repeater insertion, shielding, differential signaling, bus swizzling, and tapered interconnects are presented and compared with respect to the following

performance metrics: power, noise, area, and speed. Finally, important design and analysis metrics of these networks are reviewed.

Chapter 3

Oscillatory Clock and Signal Modulated Data Distribution Networks

Modern on-chip digital electronic circuits are composed of two types of elements: active semiconductor devices (*e.g.*, transistors) and passive networks (*e.g.*, interconnects). Clock and interconnect distribution networks are among the most important passive components in a high performance microprocessor. Both of these structures are essential in forming and synchronizing complex functional blocks. Since interconnects do not scale as well as active devices, a large performance gap has developed between the two domains. To battle this trend and reduce the scaling inferiority of interconnects, novel clock and data networks have been proposed. These networks are discussed in this chapter. To support the discussion of these networks, a brief introduction to transmission line theory is provided in section 3.1. Advanced high performance global clock and interconnect distribution networks are discussed in sections 3.2 and 3.3, respectively. A summary is provided in section 3.4.

3.1 Introduction to Transmission Line Theory

An infinitesimal segment of a transmission line is depicted in Figure 3.1 where *R*, *L*, *C*, and *G* are the resistance, inductance, capacitance, and conductance per unit length.

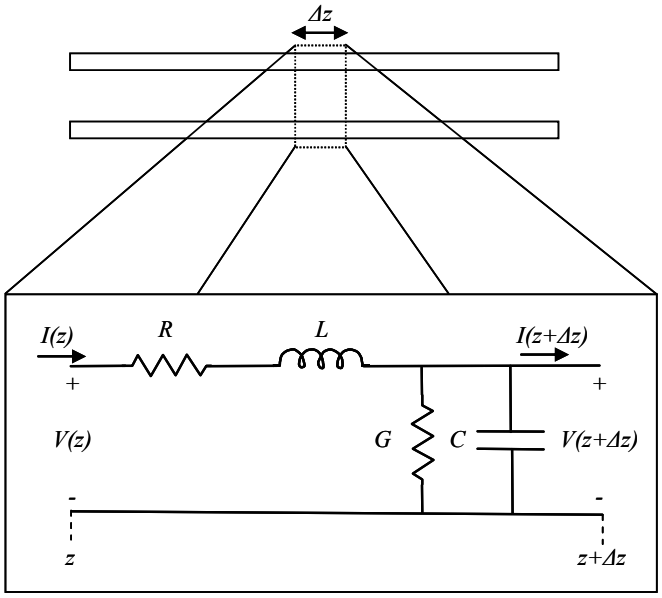


Figure 3.1 Infinitesimal segment of a transmission line

The electrical characterization of such a line is in the form of coupled, first order differential equations described in a phasor notation,

$$-\frac{\partial V}{\partial z} = \left(2R + j\omega \cdot L\right) I, \tag{3.1}$$

$$-\frac{\partial I}{\partial z} = \left(G + j\omega C\right) V. \tag{3.2}$$

The solution of this system of two differential equations is two exponential functions for the voltage and current, respectively,

$$V(z) = V^{+} e^{-\gamma z} + V^{-} e^{+\gamma z}, \tag{3.3}$$

$$I(z) = I^{+} e^{-\gamma z} + I^{-} e^{+\gamma z}, \tag{3.4}$$

where V^+ (I^+) and V^- (I^-) are the amplitude of the voltage (current) traveling wave propagating in the $+z$ and $-z$ direction, respectively. In (3.3) and (3.4), the propagation constant γ is

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)\cdot(G + j\omega C)}, \tag{3.5}$$

where α and β are the attenuation constant and wave number, respectively.

Two additional coefficients that characterize a transmission line are the characteristic impedance Z_0 and the reflection coefficient Γ. The characteristic impedance is the ratio between a voltage and current wave traveling in the same direction,

$$Z_0 = \frac{V^+}{V^-} = -\frac{V^-}{I^-} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}. \tag{3.6}$$

The voltage reflection coefficient is the ratio between the reflected and the incident voltage waves,

$$\Gamma = \frac{V^-}{V^+}. \tag{3.7}$$

From the reflection coefficient, the voltage and current waves can be expressed, respectively, as

$$V(z) = V^+\left(e^{-\gamma z} + \Gamma e^{+\gamma z}\right), \tag{3.8}$$

$$I(z) = \frac{V^+}{Z_0}\left(e^{-\gamma z} - \Gamma e^{+\gamma z}\right). \tag{3.9}$$

For example, the reflection coefficient at the load ($z = 0$), as shown in Figure 3.2, is

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}. \tag{3.10}$$

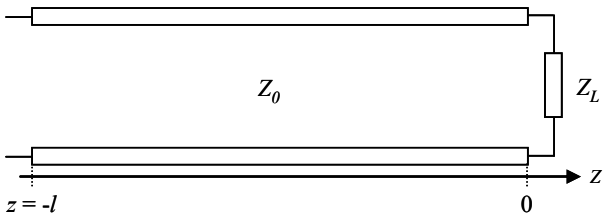


Figure 3.2 Terminated transmission line with load Z_L

3.2 Novel Global Clock Distribution Networks

Most of the work on reducing the power dissipated by the clock network has focused on reducing the voltage swing [43], optimally inserting repeaters [49], [55], [56], clock routing [57], and clock gating [65]. Clock gating has become one of the most effective and widely used techniques for mitigating power consumption in clock distribution networks, and is based on conditionally slowing or completely stopping the local clocks with respect to the global clock. These techniques, however, cannot successfully combat the increasing demand for low skew, low jitter, low power, and high speed.

The challenges related to the clock distribution network require novel approaches. Recently, three approaches for multi-gigahertz clock distribution networks have been proposed:

1) Coupled standing wave oscillators (SWO) [70];
2) Rotary traveling wave oscillators (RTWO) [71]; and
3) Resonant distributed differential oscillators (DDO) [74].

All three techniques combine a distributed oscillator with a distribution network. Hence, the clock signal is generated and distributed by the same network. A description of the three techniques is provided, followed by a comparison of these techniques.

3.2.1 Coupled Standing Wave Oscillators (SWO)

The SWO exploits standing wave phenomenon in transmission lines to distribute a clock signal. The primary advantage of the standing wave is the constant phase of the clock signal along the line. Ideally, the standing wave achieves zero skew. However, while the phase is constant, the amplitude of the clock signal varies with time. This behavior requires the use of digital clock recovery circuits at each load. In this section, the working principles of the SWO and related circuitry are described in subsections 3.2.1.1 and 3.2.1.2, respectively.

3.2.1.1 Theoretical Principles and Design Guidelines

To realize a standing wave, consider Figure 3.3 with a shorted load, *i.e.*, $Z_L = 0$. The traveling voltage signal along the line is

$$V(z) = V^+\left(e^{-\gamma z} + \Gamma e^{+\gamma z}\right). \tag{3.11}$$

Since the reflection coefficient in the shorted load case is Γ = -1, (3.11) results in

$$V(z) = V^+\left(e^{-\gamma z} - e^{+\gamma z}\right) = -2V^+ j\sin(\gamma z). \tag{3.12}$$

Converting the phasor expression (3.12) into the time domain results in

$$v(z,t) = \Re\left\{-2V^+ j\sin(\gamma z)\cdot e^{j\omega t}\right\} = \underbrace{-2V^+\cos(\omega t + \pi/2)}_{Amplitude}\cdot\underbrace{\sin(\gamma z)}_{Phase}. \tag{3.13}$$

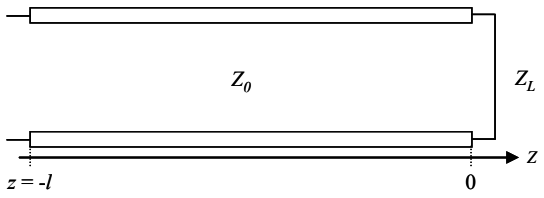


Figure 3.3 Transmission line terminated with a shorted load

In (3.13), the phase component only varies with location, while the amplitude varies with time. This characteristic suggests that the wave is effectively standing rather than traveling along the line, achieving the same phase at the same location. An illustration of standing waves in a lossless transmission line (with propagation constant β) is shown in Figure 3.4. In this graph, each curve represents a different point in time.

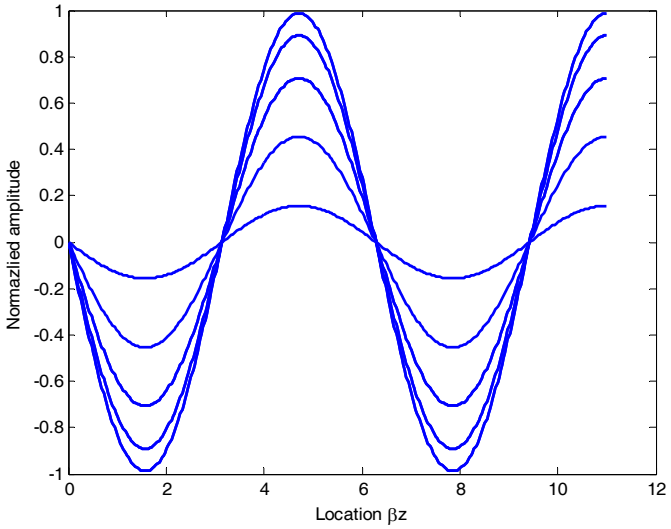


Figure 3.4 Standing waves in a lossless transmission line

On-chip transmission lines are relatively resistive due to the small size, and therefore are strongly attenuative. As a result, waves cannot travel long distances without active amplification. An example of a widely used active amplification method is repeater insertion [49], [55], [56]. To compensate for transmission line losses, distributed negative resistance circuits are inserted along the line, introducing parasitic capacitances c_d, as shown in Figure 3.5. Adding active circuits modifies the electrical characteristic of the transmission line, eliminating attenuation and enhancing the signal propagation characteristics.

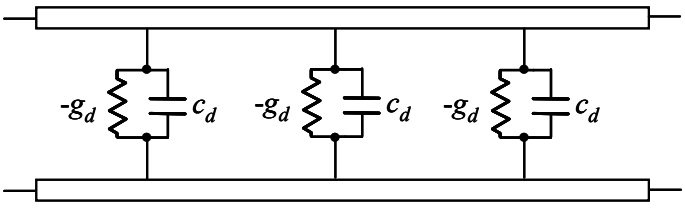


Figure 3.5 Transmission line with distributed negative transconductance

The lumped elements, g_d and c_d, can be approximated, respectively, as a distributed conductance G_d and capacitance C_d in the form,

$$G_d = \frac{ng_d}{l}, \tag{3.14}$$

$$C_d = \frac{nc_d}{l}, \tag{3.15}$$

where n is the number of inserted elements and l is the length of the interconnect. The equivalent infinitesimal segment of a compensated transmission line is shown in Figure 3.6.

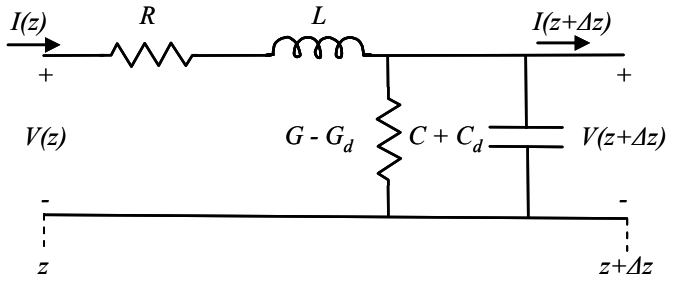


Figure 3.6 Equivalent infinitesimal segment

In this case, the propagation constant of the transmission line is

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)\left((G - G_d) + j\omega\left(C + C_d\right)\right)}. \tag{3.16}$$

The attenuation and propagation coefficients α and β can be approximated, respectively, as

$$\alpha \approx \frac{R}{2}\sqrt{\frac{C+C_d}{L}} + \frac{(G-G_d)}{2}\sqrt{\frac{L}{C+C_d}} \approx \frac{R}{2Z_0} + \frac{(G-G_d)Z_0}{2}, \tag{3.17}$$

$$\beta \approx \omega\sqrt{L(C+C_d)}. \tag{3.18}$$

Note from (3.17) that G_d can be chosen to ensure that no line attenuation exists, *i.e.*, $\alpha = 0$. The parasitic capacitance C_d increases the attenuation and decreases the effect of the transconductance G_d. The parameters, G_d and C_d, are coupled and chosen simultaneously during the design process. Analogous to the high cutoff frequency of a transistor ω_T, a figure of merit for the efficiency of the transconductance is

$$\omega_d = \frac{g_d}{c_d}. \tag{3.19}$$

The parasitic capacitance increases the phase constant, effectively reducing the propagation velocity along the interconnect. For a specific set of transmission line parameters, a value of ω_d exists that minimizes the effective interconnect loss. Note that difficulty in controlling the transconductance can result in an attenuation above zero, producing a distributed amplifier. In this operational mode, the network amplifies the non-dominant harmonics of the clock signal, producing unwanted oscillations and "ringing" [70].

To realize a standing wave oscillator, a half wavelength resonator ($\lambda/2$) is used with lossy transmission lines and distributed transconductance circuits, as shown in Figure 3.7. The Q factor of this resonator is

$$Q = \frac{\alpha}{2\beta}. \tag{3.20}$$

The conditions to realize such a resonator are

$$\alpha(\omega_{osc}) \le 0, \tag{3.21}$$

$$l = \frac{\pi}{\beta(\omega_{osc})}, \tag{3.22}$$

where l is the resonator length and α and β are expressed by (3.16).

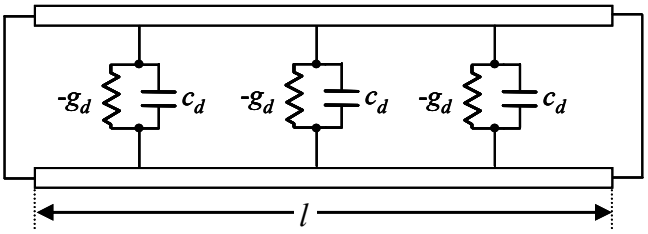


Figure 3.7 A standing wave oscillator

The final phase in the design process is to form a clock distribution network using standing wave oscillators. To construct such a grid, multiple SWOs are connected, as shown in Figure 3.8. This form of connection couples the SWOs and preserves phase locking. The coupling strength is maximum when the SWOs are connected at the center of the resonators since the signal amplitude is greatest at this location. Larger coupling strength results in both a higher mutual locking range and less phase difference between different oscillators.

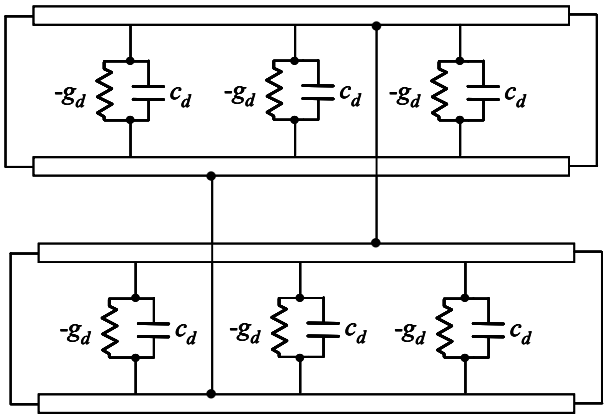


Figure 3.8 Two coupled SWOs

It is also advantageous to injection lock the SWOs to a reference signal [70]. This approach allows the clock frequency to be dictated by an external clock source such as a PLL, stabilizing the noisy signal of these low Q oscillators. Injection locking is similar to coupling; however, the former approach requires injecting current from an external source (rather than another local oscillator) to force oscillations at a specific frequency.

A resonant grid of coupled SWOs is shown in Figure 3.9 [70]. The grid is combined from multiple connected SWOs and injection locked to a single clock reference. To realize a grid formation, a portion of each SWO between the short circuit and the coupling points is folded, resulting in residue stubs where the voltage magnitude is lowest and cannot be recovered by clock buffers. An example of voltage standing waveforms for a section of the grid is shown in Figure 3.10 [70]. Note that the largest magnitude occurs around the square portion of the grid and that zero voltage amplitude of the standing waves appears at the stubs, as shown in Figure 3.10.

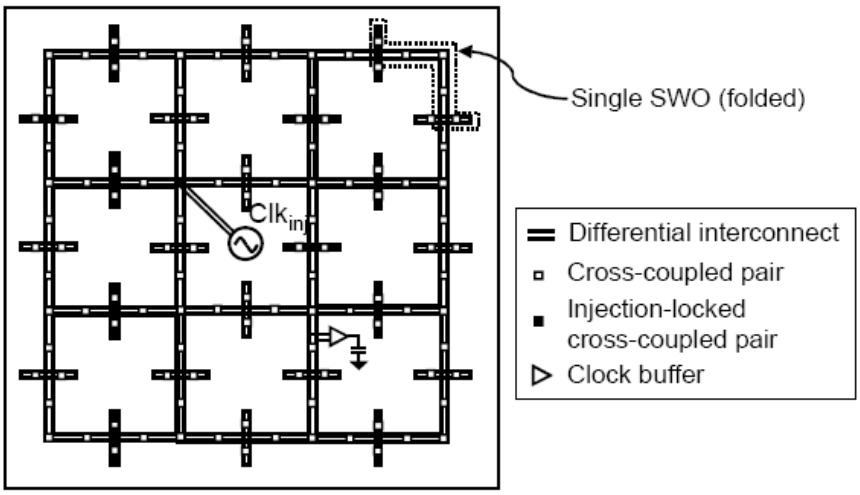


Figure 3.9 Grid of coupled SWOs [70]

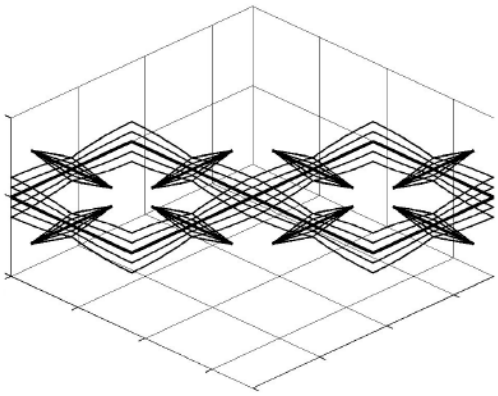


Figure 3.10 Voltage waveforms of a SWOs grid [70]

A modification to the proposed circuit in the form of an SWO with tapered transmission lines is proposed in [76]. The tapered transmission lines enhance the Q of the oscillator and lower the phase noise, *i.e*, jitter. Reduced loss is achieved by physical shaping of the transmission lines, adapted to the location-dependent amplitude of the standing waves. A micrograph of a 10 GHz SWO is shown in Figure 3.11 [70].

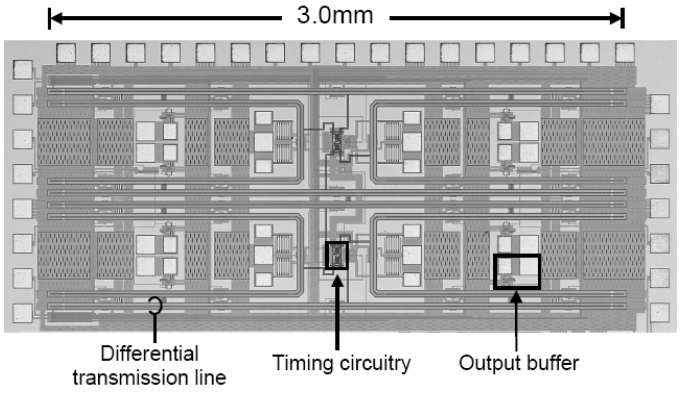


Figure 3.11 A micrograph of a 10 GHz circuit composed of eight coupled SWOs [70]

3.2.1.2 Circuit Implementation

The design of the transconductance circuit shown in Figure 3.5 is based on a pair of cross-coupled transistors, as depicted in Figure 3.12. In this circuit, the source-coupled NMOS pair determines the transconductance g_d. The load is a pair of PMOS devices that are diode-connected with large resistors. The load sets the common-mode voltage at the output terminals to $V_{dd} - V_{thp}$ while isolating the gate capacitance at high frequencies.

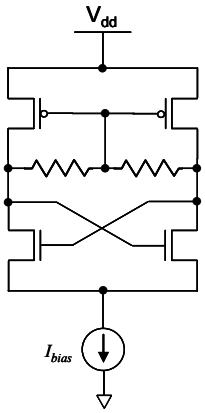


Figure 3.12 Negative resistance circuit

Clock buffers convert the sinusoidal clock to digital levels, as shown in Figure 3.13 [75]. The difference between the sinusoidal standing waves is amplified by the differential amplifier on the left portion of the circuit. The amplified inverted signals are fed into the right portion of the circuit which are further amplified by an additional differential pair. According to the difference between the two amplified signals, the positive feedback inverter pair saturates the outputs to the correct digital value. Note that this clock buffer has been demonstrated to operate at a 2 GHz sinusoidal clock signal rather then at the 10 GHz distributed standing wave reported in [70]. This circuit is predicted to be functional at higher operating frequencies in more advanced technologies [70].

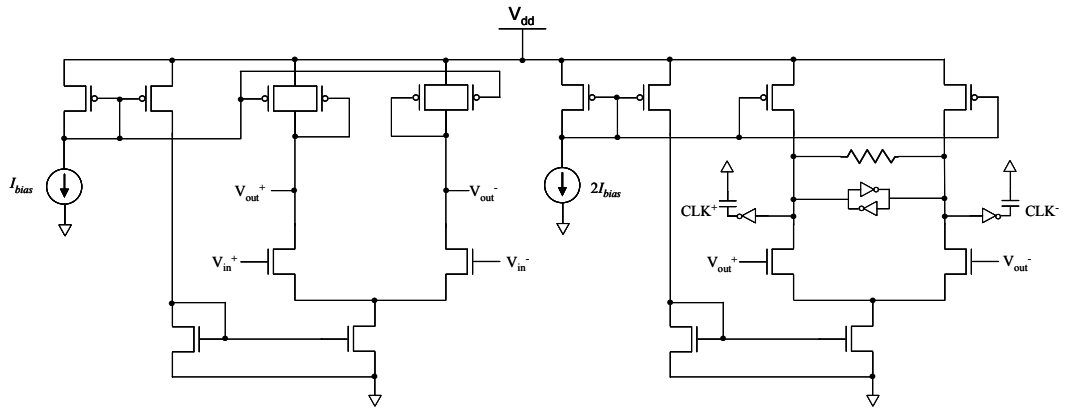
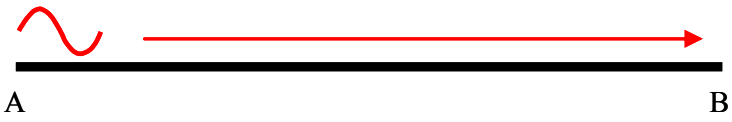


Figure 3.13 Clock buffer

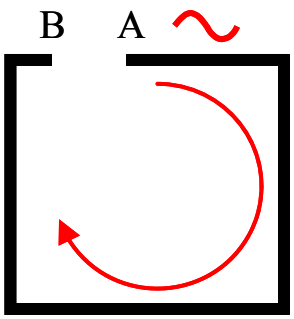
3.2.2 Rotary Traveling Wave Oscillators (RTWO)

Contrary to a standing wave clock distribution network, an RTWO exploits the nature of a traveling wave in a transmission line. Operation of the RTWO loop structure is shown in Figure 3.14. Waves travel along the transmission lines, as depicted in Figure 3.14(a). To

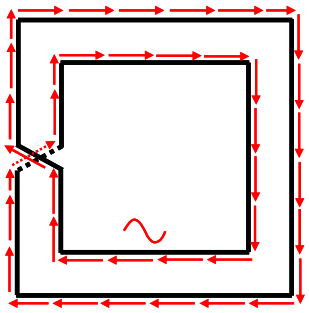
create a more efficient structure to distribute the clock, the same transmission line can be folded into a square formation, as shown in Figure 3.14(b). A closed differential loop is realized by folding the squared line, as shown in Figure 3.14(c). The circumference of the differential closed loop is approximately half of the squared line and occupies one quarter of the area of the squared structure (shown in Figure 3.14(b)).



(a)



(b)



(c)

Figure 3.14 Operation of a traveling wave loop: (a) in a transmission line, (b) in a square shaped transmission line, (c) in a differential closed loop

In a lossless transmission line, the closed loop can indefinitely sustain a traveling wave. However, due to the resistive losses of the line, a traveling wave is attenuated and cannot be sustained indefinitely. To compensate for line attenuation, inverter pairs (in a latch formation) are inserted along the ring, as shown in Figure 3.15. In this form, the electromagnetic energy circulates in a closed loop rather than dissipates as heat during the charge and discharge of the load capacitance.

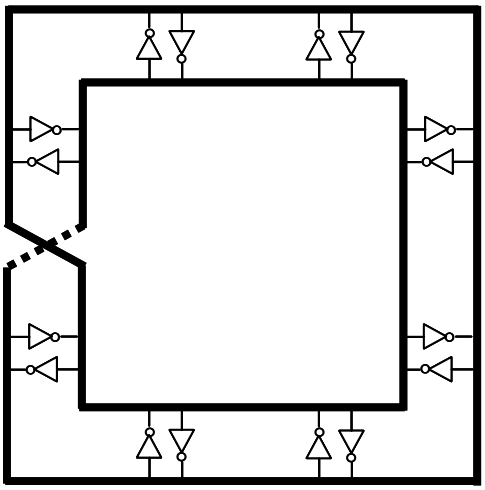


Figure 3.15 Rotary clock ring

The rotary clock ring with inverter pairs generates different phases along the lines due to the propagation of the traveling wave. The outer and inner lines are coupled by inverter pairs, resulting in a 180° phase difference. The relative phase between any two points on the loop is determined by the distance.

To analyze traveling wave propagation in a rotary loop, consider a section of the loop, as shown in Figure 3.16. The time it takes for a traveling wave to propagate from the reference point of φ_0 to an arbitrary point p on the loop is

$$t_p = l_p\sqrt{LC}, \tag{3.23}$$

where l_p is the distance between points φ_0 and p. Expression (3.23) is frequency independent since the inverter pairs effectively produce a lossless line.

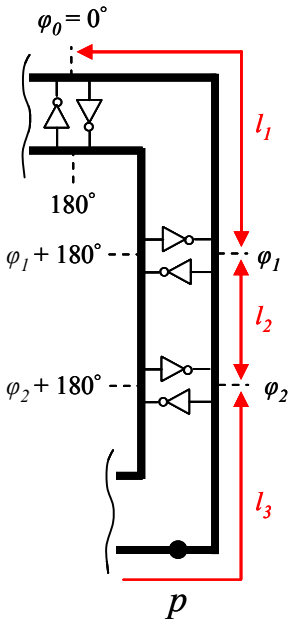


Figure 3.16 Section of a rotary clock ring

Assuming that both the outer and inner loops have approximately the same circumference, the phase of the traveling wave at point p relative to φ_0 is

$$\varphi_p = \frac{360^\circ \cdot t_p}{\sqrt{LC} \cdot \sum_{i=1}^{N} l_i}, \tag{3.24}$$

where N is the number of inverter pairs inserted in a loop and l_i is the distance between the i^{th} - 1 and the i^{th} inverter pair. Substituting (3.23) into (3.24) results in

$$\varphi_p = 360^\circ \cdot \frac{l_p}{\sum_{i=1}^{N} l_i}. \tag{3.25}$$

Note that the phase of the traveling wave is solely dependent on the length of the rotary loop and that the phase on the inner loop is $\varphi_p + 180°$.

The frequency at which the clock signal oscillates is the phase velocity divided by two times the rotary ring circumference,

$$f_{osc} = \frac{1}{2\sum_{i=1}^{N} l_i \cdot \sqrt{LC}}. \tag{3.26}$$

Multiplication by two considers the wave propagation through the outer and inner loops. The oscillating signal is a square-wave signal sustained by the inverter latches. Hence, additional circuitry is not required to convert the signal from a sinusoidal to a square waveform, as in the case of a standing wave oscillator. Another property of the RTWO is the different phases of the signal. Depending upon the location of the load, a clock signal with the same amplitude and different phase is distributed.

To form a rotary clock distribution network, multiple rotary rings are connected, as shown in the example depicted in Figure 3.17. The adjacent rings are joined together by two nodes, one in the inner loop and one in the outer loop. The clock phase of the different rings

is locked at the connection points to ensure that the rings are synchronized at the target frequency given by (3.26).

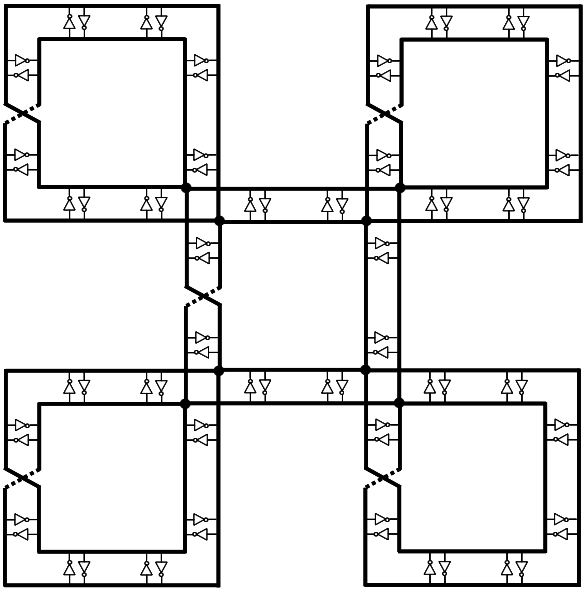


Figure 3.17 Rotary clock distribution network

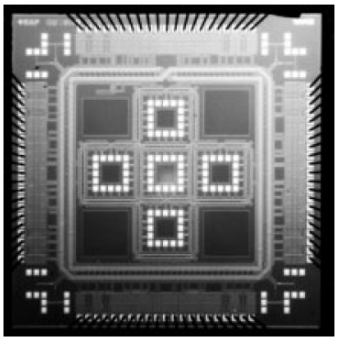


Figure 3.18 A micrograph of a 965 MHz four coupled RTWOs [71]

Optimization and design guidelines for the RTWO have been proposed in [77] and [78]. In [77], a software tool has been developed, extracting the SPICE netlist from a layout of the

rotary network. In [78], an integrated placement and skew scheduling methodology is proposed based on flip-flop assignment and cost driven skew optimization algorithms. A micrograph of an example 965 MHz RTWO is shown in Figure 3.18 [71].

3.2.3 Resonant Distributed Differential Oscillators (DDO)

As opposed to the standing wave and rotary traveling wave oscillators, the resonant distributed differential oscillator (DDO) exhibits uniform amplitude and phase across the entire distribution network. The DDO utilizes on-chip inductors to resonate at a target frequency with the transmission line capacitance in a distributed oscillator array. The oscillator is injection locked to an external reference clock to stabilize the noisy clock signal of the relatively low Q oscillator. As with the previously discussed oscillators, the DDO also operates in a differential manner, providing two opposite phase clock signals. A similar concept had been proposed earlier by Chan *at al.* in [72] and [73] in the form of resonant global clock distribution networks. A methodology for designing these networks is presented in Chapter 5.

3.2.3.1 Theoretical Principles and Design Guidelines

Four resonant oscillators are coupled to form the DDO architecture, as shown in Figure 3.19. Each quarter is composed of a spiral inductor and a negative differential transconductance to compensate for the lossy transmission lines and sustain oscillation. H-trees attached to the global clock grid distribute the clock signal across the entire physical area of the circuit. Variable capacitors are attached to the grid to tune the oscillation

frequency while analog clock buffers amplify and convert the signal from a differential to a single-ended form.

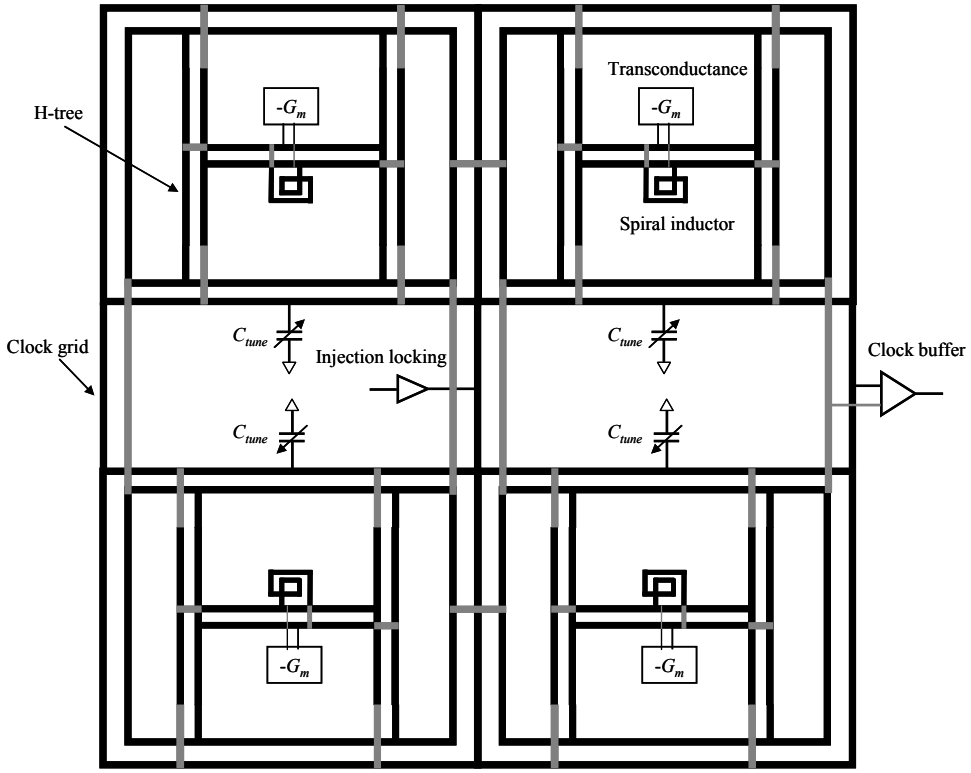


Figure 3.19 Resonant differential oscillator

As first order approximation of this complex network can be modeled by an *RLC* circuit, as shown in Figure 3.20. In this circuit, C_p, L_s, and R_s are the total network capacitance, inductance, and resistance, respectively. Note that R_s models all of the losses in the network including the line and inductor resistance, eddy current in the substrate and neighboring conductors, and the displacement current. Since the DDO is a symmetric differential network, only half of the circuit is modeled.

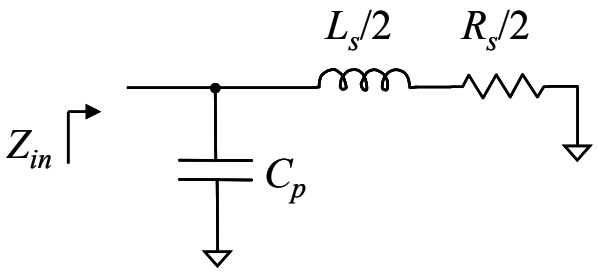


Figure 3.20 Circuit model of the DDO

The input impedance of the circuit shown in Figure 3.20 is

$$Z_{in} = \frac{R_s\left(\omega^2 C_p L_s - 2\right) + \omega^2 L_s C_p R_s}{R_s^2 + \omega^2 L_s^2} + j\left(\frac{\omega L_s\left(\omega^2 C_p L_s - 2\right) + R_s^2 \omega C_p}{R_s^2 + \omega^2 L_s^2}\right), \tag{3.27}$$

where ω is the radian frequency. From (3.27), the resonance frequency is

$$f_{osc} = \frac{1}{2\pi}\sqrt{\frac{2}{C_p L_s} - \frac{R_s^2}{L_s^2}}. \tag{3.28}$$

Consequently, the Q of the resonator is

$$Q = \frac{\omega_{osc} L_s}{R_s} = \sqrt{\frac{2L_s}{C_p R_s^2} - 1}. \tag{3.29}$$

3.2.3.2 Circuit Implementation

The transconductance circuit shown in Figure 3.21 is based on a positive feedback configuration where two inverters are connected in a latch formation. Thus, a negative transconductance $-G_m$ is produced between the output terminals. Assuming the DDO is excited by a square wave, the clock amplitude is approximated by

$$V_{clock} \approx \frac{4}{\pi} I_{tail} \left| Z_{in}\left(\omega_{osc}\right) \right|, \tag{3.30}$$

where I_{tail} is the bias current of the transconductance circuit.

The differential clock signal is amplified and converted into a single-ended form by the clock buffer shown in Figure 3.22. In this circuit, the two differential pairs on the left amplify the difference between the inverted clock signals. The amplified signals are fed to the common source amplifier on the right. The output of this circuit is sinusoidal and converted into a square wave. A micrograph of a 1.1 GHz DDO is shown in Figure 3.23 [74].

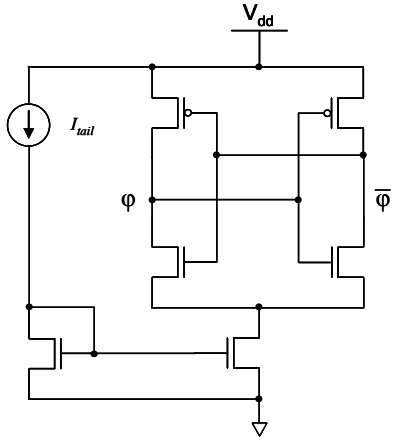


Figure 3.21 Transconductance circuit

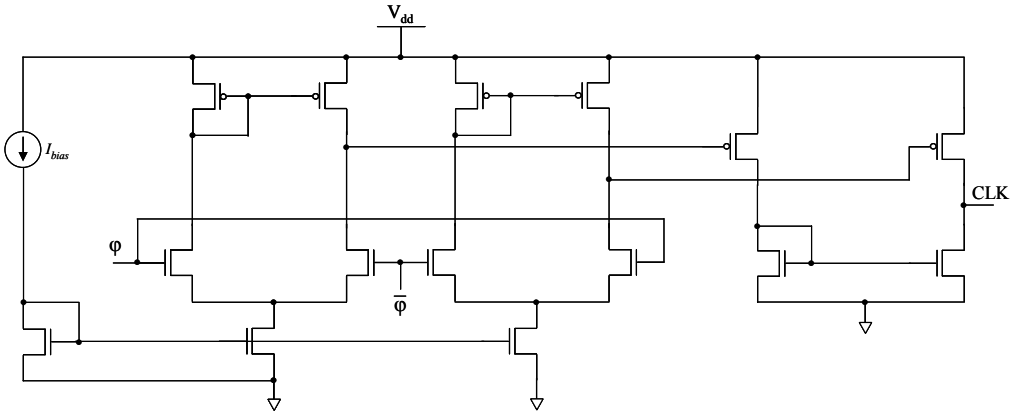


Figure 3.22 Clock buffer

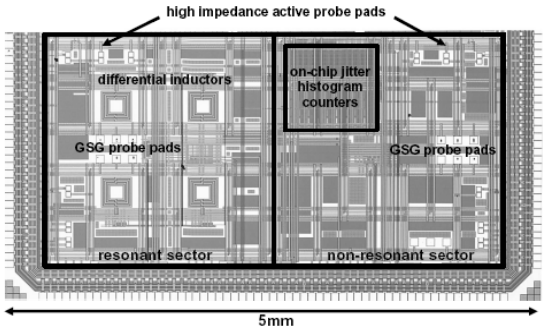


Figure 3.23 Micrograph of a 1.1 GHz DDO with four differential on-chip inductors [74]

3.2.4 Comparison of Clock Distribution Networks

The working principles of several oscillatory clock distribution networks have been described in sections 3.2.1, 3.2.2, and 3.2.3. All three techniques share similar working principles, exploiting transmission lines properties and signal oscillation. Specifically, active positive feedback circuits are utilized to initiate oscillation and compensate for resistive loses. Additionally, differential transmission lines are used to distribute the signal, eliminating the

need for shielding. In this section, the power dissipation, skew, jitter, and area of these clock distribution schemes are compared with other conventional H-tree based clock networks.

The standing wave oscillator [70], described in section 3.2.1, exhibits uniform phase and varying amplitude across the clock network. For an SWO to distribute a full swing clock signal, the use of complex clock buffers is required, as shown in Figure 3.13. The clock buffers consume power, introduce jitter and skew, and cannot operate at multi-gigahertz frequencies in current technologies. Additionally, since the SWO is based on a half wavelength resonator ($\lambda/2$), the length of the network is restricted. For example, for a 10 GHz clock frequency, a 15 mm long differential interconnect is required. To realize the SWO, the wires are folded in the form illustrated in Figure 3.9.

The rotary traveling wave oscillator [71] presented in section 3.2.2 exhibits uniform amplitude and varying phase across the clock network. Contrary to the other two clock distribution networks, the RTWO preserves a square waveform by the inverter pairs and does not require sinusoidal-to-square signal conversion. Since the rotary loops support traveling waves, different phases are exhibited along the lines. If the same phase is required, registers must be connected to the same points on the coupled rings of the RTWO, resulting in a large capacitive load concentrated at a single point. If a different phase of the clock signal is required, this topology is advantageous. As with all active circuits, the distributed inverter pairs consume power and introduce jitter and skew.

The resonant distributed differential oscillator [74], presented in section 3.2.3, exhibits uniform amplitude and phase across the clock network. As in the case with the SWO, relatively complex clock buffers are used (see Figure 3.22) to amplify and convert the differential signal into a single-ended form. These buffers also consume power and introduce jitter and skew into the system. On-chip spiral inductors are used to resonate the network with

the capacitance of the interconnects at a specific oscillation frequency. The on-chip inductors occupy area and introduce losses to the network.

Three additional examples of clock distribution networks are chosen for the sake of this comparison. Two conventional clock distribution networks based on an H-tree topology for the IBM Power4 [33] and the Intel Itanium2 [79] microprocessors are considered. A third exotic approach where a wireless clock distribution network utilizing an external antenna is also considered [80]. All five approaches are compared and listed in Table 3.1.

Table 3.1 Comparison of clock distribution networks

	IBM Power4 [33]	Intel Itanium2 [79]	Coupled SWO [70]	RTWO [71]	Resonant DDO [72]	Wireless Clock [80]
Technology [μm]	CMOS SOI	CMOS 0.13	CMOS 0.18	CMOS 0.25	CMOS 0.18	CMOS 0.13
Clock Frequency [GHz]	1.3	1.2	10	0.965	1.1	2.17
Clock Jitter [psec]	30 (p-p)	N/A	3 (p-p)	5.5 (rms)	3.1 (rms)	5 - 12 (p-p)
Clock Skew [psec]	25	52	0.6	N/A	13.4	< 18
Power Consumption [mW]	N/A	< 800	378	117	50.4	144
Area [mm^2]	20 x 20	14.1 x 18.6	4 x 4	3 x 3	2 x 2	31 x 31
Structure Complexity	Optimum H-tree	Differential H-tree with deskew buffers	Complex clock buffers	Wide differential loops	Spiral inductors and clock buffers	On-chip and off-chip antennas

It is evident from Table 3.1 that the clock distribution networks in the IBM Power4 and Intel Itanium2 microprocessors distribute the clock signal over a large area. Both networks produce large clock jitter and skew as well as dissipate high power. The coupled SWO, RTWO, and resonant DDO distribute the clock signal over a relatively smaller area due to test circuit restrictions. The resonant DDO consumes the least power while the highest

frequency and power consumption are exhibited by the coupled SWO. Similar clock jitter is exhibited by the three wave oscillators. Finally, the wireless clock network distributes the signal over a large area, dissipating a significant amount of power with high skew.

3.3 Novel Transmission Line Networks

To reduce the performance bottleneck of on-chip interconnects in high complexity integrated circuits, advanced signaling techniques have recently been proposed. Repeater insertion, the most common and widely used technique to enhance interconnect performance, can no longer satisfy the increasing demand for higher bandwidth, and lower jitter, skew, and power consumption. Advanced global data transmission networks formed in a transmitter, transmission line, and receiver topology are a promising alternative to repeater insertion. This section focuses on two families of these high speed data transmission networks. Inductance dominated interconnects are discussed in section 3.3.1 while pulse signal interconnects are presented in section 3.3.2. A comparison between the two interconnect families is provided in section 3.3.3.

3.3.1 Inductance Dominated Interconnects

By exploiting the interconnect inductance at high frequencies, data transmission at velocities near the speed of light is proposed in [81]. Principles governing inductance dominated interconnects are presented in subsection 3.3.1.1 while a circuit implementation of these networks is described in subsection 3.3.1.2.

3.3.1.1 Theoretical Principles

Consider a signal propagating as a function of time and distance along a transmission line,

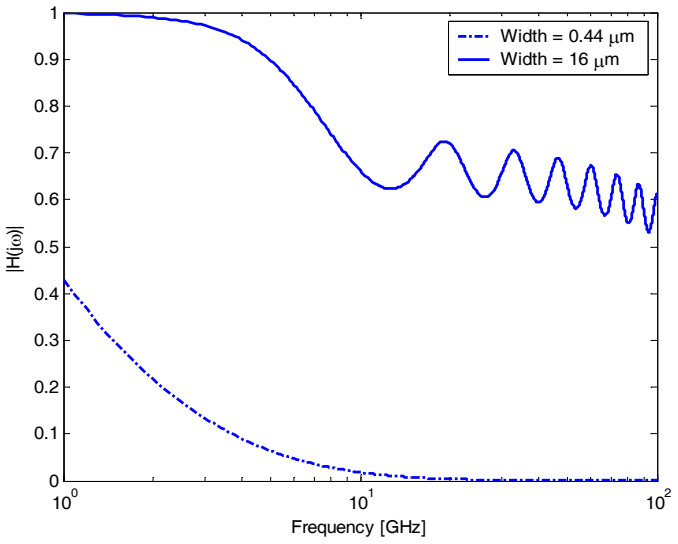
$$\frac{\partial^2 V}{\partial z^2} = RC\frac{\partial V}{\partial t} + LC\frac{\partial^2 V}{\partial t^2}, \tag{3.31}$$

where z, t, and V are the distance along the wire, time, and voltage, respectively, and R, L, and C are the resistance, inductance, and capacitance per unit length, respectively. At lower frequencies, the resistance is much greater than the inductive impedance ($R >> \omega L$), resulting in the first term of (3.31) dominating the right side of (3.31). In this RC regime, signals travel slowly and wave properties do not hold. At higher frequencies, the resistance is much smaller then the inductive component ($R << \omega L$), resulting in the second term of (3.31) dominating the right side of (3.31). In this LC regime, the wire behaves more as a waveguide, allowing for the propagation of electromagnetic waves. Ideally, in a dielectric medium, such waves can travel with a maximum phase velocity relative to the speed of light. For example, in silicon dioxide, the maximum propagation velocity is about half of the speed of light.

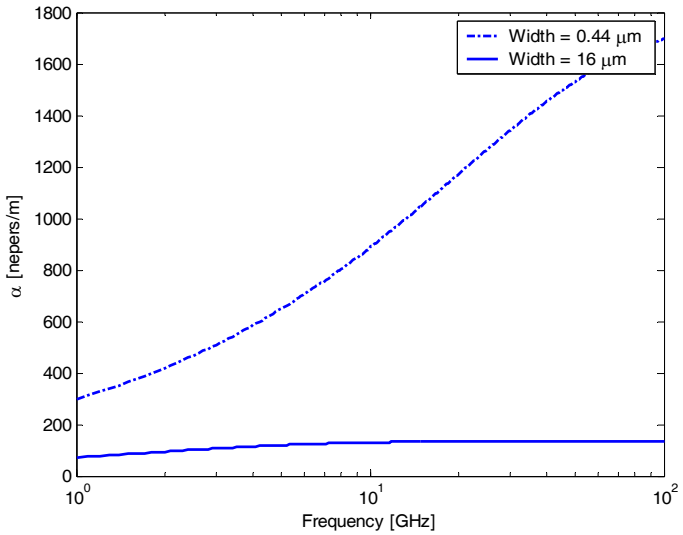
To exploit the LC characteristic of the line, the transmitted digital signal is modulated with a sufficiently high carrier frequency, such that the majority of the signal spectral components are located at higher frequencies. By eliminating the low frequency portion of the signal, the inter-symbol interference (ISI) is also reduced. Since the attenuation factor increases with frequency, the transmission lines are substantially lossier at higher frequencies. In order to minimize the resistive loss of the wires, wide interconnects are used rather than active gain circuits.

As an example, consider minimum and wide width wires (in a 0.18 μm CMOS technology), as shown in Figure 3.24. The minimum and wide interconnect widths are 0.44 μm and 16 μm, respectively. Both lines are 5 mm long, driven by a 10 Ω driver and terminated by a 50 fF capacitive load. The magnitude of the transfer function of the wider line exhibits resonances well beyond 10 GHz while the minimum width line has no gain at these frequencies, as shown in Figure 3.24(a). High frequency signals can therefore be transmitted across wide wires without active amplification.

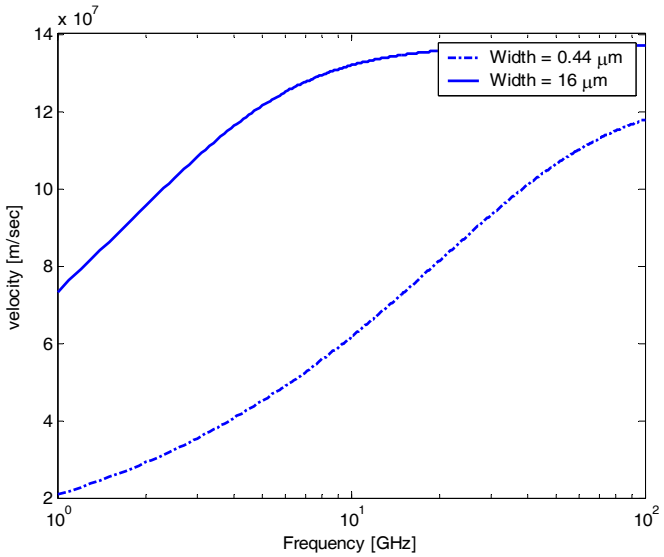
This conclusion is also supported by the attenuation coefficient α, as shown in Figure 3.24(b). For a wider line, the attenuation of the signal is significantly lower at higher frequencies as compared to the minimum width line. Note that the attenuation factor of a wide line exhibits a weak dependence on frequency. Finally, for both lines, the propagation velocity of the signal increases with frequency, reaching the speed of light, $1/\sqrt{LC}$, as shown in Figure 3.24(c).



(a)



(b)



(c)

Figure 3.24 Example of *LC* dominated interconnects: (a) transfer function, (b) attenuation coefficient, (c) phase velocity

3.3.1.2 Circuit Implementation

A block diagram of the *LC* dominated interconnect network is shown in Figure 3.25. The digital (*D*) and complementary signal are modulated by a high frequency carrier using the binary frequency shift keying (BFSK) scheme. The mixer circuit drives the differential interconnects, exploiting the *LC* characteristic of the line at high frequencies. At the receiving end, the down converting mixer demodulates the signal to the original frequency. At the output of the mixer, a sense amplifier restores the signal to a digital level. An illustration of the modulation and demodulation waveforms is shown in Figure 3.26.

At the transmitter end, signal modulation is performed by a passive ring mixer [82], as shown in Figure 3.27. When the input digital signal (D) is high, transistors M_1 and M_4 are both on and transistors M_2 and M_3 are both off. In this mode, the intermediate frequency (IF) port is connected to the local oscillator (LO) port, permitting mixing operation to occur. When the input digital signal (D) is low, transistors M_2 and M_3 are both on and transistors M_1 and M_4 are both off. In this state, the IF circuit is connected with a negative polarity to the LO. This operation achieves BPSK signal modulation.

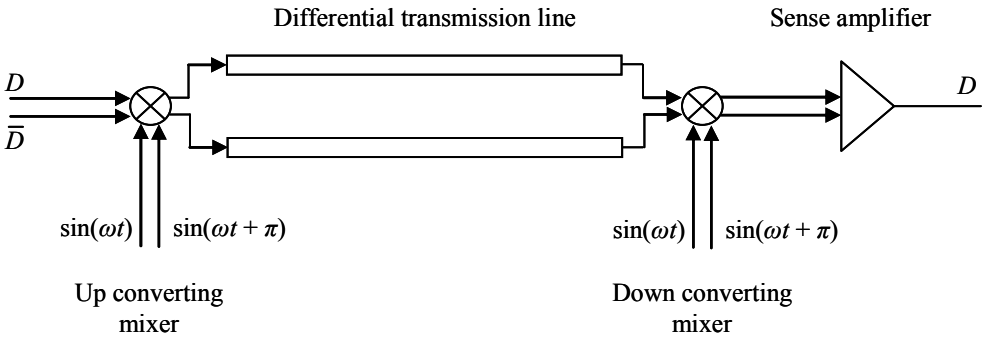


Figure 3.25 Block diagram of the near speed of light interconnect network

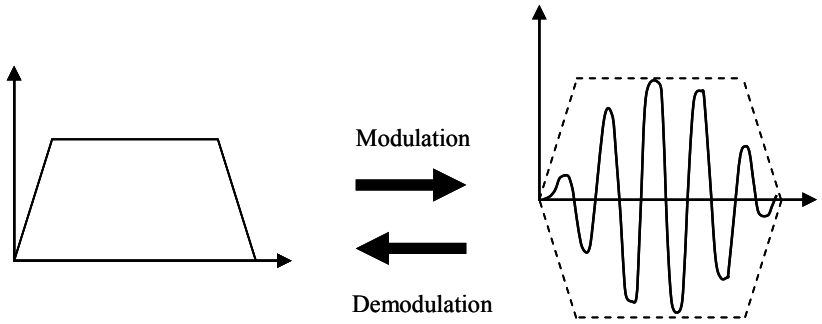


Figure 3.26 Modulation of the transmitted signal

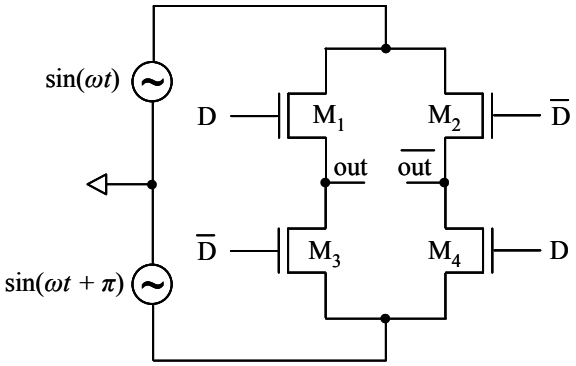


Figure 3.27 Passive ring mixer

At the receiver end, an active double-balanced mixer demodulates and amplifies the attenuated differential signal. The double-balanced mixer circuit is shown in Figure 3.28. The sense amplifier at the output of the demodulation mixer uses a self biased topology [83], as shown in Figure 3.29. In this circuit, a combination of P-type and N-type differential amplifiers are used to produce a self biasing topology. Note that additional buffers are required at the output of the amplifier to achieve a full rail-to-rail signal.

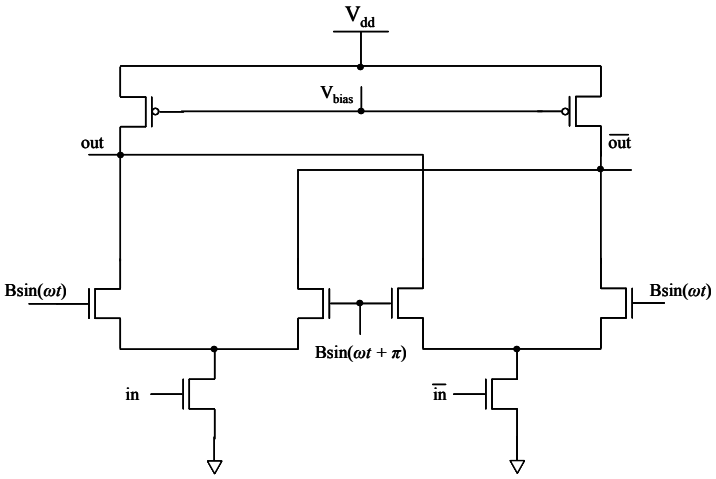


Figure 3.28 Active double-balanced mixer

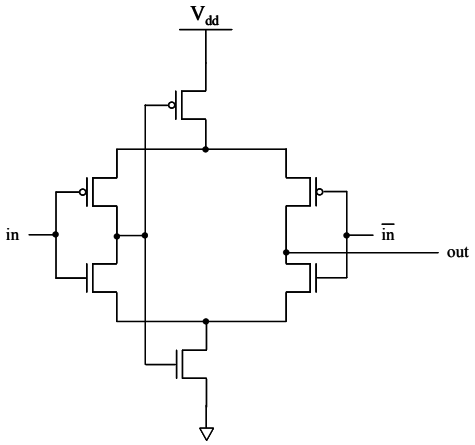


Figure 3.29 Self biased differential amplifier

A test circuit of this transmission network has been fabricated in a 0.18 μm CMOS TSMC technology, as illustrated in Figure 3.30 [81]. The differential LO carrier signals and the digital input pulse are generated off-chip, while the transmitter, interconnect, and receiver are integrated on-chip. The 5 mm spaced differential interconnect is 20 mm long and 16 μm wide.

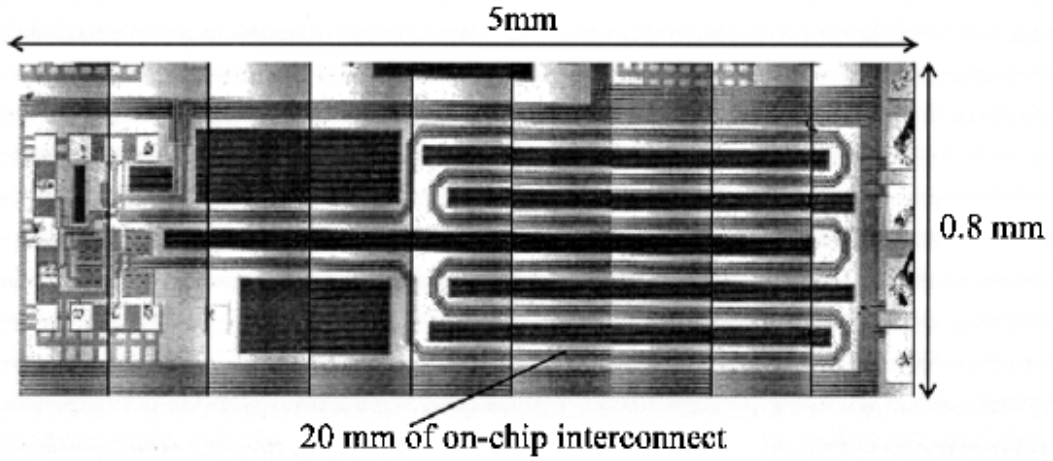
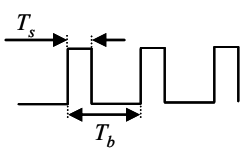


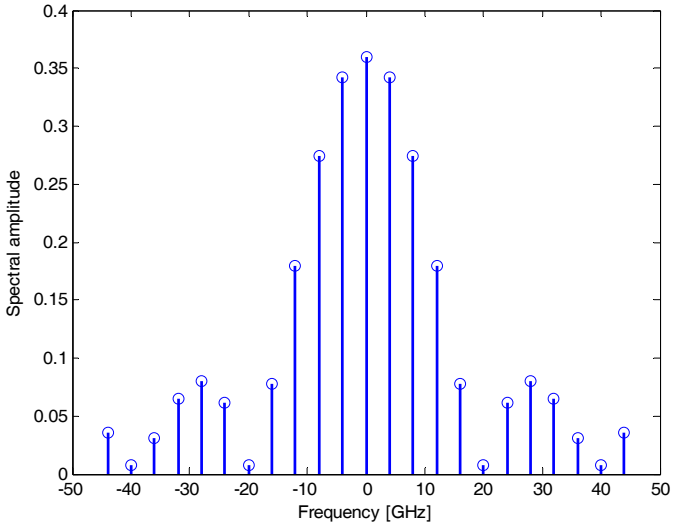
Figure 3.30 Micrograph of a 20 mm interconnect long transmission network [81]

3.3.2 Pulse Signal Interconnects

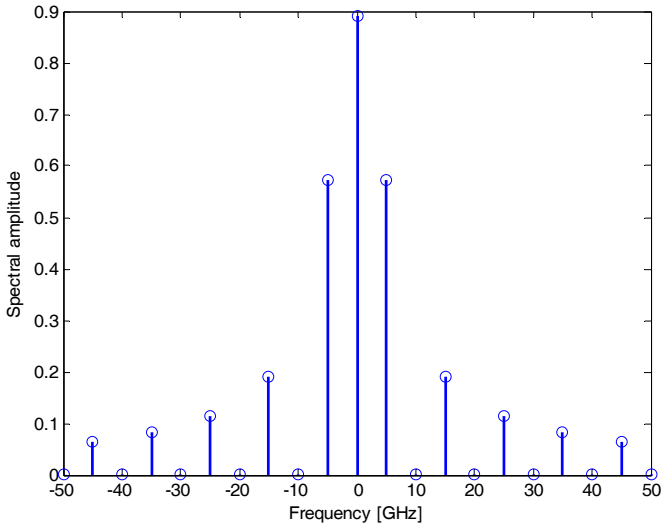
In pulsed wave interconnects, the transmitted signal is converted into short voltage or current pulses, as shown in Figure 3.31(a). By using short pulse waveforms, additional power spectral components of the signal are shifted to higher frequencies. As an example, consider the case of a short pulse signal with T_s = 50 ps and T_b = 250 ps, as shown in Figure 3.31(b), and a wide pulse signal with T_s = 100 ps and T_b = 200 ps, as illustrated in Figure 3.31(c). Note that additional, shorter spectral components appear in the frequency domain for the short pulse as compared to a wide pulse signal.



(a)



(b)



(c)

Figure 3.31 Spectrum of a periodic pulse signal: (a) periodic pulse with width T_s and a period T_b, (b) spectral amplitudes of a pulse with T_s = 50 ps and T_b = 250 ps, and (c) spectral amplitudes of a pulse with T_s = 100 ps and T_b = 200 ps

Wide interconnects suffer less attenuation at high frequencies, as described in section 3.3.1 and shown in Figure 3.24. This characteristic supports the transmission of short pulse signals with reduced power consumption. A useful metric for short pulse signals is the spectral efficiency, which is the number of bits per second transmitted for each Hertz of bandwidth [84],

$$\eta = \frac{T_s}{T_b}. \tag{3.32}$$

Shorter pulse widths result in lower spectral efficiency, as exhibited by the shorter spectral amplitudes illustrated in Figure 3.31(b). For long wires, larger pulse widths exhibit increased intersymbol interference (ISI) due to the inherent dispersion and the increased spacing between two consecutive pulses. Therefore, for optimum pulse width, a tradeoff between spectral efficiency and ISI exist [84]. To increase the bandwidth of a transmission line, time division of the data packets can be implemented [85]. To reduce pulse broadening and attenuation, nonlinear transmission lines (NLTL) can be used [86]. In NLTL networks, the interconnects are periodically loaded with non-linear energy storage elements such as variable capacitors. Examples of a pulsed current and a pulsed voltage signaling network are provided in subsections 3.3.2.1 and 3.3.2.2, respectively.

3.3.2.1 Pulsed Current-Mode Signaling

An example of a network using an 8 Gbps pulsed current signaling scheme is shown in Figure 3.32. Unipolar differential signaling is employed in the form of shielded differential interconnects. The transmitter and receiver circuits are synchronized by a global clock with a

clock offset compensation scheme [84]. Note that the transmission lines are terminated with matched resistors to reduce signal reflection. A test circuit of the transmission link fabricated in a 0.18 μm CMOS TSMC technology is shown in Figure 3.33 [84]. The differential interconnect is 3 mm long and the circuits are clocked by a 1 GHz global signal.

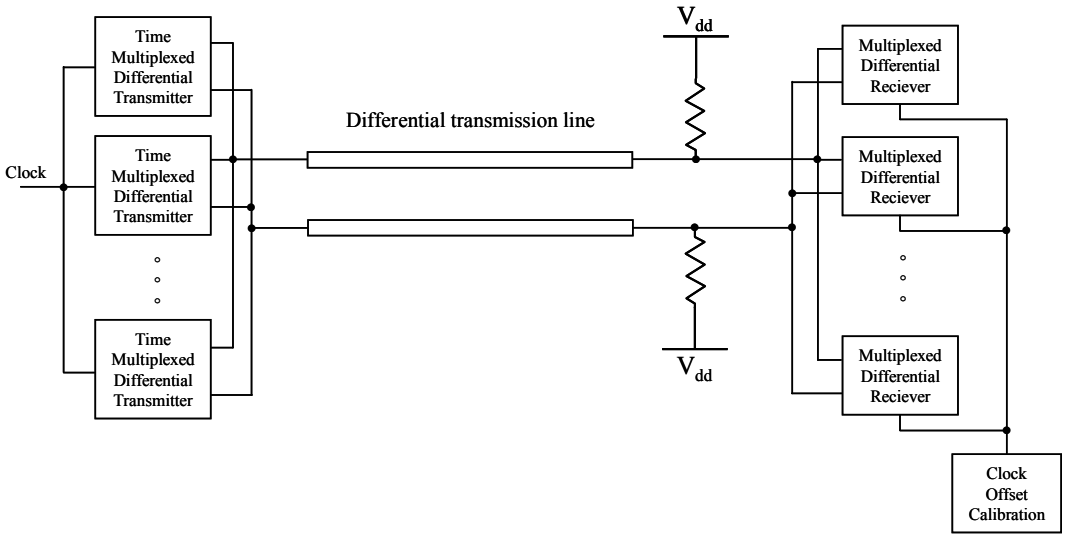


Figure 3.32 Pulsed current interconnect network

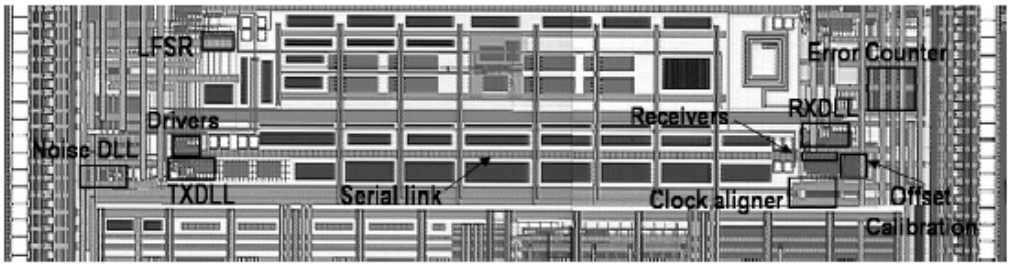
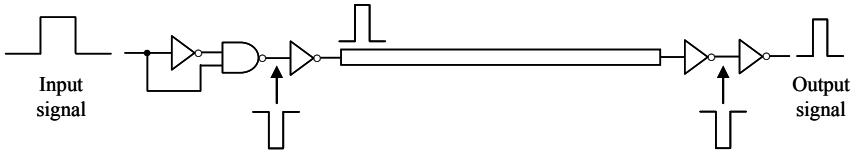


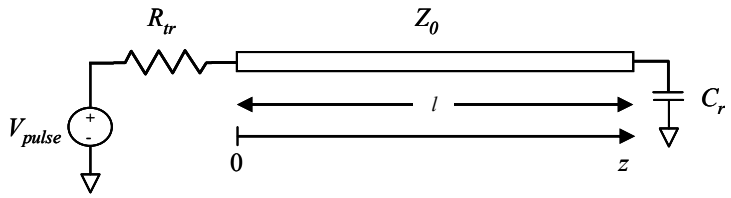
Figure 3.33 Micrograph of a 3 mm pulsed current interconnect [84]

3.3.2.2 Pulsed Voltage-Mode Signaling

A different signaling approach utilizing short width voltage pulses is shown in Figure 3.34 [86]. The short pulse is generated by NOT and NAND gates at the input of the interconnect, as shown in Figure 3.34(a). The width of the pulse is determined by the delay of the first NOT and NAND gate.



(a)



(b)

Figure 3.34 Pulsed wave interconnect: (a) generation and transmission of a pulse signal, and (b) network model

If the reflection coefficient Γ at the load of the interconnect (*i.e.*, $z = l$) equals 1, the voltage at the load is doubled. This behavior occurs since the entire signal is reflected and another signal with the same amplitude arrives at the load. In the case of the model shown in Figure 3.34(b), the reflection coefficient is

$$\Gamma = \frac{\frac{1}{j\omega C_r} - Z_0}{\frac{1}{j\omega C_r} + Z_0}. \tag{3.33}$$

To satisfy $|\Gamma| = 1$,

$$C_r << \frac{1}{2\pi f Z_0} \approx \frac{t_r}{2Z_0}. \tag{3.34}$$

Expression (3.34) assumes that the highest frequency of interest of a trapezoidal pulse is $f = 0.34 / t_r$ [87], where t_r is the rise and fall time of the pulse. To determine the required strength of the driver R_{tr}, consider a voltage wave traveling along a transmission line, as shown in Figure 3.34(b),

$$V(z) = V_{dd} \frac{Z_0}{Z_0 + R_{tr}} e^{-Rz/(2Z_0)}. \tag{3.35}$$

Assuming the voltage doubling effect and equating expression (3.35) to $V_{dd}/2$, R_{tr} is

$$R_{tr} \le Z_0 \left(4e^{-Rl/2Z_0} - 1\right). \tag{3.36}$$

Finally, the energy consumption of a pulse wave interconnect is

$$E = C_p \left(\frac{Z_0}{Z_0 + R_{tr}}\right)^2 V_{dd}^2, \tag{3.37}$$

where C_p is the total capacitance of the line, transmitter, and receiver circuits.

3.3.3 Comparison of Data Distribution Networks

For wide and long interconnects, repeater insertion is no longer an efficient technique for signal transmission. The primary reason for the paradigm shift is that wide and long interconnects are highly capacitive, requiring aggressive repeater insertion. This characteristic results in severe power consumption and signal latency. Concurrently, wide interconnects exhibit lower resistive losses and signal attenuation, permitting the use of novel techniques where transmission line properties at high frequencies and active compensation can be exploited, as described in sections 3.3.1 and 3.3.2. Consequently, the power dissipation and delay are decreased as compared to traditional techniques. A comparison of the different techniques is listed in Table 3.2. Note that in the examples listed in Table 3.2, current pulsed signaling [84] exhibits the highest delay and power consumption per unit length. Among the other three techniques, wave pulsed signal [86] exhibits the best performance per unit length.

Table 3.2 Comparison of data distribution networks

	Repeater Insertion	Signal Modulation [81]	Current Pulsed Signal [84]	Wave Pulsed Signal [86]
Technology [μm]	0.18	0.18	0.18	0.18
Frequency [GHz]	5	1	8	5
Length [mm]	5	20	3	100
Power [mW/mm]	5	0.8	9	0.27
Delay [psec/mm]	74	15	93	8.4

3.4 Summary

With advances in technology and the constant need for high performance, high complexity integrated circuits, advanced design methodologies are necessary. Current techniques and paradigms have become obsolete and cannot support the distribution of clock

and data signals at multi-gigahertz frequencies. High speed and low power clock and data distribution networks have been described in this chapter that respond to this need.

Three oscillatory clock distribution networks exploiting transmission line properties are presented in this chapter. Coupled standing wave oscillators (SWO) exploit the standing wave phenomenon in transmission lines, distributing a global clock signal with constant phase and varying amplitude. An on-chip 10 GHz SWO clock signal distribution network has been demonstrated. Contrary to the SWO, rotary traveling wave oscillators (RTWO) exploit the traveling wave phenomenon in transmission lines, distributing a global clock signal with constant amplitude and varying phase. An on-chip 965 MHz RTWO clock distribution network has been demonstrated. A resonant distributed differential oscillator (DDO) using on-chip inductors to resonate the network capacitance has also been proposed, and demonstrated by a 1.1 GHz clock distribution network. A significant reduction in power consumption (2 to 16 times) as compared to traditional clock distribution networks is shown for all three methods.

Furthermore, two novel families of data transmission networks are presented. The *LC* dominant interconnect technique exploits the properties of wide transmission lines operating at high frequencies. The transmitted signal is modulated to higher frequencies, achieving lower attenuation and higher traveling speed. Pulse signal interconnects are also described. By transmitting short width pulses, signal propagation near the speed of light is exhibited while consuming significantly lower power as compared to repeater insertion.

Chapter 4

On-Chip Inductors

Until recently, on-chip inductors have been mostly used with radio frequency (RF) integrated circuits (ICs). As technology scales and the operating frequencies of digital circuits reach the multi-gigahertz domain, on-chip inductors can be exploited in high speed digital applications. Resonant clock distribution networks and quasi-resonant interconnects, presented in Chapters 5 and 6, respectively, are two high speed digital circuits that use on-chip inductors. By intentionally placing large on-chip reactance sources, inductance is harnessed to achieve resonance in multi-gigahertz digital circuits.

RF circuits and technologies have been gaining renewed interest due to the demand and development of portable wireless products. Since most RF circuits (*e.g.*, voltage control oscillators, low noise amplifiers, and mixers) utilize inductors, the integration of inductors on-chip has become necessary. The lossy silicon substrate and the planar nature of CMOS technology, however, make realizing high quality on-chip inductors difficult. As compared to a classic solenoid structure of a discrete inductor, planar inductors do not scale well with the number of wire turns, requiring large area. Regardless, the relatively low cost of silicon fabrication processes makes CMOS technology the process of choice in many RF IC applications [88].

In this chapter, on-chip inductors are classified in section 4.1, while models of spiral planar inductors are described in section 4.2. Finally, guidelines and tradeoffs in the design of on-chip spiral inductors are provided in section 4.3.

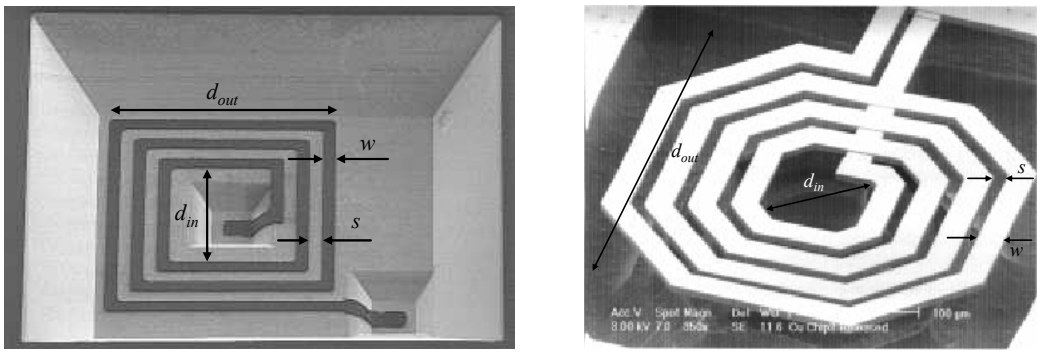
4.1 Types of On-Chip Inductors

In order to improve performance while minimizing losses in on-chip inductors, different types of inductors have been developed. The most common type are spiral (planar) inductors, which can be integrated into a standard CMOS technology. These inductors, however, are relatively lossy, exhibiting low efficiency. Advanced on-chip inductors, which require a customized CMOS process, have also been proposed. These inductors usually require a complex fabrication process with magnetic materials. These advanced inductors achieve higher performance than spiral inductors due to reduced losses. As an alternative to passive inductors, active inductors are composed of semiconductor devices (*i.e.*, transistors) which emulate the output impedance of a passive inductor. Active inductors occupy significantly less area and are not affected by the loss mechanisms of passive on-chip inductors. Active inductors, however, usually exhibit higher noise, distortion, and power consumption than passive inductors [92]. Passive and active inductors are described in the following subsections.

4.1.1 Spiral (Planar) On-Chip Inductors

The most widely used on-chip spiral inductor is the planar inductor. Different geometries of spiral inductors are illustrated in Figure 4.1. A square inductor is the most common and simplest shape due to the combination of perpendicular vertical and horizontal wires, a capability available in all physical layout tools. For hexagonal and circular inductors, less metal is required to achieve a similar number of turns. The series resistance, therefore, is reduced while the inductor performance is enhanced. A square shaped inductor, however, is

more area efficient. For example, if a square inductor utilizes 100% of an area, the octagonal and circular inductors utilize only 82.8% and 78.5%, respectively [90]. Note that the magnitude and performance of an inductor are dependent on the shape; typically, the performance can be improved by 10% [92].



(a) (b)

(c)

Figure 4.1 Planar spiral inductors: (a) square inductor [89], (b) octagonal inductor [90], (c) circular inductor [91]

On-chip inductors are typically realized on the highest metal layer. In this configuration, the inductor achieves the lowest parasitic resistance and capacitance to the substrate. To

further reduce resistive losses while increasing the efficiency of planar on-chip inductors, several enhanced topologies have been proposed.

A patterned ground shield (PGS) [93], shown in Figure 4.2, is inserted between the spiral inductor and the substrate, reducing the capacitive coupling to the lossy substrate. The slots between the ground strips reduce the magnetic flux by preventing the flow of the induced current along the path indicated in Figure 4.2. Concurrently, the shielding reduces coupling noise between the substrate and the inductor. Over a 1 to 2 GHz frequency range, the addition of the shield increases the quality factor of the inductor by 33% while reducing the substrate coupling between two adjacent inductors by 25 dB [93]. The drawback of this method, however, is the reduced self-resonance frequency of the inductor due to the increased capacitance.

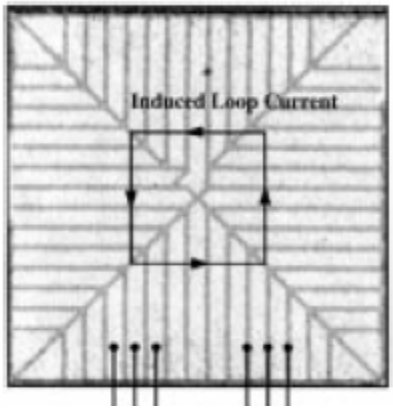
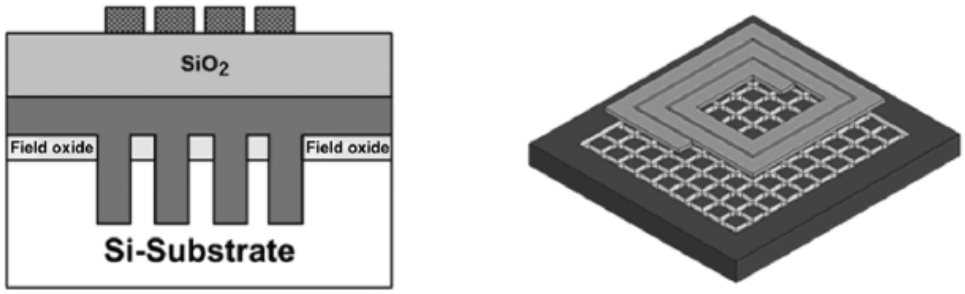


Figure 4.2 Patterned ground shield [93]

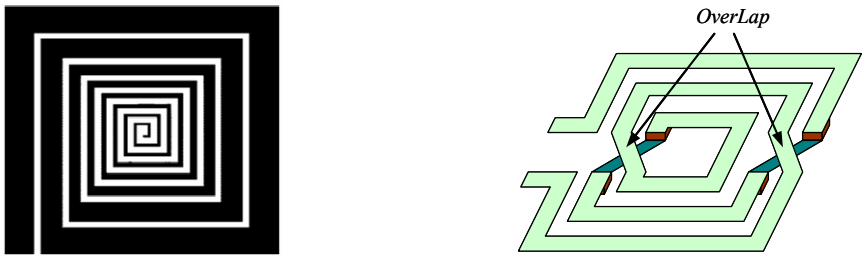
By removing portions of the substrate using a deep trench technology (see Figure 4.3), a reduction in the substrate capacitive coupling and the parasitic resistance can be achieved. A

deep trench inductor has been fabricated in a SiGe 0.35 μm BiCMOS technology, increasing the self-resonance and Q factor by 10% and 15%, respectively, as compared to a conventional on-chip inductor [96].



(a) (b)

Figure 4.3 Deep trench inductor [96]: (a) side view, (b) three dimensional view



(a) (b)

Figure 4.4 Reduced loss inductors [90]: (a) tapered spiral, (b) symmetric spiral

Two additional structures, as shown in Figure 4.4, reduce losses and improve the inductor performance [90]. The tapered inductor, shown in Figure 4.4(a), reduces the electromagnetic losses, generated mostly by the inner metal windings. By using tapered lines, the effective coupling capacitance and resistance to the substrate is reduced. The symmetric inductor, shown in Figure 4.4(b), reduces the parasitic capacitance of the inductor due to fewer overlaps between the metal lines of the inductor. Furthermore, in a symmetric

structure, the geometric center overlaps with the center of the electromagnetic field, increasing the mutual inductance.

Stacking spiral planar inductors in different metal layers can improve performance [94], as depicted in Figure 4.5. This configuration increases the effective metal length of the inductor, while requiring little area. The inductance can therefore be increased due to the positive contribution of the coupling inductance among the stacked inductors. In this manner, for an n layer stacked inductor, the nominal inductance is proportional to n^2. As an example, stacked inductors exhibit an inductance of 5 to 266 nH with a self-resonance of 0.5 to 11.2 GHz, respectively [94].

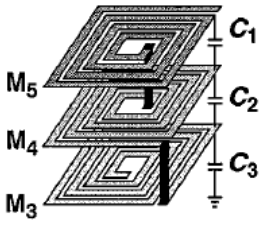
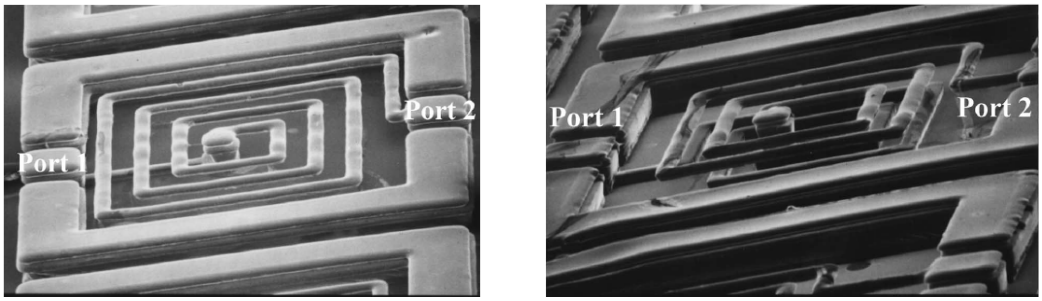


Figure 4.5 Three layer stacked inductor [94]

4.1.2 Advanced On-Chip Inductors

Different structures, materials, and fabrication processes have been used to develop high quality on-chip inductors. Three dimensional, thin film suspended inductors have been proposed [95], as shown in Figure 4.6. These on-chip inductors are realized using a sacrificial layer and electroplating within a conventional CMOS process. The suspended inductor structure reduces the capacitive coupling to the substrate, achieving a high inductance-to-area ratio. In an example, a double level suspended inductor achieves a peak quality factor of 27.7

(as compared to about 5 for a conventional on-chip inductor) with an inductance of 8.35 nH [95].



(a) (b)

Figure 4.6 Suspended inductor [95]: (a) single level, (b) double level

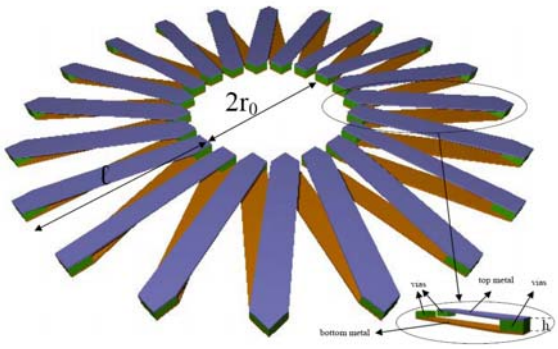
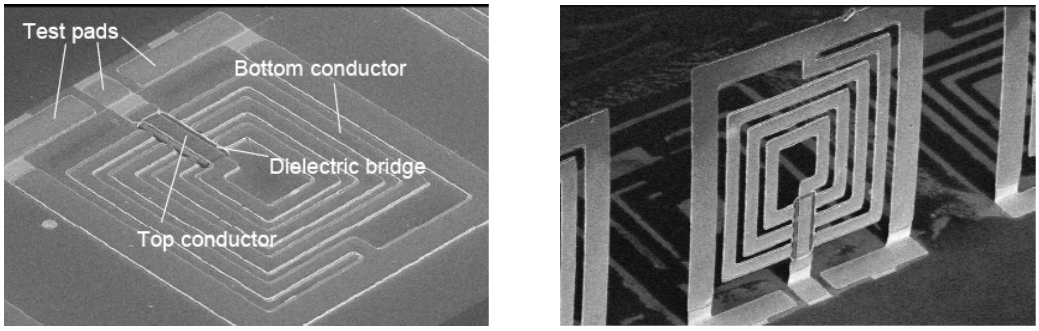


Figure 4.7 A toroidal inductor [97]

A toroidal inductor, integrated in a standard 0.13 μm CMOS technology, has been demonstrated [97]. In comparison with a conventional on-chip inductor, a toroidal inductor has fewer overlaps between the metal windings, exhibiting lower parasitic capacitance. The toroidal geometry is relatively compact and symmetric, and the inner circular area can be used for additional circuitry. The toroidal inductor achieves an inductance of 0.9 to 1.1 nH up to 20 GHz, while the quality factor reaches 10 at 15 GHz [97].

A different approach for reducing substrate related losses has been proposed in the form of vertical planar inductors [98], as illustrated in Figure 4.8. The planar spiral inductor is fabricated on a silicon substrate followed by a plastic deformation magnetic assembly (PDMA) process, which rotates the inductor from a horizontal to a vertical position [98]. This process includes three main phases.



(a) (b)

Figure 4.8 Vertical planar inductor [98]: (a) before the PDMA process, (b) after the PDMA process

In the first phase, the on-chip inductor is conventionally fabricated on a silicon substrate and coated with a magnetic material, *e.g.*, permalloy (NiFe) with an electroplating process. The second phase includes the release of the inductor structure from the substrate by etching the sacrificial layer underneath. In the third phase, an external magnetic field is applied, generating a magnetic torque which lifts the entire inductor structure. The vertical planar inductor occupies much less area than a traditional spiral inductor. Furthermore, the substrate losses are reduced due to the vertical structure. As an example, the vertical inductor exhibits a quality factor and self-resonance of 12 and 4 GHz, respectively, as compared to a conventional spiral inductor which exhibits a quality factor and self-resonance of 3 and 1 GHz, respectively [98].

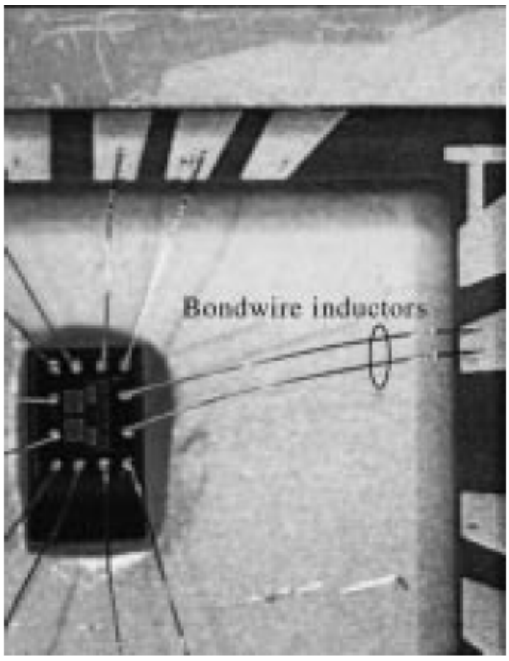


Figure 4.9 A bond wire inductor [99]

A different type of inductor can be realized by exploiting bond wires, as shown in Figure 4.9. Bond wires connect the circuit outputs to the external I/O pads. A typical diameter of a bond wire is about 25 μm [92], exhibiting a significantly larger area-to-length ratio than traditional planar inductors, which results in less resistive losses. Furthermore, since bond wires are placed far from the conductive planes, the parasitic capacitance is reduced, thereby increasing the self-resonance frequency while reducing the losses attributed to the induced image currents. The inductance of the bond wires, however, is poorly controlled due to the weak frequency dependence of the bond wires and physical geometries. The bond wire inductors can potentially reach a quality factor of 50 at 1 GHz [92].

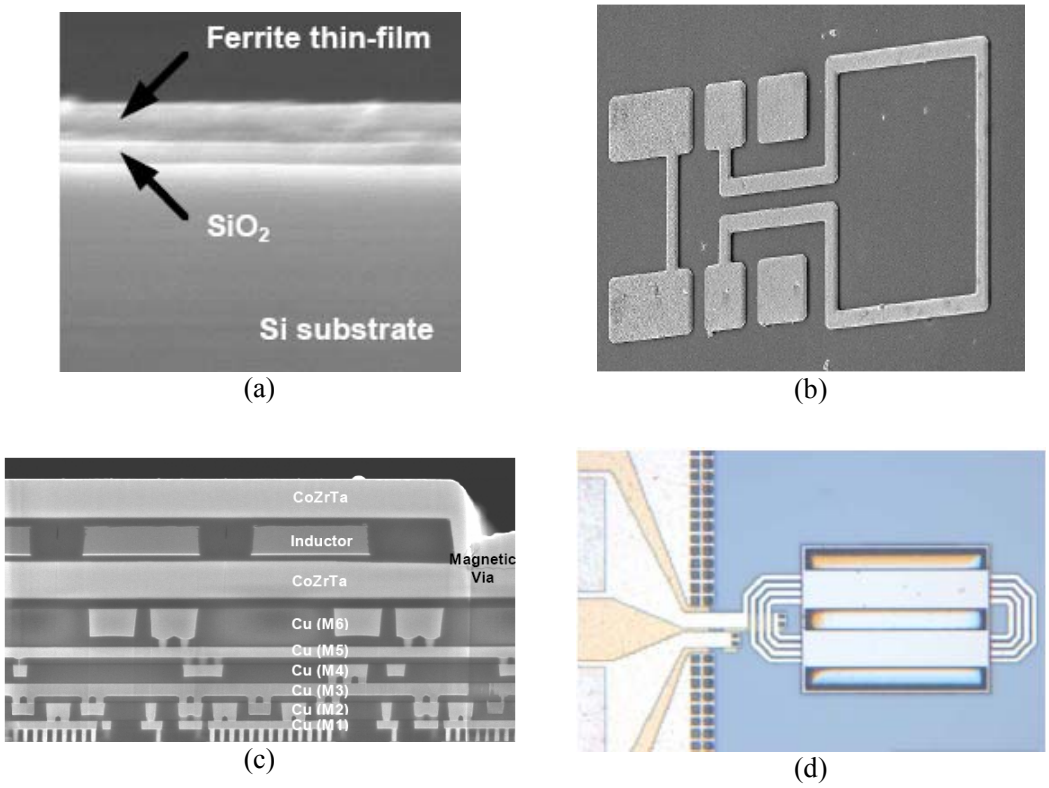


Figure 4.10 Magnetic thin film inductors: (a) thin film ferrite [100], (b) magnetic thin film inductor [100], (c) cross section of metal and magnetic layers [101], (d) spiral inductor with magnetic film [101]

A different approach to improving the performance of on-chip inductors is to combine thin magnetic films with spiral inductors, as shown in Figure 4.10. The magnetic material confines the magnetic flux within a well defined path, decreasing the magnetic losses. By increasing the flux linkage, the inductor magnitude and quality factor can be increased, while decreasing the inductor area. To realize these inductors, advanced materials with good magnetic properties in the multi-GHz region are required. Furthermore, novel inductor structures integrated with magnetic materials as well as proper fabrication processes are necessary [100]. A ferrite material composed of Co_7ZrO_9 has been integrated with a spiral

inductor, achieving improvement in the quality factor and inductance by 40% and 17%, respectively, as compared to a standard spiral inductor [100]. In [101], a magnetic material composed of an amorphous CoZrTa is used, achieving an order of magnitude increase in the inductance and quality factor in a 130 nm CMOS technology as compared to previously fabricated thin film on-chip inductors.

4.1.3 Active Inductors

In contrast to passive on-chip inductors as described in sections 4.1.1 and 4.1.2, active inductors (or synthetic inductors) are composed of active devices, *i.e.*, transistors, which emulate an inductive impedance. The working principles and a performance comparison between passive and active on-chip inductors are presented in the following subsections.

4.1.3.1 Principles of Active Inductors

Most active inductor circuits employ some kind of shunt feedback with an amplifier [102], as shown in Figure 4.11. This configuration can be generally described by the admittance matrices of the amplifier Y_a and feedback Y_f circuits, respectively,

$$Y_a = \begin{bmatrix} y_{11,a} & y_{12,a} \\ y_{21,a} & y_{22,a} \end{bmatrix}, \tag{4.1}$$

$$Y_f = \begin{bmatrix} y_{11,f} & y_{12,f} \\ y_{21,f} & y_{22,f} \end{bmatrix}. \tag{4.2}$$

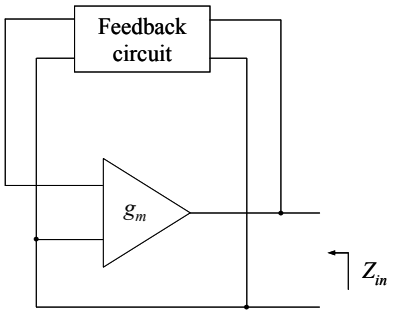


Figure 4.11 A typical configuration of an active inductor

The overall admittance of the circuit shown in Figure 4.11 is

$$Y_t = Y_a + Y_f = \begin{bmatrix} y_{11,a} + y_{11,f} & y_{12,a} + y_{12,f} \\ y_{21,a} + y_{21,f} & y_{22,a} + y_{22,f} \end{bmatrix} = \begin{bmatrix} y_{11,t} & y_{12,t} \\ y_{21,t} & y_{22,t} \end{bmatrix}, \tag{4.3}$$

while the input impedance is

$$Z_{in} = \frac{y_{11,t}}{y_{11,t} \cdot y_{22,t} - y_{12,t} \cdot y_{21,t}}. \tag{4.4}$$

For an ideal inductor, Z_{in} should satisfy

$$\Re(Z_{in}) = 0, \tag{4.5}$$

$$\Im(Z_{in}) > 0. \tag{4.6}$$

The inductance is therefore

$$L = \frac{\Im\{Z_{in}\}}{\omega}. \tag{4.7}$$

4.1.3.2 Performance Comparison between Passive and Active Inductors

Active inductors achieve a much higher quality factor (Q) than passive inductors, since active inductors consume power to emulate the inductive impedance. For example, active inductors can achieve Q values well beyond 10, while passive inductors exhibit Q values of about 5. Furthermore, in active inductors, the Q and frequency at which Q reaches a maximum are independently tunable [103]. Another significant advantage is that active inductors occupy only a fraction of the area occupied by passive inductors. Furthermore, lower electromagnetic interference (EMI) is produced by active inductors.

Unlike passive inductors, active inductors consume power to generate an inductive impedance. A tradeoff, therefore, between the quality factor and power consumption exists. Linearity is another disadvantage of active inductors. Due to the nonlinear large signal behavior of active devices, fluctuations in the impedance are exhibited. Active inductors are therefore not suitable in large signal circuits. The nonlinearity of active inductors may also result in unwanted harmonics [103]. A significant disadvantage of active inductors is the noise produced by the transistors due to thermal and flicker noise. Typically, the noise increases with a higher resistance in the feedback path and with smaller transistor sizes.

4.2 Modeling On-Chip Spiral Inductors

Since on-chip spiral (planar) inductors are widely used in ICs, models of these inductors are discussed in this section. Accurate models that capture the losses associated with spiral inductors are important for reliable circuit design and analysis. Loss mechanisms of spiral inductors, quality factor, and some common models are reviewed in the following subsections.

4.2.1 Loss Mechanisms in On-Chip Spiral Inductors

The most significant source of on-chip spiral inductor losses is the metal wires and silicon substrate. Since an inductor is constructed from metal lines, the resistance of the spiral inductor converts magnetic energy into heat. To decrease this parasitic resistance, on-chip inductors are typically formed on the top metal layers, where the metal lines are wider. Another loss mechanism associated with metal conductors occurs at high operating frequencies (*i.e*, in the gigahertz domain). As the operating frequency increases, eddy currents are induced in neighboring lines. At these high frequencies, the current flow within the metal lines is not uniformly distributed, resulting in skin and proximity effects. The magnetic field of the inductor penetrates into the metal wires, producing an opposing electric field. The induced electric field increases the impedance in the center of the conductor, resulting in an accumulation of current at the outer layer or skin of the metal conductor. Hence, this phenomenon is called the skin effect. The decreased effective cross-section of the conductor increases the current density, thereby converting additional energy into heat. The effective depth of the conductor in which current flows is

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}, \tag{4.8}$$

where ω, μ, and σ are the radian frequency, permeability, and conductivity, respectively [105]. Furthermore, the increased resistance of a conductor also contributes to the magnetic field in neighboring wires, a phenomenon called the proximity effect [104]. The magnetic field of the spiral inductor windings add, thereby increasing the effective resistance.

The other significant source of losses is the silicon substrate on which the integrated circuits and metal layers are placed. Magnetic energy generated by the inductor is converted into heat due to the conductive nature of the silicon substrate, as opposed, for example, to the insulating nature of GaAs. Silicon substrate losses can be associated with three primary sources. First, the electric field is coupled to the substrate through displacement currents [106], which flow within the substrate to a nearby ground. Second, time varying magnetic fields produce time varying electric fields which induce current flow within the substrate. This magnetic field also induces currents known as eddy current. The current generated by the induced electric fields flow vertically or laterally, perpendicular to the spiral inductor, while the currents induced by the magnetic field flow parallel to the spiral inductor. Third, radiation effects induce currents in the substrate. Note that losses associated with radiation are less prominent than losses produced by electromagnetic fields.

4.2.2 Models and Performance Metrics of Spiral Inductors

In this section, traditional and advanced models of on-chip spiral inductors are reviewed. Performance metrics in the form of a quality factor (Q) and self-resonance are also discussed in the following subsections.

4.2.2.1 Basic Model, Q, and Self-Resonance of On-Chip Inductors

The most common lumped element model of an on-chip spiral inductor is shown in Figure 4.12. This circuit models the primary parasitic capacitances and resistances associated with on-chip inductors. Capacitance C_s represents the coupling capacitance between the metal windings of the inductor. The inductance and parasitic resistance are represented by L_s and R_s, respectively. The capacitance between the metal segments and the substrate is C_{ox}, while the resistance and capacitance of the substrate are represented by R_{sub} and C_{sub}, respectively. Note that this basic model does not consider skin, proximity, and substrate eddy current effects. This circuit model is suitable for relatively narrow band frequencies.

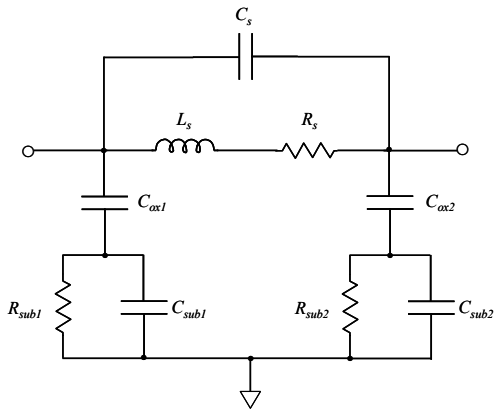


Figure 4.12 Generic lumped model of a spiral inductor

The quality factor (Q) represents the energy storing ability, or the rate at which the energy is dissipated by a physical system. Equivalently, Q is the ratio between the maximum amount of stored energy and the dissipated energy during one period of the system response,

$$Q \equiv 2\pi \cdot \frac{maximum\ energy\ stored}{total\ energy\ lost\ in\ one\ oscillation\ cycle}. \tag{4.9}$$

Specifically, for an electric system, Q is the difference between the peak magnetic and electric energy in one oscillation cycle [94],

$$Q \equiv 2\pi \cdot \frac{peak\ magnetic\ enerygy\ -\ peak\ electric\ energy}{energy\ loss\ in\ one\ oscillation\ cycle}. \tag{4.10}$$

Equivalently, since energy can only be stored by an inductor or capacitor, and dissipated only by a resistor [107], Q can be defined as

$$Q \equiv 2\pi \cdot \frac{\left[w_L(t) + w_C(t)\right]_{max}}{P_R \cdot T}, \tag{4.11}$$

where $w_L(t)$ and $w_C(t)$ are the instantaneous maximum energy stored by the inductor and capacitor, respectively, while P_R is the energy dissipated by a resistor in a time period T. Finally, a useful definition of Q based on the input impedance of a network is

$$Q \equiv \frac{\Im\{Z\}}{\Re\{Z\}}, \tag{4.12}$$

where Z is the impedance seen at one terminal, while the second terminal is grounded. Applying (4.12) to the model of Figure 4.12 results in [100]

$$Q = \underbrace{\left(\frac{\omega L_s}{R_s}\right)}_{(a)} \cdot \underbrace{\left(\frac{R_p}{R_p + R_s\left(1+\left(\omega L_s / R_s\right)^2\right)}\right)}_{(b)} \cdot \underbrace{\left(1 - \frac{R_s^2\left(C_s + C_p\right)}{L_s} - \omega^2 L_s\left(C_s + C_p\right)\right)}_{(c)}, \tag{4.13}$$

where

$$R_p = \frac{1}{\omega^2 C_{ox}^2 R_{sub}^2} + \frac{R_{sub}\left(C_{ox} + C_{sub}\right)^2}{C_{ox}^2}, \tag{4.14}$$

$$C_p = C_{ox}\frac{1+\omega^2\left(C_{ox}+C_{sub}\right)C_{sub}R_{sub}^2}{1+\omega^2\left(C_{ox}+C_{sub}\right)^2 R_{sub}^2}. \tag{4.15}$$

In (4.13), expression (a) represents the magnetic energy stored in the inductor and the ohmic loss due to the series resistance. Expression (b) represents the energy dissipated in the silicon substrate, while expression (c) is the self-resonance factor, which decreases Q due to the increase in the electric energy. By substituting (4.15) into (c), equating to zero, and solving for ω, the self-resonance frequency of the inductor can be determined [100]. Self-resonance occurs when the magnetic energy and electric energy are equal. At this energy state of the inductor, Q is zero. Equivalently, the frequency at which the imaginary portion of the input impedance equals zero is the self-resonance frequency. At this frequency, the inductor behaves as a resistive network rather than as an inductor. Beyond the self-resonance frequency, the inductor behaves as a series RC network [102].

4.2.2.2 Advanced Models of On-Chip Inductors

To improve the accuracy of on-chip inductor models and consider high frequency effects, modifications to the basic model represented by Figure 4.12 have been proposed. A circuit model of skin and proximity effects as well as substrate losses has been proposed [108], as depicted in Figure 4.13.

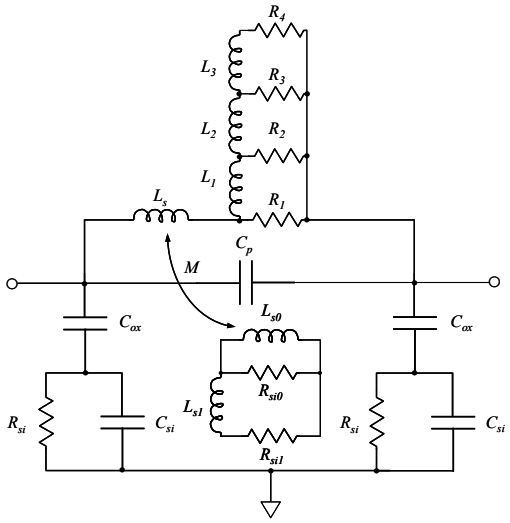


Figure 4.13 On-chip inductor model considering high frequency effects

An *RL* ladder structure (composed of L_1 to L_3 and R_1 to R_4) models skin and proximity effects in the metal lines. The effective inductance is modeled by L_s, while the capacitive coupling is represented by C_p. The dielectric capacitance is represented by C_{ox}, while the electric loss of the substrate is modeled by the parallel combination of R_{si} and C_{si}. The magnetic loss of the substrate due to the eddy current effect is modeled by an *RL* ladder structure (composed of Lsi_0, L_{si1}, R_{si0}, and R_{si1}). The effective reduction of inductance due to negative coupling between the substrate eddy current and the metal lines is modeled by the

coupling coefficient *M*. Note that this model considers different conductivity modes of the substrate, affecting the electrical behavior of the spiral inductor.

A scalable broadband model incorporating substrate and conductor losses has been proposed [109], as shown in Figure 4.14. A broadband characteristic of on-chip inductors reaching up to 20 GHz (well beyond a typical self-resonance frequency) is demonstrated by this model. This model considers three-dimensional eddy current effects (rather than two-dimensional). Two parallel *RL* networks (composed of R_{loss1}, L_{sub1}, R_{loss2}, and L_{sub2}) in series with C_{ox1} and C_{ox2} model the eddy current component normal to the substrate plane, while C_{ox} is the capacitance between the inductor and the substrate.

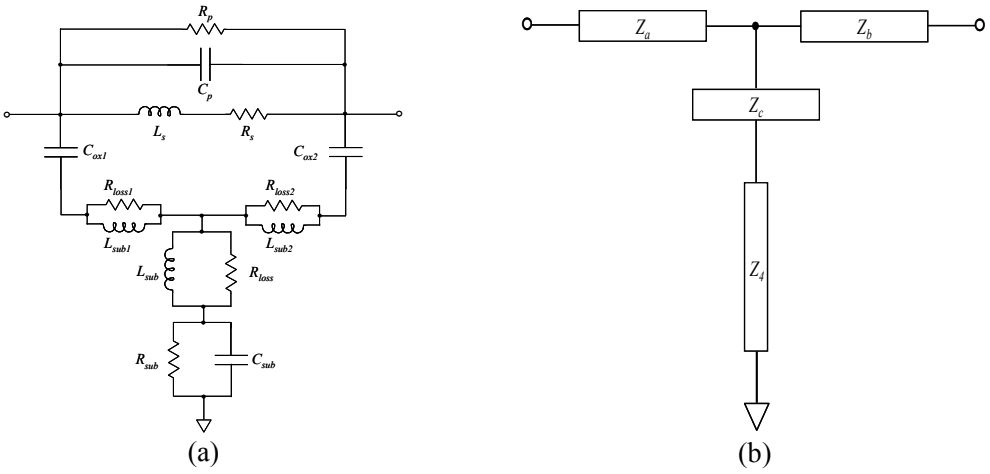


Figure 4.14 An on-chip spiral inductor model: (a) lumped model, (b) simplified T model

The parallel *RL* network (composed of R_{loss} and L_{sub}) in series with the parallel *RC* network (composed of C_{sub} and R_{sub}) represents the eddy current component within the substrate plane, while the *RC* network models the lossy substrate resistance and capacitance. The series inductance and resistance of the spiral inductor are modeled by L_s and R_s,

respectively, while the inter-coil and output terminal capacitances are modeled by C_p. Finally, skin and proximity effects within the spiral conductor due to the lossy return path in the substrate are represented by R_p. A design flow for extracting the model parameters and the equivalent T model shown in Figure 4.14(b) using scattering parameters is described in [109].

4.3 Design of On-Chip Inductors

The design of a spiral inductor is a multi-variable problem incorporating the following parameters: geometry and area, inductance magnitude, operating frequency, quality factor Q, frequency at which Q reaches a maximum value, and self-resonance frequency. These parameters are determined by the geometry and fabrication technology of the inductor. Since these parameters exhibit opposing trends, tradeoffs are frequently encountered in the design process. Guidelines for designing on-chip spiral inductors are provided in the first section. First order design equations, design methodologies, and tradeoffs are discussed in the following subsections.

4.3.1 First Order Inductance Equations

A variety of first order inductance equations have been described in the literature [92], [110], [111], [112]. The simplest inductance expressions for on-chip spiral inductors do not consider important loss mechanisms, and may exhibit up to 50% error in the inductance. More accurate expressions [110], however, which consider parasitic effects are cumbersome and require complex computations. Sufficiently accurate (exhibiting up to 5% error in the

inductance) and simple expressions can be used to approximate the magnitude of on-chip spiral inductors.

The modified Wheeler expression, originally derived for a circular coil inductor [111], can be applied to a square spiral inductor with inner and outer diameters, as depicted in Figure 4.15,

$$L = \frac{9.375\mu_r \mu_0 n^2 d_{avg}^2}{11d_{out} - 7d_{avg}}, \tag{4.16}$$

$$d_{avg} = \frac{d_{in} + d_{out}}{2}, \tag{4.17}$$

where d_{in}, d_{out}, and n are the inner and outer diameter and number of turns, respectively. The relative and absolute permeability are μ_r and μ_0, respectively.

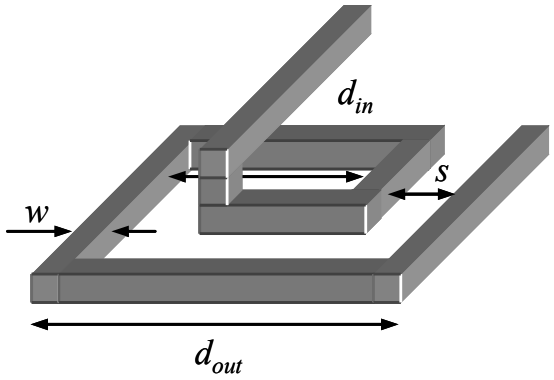


Figure 4.15 A square spiral inductor

Alternatively, the inductance of four commonly used inductor shapes is incorporated into one expression (the current sheet formula) [112],

$$L = \frac{\mu_r \mu_0 n^2 d_{avg} c_1}{2} \left(\ln\left(\frac{c_2}{\rho}\right) + c_3 \rho + c_4 \rho^2 \right), \tag{4.18}$$

$$\rho = \frac{d_{out} - d_{in}}{d_{out} + d_{in}}, \tag{4.19}$$

where ρ is the fill factor and c_1, c_2, c_3, and c_4 are a function of geometry, and listed in Table 4.1. Finally, for any polygon geometry, the inductance can be approximated by [112]

$$L = \frac{\mu_r \mu_0 n^2 d_{avg} A_{out}}{\pi d_{out}^2} \left(\ln\left(\frac{2.46 - 1.56/N}{\rho}\right) + \left(0.2 - \frac{1.12}{N^2}\right) \rho^2 \right), \tag{4.20}$$

where A_{out} is the outer area occupied by the inductor and N is the number of polygon sides. Note that (4.20) is a modified version of (4.18) with an analytic approximation of the coefficients, c_1, c_2, c_3, and c_4.

Table 4.1 Coefficients for the current sheet inductance [112]

Geometry	c_1	c_2	c_3	c_4
Square	1.27	2.07	0.18	0.13
Hexagon	1.09	2.23	0.00	0.17
Octagon	1.07	2.29	0.00	0.19
Circle	1.00	2.46	0.00	0.20

4.3.2 Design Practices and Tradeoffs

The design of an on-chip inductor is a complicated task, requiring complex electromagnetic calculations. To achieve accurate results, therefore, two- and three-dimensional electromagnetic simulation tools, such as ASITIC [113], SPIRAL [114], and HFSS [115], are often used. These electromagnetic simulators solve Maxwell's equations for a set of geometric and technology parameters. The electric and magnetic fields are obtained for different outer and inner diameters, spacings, and widths of the inductor metal windings. These electromagnetic simulations are used to minimize the area or maximize Q for a specific inductance operating at a target frequency. In order to achieve an efficient solution, simplifying assumptions to Maxwell's equations are often used, sacrificing accuracy for computational efficiency.

A less accurate and complex approach is to use models of on-chip spiral inductors, such as presented in section 4.2.2. The parameters of these models are extracted by first order equations. Subsequently, performance metrics characterizing an inductor can be approximated by the model. To demonstrate this approach, an example based on a square spiral inductor and the basic model shown in Figure 4.12 is considered.

The geometry of a square spiral inductor (as shown in Figure 4.15) can be described by

$$d_{out} = d_{in} + 2 \cdot w \cdot n + 2 \cdot s \cdot (n-1), \tag{4.21}$$

where d_{out} and d_{in} are the outer and inner diameter, respectively, and w, s, and n are the width, spacing, and number of turns, respectively. For maximum area occupied by an inductor (*i.e.*, d_{out}), different sets of inductor parameters d_{in}, w, s, and n can be obtained according to (4.21),

from which the basic model is determined. The inductance L_s can be found by using, for example, expressions (4.18) and (4.19). The series resistance R_s is obtained from

$$R_s = \frac{l}{w \cdot \sigma \cdot \delta \cdot \left(1 - e^{-t/\delta}\right)}, \tag{4.22}$$

where σ is the conductivity, l is the total length of the inductor, w and t are the width and thickness of the interconnect, respectively, and the skin depth δ is given by (4.8). The capacitance between the windings of the spiral inductor C_s is

$$C_s = n \cdot w^2 \cdot \frac{\varepsilon_{ox}}{t_{ox}}, \tag{4.23}$$

where t_{ox} is the oxide thickness between the crossunder section and the main spiral sections and ε_{ox} is the oxide dielectric. The capacitance between the spiral inductor and the substrate (assuming $C_{ox1} = C_{ox2} = C_{ox}/2$) is

$$C_{ox} = w \cdot l \cdot \frac{\varepsilon_{ox}}{t_{ox}}. \tag{4.24}$$

The resistance R_{sub} models the substrate resistance and is approximated (assuming $R_{sub1} = R_{sub2} = R_{sub}/2$) by

$$R_{sub} = \frac{2}{w \cdot l \cdot G_{fp}}, \tag{4.25}$$

where G_{fp} is a fitting parameter, which is a function of the substrate material and distance between the inductor and substrate. A typical value of G_{fp} is about 10^{-7} S/μm^2 [92]. Finally, the capacitance of the substrate C_{sub} (assuming $C_{sub1} = C_{sub2} = C_{sub}/2$) as well as other reactive elements related to the image currents is

$$C_{sub} = \frac{w \cdot l \cdot C_{fp}}{2}, \tag{4.26}$$

where C_{fp} is a fitting parameter which is a function of the substrate material and distance between the inductor and substrate. A typical range of C_{fp} is between 10^{-3} to 10^{-2} fF/μm^2 [92]. Once the parameters of the basic model are extracted, the quality factor and self-resonance of an inductor can be estimated based on (4.13).

Since the design of an on-chip spiral inductor is a complex task, general design guidelines and tradeoffs are useful. A list of suggested intuitive conventions is provided below:

1) The main component of an inductor is usually implemented on the top most metal layer, where the metal lines are widest, reducing resistive losses. Furthermore, maximizing the distance to the substrate minimizes the parasitic capacitance between the inductor and the substrate;
2) The resistance due to eddy currents in the substrate is proportional to the square of the frequency, to the square of the number of turns, and to the cube of the inductor diameter [92]. Therefore, beyond a certain threshold, losses associated with eddy currents become dominant, degrading Q. In heavily doped CMOS substrates, inductors with a smaller outer diameter and narrow wires are often used, contrary to semi-insulating substrates like GaAs;

3) The geometry of an inductor has a second order effect on the inductance and *Q*. Square spiral inductors are therefore the most commonly used;
4) Inductors with fewer inner turns are preferred, since the inner most turns contribute significant resistance and a small amount of magnetic flux. A practical ratio of three to one between the outer and inner diameters is often used [92];
5) The distance between the inductor wires is typically chosen to be the smallest possible to reduce the inductor area and increase the mutual inductance. This short spacing, however, reduces the self-resonance frequency due to the increased inductive and capacitive coupling.

4.4 Summary

Due to the demand for wireless devices, RF integrated circuits incorporating on-chip inductors have been extensively developed. These on-chip inductors are now being considered for high speed digital circuits. On-chip inductors, however, exhibit low *Q*, thereby degrading the overall performance of a circuit. To reduce the on-chip losses associated with spiral inductors, different types of inductors have been proposed and are reviewed in this chapter. Generally, an increase in the *Q* factor comes at the expense of complex and costly fabrication processes.

Accurate models of on-chip inductors are important in the design and analysis of high speed circuits. To capture the loss mechanisms of these inductors, a variety of models have been proposed and are described in section 4.2. Most of these models are based on the basic lumped element model. To incorporate losses due to eddy currents and skin and proximity

effects, additional elements have been added to the model, as depicted in Figures 4.13 and 4.14. Finally, guidelines and tradeoffs in the design of on-chip spiral inductors are described.

Chapter 5

Design Methodology for Global Resonant H-Tree Clock Distribution Networks

Design guidelines for resonant H-tree clock distribution networks are presented in this chapter. A distributed model of a two level resonant H-tree structure is described, supporting the design of low power, skew, and jitter resonant H-tree clock distribution networks. Excellent agreement is shown between the proposed model and SpectraS simulations. A case study is presented that demonstrates the design of a two level resonant H-tree network, distributing a 5 GHz clock signal in a 0.18 μm CMOS technology. This example exhibits an 84% decrease in power dissipation as compared to a standard H-tree clock distribution network.

The design methodology enables tradeoffs among design variables to be examined, such as the operating frequency, the size of the on-chip inductors and capacitors, the output resistance of the driving buffer, and the interconnect width. A sensitivity analysis of resonant H-tree clock distribution networks is also provided. The effect of the driving buffer output resistance, on-chip inductor and capacitor size, and signal and shielding transmission line width and spacing on the output voltage swing and power consumption is described.

This chapter is organized into six sections. Introduction and problem formulation of the resonant clock network are presented in sections 5.1 and 5.2, respectively. In section 5.3, design guidelines are provided. In section 5.4, a case study is presented. The sensitivity of a resonant clock distribution network to certain design parameters is examined in section 5.5.

Summary is offered in section 5.6. A derivation of the H-tree sector model is described in Appendix A.1, while the *rms* input voltage is characterized in Appendix A.2. A model of an equivalent shorted interconnect is provided in Appendix A.3.

5.1 Introduction

To combat the difficulties associated with clock distribution networks, clock generation and distribution networks based on *LC* oscillators in the form of transmission line systems have been considered. In salphasic clock distribution networks [116], a sinusoidal standing wave is established within a transmission line. Coupled standing oscillators of this type are used in [70] to distribute a high frequency clock signal. A similar approach uses traveling waves in coupled transmission line loops [71] driven by distributed cross coupled inverters. Comprehensive and systematic investigation of the impact of width, spacing, and loading of a resonant clock tree on skew and energy consumption is presented in [14], [117]. In [72], [73], a resonant global clock distribution network is described where on-chip spiral inductors and decoupling capacitors are connected to traditional clock trees.

The strategy presented in [72] is to design the H-tree sector and the clock grid for sufficient bandwidth to support distributed square waveform slew rates. Hence, distributing a square clock signal with a slew of t_r requires a bandwidth of at least $0.5/t_r$. From the Fourier theorem, as the bandwidth increases, the time domain response more closely resembles an ideal square wave signal. Unfortunately, the output resistance of the driving buffer and the rise time both have a major effect on the bandwidth.

A comprehensive and robust design methodology for resonant H-tree clock distribution networks is presented [118]. The methodology is based on the transfer function of a two level

H-tree, defined here as a sector, such that the fundamental harmonic of the input square wave is transferred to the output. The output signal at the leaf nodes exhibits a sinusoidal behavior. Inverters are placed at the leaf nodes to convert the sinusoidal waveform into a quasi-square waveform. On-chip spiral inductors and capacitors are used to resonate the clock signal around the harmonic frequency.

Significant variations are exhibited in modern high performance, high complexity integrated circuits fabricated in nanometer processes. With technology scaling, process variations have become a significant design factor that should be considered in the design process. Imperfections in the manufacturing process and environmental changes can degrade overall system performance [119]. Interconnect process variations can affect, for example, timing analysis, buffer insertion, high density SRAMs, and clock distribution networks [118]-[122]. Resonant H-tree structures tolerant to process variations are therefore important for high performance resonant clock distribution networks.

5.2 Background and Problem Formulation

The concept of exploiting resonant transmission lines was first introduced by Chi in 1994 [116]. A global resonant clock distribution network was later introduced in 2003 by Chan *et al.* [72]. In this circuit, a set of discrete on-chip spiral inductors and capacitors is attached to a traditional H-tree structure, as depicted in Figure 5.1. On-chip spiral inductors are connected at four points in the tree, while decoupling capacitors are attached to the other side of the spiral inductors. The capacitance of the clock distribution network resonates with the inductance, while the on-chip capacitors establish a mid-rail DC voltage around which the grid oscillates. This approach lowers the power consumption, since the energy alternates

between the electric and magnetic fields rather than dissipated as heat. Consequently, the number of gain stages is reduced, resulting in further reductions in power consumption, skew, and jitter.

The proposed resonant sector (such as the network shown in Figure 5.1) can be used as a building block, in a modular sense, to construct a much larger global clock distribution network, as shown in Figure 5.2. In this example, the entire network is divided into sectors of 16 leaves. Hence, the design flow is bottom to top, starting with the H-tree sectors at the leaf portion of the tree network and moving up to the central sector (or trunk).

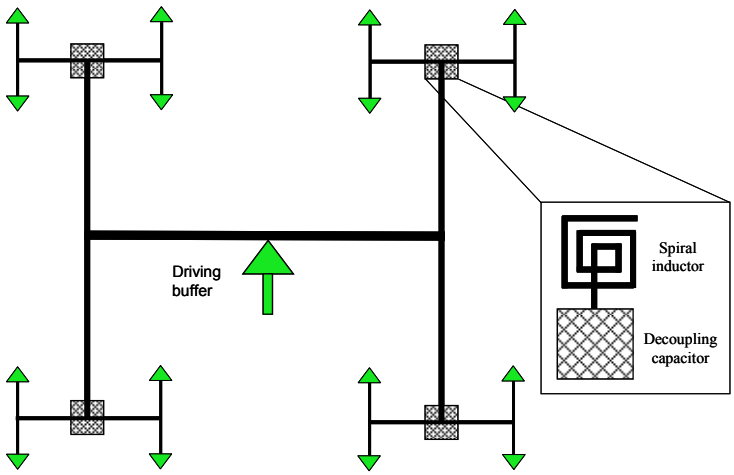


Figure 5.1 H-tree sector with on-chip inductors and capacitors

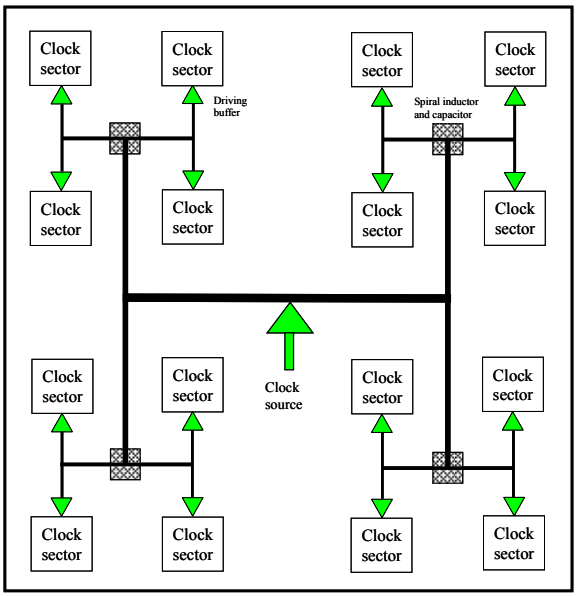


Figure 5.2 A global clock distribution network, consisting of 16 resonant clock sectors and a total of 256 leafs

The design methodology considers the physical geometry of the structure and the technology, and can be formulated as

$$\text{H-Tree Sector} = f\left(w_i, l_i, h_i, f_o, C_l\right) \ \forall i, \tag{5.1}$$

where w_i, l_i, and h_i are the width, length, and thickness of each section of the H-tree sector, respectively, f_o is the clock frequency, and C_l is the capacitive load at each leaf node. The index i varies between one and four, representing each section of the H-tree sector (see Figure 5.2). The H-tree sector function is used to determine the value of the on-chip spiral inductors (considering the effective series resistance), capacitors, and driving buffer resistance that produces the minimum power consumption.

5.3 Design Guidelines for H-Tree Sector

A methodology for designing resonant H-tree clock distribution networks is described in this section. In section 5.3.1, a distributed model is presented for the H-tree sector, while an on-chip spiral inductor model is presented in section 5.3.2. Applying the proposed models and a graphical representation of the design space, the optimum value of the on-chip inductors, capacitors, and output resistance of the driving buffer for minimum power consumption is determined as described in section 5.3.3.

5.3.1 H-Tree Sector Model

As clock signal frequencies exceed the multigigahertz regime, distributed models of interconnects are required to incorporate high frequency effects into the system behavior. The proposed model uses the classical distributed model, where an incremental section of line length Δz is modeled as a lumped element circuit, as shown in Figure 5.3. In this model, R, L, and C is the resistance, inductance, and capacitance per unit length, respectively. The lumped resistance represents the lossy component of the transmission line.

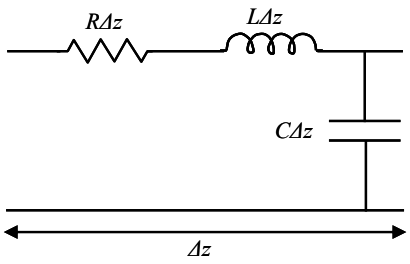


Figure 5.3 Distributed model of a transmission line

The proposed model is applied to a two level H-tree network, as depicted in Figure 5.1. The distributed *RLC* network shown in Figure 5.4 is used to model the clock tree depicted in Figure 5.1. The parameters R_i, L_i, and C_i are the resistance, inductance, and capacitance per unit length, respectively, where i varies from one through four. The parameter N is the depth of the tree, which in the example shown in Figure 5.1 equals four, while the number of leafs is 2^N.

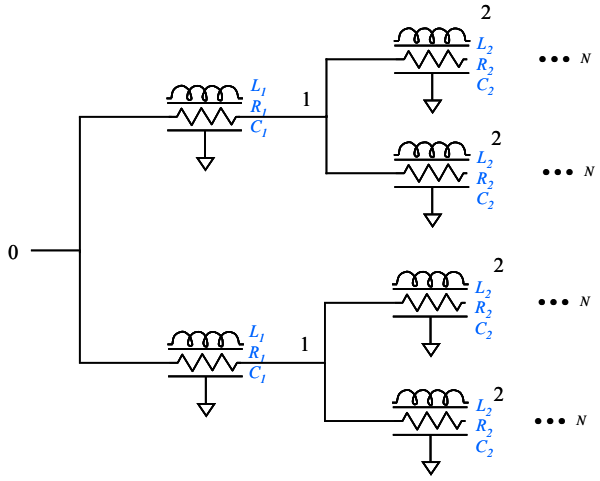


Figure 5.4 Distributed *RLC* network representation of an H-tree network

Since the two nodes labeled 1 in Figure 5.4 are symmetric, the waveforms at these nodes are assumed to be identical, and as a result, can be assumed to be shorted [123]. The same assumption applies to nodes 2, 3, and 4. This simplification is exploited to transform the circuit shown in Figure 5.4 into a distributed *RLC* transmission line as shown in Figure 5.5, making the analysis considerably simpler. Since the interconnect lines between each pair of nodes are assumed to be connected in parallel, the capacitance per unit length is increased by a factor of two, while the resistance and inductance per unit length is decreased by a factor of

two at each level of the hierarchy. A proof of the equivalent shorted interconnect is provided in Appendix A.3.

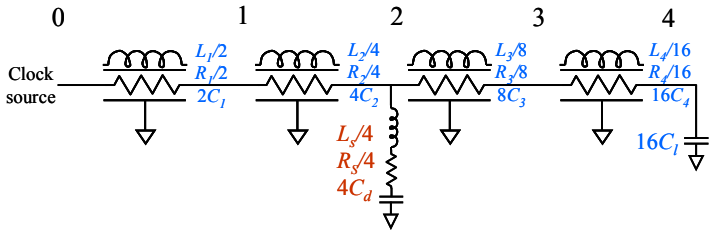


Figure 5.5 Resonant H-tree network simplified to a distributed *RLC* line

An analytic model of this structure is developed based on *ABCD* parameters [22]. From transmission line theory, the *ABCD* matrix for the overall structure is a product of the individual matrices,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = M_1 \cdot M_2 \cdot M_s \cdot M_3 \cdot M_4 \cdot M_l, \tag{5.2}$$

where M_i (i = 1..4), M_s, and M_l are the *ABCD* matrix of the four sections, the on-chip inductors and capacitors, and the load, respectively.

From the overall *ABCD* parameters, the transfer function *H(s)* and input impedance *Z(s)* of the system is, respectively,

$$H(s) = \frac{1}{A} = K(s) \cdot \frac{a_2 \cdot s^2 + a_1 \cdot s + a_0}{b_3(s) \cdot s^3 + b_2(s) \cdot s^2 + b_1(s) \cdot s + b_0(s)}, \tag{5.3}$$

$$Z_{in}(s) = \frac{A}{C} = \frac{b_3(s)\cdot s^3 + b_2(s)\cdot s^2 + b_1(s)\cdot s + b_0(s)}{d_3(s)\cdot s^3 + d_2(s)\cdot s^2 + d_1(s)\cdot s + d_0(s)}, \tag{5.4}$$

where a_0, a_1, and a_2 are constants and the parameters K, b_0, b_1, b_2, b_3, d_0, d_1, d_2, and d_3 are functions of frequency, the geometry of the structure, and the on-chip inductors and capacitors. These functions and the corresponding M_i matrices can be determined from (5.2), as described in Appendix A.1.

5.3.2 Model of On-chip Inductor

On-chip spiral inductors play an important role in the design of silicon-based RF integrated circuits (ICs). On-chip spiral inductors can be integrated into the fabrication process of a standard CMOS technology. Unfortunately, particularly in processes with a heavily doped silicon substrate, substrate losses resulting from eddy current effects can be significant [104], [124]. It is therefore important to accurately model the resistive losses of the on-chip spiral inductors.

A resistive-inductive model of the spiral inductor as shown in Figure 5.6 is used in this analysis. In order to provide a tractable solution, the parasitic capacitance is omitted from the model and treated as part of the on-chip decoupling capacitor C_d. The parasitic resistance of a spiral inductor greatly affects the behavior of the resonant clock sector, and is included in the circuit model of the inductor. To determine the value of the on-chip spiral inductors and the effective series resistance (ESR), customized extraction software, ASITIC, is used [113].

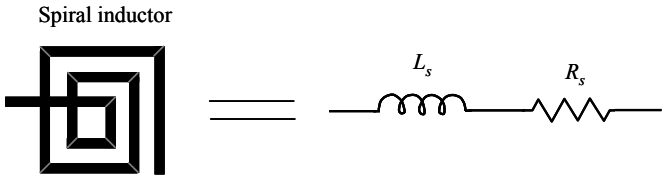


Figure 5.6 Simplified model of an on-chip spiral inductor

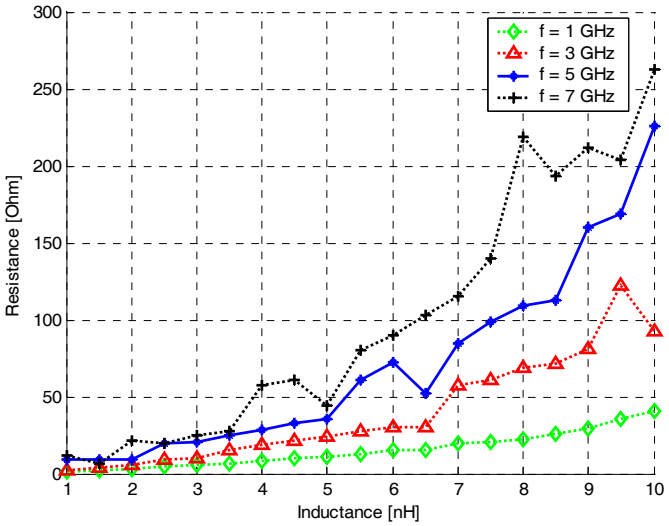


Figure 5.7 Effective series resistance of spiral inductors

In this example, the spiral inductors, occupying an area of 250 x 250 μm^2, are optimized for maximum *Q*. The ESR for a range of inductor values at certain frequencies is shown in Figure 5.7. Note that as the frequency increases, the ESR also increases. At low frequencies, the ESR exhibits a linear dependency with frequency, while at higher frequencies the relationship behaves quadratically. The values of ESR extracted at 5 GHz are used in the case study presented in section 5.4. The ESRs for the different inductors at a 5 GHz operating frequency are listed in Table 5.1.

Table 5.1 ESRs of spiral inductors at 5 GHz

L_s nH	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
ESR Ω	8	8	8	19	20	24	28	32	35	61	72	52	84	98	109	113	160	169	226

5.3.3 On-Chip Inductor, Capacitor, and Output Resistance of the Driving Buffer

Since the transmission line network of a resonant H-tree is a passive linear network (assuming the inverter at the leaf node is modeled as a constant gate capacitance), a one-port network, as depicted in Figure 5.8, is used to model the input impedance of the H-tree sector.

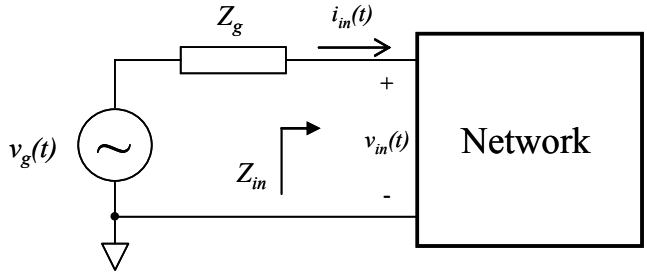


Figure 5.8 One-port network driven by a voltage source

The driving buffer of the resonant clock tree is modeled as a voltage source $V_g(t)$ with a finite output impedance. The output impedance of the voltage source Z_g, and the input impedance of the network Z_{in} can be expressed, respectively, as

$$Z_g = R_g + jX_g \tag{5.5}$$

$$Z_{in} = R_{in} + jX_{in} \tag{5.6}$$

where R_g and R_{in} is the voltage source and input resistance, respectively, and X_g and X_{in} are the voltage source and input reactance, respectively. The rate at which energy is absorbed is the power and is [125]

$$P_{net} = \frac{1}{2} V_{in,rms}^2 \cdot \Re\left\{\frac{1}{Z_{in}}\right\}, \tag{5.7}$$

where $V_{in,rms}$ is the effective or root-mean square value of any periodic voltage function and is

$$V_{in,rms} = \left(\frac{1}{T}\int_{t_0}^{t_0+T} \left|v_{in}(t)\right|^2 dt\right)^{1/2}, \tag{5.8}$$

where T is the time period of the periodic function $v_{in}(t)$. The input voltage $v_{in}(t)$ (see Figure 5.8) can be expressed in terms of the voltage source function $v_g(t)$, and is

$$v_{in}(t) = v_g(t) \cdot \frac{Z_{in}}{Z_{in} + Z_g}. \tag{5.9}$$

Substituting (5.8) and (5.9) into (5.7) results in

$$P_{net} = \frac{1}{2} V_{rms,g}^2 \cdot \left|\frac{Z_{in}}{Z_{in} + Z_g}\right|^2 \cdot \Re\left\{\frac{1}{Z_{in}}\right\} = \frac{1}{2} V_{rms,g}^2 \cdot \rho, \tag{5.10}$$

where ρ is the effective (including R_g in (5.5)) real part of the input admittance of the H-tree sector,

$$\rho = \frac{R_{in}}{\left(R_{in} + R_g\right)^2 + \left(X_{in} + X_g\right)^2}. \tag{5.11}$$

Note from (5.11) that in order to reduce the power consumption P_{net}, ρ should be made smaller. A second constraint is that the magnitude of the transfer function should be equal to or greater than 0.9 at the operating frequency in order to achieve a full swing response at the output. To justify a value of 0.9, consider a Fourier series representation of a periodic square waveform $x(t)$ with an amplitude of V_{DD},

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_0 t}, a_k = V_{DD} \frac{1}{\pi k} \sin\left(k\omega_0 T_1\right), \tag{5.12}$$

where ω_0 is the radian frequency, and T_1 is a quarter of the period of the square wave. The transfer function of the H-tree sector at the resonant frequency is designed to transfer the fundamental harmonic of the square wave, $k = \pm 1$, amplifying the spectral elements,

$$a_{\pm 1} = V_{DD} \frac{1}{\pi}. \tag{5.13}$$

From (5.13), the amplitude transferred to the output equals $2V_{DD}/\pi$. The sinusoidal amplitude at the output ranges from -0.1 volts to 1.9 volts (1 volt mean-to-peak) to allow the buffers at the leaf nodes to charge and discharge the load at frequencies as high as 5 GHz in a 0.18 μm CMOS technology. The required peak value of the magnitude of the transfer function can be derived from $(2V_{DD}/\pi) \cdot |H'(j\omega_0)| = 1$,

$$\left|H'\left(j\omega_0\right)\right| = \frac{\pi}{2V_{DD}} = \frac{\pi}{2 \cdot 1.8} \approx 0.9. \tag{5.14}$$

Since the power consumption is inversely proportional to the output resistance, (5.11) suggests that the output resistance of the driving buffer should be maximized. These design constraints for a resonant H-tree network are summarized.

$$\min\left(P_{net}\right) \tag{5.15}$$

$$\max\left(R_g\right) \tag{5.16}$$

$$\left|H'\left(j\omega_0\right)\right| \geq 0.9. \tag{5.17}$$

In (5.17), $|H'(j\omega)|$ is

$$\begin{aligned} \left|H'(j\omega)\right| &= \left|\frac{Z_{in}}{Z_{in}+Z_g}\right| \cdot \left|H(j\omega)\right| = \left|\frac{R_{in}+jX_{in}}{\left(R_{in}+R_g\right)+j\left(X_{in}+X_g\right)}\right| \cdot \left|H(j\omega)\right| \\ &= \sqrt{\frac{R_{in}^2+X_{in}^2}{\left(R_g+R_{in}\right)^2+\left(X_g+X_{in}\right)^2}} \cdot \left|H(j\omega)\right|, \end{aligned} \tag{5.18}$$

where $|H(j\omega)|$ is described in (5.3) in the S-domain.

The three conditions (5.15), (5.16), and (5.17) can be used to determine the optimal value of the on-chip inductors, capacitors, and driving buffer resistance that produces a full swing sinusoidal waveform while dissipating minimum power. Since closed-form analytic expressions for the input impedance and the transfer function, given by (5.3) and (5.4), are

somewhat cumbersome, the solution is graphically evaluated. In this manner, the design space and related tradeoffs among the different parameters can be explored.

Three design variables, L_s, C_d, and R_g, are solved simultaneously to satisfy (5.15), (5.16), and (5.17). In order to graphically represent the design space, one of the three design variables is eliminated. Equating $|H'(j\omega)|$ to 0.9 and solving for R_g (assuming $X_g = 0$),

$$R_g = \left(\sqrt{\frac{|H(j\omega)|^2}{0.9^2} \cdot \left(R_{in}^2 + X_{in}^2 \right) - X_{in}^2} \right) - R_{in}. \tag{5.19}$$

Substituting (5.19) into (5.11) yields

$$\rho = \frac{0.9^2 \cdot R_{in}}{|H(j\omega)|^2 \cdot \left(R_{in}^2 + X_{in}^2 \right)}. \tag{5.20}$$

Note from (5.20) that ρ is only a function of the on-chip spiral inductor and capacitor at a specific frequency.

Once (5.20) is obtained, ρ and R_g are plotted as a function of the possible on-chip inductor values based on (5.3), (5.19), and (5.20). The inductor value varies from 1 nH to 10 nH, considering the effective series resistance (ESR) for each value (as described in subsection 5.3.2). Two graphs corresponding to the different on-chip capacitor values are obtained. In this way, the entire design space is conveniently represented graphically. Note that the condition (5.16) is already included in (5.19) and (5.20), resulting in only two design graphs. Note also that according to (5.11), maximizing R_g results in minimizing ρ. From these

two graphs, the optimal value for the on-chip inductor, capacitor, and output resistance can be determined.

5.4 Case Study

In this section, a 5 GHz resonant H-tree sector is designed as a basic building block of a large global clock distribution network. The design guidelines and principles presented in section III are demonstrated in this case study. The resistance, inductance, and capacitance per unit length of the transmission lines are extracted using HENRY^{TM} and METAL^{TM} from the OEA software suite [114].

The layout geometry and configuration of a symmetric balanced resonant H-tree sector is schematically illustrated in Figure 5.9. Assuming a 5 GHz clock signal, the design of a resonant H-tree based on a 0.18 μm CMOS technology is described. The H-tree sector including the on-chip inductors and capacitors occupies two metal layers, metal 5 and metal 6, for a total area of 2500 × 2500 μm^2. The horizontal transmission line and capacitors are placed on metal 6, while the vertical transmission lines and spiral inductors are located on metal 5. This strategy reduces the coupling between the lines. In order to further reduce coupling noise, each signal line is shielded by two parallel ground lines. The separation distance s between the ground and signal lines is constant and is 4 μm, while the signal and ground line lengths and widths are the same in each section, *i.e.*, $w_{GND} = w_s = w_i$, where i varies between one and four. The height of each metal line is h = 0.5 μm. These values are typical for a 0.18 μm CMOS technology. The capacitors are connected to ground, while the other end of the spiral inductors is connected to a signal line, namely, to l_4 (the connection is not shown in Figure 5.9).

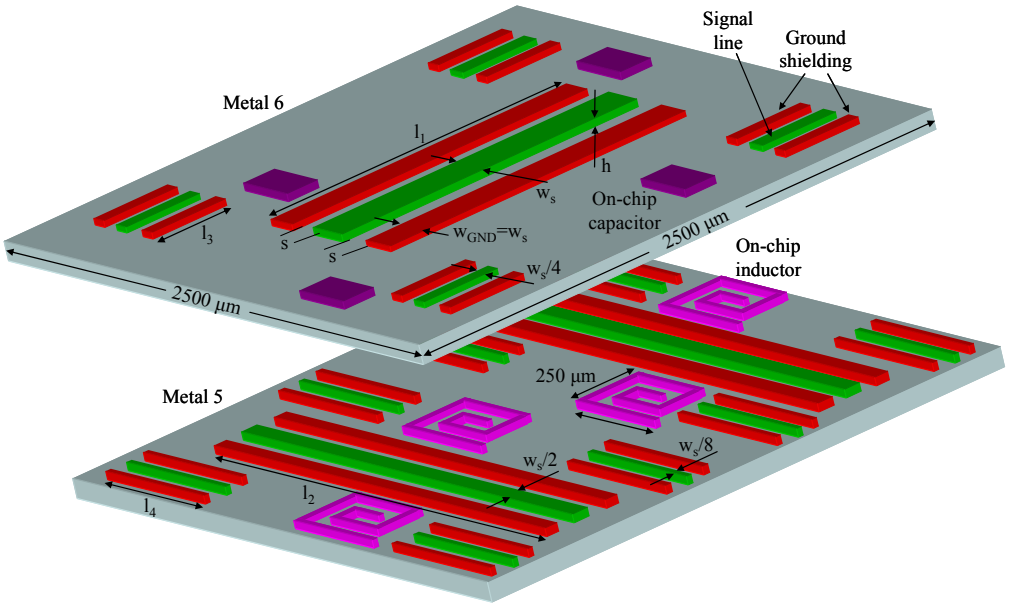


Figure 5.9 Structure of resonant H-tree sector

Finally, note that the width of the signal lines is reduced by two at each branching point in order to reduce reflections caused by differences between the characteristic impedance at the branch points. The individual interconnect widths and lengths, indicated in Figure 5.9 and used in this example, are listed in Table 5.2 .

Table 5.2 Resonant H-tree parameters

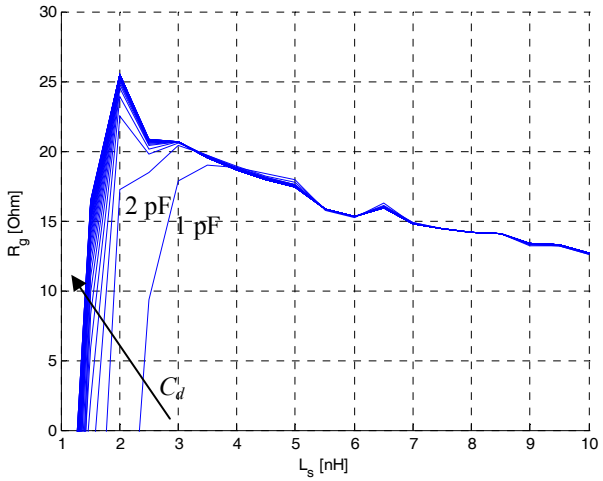
l_1, w_1 [µm]	l_2, w_2 [µm]	l_3, w_3 [µm]	l_4, w_4 [µm]
1600, 16	1600, 8	800, 4	800, 2

Table 5.3 Resonant H-tree extracted transmission line parameters

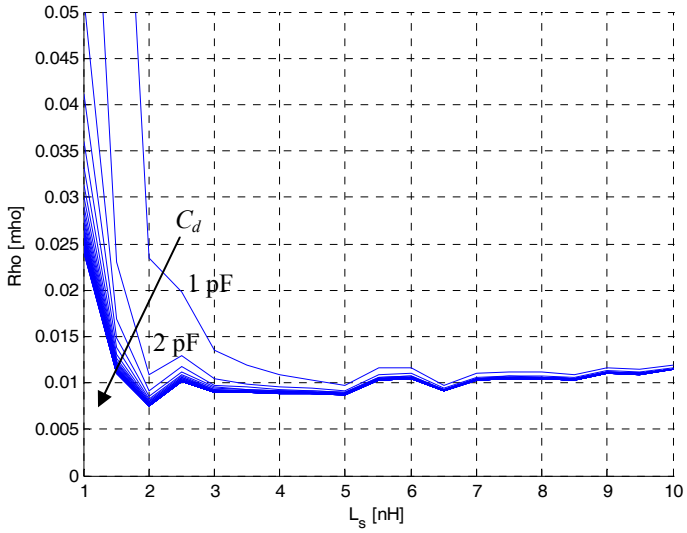
i	R_i [mΩ/µm]	L_i [pH/µm]	C_i [fF/µm]
1	3.19	1.19	0.28
2	6.37	1.43	0.30
3	12.75	1.53	0.13
4	25.50	1.78	0.14

For the technology and geometry described above (see Figure 5.9), the extracted interconnect parameters are listed in Table 5.3. The total load capacitance of the H-tree network is 2 pF. Inverters are located at the leaf nodes of the H-tree sector, each driving a capacitive load of 20 fF. A 20 fF load is chosen at the leaf nodes to satisfy a 5 GHz switching frequency of the buffers. If required, a larger load could be further partitioned into smaller loads. Expressions (5.19) and (5.20) at a 5 GHz operating frequency as a function of the spiral inductance are plotted in Figure 5.10 over a wide range of capacitance values (1 pF through 40 pF).

In order to satisfy condition (5.15), the spiral inductance is chosen to be L_s = 2 nH, thereby minimizing ρ and maximizing R_g, as evident from Figure 5.10. Consequently, the maximum output resistance is $R_g \approx 25\ \Omega$. Evident from Figure 5.10, as the on-chip capacitor increases, the curves converge. To determine the required on-chip capacitance, ρ and R_g are plotted as a function of the capacitance (see Figure 5.11). For the chosen on-chip capacitor C_d = 15 pF, ρ and R_g saturate to a constant value.



(a)



(b)

Figure 5.10 Design tradeoffs for an H-tree sector: (a) Output resistance as a function of the spiral inductor, (b) ρ as a function of the on-chip spiral inductor

The output waveform at the leaf nodes described in the time domain is shown in Figure 5.12. Note that the square clock waveform is distributed to the leaf node, achieving a full rail-to-rail voltage swing. Also note that the output waveform exhibits a quasi-sinusoidal characteristic common in non-resonant multi-gigahertz clock distribution networks [72].

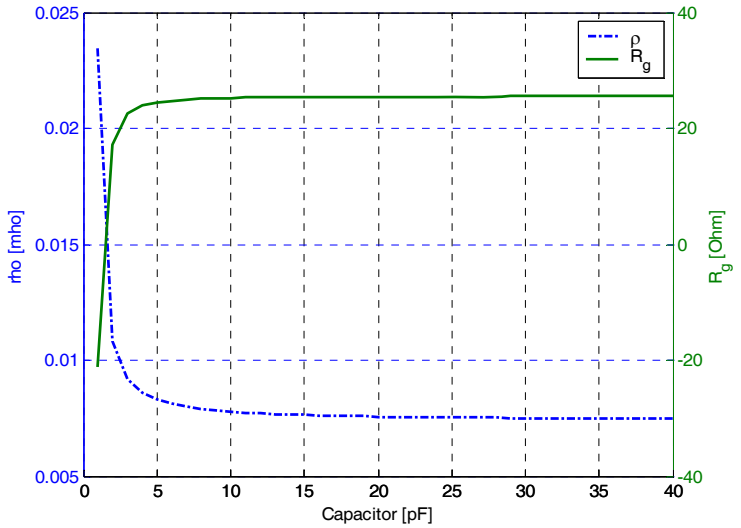


Figure 5.11 Output resistance and ρ as a function of the on-chip capacitance with L_s = 2 nH

In the frequency domain, the magnitude of the transfer function and input impedance around a 5 GHz operating frequency is shown in Figure 5.13. Good agreement between simulation and the proposed analytic expressions is achieved, exhibiting less than 5% error. Note that the magnitude of the transfer function reaches 0.9 at a 5 GHz frequency. As predicted by the design expressions and verified by simulation, the power consumption in this example is $P_{net} \approx$ 15 mW (including buffers) as compared to a nonresonant H-tree sector, where the power consumption is 93 mW (84% less than the non-resonant circuit).

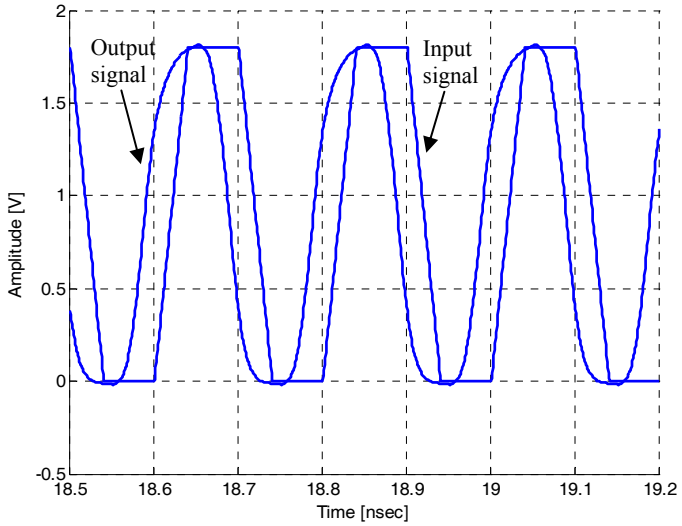


Figure 5.12 Output waveform at the leaf nodes

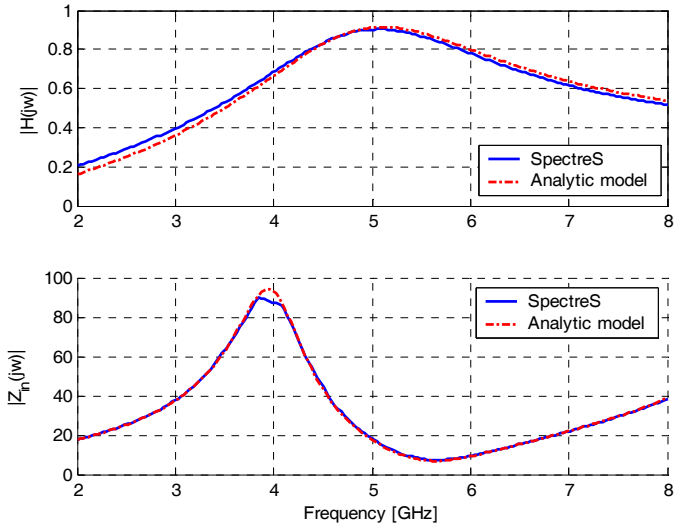


Figure 5.13 Magnitude of the transfer function and input impedance

The size, spacing between the signal and shielding lines, and the driver resistance R_g in the nonresonant network are maintained the same as in the resonant case. In order to drive the

clock signal at a 5 GHz operating frequency, buffers are added at each branching point, as shown in Figure 5.14. The size of these buffers is determined by simulations such that the rise time, 50% duty cycle, and amplitude of the clock signal at the leaf nodes are the same as the resonant network. For each buffer, the ratio of the PMOS and NMOS widths is 2.5, whereas the width of the NMOS transistors is listed in Table 5.4. The channel length of all of the buffers is 0.18 μm.

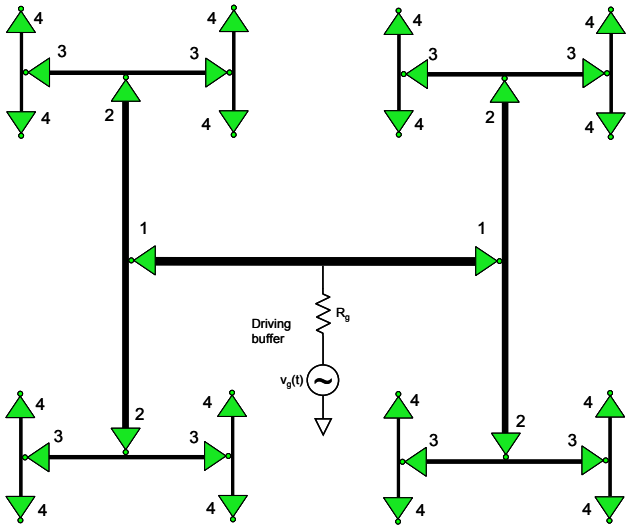


Figure 5.14 H-tree sector with buffers

Table 5.4 Buffer sizes in the non-resonant case

w_1 [μm]	w_2 [μm]	w_3 [μm]	w_4 [μm]
3	8	20	50

Note that by eliminating the need for buffers in a resonant clock network, significant power savings can be achieved. As compared to a nonresonant network, the number of buffers is reduced. The skew and jitter will therefore also be smaller. This behavior is

assumed since fewer devices will suffer process variations as well as coupling noise from the power supply.

The power consumption as a function of the size of the on-chip inductors as expressed by (5.10) is shown in Figure 5.15. Note that the maximum resistance of the output buffer is determined for each value of inductance. Good agreement between simulation and (5.10) is illustrated, exhibiting less than 10% error. As indicated in Figure 5.15, the minimum power consumption for the circuit illustrated in Figure 5.1 is about 5 mW (not including the buffers at the leaf nodes). Characterization of $V_{rms,g}$ in (5.10) is described in Appendix A.2.

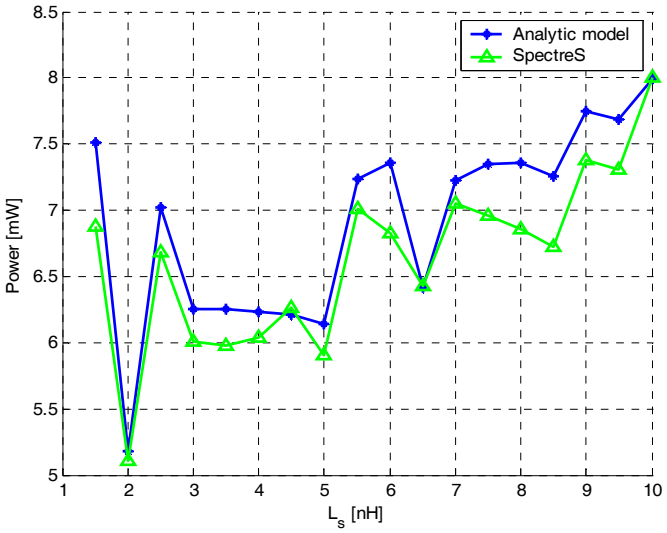


Figure 5.15 Comparison of power consumption between analytic model and SpectreS Spice simulation

To evaluate the performance and accuracy represented by this example, the normalized voltage swing as a function of R_g and the magnitude of the transfer function as a function of the on-chip inductance are shown, respectively, in Figures 5.16 and 5.17. The normalized

voltage swing is linearly decreasing with the output resistance, as shown in Figure 5.16. At an output resistance R_g = 25 Ω, the network distributes a full voltage swing to the leaf nodes. The voltage swing degrades as R_g increases beyond 25 Ω, unable to achieve a full voltage swing. This result agrees with the design constraints described by (5.15), (5.16), and (5.17).

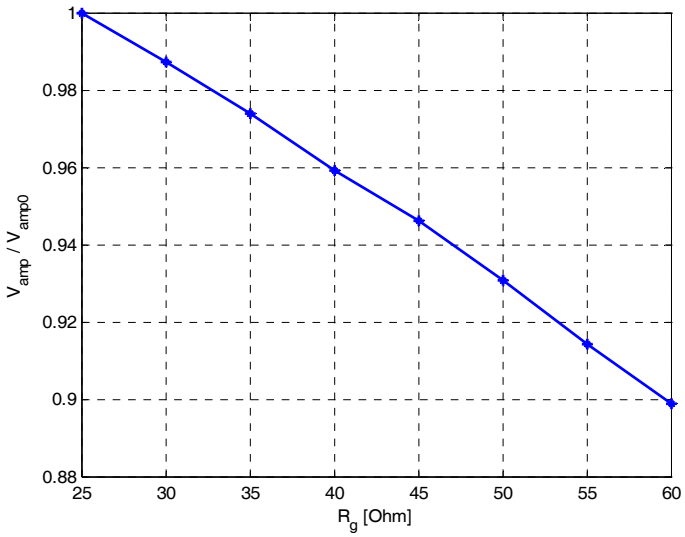


Figure 5.16 Normalized voltage swing at the leaf node

The effect of the on-chip inductor on the magnitude of a transfer function is shown in Figure 5.17. Increasing inductance degrades the magnitude of the transfer function, as indicated in Figure 5.17. The magnitude of the transfer function reaches 0.9 with a 2 nH on-chip inductance, as obtained from Figure 5.10. Note that for other values of inductance, the magnitude of the transfer function is smaller than 0.9. This behavior implies that the H-tree network can only deliver a full swing voltage waveform with an on-chip inductance of L_s = 2 nH. Also note that the model closely agrees with the simulation (exhibiting a maximum error of 1.15%), as shown in Figure 5.17.

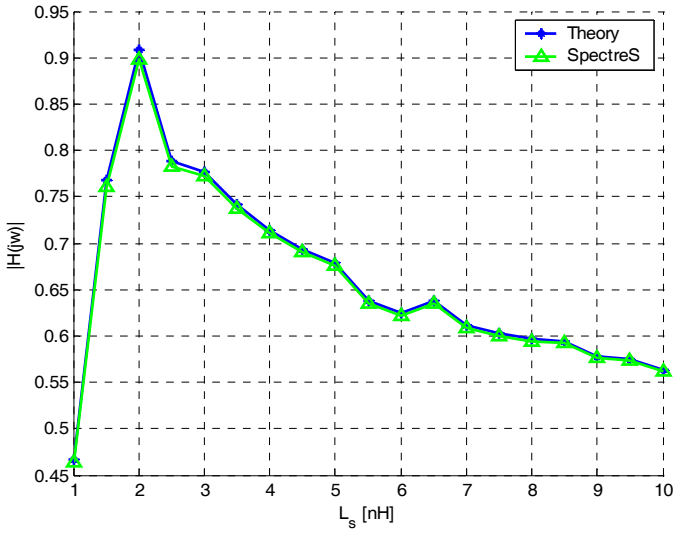


Figure 5.17 Magnitude of the transfer function as a function of the inductance at a 5 GHz operating frequency

5.5 Sensitivity of H-Tree Sector

The effect of process variations on the performance of a resonant H-tree sector is explored in this section. In particular, six types of circuit variations are considered: the driving buffer output resistance, on-chip inductor and capacitor size, and signal and shielding transmission line width and spacing. These variations are examined with respect to two performance related figures of merit: the clock signal voltage swing at the leaf nodes and the power consumption.

The buffer driving the H-tree sector is modeled as a voltage source with an effective output resistance. This simple model, however, does not consider process variations in the

channel length and transistor doping concentration. Both of these effects can change the effective output resistance and thereby the performance.

To evaluate the effect of this variation on the resonant H-tree performance, the output resistance is varied over a range of ±25% of the optimal value (25 Ω), as shown in Figure 5.18. The voltage swing at the leaves and variations in the power dissipation are considered as two performance metrics. The voltage swing is measured with respect to the clock output swing in Figure 5.12 (1.8 volts), while variations in the power consumption are measured with respect to the optimal power consumption of 15 mW.

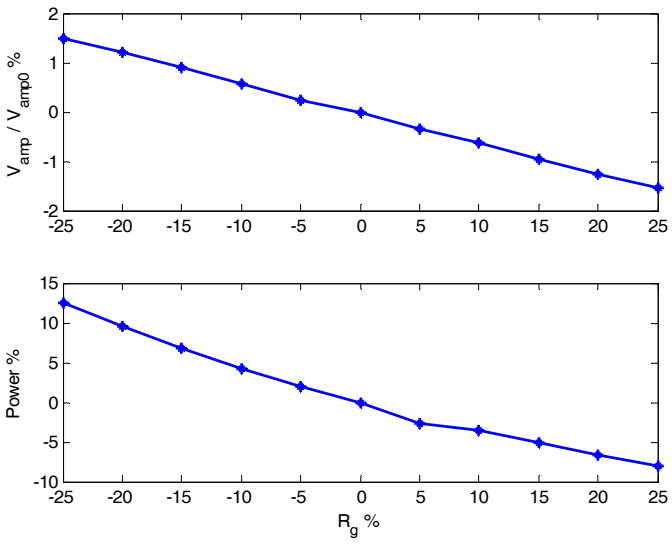


Figure 5.18 Voltage swing and power consumption as a function of variations in the output resistance

When the output resistance increases, the power consumption decreases. This behavior occurs since the power consumption is inversely proportional to the output resistance of the driving buffer. Note that the variation of the voltage swing is ±1.25%, even at the extrema of

the variations in the output resistance. The model can therefore be used to accurately represent the driving buffer.

To examine the effects of variations in the on-chip inductors and capacitors, consider Figures 5.19 and 5.20. Variations in the spiral inductor, assuming a uniform distribution, are shown in Figure 5.19. Process variations may alter the spiral wire width, changing the inductance and therefore the parasitic ESR. The change in inductance may occur due to the dependence of inductance on the specific geometry of the spiral [113].

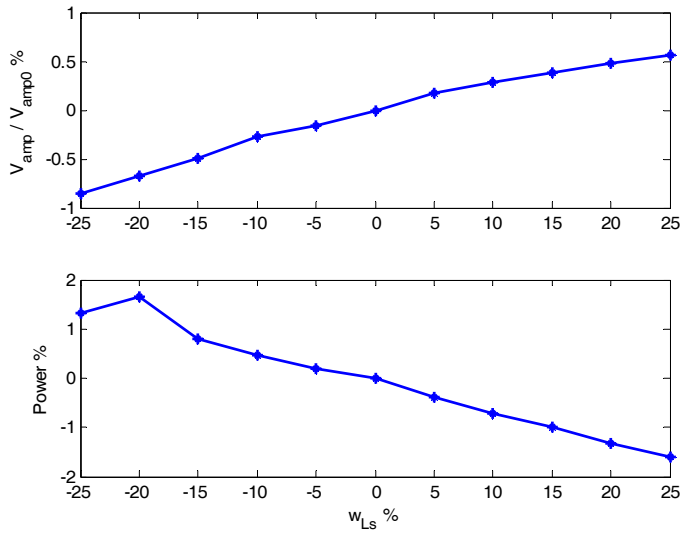


Figure 5.19 Voltage swing and power consumption as a function of variations in the on-chip inductor width

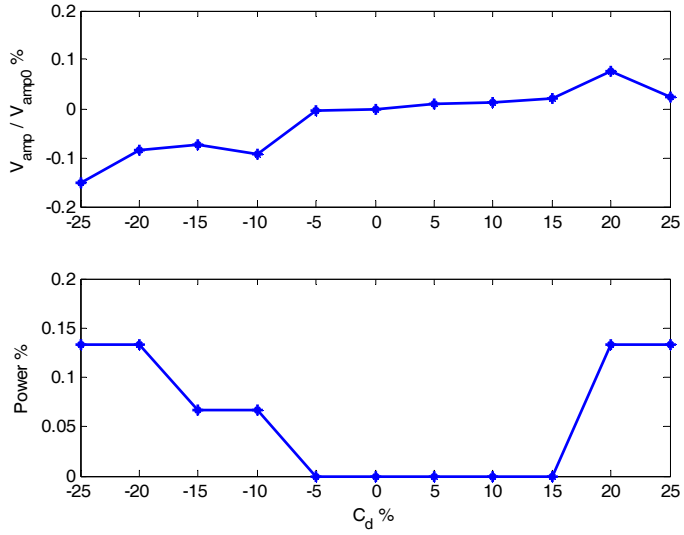


Figure 5.20 Voltage swing and power consumption as a function of variations of the on-chip capacitor

The inductance and ESR, extracted using ASITIC, are a function of the variation in spiral width. For a ±25% change in the wire width, the output voltage swing varies less than ±1%, while the power consumption varies less than ±2%. This behavior indicates that a resonant H-tree is highly tolerant to significant variations in the on-chip spiral inductor and suffers only a minor degradation in performance. Note that the inductance increases as the spiral inductor wire width increases, saving additional energy, as shown in Figure 5.19.

In modern semiconductor fabrication processes, the on-chip capacitor can vary by up to ±20%. To examine the effect of this variation on the behavior of the resonant H-tree, consider Figure 5.20. The on-chip capacitors are swept over ±25% of the optimal value (15 pF). As observed from Figure 5.20, the performance of the resonant H-tree is almost insensitive to these changes since the voltage swing varies by less than ±0.15% while the power consumption varies by less than +0.15%. This behavior occurs since the transfer function of

the entire network is a weak function of the magnitude of the on-chip capacitors. As mentioned previously, the primary purpose of the on-chip capacitors is to establish a DC voltage around which the clock signal oscillates. Another illustration of the insensitivity of a resonant H-tree network to on-chip capacitor variations is illustrated in Figure 5.11. Note in Figure 5.11 that as the on-chip capacitor increases, the characteristic become independent of C_d, hardly affecting the performance. This behavior occurs since at high frequencies and large capacitors, the impedance of C_d decreases, shunting the H-tree structure to ground.

To explore the effect of transmission line variations on the behavior of the H-tree sector, the dependence of three variations are evaluated. Specifically, the width of the signal and shield lines, and the spacing between the signal and shield lines are varied. As described by (5.1), these parameters affect the performance of the resonant H-tree. In the following investigation, it is assumed that the variations are uniformly distributed along the lines.

The impact of variations in the interconnect width (see Figure 5.9) on H-tree performance is illustrated in Figure 5.21. The H-tree interconnect width is varied by ±10%, exhibiting a voltage swing and power variations of 0 to -0.25% and -0.9% to +0.5%, respectively. Note that the resonant H-tree performance is not significantly affected, demonstrating the negligible effect of the interconnect width on a resonant H-tree sector. Note also that since the wire resistance is inversely proportional to the wire cross-sectional area, an increase in the interconnect width decreases the resistance, resulting in lower wire losses, thereby decreasing the power dissipation.

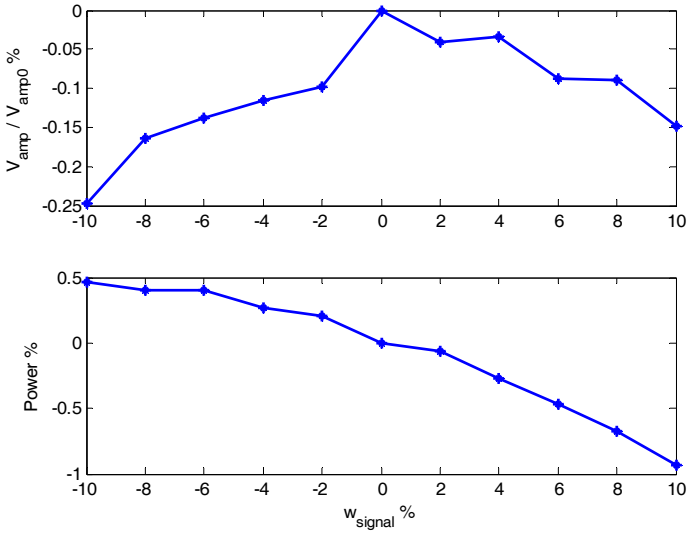


Figure 5.21 Voltage swing and power consumption as a function of variations in the signal line width

The impact of interconnect width variations on the shield lines is shown in Figure 5.22. Similar to the signal line variations, the H-tree exhibits a voltage swing and power variations of about 0.1% and -0.6% to +0.8%, respectively. Hence, the performance of the resonant H-tree network is preserved under these variations. From a power consumption perspective, note that as the shield line width increases, the power dissipation decreases. This behavior occurs since the total equivalent resistance of the structure is increasing with wider shield lines, resulting in lower power consumption. The increase and decrease in the equivalent capacitance and inductance, respectively, has an insignificant effect on the total power dissipation.

The space between the shield line and the signal lines determines the magnitude of the coupling capacitance, the mutual inductance, and the noise coupled to the signal line (see

Figure 5.9). With spacing variations, the H-tree network exhibits a voltage swing and power variations of about -0.1% to +0.03% and -0.08% to +0.2 %, respectively.

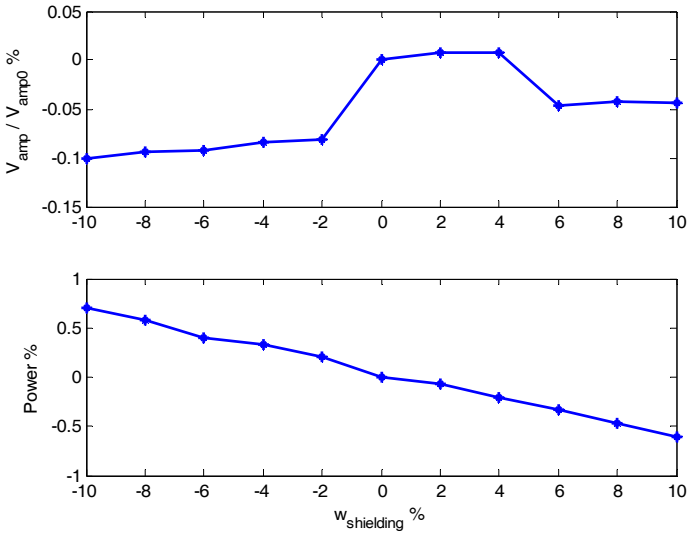


Figure 5.22 Voltage swing and power consumption as a function of variations in the shield line width

The spacing between the signal and shield lines does not significantly affect the performance of a resonant clock network. The shield lines around the H-tree interconnects should therefore be used to lower noise rather than increase the speed of the resonant H-tree network. Note also that as the spacing between the signal and shield lines increases, the capacitance of the structure decreases, resulting in increased power consumption, as depicted in Figure 5.23. The different design criteria and variation characteristics presented in this subsection are summarized in Table 5.5.

As predicted by theory and verified by OEA extraction tools, HENRY^{TM} and METAL^{TM}, and SpectreS simulations, the following trends are observed. As the width of the

signal or shield lines increases, the resistance and inductance decreases while the capacitance increases. As the separation between the signal and shield lines increases, the resistance remains constant while the inductance increases and the capacitance decreases.

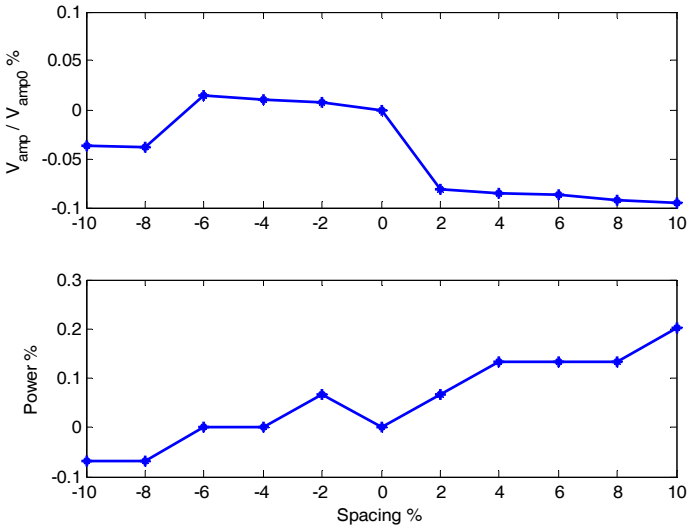


Figure 5.23 Voltage swing and power consumption as a function of spacing variations between the signal and shield lines

Table 5.5 Sensitivity evaluation of resonant H-tree network

Category	Variation range	Voltage swing variations	Power consumption variations
R_g	±25%	-1.5 to 1.5%	-8 to 13%
W_{Ls}	±25%	-0.8 to 0.6%	-1.5 to 1.9%
C_d	±25%	-0.15 to 0.1%	0 to 0.14%
W_{signal}	±10%	-0.25 to 0%	-0.9 to 0.5%
$W_{shielding}$	±10%	-0.1 to 0%	-0.6 to 0.7%
Spacing	±10%	-0.1 to 0.01%	-0.08 to 0.2%

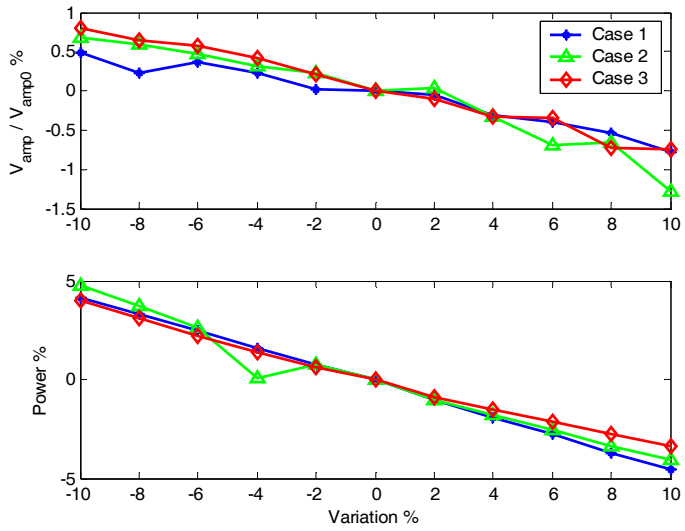


Figure 5.24 Three cases of voltage swing and power consumption as a function of four simultaneous variations: Case 1 - buffer output resistance, spiral inductors, capacitors, and signal line width are varied; Case 2 - buffer output resistance, spiral inductors, capacitors, and shield line width are varied; Case 3 - buffer output resistance, spiral inductors, capacitors, and spacing are varied

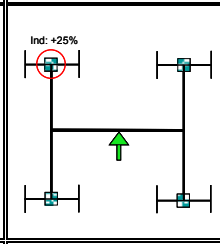
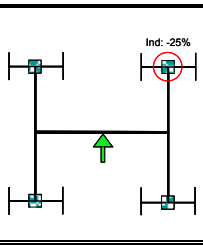
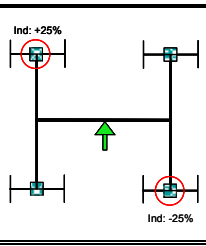
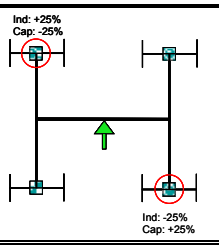
In order to explore the effects of multiple simultaneous variations on the performance of a resonant H-tree network, three simulation cases are considered, as shown in Figure 5.24. The buffer output resistance, on-chip spiral inductors, and capacitors are varied simultaneously with the signal width (case 1), shield line width (case 2), or spacing (case 3).

The H-tree network exhibits the maximum voltage swing and power variations of about ±1.3% and ±5%, respectively (see Figure 5.24) for all three cases. As expected, the dominant source of deviation (as illustrated in Figure 5.18) from the optimal behavior of the resonant H-tree sector is due to variations in the driving buffer output resistance. The other parameters have a much smaller effect on the performance of the H-tree network. These results confirm

that the nature of a resonant H-tree network is robust and relatively tolerant to process variations.

To explore the response of a resonant H-tree network to unbalanced on-chip inductor and capacitor variations, four cases are considered, as shown in Table 5.6. In the first case, only the upper left inductor width is increased by 25%. In the second case, only the width of the upper right inductor is decreased by 25%. In the third case, the left upper and right lower inductors widths are varied by 25% and -25%, respectively. Finally, in the fourth case, the left upper inductor width and capacitor are varied by 25% and -25%, respectively, while the lower right inductor width and capacitor are varied by -25% and 25%, respectively. The range of voltage swing and power consumption variations is 0.13% to 0.84% and -1.2% to -0.4%, respectively. These results demonstrate that, at least for this example, the H-tree sector is almost insensitive to highly unbalanced on-chip inductor and capacitor variations.

Table 5.6 Four cases of unbalanced inductor and capacitor variations

	Case 1	Case 2	Case 3	Case 4
				
Voltage swing	0.84%	0.13%	0.98%	0.76%
Power consumption	-1.2%	-0.4%	-0.81%	-0.79%

5.6 Summary

Resonant H-tree clock distribution networks may be a suitable alternative to traditional clock distribution networks. By exploiting the resonance behavior, a significant decrease in power consumption can be achieved as compared to standard H-tree networks. A methodology is described for designing resonant H-tree clock distribution sectors. These H-tree structures form the basic building block of a resonant network. An accurate model is developed which utilizes a transmission line model to characterize high frequency effects. The high accuracy and analytic nature of the model enables the exploration of tradeoffs in the design of a resonant H-tree sector. The optimal on-chip inductors and capacitors as well as the maximum driving buffer output resistance are determined for a specific example circuit. This set of impedances produces the minimum power at the target clock frequency.

A case study demonstrating the proposed design methodology is described. A 5 GHz H-tree network exhibits significantly improved performance in terms of power consumption and voltage swing. A comparison to a non-resonant H-tree sector is also provided, exhibiting 84% lower power consumption. The major overhead as compared to a nonresonant network is the area. This overhead is due to the four inductors connected to the network. Minimizing the area occupied by these inductors is a topic for future research.

The sensitivity of a resonant H-tree sector to six design criteria is further explored. These design criteria include the buffer output resistance, on-chip inductor and capacitor size, and the signal and shield line width and spacing. Simulations demonstrate that a resonant H-tree sector exhibits acceptable robustness and relatively low sensitivity to process and environmental variations for a 5 GHz operating frequency. The maximum voltage swing and power dissipation are +1.5% and +13%, respectively. These variations occur when the

driving buffer output resistance varies by ±25%. For the remaining parameters, the resonant H-tree sector exhibits insignificant variations as compared to the optimal performance with no variations. Simultaneous process variations demonstrate that resonant clock networks are highly tolerant and robust. Finally, unbalanced inductor width and capacitor variations illustrate the insensitivity of a resonant H-tree clock distribution network to process variations.

Chapter 6

Quasi-Resonant Interconnects: A Low Power, Low Latency Design Methodology

Design and analysis guidelines for quasi-resonant interconnect networks (QRN) are presented in this chapter. The methodology focuses on developing an accurate analytic distributed model of the on-chip interconnect and inductor to obtain both low power and low latency. Excellent agreement is shown between the proposed model and SpectraS simulations. The analysis and design of the inductor, insertion point, and driver resistance for minimum power-delay product is described. A case study demonstrates the design of a quasi-resonant interconnect, transmitting a 5 Gbps data signal along a 5 mm line in a TSMC 0.18 μm CMOS technology. As compared to classical repeater insertion, an average reduction of 91.1% and 37.8% is obtained in power consumption and delay, respectively. As compared to optical links, a reduction of 97.1% and 35.6% is observed in power consumption and delay, respectively.

This chapter is organized into seven sections. An overview on low power data transmission techniques is presented in section 6.1. On-chip resonance in integrated circuits as well as the principle of resonance for data transmission are presented in section 6.2. The interconnect and spiral inductor models are described in section 6.3. In section 6.4, a quasi-resonant interconnect design methodology is described, followed by a case study demonstrating 5 Gbps data signal distribution along a 5 mm interconnect in section 6.5. Simulation results and a comparison to other schemes are presented in section 6.6. Finally, a

summary is provided in section 6.7. In Appendix B.1, the *ABCD* matrices are defined, and a derivation of the 50% delay is described in Appendix B.2.

6.1 Introduction

To reduce significant power dissipation, increased delay, and large area of data distribution networks, low power high speed signaling techniques have been developed. In [126], current mode signaling in an ultra-low voltage environment is used to transmit high data rate signals. In this approach, a current sense amplifier at the far end of the line detects a current difference and converts it into a voltage difference. To improve both delay and energy dissipation, a transmitter generating a differential current detected by a current mode sense amplifier at the receiving end has been proposed in [65]. However, to accommodate differential operation, redundant circuitry is used.

To minimize static power dissipation associated with current mode signaling, an adaptive bandwidth bus architecture based on hybrid current voltage mode repeaters for long *RC* interconnect has been proposed in [127]. In this approach, the interconnect is divided into smaller segments. Appropriately spaced repeaters amplify the signal and drive the subsequent interconnect segment. An adaptive control unit is placed at the input of the line and connected to a control line, which is common to all the repeater stages. Contrary to the current mode signaling approach, the authors of [128], [129] suggest the use of low voltage signaling over long on-chip interconnects. In [128], a heuristic algorithm for buffer insertion that considers noise, delay, and power is proposed. The method proposed in [129] is based on the so-called swing limited interconnect accelerator. This circuit has a three stage cascade inverter configuration with keepers. The keepers limit the voltage on the interconnect to allow low

swing operation. Additional inverters are placed at the output to restore the signal level to full swing.

A different approach exploiting wire inductance at high frequencies is introduced in [81]. The authors of [81] suggest that at lower frequencies the resistive part of the interconnect is dominant, behaving as a distributed *RC* network. At higher frequencies, the inductive component of the transmission line dominates, behaving as a distributed *LC* network. In this scheme, a 1 GHz link operating with phase shift keying modulation on a 7.5 GHz sinusoidal carrier is presented. This type of modulation, however, results in relatively large power dissipation and poor spectral efficiency. Alternatively, the authors of [84] propose to mitigate dispersion related to high frequencies by utilizing a return to zero (RZ) signaling scheme in which sharp current pulses are used to transmit data. The transmitted data is modulated to higher frequencies, maximizing the effect of the wire inductance. This approach allows the interconnects to behave in a relatively dispersionless fashion.

As an alternative to electrical interconnects and related electrical limitations, optoelectronic links [130] have been considered. Optical interconnections promise to achieve high bandwidth and high density parallel communication channels. However, the interface between the electrical and optical signals is a major issue in optoelectronic applications.

In this chapter, a low power, low latency on-chip interconnect design methodology is proposed. The methodology is based on inserting an on-chip spiral inductor to resonate the interconnect around the fundamental harmonic of the transmitted signal. In this way, the interconnect capacitance resonates with the inserted on-chip inductance. The fundamental harmonic of the input signal is amplified and transmitted to the output. This approach lowers power consumption, since the energy resonates between the electric and magnetic fields

rather than dissipates as heat. Consequently, buffers are eliminated, significantly reducing power consumption and signal latency.

6.2 Principle of Quasi-Resonant Interconnect

Following the invention of the telegraph and the telephone, transmission techniques were developed to accommodate the increasing distance between the transmitter and receiver while supporting a higher transmission rate. The concept of inserting reactance sources in series with long telegraphic lines was first introduced by Pupin in 1899 [131], [132]. Pupin suggested inserting inductance coils in a network composed of a uniform conductor with a telephonic transmitter and receiver. The inductance coils are inserted in series at periodically recurring locations. The electrical line parameters and the distance among the coils are adjusted to reduce signal attenuation, enabling transmission of speech waves over very long distances. Based on Pupin's invention, the application of transmission lines with periodic inductors to microwave applications is described in [133].

The concept of exploiting standing waves in transmission lines was first introduced by Chi in 1994 [116]. A global resonant clock distribution network was later introduced in 2003 by Chan *et al.* [72]. In this circuit, a set of discrete on-chip spiral inductors and capacitors is attached to a traditional H-tree structure. The capacitance of the clock distribution network resonates with the inductance, while the on-chip capacitors establish a mid-rail DC voltage around which the grid oscillates. This approach lowers the power consumption, skew, and jitter.

The authors of [134] designed and evaluated a two-phase resonant clock generation and distribution system with layout-extracted inductor parameters in a 0.13 μm CMOS process.

The circuit includes a programmable replenishing clock generator and tunable capacitors. Skew and jitter are shown to be reduced with this approach.

A design methodology for resonant clock distribution networks is presented in [118]. Design guidelines based on an accurate distributed transmission model are proposed, supporting low power H-tree clock distribution networks. The design methodology provides tradeoffs among the operating frequency, the size of the on-chip inductors and capacitors, and the output resistance of the driving buffer. A sensitivity analysis of resonant H-tree clock distribution networks is also provided.

The proposed quasi-resonant interconnect network architecture is illustrated in Figure 6.1. The transmitter at the near end of the interconnect modulates the input data signal. The signal modulation supports a quasi-periodic signal with a single resonant frequency. The transmitter is followed by an inverter driving the interconnect. The *RLC* distributed transmission line is separated by an on-chip spiral inductor L_s inserted at a specific point to resonate the network at a desired frequency for minimum power consumption and delay. The receiver at the far end of the interconnect demodulates the transmitted signal back to the original input bit stream.

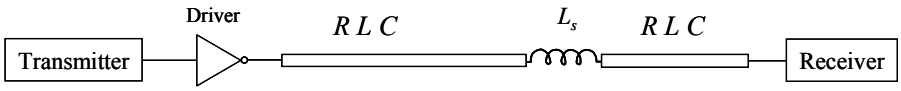


Figure 6.1 Quasi-resonant network

Since the fundamental harmonic of the input signal is amplified by the magnitude of the transfer function, the network resonates at a specific target frequency and magnitude. This behavior is required to transfer a full swing sinusoidal signal at the resonance frequency to

the far end of the interconnect. To satisfy this objective, consider the output signal in the frequency domain,

$$V_{out}(s) = H'(s) \cdot V_{in}(s), \tag{6.1}$$

where $H'(s)$ is the transfer function of the network between the transmitter and the receiver (including the driver), and $V_{in}(s)$ is the input data bit. The input data signal in the time domain can be represented by a Fourier series, assuming $V_{in}(s)$ is a periodic signal,

$$v_{in}(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_p t}, \tag{6.2}$$

where a_k is the k^{th} harmonic of the signal and ω_p is the resonant radian frequency. In the case of a periodic square waveform with period time t_p, rise and fall time t_r, and amplitude V_{dd}, the fundamental harmonic (positive and negative) is

$$a_{\pm 1} = \pm \frac{V_{dd} t_p}{2 t_r \pi^2} \left(e^{-j\omega_p t_r} - 1 \right). \tag{6.3}$$

Equation (6.1) implies that the required magnitude of the transfer function is

$$\left| H'\left(j\omega_p \right) \right| = \left(V_{dd} / 2 \right) / \left(2 \left| a_1 \right| \right) = \frac{V_{dd}}{4 \left| a_1 \right|}. \tag{6.4}$$

Substituting (6.3) into (6.4) results in

$$\left|H'\left(j\omega_p\right)\right| = \frac{t_r\pi^2}{t_p\sqrt{8\left(1-\cos\left(\omega_p t_r\right)\right)}}. \tag{6.5}$$

Equation (6.5) describes the magnitude of the transfer function at ω_p in order to transfer a full swing waveform. For example, at a 5 GHz operating frequency (t_p = 200 psec) and t_r = 20 ps, the magnitude of the transfer function is 0.8.

6.3 Interconnect and Spiral Inductor Models

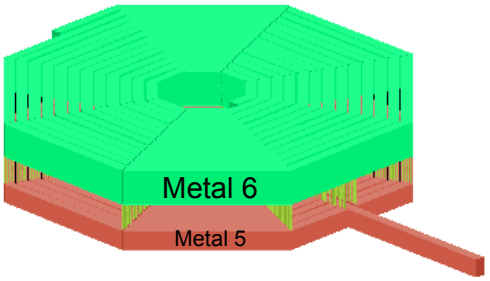
Accurate models of the quasi-resonant network are presented in this section. Transmission line and on-chip spiral inductor models are described in subsections 6.3.1 and 6.3.2, respectively.

6.3.1 Interconnect Model

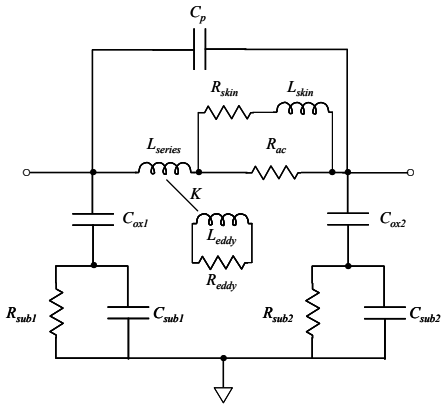
As data signal frequencies exceed the multigigahertz regime, distributed models of interconnects are required to accurately incorporate high frequency effects into the system behavior. The wire inductance can no longer be excluded from the model and traveling wave reflections should be characterized by a distributed structure. Using the classical distributed model, an incremental section of line length Δz is modeled as a lumped element circuit. In this representation, R, L, and C are the resistance, inductance, and capacitance per unit length, respectively. The lumped resistance represents the lossy component of the transmission line.

6.3.2 On-Chip Spiral Inductor Model

To accurately account for the parasitic effects of the on-chip spiral inductor, a thirteen lumped element model is used. The physical structure and a lumped model of the on-chip inductor are illustrated in Figures 6.2(a) and 6.2(b), respectively. Note that the octagonal on-chip spiral inductor shown in (a) utilizes Metal 5 and Metal 6 where the two layers are connected by vias. In this manner, the inductor can achieve a higher Q, as analyzed by SPIRAL™, a 3D spiral inductor design and synthesis tool [114].



(a)



(b)

Figure 6.2 Model of an on-chip spiral inductor: (a) structure of an octagonal on-chip spiral inductor, (b) lumped model of the spiral inductor

The capacitance C_p represents the capacitive coupling between the windings of the spiral inductor. The elements L_{series} and R_{ac} represent the inductance and parasitic resistance, respectively, while R_{skin} and L_{skin} model the skin effect. Also note that L_{series} incorporates the eddy current effect coupled to the inductor by the coefficient K. The parasitic capacitance between the lines and the substrate is modeled by C_{ox}. The parallel C_{sub} and R_{sub} combination models the parasitic resistance and capacitance to the substrate.

6.4 Design Methodology

In this section, the quasi-resonant interconnect design methodology is described. The design of the transmitter and receiver is presented in subsection 6.4.1 and 6.4.2, respectively. The input impedance and transfer function of the quasi-resonant network are analytically characterized in subsection 6.4.3. Based on these expressions, a power consumption model followed by a closed-form analytic expression is developed in subsection 6.4.4. In subsection 6.4.5, closed-form analytic expressions for the signal delay are described. Finally, design guidelines for the quasi-resonant interconnect that minimizes the power-delay product is described in subsection 6.4.6.

6.4.1 Transmitter Design

Proper operation of the network requires resonance at a single target frequency. As illustrated in Figure 6.3, the return to zero (RZ) amplitude shift keying modulation scheme is

chosen to support a single transmission frequency of the input data. In this scheme, the data is transferred at $1/t_p$ bits per second.

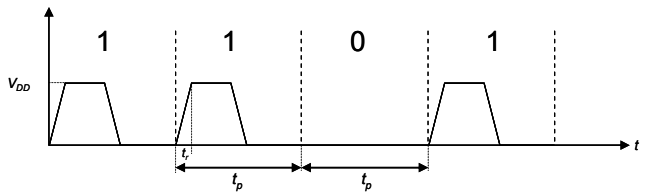


Figure 6.3 Example of transmitting a "1011" bit stream

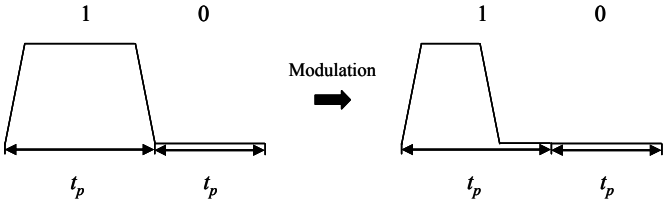
This scheme has three advantages when applied to the proposed quasi-resonant interconnect methodology. The transmitted signal has a single frequency (with amplitude variation, *i.e.*, V_{dd} and 0 volts to distinguish between logic one and zero, respectively) designed to match the resonance frequency of the network. Power is dissipated only during transmission of logic one, and only during half of the time period. Finally, no complex circuitry is required to produce this modulation scheme.

The transmitter generates the required modulated signal which is enhanced by the buffer chain that drives the driver (see Figure 6.1). The proposed modulation scheme, transmitter circuit, transistor level circuit, and simulated signal waveforms are shown in Figure 6.4.

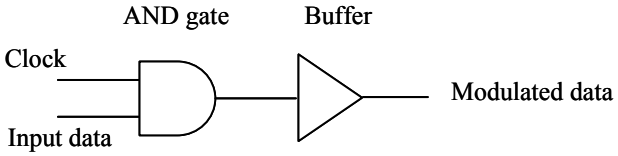
To realize the modulation scheme shown in Figure 6.4(a), an AND operation is performed between the clock and the input data followed by a buffer chain as shown in Figure 6.4(b). The high frequency data rate supported by the quasi-resonant interconnects requires a high speed transmitter and receiver. To support high speed operation, a transmission gate-based circuit (consisting of transistors M_{t1}, M_{t2}, M_{t3}, and M_{t4}) is used for the logic as shown in Figure 6.4(c). In the case where both the clock and data are at logic one, the transmission gate M_{t1} passes the logic one state to the input of the buffer chain, while the

inverter (consisting of transistors M_{t3}, and M_{t4}) turns off transmission gate M_{t2}. In all other cases, the logic zero state is passed to the input of the buffer chain.

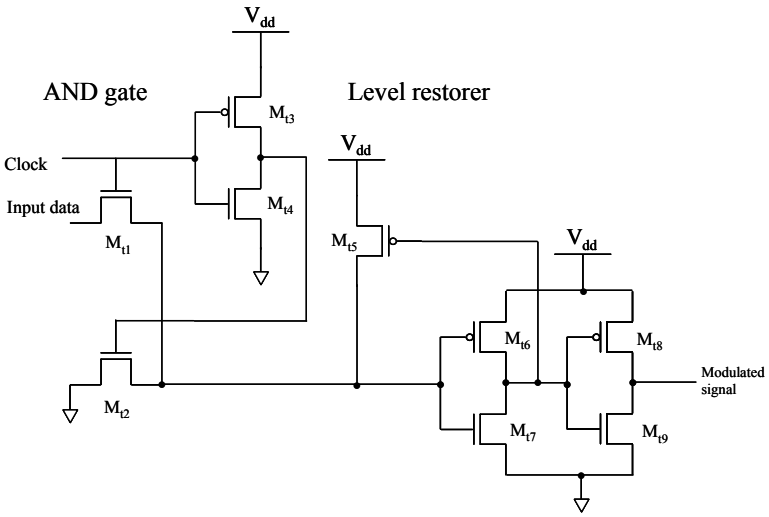
Transistor M_{t5} restores the voltage (equal to V_{th}) associated with the operation of transmission gates M_{t1} and M_{t2}. The input data as well as the modulated signal are shown in Figure 6.4(d). In this example, a "100100111" bit stream is modulated at a 5 GHz operating frequency. Note that the modulated signal follows the return to zero amplitude shift keying scheme, essential for quasi-resonant operation.



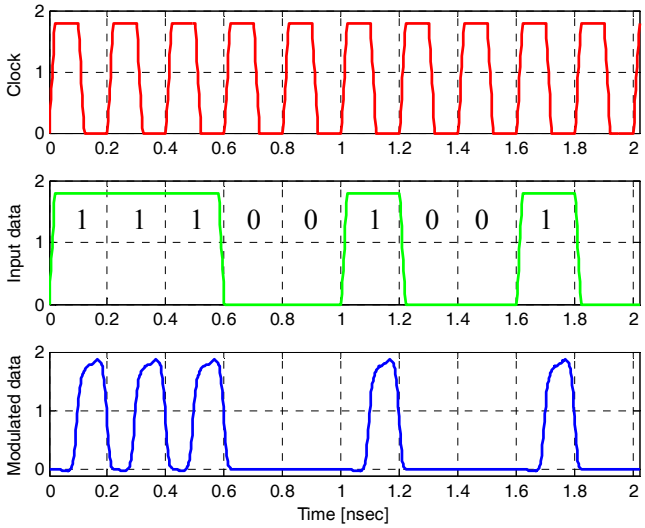
(a)



(b)



(c)

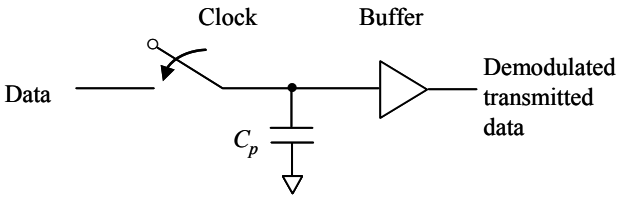


(d)

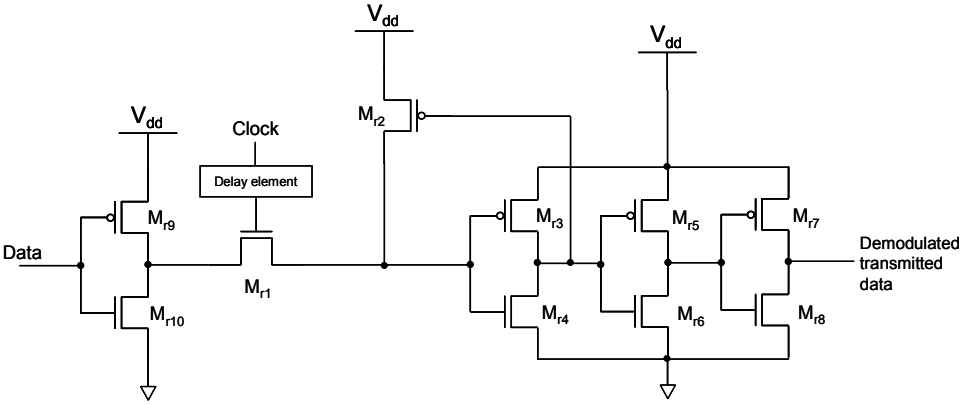
Figure 6.4 Transmitter circuit: (a) modulation scheme, (b) gate level circuit, (c) transistor level circuit, (d) signal waveforms

6.4.2 Receiver Design

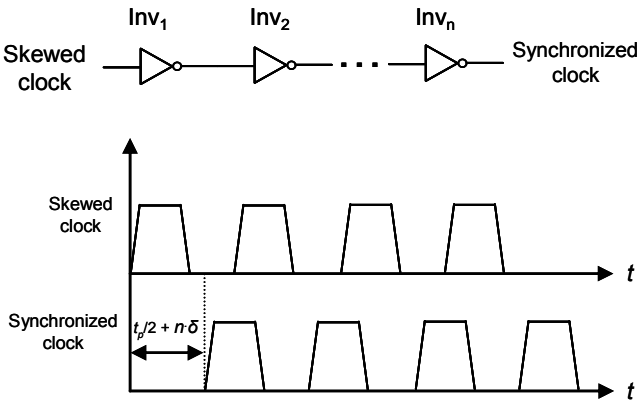
The receiver located at the far end of the interconnect (see Figure 6.1) demodulates the transmitted data into the original signal. A logic level description of the proposed receiver circuit, a transistor level circuit, and simulated signal waveforms are depicted in Figure 6.5. The principle of the demodulation scheme is based on a sample and hold (S/H) operation, as shown in Figure 6.5(a). When the clock is high, the switch closes and the data is sampled and transferred. The sampled signal charges (and discharges) the parasitic capacitance C_p. When the clock is low, the switch is open and the logic state is stored (or held) until the following clock cycle. In this manner, the S/H circuit prolongs the duration of the high state signals over the entire clock cycle t_p without altering the duration of the low state signals.


Clock
Buffer
Data
Demodulated
transmitted
data
C_p

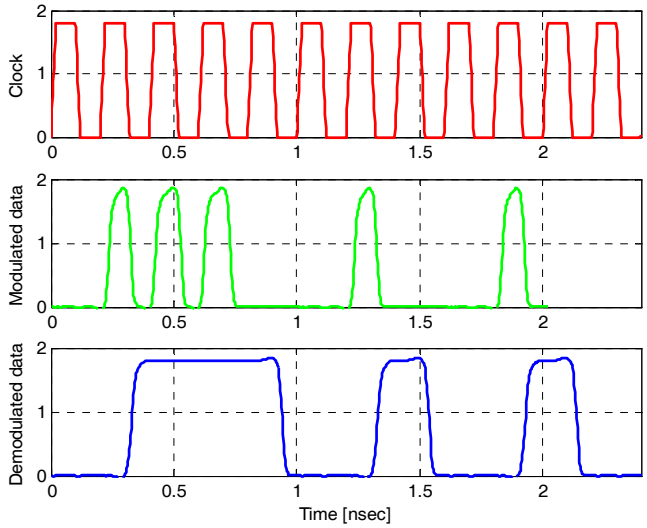
(a)


V_{dd}
V_{dd}
V_{dd}
Clock
Delay element
M_{r2}
M_{r9}
M_{r3}
M_{r5}
M_{r7}
Data
M_{r1}
Demodulated
transmitted
data
M_{r10}
M_{r4}
M_{r6}
M_{r8}

(b)


Inv_1
Inv_2
Inv_n
Skewed
clock
Synchronized
clock
Skewed
clock
t
Synchronized
clock
$t_p/2 + n\cdot\delta$
t

(c)



(d)

Figure 6.5 Receiver circuit: (a) sample and hold circuit, (b) transistor level circuit, (c) delay element circuit, (d) signal waveforms

The receiver circuit at the transistor level is shown in Figure 6.5(b). In this circuit, when the clock state is high, transistor M_{r1} turns on and the data is sampled and transferred. When the data state is low, the restorer transistor M_{r2} turns on, maintaining the high state signal at the input of the second inverter (consisting of transistor M_{r3} and M_{r4}). This mechanism serves a dual purpose. It restores the voltage associated with the transmission gate M_{r1}. And it prevents charge leakage by replenishing the charge on the parasitic capacitance of transistors M_{r3} and M_{r4} (through the feedback connection). When the clock logic is low, M_{r2} prevents the charge from leaking. When the data logic is high, transistor M_{r2} turns off and the logic low state is transferred to the output. Finally, when the clock state is low, the previous logic state is preserved until the following clock cycle.

Note that for proper operation of the receiver, the allowed skew between the clock and the incoming modulated data should be less then one quarter of the clock cycle. This constraint is required since the demodulation circuit is level sensitive. To synchronize the clock with the data, a delay element is used, as depicted in Figure 6.5(b). The delay of the arriving modulated data is obtained either by simulation or the analytic delay expressions presented in subsection 6.4.5. Accordingly, a delay element based on inverters is used as shown in Figure 6.5(c). The delay consists of a course and a fine delay. If course tuning is required (*i.e.*, the delay of a half clock cycle, $t_p/2$), an odd number of inverters should be used. If fine tuning is required (*i.e.*, the intrinsic delay of the inverters), an even number of inverters is required. In this manner, for an odd number of inverters, a total delay of $t_p/2 + n \cdot \delta$ is achieved. For an even number of inverters, a total delay of $n \cdot \delta$ is achieved. δ and n are the intrinsic delay of a single inverter and number of inverters, respectively. For example, using one inverter produces a delay of $t_p/2 + \delta$, while two inverters produce a delay of $2 \cdot \delta$.

Simulation results of this circuit are presented in Figure 6.5(d). In this example, the output of the modulation circuit shown in Figure 6.4(c) drives the demodulation circuit shown in Figure 6.5(b). As depicted in Figure 6.5(d), the output waveform is identical to the input data waveforms illustrated in Figure 6.4(d).

6.4.3 Network Input Impedance and Transfer Function

The quasi-resonant network (between the transmitter and receiver) including the driver is shown in Figure 6.6. Note that the driver is modeled as a linearized voltage source V_d serially connected with a driver resistance R_d. The load of the interconnect is modeled as a capacitor C_l.

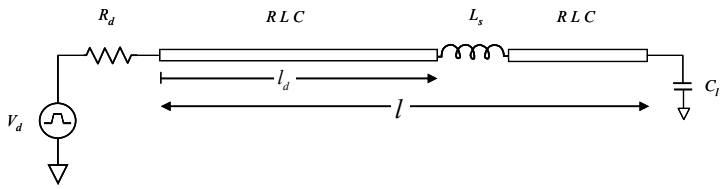


Figure 6.6 Quasi-resonant network

To analyze this type of structure, an accurate analytic model is developed based on *ABCD* parameters. From transmission line theory, the *ABCD* matrix for the entire structure is a product of the individual matrices,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = M_{\text{int}1} \cdot M_s \cdot M_{\text{int}2} \cdot M_l, \tag{6.6}$$

where M_{int1}, M_s, M_{int2}, and M_l are the *ABCD* matrix of the first interconnect section, the on-chip inductor (based on the model of Figure 6.2(b)), the second interconnect section, and the load capacitance, respectively. These matrices are described in Appendix B.1. Matrices M_{int1} and M_{int2} are based on the distributed transmission line model presented in subsection III-A. From the overall *ABCD* parameters, the transfer function $H(j\omega)$ (excluding the driver resistance) and the input impedance $Z_{in}(j\omega)$ of the system is, respectively,

$$H(j\omega) = \frac{1}{A}, \tag{6.7}$$

$$Z_{in}(j\omega) = R_{in} + jX_{in} = \frac{A}{C}. \tag{6.8}$$

The transfer function of the overall network shown in Figure 6.6 (including the driver resistance) is given by (6.9), where R_d and R_{in} are the driver and input resistance, respectively, and X_{in} is the input reactance.

$$\begin{aligned} \left|H'(j\omega)\right| &= \left|\frac{Z_{in}}{Z_{in}+R_d}\right| \cdot \left|H(j\omega)\right| = \left|\frac{R_{in}+jX_{in}}{\left(R_{in}+R_d\right)+jX_{in}}\right| \cdot \left|H(j\omega)\right| \\ &= \sqrt{\frac{R_{in}^2+X_{in}^2}{\left(R_d+R_{in}\right)^2+X_{in}^2}} \cdot \left|H(j\omega)\right|. \end{aligned} \tag{6.9}$$

The driver resistance can be extracted from (6.9). Equating $|H'(j\omega)|$ to the value obtained from (6.5) at a specific frequency ω_p and solving for R_d results in

$$R_d = \left(\sqrt{\frac{\left|H(j\omega_p)\right|^2}{\left|H'(j\omega_p)\right|^2} \cdot \left(R_{in}^2+X_{in}^2\right) - X_{in}^2} \right) - R_{in}, \tag{6.10}$$

where $H'(j\omega_p)$ and $H(j\omega_p)$ are given by (6.5) and (6.7), respectively.

6.4.4 Power Consumption Model

The average power consumed by the network shown in Figure 6.1 is

$$P_{total,avg} = P_{trx,avg} + P_{qrn,avg} + P_{rec,avg}, \tag{6.11}$$

where $P_{trx,avg}$, and $P_{rec,avg}$ are the average power consumption of the transmitter and receiver, respectively, and $P_{qrn,avg}$ is the average power consumption of the quasi-resonant network (including the driver).

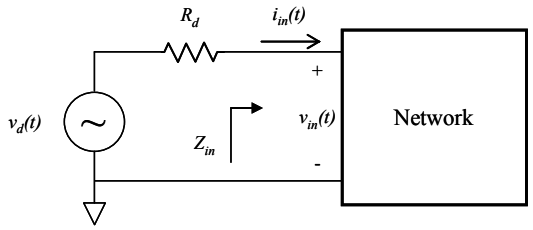


Figure 6.7 One-port network driven by a voltage source

The resonant interconnect network is a passive linear network. Therefore, a one-port network, as depicted in Figure 6.7, can be used to determine $P_{qrn,avg}$. The output impedance of the driver R_d and the input impedance of the network Z_{in} determine the power consumption of the network. Power, defined as the rate at which energy is absorbed [125], is

$$P_{qrn,avg} = \frac{1}{2} V_{in,rms}^2 \cdot \Re\left\{\frac{1}{Z_{in}}\right\}, \tag{6.12}$$

where $V_{in,rms}$ is the effective or root-mean square value of any periodic voltage and is

$$V_{in,rms} = \left(\frac{1}{T}\int_{t_0}^{t_0+T} \left|v_{in}(t)\right|^2 dt\right)^{1/2}, \tag{6.13}$$

where T is the time period of the periodic function $v_{in}(t)$. The input voltage $v_{in}(t)$ (see Figure 6.7) can be expressed in terms of the voltage source function $v_d(t)$, and is

$$v_{in}(t) = v_d(t) \cdot \frac{Z_{in}}{Z_{in} + R_d}. \tag{6.14}$$

Substituting (6.13) and (6.14) into (6.12) results in

$$P_{qrn,avg} = \frac{1}{2} V_{rms,d}^2 \cdot \left| \frac{Z_{in}}{Z_{in} + R_d} \right|^2 \cdot \Re\left\{ \frac{1}{Z_{in}} \right\} = \frac{1}{2} V_{rms,d}^2 \cdot \frac{R_{in}}{\left(R_{in} + R_d\right)^2 + X_{in}^2}. \tag{6.15}$$

Substituting (6.10) into (6.15) yields a simplified expression, depending solely on the input impedance of the network Z_{in},

$$P_{qrn,avg} = \frac{1}{2} V_{rms,d}^2 \cdot \left| \frac{H'(j\omega_p)}{H(j\omega_p)} \right|^2 \cdot \frac{R_{in}}{\left(R_{in}^2 + X_{in}^2\right)}. \tag{6.16}$$

6.4.5 Signal Delay Model

The total signal delay of the network shown in Figure 1 is

$$t_{d,tot} = t_{d,trx} + t_{d,qrn} + t_{d,rec,} \tag{6.17}$$

where $t_{d,trx}$, and $t_{d,rec}$ are the transmitter and receiver signal delay, respectively, and $t_{d,qrn}$ is the signal delay of the quasi-resonant network (including the driver).

A derivation of a closed-form expression for $t_{d,qrn}$ is based on the work of Chen and Friedman [31]. The authors of [31] based their analysis on a Fourier series, where a time-domain waveform as well as the 50% delay is approximated by the summation of several sinusoids. Although this method is designed for interconnects driven by a periodic signal, it can also be applied to a quasi-resonant network. As shown in Figure 6.3, the QRN is driven by a modulated signal periodic in time with two different symbols for logic one and zero. It can therefore be assumed that a quasi-resonant network is driven by a quasi-periodic signal.

The 50% delay can be expressed as

$$t_{d,qrn} = \frac{\arctan x_0}{\omega_p} - \frac{t_r}{2}, \tag{6.18}$$

where x_0 is a single real root of a third order polynomial described in Appendix B.2. Note that the value of $\arctan x_0$ is in the range of $[0,\pi]$; otherwise, π should be added or subtracted from x_0. In the case of three real roots, the output waveform is not shaped like a square wave and can no longer represent a logic state. Finally, note that the derivation of (6.18) assumes that the 50% delay is less than $t_p/2 - t_r/2$.

6.4.6 Design Guidelines

The QRN design process is summarized as a flow diagram in Figure 6.8. The interconnect geometry, *i.e.*, length, width, and thickness, is designed according to technological constraints

and requirements. Once the geometry and transmission frequency are specified, the next step is to determine the line resistance, inductance, and capacitance per unit length. A look-up table for different on-chip inductor magnitudes and the corresponding equivalent lump model parameters (such as in Figure 6.2) is extracted from experimental measurements or an electromagnetic field solver [114].

Expressions (6.16) and (6.18), developed in subsections 6.4.3, 6.4.4, and 6.4.5 are used to optimize the QRN to minimize the power-delay product. Simultaneously solving (6.16) and (6.18) for different on-chip inductances and insertion points along the interconnect results in a power-delay product which can be described graphically. The equivalent driver resistance can be determined from (6.10). From these graphs, it is possible to infer the required magnitude of the on-chip inductor, the insertion point, and the driver resistance to enable optimum power-delay operation.

If the on-chip inductance, insertion point, and driver resistance are not physically practical, the design cycle should be repeated choosing a different technology, transmission frequency, or interconnect geometry. Once the previous phase is completed, the driver can be designed based on the specified output resistance. Finally, the transmitter is designed based on the operating frequency and the driver input capacitance.

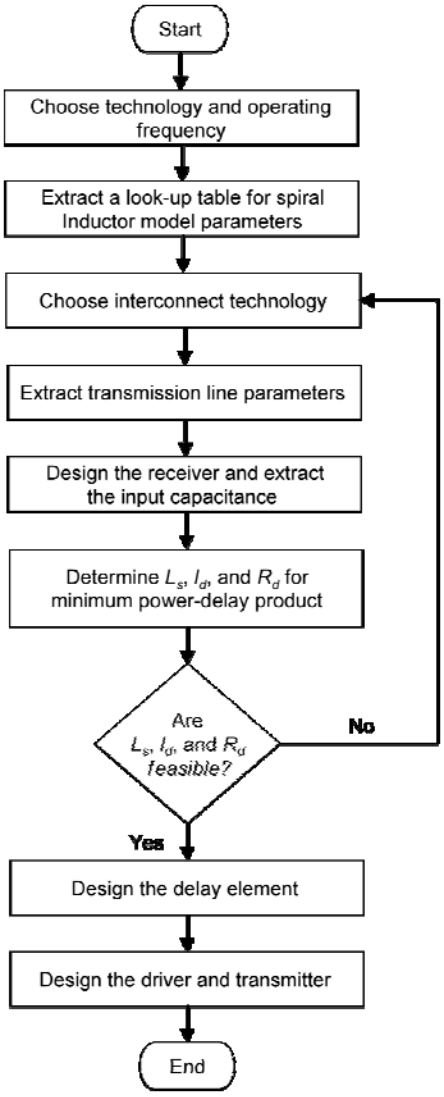


Figure 6.8 Flow diagram of QRN design process

6.5 Case Study

The design guidelines and principles presented in section 6.4.6 and illustrated in Figure 6.8 are demonstrated in a case study. This example is based on a TSMC 0.18 μm CMOS technology, transmitting data at a 5 Gbps frequency. The target time period is t_p = 200 psec

and the rise (and fall) time is t_r = 20 psec with a supply voltage V_{dd} = 1.8 volts. The on-chip octagonal inductor model parameters are extracted using SPIRAL [114]. The spiral inductors (varying between 1 to 10 nH) have been designed and optimized to achieve a maximum Q at 5 GHz.

The layout geometry and configuration of the quasi-resonant network is schematically illustrated in Figure 6.9. To reduce crosstalk and coupling noise from neighboring interconnects, the quasi-resonant interconnect is shielded by two parallel ground lines. The shield lines reduce the parasitic coupling capacitance between the two signal lines. Another strategy to mitigate crosstalk is to utilize perpendicular lines on different metal layers.

The separation between the signal and ground lines is 2 μm. The width of the signal and ground lines is 2 μm and 4 μm, respectively, while the thickness of each of the lines is 1 μm. In this example, the interconnect parameters (including the shield lines) are l = 5 mm, R = 17 mΩ/μm, L = 1.66 pH/μm, and C = 0.072 fF/μm.

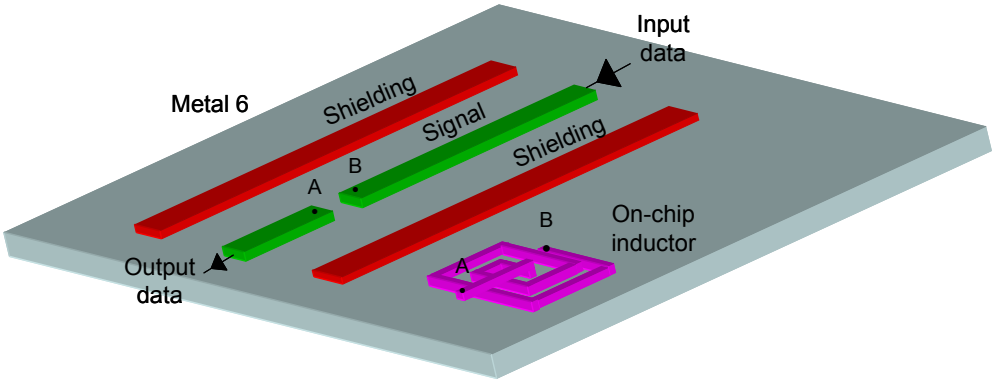


Figure 6.9 Layout of a resonant transmission line network

The receiver is based on the topology proposed in Figure 6.5(b) and the transistor widths are listed in Table 6.1. The input capacitance of the receiver circuit (*i.e.*, the first buffer shown in Figure 5(b)) is C_l = 17.4 fF.

Once the capacitive load of the interconnect is known, L_s, l_d, and R_d can be determined to minimize power and delay. The minimum power-delay product as a function of the on-chip inductance, using (6.5) to (6.18), is shown in Figure 6.10(a). The corresponding insertion point and driver resistance as a function of the on-chip inductance are shown in Figures 6.5(b) and 6.5(c), respectively. As evident from Figure 6.10, the minimum power-delay product occurs when an inductor L_s = 7.5 nH, inserted at l_d = 4.3 mm, and a driver resistance of R_d = 195 Ω is used. For these parameters, the power consumption and signal delay of the QRN are shown in Figure 6.11.

Table 6.1 Transistor width of the receiver and transmitter circuits, L = 0.18 μm channel length

Width [μm]			
Receiver		Transmitter	
M_{r1}	2.0	M_{t1}	1.0
M_{r2}	1.0	M_{t2}	1.0
M_{r3}	1.5	M_{t3}	2.5
M_{r4}	2.0	M_{t4}	1.0
M_{r5}	2.5	M_{t5}	0.5
M_{r6}	1.0	M_{t6}	2.5
M_{r7}	2.5	M_{t7}	1.0
M_{r8}	1.0	M_{t8}	0.3
M_{r9}	7.5	M_{t9}	0.2
M_{r10}	3.0		

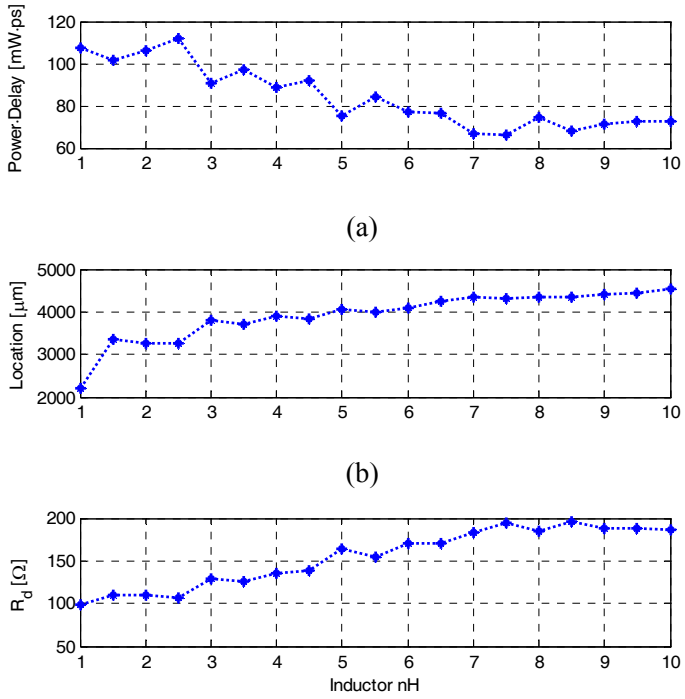
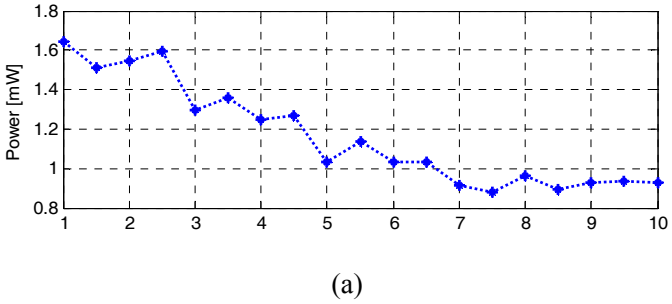
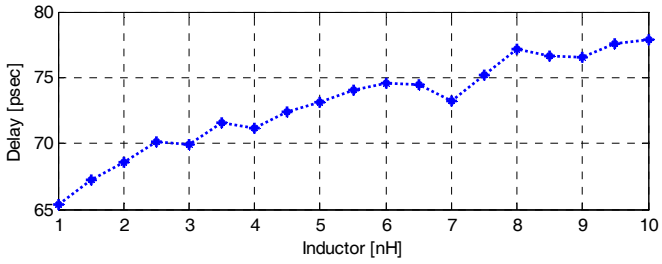


Figure 6.10 A design example of a 5 mm long interconnect operating at a 5 Gbps transmission frequency: (a) minimum power-delay product as a function of inductance, (b) insertion location as a function of inductance, (c) driver resistance as a function of inductance

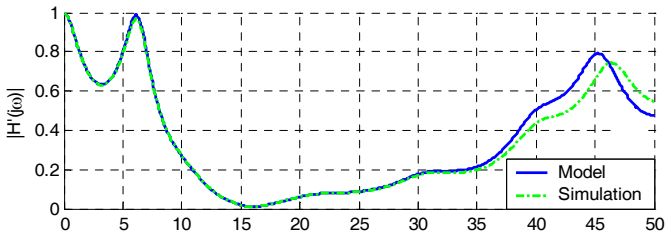




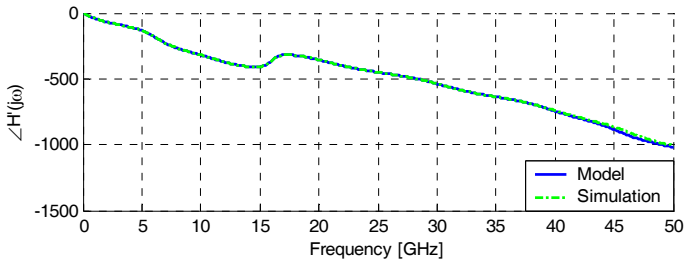
(b)

Figure 6.11 Power and delay: (a) power consumption, (b) signal delay

The magnitude and phase of the QRN transfer function are depicted in the frequency domain in Figure 6.12. Good agreement between simulations and the proposed analytic expressions for the magnitude and phase of the transfer function is achieved, exhibiting a maximum error of 19.5% and 2.8%, respectively. At a 5 GHz frequency, the magnitude of the transfer function is near resonance, here described as quasi-resonance. This behavior occurs since the addition of the parasitic capacitance of the on-chip spiral inductor to the network reduces the resonant frequency, i.e., $1/2\pi\sqrt{LC}$. The resonance peak, therefore, shifts to a higher frequency. Note that the magnitude of the transfer function reaches 0.8 at a 5 GHz frequency, as determined from (6.5).

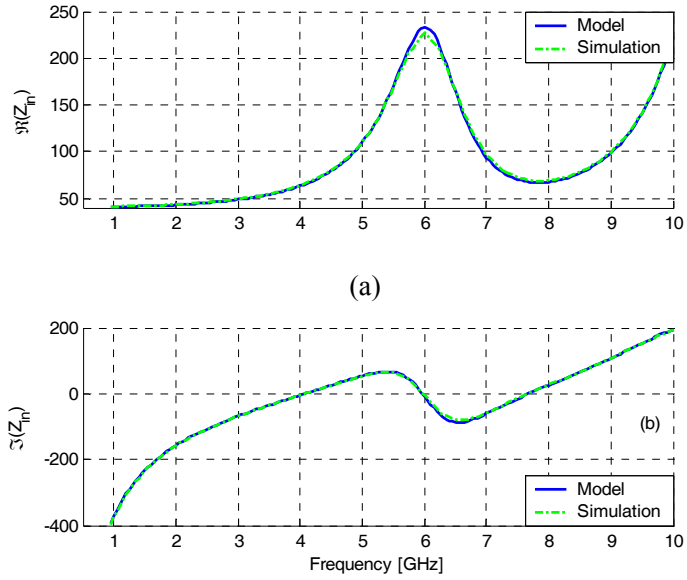


(a)



(b)

Figure 6.12 Frequency response of the transfer function: (a) magnitude, (b) phase



(a)

(b)

Figure 6.13 QRN input impedance: (a) real part, (b) imaginary part

The real and imaginary parts of the network input impedance are shown in Figure 6.13. With an on-chip inductor L_s = 7.5 nH operating at 5 GHz, the input impedance of the network is Z_{in} = 110 + j55. Good agreement between simulations and the proposed analytic

expressions for the magnitude and phase of the input impedance are achieved, exhibiting a maximum error of 4.3% and 37%, respectively.

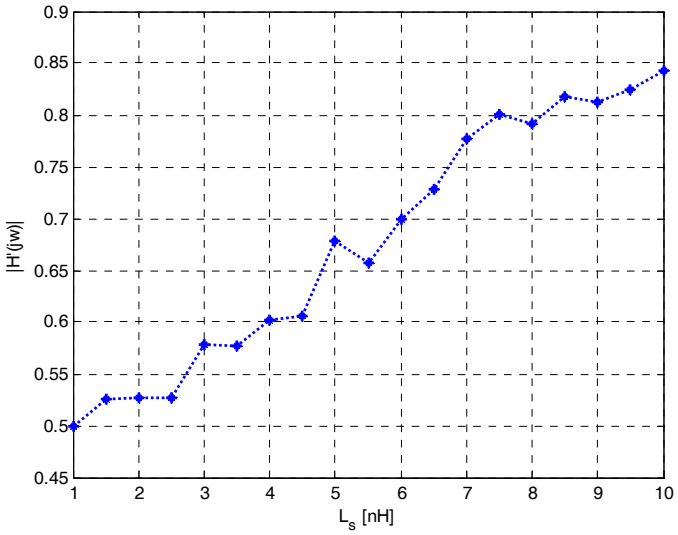
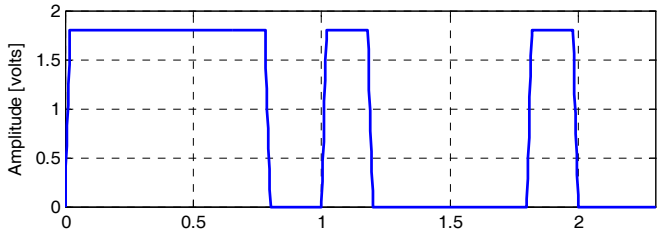
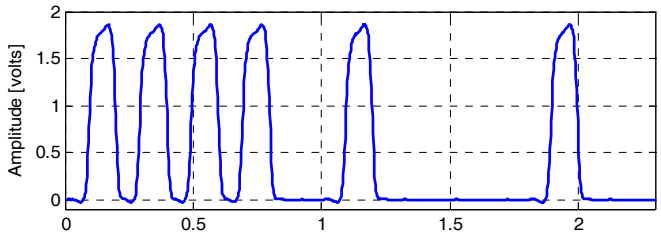


Figure 6.14 Magnitude of the transfer function as a function of the inductance at a 5 GHz operating frequency

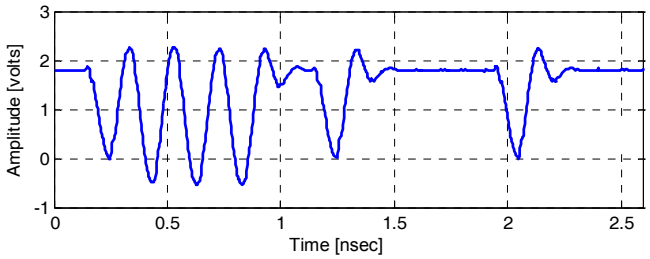
The effect of the on-chip inductor on the magnitude of the transfer function is shown in Figure 6.14. In this example, the magnitude of the transfer function reaches 0.8 with a 7.5 nH on-chip inductor. This behavior implies that a full swing voltage waveform is delivered at the far end only for this inductance.



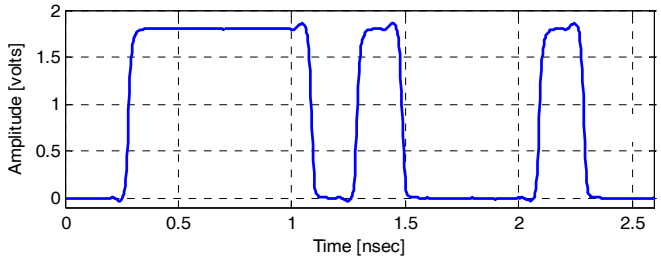
(a)



(b)



(c)



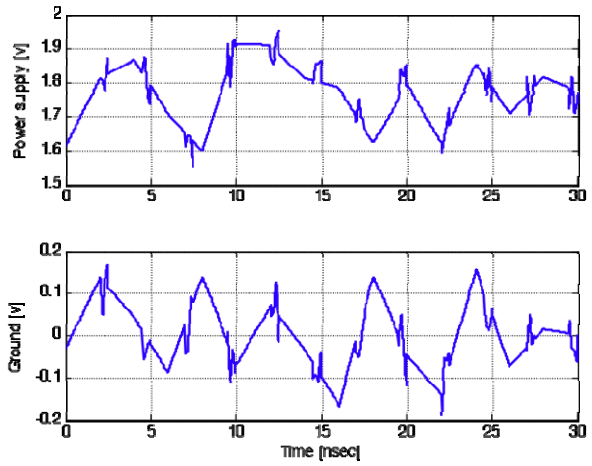
(d)

Figure 6.15 Ten bit data stream example: (a) data at input of transmitter, (b) data at output of transmitter, (c) data at input of receiver, (d) data at output of receiver

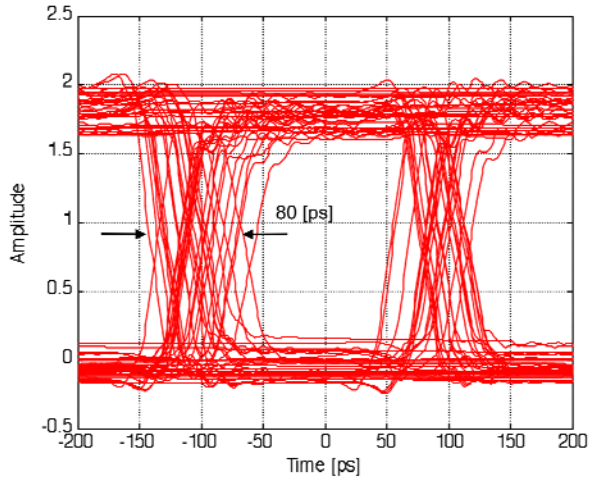
The driver in this example is a 0.18 μm CMOS inverter with P to N ratio = 20 / 8. The transmitter is based on the proposed circuit shown in Figure 6.4(c), and the transistor widths are listed in Table 6.1. The simulated input and output data signals described in the time domain are shown in Figure 6.15. Note that the square data waveform is distributed to the far end, achieving a full rail-to-rail voltage swing. In this example, a "1000101111" bit stream is transmitted at 5 Gbps. The simulated and analytic power consumption and delay of the receiver, transmitter, and QRN circuits are listed in Table 6.2.

Table 6.2 Simulated and analytic power consumption and delay of the receiver, transmitter, and QRN

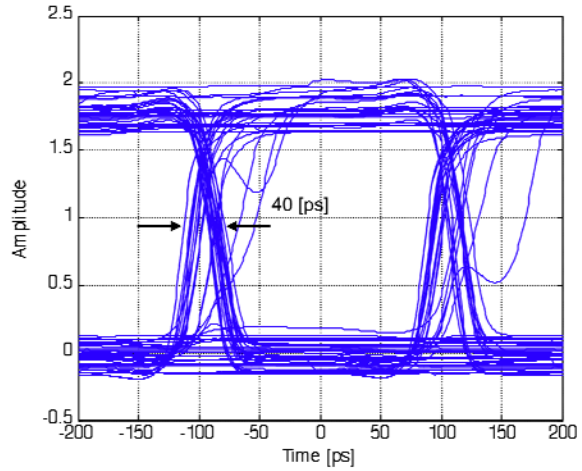
	Simulation		QRN		Total		
	Receiver	Transmitter	Model	Simulation	Model	Simulation	Error %
Power [mW]	0.45	0.34	0.88	1.18	1.67	1.90	12.11
Delay [psec]	27.50	86.00	74.00	80.00	187.50	193.50	3.10



(a)



(b)



(c)

Figure 6.16 Noise analysis: (a) noisy power supply and ground signals, (b) eye diagram of an interconnect with repeaters, (c) eye diagram of quasi-resonant interconnect

To evaluate the noise performance of the QRN for this case study, eye diagrams of non-resonant and quasi-resonant interconnects are generated by transmitting a random bit stream. To represent a noisy environment, the power supply and ground signals are corrupted by the

addition of Gaussian noise, as shown in Figure 6.16(a). The injected noise signal includes low frequencies to emulate the resonances produced by the package inductance and capacitance. From Figure Figure 6.16(c), observe that a large open eye is formed, exhibiting a high signal-to-noise ratio. The high noise rejection exhibited by the QRN is attributed to the detection and transmission of a single harmonic when a high logic level is transmitted. An eye diagram of an interconnect with repeaters is shown in Figure Figure 6.16(b). The jitter of the quasi-resonant interconnect (40 psec) is two times less than the non-resonant interconnect (80 psec).

6.6 Simulation Results and Comparison

To evaluate the accuracy of the model as compared to simulations, a 5 Gbps data signal and 0.5, 1, 3, 5, 10, 14 and 20 mm length interconnects are considered, as listed in Table 6.3. Good agreement between simulation and the model is demonstrated, exhibiting an average error of 13.1% and 6.1% for the power consumption and delay, respectively. The main reason for these discrepancies is due to the assumption of a linear driver model.

To evaluate the proposed methodology as compared to a traditional buffer insertion method, a 5 Gbps data signal and 0.5, 1, 3, 5, and 10 mm length interconnects are considered, as listed in Table 6.4. It is a challenge to design repeaters to drive long interconnects at frequencies as high as 5 GHz in a 0.18 μm CMOS technology. Hence, signal integrity in this repeater insertion case has been compromised for the sake of this comparison. From Table 6.4, the average reduction in power consumption and delay is 91.1% and 37.8%, respectively. The performance improvement is due to the repeaterless nature of the quasi-resonant interconnect.

A comparison between the proposed methodology and other aggressive approaches described in the literature is listed in Table 6.5. Interestingly, the greatest reduction in power consumption of 97.1% and delay of 35.6% occurs as compared to optoelectronic links [130]. The power consumption overhead in [130] is due to the edge emitting laser modulator at the transmitting edge and the photodiode and signal level restorer at the receiving end of the optical link. This comparison suggests that novel signaling schemes incorporating electrical interconnects outperforms optoelectronic solutions. These results are obtained despite the optical link transmitting a slower signal (3 Gbps) as compared to the resonant link (5 Gbps). With the exception of pulsed current [84], the resonant link transmits a higher frequency signal than the other methods. The proposed quasi-resonant interconnect methodology outperforms all of the other approaches described in the literature in power and, in most cases, in both power and latency.

Table 6.3 Comparison of QRN model with simulation: (a) power and delay using analytical model, (b) power and delay using simulation, (c) error between analytical model and simulation

(a)

Length [mm]	L_s [nH]	l_d [mm]	Analytic model		
			R_d [Ω]	Power [mW]	Delay [psec]
0.5	1.5	0	344	0.87	132
1	1.5	0.96	262	0.89	136
3	7.0	2.73	230	1.12	163
5	7.5	4.30	195	1.61	187
10	7.5	8.80	102	2.56	248
14	7.5	12.74	39	3.48	290
20	7.5	19.20	22	4.63	425

(b)

Length [mm]	L_s [nH]	l_d [mm]	Simulation			
			Driver		Power [mW]	Delay [psec]
			NMOS W/L	PMOS W/L		
0.5	1.5	0	5.0 / 0.18	12.5 / 0.18	0.95	136
1	1.5	0.96	6.0 / 0.18	15.0 / 0.18	0.92	148
3	7.0	2.73	7.0 / 0.18	17.5 / 0.18	1.27	168
5	7.5	4.30	8.0 / 0.18	20.0 / 0.18	1.91	193
10	7.5	8.80	11.0 / 0.18	27.5 / 0.18	2.54	263
14	7.5	12.74	14.0 / 0.18	35.0 / 0.18	2.80	309
20	7.5	19.20	16.0 / 0.18	40.0 / 0.18	3.16	368

(c)

Length [mm]	L_s [nH]	l_d [mm]	Error	
			Power %	Delay %
0.5	1.5	0	8.4	2.9
1	1.5	0.96	3.3	8.1
3	7.0	2.73	11.8	2.9
5	7.5	4.30	15.7	3.1
10	7.5	8.80	0.8	5.7
14	7.5	12.74	19.5	6.1
20	7.5	19.20	31.7	13.4
Average error			13.1	6.1

Table 6.4 Comparison of power consumption and delay

Length [mm]	This work		Repeater insertion		Improvement	
	Power [mW]	Delay [ps]	Power [mW]	Delay [ps]	Power %	Delay %
0.5	1.02	136	5.92	104	82.8	-23.5
1	0.99	148	10.79	160	90.8	7.5
3	1.34	168	16.87	255	92.1	34.1
5	1.98	193	25.20	368	92.1	47.5
10	2.61	263	51.20	1100	94.9	76.1
Average improvement					90.5	28.3

Table 6.5 Performance comparison of quasi-resonant method with different approaches

	Technology [nm]	Speed [Gbps]	Length [mm]	Power [mW]	Delay [ps]
This work [84]	180	5	3	1.34	168
Pulsed current [84][81]	180	8	3	27.12	280
Improvement				95.1%	40.0%
This work	180	5	20	3.23	368
Signal modulation [81]	180	1	20	16	300
Improvement				79.8%	-18.5%
This work	180	5	5	1.91	193
Optics (edge emitting) [130]	250	3	5	78	260
Improvement				97.5%	25.8%
This work	180	5	5	1.98	193
Optics (VCSEL) [130]	250	3	5	66	300
Improvement				97.0%	35.6%
This work	180	5	14	2.8	309
Loss compensation [135]	180	3	14	6	140
Improvement				53.3%	-54.7%

The primary tradeoff of the proposed methodology is the large area occupied by the on-chip inductor. This additional area is required to achieve a high Q inductor. However, in more advanced technologies, the area required for these inductors will be much lower. A comparison of the total area of repeaters and quasi-resonant methods for different technology nodes is listed in Table 6.6. The reduction in area overhead for a 50 nm CMOS technology is due to the requirement to insert additional repeaters to overcome the losses associated with long thin wires. Concurrently, the area of the on-chip inductor is lower due to the decrease in wire width and spacing.

Table 6.6 Area comparison of repeaters and on-chip spiral inductors for different technologies

Length [mm]	Repeater insertion [μm^2]		Quasi-resonant [μm^2]		Area overhead	
	180 [nm]	50 [nm]	180 [nm]	50 [nm]	180 [nm]	50 [nm]
0.5	19	57	9063	1007	x 480	x 17
1	32	96	9063	1007	x 288	x 10
3	44	132	6803	756	x 155	x 6
5	63	189	6722	747	x 107	x 4
10	126	378	6722	747	x 53	x 2

6.7 Summary

A methodology is described in this chapter for designing quasi-resonant interconnect networks. An accurate model is presented based on transmission line theory and a lumped high frequency model of an on-chip spiral inductor. The accuracy of the model enables the design of low power, low latency resonant communication links. The methodology can be used to determine the specific inductance L_s, insertion point l_d, and driver resistance R_d that minimizes the power-delay product.

Good agreement between the proposed model and simulation is exhibited, achieving an average error of 13.1% and 6.1% for the power consumption and delay, respectively. Quasi-resonant interconnects are shown to outperform other technologically aggressive circuit approaches. For buffered lines, an average reduction of 90.5% and 28.3% is obtained in power consumption and delay, respectively. As compared to optical links, a reduction of 97.0% and 35.6% is observed in power consumption and delay, respectively. These results show that quasi-resonant interconnects exhibit superior performance, suitable for high performance, high complexity integrated circuits.

Chapter 7

Power Supply Generation and Distribution for Three-Dimensional ICs

A design methodology for distributing a buck converter rectifier for application to three-dimensional (3-D) circuits is described. The 3-D rectifier exploits the properties of transmission lines, permitting the generation and distribution of power supplies to different planes. As compared to a conventional rectifier, the proposed rectifier circuit only requires on-chip capacitors without the use of on-chip inductors. A case study in a 0.18 μm CMOS 3-D technology demonstrates the generation of a 1.2 volt power supply delivering 700 mA peak current.

7.1 Introduction

In the era of rapid technology scaling, the performance and reliability of integrated circuits (IC) have reached limits that are difficult to surpass. As a result, novel design methodologies for high performance, high complexity ICs are required. Three-dimensional (3-D) nanoscale technology can provide the required characteristics of future state-of-the-art integrated systems. With the significant performance improvement offered by 3-D circuits [136], new design challenges arise. One primary requirement of 3-D integrated systems is diverse, high quality, and reliable power. This fundamental issue of power generation and distribution in 3-D circuits is explored in this chapter.

Three-dimensional integrated circuits are comprised of multiple planes with many circuit domains. The different planes are typically dedicated to a specific function, forming a highly heterogeneous system [136], [137], [138]. As an example, RF, analog, communications, and digital circuits are typically located on different planes, requiring several power supply voltages, as illustrated in Figure 7.1. In this example, V_{dd1}, V_{dd2}, V_{dd3}, and V_{dd4} are generated from the primary power supply V_{dd}.

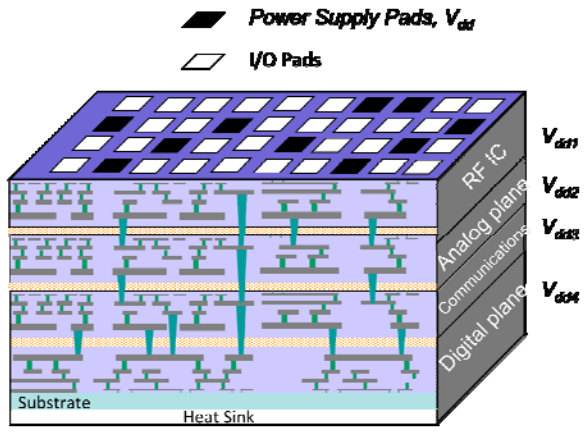


Figure 7.1 Three-dimensional circuit with multiple power supplies

Multiple circuit domains require several power supplies to reliably operate and provide sufficient and stable current. To provide circuit domains with the appropriate power supplies, DC-DC converters are distributed across each plane [139] as it is often impractical to provide external power supplies due to the limited number of input pins. Moreover, planes located far from the faces of the 3-D cube require a large number of expensive 3-D vias to distribute the power supply across the plane.

To alleviate these difficulties, DC-DC converters are distributed on-chip, generating a specific voltage required by the different circuit blocks within each plane of a 3-D system. A

suitable DC-DC converter for low power applications is a buck converter [139], [140]. A buck converter generates an output supply voltage with a magnitude smaller than the input supply voltage [139]. This property is useful for on-chip power supply generation.

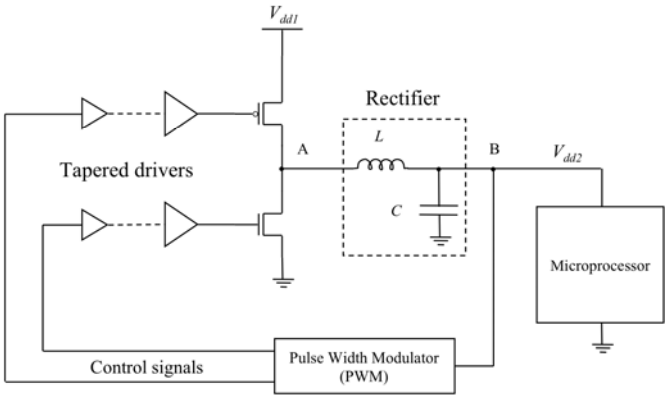
Exploiting the rectifier portion of a buck converter to generate and distribute power supplies in 3-D integrated circuits is the primary focus of this chapter. The proposed rectifier is comprised of on-chip interconnects and capacitors, eliminating the need for on-chip inductors. For a target DC voltage ripple, the distributed rectifier produces the required transfer function, passing the DC component of the input signal while attenuating the high frequency harmonics.

The chapter is organized into six sections. Background on the operation of a conventional buck converter is reviewed in section 7.2. In section 7.3, the operating principle behind a novel distributed rectifier within a buck converter is presented, followed by a methodology for designing these circuits in section 7.4. To exemplify the proposed approach, a case study is described in section 7.5, while a performance analysis of the distributed rectifier is presented in section 7.6. Some conclusions are offered in section 7.7.

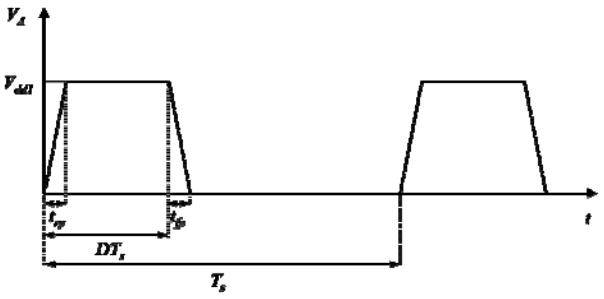
7.2 Background

A standard topology of a buck converter for high performance microprocessors is depicted in Figure 7.2(a) [139]. The power MOSFETs produce an AC signal at node A by a signal controlled by a pulse width modulator (PWM) [141], [142], as shown in Figure 7.2(b). The AC signal at node A is filtered by a rectifier composed of a second order low pass band *LC* filter. Assuming the resonance frequency $1/2\pi\sqrt{LC}$ is less than the switching frequency

of the power MOSFETs, the filter only passes the DC component of the signal and a residue composed of the high frequency harmonics.



(a)



(b)

Figure 7.2 Conventional DC-DC converter: (a) buck converter circuit [139], (b) signal at the output of the power MOSFETs (node A)

The DC component of the signal shown in Figure 7.2(b) is

$$\frac{1}{T_s}\int_0^{T_s} V_A dt = \frac{V_{dd1}\left(2DT_s - t_{rp} + t_{fp}\right)}{2T_s}, \tag{7.1}$$

where D, T_s, and V_{dd1} are the duty cycle, time period, and input power supply, respectively, and t_{rp} and t_{fp} are the rise and fall times, respectively, of the switching signal produced by a pulse width modulator (PWM) circuit. In the case of $t_{rp} = t_{fp}$, (7.1) reduces to

$$V_{dd2} = V_{DC}\big|_{t_{rp}=t_{fp}} = DV_{dd1}. \tag{7.2}$$

Hence, the buck converter produces an output voltage V_{dd2} at node B equal to DV_{dd1}.

The power transistors are typically large in physical size and are therefore driven by tapered buffers [143]. These buffers are controlled by the PWM circuit. The feedback PWM circuit senses the output voltage supply V_{dd2} at node B and modifies the control signal to ensure that the appropriate duty cycle D is produced at node A. In this manner, the output voltage is maintained at the desired value while compensating for variations in the load current and input voltage. The performance and functionality of the different circuit domains under load current and DC input voltage variations are dependent upon the maximum voltage fluctuations generated by the buck converter, which is therefore regulated to provide the target voltage [139].

The performance of a buck converter can be improved by integrating these converters on-chip. In this manner, the parasitic losses associated with the interconnects among the non-integrated components of the DC-DC converter are significantly decreased. Moreover, integrated converters benefit by advances in on-chip technologies and high operating frequencies. Monolithic fully integrated DC-DC converters can therefore achieve higher efficiency as compared to non-integrated converters [144], [145], [146], [147].

Integrating DC-DC converters on-chip in both 2-D and 3-D technologies, however, imposes challenges as the on-chip integration of large inductive and capacitive elements is

problematic. A significant issue is the poor parasitic impedance characteristics exhibited by the on-chip inductors [88], [148], which degrades the performance of the on-chip converter. To improve the quality factor of a 2 nH inductor and reduce the ripple current within the inductors, magnetic coupling between two on-chip inductors has been used for a DC-DC converter in [149]. Although the size and magnitude of the on-chip inductors and capacitors required to implement a buck converter are reduced with increasing switching frequency, the on-chip passive devices comprising the rectifier are large and cannot be practically integrated in the megahertz frequency regime [150], [151], [152].

In [153], [154], exotic inductor technologies have been developed to support on-chip integration of the DC-DC converter in a stacked-chip environment. In [153], a coupled thin film with magnetic core inductors is integrated on-chip, achieving a high effective inductance while in [154], an air core spiral inductor is integrated in a 180 nm SiGe BiCMOS technology. In both circuits, a special inductor technology is utilized to achieve higher inductance and efficiency. In [155], a 2×2 mm^2 plane within a stacked-chip technology is used for the on-chip *LC* output filter. By exploiting higher switching frequencies and the proposed distributed rectifier, the on-chip integration and distribution of these types of power supplies in 2-D and, in particular, 3-D circuits become possible. All of these circuits, however, do not consider current distribution to the target planes. The multi-plane structure of a 3-D system accommodates the nature of the proposed distributed rectifier, achieving efficient generation and distribution of the power supplies.

A three-dimensional technology provides several advantages as compared to a two-dimensional technology. When the on-chip capacitors used by a distributed rectifier are implemented with active devices, less metallization resources are required to connect these capacitors to different sections of the rectifier as compared to a 2-D technology. Since in 3-D

circuits each plane has a dedicated active device layer, routing congestion is reduced. Moreover, in a typical 2-D or 3-D system, the primary power supply is located above both circuits. In a 2-D technology, the input power supply to the power MOSFETs is routed to the lowest layer, where the active devices are located; therefore, additional routing resources are required. In a 3-D technology, however, the closest plane to the input power supply accommodates the power MOSFETs, saving metallization resources. The capacitance and inductance of the through silicon vias are exploited, reducing the required length of the interconnects and the size of the capacitors. An inherent benefit of utilizing the proposed rectifier in 3-D systems is that the generation and distribution of the power supply occur simultaneously, while in a 2-D technology, the rectifier is only used to generate the power supply.

7.3 Principle of a Distributed Rectifier

The proposed rectifier exploits the impedance characteristics of long transmission lines. By suppressing the inductive effects of long interconnects, thereby attenuating the associated resonant peaks, the transfer function of the amplitude of these *RC* dominant interconnects exhibits no overshoots and decreases rapidly with increasing frequency. Consequently, this type of interconnect behaves as a low pass filter which can be utilized as a buck converter rectifier. In this manner, the high frequency harmonics of the AC signal at node A are filtered, producing a DC signal at node B. The distributed low pass filter spans multiple planes of a 3-D circuit, providing the required power supply to a specific circuit domain.

To illustrate this impedance behavior, consider a transmission line driven by a voltage source with an output resistance R_d and terminated with a capacitive load C_l, as shown in

Figure 7.3. Transfer functions characterizing the effect of different capacitive loads are shown in Figure 7.4. In this example, the driver resistance R_d, interconnect length l, and capacitive load C_l are 100 Ω, 1 mm, and 1 pF, respectively, with transmission line parameters r = 17 mΩ/μm , l = 1.66 pH/μm, and c = 0.072 fF/μm. The *RLC* parameters are extracted using a three-dimensional electromagnetic solver, Henry and Metal, from the OEA software suite [114].

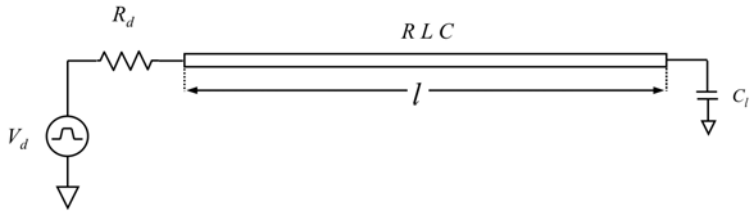


Figure 7.3 *RLC* transmission line

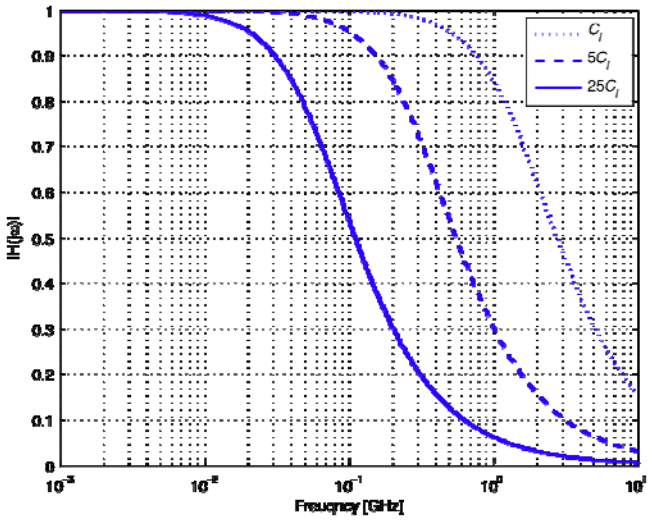


Figure 7.4 Transfer function of the magnitude of an interconnect for different capacitive loads

Note that in the example depicted in Figure 7.4, the resonant peaks associated with strong inductive effects do not appear. Additionally, as the load capacitance increases, the bandwidth of the low pass filter decreases. These observations suggest that an interconnect based rectifier can be used to generate and distribute power supplies within a 3-D system.

To obtain additional insight into the operation of a distributed rectifier as compared to a lumped *LC* rectifier, consider the transfer function of four types of rectifiers, as depicted in Figure 7.5. The transfer function of a single stage *RC* filter and a two stage *RC* filter are shown, respectively, in Figures 7.5(a) and 7.5(b). Since the roll-off slope of the single stage *RC* filter is -20 dB/decade, high frequency harmonics produced by the power MOSFETs are passed. Two cascaded *RC* filters produce a roll-off slope of -40 dB/decade. The implementation of these on-chip resistors for *RC* filters, however, requires an extremely large area since a DC-DC converter delivers a large amount of current. Wider wires are therefore required, reducing the effective resistance and increasing the physical area of the on-chip resistors. Additionally, these resistors dissipate power, reducing the overall efficiency of the converter. For these reasons, cascaded *RC* rectifiers are not a practical rectifier structure for on-chip DC-DC converters.

A second order low pass *LC* filter is used in a typical conventional DC-DC converter. When the effective output resistance R_d of the power MOSFETs is not included in the model, ideal frequency characteristics are observed in Figure 7.5(c) with a roll-off slope of -40 dB/decade and a resonance peak at $1/2\pi\sqrt{LC}$. When R_d is included, however, two poles at different frequencies are formed, resulting in a roll-off slope of -20 dB/decade in the megahertz and -40 dB/decade in the gigahertz frequency range, respectively. The frequency behavior of a distributed rectifier in the megahertz frequency range is therefore similar to a lumped *LC* rectifier, as can be observed in Figure 7.5. When R_d is included in the analysis, a

pole in the megahertz frequency is formed, resulting in a roll-off slope of -20 dB/decade, as shown in Figure 7.5(d).

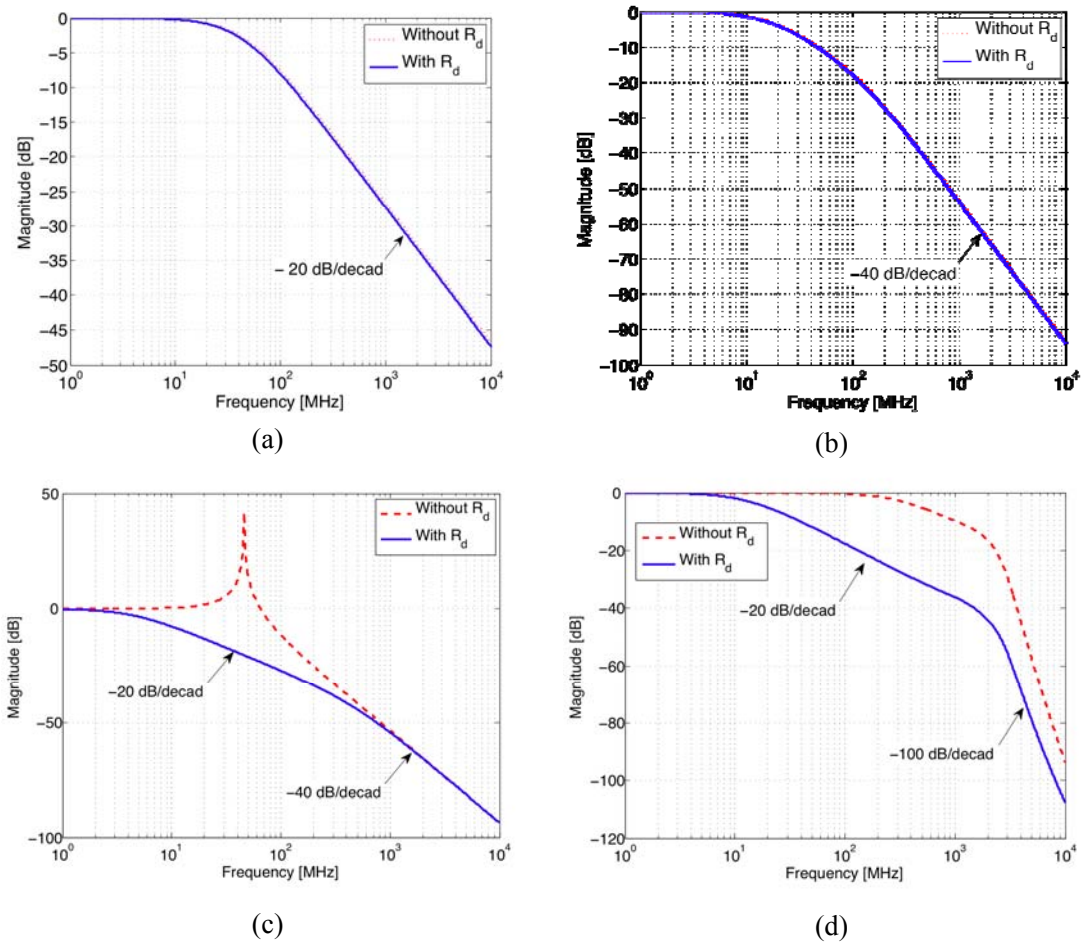


Figure 7.5 Magnitude of transfer function: (a) single stage *RC* filter, (b) two stages *RC* filter, (c) *LC* rectifier, (d) distributed rectifier

Note that in this example, a sharp -100 dB/decade roll-off slope is formed in the gigahertz frequency range, efficiently suppressing the high frequency harmonics of the AC signal produced by the power MOSFETs. The distributed nature of the proposed rectifier

forms multiple poles at approximately the same high frequency, resulting in a large negative slope.

7.4 Design Methodology

A design methodology for a distributed buck converter is described in this section. A circuit and mathematical formulation of the power delivery process in 3-D circuits are presented. Based on the expressions developed here, design guidelines are provided. The physical structure of the distributed rectifier as well as the current load characteristics are described in section 7.4.1. In section 7.4.2, the transfer function of the rectifier is used to determine the condition that satisfies a target power supply ripple while the efficiency and area of the 3-D rectifier are determined in section 7.4.3. Finally, design guidelines are described in section 7.4.4.

7.4.1 Physical Structure and Current Load Properties of a 3-D Rectifier

The proposed distributed rectifier is depicted in Figure 7.6. The rectifier is driven by power MOSFETs (see Figure 7.2(a)) which are modeled as a voltage source V_A followed by an effective resistance R_d. The voltage source V_A is assumed to be periodic, as illustrated in Fig. Figure 7.2(b). A representation of this signal in the complex frequency domain s is

$$V_A(s)=\left(\frac{V_{dd1}}{t_{rp}t_{fp}\left(e^{-sT_s}-1\right)s^2}\right)\cdot\left(e^{-sT_{rp}}st_{fp}\left(1-t_{rp}\right)+t_{rp}e^{-sDT_s}-t_{rp}e^{-s\left(DT_s+t_{fp}\right)}-t_{fp}s\right) \quad (7.3)$$

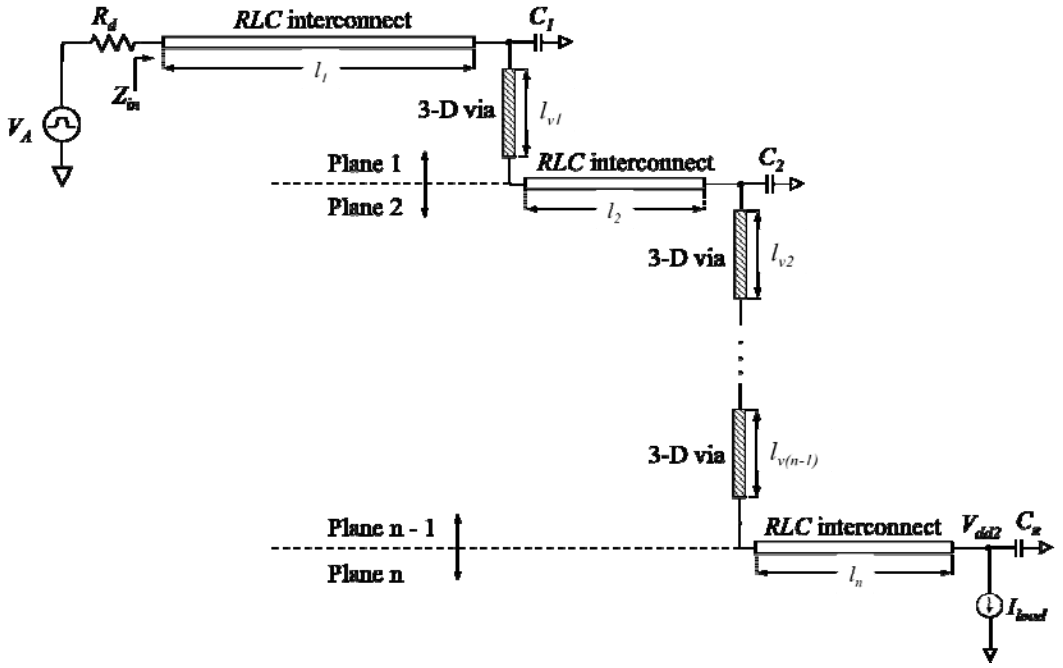


Figure 7.6 A distributed rectifier

The rectifier is composed of transmission lines terminated with lumped capacitances. The inter-plane structure is connected by 3-D vias. At the target plane *n*, the load is represented by a periodic current load and a reference clock signal, as shown in Figure 7.7. Note that the current load characterizes the approximate current profile of a specific circuit module on a plane. I_{load} remains at I_0 during clock low, demonstrating DC current flow provided by the power supply. A representation of this signal in the complex frequency domain *s* is

$$\begin{aligned} I_{load}(s) = & \left(e^{-s(t_{rc}+t_{fc})}\right)\left(a_f t_{rc} s + a_f t_{fc} s + a_f + b_f s\right) \\ & + e^{-t_{rc}s}\left(a_r t_{rc} s + a_r - a_f t_{rc} s - a_f - b_f s\right) \\ & - \left(a_r + I_0 s + I_0 s e^{-sT_{CLK}}\right)\Big)\Big/\left(s^2\left(e^{-sT_{CLK}} - 1\right)\right), \end{aligned} \tag{7.4}$$

where

$$a_r = \frac{\Delta i}{t_{rc}}, \tag{7.5}$$

$$a_f = -\frac{\Delta i}{t_{fc}}, \tag{7.6}$$

$$b_f = -a_f \left(t_{rc} + t_{fc} \right). \tag{7.7}$$

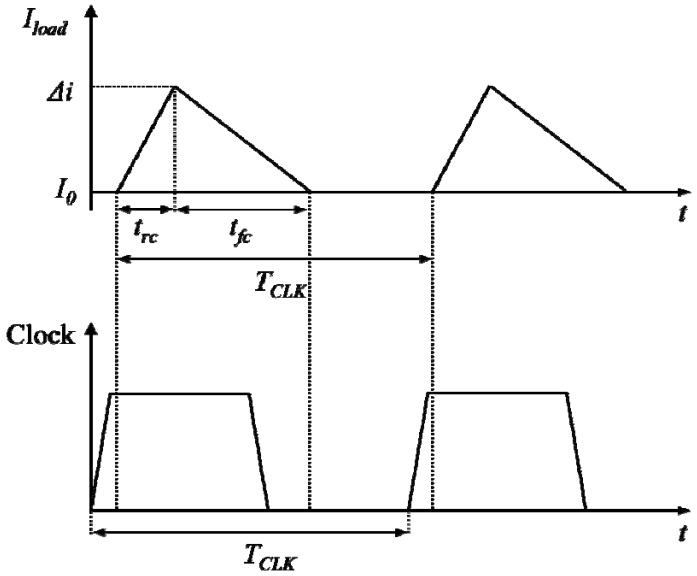


Figure 7.7 Current load profile in high speed digital circuits

As with a conventional buck converter, it is assumed that a feedback PWM circuit senses the output node of the rectifier and adjusts accordingly the duty cycle of the signal driving the power MOSFETs (see Figure 7.2(a)).

In three-dimensional circuits, the ability to deliver current is primarily limited by the 3-D vias. The maximum current that can be delivered through a single 3-D via therefore determines the current magnitude,

$$I_0 + \Delta i = J_{via,\max} \cdot A_{via} \cdot V, \tag{7.8}$$

where $J_{via,max}$, A_{via}, and V are the maximum current density, cross-sectional area, and number of 3-D vias on the same plane, respectively. Consequently, the maximum cross-sectional area of the interconnects (see Figure 7.6) distributing the current within the different planes is

$$A_{\text{int}} = \frac{I_0 + \Delta_i}{J_{\text{int},\max}} = \left(\frac{J_{via,\max}}{J_{\text{int},\max}} \right) \cdot A_{via} \cdot V, \tag{7.9}$$

where $J_{int,max}$ is the maximum current density of the interconnect.

In practical circuits, however, a significant amount of current is sunk by the load. To satisfy this requirement, multiple structures N, as depicted in Figure 7.6, are connected in parallel, delivering $N(I_0 + \Delta i)$ amperes. In this case, the number of 3-D vias within the rectifier on each plane is equal to the number of parallel connected structures $V = N$. The effective resistance and inductance per unit length of the interconnects and 3-D vias, as well as the output resistance of the driver R_d (see Figure 7.6), are N times smaller. The capacitance per unit length of the interconnects and 3-D vias, as well as the on-chip lumped capacitors, are N times larger.

7.4.2 Transfer Function of a 3-D Rectifier

To characterize the impedance of the rectifier (including R_d), the overall transfer function is determined based on the *ABCD* matrices. Hence, the overall *ABCD* matrix of a rectifier spanning *n* planes is

$$\begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{bmatrix} = \begin{bmatrix} 1 & R_d \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}. \tag{7.10}$$

The right matrix in the right hand side of (7.10) is

$$\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} = \prod_{i=1}^{n} \left(\begin{bmatrix} \cosh \gamma l_i & Z_0 \sinh \gamma l_i \\ \sinh \gamma l_i / Z_0 & \cosh \gamma l_i \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ j\omega C_i & 1 \end{bmatrix} \cdot \begin{bmatrix} \cosh \gamma_v l_{vi} & Z_{v0} \sinh \gamma_v l_{vi} \\ \sinh \gamma_v l_{vi} / Z_{v0} & \cosh \gamma_v l_{vi} \end{bmatrix} \right), \tag{7.11}$$

where Z_0 and γ are the characteristic impedance and propagation constant of the *RLC* interconnects, respectively, l_i and C_i are the interconnect length and capacitance on the i^{th} plane, respectively, Z_{v0} and γ_v are the characteristic impedance and propagation constant of the 3-D vias, respectively, and l_{vi} is the length of the 3-D via on the i^{th} plane. The transfer function of the rectifier is

$$H_{rect}(j\omega) = \frac{1}{\widetilde{A}} \left(1 - \widetilde{B} \frac{I_{load}(j\omega_{CLK})}{V_A(j\omega_s)} \right), \tag{7.12}$$

where $\tilde{A}$ and $\tilde{B}$ are obtained from (7.10), and ω_{CLK} and ω_s are the radian frequency of the clock and switching signals, respectively.

Since a practical rectifier within a buck converter does not provide ideal low pass characteristics, the signal at node B, shown in Figure 7.2(a), carries a small amount of high frequency harmonics generated by the switching power MOSFETs. Hence, the voltage at node B is

$$V_{dd2}(t) = V_{DC} + V_{ripple}(t), \tag{7.13}$$

where V_{DC} is the DC component of the output voltage described by (7.1) and $V_{ripple}(t)$ is the voltage ripple transferred by the non-ideal characteristics of the rectifier. When only the fundamental harmonic is passed, $V_{ripple}(t)$ exhibits a sinusoidal behavior,

$$V_{ripple}(t) = V_r \sin(\omega_s t). \tag{7.14}$$

To satisfy a target ripple voltage V_r (peak-to-peak), the rectifier transfer function at the switching frequency f_s has to achieve a specific magnitude. To satisfy this objective, consider the output signal in the frequency domain,

$$|V_{dd2}(s)| = |H_{rect}(s)| \cdot |V_A(s)|. \tag{7.15}$$

The periodic input signal V_A can be represented by a Fourier series,

$$v_A(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_s t}, \tag{7.16}$$

where a_k is the k^{th} harmonic of the signal. In the case of the signal illustrated in Fig. Figure 7.2(b) and assuming $t_{rp} = t_{fp} = t_r$, the fundamental harmonic (positive and negative) is

$$a_{\pm 1} = \left(\frac{V_{dd1} T_s}{4 t_r \pi^2} \right) \cdot \left(e^{\mp j\omega_s t_r} \left(1 - e^{\mp j2\pi D} \right) + e^{\mp j2\pi D} - 1 \right) \tag{7.17}$$

Equation (7.15) implies that the required amplitude of the transfer function for a specific ripple voltage V_r is

$$\left| H_{rect}(j\omega_s) \right| \le \frac{V_r/2}{2|a_1|} = \frac{V_r}{4|a_1|}. \tag{7.18}$$

Once the current profile of a circuit is determined, the interconnect length l_1 to l_n, shown in Figure 7.6, and the required ripple voltage V_r are chosen. Based on (7.12), the magnitude of the transfer function at ω_s is plotted as a function of the capacitances C_1 to C_n. The interconnect length and capacitance are chosen to satisfy (7.18). Since these design expressions are complex and unsuitable for manual calculations, numerical tools are used.

An important issue is to determine the duty cycle of the signal driving the power MOSFETs (see Figure 7.6) that produces the correct power supply voltage. Note that in this case, the duty cycle determined from (7.2) does not provide the proper DC voltage level. This behavior occurs since the signal at the input of a distributed rectifier is degraded by the resistances R_d and the input impedance of the rectifier Z_{in}, forming a voltage divider. To

obtain the duty cycle required for a specific DC voltage, consider the input impedance of the rectifier at DC,

$$Z_{in}(0)=\frac{V_{DC}A'(0)+I_{DC}B'(0)}{V_{DC}C'(0)+I_{DC}D'(0)}, \tag{7.19}$$

where A', B', C', and D' are defined in (7.11), and I_{DC} is the DC component of the current load,

$$I_{DC}=I_0+\frac{\Delta i}{2T_{CLK}}\cdot\left(t_{rc}+t_{fc}\right). \tag{7.20}$$

The DC component of the signal at the input of the rectifier (after R_d in Figure 7.6) is

$$DV_{dd1}=D_{PWM}V_{dd1}\left|\frac{Z_{in}(0)}{Z_{in}(0)+R_d}\right|, \tag{7.21}$$

the DC voltage transferred by the rectifier to the target plane. In (7.21), D_{PWM} is the duty cycle provided by the PWM feedback circuit (see Fig. Figure 7.2(a)). Consequently, to achieve a specific DC voltage at the output of the rectifier, the duty cycle in (7.2) is

$$D_{PWM}=D\cdot\left|1+\frac{R_d}{Z_{in}(0)}\right|. \tag{7.22}$$

Observe from (7.22) that D_{PWM} is always larger than the original duty cycle D obtained from (7.2), limiting the magnitude of the generated power supply. When the interconnects within the distributed rectifier are resistive, D_{PWM} approaches D (no reflections occur at the input). It is typically preferable to design the rectifier to ensure that D_{PWM} is closer to D to provide a large tuning range for the PWM circuit.

7.4.3 Efficiency and Area of a 3-D Rectifier

An important property of a buck converter is the power efficiency. The efficiency of a distributed rectifier within the buck converter is

$$\eta_{rect} = \frac{P_{load}}{P_{load} + P_{rect}} \times 100\%, \tag{7.23}$$

where P_{load} is the average power delivered to the load and P_{rect} is the average power consumed by the rectifier. The power expressions P_{load} and P_{rect} are, respectively,

$$P_{load} = V_{dd2} I_{DC}, \tag{7.24}$$

$$P_{rect} = V_{A,RMS}^2 \left| \frac{Z_{in}(j\omega_s)}{Z_{in}(j\omega_s) + R_d} \right|^2 \cdot \Re\left\{ \frac{1}{Z_{in}(j\omega_s)} \right\}, \tag{7.25}$$

where $V_{A,rms}$ is the root-mean-square value of the driving signal with a duty cycle D_{PWM}, and $Z_{in}(j\omega_s)$ are, respectively,

$$V_{A,RMS} = \sqrt{\frac{1}{T_s}\int_{t_0}^{t_0+T_s} |v_A(t)|^2 \, dt}, \tag{7.26}$$

$$Z_{in}(j\omega_s) = \frac{V_{DC}A'(j\omega_s) + I_{DC}B'(j\omega_s)}{V_{DC}C'(j\omega_s) + I_{DC}D'(j\omega_s)}. \tag{7.27}$$

The area occupied by a 3-D rectifier is

$$A_{3-D,rect} = \left(\sum_{i=1}^{n} \left(W_{metal,i} \cdot l_i + A_{C_i} \right) + \sum_{i=1}^{n-1} A_{via,i} \right) \cdot N + S_i \cdot l_i \cdot n(N-1) \tag{7.28}$$

where A_{Ci} and $A_{via,i}$ are the area occupied by the on-chip capacitors and 3-D vias on the i^{th} plane, respectively. $W_{metal,i}$ and S_i are the width and spacing of the interconnects on the i^{th} plane, respectively. As mentioned in subsection 7.4.1, n and N are the number of planes spanned by the rectifier and the number of parallel connected rectifier structures, respectively. The generation and distribution of additional power supplies, as illustrated in Figure 7.1, requires the use of additional DC-DC converters and distributed rectifiers.

7.4.4 Design Guidelines

In order to provide a comprehensive perspective of the design space, several design parameters that affect the performance of the distributed rectifier are investigated. To determine the required magnitude of the transfer function of the rectifier for different output voltage ripples and DC voltages, consider Figure 7.8. The output DC voltage ranges between

0.8 volts and 3 volts (assuming a 3.3 volt input DC voltage), while the required voltage ripple ranges between 1% and 10%. As expected, the magnitude of the transfer function increases for larger voltage ripple, since the larger amplitude of the dominant harmonic of the input signal is permitted to pass. Furthermore, as the output DC voltage becomes larger, an increase in the transfer function is evident in (7.15). Consequently, a lower capacitance and less resistive interconnects are required. Once the required output DC voltage is chosen for a specific voltage ripple, permitting the target magnitude of the transfer function to be determined, the design space of the distributed rectifier is specified.

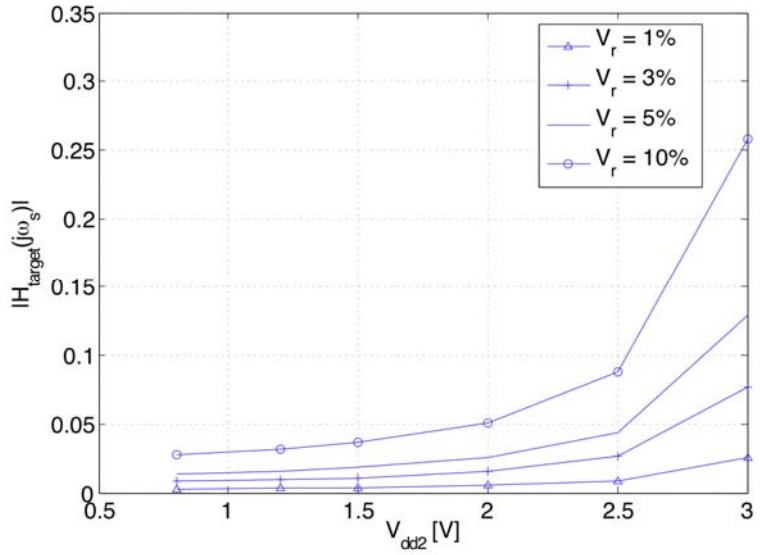


Figure 7.8 Required magnitude of the transfer function

In this example, all of the interconnect lengths are l = 1 mm with a voltage ripple V_r = 5%, input DC voltage V_{dd1} = 3.3 volts, and a switching frequency f_s = 100 MHz. Consider the design space of the rectifier, as shown in Figure 7.9. The modified duty cycle D_{PWM}, ratio between the duty cycle D (see (7.2)) and D_{PWM}, capacitance, and rectifier efficiency are

depicted as a function of the output DC voltage and series output resistance R_d of the power MOSFETs (see Figure 7.6). The modified duty cycle D_{PWM} is shown in Figure 7.9(a). Observations of Figure 7.9(a) reveal that under the specified operating conditions, this rectifier cannot convert 3.3 volts to 3 volts since for all values of R_d, D_{PWM} is greater than one. The feasible operation of the distributed rectifier is therefore limited to about 2.5 volts.

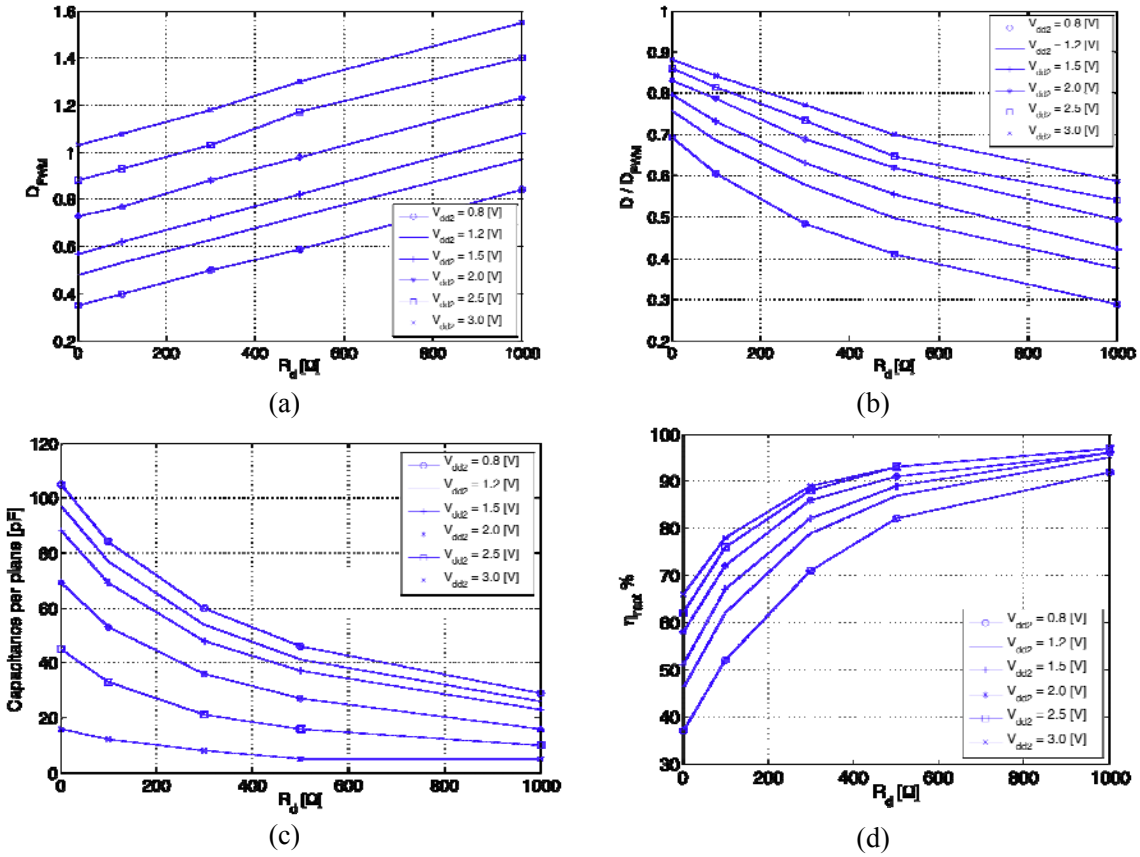


Figure 7.9 Design space of a distributed rectifier: (a) modified duty cycle as a function of R_d, (b) duty cycle ratio as a function of R_d, (c) capacitance as a function of R_d, (d) rectifier efficiency as a function of R_d

To narrow the design space, recall that D_{PWM} is always greater than D (see subsection 7.4.2). To permit a larger operational range of the feedback PWM circuit, D_{PWM} should be chosen close to D. A new design metric is therefore defined which is the ratio between D and D_{PWM}, as shown in Figure 7.9(b). Preferably, this ratio should be as closer to one as possible. As shown in Figure 7.9(b), the ratio D / D_{PWM} decreases as R_d increases. The lowest permissible ratio D / D_{PWM} is chosen to be 0.5 to accommodate a tradeoff between the ratio D / D_{PWM} and the rectifier efficiency.

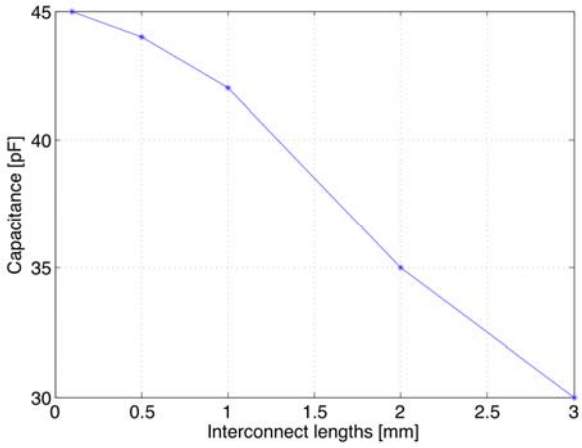
As mentioned in the beginning of this subsection, less capacitance is required for lower conversion ratios, as shown in Figure 7.9(c). As expected, with a larger series resistance R_d, the required capacitance decreases to satisfy the target magnitude of the transfer function. Finally, the rectifier efficiency η_{rect} is shown in Figure 7.9(d). The general trend for all output DC voltages is that the rectifier efficiency increases with higher R_d, as evident from (7.23) and (7.25). From Figures. 7.9(b), 7.9(c), and 7.9(d), the permissible range for R_d is

$$R_d\left(\min\left(\eta_{rect}\right)\right) < R_d < R_d\left(\frac{D_{PWM}}{D} = \frac{1}{2}\right), \tag{7.29}$$

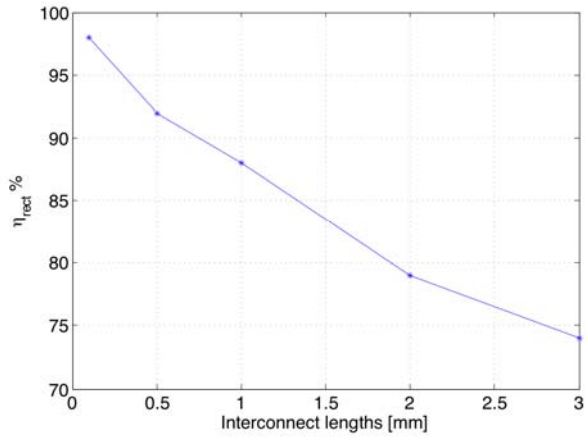
$$R_d\left(\max\left(C\right)\right) < R_d < R_d\left(\frac{D_{PWM}}{D} = \frac{1}{2}\right), \tag{7.30}$$

To evaluate the effect of the interconnect length on the performance of the distributed rectifier, consider Figure 7.10. The required capacitance to achieve a 5% voltage ripple and the corresponding efficiency is shown as a function of interconnect length. Note that these figures have been obtained assuming an output DC voltage V_{dd2} = 1.2 volts and R_d = 736 Ω, and all interconnect lengths and capacitors are equal. As expected, the efficiency decreases

for longer interconnect lengths since the resistance increases, dissipating more power. A tradeoff between the interconnect length and rectifier efficiency therefore exists. Choosing a low capacitance with longer interconnects decreases the rectifier efficiency.



(a)



(b)

Figure 7.10 Effect of interconnect length on performance: (a) capacitance as a function of interconnect length, (b) rectifier efficiency as a function of interconnect length

7.5 Case Study

To demonstrate the design methodology described in section 7.4, an example DC-DC converter based on the MIT Lincoln Lab (MITLL) 180 nm 3-D integration process [156], [160] is described in this section. Characterization of the maximum current that can flow through a single via is described in section 7.5.1. In order to quantify the advantages of the proposed distributed rectifier as compared to a conventional buck converter rectifier, a traditional rectifier is described in section 7.5.2. For the same performance requirements as a conventional rectifier, a distributed rectifier based on a 3-D technology is described and compared in section 7.5.3. SPICE simulations are performed based on the 3-D MITLL technology.

7.5.1 Current Load Characterization

The MITLL technology is a 0.18 μm low power, fully depleted silicon-on-insulator (FDSOI) CMOS process where three independent wafers are physically bonded to form a 3-D integrated structure. Each plane has three aluminum metallization layers. In this technology, the maximum current density is [160]

$$J_{Al,3-D} = 3\,mA/\mu m^2. \tag{7.31}$$

Since the cross-sectional area of a 3-D via in this technology is 1.5 × 1.5 μm^2, the maximum current that can flow through a single 3-D via, based on (7.8), is

$$I_0 + \Delta i \approx 7\ mA. \tag{7.32}$$

In this case study, the current load waveform, depicted in Figure 7.7, has the following characteristics: $1/T_{CLK}$ = 3 GHz, and t_{rc} and t_{fc} are $0.3T_{CLK}/2$ and $0.7T_{CLK}/2$, respectively.

7.5.2 Conventional Rectifier

A conventional buck converter rectifier is composed of a second order low pass *LC* filter, as shown in Figure 7.2(a). Assuming the inductor and capacitor exhibit ideal characteristics, the transfer function of the *LC* filter is

$$H_{rect,LC}(j\omega) = \frac{1}{1 - \omega^2 LC + j\omega CR_d} \cdot \left(1 - \frac{I_{load}(j\omega_{CLK})(R_d + j\omega L)}{V_A(j\omega_s)}\right), \tag{7.33}$$

where $V_A(s)$ and $I_{load}(s)$ are described by (7.3) and (7.4), respectively, and R_d, L, and C are the effective output resistance of the power MOSFETs and the inductor and capacitor of the conventional rectifier, respectively.

In this example, the DC-DC conversion is from 3.3 volts to 1.2 volts (D = 0.36) at a switching frequency f_s = 100 MHz. The effective output resistance of the power MOSFETs in a 0.18 μm CMOS technology delivering 700 mA is R_d = 7.36 Ω. To achieve 5% (V_r = $0.05V_{dd2}$ peak-to-peak) ripple voltage, the magnitude of the transfer function at 100 MHz according to (7.18) is 0.016. To satisfy these requirements, the corner frequency of an *LC* low pass filter, *i.e.*, the resonant frequency, should be about one decade less than f_s. This condition results in C = 15 nF and L = 5 nH, as indicated by the magnitude of the transfer function

shown in Figure 7.11. Note that for moderate performance requirements such as in this example, extremely large on-chip capacitance is required.

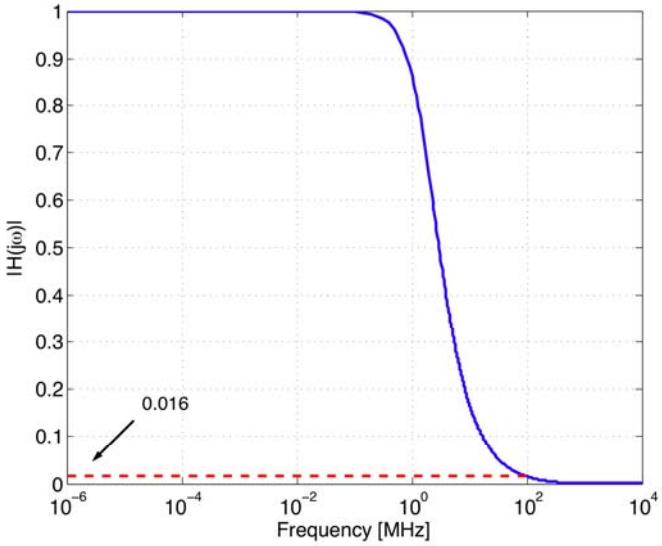


Figure 7.11 Magnitude of the transfer function of an *LC* rectifier

7.5.3 Distributed 3-D Rectifier

To overcome the difficulty described in subsection 7.5.2, a distributed rectifier circuit has been developed that generates and distributes the power supply to a target plane within a 3-D structure. This rectifier is described in this section. The resistance, inductance, and capacitance per unit length of the interconnects (Metal 3) and 3-D vias are extracted based on the predictive technology model (PTM) [157], [158], as listed in Table 7.1. The width of the interconnects is determined by the maximum current density of the MITLL 3-D technology. Assuming that both the interconnect and 3-D vias support the same current density of 3

mA/μm^2, the maximum cross-sectional area of the interconnect determined from (7.9) is 1.5 × 1.5 μm^2. The thickness of the interconnect for this technology is 630 nm, resulting in an approximately 4 μm wide line. Note that each 3-D via is 7.34 μm long, connecting three planes, as illustrated in Figure 7.6.

Table 7.1 *RLC* Interconnect and 3-D via impedances

	R [mΩ/μm]	L [pH/μm]	C [fF/μm]
Interconnects	14.5	1.3	0.5
3-D via [158]	20.4	0.55	0.37

For the same performance requirements as a conventional rectifier, the magnitude of the transfer function at ω_s for different capacitances and interconnect lengths, assuming $l = l_1 = l_2 = l_3$ and $C = C_1 = C_2 = C_3$, is depicted in Figure 7.12. An output resistance of the power MOSFETs R_d = 736 Ω is assumed in this case study.

As the interconnect line length increases, less capacitance is required, as evident in Figure 7.12. To satisfy the required voltage ripple, the target interconnect length in this example is l = 1 mm with a capacitance C = 42 pF. The magnitude of the transfer function of the distributed rectifier is shown in Figure 7.13. Note that the simulated rectifier accurately verifies the physical model described in section 7.4. The DC voltage at the output of the rectifier is shown in Figure 7.14. The simulated ripple of the output DC voltage is 52 mV (4.3% of the output DC voltage), satisfying the design objective of a maximum 5% voltage ripple. In this example, the required duty cycle of the signal driving the power MOSFETs is D_{PWM} = 0.73, and the efficiency of the rectifier is about 88%.

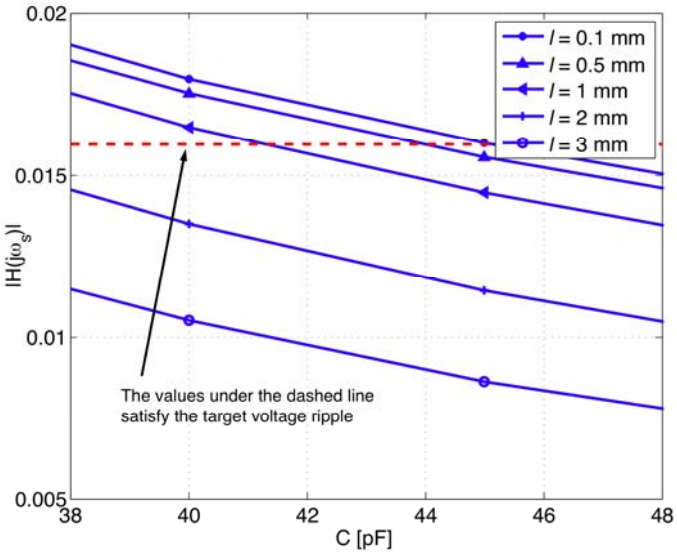


Figure 7.12 Magnitude of the transfer function at ω_s for different lengths and capacitances

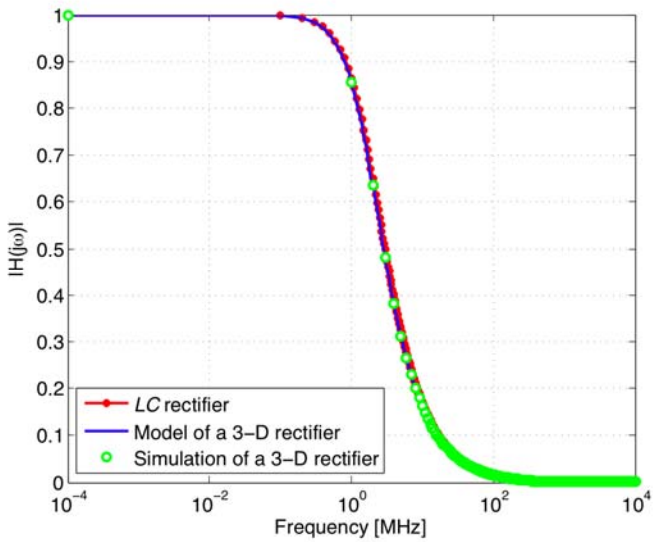


Figure 7.13 Simulation and model of the magnitude of the transfer function of a 3-D rectifier and a conventional *LC* rectifier circuit

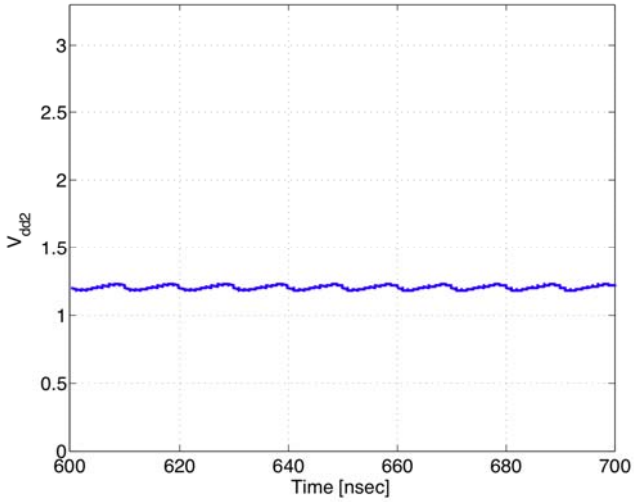


Figure 7.14 DC voltage at the output of a distributed rectifier

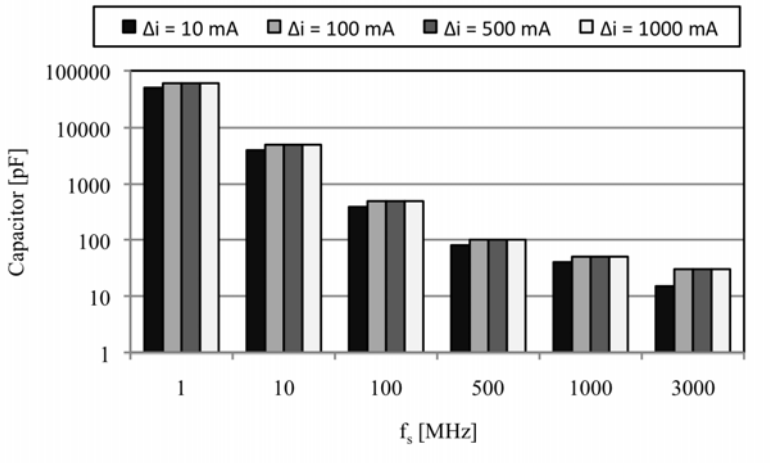
If additional current is required, the distributed rectifier can be extended by connecting multiple structures in parallel. For example, to enlarge this structure to produce 700 mA, 100 parallel structures are required. In this case, the effective resistance and inductance per unit length, listed in Table 7.1, and R_d are N times smaller, while the capacitance per unit length is N times larger. The multiple parallel rectifiers produce the same transfer function magnitude, as shown in Figure 7.12. The same interconnect lengths and capacitances are therefore chosen, producing a 1.2 volt power supply with 700 mA maximum current.

7.6 Performance Analysis

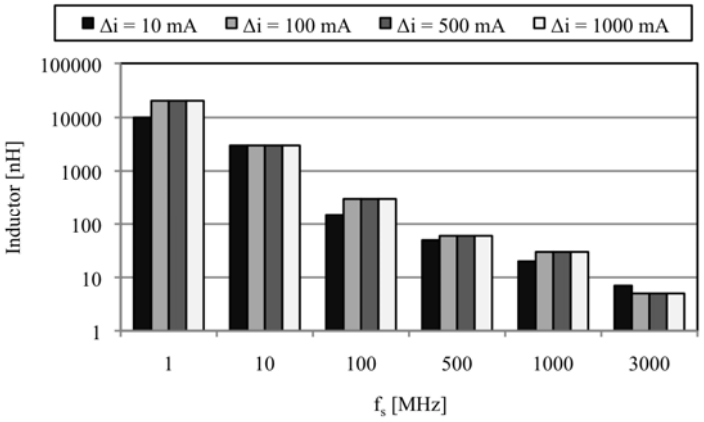
To obtain deeper insight into the performance and characteristics of the proposed rectifier, an analysis over a wide range of design parameters are presented in this section. In all cases, the rectifiers are designed to satisfy less than 5% voltage ripple with an input 3.3

volt DC voltage. In the simulation of a 3-D rectifier, the parameters and structure presented in the case study in section 7.5 are assumed.

To place these results in the context of a conventional rectifier, inductor and capacitor values as a function of switching frequency for different current loads are shown in Figure 7.15. As expected, the magnitude of the lumped inductor and capacitor decreases as the switching frequency increases. However, in the megahertz frequency region, the magnitude of the capacitor is between 100 pF and 100 nF, while the magnitude of the inductor is between 60 nH and 10 μH. In current technologies, these large inductors cannot be implemented on-chip. When the switching frequency shifts to gigahertz frequencies, the magnitude of the on-chip inductor becomes realizable on-chip. Note that in this case, the current load is assumed to affect only the output resistance of the power MOSFETs. Therefore, at higher current loads (above 100 mA), insignificant changes in the capacitance and inductor magnitudes can be observed in Figure 7.15.



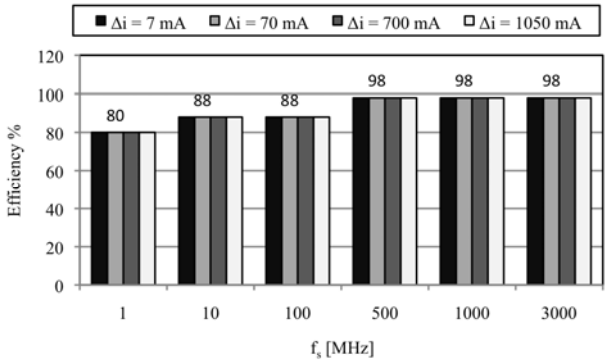
(a)



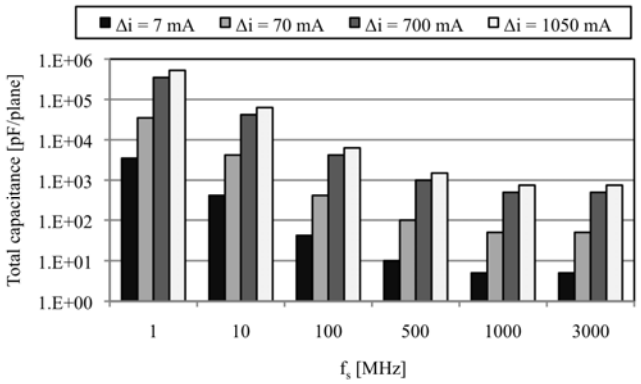
(b)

Figure 7.15 Conventional *LC* rectifier: (a) capacitance, (b) inductance

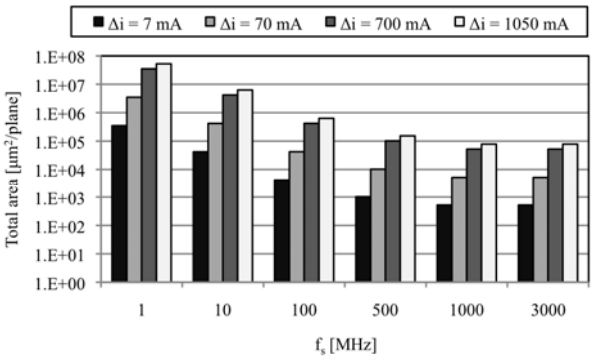
The efficiency, total capacitance, and area per plane of the distributed rectifier as a function of the switching frequency for different current loads are shown in Figure 7.16. At a 1 MHz switching frequency, the efficiency of the rectifier reaches 80%. Once the frequency is increased, the efficiency is significantly greater, reaching 98% at 500 MHz, 1 GHz, and 3 GHz. Since only the rectifier efficiency is considered, excluding the power MOSFETs, the efficiency at higher frequencies is only dependent on the input impedance of the distributed rectifier which does not change significantly at higher frequencies. The input impedance remains approximately the same at higher frequencies since the capacitors C_1 to C_n of the distributed rectifier contribute less to the input impedance at higher frequencies. As expected, the total capacitance and area occupied by the distributed rectifier decreases as the frequency increases. This behavior occurs since less capacitance and shorter interconnects are required to satisfy the target voltage ripple at higher frequencies.



(a)



(b)



(c)

Figure 7.16 3-D distributed rectifier: (a) efficiency, (b) total capacitance per plane, (c) total area per plane

The efficiency and area occupied by a conventional *LC* filter is omitted from this comparison. As shown in Figure 7.15(b), the magnitude of the required inductor within an *LC* filter varies between 50 and 10,000 nH for switching frequencies up to 1 GHz. At 3 GHz, the required inductor size is below 10 nH. This behavior implies that at switching frequencies used in current technologies, *i.e.*, up to several Megahertz, a conventional *LC* filter cannot be implemented on-chip.

7.7 Conclusions

Integrating DC-DC converters on-chip can significantly enhance the performance and reliability of integrated circuits. Conventional rectifiers, however, are difficult to integrate due to the large magnitude of the on-chip inductors and capacitors. A distributed buck converter rectifier for application to three-dimensional circuits is proposed in this chapter. By exploiting the low pass bandwidth properties of transmission lines, the efficient generation and distribution of power supplies to different planes is possible. An example converter based on the MITLL 3-D CMOS technology is described, demonstrating a DC-DC converter that can generate a 1.2 volt power supply while delivering 700 mA.

A performance analysis considering the rectifier efficiency, total capacitance, and physical area is performed for a wide range of switching frequencies and current loads. A conventional on-chip *LC* rectifier in the megahertz frequency range is not practical due to the large inductors. Alternatively, an on-chip distributed rectifier can be implemented in the megahertz frequency range, achieving high efficiency, suitable for 3-D integration.

Chapter 8

3-D Distributed Rectifier Test Circuit

Based on the design methodology of a distributed rectifier presented in Chapter 7, a test circuit has been developed and manufactured in the MIT Lincoln Laboratories (MITLL) 3-D SOI 0.15 μm CMOS technology. The objective of this test circuit is to demonstrate on-chip power conversion and distribution in 3-D circuits. In this chapter, the process used to design the test circuit is described. The analysis, design, and simulation of a distributed rectifier to achieve a target performance are described in section 8.1. In section 8.2, auxiliary circuits used to emulate significant current load requirements and sense the power supply noise are described. A review of the MITLL 3-D fabrication process is presented in section 8.3 followed by the physical layout of the distributed rectifier in section 8.4. The chapter is concluded in section 8.5.

8.1 Analysis and Design of a Distributed Rectifier

The target performance of the on-chip DC-DC converter test circuit is strongly dependent on the allocated area, which in this test circuit is 500 × 500 μm^2. The intended voltage conversion is from 3.3 to 1.5 volts switching at a frequency f_s = 100 MHz, while the required maximum current load is 100 mA. The voltage ripple is chosen to be ± 2.5% of the output power supply. The distributed rectifier circuit schematic, shown in Figure 8.1, is composed of a network of interconnects and capacitors, spanning from the top, plane C, to the bottom, plane A. Each interconnect is followed by a shunted connected capacitor on each

plane. The three planes are connected by 3-D vias (vertical interconnects). The distributed rectifier filters the high frequency harmonics of the input AC signal, producing a DC voltage with residual harmonics at the output.

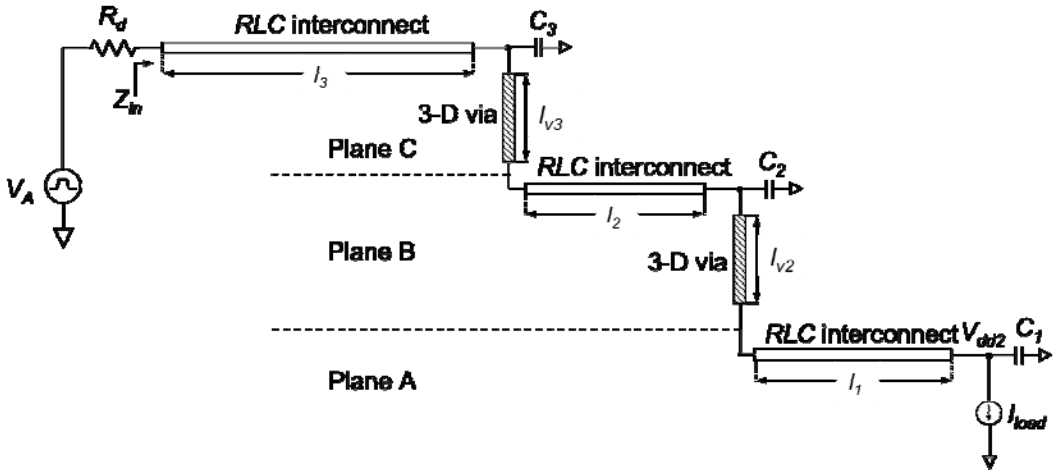


Figure 8.1 A distributed rectifier of a DC-DC power converter in a three-plane structure

The maximum current density in this technology is [156]

$$J_{\max} = 3\,\text{mA}/\mu\text{m}^2. \tag{8.34}$$

The cross sectional area of a 3-D via in this 150 nm CMOS technology is 1.25 × 1.25 μm^2, resulting in 5 mA maximum current flowing through a single via. Since in this technology the interconnect thickness is 630 nm, a current of 5 mA requires a 2.5 μm interconnect width. Note that the 3-D via is 7.34 μm long, connecting the three planes as illustrated in Figure 8.1. In the initial design phase, the current load profile is assumed to be as shown in Figure 8.2 with the following characteristics: $1/T_{CLK}$ = 1 GHz, and t_{rc} and t_{cf} are $0.3T_{CLK}/2$ and $0.7T_{CLK}/2$,

respectively. The resistance, inductance, and capacitance per unit length of the interconnect are extracted based on the predicative technology model (PTM) [157], while the electrical parameters of the 3-D vias are extracted in [158], as listed in Table 8.1.

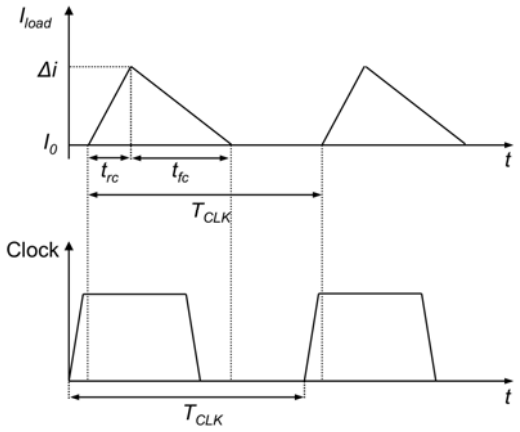


Figure 8.2 Current load profile

Table 8.1 Electrical parameters of interconnects and 3-D vias

	R [mΩ/μm]	L [nH/μm]	C [pF/μm]
Interconnect [157]	21	1.4	207
3-D via [158]	20.4	0.00055	0.00037

To ensure the voltage ripple of the output power supply is 75 mV (± 2.5%), the magnitude of the transfer function at 100 MHz has to be equal or less than 0.0183 (see Chapter 7) according to

$$\left|H_{rect}\left(j\omega_s\right)\right| = \frac{v_r \cdot V_{dd1}}{4\left|a_1\right|}, \tag{8.35}$$

$$a_1 = \left(\frac{V_{dd1}T_s}{4t_r\pi^2}\right)\cdot\left(e^{-j\omega_s t_r}\left(1-e^{-j2\pi D}\right)+e^{-j2\pi D}-1\right) \tag{8.36}$$

The magnitude of the transfer function at a switching frequency of 100 MHz for different capacitances and interconnect lengths, assuming $l = l_1 = l_2 = l_3$ and $C = C_1 = C_2 = C_3$, is shown in Figure 8.3. Note that in this case, R_d is 26.8 Ω, respectively. The area below the dashed line in Figure 8.3 indicates the acceptable interconnect length and capacitance to satisfy the required voltage ripple. Due to area constraints, the interconnect lengths and capacitors are l = 400 μm and C = 50 pF, respectively. In this example, the required duty cycle is 0.65 (see Chapter 7) according to

$$D_{PWM} = D \cdot \left| 1 + \frac{R_d}{Z_{in}(0)} \right|. \tag{8.37}$$

Since this circuit is designed to provide 5 mA, 20 rectifier circuits on each plane are connected in parallel to provide 100 mA current to the third plane.

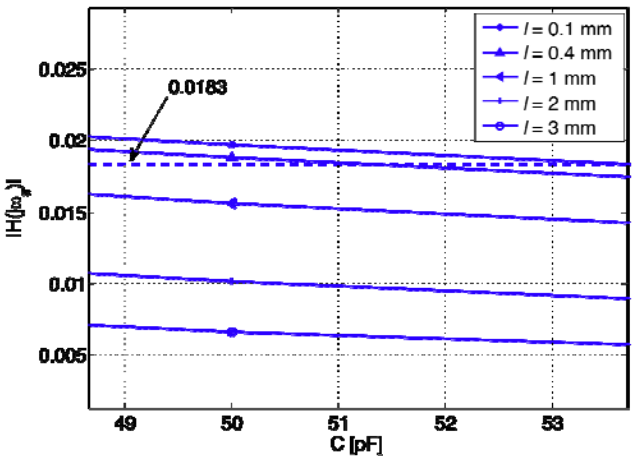


Figure 8.3 Magnitude of the transfer function for different line lengths and capacitor magnitudes

8.2 Design of Current Load, Control, and Noise Measurement Circuits

To emulate a practical circuit with a maximum 100 mA current load, several switched current load circuits have been developed. Switches, as shown in Figure 8.4, control these current loads. Note that the bias current I_{bias} is supplied off-chip. The pseudo–random current activity is produced by pseudo-random number generators (PRNG), activating different switches at random times, as depicted in Figure 8.5. To reduce the complexity of the test circuit and avoid the design of a complicated clock distribution network, 250, 500 and 1000 MHz ring oscillators have also been developed. These ring oscillators provide three clock signals to the pseudo-random number generators. Four PRNG have been designed with different bit lengths. Each PRNG spans $2^n - 1$ different states, where n is the number of bits.

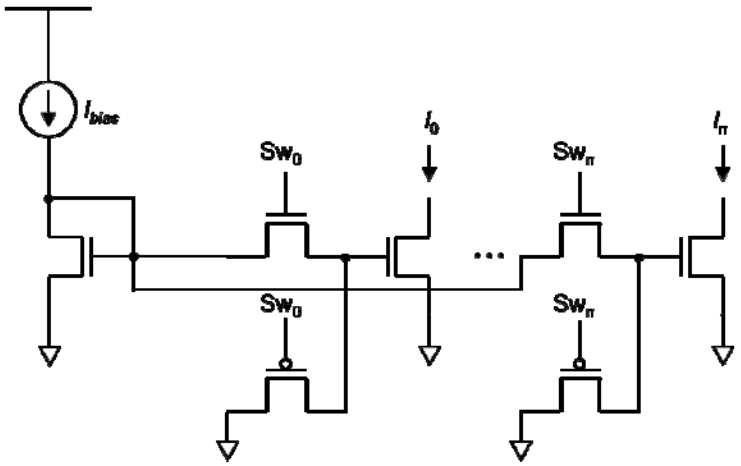


Figure 8.4 Current load controlled by switches

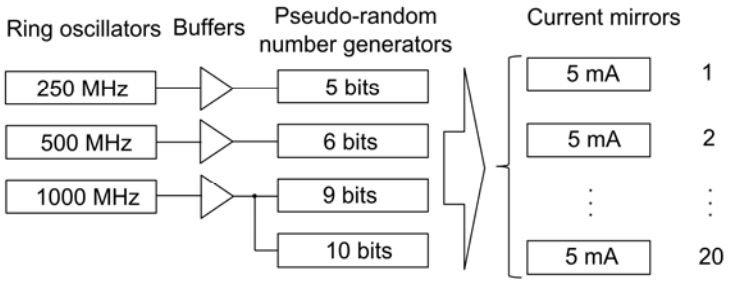


Figure 8.5 Control logic for the current mirrors

To consider post-fabrication process variations and enable adjustment of the output DC voltage level V_{dd2}, a control circuit has been designed, as shown in Figure 8.6. The control signal V_{cont} adjusts the additional amount of current drawn from the primary 3.3 volt power supply to ensure that the output voltage remains at the required 1.5 volt DC voltage. An additional mechanism to control the output power supply of the distributed rectifier is to vary the duty cycle of the input signal V_A (see Figure 8.1).

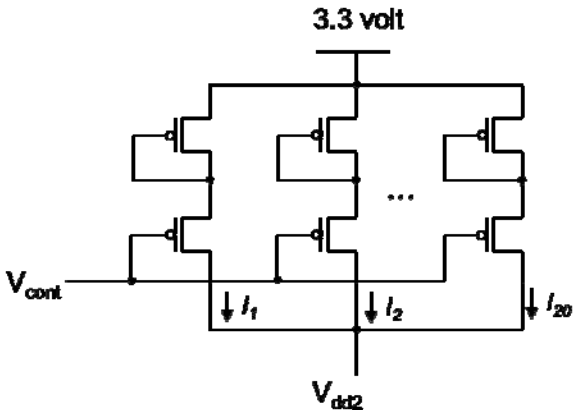
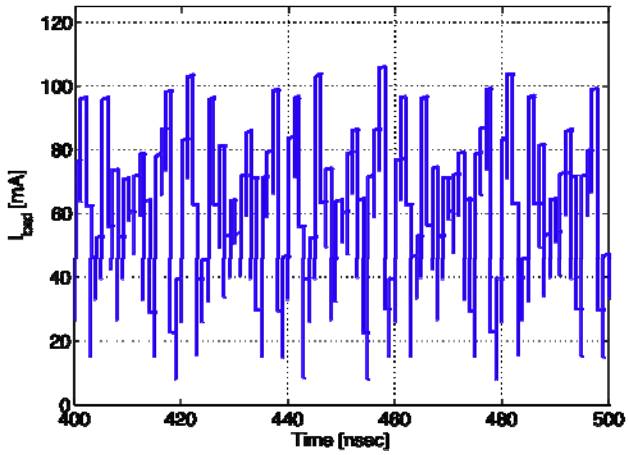
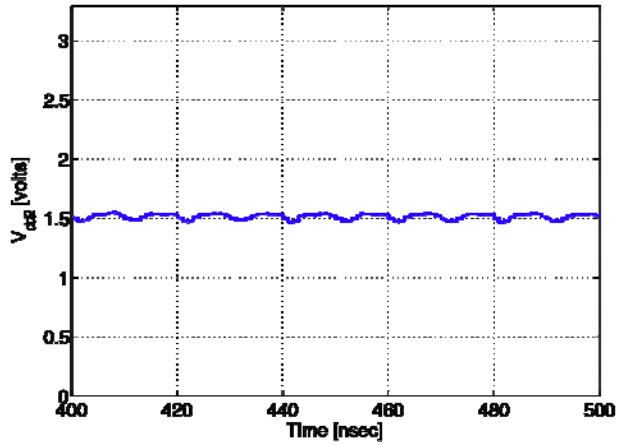


Figure 8.6 Post-fabrication control circuit



(a)



(b)

Figure 8.7 Distributed rectifier: (a) current load, (b) output voltage supply

The simulated current load and output voltage power supply are shown in Figure 8.7. As evident from this simulation, at a maximum current load of 100 mA, the rectifier provides a power supply voltage within the required 75 mV voltage ripple. In this case, the voltage ripple of the output power supply is 67 mA.

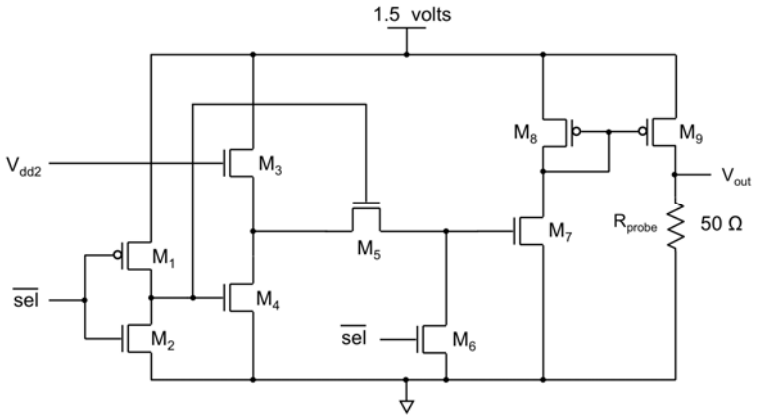


Figure 8.8 Power supply noise detector

The on-chip power supply noise measurement circuit is shown in Figure 8.8 [159]. When the circuit is enabled, transistors M_2 and M_6 turn off while transistors M_1 and M_5 turn on, biasing the source follower amplifier (transistors M_3 and M_4). Transistor M_3 is connected to the output of the rectifier (V_{dd2} in Figure 8.1), sensing the voltage fluctuations. The output of the source follower is fed to the input of the common source amplifier (transistor M_7 and M_8). Note that the source follower biases the common source amplifier at the required bias voltage while providing the sensed voltage signal. A current flowing through transistor M_8 expresses the amplified voltage signal mirrored to the output by transistor M_9. This current is converted to an output voltage signal by the 50 Ω resistance of the probe, representing the measured power supply noise. In order to determine the amplification of this circuit, an identical power supply noise detector circuit with transistor M_3 is connected to an external sinusoidal signal with a 1.5 volt DC voltage. The accuracy of this noise detector circuit is less then 1 mV.

8.3 MIT Lincoln Laboratories 3-D Fabrication Process

MIT Lincoln Laboratory has developed a three-dimensional (3-D) integrated circuit technology in which circuit structures formed on several silicon-on-insulator (SOI) substrates are combined into a 3-D integrated circuit. This 3-D circuit integration technology is comprised of fully depleted SOI circuit fabrication, low temperature wafer-to-wafer oxide bonding, precision wafer-to-wafer alignment, and electrical connection of the circuit structures with vertical interconnections (3-D vias) [156].

The first phase in this 3-D circuit integration process is the fabrication of three fully depleted SOI tiers (or planes), as shown Figure 8.9. In the second phase, wafer 2 is inverted, aligned, and bonded to wafer 1, as shown in Figure 8.10. The handle silicon is removed from tier 2 and the 3-D vias are etched through the oxide layers of tiers 1 and 2, landing on metal pads in tier 1. Tungsten vias are deposited and planarized using chemical mechanical polishing (CMP). The structure shown in Figure 8.11 is a two-tier assembly with electrical connections between the top-level metal of tier 2 and the top-level metal of tier 1. The following phase is the inter-tier via, back side via, and back side metal formation, as shown in Figure 8.12. In the next phase, 3-D vias connect the top-level metal of tier 3 to the back side metal of tier 2, as shown in Figure 8.13. In the final phase, back side metal, bond pads, and heat sink cuts are formed on the top of tier 3. The back side metal on tier 3 permits probing, wire bonding, and cooling of the entire structure. The completed 3-D assembly is shown in Figure 8.14. A list of layer thicknesses is presented in Figure 8.15 [156].

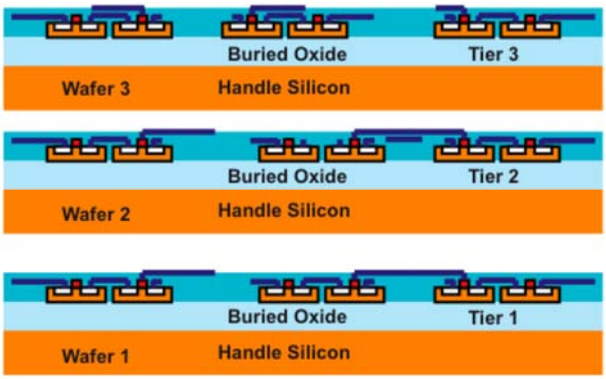


Figure 8.9 Three independent tiers fabricated using a conventional 2-D technology [156]

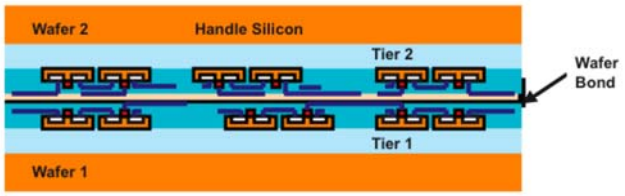


Figure 8.10 Tier 2 aligned and bonded to tier 1 [156]

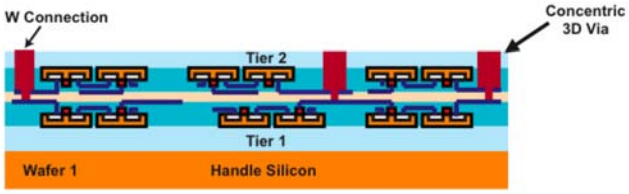


Figure 8.11 Tier 2 electrically connected to tier 1 with Tungsten 3-D vias [156]

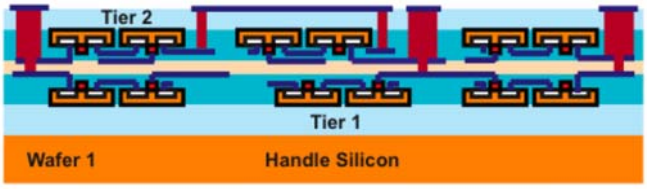


Figure 8.12 Formation of the back side vias and metal on tier 2 [156]

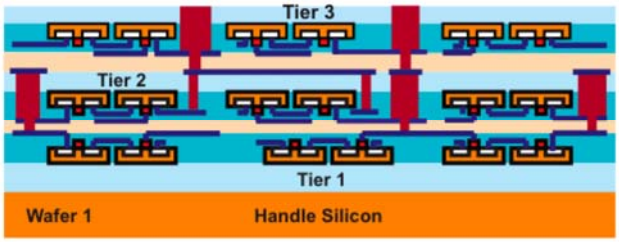


Figure 8.13 Tier 3 is aligned and bonded to the back side metal of tier 2 [156]

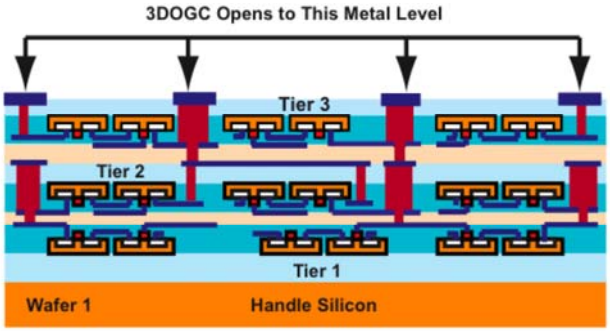


Figure 8.14 Formation of the back side metal on tier 3 [156]

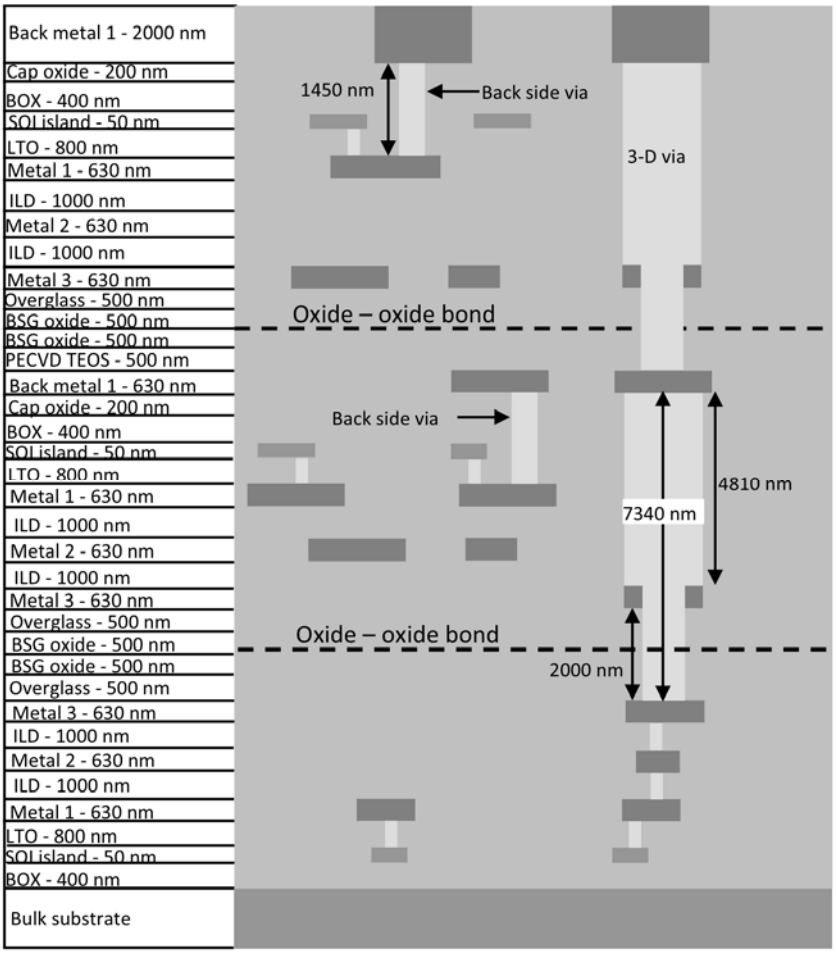


Figure 8.15 Layer thicknesses within the three-tier structure [156]

8.4 Rectifier Layout Design

Based on the results described in section 8.1, a three-plane layout of the distributed rectifier has been designed. The test circuit layout (note that the three planes are overlaying) is shown in Figure 8.16, occupying a total area of 2 × 2 mm^2. The top two and the bottom left blocks have been dedicated to evaluating 3-D power distribution networks, while the bottom right block is the 3-D rectifier. An illustration of the distributed rectifier, as well as the enlarged layout circuit is shown in Figures 8.17 and 8.18, respectively. The power distribution networks are not discussed in this thesis.

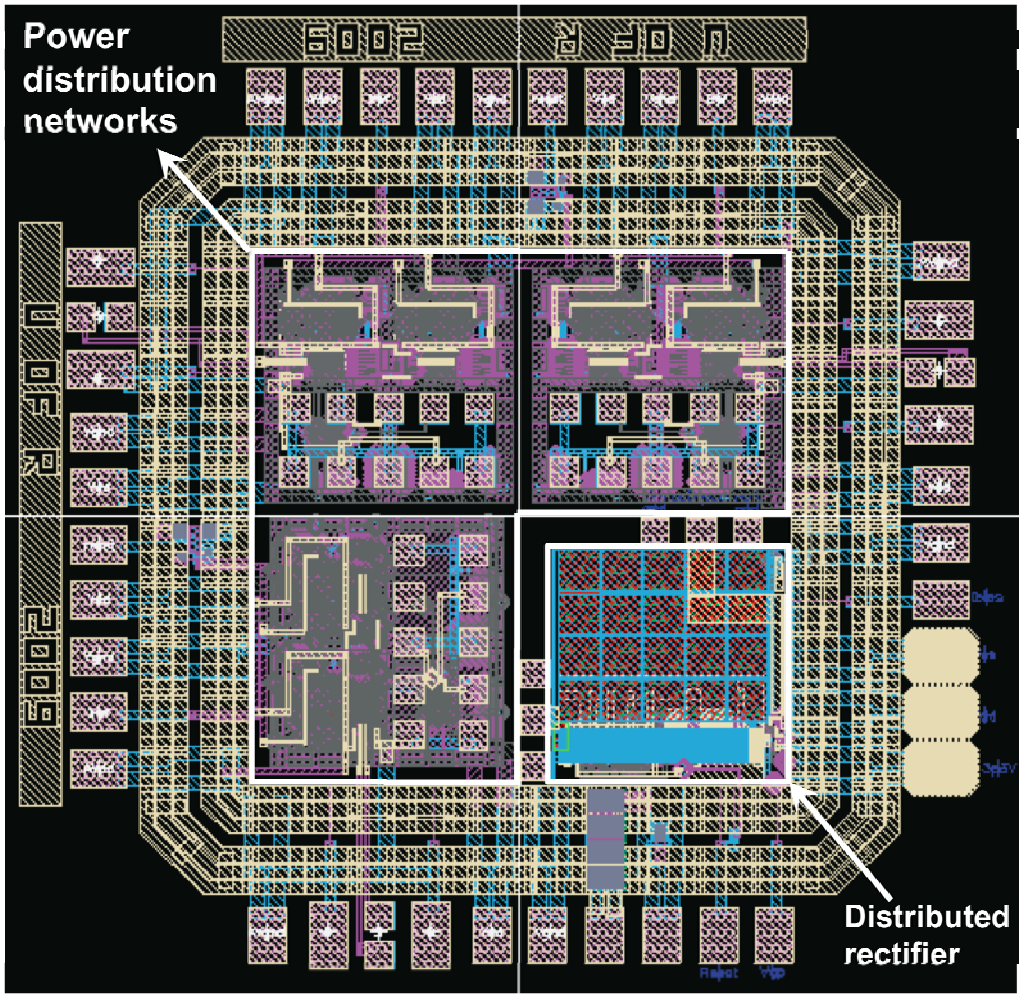


Figure 8.16 Four blocks within the power delivery test circuit

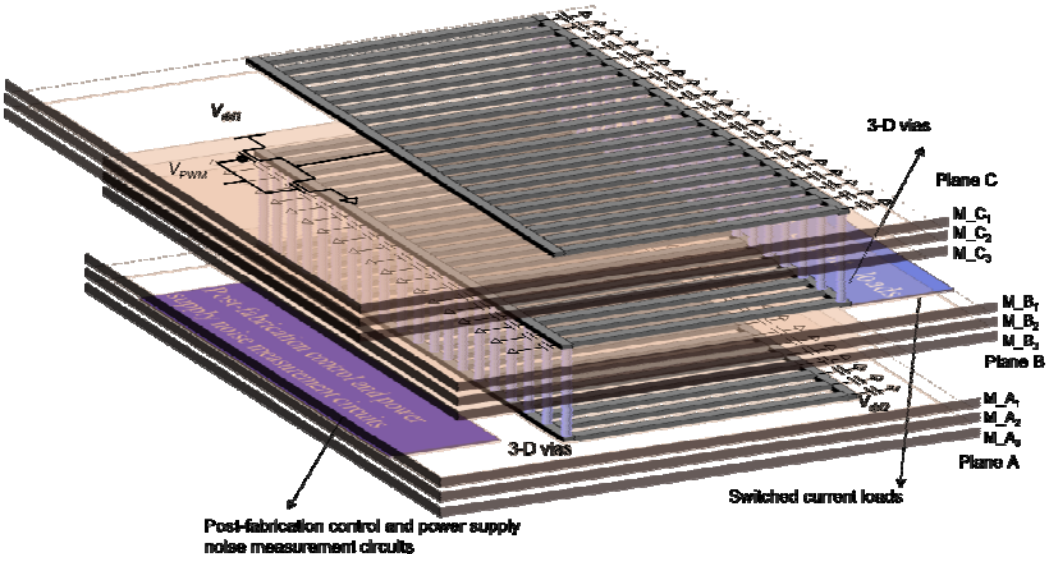


Figure 8.17 Schematic illustration of the distributed rectifier circuit

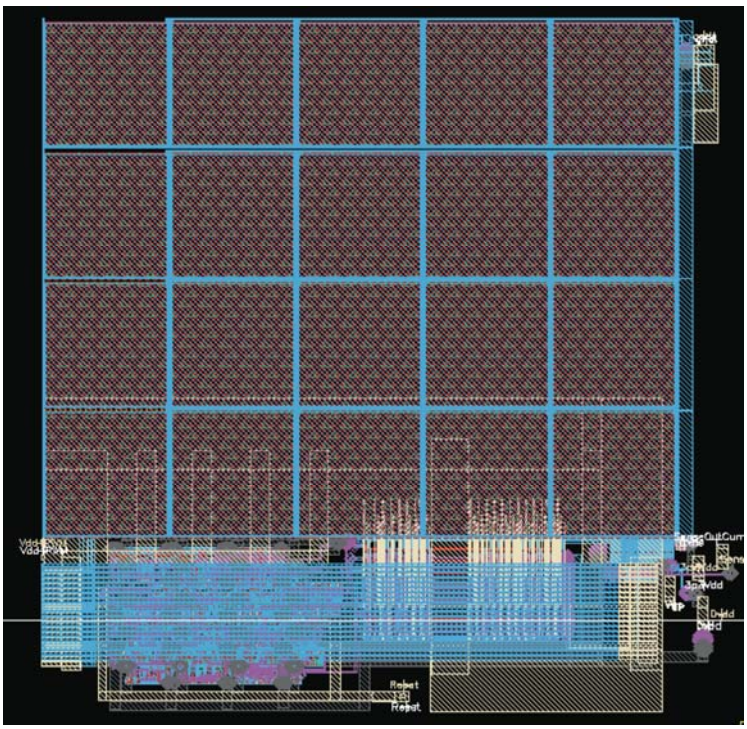
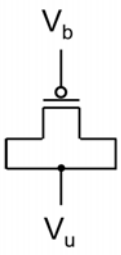


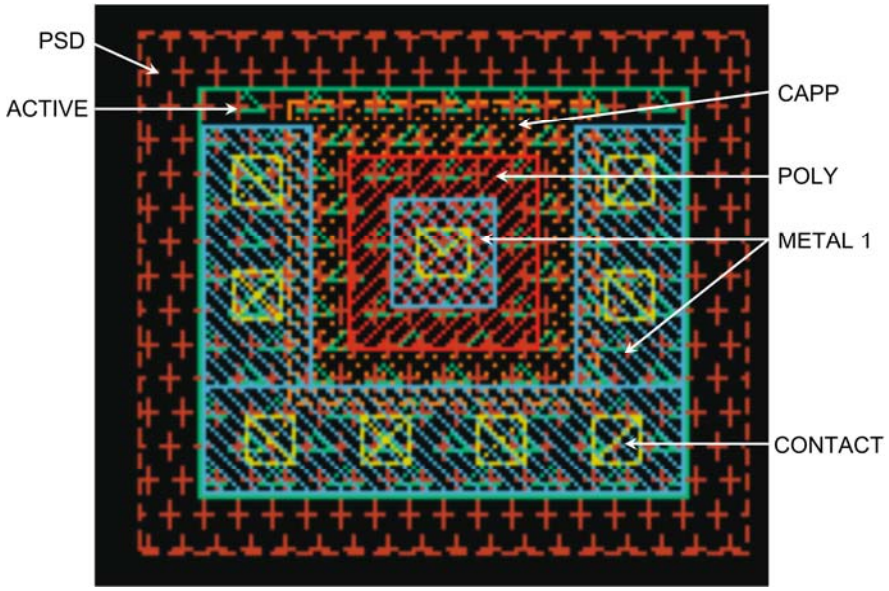
Figure 8.18 Physical layout of the 3-D rectifier circuit (all three planes)

As shown in the upper part of Figure 8.18, a large portion of the area is occupied by capacitors. On each plane, twenty 50 pF capacitors are connected in parallel, where each square illustrated in Figure 8.18 is a single 50 pF capacitor occupying an area of 88 × 88 μm^2. Doped silicon and polysilicon layers are used to form capacitors in this technology, similar to forming active devices. A model and layout of a p-type capacitor are shown in Figure 8.19. A p^+ implant layer (PSD) is used to dope the polysilicon layer of the capacitor. The PSD is also used to form p-type transistors. The next layer is a p-type capacitor bottom plate implant layer (CAPP) with a doping density of about 5×10^{18} cm^{-3}. This layer is suitable for forming the bottom plate of the low bias voltage capacitors and high value resistors. The bottom (gate) plate of the capacitor is formed by the polysilicon layer, which is connected to Metal 1. The upper plate of the capacitor is formed by the p^+ implant layer (PSD) and connected to Metal 1, resulting in a transistor-like structure with the drain and source terminals shorted. The capacitor value is determined from the voltage difference between the bottom and upper plates. The relationship between the voltage difference and the resulting capacitance has been graphically provided based on measurements in [160].

The three planes of the rectifier, C (upper), B (middle), and A (bottom), are shown in Figure 8.20. Different circuitry is located on each plane, adjacent to the capacitors,. Twenty 400 μm long interconnects and analog sensing circuitry are located on plane A (Figure 8.20(a)). Ring oscillators and current mirrors are located on plane B (Figure 8.20(b)) while the interconnects have been removed due to insufficient area. Alternatively, back side metal on plane B is used to distribute the current to plane A. Twenty 400 μm long interconnects connected in parallel are located on plane C (Figure 8.20(c)). The enlarged portion of the control circuit and power supply noise detector on plane A is shown in Figure 8.21 while the ring oscillators and current loads on Plane B are shown in Figure 8.22.

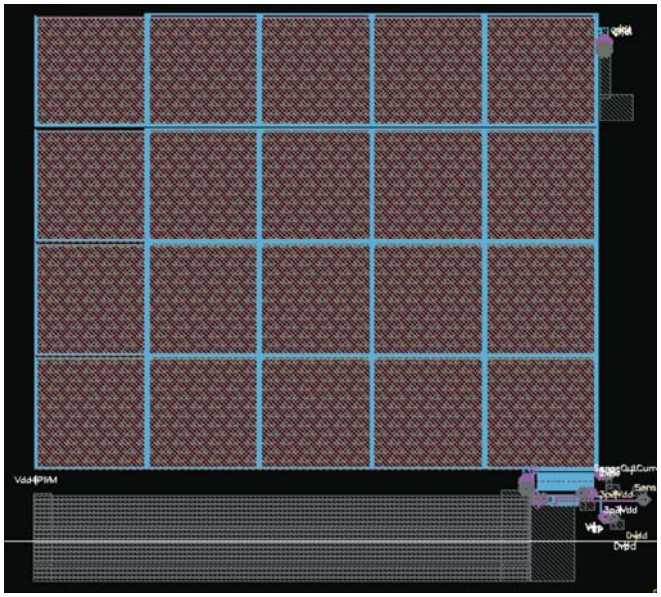


(a)

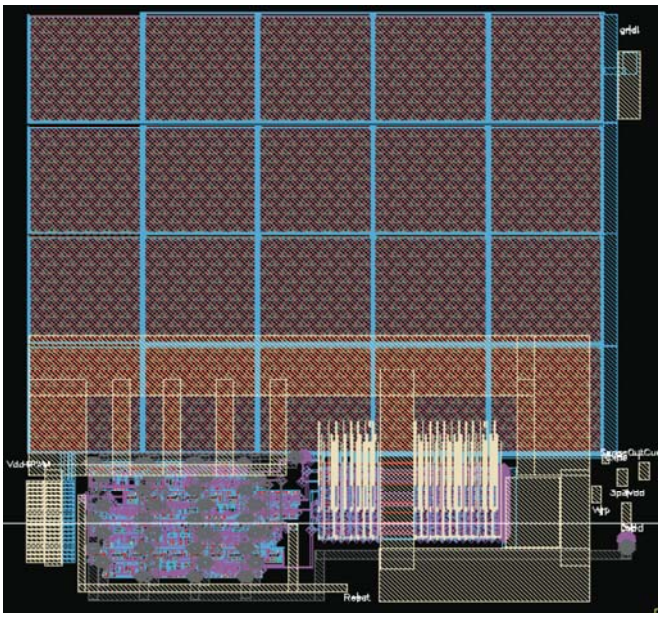


(b)

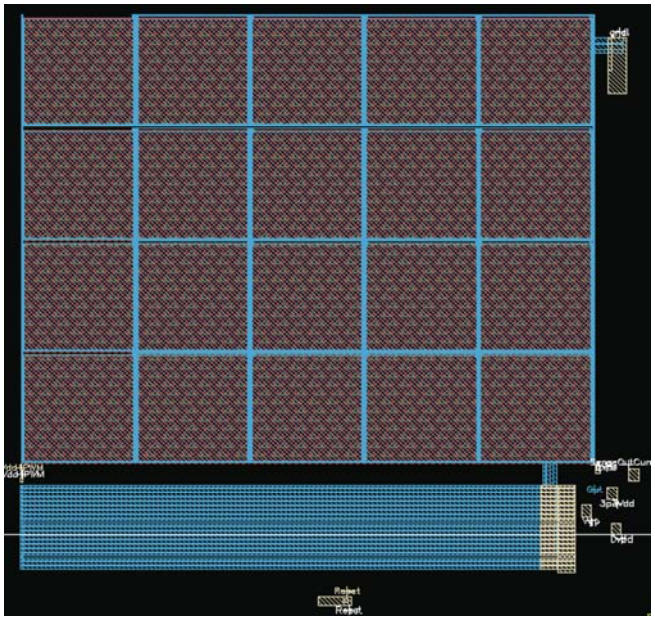
Figure 8.19 P-type capacitor: (a) model, (b) layout



(a)

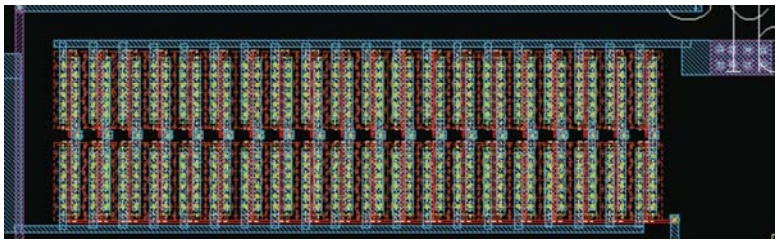


(b)

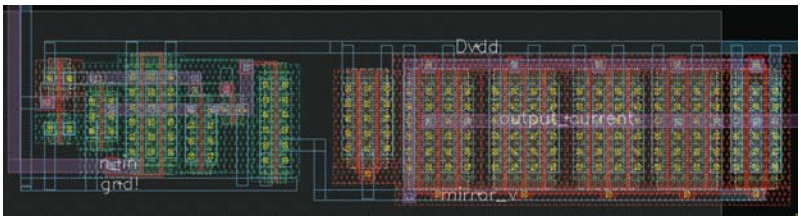


(c)

Figure 8.20 Distributed rectifier structure: (a) plane A (bottom), (b) plane B (middle), (c) plane C (upper)

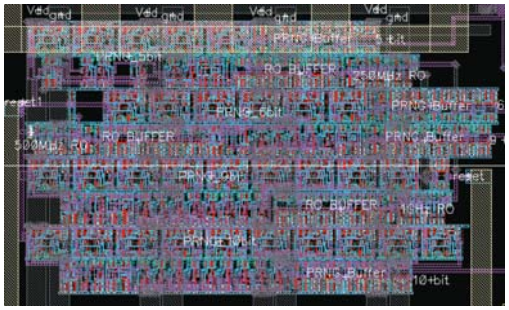


(a)

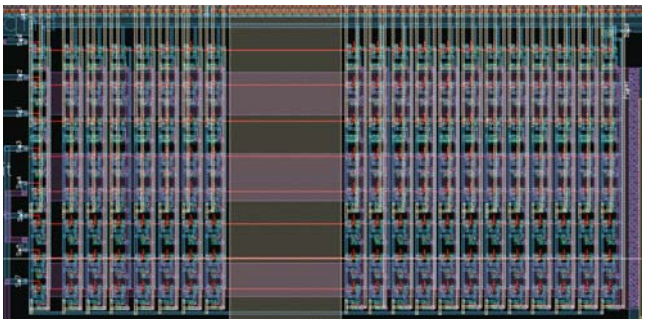


(b)

Figure 8.21 Circuitry on plane A: (a) post-fabrication control circuit, (b) power supply noise measurement circuit



(a)



(b)

Figure 8.22 Circuitry on plane B: (a) ring oscillators and buffers, (b) switched current loads

8.5 Summary

A distributed rectifier for use within a DC-DC converter has been designed in the MITLL 3-D CMOS technology. Using the methodology developed in Chapter 7, the distributed rectifier design is demonstrated based on an existing 3-D technology. Current loads that emulate the random current activity of a processor and power supply noise detector circuits have also been incorporated, enabling on-chip noise detection and measurement. A review of the MITLL 3-D fabrication process has also been provided in this chapter followed by the layout design of the different components within the rectifier circuit. The on-chip capacitor fabrication process technology is also described. The measurement of this test circuit will demonstrate the proposed 3-D rectifier concept, providing insight into exploiting a 3-D integration process. Once the concept is proven, the distributed rectifier circuit will support the on-chip integration of DC-DC power converters.

Chapter 9

A Hybrid Methodology for Wide Range DC-DC Conversion in 3-D Circuits

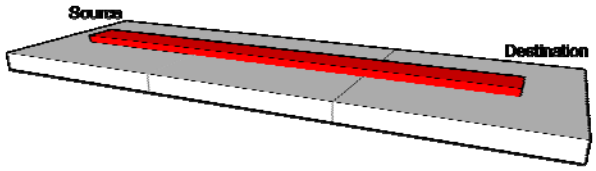
A novel methodology for DC-DC conversion in 3-D circuits is described in this chapter. The proposed approach exploits both linear and switching buck converters within different conversion ranges, thereby increasing the power efficiency. Depending upon the required voltage conversion, an appropriate converter is used. By replacing the traditional *LC* filter of a switching converter with a distributed rectifier, a further increase in efficiency is demonstrated. Design guidelines and expressions are developed, achieving good agreement with simulations based on the MIT Lincoln Laboratories CMOS/SOI 180 nm 3-D technology. The proposed converter distributes 2.5 amperes maximum current load, achieving conversion from 3.3 volts to 2.5 volts and 1 volt with 74% and 44% power efficiency, respectively.

9.1 Introduction

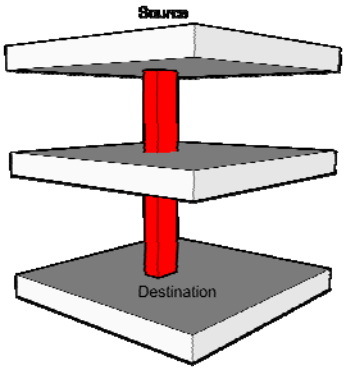
Power dissipation is a primary performance bottleneck in high complexity nanoscale integrated circuits which is exacerbated by the increase in current loads and clock frequencies. Current levels, transistor densities, and transient waveforms of high performance microprocessors have increased by 50% with every advance in technology node despite scaling power supply voltages [153]. To battle the significant increase in power dissipation, a multiple voltage power supply paradigm has been incorporated in large scale circuits. Multi-

voltage circuits exploit the delay differences among the different signal propagation paths by selectively lowering the supply voltage of the gates on the non-critical delay paths while maintaining a higher supply voltage on the critical delay paths in order to satisfy a target clock frequency [139], [161].

To fundamentally reduce power dissipation and latency in highly complexity circuits, an emerging three-dimensional (3-D) integrated circuit technology is under development. In a 3-D technology, individual planes of 2-D integrated circuits are combined into cubes of 3-D structures, increasing density and functionality, as illustrated in Figure 9.1. In this example, the distance between the source and destination in a 2-D circuit is significantly reduced once a plane is partitioned and stacked into a 3-D structure.



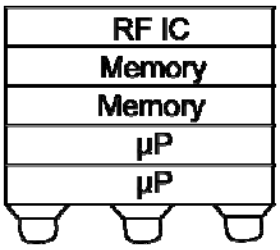
(a)



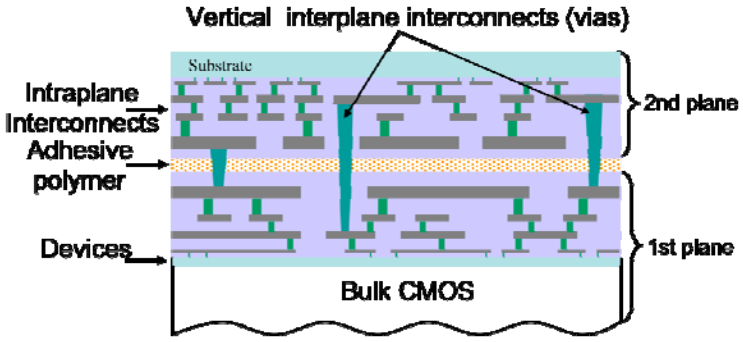
(b)

Figure 9.1 2-D and 3-D cells: (a) 2-D horizontal plane, (b) three stacked planes

Three-dimensional integration offers novel architectural structures for microprocessors, as shown in Figure 9.2. A multi-plane structure enables the efficient distribution of different portions of high performance, high complexity systems, such as memory, RF, and digital circuits. Additionally, heterogeneous technologies can be integrated onto different planes on the same die, thereby improving performance.



(a)



(b)

Figure 9.2 3-D circuit: (a) multi-core processor combined from stacked planes, (b) cross section of a 3-D circuit

The need for multiple on-chip supply voltages in 2-D circuits has become important due to the integration of low power, high speed digital circuits with analog/RF circuits on the same die. System heterogeneity offered by 3-D circuits has exacerbated the requirement for

multiple, wide range, and well controlled power supplies. Each plane will support a variety of functions and applications, such as MEMs, RF, analog, memory, and high speed digital circuits. A system of distributed on-chip power supplies is therefore required, as shown in Figure 9.3. In this example, V_{dd1} and V_{dd3} are delivered to the first and third planes, respectively. The second plane, however, is partitioned into three different voltage domains: $V_{dd2,a}$, $V_{dd2,b}$, and $V_{dd2,c}$. To support a multiple voltage domain system, efficient on-chip DC-DC converters are required.

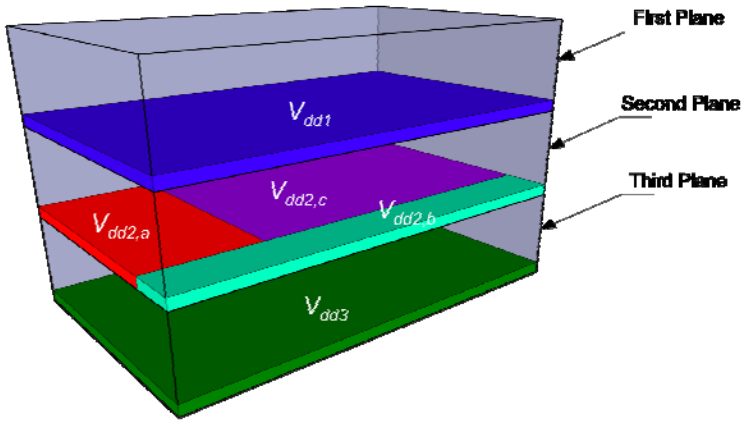


Figure 9.3 Different partitions of voltage domains within a 3-D circuit

The on-chip integration of DC-DC converters is highly efficient from a power and I/O perspective, particularly in 3-D circuits [154]. By integrating the entire buck converter on-chip, the number of I/O pins is reduced, decreasing the cost and size of a circuit package. The large external resonant voltage swing, produced by parasitic impedances from the package, is also eliminated [162]. Supply voltages locally produced on-chip can be dynamically controlled to compensate for on-chip temperature, process, and load variations. Furthermore,

in a typical off-chip DC-DC converter, significant power is dissipated by the parasitic impedances of the interconnect among the non-integrated devices.

A methodology for designing on-chip DC-DC converters for application to 3-D circuits is described in this chapter. The proposed scheme is a combination of linear and switching converters, forming a high efficiency hybrid conversion system. The hybrid circuit exploits the benefits of both types of DC-DC converters, depending upon the required power supply.

This chapter is organized into five sections. A description of three common types of buck converters as well as recent work on on-chip converter circuits are described in section 9.2. In section 9.3, a methodology for designing novel linear and switching converters in 3-D circuits is described, followed by two case studies in section 9.4. In section 9.5, a performance comparison between the proposed circuit and other on-chip DC-DC converters is presented. Finally, the chapter is concluded in section 9.6.

9.2 Background

Converting AC or DC power supplies into lower or higher DC power supplies has been a topic of interest since the invention of modern electrical networks. For integrated circuit applications, converting a high DC power supply into a lower DC power supply (called buck conversion) is of primary interest. Three common types of DC-DC converters are linear, switched-capacitor, and switching converters.

Linear Converters

A linear converter, also called a low dropout regulator (LDO), is effectively a voltage divider, as shown in Figure 9.4. The output voltage V_{dd2} is determined by the ratio of the

impedances of the power MOSFET M_1 and the load. An amplifier in a feedback loop senses the output DC voltage V_{dd2} and appropriately adjusts the gate voltage of transistor M_1 to provide the required current to the load. Note that transistor M_1 is typically very wide in order to conduct a large amount of current.

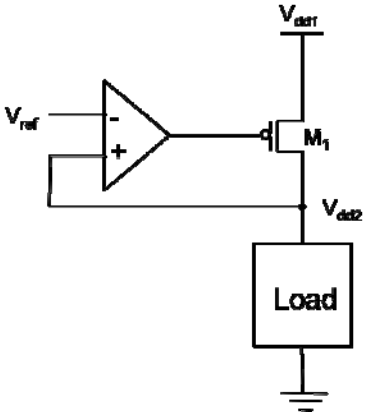


Figure 9.4 Linear converter

The design simplicity of a linear converter is attractive; however, at lower conversion ratios, the power efficiency exhibited by this type of converter is relatively low since the maximum power efficiency $\eta_{linear,max}$ is determined by the ratio between the output and the input DC voltages,

$$\eta_{linear,\max} = \frac{V_{out}}{V_{in}} = \frac{V_{dd2}}{V_{dd1}}. \tag{9.1}$$

Due to the maximum efficiency imposed by (9.1), the current efficiency is of interest in linear converters,

$$\eta_{linear} = \eta_{linear,\max} \cdot \eta_{current}, \tag{9.2}$$

where

$$\eta_{current} = \frac{I_{out}}{I_{in}}. \tag{9.3}$$

For high conversion ratios, however, an LDO is an attractive approach. In [163], conversion from 1.2 volts to 0.9 volts with 94% current efficiency is demonstrated, resulting in 71.25% power efficiency while delivering 100 mA maximum current. In this circuit, the LDO achieves a small dropout voltage and fast load regulation due to a single-stage feedback loop, rapidly adjusting the PMOS power transistor. Alternatively, an LDO with an impedance-attenuated buffer is described in [164]. Dynamically-biased shunt feedback is used, achieving stability of the LDO over a wide range of current loads. In this circuit, the input voltage ranges from 2 volts to 5.5 volts, while the minimum output DC voltage is 1.8 volts. With 99.8% current efficiency, the highest and lowest power efficiencies are 89% and 32%, respectively, delivering 200 mA maximum current.

Switched-Capacitor Converters

The operation of a switched-capacitor converter is based on periodically charging/discharging the charge pump capacitors through resistive switches. These switches contribute to the power losses of the converter. Additionally, a switched-capacitor converter exhibits poor output voltage regulation. To maintain a stable output voltage, the use of high power feedback control circuitry is required which decreases the power efficiency. A

switched-capacitor convertor can typically provide moderate levels of efficiency for small conversion ratios, while providing low current loads. The output resistance limits the peak power efficiency of a capacitive converter; therefore, the efficiency of a switched-capacitor converter increases as the current load decreases [165].

A controllable DC-DC converter that combines a switched-capacitor voltage divider and a linear regulator is described in [166]. The proposed circuit converts from 2.5 volts to as low as 0.65 volts. A hybrid switched-capacitor and linear converter achieves higher power efficiency than a traditional linear converter. In this circuit, the switched capacitor converter generates a $V_{dd}/2$ power supply with relatively high efficiency. Depending upon the required DC output voltage, a high voltage (with V_{dd} input power supply) or low voltage (with $V_{dd}/2$ input power supply) linear converter is chosen using programmable control circuitry. The highest and lowest power efficiencies achieved by the proposed circuit when converting from 2.5 volts to 2.35 volts and 0.6 volts are 93% and 25%, respectively, delivering 100 mA maximum current.

Switching Converters

A switching converter is the most common converter implemented off-chip due to the high efficiency produced by this converter. The working principle of this converter is based on an AC signal produced by the power MOSFETs. Depending upon the duty cycle and the input power supply, a specific DC component of the signal is passed to the output by a second order *LC* filter. Due to the nonidealities of the filter, residual high harmonics are also passed. The fundamental difficulty to fully integrate a switching converter on-chip is the high magnitude and area occupied by the *LC* filter. A detailed description of a switching buck

converter as well as a novel distributed rectifier that eliminates the need for an on-chip inductor is provided in Chapter 7.

9.3 Design Methodology

The proposed design methodology generates and distributes numerous power supplies based on a single input power supply within a 3-D cube, as illustrated in Figure 9.3. By utilizing a combination of linear and switching buck converters, high efficiency is achieved over a wide range of conversion ratios. A distributed rectifier (see Chapter 7) within the switching buck converter is used to further improve power efficiency.

In section 9.3.1, the efficiency of several 2-D and 3-D converters is evaluated. Based on the observations drawn from section 9.3.1, a novel linear converter is proposed in section 9.3.2. In section 9.3.3, a switching converter with a distributed rectifier is described.

9.3.1 Evaluation of Switching Converter Efficiency

Since energy losses during power conversion is a significant concern, the primary focus of the proposed methodology is on increasing converter efficiency over a wide range of conversion ratios. Based on [167], the efficiency of a switching converter in a 180 nm (curve a) and 130 nm (curve b) 2-D technology as well as an ideal linear converter (curve c) and a switching converter with a distributed filter (curve d) based on the MIT Lincoln Laboratory 180 nm 3-D technology [156] are depicted in Figure 9.5. The efficiency of the switching converter excluding the filter losses is also shown as a reference (curve e). In the 180 nm

CMOS 2-D technology (curve a), the efficiency of an ideal linear converter (curve c) surpasses the efficiency of a conventional switching converter above a 0.5 conversion ratio. In the 130 nm technology (curve b), however, the efficiency of an ideal linear converter is higher than a conventional switching converter above a 0.65 conversion ratio. The low efficiency of a switching converter can be attributed to the dynamic power losses of the power MOSFETs (and the drivers) and the inductor losses within the *LC* filter. With advances in technologies, however, switching converters tend to achieve higher efficiencies than a linear converter over a wider range of conversion ratios due to the decreased parasitic capacitance of the power MOSEFTs and the improvement in the on-chip inductor.

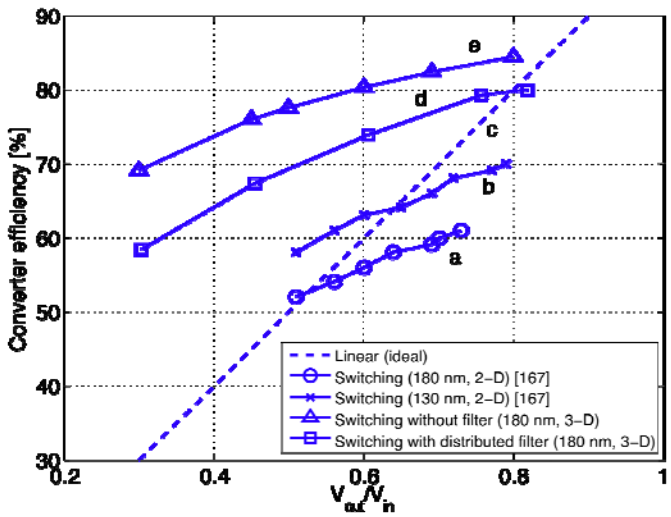


Figure 9.5 Efficiency as a function of conversion ratios in different technologies

As a reference, consider the efficiency of an optimal switching converter excluding the filter losses (curve e). As evident from Figure 9.5, the associated filter losses significantly degrade the converter performance. The efficiency of an optimal switching converter

employing the distributed 3-D rectifier (see Chapter 7) is also shown in Figure 9.5 (curve d). The distributed rectifier increases the efficiency of a switching rectifier as compared to an ideal linear and a conventional switching converter in both a 180 nm and 130 nm technology. Note that at a 0.8 conversion ratio, the power efficiency of the linear converter surpasses the efficiency of the switching converter (curves c and d).

From the above discussion, the following conclusions can be drawn. A switching converter with a distributed rectifier is superior in power efficiency as compared to on-chip linear and conventional switching converters. Due to the losses of a switching converter with a distributed rectifier operating at higher switching ratios (in this example 0.8), a linear rectifier is more efficient for these conversion ratios. To obtain the highest efficiency over a wide range of conversion ratios (in 3-D circuits), a hybrid methodology that uses a combination of linear and switching converters is proposed. As evident from Figure 9.5, a switching converter with a distributed rectifier should be used up to a critical conversion ratio (in this example, the ratio is 0.8). Above this critical ratio, a linear rectifier provides the highest efficiency.

Alternatively, a dual-V_{dd} approach can further improve efficiency. In this approach, two external power supplies V_{dd1H} and V_{dd1L} are provided where V_{dd1H} is higher than V_{dd1L}. As observed from Figure 9.5, the highest efficiency is obtained at higher conversion ratios. By separating the entire conversion range into three sections, overall performance is improved, as shown in Figure 9.6. In this example, V_{dd1L} / V_{dd1H} equals 0.4. The highest power efficiency depends upon the required conversion ratio. For conversion ratios between 0.2 to 0.4 and 0.8 to 1, a linear 3-D converter is preferable, while for conversion ratios between 0.4 to 0.8, a switching converter with a distributed rectifier provides the highest power efficiency. Over

the entire conversion range (0.3 to 1), the power efficiency ranges from 65% to about 95% (at a 0.4 conversion ratio).

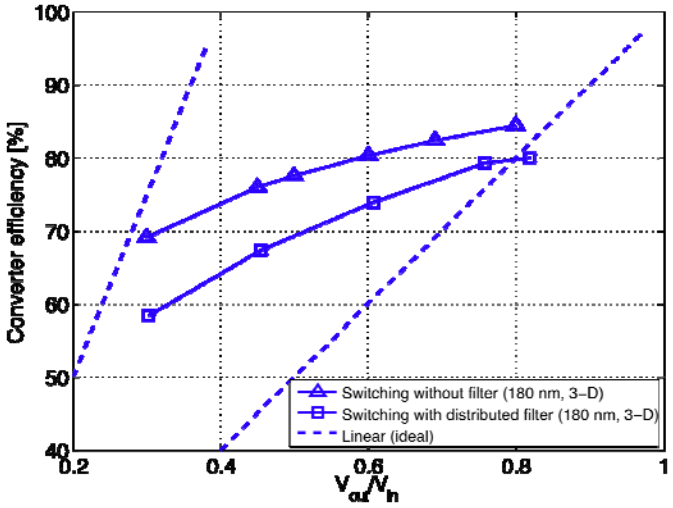


Figure 9.6 Efficiency of linear and switching converters with dual-V_{dd} input voltage (the two dashed lines represent the two input voltages)

9.3.2 A Linear Converter in 3-D Technology

A novel 3-D based linear converter is illustrated in Figure 9.7. The current is distributed within the interconnect network. The *RLC* transmission lines are connected by 3-D vias, forming a network that spans from the first plane to a target plane. To adhere to electromigration and current density rules, several interconnect networks, as shown in Figure 9.7, are connected in parallel, providing the required current load.

On the first plane, current mirror transistors (M_7 and M_8) conduct the primary portion of the current load. These transistors are biased by the circuit shown in Figure 9.8. The current mirror produces a high output impedance at the input to the first interconnect on the first

plane. The same amount of current will therefore flow within the interconnect network regardless of the voltage ripple. The remaining portion of the current flows through the PMOS transistor M_9 which accommodates the input voltage, process, and DC current load variations.

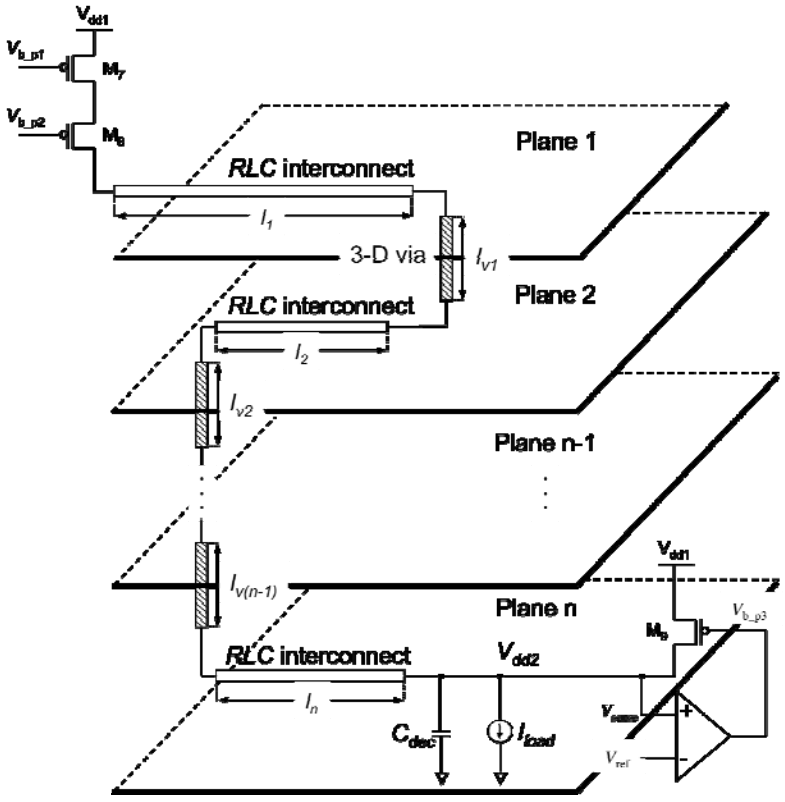


Figure 9.7 Linear converter within a 3-D circuit

An amplifier connected in a negative feedback loop controls the gate voltage of transistor M_9. By comparing the voltage at the output of the network to a reference voltage (V_{dd2}), the amplifier adjusts the gate voltage of transistor M_9 to maintain a stable output power supply. The amplifier circuit is shown in Figure 9.9. Note that the voltages used to bias this

amplifier are also provided by the circuit shown in Figure 9.8. The decoupling capacitor at the output C_{dec} reduces the voltage ripple produced by the transient current load.

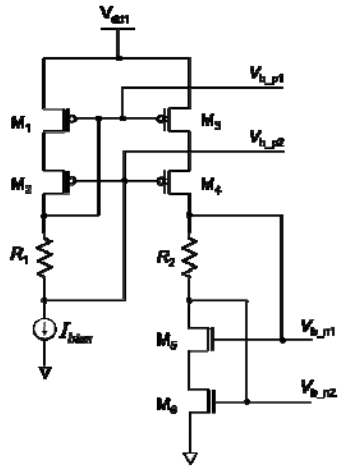


Figure 9.8 Current mirror bias circuit

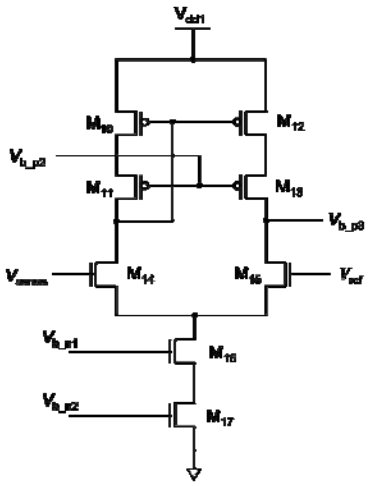


Figure 9.9 A single-ended differential amplifier

As described in section 9.2, the maximum power efficiency of a linear converter is determined by the ratio of the output voltage to the input voltage. The current efficiency is

therefore a useful metric for linear converters. In this linear converter circuit (see Figure 9.7), three currents have a direct DC path to ground (see Figures 9.8 and 9.9) which reduce the current efficiency. The total lost current is

$$I_{linear,lost} = 2I_{bias} + m \cdot I_{bias} = (2+m)I_{bias}, \tag{9.4}$$

where m is the ratio between the width of transistor M_{16} (M_{17}) and M_5 (M_6). Thus, assuming $m = 1$, the power efficiency of a linear converter is

$$\eta_{linear,3-D} = \left(\frac{V_{dd2}}{V_{dd1}}\right) \cdot \left(\frac{I_{DC}}{I_{DC} + 3I_{bias}}\right), \tag{9.5}$$

where I_{DC} is the DC current load,

$$I_{DC} = I_0 + \frac{\Delta i}{2T_{CLK}}\left(t_{rc} + t_{fc}\right). \tag{9.6}$$

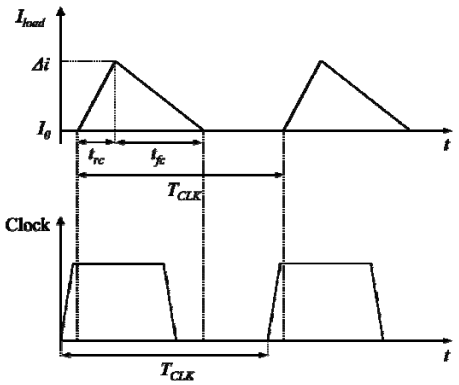


Figure 9.10 A typical current load profile in a high performance processor

In (9.6), T_{CLK} is the time period of the current load, which corresponds to the clock signal frequency, as depicted in Figure 9.10.

9.3.3 A Switching Converter in 3-D Technology

A switching converter with a distributed rectifier is illustrated in Figure 9.11. The distributed rectifier passes the DC and residual higher order harmonic components of the signal at the output of the power MOSFETs (PM_1 and PM_2). The AC signal produced by PM_1 and PM_2 is shown in Figure 9.12. The rectifier is comprised of an interconnect network that spans multiple planes. Capacitors C_1 to C_n are connected to each interconnect while 3-D vias connect the interconnects throughout the planes, forming a distributed filter structure. Analysis and design guidelines of the distributed rectifier are described in Chapter 7 and the power losses are included in this chapter for completeness.

Power Losses of Switching Converter with Distributed Rectifier

The input power dissipated by the distributed rectifier (see Figure 9.11) is

$$P_{rect} = V_{A,rms}^2 \left| \frac{Z_{in}}{Z_{in} + R_{0,eff}} \right|^2 \Re\left\{ \frac{1}{Z_{in}} \right\}, \tag{9.7}$$

where $V_{A,rms}$ is the rms voltage of the input signal, $R_{0,eff}$ is the effective output resistance of the power MOSFETs PM_1 and PM_2, and Z_{in} is the input impedance of the rectifier network, as described in section 7.4.3. The output power at the load of the rectifier is

$$P_{out,rect} = V_{dd2} I_{DC} . \tag{9.8}$$

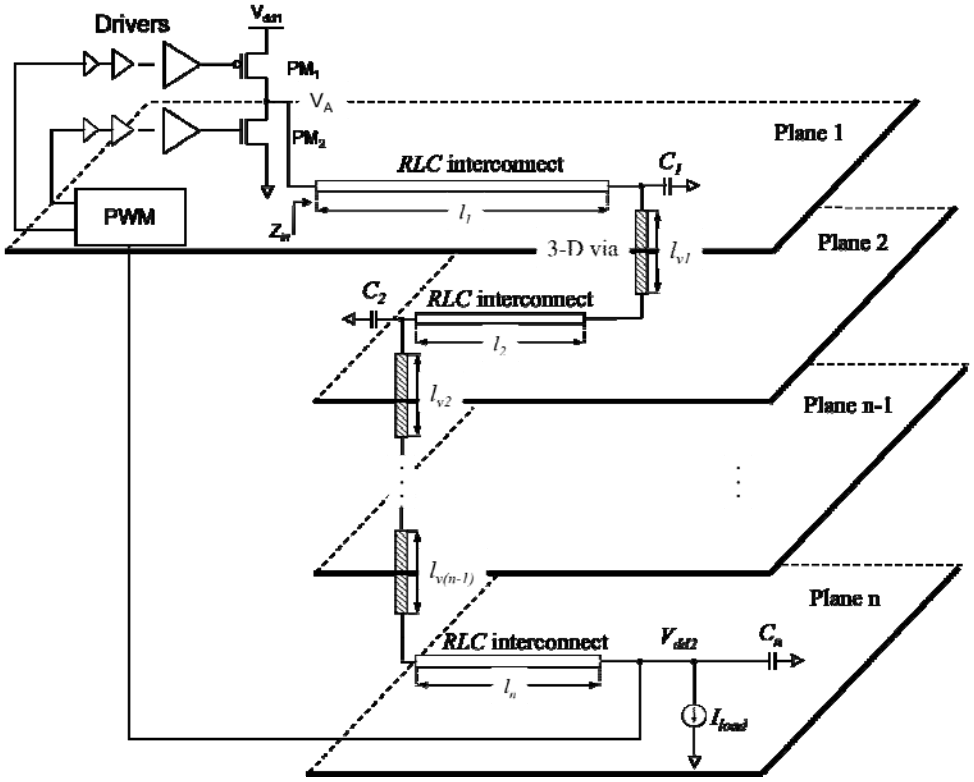


Figure 9.11 A switching converter with the 3-D distributed rectifier

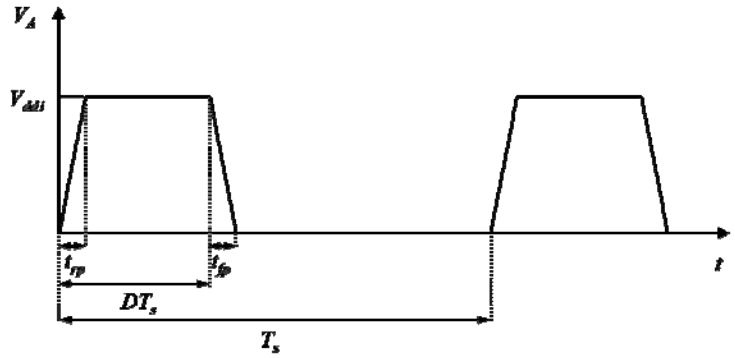


Figure 9.12 Voltage signal at the input of the rectifier

The power MOSFETs are driven by buffers controlled by a pulse width modulation (PWM) circuit. The PWM circuit senses the output voltage V_{dd2}, generating the required duty cycle to maintain a stable output voltage. The design of the power MOSFETs and the drivers for maximum power efficiency is adopted from [139].

The average power consumed by a power MOSFET (PMOS or NMOS) and the driving buffers is composed of resistive and dynamic power losses, respectively,

$$P_{MOS,res} = \frac{R_0}{W} I_{rms}^2, \tag{9.9}$$

$$P_{MOS,dyn} = EWf_s, \tag{9.10}$$

where R_0 is the effective resistance of a 1 μm wide transistor, W is the width of the power MOSFET, I_{rms} is the root-mean-square current passing through the power MOSFET, f_s is the switching frequency, α is the tapering factor of the buffers, and C_{ox}, C_{gs}, C_{gd}, and C_{db} are the gate oxide, gate-to-source, gate-to-drain, and drain-to-body capacitances, respectively, of a 1 μm wide transistor. E is the unit energy per 1 μm wide power MOSFET consumed during one switching cycle,

$$E \cong \frac{\alpha}{\alpha - 1}\left(C_{ox} + C_{gs} + 2C_{gd} + C_{db}\right)V_{dd1}^2. \tag{9.11}$$

As evident from (9.9) and (9.10), increasing the width of the power MOSFET reduces the resistive losses and increases the dynamic power losses, respectively. An optimum MOSFET width, therefore, exists that minimizes the total power of the drivers and the power

MOSFET. The optimal width and the related power consumption expressions are, respectively,

$$W_{opt} = \sqrt{\frac{R_0 I_{rms}^2}{f_s E}}\ [\mu\text{m}], \tag{9.12}$$

$$P_{MOS} = 2\sqrt{R_0 I_{rms}^2 f_s E}. \tag{9.13}$$

The effective output resistance of the optimal PMOS and NMOS transistors are, respectively,

$$R_{PMOS,eff} = \frac{R_0(PMOS)}{W_{opt}(PMOS)}, \tag{9.14}$$

$$R_{NMOS,eff} = \frac{R_0(NMOS)}{W_{opt}(NMOS)}, \tag{9.15}$$

and the total effective resistance of the power MOSFETs $R_{0,eff}$ is the average of $R_{PMOS,eff}$ and $R_{NMOS,eff}$. The rms currents through PM_1 and PM_2 (assuming the ratio between the time when both transistors are off to the switching period T_s as compared to the duty cycle D is small) are

$$I_{rms,PMOS} = \sqrt{(1-D)\left(I_{DC}^2 + \frac{\Delta i^2}{3}\right)}, \tag{9.16}$$

$$I_{rms,NMOS} = \sqrt{D\left(I_{DC}^2 + \frac{\Delta i^2}{3}\right)}, \tag{9.17}$$

where I_{DC} is the DC current supplied to the load. Applying (9.13) for both PM_1 and PM_2, and substituting the rms current expressions, (9.16) and (9.17), results in the total optimal power consumption of the drivers and the power MOSFETs, respectively,

$$P_{total,MOS} = a\sqrt{\left(I_{DC}^2 + \frac{\Delta i^2}{3}\right)f_s}, \tag{9.18}$$

$$a = 2\left(\sqrt{R_{0,NMOS}(1-D)E_{NMOS}} + \sqrt{R_{0,PMOS}DE_{PMOS}}\right). \tag{9.19}$$

The overall efficiency of the 3-D switching converter is therefore

$$\eta_{switching,3-D} = \frac{I_{DC}\cdot V_{dd2}}{I_{DC}\cdot V_{dd2} + P_{total,MOS} + P_{rect}}. \tag{9.20}$$

The optimal design of the drivers and power MOSFETs does not consider the power consumed by the distributed rectifier. As evident from (9.9) and (9.10), increasing the width of the power MOSFETs decreases the resistive power component while increasing the dynamic power. When both components are equal, minimum power is dissipated. Consequently, decreasing the width of the power MOSFETs increases the effective output resistance of the power MOSFETs, reducing the power dissipated by the distributed rectifier, as evident in (9.7). Minimizing the power dissipation of the entire network may result in

differently sized power MOSFET than the optimal width (minimum dynamic and resistive loss) determined by (9.12). To illustrate this issue, the MOSFETs, rectifier, and total power dissipation as a function of the ratio between the effective output resistance of the power MOSFETs and the optimal power MOSFETs is shown in Figure 9.13. In this example, the minimum power consumption is achieved when twice $R_{MOS,opt}$ is used.

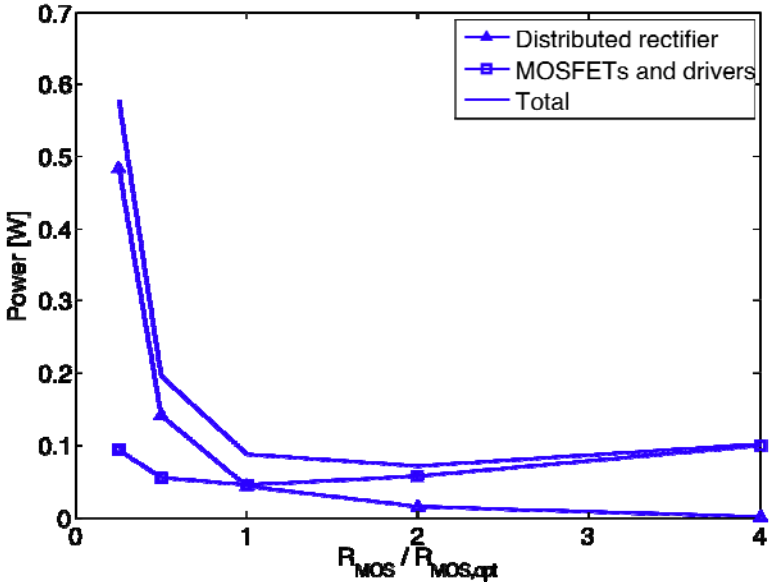


Figure 9.13 Power breakdown of the power MOSFETs and rectifier filter, distributing a 200 mA peak current

9.4 Case Study

Based on the methodology developed in section 9.3 and illustrated in Figure 9.5, two case studies are presented in this section. Conversion from 3.3 volts to 2.5 volts and 1.0 volt, based on the MIT Lincoln Laboratories (MITLL) 180 nm CMOS 3-D technology [156], is

demonstrated, assuming current distribution from the first plane to a third plane. In both cases, the target voltage ripple at the output is ± 2.5%. Additionally, the current load profile is assumed to behave as shown in Figure 9.10 with the following characteristics: $1/T_{CLK}$ = 3 GHz, t_{rc} and t_{cf} are $0.3T_{CLK}/2$ and $0.7T_{CLK}/2$, respectively, and I_0 and Δi are 0.5 and 2 amperes, respectively, resulting in 1 ampere DC current. A hybrid approach is employed to support a wide range of conversion ratios (in this example 3.3 volts to 1.0 volt).

The maximum current density in the MITLL 3-D technology is

$$J_{3-D,\max} = 3\ \text{mA}/\mu\text{m}^2. \tag{9.21}$$

To adhere to electromigration rules in this technology, the width of a single interconnect is chosen to be 2.5 μm, supporting a DC current of 5 mA. To conduct a maximum 2.5 amperes current ($I_0 + \Delta i$), 500 interconnects are connected in parallel. In section 9.4.1, a linear 3-D converter is used to convert from 3.3 volts to 2.5 volts. In section 9.4.2, a switching converter with a distributed rectifier is used to convert from 3.3 volts to 1.0 volt.

9.4.1 3.3 Volts to 2.5 Volts Conversion

The length of the interconnects on each plane are chosen arbitrarily to be 1 mm long. The current mirror bias (Figure 9.8) and the amplifier (Figure 9.9) circuits have been designed with the transistor widths listed in Table 9.1. All of the transistor channel lengths are 360 nm. Both resistors, R_1 and R_2, and the current bias I_{bias} are 5 kΩ and 100 μA, respectively. Note that transistors M_7 and M_8 conduct the majority of the current (800 mA) while transistor M_9 conducts the rest of the DC current (200 mA). The area occupied by all of

the transistors is 0.0176 mm^2, while the area of 500 minimum-spaced (0.35 μm) parallel interconnects is 0.00122 mm^2 per plane. To support an output voltage ripple of ± 2.5%, a 300 pF decoupling capacitor C_{dec} is chosen, achieving 104 mV ripple (± 2.1% of a 2.5 output voltage). The power efficiency of this linear converter is 74.2%, achieving 98.3% current efficiency. The simulated output voltage is shown in Figure 9.14.

To examine the response of a linear 3-D converter to abrupt changes in the current load, consider Figure 9.15. The response of the linear converter to an increase of 200 mA current load with a 20 A/μs slew rate is shown in Figure 9.15(a). The output signal exhibits a voltage undershoot of 386 mV, settling to 2.5 volts after 156 ns. The response of the linear converter to a decrease of 100 mA in the current load with a 20 A/μs slew rate is shown in Figure 9.15(b). In this case, the output voltage exhibits an overshoot of 187 mV over a period of 523 ns. Note that the settling time when the current load is decreased by 100 mA is long due to the relatively low gain of the amplifier.

Table 9.1 Width of the transistors within the 3-D linear converter

Transistor	Width [μm]
M_1, M_2, M_3, M_4	5
M_5, M_6, M_{16}, M_{17}	2
M_7, M_8	40000
M_9	8800
M_{10}, M_{12}	1.1
M_{11}, M_{13}	13.3
M_{14}, M_{15}	5

9.4.2 3.3 Volts to 1.0 Volt Conversion

To achieve maximum power efficiency for a conversion ratio of 0.3, the use of a switching converter with a distributed rectifier is required (see Figure 9.5). The same

specifications as for the linear converter are assumed in this example. The methodology described in Chapter 7 is used to design the distributed rectifier.

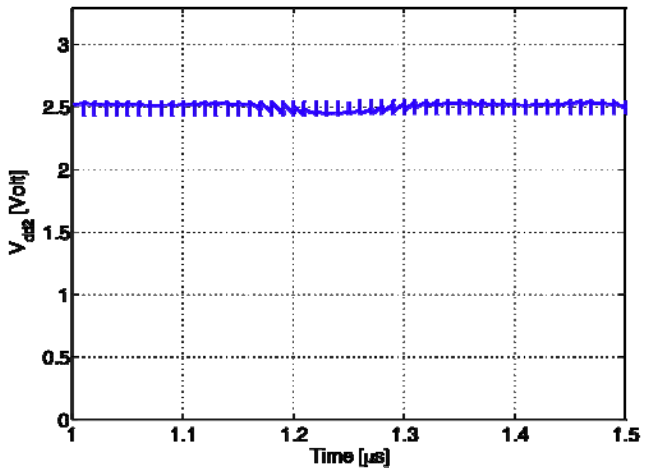
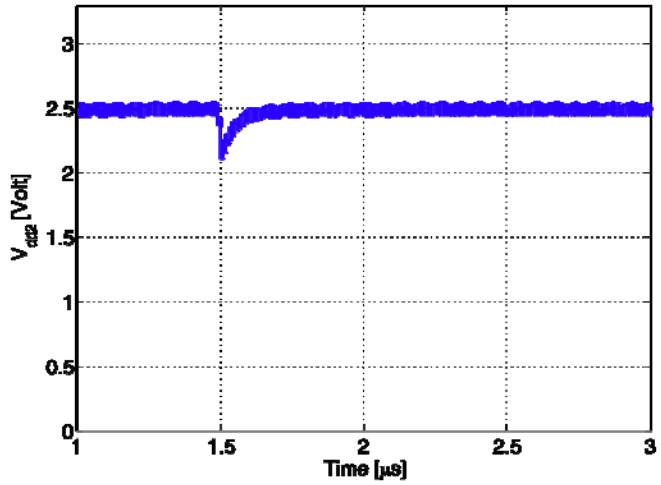
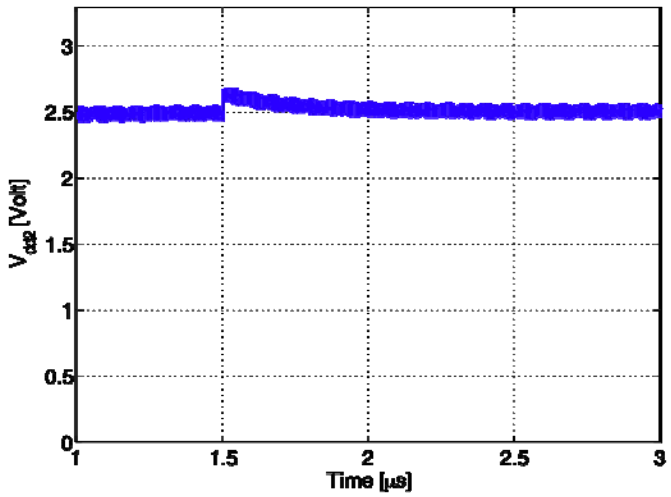


Figure 9.14 Output voltage of the linear 3-D converter

Designing a switching converter with a distributed rectifier requires an iterative approach, since the effective output resistance of the power MOSFETs affects the magnitude of the duty cycle D_{PWM} to achieve the required output voltage. Concurrently, the duty cycle D_{PWM} affects the design of the drivers and power MOSFETs. The iterative approach converges within a few iterations. To convert from 3.3 volts to 1.0 volt, the optimal PMOS and NMOS transistor widths (using (9.12)) are, respectively, 22.9 and 15.9 mm. The corresponding total output resistance is 0.11 Ω. Note that although the width of the power MOSFETs are extremely large, the corresponding area occupied by these transistors is only 0.0041 and 0.0023 mm^2.



(a)



(b)

Figure 9.15 Transient output voltage response: (a) 200 mA increase in current load, (b) 100 mA decrease in current load

To ensure the voltage ripple of the output power supply is 50 mV (± 2.5% of 1 volt), the magnitude of the transfer function of the distributed rectifier, (9.22), switching at 100 MHz has to be equal to or less than 0.0149 (see Chapter 7)

$$\left|H_{rect}\left(j\omega_s\right)\right| = \frac{v_r \cdot V_{dd1}}{4\left|a_1\right|}, \tag{9.22}$$

where

$$a_1 = \left(\frac{V_{dd1} T_s}{4 t_r \pi^2}\right) \cdot \left(e^{-j\omega_s t_r}\left(1 - e^{-j2\pi D}\right) + e^{-j2\pi D} - 1\right) \tag{9.23}$$

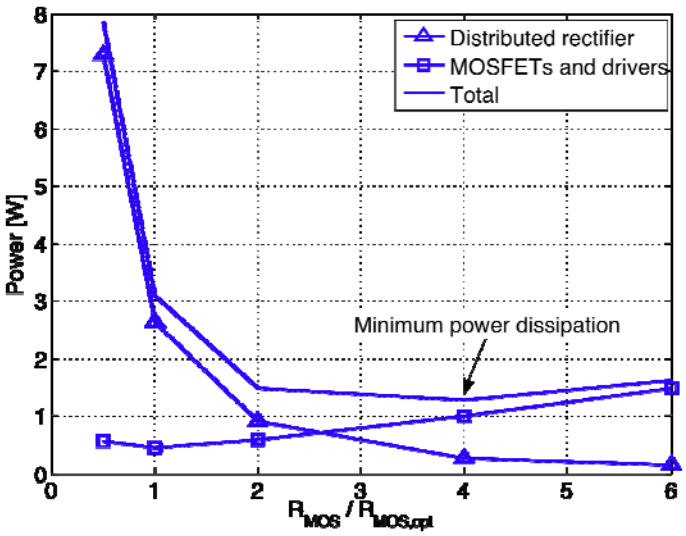


Figure 9.16 Power components of the 3-D switching converter, distributing 2.5 Amperes of peak current

To determine the width of the power MOSFETs and the distributed rectifier structure, the MOSFETs, rectifier, and combined power dissipation as a function of the ratio between the effective output resistance of the power MOSFETs and the optimal power MOSFETs is shown in Figure 9.16.

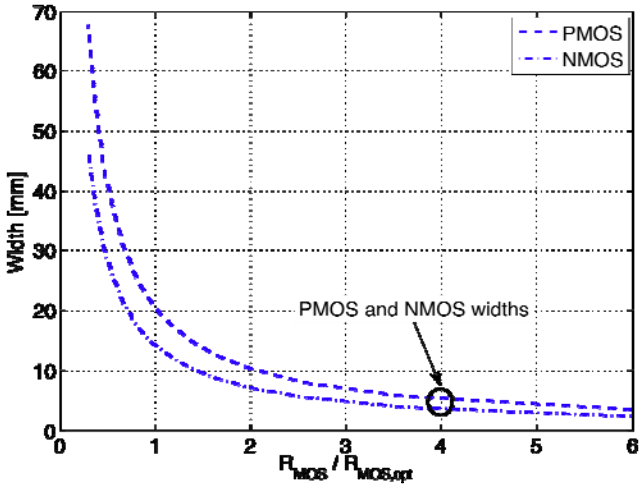


Figure 9.17 Power MOSFET widths as a function of $R_{MOS}/R_{MOS,opt}$

To achieve minimum power dissipation, the effective output resistance of the power MOSFETs is four times $R_{MOS,opt}$. To determine the corresponding width of the power MOSFETs, consider Figure 9.17. The PMOS and NMOS widths provide a four times $R_{MOS,opt}$ effective resistance of 5.3 and 3.7 mm, respectively. In this example, the required duty cycle is 0.46 (see Chapter 7) according to

$$D_{PWM} = D \cdot \left| 1 + \frac{R_{0,eff}}{Z_{in}(0)} \right|. \tag{9.24}$$

Since this circuit is designed to provide 2.5 amperes maximum current, 500 rectifier circuits on each plane with 0.5 mm long interconnects are connected in parallel, resulting in 67 nF capacitance per plane. Assuming a 10 fF/μm^2 capacitance density in the MITLL 3-D technology [156], the area of the switching converter is about 2.6 × 2.6 mm^2 (mostly occupied by on-chip capacitors) and the efficiency is 44%. Although this is a relatively low

power efficiency, note that the aggressive operation of on-chip conversion from 3.3 volts to 1.0 volt, while distributing 2.5 amperes maximum current, has not been reported in the literature, and is shown to be possible here for the first time. As compared to other converters, the proposed 3-D converter exhibits superior performance, as described in section 9.5. The output voltage produced by this converter, exhibiting 49.5 mV voltage ripple, is shown in Figure 9.18.

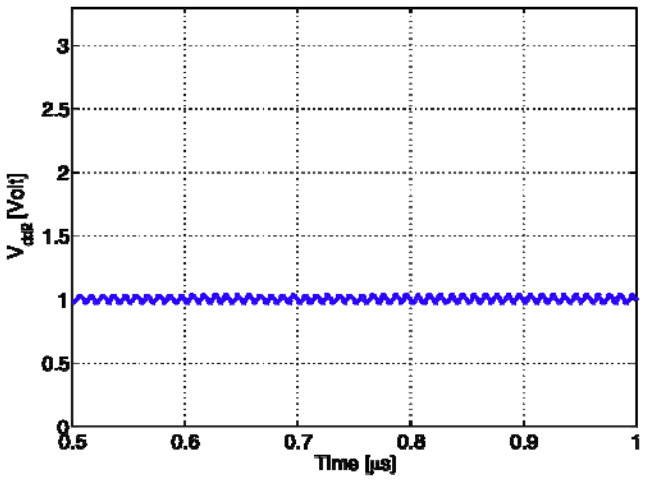


Figure 9.18 Output voltage of the switching converter

9.5 Performance Comparison

The switching converter occupies a relatively large area due to the significant amount of capacitance required to distribute a stable power supply while delivering several Amperes of current. The area occupied by the on-chip capacitors, however, is predicted to decrease with advanced technologies. The MOS capacitor density can be estimated as $\varepsilon_{ox} / t_{ox}$, where ε_{ox} and t_{ox} are the dielectric constant and MOS oxide thickness, respectively [168]. MOS capacitor

densities as a function of technology are shown in Figure 9.19. Furthermore, the feasibility of fabricating extremely large on-chip capacitances has been demonstrated [169], where a 250 nF on-chip MIM capacitor in 90 nm technology has been implemented. Alternatively, by exploiting the Miller effect, small area active capacitors can also be used, albeit requiring additional power [170].

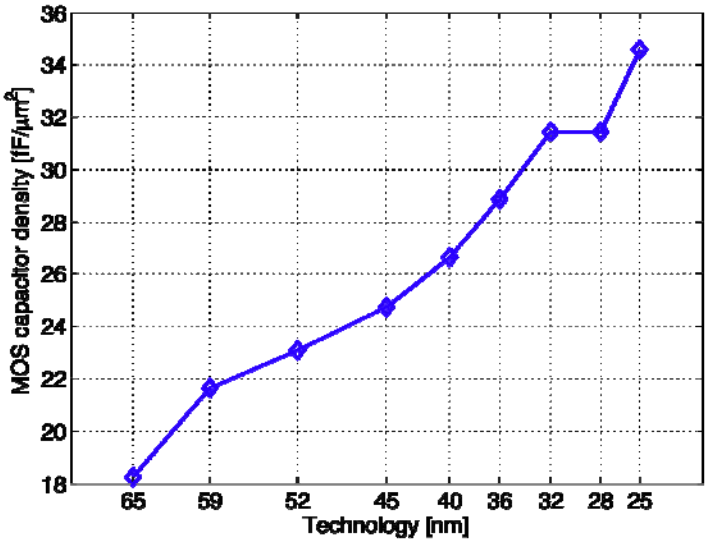


Figure 9.19 MOS capacitor density for different technologies

The multiple plane structure in 3-D circuits suggests that area is less of a concern. Thus, several planes can be dedicated to accommodate the required capacitors. Note that a conventional *LC* filter not only requires integration of an inductor but also a significant amount of capacitance. For example, an on-chip 5 nH inductor requires a 35 nF capacitor to achieve 5% output voltage ripple at 100 MHz with 0.46 Ω output resistance of the power MOSFETs (corresponds to conducting 2.5 amperes maximum current). The capacitance of the proposed switching converter for a wide output voltage range while distributing 200 mA

is shown in Figure 9.20, demonstrating that less capacitance is required for higher conversion ratios.

To evaluate the capability of providing a wide voltage conversion range, as required in 3-D circuits, the proposed converter is compared to three on-chip converters, as shown in Figure 9.21. The proposed converter in this example distributes 200 mA maximum current. As evident from Figure 9.21, the power efficiency of the proposed circuit is greater than the other converters. A primary reason for the superior power efficiency is that the proposed circuit eliminates the need for an on-chip inductor and any associated losses. Additionally, the hybrid circuit is composed of both a linear or switching converter with the distributed rectifier (depending on the required conversion ratio), further increasing the efficiency. Note that the power efficiency of the hybrid (linear and switched-capacitor converter) [166] achieves a maximum efficiency of 40% at a 0.38 conversion ratio due to the increased efficiency provided by the low voltage regulator system [166].

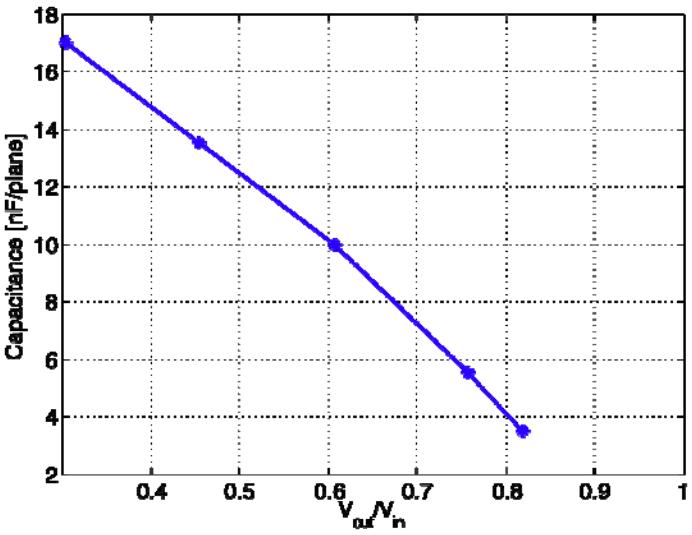


Figure 9.20 Capacitance per plane as a function of V_{out}/V_{in}

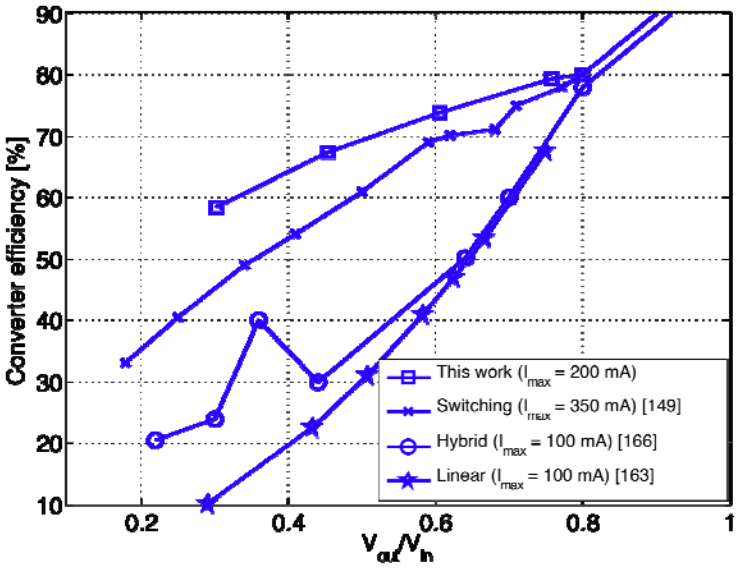


Figure 9.21 Power efficiency of different approaches as a function of V_{out}/V_{in}

A performance comparison between the hybrid 3-D converter and six different approaches for on-chip DC-DC conversion is listed in Tables 9.2 and 9.3. In Table 9.2, the performance characteristics of state-of-the-art linear, switched-capacitor, and switching converters in 2-D and 3-D technologies are compared. The performance of the hybrid 3-D converter with a 3.3 volt input power supply, 100 MHz switching frequency, and 5% voltage ripple based on the MITLL CMOS/SOI 180 nm technology is listed in Table 9.3.

As can be observed from this comparison, the proposed circuit achieves significantly superior performance. The proposed buck converter efficiently distributes current in the ampere range, as compared to the other approaches that only provide sub-ampere current loads while maintaining reasonable power efficiency.

The proposed converter enables a wide voltage conversion range, achieving the highest power efficiency as compared to other approaches. The area required by the proposed circuit is comparable to other approaches. These properties and the distributed nature of the

converter make the proposed converter highly suitable for 3-D integration, where multiple voltage domains are required.

Table 9.2 Performance comparison of different on-chip DC-DC converters in the literature

	Switching converter [149]	3-D switching [154]	Stacked switching [155]	Linear converter [163]	Linear converter [164]	Hybrid converter [166]
Technology [nm]	130, 2-D	180, 3-D	350, 3-D	90, 2-D	350, 2-D	250, 2-D
V_{out} / V_{in} [V/V]	0.9 / 1.2	0.9 / 1.8	2.3 / 3.3	0.9 / 1.2	3.15 / 5.5	1.62 / 2.5
Conversion ratio	0.75	0.50	0.70	0.75	0.57	0.65
I_{max} [mA]	350	500	70	100	200	100
Power efficiency [%]	77.9	64.0	62.0	70.7	57.1	61
Voltage ripple [%]	4.4	4.5	10	10	2	19
Switching frequency [MHz]	170	200	200	–	–	–
Area [mm^2]	1.5	6.9	4	0.098	0.264	0.42

Table 9.3 Performance summary of the proposed 3-D hybrid converter (this work)

	I_{max} = 0.2 amperes					I_{max} = 2.5 amperes	
V_{out} [volt]	1.0	1.5	2.0	2.5	2.7	1.0 (switching)	2.5 (linear)
Conversion ratio	0.30	0.45	0.60	0.75	0.82	0.30	0.75
Power efficiency [%]	58	67	74	79	80	44	74
Area [mm^2/plane]	1.7	1.34	1.0	0.55	0.35	6.65	0.048

9.6 Conclusions

An analysis of on-chip linear and switching DC-DC converters is provided in this chapter. From this analysis it can be observed that a hybrid approach where both linear and

switching converters are combined achieves the highest power efficiency as compared to other methods. By incorporating the concept of an inductorless distributed rectifier in 3-D circuits, the efficiency is further increased. The proposed circuit provides a wide voltage conversion range, suitable for the integration of 3-D heterogeneous systems, distributing 2.5 amperes maximum current to the load. With this current load, 0.75 and 0.3 conversion ratios with a 74% and 44% power efficiency are, respectively, demonstrated.

Chapter 10

Conclusions

Scaling is the most significant driving force in the development of CMOS integrated circuit technology. As the channel length of a transistor approaches a few tenths of a nanometer, physical and technological boundaries pose significant difficulties to advancing beyond subnanometer geometries. In current technologies, interconnect networks restrict the advance in performance of high complexity integrated circuits. This degradation in performance is a result of the effects of scaling on-chip interconnects.

Interconnect networks tend to be highly dense and long, reaching a few millimeters in length and distributing signals that often exceed Gigahertz frequencies. The rapid increase in current loads, exceeding 100 Amperes in state-of-the-art microprocessors, has a profound impact on the operation of these networks. While transistor scaling has significantly improved the performance of the active devices, interconnect networks degrade performance, undermining the overall advancements. Performance bottlenecks have therefore shifted from the active devices to the interconnect networks.

Clock and interconnect distribution networks are primary bottlenecks in high performance circuits. In order to improve performance, several novel clock and data distribution networks have been proposed, as described in Chapters 5 and 6. As technology scales, new design challenges have emerged, making the design of these networks an intricate and complicated task. Existing methodologies for designing clock and data distribution networks are being reevaluated to meet the challenges of future nanoscale integrated circuits.

Due to the large load capacitance at the local level of a clock distribution network, the portion of the clock tree closest to the registers dissipates a significant fraction of the total on-chip power. Inserting on-chip inductors, thereby eliminating the need for inserting buffers within these local networks, can significantly reduce power consumption, clock skew, and jitter. Energy resonates between magnetic and electrical fields rather than dissipates as heat, forming a highly power efficient clock distribution network. A design methodology for these resonant clock distribution networks has been described. A 5 GHz resonant H-tree clock distribution network based on a 180 nm CMOS technology is described to demonstrate this proposed design methodology.

Until recently, inductive effects in on-chip interconnects have been primarily treated as a detrimental parasitic phenomenon. Due to the difficulties in analyzing inductive effects, inductance has been mostly ignored during the design process. By enhancing the magnetic properties of on-chip interconnects, remarkable improvements in the interconnect bandwidth and signal transmission characteristics can be achieved as compared to repeater insertion. A novel quasi-resonant interconnect network for global data distribution has been proposed. By exploiting resonance, distribution of a 5 Gbps data signal within a 5 mm long interconnect is demonstrated which significantly reduces power dissipation and noise.

To significantly enhance the performance of interconnect networks, an emerging three-dimensional (3-D) integration technology has recently been developed. Three-dimensional technology offers the traditional benefits of large-scale integration as well as addresses the important issue of interconnect in integrated circuits. By vertically stacking two-dimensional (2-D) planes, 3-D integration can lead to a drastic decrease in interconnect length, resulting in increased speed and reduced power.

A primary difficulty in 3-D systems is power delivery, specifically, power generation and distribution. This difficulty originates from the technological heterogeneity and multi-plane structure offered by 3-D circuits. Since different planes occupy different circuit domains, *e.g.*, digital, mixed signal/analog, RF, and MEMs, distributed, reliable, and stable power supplies are required. To provide circuit domains with the required power supplies, multiple on-chip DC-DC converters should be distributed across each plane. The difficulty in distributing these power supplies in 3-D circuits arises due to the integration of several power distribution networks with different characteristics, all of which need to be integrated within a single multi-plane system. Another significant issue is the vertical interconnects, which provide communications among the planes. These vertical interconnects must be considered in determining the overall impedance characteristics of the power distribution networks. Since a significant amount of current flows within the vertical interconnects in a 3-D power distribution network, the electrical characterization and deliverable current properties of these vertical interconnects are of paramount importance.

Exploiting interconnect properties in 3-D technology to produce power generation and distribution networks is also described in this thesis. A methodology for designing a rectifier portion of a buck converter is described. As compared to a conventional filter within a buck converter, the proposed rectifier is composed of only interconnects and capacitors. By eliminating the need for an on-chip inductor, on-chip implementation of DC-DC converters for 3-D circuits becomes feasible. The proposed design methodology exploits the current delivery network to a target plane, producing the required voltage conversion. To demonstrate and evaluate the proposed rectifier, a test circuit has been designed in the MITLL 150 nm CMOS/SOI technology. In this three-plane structure, a distributed rectifier

demonstrating 3.3 to 1.5 volt conversion while delivering a peak current of 100 mA has been implemented.

Additionally, a comprehensive design methodology for on-chip 3-D buck converters is described. The proposed rectifier is a novel hybrid DC-DC converter, combining switching and linear buck converters utilizing the distributed rectifier. A wide conversion range with high efficiency is demonstrated in the MITLL 180 nm CMOS/SOI technology.

The primary contribution of this thesis is the study and exploration of three fundamental on-chip interconnect structures: clock, data, and power generation and distribution networks. To alleviate the difficulties associated with these networks, novel and innovative approaches exploiting on-chip resonance and three-dimensional integration have been proposed, delivering a significant improvement in performance over existing networks.

Clock, data, and power generation and distribution networks are the pillars on which high performance, high complexity integrated circuits are positioned. By providing novel circuit solutions and design methodologies, another chapter has been added to the chronicles of the extraordinary microelectronics revolution.

Chapter 11

Future Research

Challenges in high performance 2-D and 3-D integrated circuits have been addressed in this thesis. Existing methodologies for designing global clock, data, and power delivery networks can no longer efficiently support billions of transistors on a single die. To further improve the performance of future integrated circuits and maintain the current pace of advancement, revolutionary approaches must be explored and adopted. Novel materials, devices, integration techniques, and design methodologies are of future research interest in nanoscale integrated circuits. Future research topics are proposed in this chapter as an extension to the proposed methodologies and design techniques described in this thesis.

The spectral analysis of jitter associated with resonant clock distribution networks is proposed in section 11.1. The objective of the following two research topics is the development of novel circuits and design methods to improve the efficiency of DC-DC converters. A resonant on-chip DC-DC converter circuit is proposed in section 11.2 while a high efficiency buck converter filter is proposed in section 11.3. Design techniques to increase the efficiency and reduce the area of on-chip inductors are discussed in section 11.4. A summary is offered in section 11.5.

11.1 Spectral Analysis of Jitter in Resonant Clock Distribution Networks

Resonant clock distribution networks have demonstrated reduced power dissipation, clock skew, and signal jitter. Since resonant networks exhibit limited bandwidth as compared to a conventional clock distribution network, jitter noise propagated by the clock driver is attenuated.

To rigorously evaluate the effectiveness of jitter reduction in resonant networks, spectral analysis of jitter noise is an issue of interest. The distribution and characterization of jitter power between the fundamental and high order harmonic frequencies are important topics. If the majority of the power is carried by the higher order harmonics, jitter noise can be significantly reduced. If the fundamental harmonic carries the majority of the jitter noise, methods to spread this noise to higher frequencies should be explored.

11.2 On-Chip Resonant Buck Converters

Techniques for exploiting on-chip resonance to reduce power consumption and improve the performance of buck converters are proposed, as illustrated in Figure 11.1. The savings in area and power of the resonant buck converter is three-fold. As compared to a conventional buck converter, the large size MOSFETs as well as the tapered drivers can be eliminated, thereby reducing both power and area. Moreover, the PWM circuit is reduced in complexity, requiring only one control signal. The energy also resonates adiabatically between the electrical and magnetic fields rather than dissipates as heat.

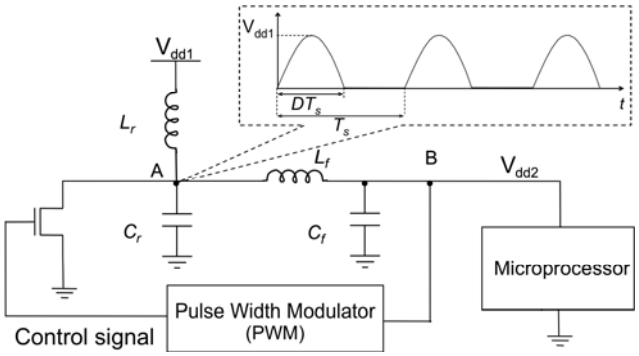


Figure 11.1 Resonant buck converter

The resonant buck converter operates in the following manner. When the control signal from the PWM circuit is high, the NMOS transistor turns on, shorting node A to ground. At this phase, current from V_{dd1} charges inductor L_r (see Figure 11.1). Once the control signal is at logic low, the NMOS transistor turns off, forming an LC tank circuit that resonates at the $1/2\pi\sqrt{LC}$ frequency. This periodic operation generates a half sinusoidal signal at node A with period T_s. The L_fC_f pass band filter only passes the DC component of the signal at node A, producing a voltage DV_{dd1} at node B. Therefore, low power and high efficiency operation can be provided by resonant buck converters. Design methodologies for resonant buck converters are therefore of interest. Furthermore, novel circuit topologies to implement resonant buck converters within a three-dimensional structure should be of interest, particularly those structures that exploit the characteristics and properties offered by three-dimensional integration.

11.3 On-Chip Buck Converter Rectifier (Filter)

An important issue with buck converters is the large magnitude and physical size of the

passive *LC* filter components required to achieve a low ripple in the output DC voltage. In order to integrate buck converters on-chip, the size of these passive inductors must be reduced. A novel method to achieve size reduction is suggested, as shown in Figure 11.2.

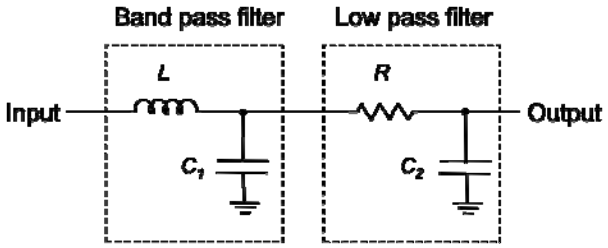


Figure 11.2 Proposed filter within the buck converter

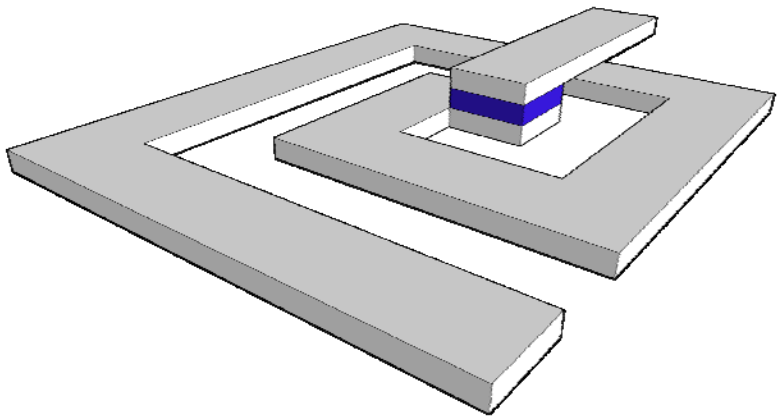
In this circuit, the LC_1 band pass filter is designed to only pass the DC component and a higher harmonic at a resonant frequency significantly greater than the fundamental harmonic of the input signal. The remaining harmonics are attenuated, resulting in a high frequency quasi-sinusoidal signal. This signal is applied to a low pass filter RC_2 that only passes the DC component of the signal to the output. With this strategy, the low frequency AC input signal is modulated to a higher frequency (with the same DC component), while ensuring that the required passive devices, L, C_1, R, and C_2, remain relatively small, suitable for on-chip integration.

11.4 On-Chip Inductors in 3-D Circuits

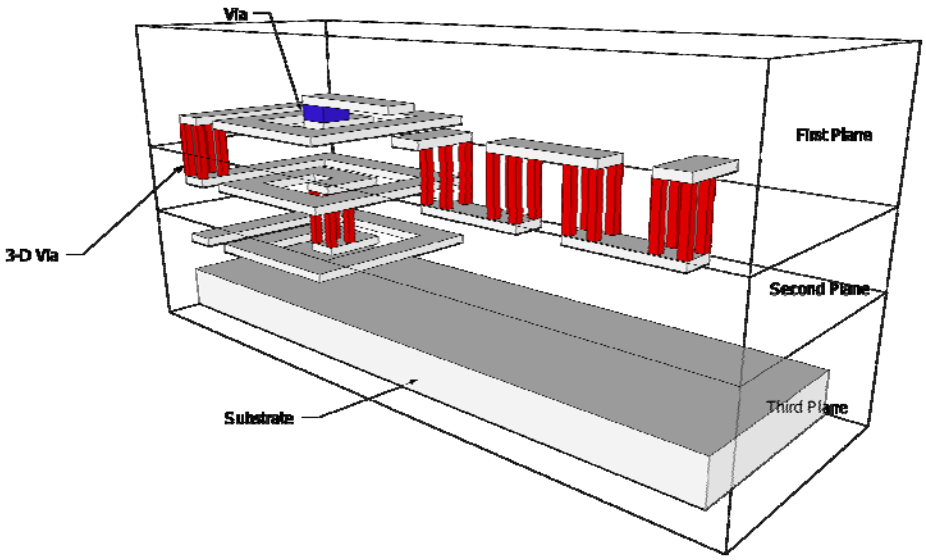
The importance of on-chip inductors for RF and, recently, digital circuits has been discussed in Chapters 4, 5, and 6. Unfortunately, the performance of on-chip inductors is greatly affected by the inherent parasitic losses of the inductor; moreover, spiral inductors

occupy a large area. In deeply scaled technologies, the performance and area of these spiral inductors are of significant importance.

Undoubtedly, on-chip inductors will be required in 3-D circuits. Novel inductor structures integrated within a three-dimensional technology should be explored. Specifically, the design of different on-chip inductors exploiting both the vertical and horizontal degrees of freedom are of interest. Three spiral inductor topologies based on a 3-D technology are identified, as shown in Figure 11.3. As opposed to planer two-dimensional technologies, where the inductors are horizontal to the substrate (Figure 11.3 (a)); in a three-dimensional structure, the inductors can be positioned perpendicular to the substrate, significantly reducing the losses associated with the substrate. This approach can improve the efficiency of the spiral inductors and, consequently, the performance of the circuit. Stacked 2-D inductors and vertical 3-D inductors are two topologies of research interest, as shown in Figure 11.3(b).



(a)



(b)

Figure 11.3 On-chip inductor topologies in 3-D circuits: (a) planar 2-D inductor, (b) stacked planar 2-D and vertical 3-D inductors

An emerging approach to improve the performance of on-chip inductors is to integrate on-chip magnetic materials [101]. In these inductors, the magnetic material confines the magnetic field, thereby reducing the losses associated with conventional metal on-chip inductors. Significant improvements in the Q factor, self-resonance, and magnitude of the inductors incorporating thin magnetic films have been demonstrated [101]. The integration of magnetic inductors for application to DC-DC converters within a three-dimensional structure is an issue of significant interest. Since multiple planes are available in three-dimensional technology and area is less of a concern as compared to a two-dimensional technology, a large portion or perhaps an entire plane can be dedicated to the integration of these magnetic inductors. Techniques to efficiently partition the planes, integrate magnetic materials, and utilize the magnetic inductors with buck converters are topics of importance.

11.5 Summary

As technology advances, new challenges arise, requiring novel solutions and innovative approaches. Some of these approaches have been addressed in this thesis. Several open research directions related to on-chip resonant networks and 3-D integrated circuits are described in this chapter.

Specifically, research problems related to jitter analysis and characterization in resonant distribution networks, as well as resonant on-chip buck converters are suggested. Additionally, a novel approach to increase the efficiency and reduce the area of an on-chip buck converter rectifier is described. Finally, novel topologies for on-chip inductors for application to 3-D integrated circuits are described as topics for potential research.

Bibliography

[1] http://www.pbs.org/transistor

[2] J. D. Meindl, Retrospect and Prospect, *Proceedings of the IEEE International Conference on Very Large Scale Integration* (keynote speech), October 2007

[3] http://www.itrs.net

[4] R. G. Beausoleil, P. J. Kuekes, G. S. Snider, S. Y. Wang, and R. S. Williams, "Nanoelectronic and Nanophotonic Interconnect," *Proceedings of the IEEE*, Vol. 96, No. 2, February 2008

[5] D. Chapiro, *Globally-Asynchronous Locally-Synchronous Systems*, Doctoral Dissertation, Stanford University, 1984

[6] M. Krstic, E. Grass, F. K. Gurkaynak, and P. Vivet, "Globally Asynchronous, Locally Synchronous Circuits: Overview and Outlook," *IEEE Design and Test of Computers*, Vol. 24, No. 5, pp. 430-441, September/October 2007

[7] E. G. Friedman, "Clock Distribution Networks in Synchronous Digital Integrated Circuits," *Proceedings of the IEEE,* Vol. 89, No. 5, pp. 665-692, May 2001

[8] E. G. Friedman, *High Performance Clock Distribution Networks*, Kluwer Academic Publishers, 1997

[9] A. V. Mule, E. N. Glytsis, T. K. Gaylord, and J. D. Meindl, "Electrical and Optical Clock Distribution Networks for Gigascale Microprocessors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 10, No. 5, pp. 582-594, October 2002

[10] P. Gronowski, W. J. Bowhill, R. P. Preston, M. K. Gowan, and R. L. Allmon, "High-Performance Microprocessors Design," *IEEE Journal of Solid-State Circuits,* Vol. 33, No. 5, pp. 676-686, May 1998

[11] D. R. Gonzales, "Micro-RISC Architecture for the Wireless Market," *IEEE Micro*, Vol. 19, No. 4, pp. 30-37, July/August 1999

[12] The International Technology Roadmap for Semiconductors (ITRS), 2005

[13] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Figures of Merit to Characterize the Importance of On-Chip Inductance," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 7, No. 4, pp. 442-449, December 1999

[14] C. Juang-Ying, C. H. Ziesler, and M. C. Papaefthymiou, "Empirical Evaluation of Timing and Power in Resonant Clock Distribution," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 249-252, May 2004

[15] D. Wann and M. Franklin, "Asynchronous and Clocked Control Structures for VLSI Based Interconnections Networks," *IEEE Transactions on Computers*, Vol. C-32, No. 3, pp. 284-293, March 1983

[16] R. Navid, T. H. Lee, and R. W. Dutton, "An Analytical Formulation of Phase Noise of Signals With Gaussian-Distributed Jitter," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* , Vol. 52, No. 3, pp. 149-153, March 2005

[17] F. Herzel and B. Razavi, "A Study Oscillator Jitter Due to Supply and Substrate Noise," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 46, No. 1, pp. 56-62, January 1999

[18] B. Nikolic, A. Chandrakasan, and J. M. Rabaey, *Digital Integrated Circuits*, Prentice Hall Electronics and VLSI Series, 2003

[19] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-Power Dissipation in Microprocessor," *Proceedings of the ACM International Workshop on System Level Interconnect Prediction*, pp. 7-13, February 2004

[20] M. Donno, E. Macii, and L. Mazzoni, "Power-Aware Clock Tree Planning," *Proceedings of the ACM International Symposium on Physical Design*, pp. 138-147, April 2004

[21] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*, John Wiley & Sons Press, 2006

[22] D. M. Pozar, *Microwave Engineering*, Addison-Wesley Publishing Company, 1990

[23] P. J. Restle *et al.*, "A Clock Distribution Network for Microprocessors," *IEEE Journal of Solid-State Circuits,* Vol. 36, No. 5, pp. 792-799, May 2001

[24] J. A. Davis and J. D. Meindl, "Compact Distributed *RLC* Interconnect Models—Part I: Single Line Transient, Time Delay, and Overshoot Expressions," *IEEE Transactions on Electron Devices*, Vol. 47, No. 11, pp. 2068–2077, November 2000

[25] R. Venkatesan, J. A. Davis, and J. D. Meindl, "A Physical Model for the Transient Response of Capacitively Loaded Distributed *RLC* Interconnects," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 763-766, June 2002

[26] B. Kahng and S. Muddu, "An Analytical Delay Model for *RLC* Interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 16, No. 12, pp. 1507–1514, December 1997

[27] B. Kahng, K. Masuko, and S. Muddu, "Analytical Delay Models for VLSI Interconnects Under Ramp Input," *Proceedings of the ACM/IEEE International Conference on Computer-Aided Design*, pp. 30–36, November 1996

[28] K. Banerjee and A. Mehrotra, "Analysis of On-Chip Inductance Effects for Distributed *RLC* Interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, No. 8, pp. 904–914, August 2002

[29] Y. Ismail, E. G. Friedman, and J. L. Neves, "Equivalent Elmore Delay for *RLC* Trees," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 19, No. 1, pp. 83–97, January 2000

[30] L. T. Pillage and R. A. Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 9, No. 4, pp. 352–366, April 1990

[31] G. Chen and E. G. Friedman, "An *RLC* Interconnect Model Based on Fourier Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 2, pp. 170–183, February 2005

[32] P. J. Restle and A. Deutsch, "Designing the Best Clock Distribution Network," *Proceedings of the IEEE Symposium on Very Large Scale Integration (VLSI) Circuits*, pp. 2-5, June 1998

[33] P. J. Restle *et al.*, "The Clock Distribution of the Power4 Microprocessor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 144-145, February 2002

[34] F. E. Anderson, J. S. Wells and E. Z. Berta, "The Core Clock System on the Next Generation Itanium Microprocessors," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 146-147, February 2002

[35] X. McMahan *et al.*, "A 600 MHz NT3 Network Processor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 256-257, February 2003

[36] S. Tam *et al.*, "Clock Generation and Distribution for the First IA-64 Microprocessor," *IEEE Journal of Solid-State Circuits,* Vol. 35, No. 11, pp. 1545-1552, November 2000

[37] T. Xanthopoulos *et al.*, "The Design and Analysis of the Clock Distribution for a 1.2 GHz Alpha Microprocessors," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 402-403, February 2001

[38] N. A. Kurd *et al.*, "Multi-GHz Clocking Scheme for Intel Pentium4 Microprocessors," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 402-403, February 2001

[39] N. Bindal *et al.*, "Scalable Sub-10 ps Skew Global Clock Distribution for a 90 nm Multi-GHz IA Microprocessor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 346-347, February 2003

[40] G. Geannopoulos and X. Dai, "An Adaptive Digital Deskewing Circuit for Clock Distribution Networks," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 400-401, February 1998

[41] Y. Nakagome *et al.*, "Sub 1-V Swing Internal Bus Architecture for Future Low-Power ULSI's," *IEEE Journal of Solid-State Circuits*, Vol. 28, pp. 414-419, April 1993

[42] R. Golshan and B. Haroun, "A Novel Reduced Swing CMOS Bus Interface Circuit for High Speed Low Power VLSI Systems," *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 4, pp. 351-354, May 1994

[43] D. Markovic, J. Tschanz, and V. De, "Feasibility Study of Low-Swing Clocking," *Proceedings of the IEEE International Conference on Microelectronics*, Vol. 2, pp. 547–550, May 2004

[44] H. Zhang, V. Goerge, and J. M. Rabaey, “Low-Swing On-Chip Signaling Techniques: Effectiveness and Robustness,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 8, No. 3, pp. 264-272, June 2000

[45] V. Kursun, R. M. Secareanu, and E. G. Friedman, "CMOS Voltage Interface Circuit for Low Power Systems," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 3.667-3.670, May 2002

[46] T. Burd, *Energy Efficient Processor System Design*, Doctoral Dissertation, University of California, Berkeley, 1998

[47] P. Gupta and A. B. Kahng, “Wire Swizzling to Reduce Delay Uncertainty Due to Capacitive Coupling,” *Proceedings of the IEEE International Conference on VLSI Design*, pp. 431-436, 2004

[48] B. Soudan, “The Effects of Swizzling on Inductive and Capacitive Coupling for Wide Signal Buses,” *Proceedings of the IEEE International Conference on Microelectronics*, pp. 300-303, December 2003

[49] T. Sakurai, “Closed-Form Expressions for Interconnections Delay, Coupling, and Crosstalk in VLSI’s,” *IEEE Transactions on Electron Devices*, Vol. 40, No. 1, pp. 118-124, January 1993

[50] V. Adler and E. G. Friedman, "Uniform Repeater Insertion in *RC* Trees," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* Vol. 47, No. 10, pp. 1515-1523, October 2000

[51] Y. I. Ismail and E. G. Friedman, "Effects of Inductance on the Propagation Delay and Repeater Insertion in VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 8, No. 2, pp. 195-206, April 2000

[52] H. B. Bakoglu and J. D. Meindl, “Optimal Interconnection Circuits for VLSI,” *IEEE Transactions on Electron Devices*, Vol. ED-32, No. 5. pp. 903-909, May 1985

[53] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Repeater Insertion in Tree Structured Inductive Interconnect," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* Vol. 48, No. 5, pp. 471-481, May 2001

[54] N. Ahmed, M. H. Tehranipour, D. Zhou, and M. Nourani, “Frequency Driven Repeater Insertion for Deep Submicron,” *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 5, pp. 181–184, May 2004

[55] G. Chen and E. G. Friedman, "Low Power Repeaters Driving *RC* and *RLC* Interconnects with Delay and Bandwidth Constraints," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 14, No. 2, pp. 161-172, February 2006

[56] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Repeater Insertion in Tree Structured Inductive Interconnect," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* Vol. 48, No. 5, pp. 471-481, May 2001

[57] E. G. Friedman, “Clock Distribution Design in VLSI Circuits: An Overview,” *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1475–1478, May 1994

[58] P. Yuantao and L. Xun, “An Efficient Low Power Repeater Insertion Scheme,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 12, pp. 2726-2736, December 2006

[59] P. Yuantao and L. Xun, “Low Power Repeater Insertion with Both Delay and Slew Rate Constraints,” *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 302-307, July 2006

[60] A. B. Kahng *et al.*, "Interconnect Tuning Strategies for High-Performance ICs," *Proceedings of the IEEE Design*, *Automation and Test in Europe Conference*, pp. 471-478, February 1998

[61] K. J. Chun and Y. Ismail, "Thermal-Aware Methodology for Repeater Insertion in Low-Power VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 8, pp. 963-970, August 2007

[62] H. Kaul, D. Sylvester, and D. Blaauw, "Performance Optimization of Critical Nets Through Active Shielding," *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 51, No. 12, pp. 2417-2435, December 2004

[63] M. Ghoneima *et al.*, "Formal Derivation of Optimal Active Shielding for Low-Power On-Chip Buses," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 5, pp. 821-836, May 2006

[64] H. Kaul, D. Sylvester, and D. Blaauw "Clock Net Optimization Using Active Shielding," *Proceedings of the European Solid-State Circuits Conference*, pp. 265-268, September 2003

[65] N. Tzartzanis and W. W. Walker, "Differential Current-Mode Sensing for Efficient On-Chip Global Signaling," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 11, pp. 2141-2147, November 2005

[66] E. Mensink, D. Schinkel, E. A. M. Klumperink, E. van Tuijl, and B. Nauta, "Optimal Positions of Twists in Global On-Chip Differential Interconnects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 4, pp. 438-446, April 2007

[67] M. A. El-Moursy and E. G. Friedman, "Wire Shaping of *RLC* Interconnects," *Integration, the VLSI Journal*, Volume 40, Issue 4, pp. 461 - 472, July 2007

[68] M. A. El-Moursy and E. G. Friedman, "Exponentially Tapered H-Tree Clock Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 13, No. 8, pp. 971-975, August 2005

[69] L. Benini, P. Siegel, and G. De Micheli, "Saving Power by Synthesizing Gated Clocks for Sequential Circuits," *IEEE Design and Test of Computers*, Vol. 11, No. 4, pp. 32-41, Winter 1994

[70] F. O. Mahony, C. P. Yue, M. A. Horowitz, and S. S. Wong, "A 10-GHz Global Clock Distribution Using Coupled Standing-Wave Oscillators," *IEEE Journal of Solid-State Circuits*, Vol. 38, No. 11, pp. 1813-1820, November 2003

[71] J. Wood, T. C. Edwards, and S. Lipa, "Rotary Traveling-Wave Oscillator Arrays: A New Clock Technology," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 11, pp. 1654-1665, November 2001

[72] S. C. Chan, K. L. Shepard, and P. J. Restle, "Design of Resonant Global Clock Distributions," *Proceedings of the IEEE International Conference on Computer Design*, pp. 248-253, October 2003

[73] S. C. Chan, K. L. Shepard, and P. J. Restle, "Uniform-Phase Uniform-Amplitude Resonant-Load Global Clock Distributions**,**" *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 1, pp. 102-109, January 2005

[74] S. C. Chan, K. L. Shepard, and P. J. Restle, "Distributed Differential Oscillators for Global Clock Networks**,**" *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 9, pp. 2083-2094, September 2006

[75] M. B. Scott, E. M. Schneider, M. L. Hagge, S. Chacko and D. Stiurca, "A Low Jitter 125-1250 MHz Process Independent and Ripple-Poleless 0.18 μm CMOS PLL Based

on a Sample-Reset Loop Filter," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 11, pp. 1673-1683, November 2001

[76] W. F. Andress and D. Ham, "Standing Wave Oscillators Utilizing Wave-Adaptive Tapered Transmission Lines," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 3, pp. 638-651, March 2005

[77] Z. Yu and X. Liu, "Low-Power Rotary Clock Array Design," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 1, pp. 5-12, January 2007

[78] G. Venkataraman, J. Hu, and F. Liu, "Integrated Placement and Skew Optimization for Rotary Clocking," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 2, pp. 149-158, February 2007

[79] F. E. Anderson, J. S. Wells and E. Z. Berta, "The Core Clock System on the Next Generation Itanium Microprocessor," *Proceeding of the IEEE International Solid-State Circuits Conference,* pp. 146-147, February 2002

[80] R. Li, X. Guo, D. Yang, and K. K. O, "Wireless Clock Distribution System Using an External Antenna," *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 10, pp. 2283-2292, October 2007

[81] R. T. Chang, N. Talwalker, C. P. Yue, and S. S. Wong, "Near Speed-of-Light Signaling Over On-Chip Electrical Interconnects," *IEEE Journal of Solid-State Circuits*, Vol. 38, No. 5, pp. 834-838, May 2003

[82] Shahini, D. Shaeffer, and T. Lee. "A 12-mW Wide Dynamic Range CMOS Front-End for a Portable GPS Receiver," *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 12, pp. 2061-2070, February 1997

[83] M. Bazes. “Two Novel Fully Complementary Self-Biased CMOS Differential Amplifiers,” *IEEE Journal of Solid-State Circuits*, Vol. 26, No. 2, pp. 165-168, February 1991

[84] A. P. Jose, G. Patounakis, and K. L. Shepard, “Near Speed-of-Light On-Chip Interconnects Using Pulsed Current-Mode Signaling,” *Proceedings of the IEEE Symposium on VLSI Circuits*, pp. 108-111, June 2005

[85] A. J. Joshi, G. G. Lopez, and J. A. Davis, “Design and Optimization of On-Chip Interconnects Using Wave-Pipelined Multiplexed Routing,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 9, pp. 990-1002, September 2007

[86] P. Wang and E. C. Kan, “Pulsed Wave Interconnect,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 12, No. 5, pp. 453-463, May 2004

[87] C. Cheng, J. Lillis, S. Lin, and N. Chang, *Interconnect Analysis and Synthesis*, Wiley Interscience, 2000

[88] A. M. Niknejad and R. G. Meyer, “Analysis, Design, and Optimization of Spiral Inductors and Transformers for Si RF IC’s,” *IEEE Journal of Solid-State Circuits*, Vol. 33, No. 10, pp. 1470-1481, October 1998

[89] http://hitec.ewi.tudelft.nl

[90] http://ewh.ieee.org/r3/cnc/ed-mtt-ssc

[91] http://www.emeraldinsight.com

[92] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 2004

[93] P, Yue and S. Wong, "On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RFIC's," *IEEE Journal of Solid-State Circuits*, Vol. 33, No. 5, pp. 743-752, May 1998

[94] A. Zolfaghari, A. Chan, and B. Rezavi, "Stacked Inductors and Transformers in CMOS Technology," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 4, pp. 620-628, April 2001

[95] C. M. Tai and C. N. Liao, "Multilevel Suspended Thin-Film Inductors on Silicon Wafers," *IEEE Transactions on Electron Devices*, Vol. 54, No. 6, pp 1510-1514, June 2007

[96] H. L. Tu, I. S. Chen, P. C. Yeh, and H. K. Chiou, "High Performance Spiral Inductor on Deep-Trench-Mesh Silicon Substrate," *IEEE Microwave and Wireless Components Letters*, Vol. 16, No. 12, pp. 654-656, December 2006

[97] L. Vandi *et al.*, "A Toroidal Inductor Integrated in a Standard CMOS Process," *Analog Integrated Circuits and Signal Processing*, Vol. 50, No. 1, pp. 39-46, January 2007

[98] J. Zou, J. G. Nickel, D. Trainor, C. Liu, and J. E. Schutt-Aine, "Development of Vertical Planar Coil Inductors Using Plastic Deformation Magnetic Assembly (PDMA)," *Proceedings of the IEEE Microwave Symposium*, pp. 193-196, May 2001

[99] D. F. Svelto and R. Castello, "A Bond-Wire Inductor-MOS Varactor VCO Tunable From 1.8 to 2.4 GHz," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 1, pp. 403-410, January 2002

[100] C. Yang, T. L. Ren, F. Liu, L. T. Liu, G. Chen, X. K. Guan, and A. Wang, "On-Chip Integrated Inductors with Ferite Thin-Films for RF IC," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 1-4, December 2006

[101] D. S. Gardner *et al.*, "Integrated On-Chip Inductors with Magnetic Films," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 1-4, December 2006

[102] F. E. van Vliet, F. L. M. van den Boggart, J. L. Tauritz, and R. G. F. Baets, "Systematic Analysis, Synthesis, and Realization of Monolithic Microwave Active Inductors," *Proceedings of the IEEE International Microwave Symposium*, pp. 17-21, June 1996

[103] R. Mukhopadhyay, S. W. Yoon, Y. Park, C. H. Lee, S. Nuttinck, and J. Laskar, "Investigation of Inductors for Digital Si-CMOS Technologies," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 3750 - 3753, May 2006

[104] A. M. Niknejad and R. G. Meyer, *Design, Simulations and Applications of Inductors and Transformers for Si RF ICs*, Kluwer Academic Publishers, 2000

[105] L. N. Dworsky, *Modern Transmission Line Theory and Applications*, John Wiley & Sons, 1979

[106] J. C. Maxwell, "On Physical Lines and Forces," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Fourth series, March 1861

[107] W. H. Hayt, J. E. Kemmerly, and S. M. Durbin, *Engineering Circuit Analysis*, McGraw-Hill, 2002

[108] X. Huo, P. C. H. Chan, K. J. Chen, and H. C. Luong, "A Physical Model for On-Chip Spiral Inductors With Accurate Substrate Modeling," *IEEE Transactions on Electron Devices*, Vol. 53, No. 12, pp. 2942-2949, December 2006

[109] J. C. Guo and T. Y. Tan, "A Broadband and Scalable On-Chip Inductor Model Appropriate for Operation Modes of Varying Substrate Resistivities," *IEEE Transactions on Electron Devices*, Vol. 54, No. 11, pp. 3018-3029, November 2007

[110] H. M. Greenhouse, “Design of Planar Rectangular Microelectronic Inductors,” *IEEE Transactions on Parts, Hybrids, and Packaging*, Vol. 10, No. 2, pp. 101-109, June 1974

[111] H. A. Wheeler, “Simple Inductance Formulas for Radio Coils,” *Proceedings of the IRE*, Vol. 16, No. 10, pp. 1398-1400, October 1928

[112] S. S. Mohan, M. del Mar Hershenson, S. P. Boyd, and T. H. Lee, “Simple Accurate Expressions for Planar Inductors,” *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 10, pp. 1419-1424, October 1999

[113] http://rfic.eecs.berkeley.edu/~niknejad/asitic.html

[114] http://www.oea.com

[115] http://www.ansoft.com

[116] V. L. Chi, “Salphasic Distribution of Clock Signals for Synchronous Systems,” *IEEE Transaction on Computers*, Vol. 43, No. 5, pp. 597-602, May 1994

[117] C. Juang-Ying, C. H. Ziesler, and M. C. Papaefthymiou, “Experimental Evaluation of Resonant Clock Distribution,” *Proceedings of the IEEE Computer Society Annual Symposium on VLSI,* pp. 135-140, February 2004

[118] J. Rosenfeld and E. G. Friedman, "Design Methodology for Global Resonant H-Tree Clock Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 15, No. 2, pp. 135-148, February 2007

[119] J. Rosenfeld and E. G. Friedman, “Sensitivity Evaluation of Global Resonant H-Tree Clock Distribution Networks,” *Proceedings of the IEEE/ACM Great Lakes Symposium on VLSI*, pp. 192-197, April/May 2006

[120] S. Borkar *et al.*, "Parameter Variations and Impacts on Circuits and Microarchitecture," *Proceedings of the IEEE/ACM Conference on Design Automation*, pp. 338-342, June 2003

[121] D. Velenis, R. Sundaresha, and E. G. Friedman, "Buffer Sizing for Delay Uncertainty Induced by Process Variations," *Proceedings of the IEEE International Conference on Electronics, Circuits and Systems*, pp. 415-418, December 2004

[122] W. Dai, and H. Ji, "Timing Analysis Taking into Account Interconnect Process Variations," *Proceedings of the IEEE International Workshop on Statistical Methodology*, pp. 51-53, June 2001

[123] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI*, Addison-Wesley Publishing Company, 1990

[124] L. Sun, J. Wen, J. Yan, and J. Hu, "Modeling and Parameters Extraction of Spiral Inductors for Silicon-Based RFICs," *Proceedings of the IEEE Conference on Solid-State and Integrated Circuits Technology*, Vol. 1, No. 1, pp. 224-227, October 2004

[125] M. E. Van Valkenburg, *Network Analysis*, Prentice-Hall Inc., 1974

[126] A. Valentian and A. Amara, "On-Chip Signaling for Ultra Low-Voltage 0.13 μm CMOS SOI Technology," *Proceedings of the IEEE Northeast Workshop on Circuits and Systems*, pp. 169-172, June 2004

[127] R. Bashirullah, W. Liu, R. Cavin, and D. Edwards, "A Hybrid Current/Voltage Mode On-Chip Signaling Scheme With Adaptive Bandwidth Capability," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 12, No. 8, pp. 876-880, August 2004

[128] I. Ben Dhaou, V. Sundararajan, H. Tenhunen, and K. K. Parhi, "Energy Efficient Signaling in Deep Submicron CMOS Technology," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 319-324, March 2001

[129] V. Venkatraman, M. Anders, H. Kaul, W. Burleson, and R. Krishnamurthy, "A Low-Swing Signaling Technique for 65nm On-chip Interconnects," *Proceedings of the IEEE SoC International Conference*, pp. 289-292, September 2006

[130] E. D. Kyriakis-Bitzaros, N. Haralabidis, M. Lagadas, A. Georgakilas, Y. Moisiadis, and G. Halkias, "Realistic End-to-End Simulation of the Optoelectronic Links and Comparison with the Electrical Interconnections for System-on-Chip Applications," *IEEE Journal of Lightwave Technology*, Vol. 19, No. 10, pp. 1532-1542, October 2001

[131] M. I. Pupin, "Propagation of Long Electrical Waves," *American Institute of Electrical Engineers Transactions*, Vol. XV, pp. 93-142, March 1899

[132] M. I. Pupin, "Art of Reducing Attenuation of Electrical Waves and Apparatus Therefore," United States Patent, No. 652,230, June 19, 1900

[133] A. Marinčić, "Analysis of Transmission Line with Periodical Inductance Loading," *Microwave Review*, Vol. 10, No. 2, pp. 43-48, November 2004

[134] J. Chueh, M. C. Papaefthymiou, and C. H. Ziesler, "Two-Phase Resonant Clock Distribution," *Proceedings of the IEEE Annual Symposium on VLSI,* pp. 65-70, May 2005

[135] A. P. Jose and K. L. Shepard, "Distributed Loss Compensation for Low-Latency On-Chip Interconnects," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 392-393, February 2006

[136] V. F. Pavlidis and E. G. Friedman, *Three-Dimensional Integrated Circuit Design*, Morgan Kaufmann, 2009

[137] R. G. Gutmann *et al.*, "Three-Dimensional ICs: A Technology Platform for Integrated Systems and Opportunities for New Polymeric Adhesives", *Proceedings of the IEEE International Conference on Polymers and Adhesives in Microelectronics and Photonics*, pp. 173-180, October 2001

[138] C. S. Tan, R. J. Gutmann, and L. R. Reif, *Wafer Level 3-D ICs Process Technology*, Springer, 2008

[139] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*, John Wiley & Sons, 2006

[140] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*, Kluwer, 2001

[141] H. Kim, H.-J. Kim, and W.-S. Chung, "Pulsewidth Modulation Circuits Using CMOS OTAs," *IEEE Transaction on Circuits and Systems I: Regular Papers*, Vol. 54, No. 9, pp. 1869-1878, September 2007

[142] B. Sahu and G. A. Rincon, "An Accurate, Low-Voltage, CMOS Switching Power Supply With Adaptive On-Time Pulse-Frequency Modulation (PFM) Control," *IEEE Transaction on Circuits and Systems I: Regular Papers*, Vol. 54,No. 2, pp. 312-321, February 2007

[143] B. S. Cherkauer and E. G. Friedman, "A Unified Design Methodology for CMOS Tapered Buffers," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 3, No. 1, pp. 99-111, March 1995

[144] R. S. Kagan and M. Chi, "Improving Power Supply Efficiency with MOSFET Synchronous Rectifiers," *Proceedings of the IEEE International Conference on Polymers and Adhesives in Microelectronics and Photonics*, pp. D4.1-D4.9, July 1982

[145] V. Kursun, S. G. Narendra, V. K. De, and E. G. Friedman, "Low-Voltage-Swing Monolithic DC-DC Conversion," *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 51, No. 5, pp. 241-248, May 2004

[146] V. Kursun, S. G. Narendra, V. K. De, and E. G. Friedman, "Cascode Monolithic DC-DC Converters for Reliable Operation at High Input Voltages," *Analog Integrated Circuits and Signal Processing*, Vol. 42, No. 3, pp. 231-238, May 2005

[147] V. Kursun, V. K. De, E. G. Friedman, and S. G. Narendra, "Monolithic Voltage Conversion in Low-Voltage CMOS Technologies," *Microelectronics Journal*, Vol. 36, No. 9, pp. 863-867, May 2005

[148] W.-Z. Chen, W.-H. Chen, and K.-C. Hsu, "Three-Dimensional Fully Symmetric Inductors, Transformer, and Balun in CMOS Technology," *IEEE Transaction on Circuits and Systems I: Regular Papers*, Vol. 54, No. 7, pp. 1413-1423, July 2007

[149] J. Wibben and R. Harjani, "A High-Efficiency DC-DC Converter Using 2 nH Integrated Inductors," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 4, pp. 844-854, April 2008

[150] T. Karnik *et al.*, "High-Frequency DC-DC Conversion: Fact or Fiction," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 245-248, May 2006

[151] T. Suetsugu and M. K. Kazimierczuk, "Integration of Class DE Inverter for On-Chip DC-DC Power Supplies," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 3133-3136, May 2006

[152] P. Hazucha *et al.*, "A 233-MHz 80%-87% Efficient Four-Phase DC-DC Converter Utilizing Air-Core Inductors on Package," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 4, pp. 838-845, April 2005

[153] G. Schrom, P. Hazucha, J.-H. Hahn, V. Kursun, D. Gardner, S. Narendra, T. Karnik, and V. De, "Feasibility of Monolithic and 3D-Stacked DC-DC Converters for Microprocessors in 90nm Technology Generation," *Proceedings of the IEEE International Symposium on Low Power Electronics and Design*, pp. 263-268, August 2004

[154] J. Sun, J.-Q. Lu, D. Giuliano, T. P. Chow, and R. J. Gutmann, "3D Power Delivery for Microprocessors and High-Performance ASICs," *Proceedings of the IEEE Applied Power Electronics Conference*, pp. 127-133, February 2007

[155] K. Onizuka, K. Inagaki, H. Kawaguchi, M. Takamiya, and T. Sakurai, "Stacked-Chip Implementation of On-Chip Buck Converter for Distributed Power Supply System in SiPs," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 11, pp. 2404-2410, November 2007

[156] MIT Lincoln Laboratories, *MITLL Low-Power FDSOI CMOS Process Design Guide*, 2008

[157] Predicative Technology Model: http://www.eas.asu.edu/~ptm

[158] I. Savidis and E. G. Friedman, "Electrical Characterization and Modeling of 3-D Vias," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 784-787, May 2008

[159] M. Nagata, T. Okumoto, and K. Taki, "A Built-in Technique for Probing Power Supply and Ground Noise Distribution Within Large-Scale Digital Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 4, pp. 813-819, April 2005

[160] MIT Lincoln Laboratories, *MITLL Low-Power FDSOI CMOS Process Device Modeling*, 2008

[161] K. Usami *et al.*, "Automated Low-Power Technique Exploiting Multiple Supply Voltages Applied to a Media Processor," *IEEE Journal of Solid-State Circuits*, Vol. 33, No. 3, pp. 463-472, March 1998

[162] J. Gu, R. Harjani, and C. H. Kim, "Design and Implementation of Active Decoupling Capacitor Circuits for Power Supply Regulation in Digital ICs," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 17, No. 2, pp. 292-301, February 2009

[163] P. Hazucha *et al.*, "Area-Efficient Linear Regulator With Ultra-Fast Load Regulation," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 4, pp. 933-940, April 2005

[164] M. Al-Shyoukh, H. Lee, and R. Perez, "A Transient-Enhanced Low-Quiescent Current Low-Dropout Regulator With Buffer Impedance Attenuation," *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 8, pp. 1732-1742, August 2007

[165] R. Balczewski and R. Harjani, "Capacitive Voltage Multipliers: A High Efficiency Method to Generate Multiple On-Chip Supply Voltages," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 508-511, May 2001

[166] G. Patounakis, Y. W. Li, and K. L. Shepard, "A Fully Integrated On-Chip DC-DC Conversion and Power Management System," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 3, pp. 443-451, March 2004

[167] J. Lee, G. Hatcher, and L. Vandenberghe, "Evaluation of Fully-Integrated Switching Regulators for CMOS Process Technology," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 15, No. 9, pp. 1017-1027, September 2007

[168] P. Zhou, K. Sridharan, S. S. Sapatnekar, "Congestion-Aware Power Grid Optimization for 3D Circuits Using MIM and CMOS Decoupling Capacitors," *IEEE Asia and South Pacific Design Automation Conference*, pp. 179-184, January 2009

[169] D. Roberts *et al.*, "Application of On-Chip MIM Decoupling Capacitor for 90 nm SOI Microprocessor," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 72-75, December 2005

[170] J. Gu, R. Harjani, and C. H. Kim "Design and Implementation of Active Decoupling Capacitor Circuits for Power Supply Regulation in Digital ICs," *IEEE Journal of Solid-State Circuits*, Vol. 17, No. 2, pp. 291-301, February 2009

Appendix A

Modeling of a Resonant H-Tree Clock Network

The design of resonant clock distribution networks requires the development of an accurate model. Such a model is described in this Appendix. This model is based on characterizing the impedances of an H-tree sector using *ABCD* matrices. The model accurately characterizes an *RLC* transmission line, as well as incorporates a resistive model of an on-chip inductor. A derivation of the H-tree sector model is described in section A.1, while the *rms* input voltage is characterized in section A.2. To convert the H-tree sector into a single transmission line, a model of an equivalent shorted interconnect is provided in section A.3.

A.1 Model of an H-Tree Sector

An *ABCD* matrix represents each section of the structure illustrated in Figure 6. From transmission line theory, the *ABCD* matrix for the entire structure is a product of the individual matrices. Hence, the *ABCD* matrix for each section illustrated in Figure 5.5 is

$$M_i = \begin{bmatrix} \cosh \gamma_i l_i & Z_{0i} \sinh \gamma_i l_i \\ (1/Z_{0i}) \sinh \gamma_i l_i & \cosh \gamma_i l_i \end{bmatrix} \forall i = 1,2,3,4, \tag{A.1}$$

where Z_{0i}, γ_i, and l_i are the characteristic impedance, propagation constant, and length, respectively, of the i^{th} transmission line. The propagation constant and characteristic impedance are given by, respectively,

$$\gamma_i = \sqrt{\left(R_i/2^i + s\left(L_i/2^i\right)\right) s\left(2^i \cdot C_i\right)}, \tag{A.2}$$

$$Z_{0i} = \sqrt{\frac{R_i/2^i + s\left(L_i/2^i\right)}{s\left(2^i \cdot C_i\right)}}. \tag{A.3}$$

The *ABCD* matrix of the lumped series inductance, resistance, and capacitance at node 2 (see Figure 5.5) is

$$M_s = \begin{bmatrix} 1 & 0 \\ Y_s & 1 \end{bmatrix}, \tag{A.4}$$

where the admittance Y_s is

$$Y_s = \frac{s}{\left(L_s/4\right) \cdot s^2 + \left(R_s/4\right) \cdot s + 1/\left(4C_d\right)}. \tag{A.5}$$

The *ABCD* matrix of the effective lumped load at node 4 is

$$M_l = \begin{bmatrix} 1 & 0 \\ Y_l & 1 \end{bmatrix}, \tag{A.6}$$

where the admittance Y_l is

$$Y_l = 16 \cdot C_l \cdot s. \tag{A.7}$$

The *ABCD* matrix of the entire structure is given by (5.2) and repeated here for convenience.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = M_1 \cdot M_2 \cdot M_s \cdot M_3 \cdot M_4 \cdot M_l. \tag{A.8}$$

The constants a_0, a_1, and a_2 and the functions K, b_0, b_1, b_2, b_3, d_0, d_1, d_2, and d_3 in (5.3) and (5.4) are extracted from the *ABCD* matrix in (A.8). In order to represent (A.8) in a convenient manner, additional auxiliary functions, k_1 to k_6, are provided. The constants a_0, a_1, and a_2 represent the on-chip capacitor, effective series resistance, and on-chip inductor, respectively,

$$a_2 = L_s, \tag{A.9}$$

$$a_1 = R_s, \tag{A.10}$$

$$a_0 = 1/C_d. \tag{A.11}$$

The auxiliary functions k_1 - k_6 are defined as

$$k_1 \equiv \cosh\gamma_1 l_1 \cdot \cosh\gamma_2 l_2 + \left(Z_{01}/Z_{02}\right) \cdot \sinh\gamma_1 l_1 \cdot \sinh\gamma_2 l_2, \tag{A.12}$$

$$k_2 \equiv Z_{02} \cdot \cosh\gamma_1 l_1 \cdot \sinh\gamma_2 l_2 + Z_{01} \cdot \sinh\gamma_1 l_1 \cdot \cosh\gamma_2 l_2, \tag{A.13}$$

$$k_3 \equiv \cosh\gamma_3 l_3, \tag{A.14}$$

$$k_4 \equiv \cosh\gamma_4 l_4, \tag{A.15}$$

$$k_5 \equiv \sinh\gamma_3 l_3, \tag{A.16}$$

$$k_6 \equiv \sinh\gamma_4 l_4, \tag{A.17}$$

and an additional set of functions is defined as

$$p_1 \equiv \left(1/Z_{01}\right) \cdot \sinh\gamma_1 l_1 \cdot \cosh\gamma_2 l_2 + \left(1/Z_{02}\right) \cdot \cosh\gamma_1 l_1 \cdot \sinh\gamma_2 l_2, \tag{A.18}$$

$$p_2 \equiv \left(Z_{02}/Z_{01}\right) \cdot \sinh\gamma_1 l_1 \cdot \sinh\gamma_2 l_2 + \cosh\gamma_1 l_1 \cdot \cosh\gamma_2 l_2. \tag{A.19}$$

Finally, K, b_0, b_1, b_2, b_3, d_0, d_1, d_2, and d_3 in (5.3) and (5.4) are

$$K \equiv 1/\left(Z_{03} \cdot Z_{04}\right), \tag{A.20}$$

$$b_3 \equiv C_l \cdot a_2 \cdot Z_{03} \cdot Z_{04} \cdot \left(1 + Z_{04} \cdot k_1 \cdot k_3 \cdot k_6 + \left(Z_{04}/Z_{03}\right) \cdot k_2 \cdot k_5 \cdot k_6 + Z_{03} \cdot k_1 \cdot k_4 \cdot k_5\right), \tag{A.21}$$

$$
\begin{aligned}
b_2 \equiv\ & Z_{04}\cdot a_2\cdot k_4\cdot\left(k_2\cdot k_5+Z_{03}\cdot k_1\cdot k_3\right)+\\
& \mathrm{C}_l\cdot Z_{04}\cdot k_2\cdot k_5\left(Z_{03}^2\cdot k_4+Z_{04}\cdot a_1\cdot k_6\right)+\\
& \mathrm{C}_l\cdot Z_{03}\cdot Z_{04}^2\cdot k_3\cdot k_6\cdot\left(a_1\cdot k_1+k_2\right)+\\
& Z_{03}\cdot\left(k_2\cdot k_3+Z_{03}\cdot k_1\cdot k_5\right)\cdot\left(k_6\cdot a_2+\mathrm{C}_l\cdot Z_{04}\cdot a_1\cdot k_4\right),
\end{aligned}
\tag{A.22}
$$

$$
\begin{aligned}
b_1 \equiv\ & C_l\cdot Z_{04}\cdot a_0\cdot k_5\left(Z_{04}\cdot k_2\cdot k_5+Z_{03}^2\cdot k_1\cdot k_4\right)+\\
& C_l\cdot Z_{03}\cdot Z_{04}\cdot a_0\cdot k_3\cdot\left(k_2\cdot k_4+Z_{04}\cdot k_1\cdot k_6\right)+\\
& Z_{03}\cdot k_2\cdot k_6\cdot\left(Z_{03}\cdot k_5+a_1\cdot k_3+\left(Z_{04}\cdot k_3\cdot k_4\right)/k_6\right)+\\
& a_1\cdot\left(Z_{04}\cdot k_4\cdot\left(k_2\cdot k_5+k_1\cdot k_3\right)+Z_{03}^2\cdot k_1\cdot k_5\cdot k_6\right),
\end{aligned}
\tag{A.23}
$$

$$
b_0 \equiv Z_{04}\cdot a_0\cdot k_4\cdot\left(Z_{03}\cdot k_1\cdot k_3+k_2\cdot k_5\right)+Z_{03}\cdot a_0\cdot k_6\cdot\left(k_2\cdot k_3+k_1\cdot k_5\right),
\tag{A.24}
$$

$$
d_3 \equiv C_l\cdot Z_{04}^2\cdot a_2\cdot k_6\cdot\left(p_2\cdot k_5+Z_{03}\cdot p_1\cdot k_3\right)+C_l\cdot Z_{03}\cdot Z_{04}\cdot a_2\cdot k_4\left(Z_{03}\cdot p_1\cdot k_5+p_2\cdot k_3\right),
\tag{A.25}
$$

$$
\begin{aligned}
d_2 \equiv\ & \mathrm{C}_l\cdot Z_{04}\cdot a_1\cdot p_2\cdot\left(Z_{04}\cdot k_5\cdot k_6+Z_{03}\cdot k_3\cdot k_4\right)+\\
& Z_{04}\cdot a_2\cdot k_4\cdot\left(Z_{03}\cdot p_1\cdot k_3+p_2\cdot k_5\right)+\\
& Z_{03}\cdot a_2\cdot k_6\cdot\left(Z_{03}\cdot p_1\cdot k_5+p_2\cdot k_3\right)+\\
& C_l\cdot Z_{03}\cdot Z_{04}\cdot\left(a_1\cdot p_1+p_2\right)\cdot\left(Z_{04}\cdot k_3\cdot k_6+Z_{03}\cdot k_4\cdot k_5\right),
\end{aligned}
\tag{A.26}
$$

$$
\begin{aligned}
d_1 \equiv\ & C_l\cdot Z_{04}^2\cdot a_0\cdot k_6\cdot\left(Z_{03}\cdot p_1\cdot k_3+p_2\cdot k_5\right)+\\
& C_l\cdot Z_{03}\cdot Z_{04}\cdot a_0\cdot k_4\cdot\left(Z_{03}\cdot p_1\cdot k_5+p_2\cdot k_3\right)+\\
& \mathrm{a}_1\cdot p_2\cdot\left(Z_{04}\cdot k_4\cdot k_5+Z_{03}\cdot k_3\cdot k_6\right)+\\
& Z_{03}\cdot\left(a_1\cdot p_1+p_2\right)\cdot\left(Z_{04}\cdot k_3\cdot k_4+Z_{03}\cdot k_5\cdot k_6\right),
\end{aligned}
\tag{A.27}
$$

$$
d_0 \equiv Z_{04}\cdot a_0\cdot k_4\cdot\left(p_2\cdot k_5+Z_{03}\cdot p_1\cdot k_3\right)+\mathrm{Z}_{03}\cdot a_0\cdot k_6\cdot\left(\mathrm{Z}_{03}\cdot p_1\cdot k_5+p_2\cdot k_3\right).
\tag{A.28}
$$

A.2 RMS Input Voltage

The input clock signal driving a resonant H-tree clock sector is a square periodic waveform,

$$V_{in}(t) = \begin{cases} \frac{t}{\tau} V_{dd} & nT \le t < nT + \tau \quad \text{(A.29)} \\ V_{dd} & nT + \tau \le t < \left(n + \frac{1}{2}\right)T \quad \text{(A.30)} \\ \left(1 - \frac{t}{\tau} + \frac{T}{2\tau}\right)V_{dd} & \left(n + \frac{1}{2}\right)T \le t < \left(n + \frac{1}{2}\right)T + \tau \quad \text{(A.31)} \\ 0 & \left(n + \frac{1}{2}\right)T + \tau \le t < (n+1)T, \quad \text{(A.32)} \end{cases}$$

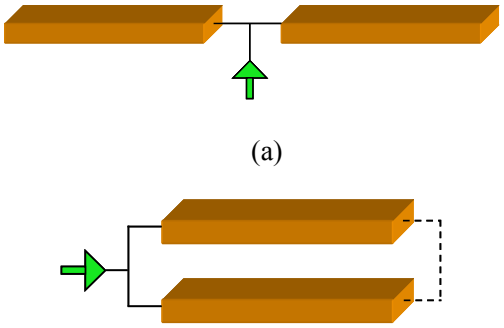
where T is the period of $V_{in}(t)$, n is an integer, and τ is the transition time. In the case study described in section 5.4, T and τ are 200 psec and 40 psec, respectively, and V_{dd} = 1.8 volts. Substituting (A.29) to (A.32) into (5.8),

$$V_{rms,g} = \sqrt{1.404} \text{ volts.} \quad \text{(A.33)}$$

A.3 Model of an Equivalent Shorted Interconnect

Assuming that the interconnect between each pair of nodes are effectively in parallel (see Figures 5.4 and A.1), the capacitance per unit length is increased by a factor of two,

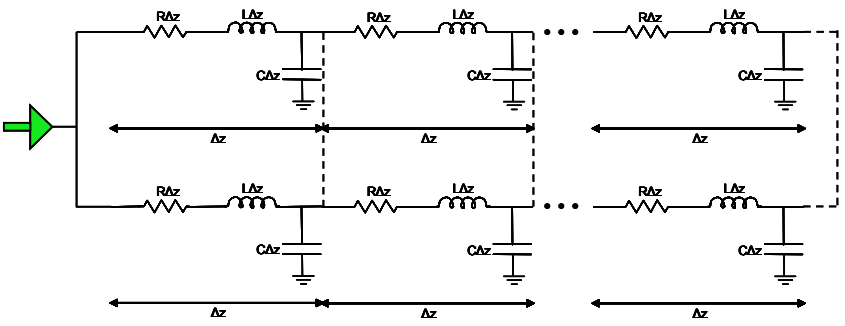
while the resistance and inductance per unit length is decreased by a factor of two at each level of the hierarchy. To justify this assumption, consider the interconnect model, as depicted in Figure A.2.



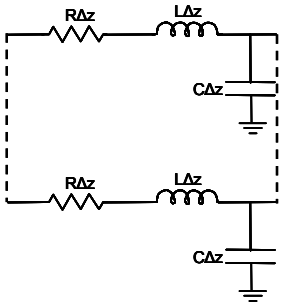
(a)

(b)

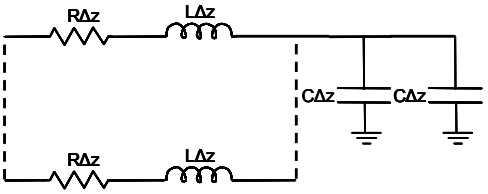
Figure A.1 First level of the resonant H-tree sector: (a) Two identical interconnects driven by a buffer, (b) Two symmetric interconnects effectively shorted



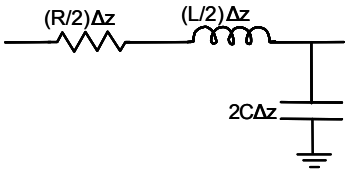
(a)



(b)



(c)



(d)

Figure A.2 Distributed model of two shorted interconnects: (a) Model of two identical shorted interconnects, (b) Two shorted infinitesimally small sections, (c) Two shorted sections in a different configuration, (d) Equivalent model of two shorted sections

Since all of the nodes depicted in Figure A.2(a) are symmetric, the waveforms at these nodes are identical. As a result, the infinitesimally small sections can be assumed to be effectively shorted (as illustrated by the dashed lines). The total impedance of each of the two shorted sections is determined from the circuit depicted in Figure A.2(c). The total shunted capacitance is

$$C_{tot} = (2C)\Delta z, \tag{A.34}$$

while the total impedance is

$$Z_{R,L,tot} = 1 \Big/ \left(\frac{1}{R\Delta z + sL\Delta z} + \frac{1}{R\Delta z + sL\Delta z} \right) = \left(\frac{R}{2} \right) \Delta z + s \left(\frac{L}{2} \right) \Delta z. \tag{A.35}$$

Appendix B

Modeling of a Quasi-Resonant Interconnect

Power and delay expressions of a quasi-resonant transmission network are derived based on the model described in this Appendix. The transmission portion of the network composed of an *RLC* transmission line and an on-chip inductor are modeled based on *ABCD* matrices. In section B.1, the *ABCD* matrices characterizing the transmission line and on-chip inductor are defined, and a derivation of the 50% delay is described in section B.2.

B.1 Definition of *ABCD* matrices

The *ABCD* matrix of the transmission line depicted in Figure 6.6 is

$$M_{\text{int}1} = \begin{bmatrix} \cosh\gamma l & Z_0 \sinh\gamma l \\ \frac{1}{Z_0}\sinh\gamma l & \cosh\gamma l \end{bmatrix}, \tag{B.1}$$

where Z_0, γ, and l are the characteristic impedance, propagation constant, and line length, respectively. Similarly, M_{int2} is

$$M_{\text{int}2} = \begin{bmatrix} \cosh\gamma(l-l_d) & Z_0 \sinh\gamma(l-l_d) \\ \frac{1}{Z_0}\sinh\gamma(l-l_d) & \cosh\gamma(l-l_d) \end{bmatrix}. \tag{B.2}$$

To derive the *ABCD* matrix M_s, the spiral inductor depicted in Figure 6.2 is simplified into the network shown in Figure B.1.

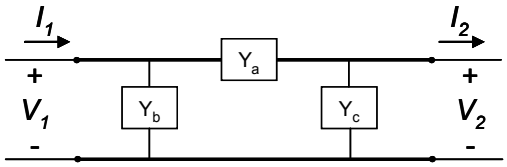


Figure B.1 Simplified admittance network of the spiral inductor model

From this network, the *ABCD* matrix is

$$M_s = \begin{bmatrix} 1+\dfrac{Y_c}{Y_a} & \dfrac{1}{Y_a} \\ Y_b + Y_c + \dfrac{Y_b Y_c}{Y_a} & 1+\dfrac{Y_1}{Y_3} \end{bmatrix}, \tag{B.3}$$

where

$$Y_a = \frac{1}{j\omega L_{series} + \dfrac{R_{ac}\left(R_{skin} + j\omega L_{skin}\right)}{R_{ac} + R_{skin} + j\omega L_{skin}}} + j\omega C_p, \tag{B.4}$$

$$Y_b = \frac{1}{1/\left(j\omega C_{ox1}\right) + \left(R_{sub1}/j\omega C_{sub1}\right)/\left(R_{sub1} + 1/j\omega C_{sub1}\right)}, \tag{B.5}$$

$$Y_c = \frac{1}{1/\left(j\omega C_{ox2}\right) + \left(R_{sub2}/j\omega C_{sub2}\right)/\left(R_{sub2} + 1/j\omega C_{sub2}\right)}. \tag{B.6}$$

Finally, the *ABCD* matrix of the capacitive load is

$$M_l = \begin{bmatrix} 1 & 0 \\ j\omega C_l & 1 \end{bmatrix}. \tag{B.7}$$

B.2 Delay of a Quasi-Resonant Interconnect

The transfer function of the QRN at each angular frequency can be represented as

$$H'(j\omega) = \left|H'(\omega)\right| e^{j\beta(\omega)}, \tag{B.8}$$

where $\beta(\omega)$ is the phase of the transfer function. The output waveform in the time domain of a linear system, described by a transfer function such as (B.8) and stimulated by a periodic square signal, can be approximated by

$$V_{out}(t) \approx \frac{V_{dd}}{2} + A_1 \sin\left(\omega_p t + \phi_1\right) + A_3 \sin\left(3\omega_p t + \phi_3\right), \tag{B.9}$$

where

$$A_{m=1,3} = \frac{2t_p V_{dd}}{t_r m^2 \pi^2} \left| \sin\left(\frac{m\omega_p t_r}{2}\right) \right| \left| H'\left(m\omega_p\right) \right| \tag{B.10}$$

and

$$\phi_{m=1,3} = -\frac{m\omega_p t_r}{2} + \beta\left(m\omega_p\right). \tag{B.11}$$

Note that the approximated closed-form solution in (B.9) is available only when considering the first three harmonics of the input signal. To determine the 50% delay, (B.9) is set to V_{dd} / 2. A third-order trigonometric expression can be obtained as

$$a_3 x^3 + a_2 x^2 + a_1 x + a_0 = 0, \tag{B.12}$$

where $x = \tan(\omega_p t)$ and

$$a_0 = A_1 \sin\phi_1 + A_3 \sin\phi_3, \tag{B.13}$$

$$a_1 = A_1 \cos\phi_1 + 3A_3 \cos\phi_3, \tag{B.14}$$

$$a_2 = A_1 \sin\phi_1 - 3A_3 \sin\phi_3, \tag{B.15}$$

$$a_3 = A_1 \cos\phi_1 - A_3 \cos\phi_3. \tag{B.16}$$

A third order polynomial such as (B.12) has either one or three real roots, and a closed-form solution exists. In the case of a single real root x_0, a 50% delay can be determined from (6.17).