DESIGN AND ANALYSIS OF ACTIVE AND PASSIVE DECOUPLING CAPACITORS FOR ON-CHIP POWER SUPPLY NOISE MANAGEMENT

by

XIONGFEI MENG

B.A.Sc., The University of British Columbia, 2004 M.A.Sc., The University of British Columbia, 2006

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES (Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA (Vancouver)

July 2009

© Xiongfei Meng, 2009

ABSTRACT

On-chip decoupling capacitors (decaps) in the form of MOS transistors are widely used to reduce power supply noise in both standard-cell blocks and white spaces between blocks. This research provides guidelines for layouts of decaps that properly tradeoff high-frequency response, electrostatic discharge (ESD) reliability and gate tunneling leakage for use within standard-cell blocks in ASIC designs in 90nm and 65nm CMOS technologies. A simple but effective metric is developed to determine the optimal decap layout based on the frequency response. Novel active designs are also presented.

If an *IR*-drop violation (hot spot) is found after the physical design is completed, it is usually difficult to implement a quick fix to the problem. In this dissertation, the use of an active decap in white-space areas as a drop-in replacement for passive decaps is investigated to provide noise reduction for these "hot-spot" problems found late in the design process. A modified active decap design is proposed for ASIC applications operating up to 1GHz, and the use of latch-based comparators provides a better power-delay trade-off. Measurement results from a test chip show that the noise reduction using active decaps improves as operating frequency increases, and provides between 10%-20% noise reduction at 200MHz-1GHz over its passive counterpart.

The concept of active decap is further extended to achieve lower supply noise. It is found that an active decap with a stack height of three (i.e., number of pieces switching) provides the best noise reduction if the supply noise level is between 7%-14%, but a stack height of two is best if the noise level is between 14%-16%. In addition, a novel charge-borrowing decap circuit is introduced which outperforms all forms of active decaps for a fixed area in terms of removing local hot spots.

TABLE OF CONTENTS

Abstractii
Table of Contents iii
List of Tables
List of Figures
Acknowledgments xii
Chapter 1 Introduction
1.1 Motivation1
1.2 Research Objectives
1.3 Organization of the Dissertation
Chapter 2 Background
2.1 Introduction
2.2 Decoupling Capacitor Basics and Design Challenges
2.3 Thin-Oxide Gate Tunneling Leakage12
2.4 Electrostatic Discharge Reliability in Decap Design
2.5 Standard-Cell Decap Layout and Placement
2.6 Metric for Power Supply Noise Management
2.7 Summary
Chapter 3 Passive Decoupling Capacitor Design
3.1 Introduction
3.2 High-Frequency Response of Decoupling Capacitors
3.3 Cross-Coupled Decoupling Capacitor Designs

3.4 Summary	
Chapter 4 Active Decoupling Capacitor Design	50
4.1 Introduction	50
4.2 Active Decoupling Capacitor Analysis and Design	
4.2.1 Active Decap Concept and Design Considerations	
4.2.2 Overall Active Decap Architecture	
4.2.3 Design Specifications	61
4.2.4 Latch-Based Comparator Design	63
4.3 Chip Design and Experimental Results	
4.3.1 Test Chip Setup	
4.3.2 Test Chip Simulations	75
4.3.3 Test Chip Measurements	
4.3.4 Measurement Results on One Typical Chip	
4.4 Active Decap Size and Placement	
4.5 Summary	
Chapter 5 Generalized Active Decap and Charge-Borrowing Decap	
5.1 Introduction	
5.2 Extended Active Decoupling Capacitor	
5.2.1 Optimal Stack Height <i>n</i>	
5.2.2 Design and Layout of <i>n</i> =3 Extended Active Decap	
5.2.3 Simulation Results	
5.3 Charge-Borrowing Decap (CBD)	
5.3.1 Charge-Borrowing Decap Concept	
5.3.2 "Clk" Signal Generation	
5.3.3 Design of Charge-Borrowing Decap	
5.3.4 Simulation Results	
5.4 Test Chip Setup and Measurement Results	

5.5 Summary		
Chapter 6 Conclusions and Future Work	130	
6.1 Summary and Conclusions		
6.2 Contributions in this Dissertation		
6.3 Future Work		
Appendix: Comparator Design Fundamentals135		
References		

5

LIST OF TABLES

Table 1.1: Comparison on active and passive decap implementations
Table 3.1: Optimal number of fingers for different frequency ranges
Table 3.2: Comparison of the passive decap designs and their gate leakage current. 45
Table 4.1: Design specifications of the active decap
Table 4.2: Transistor sizes of the comparators. 65
Table 4.3: Simulated switching circuit design specification comparison
Table 4.4: Comparator delay t_d and delay difference Δt_d in different corners
Table 4.5: Measured active decap performance for different process corners. 82
Table 4.6: Comparison between equation and simulated result after correlation
Table 4.7: Active decap bandwidth versus average comparator delay under process corners 83
Table 5.1: Optimal stack height <i>n</i> selection based on the supply noise <i>k</i> (from formula)
Table 5.2: Optimal stack height n selection based on the supply noise k (from simulation) 105
Table 5.3: High-level comparison among passive, active, and charge-borrowing decaps 108

LIST OF FIGURES

Figure 1.1: Effectiveness of on-chip decaps: (a) without decaps, the power supply noise level can
be high, compared to (b) where the noise level is low due to the use of decaps
Figure 2.1: Decoupling capacitor implemented using an NMOS device
Figure 2.2: Cross-coupled decap schematic [21]10
Figure 2.3: Active decap concept shown in (a) and two previous designs from (b) [37][38] and (c)
[39]11
Figure 2.4: Gate leakage current versus gate area
Figure 2.5: Gate leakage current density J _{leak} versus oxide thickness t _{ox}
Figure 2.6: Complete ESD protection scheme
Figure 2.7: Simulation setup for ESD analysis [24]
Figure 2.8: Sample layout of standard-cell N+P decap (a) with one finger and (b) with two
fingers
Figure 2.9: DVD _{avg} and DVD _{max} : metric used to evaluate DVD profiles
Figure 3.1: (a) A standard cell decap can be implemented as a NMOS in parallel with a PMOS
device. The corresponding layout is shown in (b)
Figure 3.2: A decap can be implemented as a NMOS device and modeled as a lumped-RC circuit
with effective resistance and effective capacitance as functions of frequency, f
Figure 3.3: The circuit setup to extract the effective resistance and the effective capacitance
values from an ac analysis
Figure 3.4: Plots of C_{eff} and R_{eff} for three device sizes (W x L): 10x10µm, 15x5µm, and 5x15µm.
Figure 3.5: Plots of C _{eff} and R _{eff} for three NMOS devices (HSPICE versus model)

Figure 3.6: The effective capacitance, $C_{eff}(f)$, of NMOS and PMOS decaps in 90nm for different			
numbers of fingers in a fixed area of Y=2µm and X=9µm			
Figure 3.7: The effective capacitance, $C_{eff}(f)$, of NMOS and PMOS decaps in 90nm for different			
numbers of fingers in a fixed area of Y=2 μ m and X=9 μ m			
Figure 3.8: $C_{eff}(f)$ and $R_{eff}(f)$ comparison of fixed-area standard decap and cross-coupled decap:			
same MOS device sizes but different poly connections			
Figure 3.9: Sample layouts of cross-coupled decap cells for (a) 3N-4P (b) 8N-9P 43			
Figure 3.10: Sample layouts of improved decap cells for (a) 1N-9P (b) 1N-16P			
Figure 3.11: Frequency response of various cross-coupled designs			
Figure 4.1: Active decap concept and its MOS implementation			
Figure 4.2: The reductive factors f and g for the boosted voltage as a function of (a) "on"			
resistances of the switches, R_{on} , and (b) leakage due to the size of decap C_{decap}			
Figure 4.3: Active decap architecture			
Figure 4.4: Difference in the input voltages of the comparators as a function of supply noise k . 58			
Figure 4.5: The reductive factor h for the boosted voltage as a function of delay difference Δt_d . 61			
Figure 4.6: Complementary comparator design: (a) n-type input for the top comparator, and (b)			
p-type input for the bottom comparator			
Figure 4.7: (a) DC and (b) AC (compensated) characteristic curves for the two-stage latch-based			
comparator design (n-type input shown)			
Figure 4.8: AC characteristic curves for the three designs			
Figure 4.9: Test chip setup			
Figure 4.10: Layout of active decap showing the relative size of the components			
Figure 4.11: Final voltage a as a function of sensing and switching circuitry area overhead x 72			

.

Figure 4.12: Annotated test chip microphotograph
Figure 4.13: Simulated V_{DD} voltage (on a 500MHz clock) with active decap on and off
Figure 4.14: Simulated V_{DD} voltage with active decap on for different process corners
Figure 4.15: Grouped scatter plot for active decap noise reduction for the tested sample chips 79
Figure 4.16: Comparator delay t_d and delay difference Δt_d as a function of supply noise k 80
Figure 4.17: Simulated average V_{DD} voltage with active decap on and off versus clock frequency
for (a) slow, (b) typical, and (c) fast process corners
Figure 4.18: Measured results (on a 500MHz clock) for (a) active decap on and (b) plotted
comparison between active decap on and off
Figure 4.19: Measured (0.2 - 1GHz) and simulated (1 - 2.5GHz) average V_{DD} voltage with active
decap on and off versus clock frequency
Figure 4.20: Simulated average V_{DD} noise per clock cycle versus normalized decap size90
Figure 4.21: Power supply noise reduction difference from active decap and passive decap with
area overhead from switching circuit of active decap
Figure 4.22: Improvement on average V_{DD} noise for using active decaps in different placement
locations by varying Rdist and Rmesh92
Figure 5.1: Active decap concept showing different stack height: (a) $n=2$, (b) $n=3$, and (c) $n=4.96$
Figure 5.2: MOS implementation of the extended active decap (<i>n</i> =3)
Figure 5.3: Final voltage aV_{DD} as a function of stack height <i>n</i> with <i>k</i> varying (fixed area)
Figure 5.4: Final voltage aV_{DD} as a function of k with different stack height n (from formula).

Figure 5.6: Layout of extended active decap $(n=3)$ showing the relative size of the components.
Figure 5.7: Simulated V _{DD} voltage with extended active decap ($n=3$) on for two different k levels.
Figure 5.8: Average V_{DD} voltage as a function of k with different stack height n (from
simulation)
Figure 5.9: Charge-borrowing decap concept shown in (b), compared to a passive decap in (a).
Figure 5.10: Diode inserted charge-borrowing decap showing the states of the clocking signal: (a)
Clk at 0, (b) Clk rises to V_{DD} , and (c) Clk falls back to 0
Figure 5.11: Possible implementations of charge-borrowing decap with (a) NMOS-formed
diodes and (b) PMOS-formed diodes
Figure 5.12: Improved implementation of charge-borrowing decap with PMOS formed diodes.
Figure 5.13: Generation of the boosted voltage on node B2
Figure 5.14: "Clk" signal generation using ring oscillator
Figure 5.15: Buffer size determines the edge delay while the buffer chain (with ring oscillator)
consumes dynamic power115
Figure 5.16: Complete circuit diagram of charge-borrowing decap 117
Figure 5.17: Simulated V_{DD} voltage (on a 1GHz clock) with a CBD on and off (best case) 119
Figure 5.18: Simulated V_{DD} voltage (on a 1GHz clock) with a CBD on and off (worst case) 120
Figure 5.19: Simulated average V_{DD} voltage as a function of k showing the case of CBD 121
Figure 5.20: Test chip 2 setup 122

Figure 5.21: Layout of charge-borrowing decap showing the relative size of the components. 122
Figure 5.22: (a) Annotated microphotograph and (b) layout of the second test chip 123
Figure 5.23: Scatter plot comparing average V_{DD} for the tested sample chips
Figure 5.24: Superimposed waveforms showing a CBD and a passive decap on a 1GHz clock. In
the top panel of the figure, the upper trace is when the CBD is on, while the lower trace is
when the passive decap is on 126
Figure 5.25: Measured average V_{DD} voltages at different clock frequencies
Figure A.1: A differential input, single-ended output comparator symbol
Figure A.2: DC transfer characteristics of (a) an ideal comparator and (b) a practical comparator
with finite gain and offset voltage
Figure A.3: Circuit schematics showing positive feedback: (a) a basic differential pair with
diode-connected load, (b) a gain-enhanced differential pair, and (c) a differential pair using
positive feedback to provide increased gain

ACKNOWLEDGMENTS

I would like to express my gratitude to my academic supervisor, Dr. Resve Saleh, whose expertise, understanding, encouragement, and support, added significantly to my graduate experience. I appreciate his profound and vast knowledge in many areas, both within and outside of the scope of my research. I would like to thank the exam committee, Dr. Steve Wilton, Dr. David Pulfrey, Dr. Mark Greenstreet, Dr. John Madden, Dr. Lutz Lampe and Dr. Tim Salcudean, for reviewing this dissertation and providing valuable feedback.

A very special thanks goes to my colleagues and friends in the SoC group, for their technical advice and kindness. More specifically, I would like to thank Dr. Shahriar Mirabbasi, Dr. Roberto Rosales, Dipanjan Sengupta, Dr. Mehdi Alimadadi, Dr. Samad Sheikhaei, Jeff Mueller, and Sohaib Majzoub. I also acknowledge Dr. Karim Arabi of Qualcomm and Asad Shayan of PMC-Sierra for their suggestions and help in this study.

I recognize that this research would not have been possible without the financial support from NSERC and PMC-Sierra Inc., and express my gratitude to them. I also thank CMC Microsystems for providing chip fabrication and the CAD tools.

Last but not least, I would especially like to thank my family for both giving and encouraging me to seek for myself a demanding and meaningful education. In particular, I must acknowledge my wife, Liming, for her love, caring and patience through many years of my life. I would not have accomplished this research without her support. The appreciation extends beyond any words at my command.

xii

Chapter 1 Introduction

1.1 Motivation

Scaling of CMOS technology allows higher speed and higher functional density. As the clock frequency increases and the supply voltage decreases to about 1V, maintaining the quality of the power supply has become a primary issue. Power supply noise in the form of voltage variations arises due to *IR* drop and *Ldi/dt* effects [1]. The *IR* drop has been increasing over time due to increased resistances in the power grid as the metal widths continue to shrink with each successive technology generation. The inductive *Ldi/dt* effects are also increasing due to the higher current demands in more complex chips. However, the pin and package inductance overwhelms the inductance of the on-chip power distribution network, and therefore the on-chip inductance can be neglected [2], although the on-chip inductance may be considered in certain applications [3][4]. Having the two components together, the overall voltage drop ΔV_{drop} at any point in the power grid is [5][6][7]:

$$\Delta V_{drop} = I_{\text{supply}} \cdot R_{\text{mesh}} + L_{\text{pack}} \frac{dI_{\text{supply}}}{dt}$$
(1.1)

where R_{mesh} is the power grid (mesh) resistance, L_{pack} is the package and pin inductance, and I_{supply} is the current flow through the user logic circuits.



Figure 1.1: Effectiveness of on-chip decaps: (a) without decaps, the power supply noise level can be high, compared to (b) where the noise level is low due to the use of decaps.

A variety of different methods can be used to manage supply voltage drops. Among them, the most popular is to use on-chip decoupling capacitors (decaps) to maintain the power supply within a certain percentage (e.g., 10%) of the nominal supply voltage [1][6]. Decaps are typically placed in regions between areas of high current demand and the power pads and I/O pins [7][8][9][10][11]. The effectiveness of decaps can be illustrated conceptually in Fig. 1.1, where the power system can be modeled as a distributed *RLC* network [8][12][13][14]. The power supply noise level of Fig. 1.1(a) is reduced in Fig. 1.1(b) due to the use of decaps.

In Application Specific Integrated Circuit (ASIC) designs, two types of decaps can be identified: white-space decaps and standard-cell decaps. White-space decaps are placed in the open areas of the chip between intellectual property (IP) blocks, and are made from passive decaps or active decaps. Standard-cell decaps are always passive decaps placed within the IP blocks themselves [9], typically as filler cells.

Passive decaps can be implemented with MOS transistors or metal-insulator-metal (MIM) capacitors. White-space decaps are usually implemented with NMOS transistors since they

provide a high capacitance value per unit area. In certain applications, PMOS devices and MIM capacitors [15][16][17][18] are two alternatives for white-space decaps. Standard-cell decaps normally use both NMOS and PMOS devices. A more recent approach is the active decap which requires dynamic switching of passive decaps to boost up the power rail voltages when excessive voltage drop is detected. Due to its area requirement, it can only be used in the open areas between blocks. The implementation of active and passive decaps in terms of NMOS and PMOS devices in either white-space areas or standard cells can be illustrated in Table 1.1. This research addresses design and implementation issues for both active and passive decaps.

	Active decap	Passive decap
White space	NMOS and PMOS	NMOS or PMOS (or MIM)
Standard cell	Not used	NMOS and PMOS

Table 1.1: Comparison on active and passive decap implementations.

The lack of sufficient decaps can result in unsatisfactory timing and even functional failure for the logic circuits and memory cells [19]. On the other hand, overdesign may cost too much area. It is necessary to develop a metric to evaluate the decap effectiveness in terms of power supply noise management [20]. Starting from 90nm, a number of relatively new issues [21] must be addressed that impact the design and layout of decaps. This research addresses three important decap problems including frequency response [22][23], electrostatic discharge (ESD) protection [24], and gate tunneling leakage [6][25][26][27][28][29][30]. The frequency response controls the performance of decaps at increasingly higher operating frequencies. A potential ESD event across a thin gate oxide increases the likelihood that a chip will be permanently damaged due to a short circuit in the decap itself. Higher gate leakage significantly increases the total static power consumption of the chip.

In white-space areas, prior to the 90nm technology, the use of passive decaps was sufficient. At 90nm, the oxide thickness has been reduced to 2nm or less. Therefore, decaps have been redesigned into a cross-coupled form [21] to protect the device from potential electrostatic discharge (ESD) induced oxide breakdown [24]. However, the additional series resistance considerably reduces the transient response of the decap [23]. As a result, large *IR*-drop levels in localized regions (usually called "*hot-spot*" *IR*-drop violations) may unexpectedly be present in high-speed ASICs. These unresolved hot spots cause timing closure problems or result in functional failures in extreme cases. To remove them, designers must consider many options such as moving the logic blocks, adding or rearranging the power pins, and/or modifying power grid design. Near the tapeout deadline, such time-consuming design iterations may not always be feasible. In these situations, an active decap design can play an important role in reducing supply noise without major changes to the design [31]. That is, the active decap can be a drop-in replacement of the passive decap in an attempt to remove any remaining hot spots, thereby saving time and effort [32].

The concept of active decap can be extended with more switches and decaps to ideally achieve lower supply noise. These extended active decap designs need to be evaluated for advantages and limitations so that an optimal design can be made practical. In addition, a novel circuit called charge-borrowing decap that transfers charge from a clean supply node to a noisy supply node will be introduced to produce superior performance to the basic and the extended active decaps.

1.2 Research Objectives

Research Statement:

To investigate new designs and proper placement of active and passive decoupling capacitors to efficiently manage on-chip power supply noise in white-space areas and standard cells for ASIC applications.

Specific Research Goals and Contributions:

- Design passive decaps in standard-cell arrays that properly trade off gate leakage, ESD and transient response, and provide decap design metrics to determine the optimal layout to obtain a desired capacitance level over a target operating frequency. Develop empirical formulae with only a few parameters to capture the frequency responses for 90nm and 65nm CMOS technologies.
- Design the basic active decap to provide better power-delay trade-off than prior approaches for white-space hot-spot removal. Identify and resolve limitations of the active decap in terms of suitability for ASIC applications in deep-submicron technologies. Explore the placement of active decaps to remove late-stage *IR*-drop violations. Validate the design of the active decap using a chip design that provides testing mechanisms to evaluate improvements in dynamic voltage drops.
- Extend the concept of active decap by modifying its design for improved supply noise management. Achieve a better active decap design that provides a higher level of power supply noise reduction than the basic form of active decaps within a fixed area. Propose a novel circuit that outperforms both the basic and the extended active decaps with help from a clean supply rail.

1.3 Organization of the Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 provides the necessary background on decap design basics and challenges, gate tunneling leakage through decaps, ESD reliability of thin-oxide gates, standard-cell layout and placement of decaps, and metrics for power supply noise management.

Chapter 3 develops a set of new passive decap designs based on the cross-coupled decap. The modeling of the new designs is described and design metrics are provided to allow hand calculations and analyses to be carried out. Based on the simulation results, the proper layout of these designs is described.

Chapter 4 proposes a modified active decap design for hot spot removal in ASIC applications. The design advantages and disadvantages are compared against prior work. Measurement results from a test chip are used to validate the design. After correlation with measurement results, further simulation is carried out to explore the efficiency of active decap placement.

Chapter 5 extends the concept of the active decap to achieve a better design that has a higher level of power supply noise reduction. Also presented in the chapter is a novel charge-borrowing decap circuit that outperforms the basic and the extended active decaps.

Chapter 6 summarizes the results of the dissertation and provides conclusions. Future research directions are provided.

Chapter 2 Background

2.1 Introduction

The topics in this chapter provide the necessary background for the rest of the dissertation. Some fundamental and practical decap design issues are also highlighted to motivate the topics in the remainder of the dissertation. This chapter begins with an overview of design challenges and problems associated with decoupling capacitors in 90nm and 65nm. The overview includes gate tunneling leakage, electrostatic discharge phenomenon and protection, and standard-cell decap placement. Gate leakage is introduced from a physical point of view, and useful information from recent technologies is given. ESD reliability is presented and typical phenomena during an ESD event are discussed. Primary and local ESD protection schemes are briefly illustrated. Since ASIC designs typically utilize standard cells, the decap insertion and placement procedure within standard-cell blocks is briefly introduced. A simple metric for supply noise management is proposed to assess the profiles showing power supply noise and to compare the designs providing different decap performance.

2.2 Decoupling Capacitor Basics and Design Challenges

A passive decap in the white spaces of a chip can be implemented using an NMOS transistor with the gate connected to V_{DD} and both source and drain connected to V_{SS} , as shown in Fig. 2.1. This approach is considered effective because the thin-oxide capacitance of the transistor gate provides a higher capacitance than any other oxide capacitance available in a standard CMOS fabrication process [21]. For design purposes, an approximation for hand calculating the capacitance of this MOS decap can be given by [7]:

$$C_{decap} \approx W \cdot L \cdot C_{OX} \tag{2.1}$$

where W is the transistor width, L is the transistor length, and C_{OX} is the oxide capacitance per unit area. A more accurate capacitance model needs to include the parasitic fringing and overlap capacitance of the transistor, and will be discussed in greater detail in Chapter 3. Passive whitespace decaps can also be implemented with thick-oxide MOS devices, depending on the requirements on ESD and leakage, knowing that there is a capacitance density penalty.



Figure 2.1: Decoupling capacitor implemented using an NMOS device.

In the past, the analysis techniques and design metrics dealing with power supply voltage drop were overly simplistic [33][34][35]. Designers analyzed power supply noise with static voltage drop (SVD) analysis, which might not reflect the true nature of power supply fluctuations, leading to either unnecessary overdesign or risk of timing failures [20]. Although SVD analysis can provide useful feedback in terms of certain glaring errors in the power grid design, it does not take into account the impact of decaps and many other important factors. Dynamic voltage drop (DVD) analysis is emerging as a replacement of SVD analysis to capture the impact of decaps, package inductance, and simultaneous switching events. The drawback is that DVD analysis does not return a fixed value that can assess the degree of improvement. Currently, there is no signoff or analysis metric to characterize a DVD profile [36]. Therefore, a good metric for DVD analysis is desired to evaluate decap design and placement, for the purposes of this research.

At 90nm, the oxide thickness has been reduced to about 2nm or less. Oxide thickness reduction causes two problems: possible oxide breakdown during an ESD event and increased leakage current. ESD is a transient process of static charge transfer that can typically arise from human contact with any IC pin [24]. Additional input resistance can be inserted in series with passive decaps to protect from ESD. However, this input resistance causes the decap to suffer from the degraded frequency response, resulting in a poor performance in terms of managing power supply noise. Moreover, increased gate leakage should be considered. If decaps can be disconnected from the power rails when they are not needed (e.g., the logic circuit nearby is quiet), gate leakage reduction can be achieved. Therefore, overdesign of decaps should be avoided.

Standard-cell layouts of an IP block consist of rows of fixed-height cells in the ASIC design flow. Typically, decap filler cells have both NMOS and PMOS devices. From the 90nm node, a crosscoupled decap design has been proposed by cell library developers [21] to address the issue of ESD reliability, as shown in Fig. 2.2. The cross-coupled design provides additional ESD protection to the thin-oxide gate of the device [21]. Standard-cell decaps are generally implemented with thin-oxide MOS devices.



Figure 2.2: Cross-coupled decap schematic [21].

The major concern for white-space decaps at 90nm and 65nm is a reduced budget for power supply noise as the supply voltage decreases to about 1V [19]. In certain situations, the use of passive decaps only in white-space areas is not sufficient and hot spots at certain locations may appear at a late design stage. Active decaps in the white space may be used to remove hot spots. The goal is to investigate strengths, limitations, design issues and placement strategies for active decaps in ASIC applications.

Active decaps were originally intended for custom designs. Our goal is to optimize them for ASIC designs. Two active decap circuits using switched capacitances were proposed in the past to regulate the supply voltages [37][38][39]. The basic concept of the active decap is to switch a pair of passive decaps C_{decap} either in parallel or in series to increase charge delivery capability, as shown in Fig. 2.3(a).









Figure 2.3: Active decap concept shown in (a) and two previous designs from (b) [37][38] and (c) [39].

The two designs in [37] and [39] are effective for power supply noise reduction, but they also have certain limitations. The design in [37] can respond quickly to supply noise but dissipates large power, whereas [39] saves power but experiences long switching delays. Therefore, an improved design with a better power-delay trade-off is desired. The two previous designs [37][38][39] are illustrated in Fig. 2.3(b) and 2.3(c), respectively. Ref. [37][38][39] mitigate the effects of *LC* resonance typically in the 20-400MHz band [40]. Recent work has been done to reduce *LC* resonance using a switched decap technique [41]. Researchers have also reported on a distributed active decap to greatly boost the effective decap value while reducing the area requirements [41]. In addition, the issue of directly replacing an area occupied by a passive decap with an active decap needs to be addressed to determine the degree of noise reduction that can be obtained and the associated tradeoffs of area and power.

2.3 Thin-Oxide Gate Tunneling Leakage

In 90nm and 65nm processes, a new design issue for decaps due to oxide thickness reduction is the thin-oxide gate tunneling current. The current is in the form of tunneling electrons or holes from substrate to gate or from gate to substrate through the gate oxide, depending on the voltage biasing conditions [26]. Two forms of gate tunneling exist: Fowler–Nordheim (FN) tunneling and direct tunneling. For normal operations on short-channel devices, FN tunneling is negligible, and direct tunneling is dominant [26]. In the case of direct tunneling, the gate leakage current in PMOS is much less than in NMOS, and it has been shown experimentally that PMOS gate leakage is roughly three times smaller than NMOS gate leakage for same size transistors [27][43]. The gate leakage simulations can be carried out by using BSIM4 SPICE models [44][45]. Assuming a 90nm technology with 1.7nm oxide thickness and 1.0V power supply, the gate leakage current I_{leak} is shown in Fig. 2.4.



Figure 2.4: Gate leakage current versus gate area.



Figure 2.5: Gate leakage current density J_{leak} versus oxide thickness t_{ox} .

Clearly, as indicated from simulation results, the gate leakage is proportional to the transistor area. That is,

$$I_{leak} = J_{leak} \cdot WL \tag{2.2}$$

where J_{leak} is the gate leakage current density. From simulation, the decoupling capacitance values of the transistor, C_{decap} , are also shown in the figure. As described earlier, C_{decap} is equal to WLC_{OX} to the first order. Since C_{OX} is fixed in this case, C_{decap} is proportional to the decap area of WL.

The gate leakage current density J_{leak} and the oxide thickness t_{OX} have an empirical relationship as follows, assuming the voltage across the oxide V_{OX} is fixed [43]:

$$\log J_{leak} = K_1 - K_2 \cdot t_{OX} \tag{2.3}$$

where K_1 and K_2 are non-negative experimental constants and are process dependent. Equation (2.3) implies that the gate leakage current is exponentially related to the oxide thickness. A typical J_{leak} and t_{OX} relationship for a fixed V_{OX} at 1V is illustrated in Fig. 2.5.

It is evident that at 90nm and 65nm technologies, the gate leakage from decaps will be significant [27]. The gate leakage contributes to the total static power consumption, and decaps usually occupy a large on-chip area. The use of PMOS devices exclusively is not a viable solution for high-frequency circuits since they have a poor frequency response relative to the NMOS devices for 90nm and 65nm.

In addition, the amount of gate leakage is also a strong function of the applied bias [28]. If the transistor has a voltage across the oxide, V_{OX} , roughly equal to V_{DD} , the leakage current density

is largest. If the transistor has a V_{OX} set to close or below the threshold voltage V_T , it leaks significantly less. Indeed, under such a condition, the gate leakage current is typically 3-6 orders of magnitude less, depending on the values of V_{DD} and t_{OX} [28]. Thus, the gate leakage in the second condition can be roughly considered to be zero. In decaps, the gate is at V_{DD} and the source and drain of a transistor are tied together. Therefore, decaps would experience the highest levels of leakage, as a function of V_{OX} .

The oxide capacitance C_{OX} is a critical factor to many physical properties of MOS transistors since the drain current of a transistor is proportional to C_{OX} . A larger C_{OX} results in a larger drain current and hence a faster transition or a shorter gate delay. On the other hand, the subthreshold current is related to C_{OX} : a smaller C_{OX} (or a larger t_{OX}) corresponds to a higher threshold voltage V_T and therefore smaller subthreshold current [26]. Each technology generation attempts to increase C_{OX} by roughly 1.4x while reducing the channel length L to 0.7x of the previous technology's channel length [7]. The result is that the product of $C_{OX}L$ has been maintained relatively constant as technology scales. The proper C_{OX} selection balances the trade-off between the drain current and the subthreshold current in each technology node.

From Equation (2.3), the gate leakage density is inversely related to t_{OX} . A smaller t_{OX} leads to exponentially increasing gate leakage. From a gate leakage perspective, the oxide thickness t_{OX} should be kept large. However, C_{OX} is determined by:

$$C_{OX} = \frac{\varepsilon_{OX}}{t_{OX}}$$
(2.4)

where ε_{OX} is the permittivity of the oxide and is fixed for a given oxide material. Equation (2.4) suggests that if ε_{OX} is kept unchanged, the increase in C_{OX} will lead to a decrease in t_{OX} and hence an exponential growth in gate leakage [46].

Knowing that the gate leakage increase may be excessive for 90nm and 65nm, in order to keep t_{0X} thick while increasing C_{0X} , one can adjust the dielectric constant, k, where $\varepsilon_{0X} = k \cdot \varepsilon_0$, and ε_0 is the vacuum permittivity. If a high permittivity (high-k) dielectric can be used instead of the normal SiO₂ oxide, the physical oxide thickness t_{0X} would no longer be limited by its electrical property C_{0X} . This concept of using high-k dielectrics was presented in [47], and researchers and process engineers have continued to pursue better high-k materials [48]. One of the main goals of using high-k gate dielectrics is to keep the gate leakage under control [48]. Commonly suggested high-k materials include HfO₂, HfSiON, ZrO₂, and Al₂O₃, whose permittivity ranges from 10 to 30 [48][49], compared to 3.97 of SiO₂. Commercial microprocessors fabricated at a 45nm technology have been developed using high-k materials, and it has been reported that a 10X gate leakage reduction was achieved [50]. Starting from 45nm, most fabrication facilities are anticipated to shift to high-k technology to reduce gate leakage [51]. However, for 90nm and 65nm, the concerns on gate leakage are still significant [27][51].

2.4 Electrostatic Discharge Reliability in Decap Design

ESD protection due to the thin oxide has become an important concern starting from the 90nm technology node. ESD is the process of static discharge that can typically arise from human contact with any IC pin. Approximately 0.6μ C of charge is carried on a body capacitance of 100pF, generating a potential of 2kV or higher to discharge from the contacted IC pin to ground

for a duration of more than 100ns [24]. Under such an event, the peak discharge current is in the ampere range, leading to permanent damage on certain transistors in the chip if not properly protected. The damage can be in one of two forms, or a combination of the two. The first is thermal burnout in devices or interconnects, while the other is oxide breakdown of devices due to the high voltage across the oxide [24]. When running simulations for an ESD event, the maximum current density J_{max} of devices and interconnects is measured to check for potential thermal damage. The oxide voltage also needs to be measured to compare with the oxide breakdown voltage of a device for a given fabrication process. The oxide breakdown voltage is almost linearly proportional to the oxide thickness [24]. For instance, assuming a 90nm process uses 1.7nm of oxide thickness, the corresponding oxide breakdown voltage is just below 5V. If the thickness is doubled, the oxide breakdown voltage is also doubled to around 10V [24].

An ESD event can be delivered between any two pins of an IC. To properly protect an IC from ESD damage, an ESD circuit must shunt ESD current between these two pins [24]. In the case of decaps within standard cells, the only two pins that the decaps have access to are the two local power rails, namely V_{DD} and V_{SS} . Primary and local (sometimes called *secondary*) protection elements are needed to protect the two rails by limiting the voltage difference between the two rails to a value below the oxide breakdown voltage. The primary element will shunt most of the ESD current, whereas the local element serves to limit the voltage or current at the local circuit until the primary element is fully operational [24]. A primary element can be a thick oxide transistor, a silicon-controlled rectifier, an open-gate, grounded-gate or coupled-gate NMOS transistor [24].

A typical ESD protection scheme is illustrated in Fig. 2.6. In addition to the primary and local elements, a resistor R_{in} is required to limit the maximum current flow to the decap and to limit the voltage seen from the gate of the decap. For better ESD protection, this resistance is normally large and can be in the forms of polysilicon, diffusion, n-well, or channel resistance [24]. The resistance is generally not implemented together with primary and local protection devices. Rather, it is usually inserted within standard cells where ESD damage is a concern.



Figure 2.6: Complete ESD protection scheme.

Previous decap designs (typically before 90nm technology) did not consider ESD performance for two reasons. First, the transistor's oxide thickness was large and the oxide breakdown voltage was high enough that the transistor was likely to survive during an ESD event with adequate protection circuits. Second, insertion of the large resistance R_{in} dramatically reduces the transient response of the decap. However, starting from 90nm, the gate oxide is so thin that the designer cannot ignore the increased ESD risk. A large resistance is therefore recommended to be placed inside the decap cells to protect the circuit from potential ESD damage. Hence, this tradeoff between ESD reliability and transient response becomes the one of major decap design challenges in 90nm and 65nm technologies. The ESD simulation requires an ESD generation model. Among all the existing models, the *human body model* (HBM) was adopted for simplicity. Following the standard MIL-STD-883x method 3015.7 [24], a human body can be simulated as a series of $1.5k\Omega$ resistance R_{HBM} and 100pF capacitance C_{HBM} . The capacitor C_{HBM} is initially charged to 2kV that needs to be discharged through some primary elements. The primary element is arbitrarily chosen to be an ESD diode plus a gate-coupled NMOS device (GCNMOS) with an n-well resistor R_{nwell} (~15k Ω) and an NMOS bootstrap capacitor C_b . Two identical primary elements are used to protect the circuit placed in between the HBM generation and the elements, as shown in Fig. 2.7. For simplicity, no secondary element is used.



Figure 2.7: Simulation setup for ESD analysis [24].

Since the primary elements are designed to handle large current flow, the maximum current density, J_{max} , is assumed to be within the safe range and is not measured. HBM generation raises the voltage level at node V_{DD} , and hence turns on the primary elements to discharge. For device protection from oxide breakdown, the voltage across the oxide V_{OX} of each decap transistor needs to be observed in simulation. The V_{OX} voltages should to be kept as low as possible, given that the oxide breakdown voltage for a typical 90nm is below 5V.

2.5 Standard-Cell Decap Layout and Placement

In the white spaces around the chip, decaps are usually made of NMOS devices, as described earlier. However, within standard cells, it is more convenient to make decaps using both NMOS and PMOS transistors to form a decap filler cell. This is because the n-well is already implemented and usually reserved for PMOS devices. Only about a half-cell area is for NMOS devices. One sample standard-cell decap layout is illustrated in Fig. 2.8. In the figure, the NMOS decap occupies roughly the bottom half of the cell area, whereas the PMOS decap is located in the n-well. The capacitor areas are the polysilicon gates placed on top of the channel regions of the MOS transistors. For standard cells, the height of the cell is always fixed, and the designers can only adjust the cell width. Once the cell width is determined, the size of the decap and the capacitance of the decap are established. Fig. 2.8(a) illustrates a large decap cell (measured in cell width) with long channel transistors. A fingering technique is commonly used to have a smaller effective channel length to improve the decap frequency response. Fig. 2.8(b) depicts the same decap cell but with two fingers.





During the placement procedure, computer-aided design (CAD) tools place standard cells into rows. Because the height of each cell is always the same, when cells are placed adjacent to each other, the n-well region and the V_{DD} and V_{SS} lines are automatically aligned. The cells for placement are obtained from the standard-cell library, where all the cells are predefined in width and driving strength. Since the total width of the row is fixed and the individual cell widths are fixed, some empty spaces (typically small) between the cells are left after placing cells. Those empty spaces are good candidates for the placement of decap cells [9]. In fact, a set of decap cells with different cell widths is often included in the standard-cell library.

Decap insertion is considered as a part of the complete design flow. In a typical ASIC design flow, once the standard-cell blocks are synthesized, placed and routed by CAD tools, the decap cells are placed into the empty spaces. Generally, since the spaces are filled using a library of decap cells with various sizes, the decap placement is done without affecting the placement of other logic cells. After placement and routing, chip-level timing is analyzed and timing violations will be fixed by replacement and/or rerouting. Then, chip-level *IR*-drop analysis is carried out by a CAD tool (e.g., Apache Redhawk) such that the hot spots of severe voltage-drop areas are identified [52]. If the voltage drop at the hot spots exceeds the noise budget, more decaps will be inserted into the violation regions and a modification of the placement of other logic cells may have to be done. The logic cell movement requires additional timing and routability analysis before moving on to next step. Then, the chip's *IR* drop is analyzed again for the remaining hot spots. These steps in the design flow are iterated until all the hot spots are eliminated and all the logic circuits pass timing analysis. Typically, it may take one or two (occasionally even more) iterations to eliminate all the hot spots [7][9]. In addition, the potential problem of electromigration is also checked alongside the *IR*-drop analysis [7].

This commonly used decap placement approach is not optimal simply because the empty cells may not be located near the high *IR*-drop regions. After the hot spots are first identified, the remaining empty spaces near the hot spots may not be large enough. Hence, the logic cells may have to be shifted, resulting in a need for additional timing analysis. In order to improve the placement efficiency, researchers suggest a few approaches including: global decap placement between standard-cell blocks [53], decap placement using activity [1], standard-cell decap placement not affecting relative placement of logic cells [9], and earlier-stage decap placement methods considering leakage current are proposed in [6] and [25].

2.6 Metric for Power Supply Noise Management

Static Voltage Drop (SVD) analysis and verification has traditionally been an essential part of the overall physical design and verification flow in semiconductor industry for over the past ten years [20]. In this approach, the *IR* drops across the chip are computed by averaging the current draw by the transistors and blocks from the power grid. The computed fixed values can be fed to timing verifiers to assess the impact on delay. In the past, SVD analysis provided useful feedback in terms of major errors in the power grid design. Going forward, at 90 nm and smaller technologies, SVD verification is not enough to ensure power integrity. SVD does not take into account the contribution of power density, variations in the switching activity profile and impact

of inductance and decoupling capacitors (including *LC* resonance effects) [2]. Therefore, SVD is not an adequate approach to analyze and optimize power delivery networks in SoC designs.

Recently, industry began to use Dynamic Voltage Drop (DVD) analysis as a way to capture the impact of decaps, inductance, and spatial and temporal switching events in the design [20]. DVD is emerging as a replacement to SVD to capture the impact of power supply noise on timing behavior of logic and memory cells. In order to evaluate the design of active or passive decoupling capacitors on a DVD profile and to qualify its impact on a design's timing performance, two quantities can be used as a metric: DVD_{avg} and DVD_{max}, as shown in Fig. 2.9. DVD_{avg} is the DVD profile's average value in the timing cycle, whereas DVD_{max} is the DVD profile's peak value in the same timing cycle. A design is considered better if it has smaller DVD_{avg} and DVD_{max}. Users should add design margins to these metrics to account for the metric's simplifications, or they should perform the final signoff process with actual DVD profiles.



Figure 2.9: DVD_{avg} and DVD_{max}: metric used to evaluate DVD profiles.

Ref. [20] validates the use of this metric. It can be established that the impact of the voltage drop profile on the timing performance of a digital path is equivalent to applying a fixed supply voltage of V_{DD} - DVD_{avg} to the same path. To show this, a logic path was first simulated in presence of a true dynamic voltage profile on the power supply, and then a DC voltage equal to the average of the voltage profile was used on the power supply [20]. The results show that the timing behaviour of the two cases matches. The intuitive reason for this relationship can be illustrated in Fig. 2.9. Considering the delay of the critical path of the circuit between the two flops, gate delay is reduced when the supply voltage overshoots ($V_{DD}(t) > V_{DD_nominal}$) and increased when the supply voltage undershoots ($V_{DD}(t) < V_{DD_nominal}$). When the supply voltage fluctuates, gates that see a voltage drop higher than the V_{DD} - DVD_{avg} accelerate and gates that see a voltage drop lower than V_{DD} - DVD_{avg} . Therefore, DVD_{avg} is a good measure of the average effect of the *IR* drop on delay. The use of DVD_{avg} as a metric for *IR*-drop analysis has been shown to be valid on several industrial designs [20].

Applying the metric, any approach (e.g., decap insertion) that reduces the *average* dynamic voltage drop (DVD_{avg}) or raises the *average* supply voltage $(V_{DD}-DVD_{avg})$ can be considered as a valid solution for reducing power supply noise to improve timing performance, although the instantaneous supply voltage drop may not be affected.

The DVD_{max} value represents the worst-case voltage drop, with a safety margin, that causes a failure in logic circuit and memory cells. That is, if the voltage drop exceeds DVD_{max} by the
margin of safety, the behaviour of standard cells or memory cells will be unpredictable. The value of DVD_{max} depends on the tolerance of individual IP blocks to power supply noise.

Ref. [20] suggests that DVD_{avg} should not be bigger than 10% of $V_{DD}-V_{SS}$ and DVD_{max} should not be bigger than 20% of $V_{DD}-V_{SS}$. These percentage values are commonly used in the industry [20]. These limits are considered pessimistic enough to account for the simplified nature of the metric.

By analyzing different DVD profiles using the metric of DVD_{avg} and DVD_{max} [20], an important conclusion can be made: Ldi/dt contribution of voltage drop is not as critical as the *IR* contribution. Ldi/dt only affects the DVD_{max} and therefore has minimal impact on DVD_{avg} as long as the transfer of charge is completed within the cycle. Therefore, Ldi/dt voltage drop may not significantly affect the timing performance of the chip, as long as it does not create supply fluctuations that exceed the DVD_{max} . The above observation is consistent with [2].

2.7 Summary

This chapter summarized a number of decap design issues including gate tunneling leakage, ESD protection, standard-cell layout and placement requirements, and the lack of useful metrics evaluating the results from DVD analysis. The decap design challenges for 90nm and 65nm were described. A simple metric, DVD_{avg} and DVD_{max} , was proposed to interpret the DVD results from CAD tools. The metric is best used to compare and evaluate different decap designs.

Chapter 3

Passive Decoupling Capacitor Design

3.1 Introduction

In an ASIC design flow, after placement and routing, empty spaces naturally exist within standard cells. Passive decoupling capacitors, as filler cells, are usually used to fill these empty spaces to reduce IR drop problems locally. This chapter addresses the design and layout of passive decaps [9][23][24] for standard cells at the 90nm technology node. As described in the previous sections, the IR drop has been increasing over time due to increased resistances in the power grid as the metal widths continue to shrink with each successive technology generation. The inductive Ldi/dt effects are also increasing due to the high current demands of ASIC designs in deep submicron technologies [7][10][11]. The increased supply noise level challenges the design and layout of passive decaps.

A number of relatively new issues for standard-cell decaps must be addressed that impact the design and layout of these cells at scaled technology nodes. Two important problems of decap frequency response and electrostatic discharge (ESD) protection [24] will be addressed. Since decaps are required to perform at increasingly higher operating frequencies, the frequency response [10][12][54] of passive decaps will be investigated first to propose improvements to

optimize decap layouts. Next, the problems of reduced oxide thickness of a transistor, namely, ESD [24] and thin-oxide gate leakage [6][11], will be explored in the context of decap design. A potential ESD event across a thin gate oxide increases the likelihood that a chip will be permanently damaged due to a short circuit in the decap itself. Higher gate leakage increases the total static power consumption of the chip.

A cross-coupled standard-cell design was proposed [21] to address the issue of ESD performance. The design provides sufficient ESD protection, but does not offer any savings in gate leakage and it may compromise the frequency response. This chapter aims to suggest improved layouts of the cross-coupled design that properly tradeoff frequency response and ESD performance, while greatly reducing gate leakage current.

3.2 High-Frequency Response of Decoupling Capacitors

Standard-cell layouts of an IP block consist of rows of fixed-height cells in the ASIC design flow. After cell placement is completed, there are a number of empty cells that can be filled with decaps of various sizes depending on the space available. Previous work has addressed the automatic placement and sizing of decap cells [9]. The focus in this chapter is on optimal layout of each decap filler cell. Typically, these standard cells have both NMOS and PMOS devices as shown in Fig. 3.1(a), with a corresponding layout in Fig. 3.1(b). Thin-oxide MOS devices are generally used for standard-cell decap implementation.



Figure 3.1: (a) A standard cell decap can be implemented as a NMOS in parallel with a PMOS device. The corresponding layout is shown in (b).

As the frequency of operation increases, a fingering approach is required to implement the layout. That is, a single transistor is split into a number of parallel transistors with the same width, but smaller channel lengths. The overhead of this approach is additional spacing for source/drain contacts and an overall reduction in the low-frequency capacitance. However, the average capacitance of the decap over a given frequency range improves as the number of fingers increases. Therefore, the problem of how many fingers to use given a fixed area of a filler cell and fixed gate-oxide thickness needs to be addressed. The objective here is to develop a useful metric to capture the frequency response characteristics in order to choose the optimal number of fingers.

To derive the needed equations, an NMOS decoupling capacitor is depicted first in Fig. 3.2. Non-idealities associated with MOS devices are modeled as a lumped-RC circuit [12] where both the effective resistance, R_{eff} , and effective capacitance, C_{eff} , are functions of frequency, f, as shown in Fig. 3.2.



Figure 3.2: A decap can be implemented as a NMOS device and modeled as a lumped-RC circuit with effective resistance and effective capacitance as functions of frequency, *f*.

The DC capacitance $C_{eff,0}$ and resistance $R_{eff,0}$ are given by [12][24][55] (where the subscript 0 indicates zero frequency at DC):

$$C_{eff,0} = C_{OX} W L + 2C_{OL} W$$
(3.1)

$$R_{eff,0} = \frac{1}{12} \frac{L}{\mu C_{OX} W(V_{GS} - V_T)}$$
(3.2)

where C_{OX} is the oxide capacitance per unit area, C_{OL} is the overlap and fringing capacitances per unit width of the device, μ is the channel mobility, V_{GS} is the voltage across the oxide, and V_T is the threshold voltage.

Assuming that a given filler cell has a horizontal dimension X and vertical dimension Y, the channel length of each device in a fingered layout is:

$$L_m = [X - (m-1)x_{contact}]/m$$
(3.3)

where *m* is the number of fingers and $x_{contact}$ is the distance between fingers required by contact spacing rules. Modified expressions for $C_{eff,0}$ and $R_{eff,0}$ can be derived as a function of the number of fingers. Thus, the effective capacitance at DC is given by:

$$C_{eff,0(m)} = m \cdot (C_{OX} Y L_m + 2 \cdot C_{OL} Y)$$
(3.4)

For capacitance, each additional finger adds extra overlap and fringing capacitances but loses area due to the contact spacing. Therefore, the capacitance actually decreases linearly as the number of fingers increases. The corresponding equation for $R_{eff,0}$ with *m* fingers in a parallel combination is:

$$R_{eff,0(m)} = \frac{1}{12} \cdot \frac{L_m}{\mu C_{ox} W(V_{GS} - V_T)} \cdot \frac{1}{m} \approx \frac{R_{eff,0}}{m^2}$$
(3.5)

In previous work [12], the resistance was used to select the channel length and there were no area constraints involved. However, since the resistance drops off as m^2 , it is not as important in the selection of a suitable m. In fact, the goal of an optimal layout should be to provide the highest capacitance value in the given area over a desired operating frequency, 0 to f_0 , while delivering a low resistance. A simple metric is needed to evaluate layouts with differing number of fingers. The easiest choice for a metric is to use the average capacitance over this frequency response up to f_0 , as follows:

$$C_{avg(m)} = \frac{C_{eff,0(m)} + C_{eff(m)}(f_o)}{2}$$
(3.6)

where $C_{eff,0 (m)}$ is obtained from Equation (3.4) and $C_{eff(m)}(f_0)$ is the effective capacitance with *m* fingers at frequency f_0 . A weighted average is also feasible, but it was observed that the simple average works well in practice.

The main issue with the metric is that $C_{eff(m)}(f_o)$ is difficult to compute without the aid of HSPICE or an equivalent simulation tool. To facilitate the process, simple frequency-dependent models for both C_{eff} and R_{eff} are developed. Also, the characteristics of both functions need to be accurate as technology scales. First, a number of AC simulations were performed in HSPICE for a 90nm CMOS technology using *non-quasi-static* (NQS) models, which are essential when simulating decaps in the gigahertz frequency range of operation. Two parameters, ACNQSMOD and TRNQSMOD, were set to "1" in BSIM4 [55]. The circuit in Fig. 3.3 was used to extract the effective resistance and capacitance from HSPICE results as follows [52]:

$$R_{eff}(f) = \frac{\operatorname{Re}(I_{RC})}{\operatorname{Mag}^{2}(I_{RC})}$$
(3.7)

$$C_{eff}(f) = \frac{\operatorname{Mag}^{2}(I_{RC})}{2\pi f \operatorname{Im}(I_{RC})}$$
(3.8)

where $\operatorname{Re}(I_{RC})$, $\operatorname{Im}(I_{RC})$, and $\operatorname{Mag}(I_{RC})$ are the real, imaginary, and magnitude components of I_{RC} , respectively. It is assumed in Equation (3.7) and (3.8) that the applied AC voltage V_{ac} is $1 \angle 0^{\circ} V$.



Figure 3.3: The circuit setup to extract the effective resistance and the effective capacitance values from an ac analysis.

Both NMOS and PMOS decaps were simulated with W x L sizes as follows: 15µm x 5µm, 10µm x 10µm, and 5µm x 15µm. The simulation frequency ranged from 0 to 10GHz. Typical ASIC clock rates today are in the range of 500MHz to 1GHz, but it is important to study frequency response well beyond the clock frequency. Most of the spectral power density of digital signals lies within frequencies of up to $f_{\text{knee}} = 1/(2t_{\text{rise}})$, where t_{rise} is a signal's rise time (which can be on the order of 50ps or less), and f_{knee} is the 3-dB cutoff frequency of the spectral power density [56]. It was assumed conservatively that $t_{\text{rise}} = 50$ ps, and the analysis was carried out up to 10GHz.





Large Length R_{eff} NMOS



Large Length C_{eff} PMOS



Figure 3.4: Plots of C_{eff} and R_{eff} for three device sizes (W x L): 10x10µm, 15x5µm, and 5x15µm.

The results of the simulations are shown in Fig. 3.4. As f increases, there is a noticeable roll-off in the curves due to finite transit time effects. Devices with large L's have a more pronounced effect. In fact, the C_{eff} curve for 5µm x 15µm quickly decays in value relative to 15µm x 5µm. A general observation on NMOS and PMOS decaps can also be made: NMOS is superior to PMOS in its high-frequency behavior since it has a larger C_{eff} and a smaller R_{eff} at high frequencies, assuming the area is fixed. Although standard cells employ both NMOS and PMOS devices for decaps, these results show that NMOS decaps would provide better frequency response characteristics.

For modeling purposes, based on the frequency responses of C_{eff} and R_{eff} , it is suggested [55] that the functions can be postulated into the form:

$$C_{\text{eff}}(f) = \frac{C_{\text{eff},0}}{1 + (f/f_T)^2 \tau_1^2}$$
(3.9)

$$R_{eff}(f) = \frac{R_{eff,0}}{1 + (f/f_T)^2 \tau_1^3}$$
(3.10)

where $\tau_1 = 1/12$ [55] and the transition frequency $f_T = \frac{\mu(V_{GS} - V_T)}{2\pi L^2}$. Shown in Fig. 3.5 are the results of curve-fitting using Equations (3.9) and (3.10) against HSPICE for the NMOS device. The results are very close. A factor of 1/2 was applied to f_T in order to produce results shown in Fig. 3.5. That is, the equation for f_T must be adjusted by a fitting factor of 0.5 in order to obtain good results. Similar results were obtained for PMOS devices. This demonstrates that the first-order equations for $C_{eff}(f)$ and $R_{eff}(f)$ are reasonably accurate and, perhaps more importantly, that $C_{eff(m)}(f_o)$ can be easily computed for the metric without running HSPICE.



Large Length Ceff NMOS

Large Length $R_{\rm eff}$ NMOS



Figure 3.5: Plots of C_{eff} and R_{eff} for three NMOS devices (HSPICE versus model).

At this point, all the necessary information is obtained to determine the number of fingers based on the frequency response. From Equation (3.9), the effective capacitance in Equation (3.6) at f_0 with *m* fingers is:

$$C_{eff(m)}(f_o) = \frac{C_{eff,0(m)}}{1 + (f_o / f_{T(m)})^2 \tau_1^2} \text{ where } f_{T(m)} = \frac{\mu(V_{GS} - V_T)}{2\pi L_m^2}$$
(3.11)

To demonstrate the efficacy of the metric, it was applied to the layout of a standard-cell decap in an available area of $2\mu m \ge 9\mu m$. Using Equation (3.6), Table 3.1 lists the $C_{avg(m)}$ metric values for the NMOS or PMOS devices for different frequency ranges. The optimal number of fingers corresponds to the largest entries in bold. For example, if the frequency range of interest is 0 to 10GHz, then 3 NMOS fingers and 4 PMOS fingers are optimal relative to the metric. Of course, if the range is 0 to 2GHz, two fingers are sufficient for both N or P devices. Note that PMOS devices typically require one more finger than NMOS devices at higher frequencies of operation.

Frequency Range		Cavg Metric versus Number of Fingers					
$0 - f_0$		<i>m</i> =1	<i>m</i> =2	<i>m</i> =3	<i>m</i> =4	<i>m</i> =5	
0 - 2 GHz	N	180 fF	187 fF	182 fF	177 fF	171 fF	
	Р	150 fF	183 fF	182 fF	177 fF	171 fF	
0 - 5 GHz	N	150 fF	184 fF	182 fF	177 fF	171 fF	
	Р	110 fF	165 fF	178 fF	176 fF	171 fF	
0 - 10 GHz	N	120 fF	173 fF	180 fF	176 fF	171 fF	
	Р	100 fF	135 fF	166 fF	172 fF	169 fF	

Table 3.1: Optimal number of fingers for different frequency ranges.

Table 3.1 was shown to illustrate the use of the metric in determining the optimal number of fingers. In practice, the design process would be as follows. First, the area of a filler cell (in particular, the X dimension of the cell) and frequency range of operation are used as input parameters. Then, the capacitance value as a function of m is computed using Equation (3.6). Finally, the value of m producing the highest capacitance is used to implement the layout.

The results in Table 3.1 can be validated by using Equations (3.4) and (3.9) to generate $C_{eff(m)}$ plots for both NMOS and PMOS devices, as shown in Fig. 3.6. The results in the plot were verified with HSPICE to ensure consistency. As an example, consider the cases with 1 finger and $f_0=10$ GHz. For the NMOS case in Fig. 3.6, $C_{avg(1)N}=(C_{eff,0}+C_{eff(1)}(f_0))/2=(190fF+50fF)/2=120fF$, whereas for PMOS, $C_{avg(1)P}=(C_{eff,0}+C_{eff(1)}(f_0))/2=(190fF+10fF)/2=100fF$. These are the same values that are found in the last row of the table with m=1. The rest of the table is produced in the same manner for different values of m and frequency range, $0 - f_0$.

By inspection, the plots indicate that 3 fingers would be optimal for NMOS decaps and 4 fingers would be optimal for PMOS decaps, based on the flatness of the lines and the initial value of the capacitance. This conclusion is consistent with Table 3.1. However, by using the metric, designers can quickly obtain the optimal number of fingers for a target operating frequency, without the need for such plots or SPICE simulations.

Calculated Ceff NMOS



Figure 3.6: The effective capacitance, $C_{eff}(f)$, of NMOS and PMOS decaps in 90nm for different numbers of fingers in a fixed area of Y=2µm and X=9µm.



Figure 3.7: The effective capacitance, $C_{eff}(f)$, of NMOS and PMOS decaps in 90nm for different numbers of fingers in a fixed area of Y=2µm and X=9µm.

Fig. 3.7 illustrates how standard cell layouts would be implemented using the above results, assuming a 10GHz operating range. These layouts would be automatically created by a decap filler cell generator. Two possible layouts are shown: (a) uses the N and P devices and (b) is NMOS only. Fig. 3.7(a) uses 3 fingers for the NMOS device and 4 fingers for the PMOS device. From an average capacitance perspective, the NMOS-only layout style of Fig. 3.7(b) is better, and this is also reflected in Table 3.1. To implement this type of layout in a standard cell, the p-well region must be extended to cover the entire area, which is not typical of standard cell design. This approach can be used as long as the design rules at the boundaries of adjacent standard cells are satisfied.

3.3 Cross-Coupled Decoupling Capacitor Designs

At the 90nm technology node, there is the possibility of oxide breakdown during an ESD event. A simple ESD protection scheme for decaps is to insert a relatively large resistance in series to limit the maximum voltage seen at the gate of the decap [24]. A minimum $R_{eff,0}$ is needed to ensure ESD reliability for decap cells. A cross-coupled decap design has been proposed by cell library developers [21] to address the issue of ESD reliability. As shown previously in Fig. 2.2, the drain of the PMOS device is connected to the gate of the NMOS, and vice-versa [21]. The cross-coupled design provides additional series resistance to the inherent decap resistance to increase $R_{eff,0}$.

The frequency response characteristics of this new configuration can be evaluated to determine if the results obtained in the last section can be applied directly to the new circuit. A standard 3N-4P decap of Fig. 3.7(a) is first compared to a same-area cross-coupled decap using HSPICE ac analysis in Fig. 3.3, and the results are shown in Fig. 3.8.

The standard 3N-4P decap has a very low resistance (around 30 Ω), which makes it prone to ESD failure. The cross-coupled 3N-4P design has a much higher DC $R_{eff, 0}$ (around 3500 Ω) but a poorer frequency response for C_{eff} . Consequently, the tradeoff between ESD reliability and frequency response must be considered in the design process and decap layout. To improve the frequency response, additional fingers must be used. The target resistance $R_{eff,0_target}$ for ESD protection in our case is a minimum of 500 Ω . The number of fingers was increased to reduce the resistance from 3500 Ω down to that required by ESD. According to Equation (3.5), the scale factor on the 3N-4P design can be found as follows:

$$R_{eff,0_target} = \frac{3500}{m^2} = 500$$

$$\therefore m = \sqrt{3500/500} = 2.6$$
(3.12)

Scaling the 3N-4P by approximately this amount, a cross-coupled 8N-9P decap was produced. Similarly, for an ESD target $R_{eff,0_target} = 1000\Omega$, it was found that m=1.9, so a cross-coupled 6N-7P decap was chosen. The plots for 6N-7P and 8N-9P fingers are also illustrated in Fig. 3.8. The results show that the 8N-9P cross-coupled version is the best configuration to address both frequency response and ESD protection.



Figure 3.8: $C_{eff}(f)$ and $R_{eff}(f)$ comparison of fixed-area standard decap and cross-coupled decap: same MOS device sizes but different poly connections.

From a layout perspective, the cross-coupled decaps can be realized by simply rerouting the poly connections of the standard decaps, while keeping the MOS devices the same. The layouts of two cases, 3N-4P and 8N-9P, are shown in Fig. 3.9.



Figure 3.9: Sample layouts of cross-coupled decap cells for (a) 3N-4P (b) 8N-9P.

It is important to address one other issue of thin-oxide, gate leakage current of the decap, which contributes to the chip's total static power. Using HSPICE, the standard and cross-coupled decap circuits were found to have almost identical gate leakage. That is, since the cell area is fixed and only the poly terminal connections are swapped, the cross-coupled design provides no inherent savings in gate leakage as compared to the standard design. There exists a simple design approach to save gate leakage. Simulations using BSIM4 SPICE models [44] indicated that

PMOS gate leakage is roughly 3 times smaller than NMOS gate leakage for same size transistors [27][43]. Therefore, PMOS devices are preferred from a leakage perspective. Since PMOS devices have a poor frequency response, more fingers can be used to obtain the desired result. But this must be carried out in the context of the cross-coupled design to preserve ESD protection.

The basic idea to control leakage is to have the smallest possible NMOS device cross-coupled with the largest possible multi-fingered PMOS device. This way, the advantages of PMOS leakage and cross-coupling ESD protection are preserved. The layouts of two configurations are illustrated in Fig. 3.10. A small NMOS device is used in both cases. Note the n-well regions have been expanded in both layouts to accommodate the larger PMOS device. Fig. 3.10(a) uses 9 PMOS fingers while Fig. 3.10(b) has a total of 16 fingers. The same cell area as before (2µm x 9µm) was used for the two designs.



Figure 3.10: Sample layouts of improved decap cells for (a) 1N-9P (b) 1N-16P.

Table 3.2: Comparison of the p	passive decap designs and	their gate leakage current.
--------------------------------	---------------------------	-----------------------------

Decap Cell Layout		Description	Gate Leakage
Std. 3N-4P		Std. decap with 3 fingers for N and 4 fingers for P	262.4 nA
Cross-Coupled 3N-4P		Cross coupled with 3 fingers for N and 4 fingers for P	260.8 nA
Cross-Coupled 8N-9P		Cross coupled with 8 fingers for N and 9 fingers for P	206.8 nA
Modified	9 fingers	Cross coupled with smallest N and 9 fingers for P	119.1 nA
	16 fingers	Cross coupled with smallest N and 16 fingers for P	99.7 nA

Table 3.2 summarizes the leakage values for the different cases. The standard and cross-coupled 3N-4P decaps have roughly the same leakage. It is somewhat reduced for the 8N-9P case since there is less area for leakage. However, for the two layouts with the small NMOS devices, the leakage is cut in half. In fact, the case with 1N-16P, the leakage is 62% less than the standard decap 3N-4P.

The $R_{eff,0_target}$ of the cross-coupled design must be set based on ESD considerations, but that also controls the maximum number of fingers permitted, m_{max} . Since the NMOS device is fixed while the PMOS device is multi-fingered, the following equation can be used to determine the resistance:

$$R_{eff,0_target} = R_N / / \frac{R_P}{m_{max}^2}$$
(3.13)

where R_N and R_P are the resistance of the decaps without fingers, and "//" means "in parallel with." This target $R_{eff,0}$ sets up the equation for a maximum number of fingers, m_{max} . That is,

$$m_{\max} = \sqrt{\left(\frac{1}{R_{eff,0_target}} - \frac{1}{R_N}\right)R_P}$$
(3.14)

As described in the previous section, the optimal m depends on the frequency response (i.e., $C_{avg(m)}$), but the number of fingers selected should not exceed to m_{max} to satisfy ESD requirements.



Figure 3.11: Frequency response of various cross-coupled designs.

Fig. 3.11 illustrates the frequency response of the various designs from 0-10GHz. All of the configurations provide similar $C_{eff,0}$ values but are dramatically different in the frequency response characteristics. The standard 3N-4P case is the best, followed by the modified cross-coupled 1N-16P. The $R_{eff,0}$ are different in all cases but only the standard 3N-4P case is unsuitable for ESD protection. However, it is desirable to select the configuration with the lowest $R_{eff,0}$ that satisfies the ESD criteria (500 Ω in this case) for a rapid time-domain response. Overall, the cross-coupled 1N-16P layout is recommended because it provides the required $R_{eff,0}$ for ESD reliability and saves at least 50-60% on gate leakage.

To summarize, for 90nm and 65nm, standard-cell passive decap design should follow the layout strategy shown in Fig. 3.10. By using the smallest NMOS device and the largest multi-fingered PMOS device in the cross-coupled form, the decap has the lowest leakage and is able to satisfy the ESD requirements.

3.4 Summary

This chapter investigated the tradeoffs between high-frequency performance of decaps and ESD protection and its impact on the layout of standard-cell passive decaps. A design metric was introduced to determine the optimal number of fingers to use in the standard-cell layout to obtain a desired capacitance level over a target operating frequency. Models were developed to capture the frequency responses of R_{eff} and C_{eff} for a given technology with only a few parameters. As a result, the models can be used to predict the same characteristics of future technologies.

For ESD protection, a cross-coupled design was proposed by cell library developers to provide a large series resistance, but it suffers from reduced frequency response and provides no savings in gate leakage. This chapter demonstrated that more fingers are needed with the cross-coupled standard-cell layouts to provide the target resistance value for ESD protection. The design of the target resistance can follow the formulae provided in this chapter. The layout with the smallest NMOS device and a multi-fingered PMOS device delivers acceptable frequency response and ESD reliability, while providing the lowest leakage.

Active Decoupling Capacitor Design

4.1 Introduction

Passive decaps described previously have a small layout and are useful within the block of standard cells. However, for large global decaps (i.e., outside the block), other approaches can be used. This chapter addresses a novel application of the active decoupling capacitor. The objective is to investigate the effectiveness of removing local *IR*-drop violations (usually called "hot spots") by replacing passive decaps with active decaps. Starting from 90nm, large power supply noise levels in localized regions may unexpectedly be present in high-speed ASICs [19]. These unresolved hot spots cause timing closure problems or result in functional failures in extreme cases. These hot spots are often detected late in the design cycle so they become problematic and difficult to remove.

To remove them, designers must consider many options such as moving the logic blocks, adding or rearranging the power pins, and/or modifying power grid design. Near the tapeout deadline, such time-consuming design iterations may not always be feasible. In these situations, an active decap design can play an important role in reducing supply noise without major changes to the design [31]. That is, the active decap can be a drop-in replacement of the passive decap in an attempt to remove any remaining hot spots, thereby saving time and effort. In this chapter, to explore the effectiveness of an active decap, quantitative data will be provided on the expected improvements, sizing considerations and placement of an active decap relative to the hot spot.

Two active decap circuits using switched capacitances were proposed in the past to regulate the supply voltages [37][38][39]. By increasing charge delivery capability, the two designs in [37] and [39] are quite effective at reducing supply noise, but they also have certain limitations. The design in [37] can switch quickly but dissipates large power, whereas [39] saves power but experiences long switching delays. They both mitigate the effects of LC resonance [40], which is typically in the 20-400MHz band. Recent work has been done to reduce LC resonance using an improved switched decap technique [41]. Further work has also been reported on a distributed active decap to greatly boost the effective decap value while reducing the area requirements [41]. In this chapter, a modified active decap design that has lower power than [37] and a better response time than [39] is proposed, targeting ASIC applications up to 1GHz. The issue of directly replacing an area occupied by a passive decap with an active decap is addressed to determine the degree of noise reduction that can be obtained and the associated tradeoffs of area and power.

4.2 Active Decoupling Capacitor Analysis and Design

4.2.1 Active Decap Concept and Design Considerations

The basic idea of an active decap is to switch a pair of passive decaps, C_{decap} , from parallel to series to provide a local boost in the supply voltage [37][38][39]. As illustrated in Fig. 4.1(a), the decaps are initially in a parallel configuration with a full charge developed across both capacitors.

In this standby state, the equivalent capacitance is $2C_{decap}$. When placed in a series stack, as in Fig. 4.1(b), the boosted voltage is ideally $2V_{DD}$ while the equivalent capacitance is reduced to $C_{decap}/2$. When switched back in parallel, the voltage returns to the original value of V_{DD} . In this case, the stacking level *n* is 2.



Figure 4.1: Active decap concept and its MOS implementation.

The active decap circuit is depicted in Fig. 4.1(c), with C_{decap} and the switches implemented using NMOS and PMOS transistors [37][38][39]. When the capacitors are in parallel, both Mn1 and Mp1 are on while Mn2 and Mp2 are off (i.e., subthreshold). When the capacitors are in series, both Mn1 and Mp1 are off while Mn2 and Mp2 are on. The switches exhibit finite "on" resistances, indicated as R_{on} and R'_{on} , and there is also thin-oxide gate leakage through the decaps,

 I_{leak} , especially in 90nm and 65nm CMOS technologies. Both of these effects reduce the performance of the active decap, as described below.

For the general case of stacking *n* parallel decaps into a series chain, the maximum improvement can be characterized in terms of a gain, G [37]. If *k* is the voltage regulation tolerance, where kV_{DD} is the permissible drop in voltage, then the charge delivered by *n* parallel capacitors is:

$$Q_{parallel} = kV_{DD} \cdot nC_{decap} \tag{4.1}$$

When the capacitors are stacked in series, the charge delivered for the same voltage drop is:

$$Q_{series} = [nV_{DD} - (1-k)V_{DD}] \cdot C_{decap} / n$$
(4.2)

The overall charge gain is:

$$G = \frac{Q_{series}}{Q_{parallel}} = \frac{[nV_{DD} - (1 - k)V_{DD}] \cdot C_{decap} / n}{kV_{DD} \cdot nC_{decap}}$$
(4.3)

Therefore, as given in [37], the gain is controlled by *n* and *k*:

$$G = \frac{n+k-1}{kn^2} \tag{4.4}$$

There exists a value of k such that the regular decap outperforms the active decap. For example, setting G=1 and n=2, it is found that k=1/3. For values of k > 1/3, the active decap is of no value. However, if k is below this value, the active decap is able to deliver more charge. For example, if k=0.1 and n=2, then G=2.75. This implies that 2.75 times more charge can be delivered by the active decap before its output voltage drops to the same level as the passive decap.

Previous research [37][38][39] provides no information on practical limitations when using Equation (4.4). For design purposes, this level of improvement is not possible due to the switch

resistances and leakage currents. In fact, the boosted voltage cannot reach nV_{DD} but instead reaches a lower voltage of bV_{DD} . Therefore, the gain equation should be rewritten as:

$$G = \frac{[bV_{DD} - (1 - k)V_{DD}] \cdot C_{decap} / n}{kV_{DD} \cdot nC_{decap}} = \frac{b + k - 1}{kn^2}$$
(4.5)

where

$$b = n \cdot f(R_{on}) \cdot g(C_{decap}) \tag{4.6}$$

The reduction factors, $f(R_{on})$ and $g(C_{decap})$, depend on the switch resistance, R_{on} , and the leakage current which, in turn, is proportional to C_{decap} . Using circuit simulation, normalized plots of $f(R_{on})$ and $g(C_{decap})$ are provided in Fig. 4.2. The switch resistance has a more pronounced effect on b as compared to the leakage current. For example, with $R_{on}=10\Omega$ and $C_{decap}=700$ pF, it can be obtained that f=0.9 and g=0.95 from Fig. 4.2. If the two effects are combined, then $b=2(0.9)(0.95)\approx1.7$ instead of 2. With k=0.1, the achievable gain is now reduced to G=2.0. The actual final voltage value, aV_{DD} , when the active decap supplies the same charge as the passive decap is determined by setting G=1 and solving for a in the following equation:

$$G = \frac{(bV_{DD} - aV_{DD}) \cdot C_{decap}}{kV_{DD} \cdot nC_{decap}} = \frac{b - a}{kn^2}$$
(4.7)

In this case, with b=1.7, k=0.1 and n=2, one can obtain a=1.3, which implies that the active decap will be boosted initially to 1.7V (instead of 2V) and then falls back to 1.3V due to the charge demand of a nearby logic circuit. In the passive case, the initial voltage of 1V would be reduced to 0.9V, so the active decap is still superior even with the nonidealities included.



Figure 4.2: The reductive factors f and g for the boosted voltage as a function of (a) "on" resistances of the switches, R_{on} , and (b) leakage due to the size of decap C_{decap} .

To design the sizes of the MOS switches, a number of issues must be considered. From the above analysis, a small resistance value is preferable to increase the voltage boosting capability of the active decap, and to improve transient response times. The "on" resistances also provide ESD protection because they are in series with the decaps. Any large voltage fluctuations are absorbed by the resistors to reduce the drop across the thin-oxide gates of the decaps, similar to the effect of cross-coupling decaps [22]. Therefore, this resistance must be large enough to safely protect the thin-oxide gates. Considering the factors of boosted voltage level, decap performance, and ESD reliability, the "on" resistances should be designed to be in the range of 10-20 Ω by proper selection of transistor widths. This will require rather large switches. Once their sizes are determined, the buffers generating the switching signals must supply enough current to drive the large capacitances resulting in a large sensing and switching circuitry that consumes a considerable amount of power and area. Therefore, these active decaps should be used sparingly in ASIC designs but are particularly suitable for localized hot-spot removal.



4.2.2 Overall Active Decap Architecture

Figure 4.3: Active decap architecture.

Fig. 4.3 illustrates the complete active decap design containing four blocks: a reference voltage generator, a pair of high-pass filters, two comparators, and the switched decaps. Compared to the previous work [37][38][39], the key difference in this new approach is the use of latch-based comparators. The user logic circuit block shown in the figure is considered to be the main cause of power supply noise violation. The switch control circuit for the active decap is realized using two comparators. The differential inputs of each comparator decide the standby voltage level at the outputs of the comparators. In the standby mode, the top comparator has an output at V_{DD} , whereas the bottom comparator is set to V_{SS} . When the power grid discharges, V_{DD} will drop and V_{SS} will rise. The voltage variations are passed through the high-pass filters to the comparator inputs causing the comparators to reverse their output values, and switching the decaps from parallel to series. Later, when the power grid charges up, the comparator inputs and outputs switch back to their original values. The use of latch-based comparators with hysteresis to switch the decaps is one of the main contributions to this work. An *enable* signal is provided for testing purposes to allow the active decap circuitry to be turned on or off. When off, the design behaves purely as a passive decap. This allows for a comparison between active and passive decaps.

The trigger voltage for the circuit is set by the comparators and the resistor R_1 . In Fig. 4.3, the reference voltages are generated by a simple voltage divider and are set to roughly $V_{DD}/2$. However, depending on the comparator design, the absolute input levels of $V_{DD}/2$ are somewhat flexible due to the differential nature of the inputs. The diode-connected transistors in the reference generator should have large length and small width to control the static current. Inserting a small resistor, R_1 , between the two transistors is intended to separate the reference voltages by approximately 30mV. Then, if the comparators are designed to switch when the voltage difference at the inputs is 10-15mV (plus an additional 5mV of hysteresis), the overall design will trigger at approximately 50mV. If R₁ is chosen to be smaller, the sensitivity of the active decaps is improved [41], at a cost of significantly increased dynamic power because the active decap is triggered more often. If R₁ is designed to be larger, the resulting supply noise *k* will increase, as shown in Fig. 4.4. When the two input signals of the comparators are separated by a level that exceeds the supply noise generated by the nearby logic, the comparators will not switch, making it a passive decap. In the plot, the active decap stops switching when the input voltages differ by approximately 130~150mV.



Figure 4.4: Difference in the input voltages of the comparators as a function of supply noise k.

The targeted supply noise k value is at 0.09-0.1 (i.e., kV_{DD} value is 90-100mV), resulting in a triggering input voltage between 20-70mV. Thus, the comparator can be designed with a 20mV

switching threshold then the input voltage difference can be adjusted by varying R_1 to the midpoint of about 50mV. From another perspective, since the maximum *IR* drop is allowed to be 100mV, the active decap should trigger at about a half of 100mV, which is 50mV, because there is a delay between the time that the active decap is switched and the time that it actually boosts the supply voltage. Of course, a 20mV of input triggering voltage would be better in that sense, but the active decap will be switching too often, resulting in an increased concern on large power consumption. To design the triggering voltage at 50mV seems to be a good tradeoff between power and performance. Therefore, the comparators are required to switch once the voltage ($V_{DD}-V_{SS}$) discharges by about 50mV, which can be considered as the input sensitivity of the active decap.

The delay between detection and activation of the switched decap, t_d , and the delay difference between the two outputs of the comparators, Δt_d , impact the bandwidth of the active decap and the boosted voltage, respectively. Specifically, the delay of the comparators t_d is inversely proportional to the bandwidth of the active decap, BW. That is,

$$BW \approx \frac{1}{t_d} \tag{4.8}$$

If the operating frequency is below the bandwidth, the active decap reduces the supply noise relative to a same-area passive decap. On the other hand, if the clock frequency is beyond the bandwidth, the active decap may result in equal or more supply noise than the passive decap. If it takes an entire clock period to switch the decaps, then they are just like the passive case because they will not switch during the whole clock cycle. Before they actually switch, the supply voltage goes back to the right level and they are forced not to switch any more. The same situation happens on the next cycle. Therefore, as the switching frequency of the logic increases,

the active decap is less and less effective. Then, at $1/t_d$, it looks like a passive decap. Beyond that point, due to larger static power consumption and varying "on" resistance of the switches, the active decap becomes worse than the passive decap. The above observation will be validated in the next section.

Ideally, the delay of the top and the bottom comparators should be the same. In practice, a difference in delay Δt_d between the top and the bottom comparators exists and can be defined as:

$$\Delta t_d = t_{d_top} - t_{d_bottom} \tag{4.9}$$

The delay difference will not result in a short connection between power and ground or an open circuit where no decoupling capacitance is present. The effect is, however, the boosted voltage bwill be degraded due to leakage current when the switches are not turned on/off at the same time. A function, $h(\Delta t_d)$, can be used to capture this effect, as shown in Fig. 4.5. Therefore, Equation (4.6) should be re-written as:

$$b = n \cdot f(R_{on}) \cdot g(C_{decan}) \cdot h(\Delta t_d)$$
(4.10)

It is desired to keep the delay difference of the two comparators small even under process/ voltage/temperature (PVT) variations to ensure sufficient improvement of the boosted voltage.


Figure 4.5: The reductive factor h for the boosted voltage as a function of delay difference Δt_d .

4.2.3 Design Specifications

When designing the active decap, a certain value for C_{decap} is needed to keep the supply noise small. However, in the case of drop-in replacement, this design freedom is not available. Therefore, as the first step, designers should check and see through simulation if the replacement of active decap can remove the hot spot nearby or not. The next step is to select a proper "on" resistance of the switches R_{on} . For a high boosted voltage *b*, R_{on} should be kept low enough. On the other hand, R_{on} must also be high enough for ESD protection. Designers should make this tradeoff for R_{on} selection to design the switches. The size of the switches and the passive decaps determines the load capacitance on the comparators. The comparator design should generally satisfy low power and high speed requirements of ASIC applications. For example, a static power of 5mW or below can be achieved while the bandwidth of the active decap should be able to handle a 1GHz clock. This sets the average comparator delay to approximately 1ns. Note that the delay requirement should be fulfilled under PVT variations. Therefore, the worst-case delay should be used here. The output of the comparators in the standby state should be close to V_{DD} or V_{SS} to reduce leakage current through the passive decaps and switches to save power. The comparator should be designed to provide high gain in the switching region such that the output will swing from low to high when the input varies 10-15mV. Higher gain can ease the requirement on input DC biasing and lower the static power consumption. From a stability perspective, a certain amount of hysteresis is desired to reduce the risk of oscillation. A practical value of 5mV for hysteresis is reasonable. Overall, the design specifications can be summarized in Table 4.1.

	Specifications
Worst-case switching delay	< 1 ns
Bandwidth	> 1 GHz
Static power	< 5 mW

 Table 4.1: Design specifications of the active decap.

Specific design values are as follows. In the reference voltage generation circuit, the size of the transistors is chosen to provide a branch current of 5-6 μ A. R₁ can then be implemented with a value of 5k Ω to produce a separation voltage of the comparator inputs for ~30mV. For the *RC*-based high-pass filters, a cut-off frequency above 10MHz is used to filter out low-frequency

supply noise to save power. However, if the cut-off frequency is set too high, it may cause oscillation at the supply rails. A cut-off frequency of 16MHz was finally selected. The two resistors have the value of $R_2=R_3=10k\Omega$, while the capacitors (C_2 and C_3) are implemented with the same value of 1pF. These *RC* values are somewhat flexible, unless the cut-off frequency is set exceedingly high. For instance, it was observed that for a cut-off frequency of 1.6GHz or above, oscillation on the supply rail will occur. Therefore, it is a good approach to have the cut-off frequency.

4.2.4 Latch-Based Comparator Design

There are a wide variety of ways to design a comparator [57][58][59][60]. Also, the Appendix section of this dissertation provides the fundamentals of comparator designs. In this specific application, the two comparators must be able to sense voltage variations that exceed the prespecified sensitivity level (i.e., 10-15mV in this case). When the decaps are in parallel, the subthreshold leakage from the switches consumes considerable power due to the large sizes of the switch transistors. To reduce leakage current, the outputs of the comparators should be as close as possible to either V_{DD} or V_{SS}. The supply noise budget for a 1V power supply is normally less than 50-100mV and the output is full swing, indicating the need for high gain in the switching region.

With the above considerations, a latch-based comparator was selected for this application, as shown in Fig. 4.6. The exact transistor sizes are listed in Table 4.2. The branch current of the comparators is as follows: I_{b1} = 826µA, I_7 = 377µA, I_8 = 135µA, I_{b2} = 900µA, I_{17} = 339µA, and I_{18} = 136µA.





Figure 4.6: Complementary comparator design: (a) n-type input for the top comparator, and (b) p-type input for the bottom comparator.

Transistors	Width/Length	Transistors	Width/Length
Mb1 (NMOS)	75μm/0.1μm	Mb2 (PMOS)	150µm/0.1µm
M1/M2 (NMOS)	37.5µm/0.1µm	M11/M12 (PMOS)	75µm/0.1µm
M3/M4 (PMOS)	37.5µm/0.1µm	M13/M14 (NMOS)	18.75µm/0.1µm
M5/M6 (PMOS)	48µm/0.1µm	M15/M16 (NMOS)	24µm/0.1µm
M7/M8 (NMOS)	37.5µm/0.1µm	M17/M18 (PMOS)	75µm/0.1µm
M9/M10 (PMOS)	75μm/0.1μm	M19/M20 (NMOS)	18.75µm/0.1µm

Table 4.2: Transistor sizes of the comparators.

This two-stage architecture satisfies the need for high gain and full swing, but must be designed to avoid any potential stability or oscillation problems. For a latch-based first stage, introducing a certain amount of hysteresis will prevent the comparator from switching back to the standby state in the presence of small variations around the switching region. For the n-type input comparator shown in Fig. 4.6(a), the hysteresis voltage, V_{hys} , is given by [57]:

$$V_{hys} \approx 2\sqrt{\frac{2I_{b1}}{\mu_n C_{OX} \cdot (W/L)_{M1}}} \cdot \frac{\sqrt{\lambda} - 1}{\sqrt{1 + \lambda}}$$
(4.11)

where I_{b1} is the bias current. In Equation (4.11), λ is the size ratio [(W/L)_{M5}/(W/L)_{M3}] and $\lambda > 1$ for a latch. Once the slew rate and the bias current is determined, both I_{b1} and (W/L)_{M1} are fixed, leaving λ as the only parameter for V_{hys}. In this case, a λ value of 1.28 was chosen, producing a hysteresis voltage of around 5mV.

The second stage converts the differential signals into a signal-ended output and provides the requisite level shifting. The second stage is also used as an output buffer to drive the large switches, where the desired slew rate can be achieved by adjusting the bias currents and transistor sizes. Complementary designs are used for the top and the bottom comparators to have roughly equal switching delays. The bias voltages for the comparators are generated by simple current mirrors. PVT variations on the comparator and the bias generation can cause delay differences in the comparator outputs. This delay difference acts to further reduce the boosted voltage, as illustrated earlier in Fig. 4.5. During the design stage, great care has been given to ensure that the delay differences are within 100ps under all PVT variation simulations, which results in an additional 5% loss in the boosted voltage (i.e., 1.6V rather than 1.7V).

The dominant poles of this two-stage comparator were identified for stability compensation since there is a feedback path through the supply rails back to the comparator inputs. Therefore, the output resistance and the load capacitance of the comparator need to be carefully designed to properly position the dominant pole. In this case, a Miller compensation capacitance C_C is added to shift the dominant pole to a low frequency to improve stability. Also, a nulling resistor R_Z is present to cancel the right-half-plane zero [59].

The simulated large-signal DC characteristics of the n-type input comparator are illustrated in Fig. 4.7(a), where the curves with hysteresis are shown. Here, the switching region of the comparator is in the range of ± 10 mV. A λ value of 1.28 from Equation (4.11) was selected to produce about 5mV of hysteresis. The peak DC gain is around 48dB. The AC curve for the comparator is shown in Fig. 4.7(b) where the phase margin (PM) at unity gain is indicated as 39°.



Figure 4.7: (a) DC and (b) AC (compensated) characteristic curves for the two-stage latch-based comparator design (n-type input shown).

The active decap must be able to boost the supply voltage within one clock cycle such that the average supply noise per clock cycle is reduced, since this factor controls the path delay of the logic blocks [20][61]. In this case, our design goal was set to a maximum clock speed of about 1GHz, which makes it suitable for today's high-end ASICs, and even medium-speed custom designs. When the supply voltage drops to 0.9V (implying 100mV of noise, i.e., k=0.1), the average switching delay for a full output swing was designed to be 0.5ns, which should allow proper operation up to 2GHz. The boosted voltage, based on prior considerations, should be in the range of 1.6V. The charge demand of the logic circuit itself will cause an additional voltage drop of $kV_{DD}n^2 = 0.1 \cdot 1 \cdot 2^2 = 0.4$ V, resulting in an expected final voltage of 1.2V. In addition, the current drive of the comparators will act to reduce the supply voltage further, but hopefully keep the value above 0.9V, which is the noise budget.

The active decap was simulated and compared to prior architectures that were also redesigned in 90nm CMOS to quantify the improvement and design tradeoffs. The circuit proposed in [39] was first implemented, where it uses opamps in place of comparators, followed by a chain of inverters to drive the switches. The inverters were optimally-sized according to logical effort [39]. However, its minimum delay was about 0.9ns (1.2ns for the slow process corner) which is almost unsuitable for typical ASIC speeds, although its power dissipation was only 0.8mW. In the second design in [37][38], the sensing circuitry is formed by a pseudo-cascode amplifier delivering high speed at the cost of high power. The original design [37] was implemented in a 0.15µm process. The design was adapted to the 90nm process and it was found that, by replacing their comparator design with the latch-based version, the static power consumption of the switching circuitry is reduced from 13mW to 2.8mW, an improvement factor of almost 5X,

while the delay only increases slightly from 0.4ns to 0.5ns. The comparison of the three designs is provided in Table 4.3, where the design parameters that fail to satisfy the specifications are shown in italic. Note that the new design features hysteresis while the other two do not. The small-signal ac characteristics of the three designs are shown in Fig. 4.8. For the circuit in [39], the chain of inverters was removed for small-signal analysis.

	Specifications	[37][38]	[39]	This work
Process	1V-core 90nm STM			
Switching delay (typical)		0.4 ns	0.9 ns	0.5 ns
Switching delay (slow)	< 1 ns	0.5 ns	1.2 ns	0.75 ns
Bandwidth	>1 GHz	2 GHz	0.8 GHz	1.5 GHz
Static power*	< 5 mW	13 mW	0.8 mW	2.8 mW
Hysteresis voltage		0	0	4.2 mV

Table 4.3: Simulated switching circuit design specification comparison.

*Switching circuitry only



Figure 4.8: AC characteristic curves for the three designs.

4.3 Chip Design and Experimental Results

4.3.1 Test Chip Setup

A test chip was fabricated in a standard 90nm 1V-core CMOS process with seven metal layers to validate the results and to quantify the degree of improvement as operation frequency increases when an active decap is used as a drop-in replacement for a passive decap. The test chip setup is shown in Fig. 4.9, where an active decap, a passive decap, and some user logic are implemented. The layout of the active decap is shown in Fig. 4.10. The switching circuitry is located in the center, with the two parallel decaps on either side. The decap on the left is PMOS, and the one on the right is NMOS. The total layout area for the active decap is 600 μ m x 142 μ m = 0.085mm², in

which the two decaps on either side combine for an area of 0.077mm². The switching circuitry, including switch transistors (Mn1, Mn2, Mp1 and Mp2), accounts for only 10% of the area and this does not greatly affect the final voltage drop, as shown below.









Figure 4.10: Layout of active decap showing the relative size of the components.

The area overhead consumed by the sensing and switching circuitry should be considered as an additional penalty for the active decap performance. The percentage area overhead can be defined as x, then the charge provided by the active decap is:

$$Q_{series} = [bV_{DD} - (1-k)V_{DD}] \cdot (1-x)C_{decap} / n$$
(4.12)

Thus, Equation (4.7) should be re-written as:

$$G = \frac{(bV_{DD} - aV_{DD}) \cdot (1 - x)C_{decap} / n}{kV_{DD} \cdot nC_{decap}} = \frac{(b - a)(1 - x)}{kn^2}$$
(4.13)

Assuming k=0.1, n=2, and b=1.5, the final voltage *a* can be plotted as a function of the area overhead *x*, as shown in Fig. 4.11. From the figure, it is clear that the area overhead should be limited to within 30% to achieve a reasonable final voltage. If the area overhead is above 43%, using the active decap brings no benefit. In our case, x=10% so that the penalty is only 50mV.



Figure 4.11: Final voltage a as a function of sensing and switching circuitry area overhead x.

The switch sizes were chosen to have a suitable parasitic series resistance to provide ESD protection, sufficient transient response [12] and good damping for potential LC resonance [13]. In our case, the two parallel decaps are formed using thin-oxide transistors to improve area

efficiency, since ESD is not a major concern due to switch resistances inherent in the circuit. The decoupling capacitance values in the standby mode are 0.34nF each, resulting in a total of 0.68nF in parallel.

The extra passive decap of Fig. 4.9 was used to represent fixed decap that is always present in the neighborhood of the active decap. It cannot be shut off. It also employs a series PMOS device to protect it from ESD risks. Both active and passive decaps are placed about 600 μ m away from the user logic. Ref. [41] uses a linear feedback shift register (LFSR) as the user logic to generate power supply noise because the resulting noise pattern is somewhat randomized. In our design, simply a large buffer with a large capacitive load was used to create supply noise, with the input controlled by an external signal. This way, the switching frequency can be controlled and modified directly from the input. The size of the decaps was chosen to be only a few times larger than the capacitive load to create a ~100mV voltage drop for the experiments. The three resistors (R₁, R₂ and R₃ shown in Fig. 4.3) were implemented using p+ poly resistances. The two capacitors C₂ and C₃ in Fig. 4.3 were implemented using MOS transistors to minimize area overhead. A test chip microphotograph is illustrated in Fig. 4.12. The test chip area is totally 1.2 x 0.8 mm².



Figure 4.12: Annotated test chip microphotograph.

To measure the on-chip supply noise, a packaged die was not used because it was intended to observe internal voltages near the logic block. Thus, the supply variations were measured directly with probes. Power supply noise comes from both *IR* drop and *Ldi/dt* effects. The inductance *L* in the *Ldi/dt* effect is mainly due to the package, as the on-chip wire inductances are normally negligible [2]. Since an actual package was not used, two on-chip spiral inductors were implemented to mimic the package inductances, one on the supply path and the other on the return path. The value of the spiral inductors is close to a typical wire-bond package inductance. The user logic and the decaps were placed far away from the supply/ground pads (about 600 μ m) to create a large mesh resistance. Effectively, the pad and mesh of the test chip were designed to

produce a measurable amount of power supply noise such that any improvements of using active decaps could be easily observed.

4.3.2 Test Chip Simulations

Before showing the measurement results of the test chips, it is desired to illustrate the simulation results first to demonstrate the close relationship of the two and to provide a better understanding of how the active decap actually behaves when large supply noise is present. The simulation setup follows exactly the test chip in Fig. 4.9. When a clock signal is fed into the test chip, the buffer, as the user logic circuit, switches to draw certain current from the supply rails. The current perks created from the buffer cause the supply voltage V_{DD} to drop, resulting in the two input voltages of the comparators to swap. The comparators then switch their output levels according to the input swap. After certain delay, the outputs of the comparators switch, causing a local boost at the supply voltage V_{DD}. Once the supply is boosted, the inputs of the comparators move back to their nominal values. As a result, the outputs of the comparators switch back after some delay and the active decap waits for the next voltage drop. When the active decap is turned off, the circuit behaves like a passive decap. In such a scenario, the supply voltage follows the current draw of the buffer, and no boost in the supply voltage is achieved. From post-layout HSPICE simulation, the current draw, the supply voltage V_{DD} and the internal signal switching are shown in Fig. 4.13, where the clock frequency is set at 500MHz.



Figure 4.13: Simulated V_{DD} voltage (on a 500MHz clock) with active decap on and off.

Random process variations exist to affect the effectiveness of the active decap. At the slow corner, the delay of the comparators is increased, resulting in a later boosting point at the supply rail. On the other hand, for the fast process, the comparators take a shorter delay to switch so that the local supply boost occurs earlier. As a consequence, the supply voltage remains high for a longer time for the fast corner, and for a shorter time for the slow corner, as illustrated in Fig. 4.14. Therefore, the average supply noise per clock cycle should be less for a test chip at the fast process corner. Alternatively, a larger average supply noise level can be expected if a test chip happens to be at the slow process corner. Designers should make sure that the active decap provides satisfactory improvement to remove hot spots under process variations, especially at the slow process corner.



Figure 4.14: Simulated V_{DD} voltage with active decap on for different process corners.

4.3.3 Test Chip Measurements

This section describes the measurement results obtained on 15 test chips and validates the simulation results, as illustrated in Fig. 4.15. An Agilent 86130A bit error rate tester was used to drive the inputs, while an Agilent DSO81304A oscilloscope was used to observe the results. As mentioned earlier, the passive decap shown in the figure is always present in the test chip. The *enable* signal is used to selectively turn on or off the active decap by applying a high or low voltage. When disabled, the decaps are biased in parallel, utilizing a maximum standby capacitance. In that case, the active decap behaves purely as a passive decap. When enabled, the active decap is triggered by voltage drops of about 50mV. By turning on and off the active decap, the average V_{DD} voltage improvement can be measured. A collection of 15 sample chips are tested, where the clock frequency is fixed at 1GHz to observe the improvement across the sample space. The test results from these sample chips were categorized into three groups: slow, typical, and fast, to reflect the nature of random process variation on silicon. Note that in all cases, the active decap moved the *IR* drop inside the 100mV noise budget. The average V_{DD} voltages of each group line up closely with simulation under process variations.



Figure 4.15: Grouped scatter plot for active decap noise reduction for the tested sample chips.

In Fig. 4.15, when the active decap is off, the average supply voltage varies significantly from 873mV to 910mV. This is due to the random process variation on R_{mesh} and L_{pack} , which determines the *IR* drop on the supply rails. But more importantly, a higher average V_{DD} value when the active decap is off results in a smaller improvement when turning on the active decap. This fact is caused by the nature of the comparators. Specifically, if the input of the comparator varies with a large swing, then the comparator will generate an output with a shorter delay. That is, for a larger supply noise *k*, the delay of the comparator t_d is smaller. This effect can be illustrated in Fig. 4.16. Longer delay results in lower bandwidth, which affects the active decap performance at high frequencies. Note that the delay difference Δt_d also increases as *k* increases, which will reduce the boosted voltage *b* slightly.



Figure 4.16: Comparator delay t_d and delay difference Δt_d as a function of supply noise k.

Corners	t _d (top)	<i>t</i> d (bottom)	$\Delta t_{\rm d}$	Average t _d
Slow	0.77 ns	0.72 ns	0.05 ns	0.75 ns
Typical	0.53 ns	0.48 ns	0.05 ns	0.51 ns
Fast	0.32 ns	0.30 ns	0.02 ns	0.31 ns

Table 4.4: Comparator delay t_d and delay difference Δt_d in different corners.

The different process corners have direct impact on comparator delay and delay difference. For a fixed value of k=0.1, the comparator delay and delay difference under process variation is highlighted in Table 4.4. The average delay between the top and the bottom comparator varies

from 0.75ns (slow) to 0.31ns (fast), more than 140% of variation. This effect explains the on-die variation for the active decap performance measurements. However, the delay difference of the comparators only varies slightly, indicating that the boosted voltage b will not be affected greatly by process variations.

By averaging the average supply voltage from each group in Fig. 4.15, the overall improvement of using the active decap can be assessed, as highlighted in Table 4.5. Note that in the test chip the passive decap is always connected. To illustrate the improvement solely due to use of active decap, simulations were carried out with the passive decap completely removed. The simulated results showing active decap only versus passive decap is summarized in Table 4.6, where the case for the active decap provides a higher average supply voltage across the process corners. Due to the dynamic power consumption of the comparators during switching and other non-idealities, the active decap cannot practically reach the final voltage value described in Equation (4.13). Given a level of the supply noise when the active decap is off, the final voltage when the active decap is turned on, the two values are close (~100mV of difference) for the cases of typical and fast process corners, as summarized in Table 4.6.

Corners	Average V	Improvement	
	active decap ON	active decap OFF	
Slow	914 mV	903 mV	11 mV
Typical	904 mV	878 mV	26 mV
Fast	917 mV	872 mV	45 mV

Table 4.5: Measured active decap performance for different process corners.

Table 4.6: Comparison between equation and simulated result after correlation.

Corners	Final voltage (from eqn.)	Simulated avgSinal voltage(from eqn.) aV_{DD} (correlated)	Measured average V _{DD} voltage	
	$a \mathrm{V}_\mathrm{DD}$		active decap ON	active decap OFF
Slow	1169 mV	932 mV	914 mV	903 mV
Typical	1058 mV	909 mV	904 mV	878 mV
Fast	1031 mV	921 mV	917 mV	872 mV

It is now possible to validate Equation (4.8). Although the test equipment has a limited bandwidth of less than 1.5GHz, the test results for a 1GHz clock can be used to correlate simulations for high frequency effects. The three process corners are used here. Simulations were

carried out from 1GHz to 3GHz, showing the cross points of active decap and passive decap for the average supply voltage. The simulation results are then correlated with measurement results for the three process corners. The correlated results are shown in Fig. 4.17. The relationship between the active decap bandwidth (crossing point) and the average comparator delay are summarized in Table 4.7. Clearly, Equation (4.8) captures the effect of process variation corners.

Table 4.7: Active decap bandwidth versus average comparator delay under process corners.

Corners	Bandwidth	Average delay
Slow	1.55 GHz	0.75 ns
Typical	2.4 GHz	0.51 ns
Fast	> 3 GHz	0.31 ns





Figure 4.17: Simulated average V_{DD} voltage with active decap on and off versus clock frequency for (a) slow, (b) typical, and (c) fast process corners.

4.3.4 Measurement Results on One Typical Chip

Three sample chips were found to be at the typical process corner. The average V_{DD} voltages when the active decap is on and off can be found back in Table 4.5. A sample chip that has a close value to the average voltage of the three typical chips was used for further analysis in this section. As before, by turning on and off the active decap, the average V_{DD} voltage improvement was measured. This is shown in Fig. 4.18, where the input is set at 500MHz, typical of ASIC designs. In Fig. 4.18(a), an actual screen shot of the input and supply voltages are provided with the active decap enabled. In Fig. 4.18(b), the supply waveforms for the passive and active cases are superimposed. The average supply voltage increases from 900mV to 914mV. Therefore, the noise level dropped from 100mV to 86mV, an improvement of 14mV (or about 14% less noise). This improvement can be expected to be almost doubled for an isolated active decap, as illustrated later using simulation.



(a)



Figure 4.18: Measured results (on a 500MHz clock) for (a) active decap on and (b) plotted comparison between active decap on and off.

Fig. 4.19 shows the measured points as the external input frequency increased from 200MHz to 1GHz. The measurements at 500MHz described above are circled. Two solid trend lines are provided corresponding to active decap on and off. The gap between the two trend lines initially widens indicating that the benefit of active decaps increases as frequency increases. The test chip validated that the active decap has a maximum improvement of 23mV (or about 20% less noise) for a 1GHz design. Circuit simulation was used to further study the effect of higher clock frequencies, also shown in Fig. 4.19 as dashed lines. Below 2GHz, the active decap can provide more charge than the passive decap. Above 2GHz, its performance diminishes because of the fixed switching speed. The crossing point of the two trend lines at about 2.4GHz indicates the

bandwidth limits due to the switching nature of this active decap design. However, today's highend ASIC designs still run at below 1GHz, so this active decap is quite acceptable.



Figure 4.19: Measured (0.2 - 1GHz) and simulated (1 - 2.5GHz) average V_{DD} voltage with active decap on and off versus clock frequency.

In Fig. 4.19, the maximum difference between the active decap on and off occurs at about 1GHz, about a half of the active decap bandwidth. As described in the previous section, the high-frequency cross point of the active decap and the passive decap is at about 1/0.5ns=2GHz, where 0.5ns is the comparator delay. At low frequencies, the active decap and the passive decap have similar average V_{DD} values because the clock period is long enough to eliminate the advantage of using the active decap when taking an average V_{DD} level per clock cycle. As a result, since the difference is almost zero at low frequency and at the bandwidth frequency, the maximum

difference (or benefit) can be considered to occur at roughly 1/2 of the active decap bandwidth, in this case, 2GHz/2=1GHz. Therefore, to maximize the effectiveness of the active decap, designers should ensure that the comparator delay t_d is always less than 1/2 of the clock cycle.

4.4 Active Decap Size and Placement

While the test chips are useful in quantifying active decap improvement over passive decap as frequency increases, the proper sizing and placement of the active decap determines the effectiveness of the drop-in replacement approach. When converting a fixed area from passive decaps into active decaps, the standby capacitance is always smaller because of the area overhead of the switches and comparators in the active decap. The actual noise improvement depends on the area available and the location of the active decap relative to the hot spot. Intuitively, if the area available for active decaps is small and the overhead area is a large percentage of the total area, active decaps may not be an effective replacement for passive decaps. On the other hand, if the area available is too large, the noise reduction for active and passive decaps may be similar because of excessive delays in the switching circuitry trying to switch the oversized C_{decap} 's. Therefore, there exists an optimal area for active decaps where they are most effective. Similarly, the instantaneous boost provided by active decap must be close to the hot spot to be effective, but this should be traded off against the distance from the supply to replenish the charge.

To explore these aspects, extensive circuit simulation was used to first calibrate the test chip measurements with HSPICE simulation using exactly the same setup in Fig. 4.9. As a calibration

metric, the average V_{DD} noise per clock cycle, $V_{DD_{noise}}$, was used since it is known to be the controlling factor of the critical path delay of logic circuits [20]:

$$V_{DD_noise} = V_{DD_nominal} - V_{DD_avg} = 1 V - V_{DD_avg}$$
(4.14)

With a 500MHz input and the active decap turned on, the waveforms from circuit simulation of the supply voltage and internal signals were previously shown in Fig. 4.13. These results are not identical to the measured results of Fig. 4.18 but they do, in fact, have a similar average value. For one clock cycle, $V_{DD_avg} = 918$ mV, which is fairly close to the measured average of 914mV. It was found that, for other frequencies, V_{DD_avg} closely matched the running average from measurement. Since the measured and simulated values are correlated in this way, the average was used for the rest of the analysis. The fixed passive decap was also removed from the circuit in further analysis to study the improvements derived from the stand-alone active decap.

While V_{DD_noise} is determined by many other factors including package/pad/power grid design and clock frequency, for the purpose of this analysis the same power grid design and the 500MHz input were kept the same, and only the decap size varied. The average noise for samearea passive and active decaps was compared. The size of the center switching circuitry of Fig. 4.10 also remained constant. In the passive decap simulations, standard cross-coupled designs were used. In Fig. 4.20, the average noise V_{DD_noise} is plotted versus passive and active decap size varying from $85\mu m^2$ to $8.5mm^2$. The plot shows that the active decap reduces the noise relative to passive decaps for sizes between $0.001mm^2$ and $0.6mm^2$. However, if the available area for decap insertion is smaller than $0.001mm^2$ or greater than $0.6mm^2$, there is little difference between the two. When the area is small, the active decap is not as effective since the amount of capacitance switched in series is small. When the area is large, the fixed switching circuit in the active decap cannot switch the decaps effectively because the capacitive load exceeds its capability.



Figure 4.20: Simulated average V_{DD} noise per clock cycle versus normalized decap size.

Fig. 4.21 is used to illustrate the optimal size for the active decap design, where the solid curve indicates the noise reduction difference between passive and active decaps. The maximum difference occurs in the range of 0.01mm^2 to 0.1mm^2 . If the active decap is designed in this range, it has the greatest advantage over passive decaps in terms of average supply noise reduction. The test chip was designed to be 0.085mm^2 to obtain close to optimum improvement of 23 mV, as described in the previous section. In Fig. 4.21, it is also shown that the area overhead of the switching circuitry in the design is only 10% in the region around the optimal value.



Figure 4.21: Power supply noise reduction difference from active decap and passive decap with area overhead from switching circuit of active decap.

As mentioned earlier, another important factor in the resulting improvement is the actual placement of active decaps relative to the hot spot. Referring back to Fig. 4.9, the effective distance from the hot spot can be adjusted by varying R_{dist} . Similarly, the distance from the charge re-supply path can be controlled by varying R_{mesh} . Simulations were carried out to observe the voltage drops while changing only R_{mesh} and R_{dist} . The simulation results are shown in Fig. 4.22. The decap size was fixed at the optimal value of 0.02 mm^2 from Fig. 4.20 so that the maximum improvement could be observed. As the distance between the decap and the user logic is varied from $10R_{dist}$ to $0.1R_{dist}$, the average noise level in the passive case changes from 134mV to 124mV. However, for the active case, the average noise level reduces from 133mV to 74mV.

Therefore, the active decap is more sensitive to placement than the passive decap. This makes intuitive sense because the active decap provides a short-term boost in the charge which acts in a small, localized neighborhood. However, the passive and active decaps exhibit similar characteristics as a function of R_{mesh} , according to the results in Fig. 4.22. As a result, the active decap should be placed as close as possible to the hot spot to be most effective.



Figure 4.22: Improvement on average V_{DD} noise for using active decaps in different placement locations by varying Rdist and Rmesh.

4.5 Summary

This chapter described the effectiveness of active decaps as a late-stage drop-in replacement for passive decaps so that a completed chip layout need not be disrupted near the tapeout deadline. Improvements to the design of an active decoupling capacitor were described for removal of hot-

spot power supply noise in ASIC designs up to 1GHz operation. The modified active decap using latch-based comparators in 90nm CMOS is able to switch in 0.5ns and consumes a relatively low power of 2.8mW, which is about 5X lower than a previous design running at approximately the same speed. This reduced power makes it more suitable for use in ASIC designs. Measurement results from test chips indicate improvement over passive decaps of 10% - 20%, operating from 200MHz to 1GHz. The optimal active decap size to maximally remove hot-spot noise was identified. Placement analysis was also carried out and it was found that the active decap is most effective when placed in close proximity to the hot spot, as compared to the passive decaps which is not as sensitive to the exact location. In summary, if sized and placed properly, active decaps can be up to 20% better when used as drop-in replacements of passive decaps for power supply noise reduction.

Chapter 5

Generalized Active Decap and Charge-Borrowing Decap

5.1 Introduction

The previous chapter explored the effectiveness of using active decaps to remove hot-spot IRdrop violations. This chapter investigates advanced versions of the active decoupling capacitor and proposes a novel design of a charge-borrowing decap (CBD). The extension of the active decap concept is derived by increasing the stack height n to a larger value to ideally achieve a higher boosted voltage than the basic n=2 active decap. The optimal number of n will be evaluated in theory and practical applications [62]. The CBD design is a completely different approach to addressing the hot-spot removal problem [63]. The new design aims to provide enough charge during every cycle to reduce IR drop with only a minimum power overhead. However, the location of the IR-drop problem must be known in advance and sufficient area must be available with a relatively clean supply to implement the solution. The applications and limitations of the charge-borrowing decap will be evaluated in this chapter.

5.2 Extended Active Decoupling Capacitor

5.2.1 Optimal Stack Height n

The concept of the extended active decap is simply to increase the stack height n of the basic active decap described in Chapter 4. The motivation to have a larger stack height (n>2) is to generate a higher boosted voltage nV_{DD} to potentially achieve a better improvement when applied to reduce supply noise. For example, the ideal boosted voltage is $3V_{DD}$ for n=3, $4V_{DD}$ for n=4, and so on. Therefore, it seems that the stack height should be designed as large as possible to obtain a high enough local boost so that the supply noise can be reduced to an arbitrarily small level. However, this is not true in practice. The higher boosted levels cannot be reached due to the nonidealities of the circuit. Also, by increasing the stack height, more switches will be required to turn the decaps in parallel or in series. The active decap circuits for n=2, n=3 and n=4are illustrated in Fig. 5.1(a), 5.1(b) and 5.1(c), respectively. In the figure, it is assumed that the total area available for the decaps is fixed at $2C_{decap}$. Therefore, the decap occupies an area of C_{decap} for n=2, while the cases of n=3 and n=4 have an area of $(2/3)C_{decap}$ and $(1/2)C_{decap}$, respectively.









decaps in series



(b)

(c)

Figure 5.1: Active decap concept showing different stack height: (a) n=2, (b) n=3, and (c) n=4.
The practical constraints of stacking the decaps can be illustrated by first showing the actual transistor implementation of the n=3 case in Fig. 5.2, where the switches are implemented using MOS transistors. Note that the number of switches required is increased by three every time the stack height is increased by one. That is, for $n\geq 2$, the number of switches = 3(n-1). In the figure, each horizontal switch is implemented with two transistors (NMOS and PMOS), whereas each vertical switch requires only one transistor (NMOS or PMOS). The design of the stack height of n=4 can be realized in a similar manner, although not shown here.



Figure 5.2: MOS implementation of the extended active decap (n=3).

Due to the additional switches, two methods of design exist: one is to expand the area occupied by the switch transistors, and the other is to reduce the size of each switch transistor if a relatively constant total area needs to be maintained. The *first* method increases the area overhead x of the active decap, and also causes a longer comparator delay because of the additional loading capacitance. The increased area overhead reduces the final voltage aV_{DD} to which the active decap can boost the supply voltage. But more importantly, a longer delay results in a reduced bandwidth, which lowers the operating frequency, as described in the previous chapter. The *second* method uses a fixed total area occupied by the switch transistors. By this approach, the delay of the comparators should remain roughly the same since there is only a minimum change in the loading capacitance. Therefore, the operating frequency is not affected. However, packing more switch transistors into the same area can cause the "on" resistance R_{on} of each transistor to increase, which reduces the boosted voltage bV_{DD} .

The second method to implement the extended active decap was used since the active decap bandwidth should not be compromised for high-end ASIC applications. That is, different designs are designed to have the same operating bandwidth. The final voltage aV_{DD} can be obtained by varying the stack height, n, as illustrated in Fig. 5.3, with different supply noise levels, k. The large dots indicate the highest final voltage points on the different k curves. Note that the starting point of n=1 implies a passive decap. Clearly, for k=0.05, the optimal n value that produces the highest final voltage is 4, whereas for k=0.1, the optimal value is n=3. For an increased level of k=0.15, the optimal n value reduces to 2. As the supply noise k further increases, the use of passive decaps (n=1) is recommended, as for the case of k=0.2. From the figure, one can conclude that a higher n should be used if k is small, and vice-versa. Therefore, it is important to select n based on the k range where the resulting final voltage aV_{DD} is the highest. In order to do this, the optimal k ranges for each n and the crossover points between ranges must be determined.



Figure 5.3: Final voltage aV_{DD} as a function of stack height *n* with *k* varying (fixed area).

The supply noise crossover point, $k_{n1,n2}$, for two different stacking levels, $n=n_1$ and $n=n_2$, is defined when both cases produce the same final voltage. This can be used to identify suitable ranges for each stacking level. To obtain $k_{n1,n2}$, Equation (4.13) can be rearranged into the following form:

$$k_{n1,n2} = \frac{(b|_{n2} - b|_{n1})(1 - x|_{n1})(1 - x|_{n2})}{(n_2^2 - n_1^2) + n_1^2 x|_{n2} - n_2^2 x|_{n1}}$$
(5.1)

where $n_1 \neq n_2$ and $n_2 > n_1$. Plugging in numbers, the crossover noise value from n=2 to n=3 is $k_{2,3}=0.12$, and from n=3 to n=4 is $k_{3,4}=0.08$. Effectively, the crossover point $k_{4,5}=0.05$ determines the boundary where a passive decap should be used since the case of n=5 would not be used if

the noise was 5% or less (i.e., acceptable level). Similarly, $k_{1,2}=0.17$ produces the same final voltage for n=2 and n=1 (passive decap). When k is above 0.17, the passive decap should be used. The results are presented in a graphical form in Fig. 5.4. The four lines represent n=1, 2, 3 and 4, respectively. The line with the highest value in each region is the optimal value for n. For low values of k, the best choice is n=4. Starting at $k_{3,4}$, the best choice is n=3. At $k_{2,3}$, the best choice becomes n=2. At $k_{1,2}$, the best choice is n=1 from that point onward. The results are summarized in Table 5.1.

Condition	Optimal <i>n</i>	
k < 0.05	n=1 (use passive decap)	
0.05 < k < 0.08	<i>n</i> =4	
0.08 < k < 0.12	n=3	
0.12 < k < 0.17	n=2	
k > 0.17	n=1 (use passive decap)	

Table 5.1: Optimal stack height *n* selection based on the supply noise *k* (from formula).



Figure 5.4: Final voltage aV_{DD} as a function of k with different stack height n (from formula).

As described earlier, if the supply noise is above 0.17, the use of any form of the active decap cannot boost the supply voltage to a satisfactory level. Other design approaches to reduce the supply noise may have to be used in that situation. However, the more interesting range of k is from 0.08 to 0.17, since this noise level is typically unacceptable. If the supply noise k is below 0.12 but above 0.08, then the active decap should be designed with n=3 to produce the minimum noise. Similarly, if k is above 0.12 but below 0.17, the basic active decap with n=2 is optimal.

5.2.2 Design and Layout of *n*=3 Extended Active Decap

To validate the results of optimal *n* selection, the extended active decaps with n=3 and n=4 were implemented. For simplicity, only the design of n=3 is illustrated in this section. Similar to the basic active decap design, Fig. 5.5 illustrates the extended active decap for n=3 containing four blocks: a reference voltage generator, a pair of high-pass filters, two comparators, and the switched decaps. The operation of the circuit remains the same: the differential inputs of each comparator decide the standby voltage level at the outputs of the comparators. When the power grid discharges, V_{DD} will drop and V_{SS} will rise. The voltage variations are passed through the high-pass filters to the comparator inputs causing the comparators to reverse their output values, and switch the three-piece decaps from parallel to series. Later, when the power grid charges up, the comparator inputs and outputs switch back to their original values. An *enable* signal is provided for testing purposes to allow the active decap circuitry to be connected to the global supply rail. Unlike the previous design, when off, the extended active decap is disconnected from the rest of the circuit. This allows for a comparison between the basic and the extended decaps.



Figure 5.5: Extended active decap (n=3) architecture.



Figure 5.6: Layout of extended active decap (n=3) showing the relative size of the components.

The layout of this extended active decap is shown in Fig. 5.6. The switching circuitry is located offset from the center, with the three parallel decaps on either side. The two decaps on the left are PMOS, and the one on the right is NMOS. The total layout area for the active decap is $600\mu \text{m} \times 142\mu \text{m} = 0.085\text{mm}^2$, in which the three decaps combine for an area of 0.077mm^2 . Each switch transistor was layed out with only a half of the area as before, resulting in an almost doubling of the "on" resistance. As a result, the switching circuitry, including the switch transistors, still accounts for only 10% of the area overhead. Note that the total area is the same as the basic active decap so that a comparison between the two can be carried out. Although not shown, the design and layout of the *n*=4 case is similar to the case of *n*=3.

5.2.3 Simulation Results

The next step was to use simulations to obtain the supply voltage waveforms under different supply noise levels, k. By increasing the size of the user logic buffer, k can be varied and the supply voltages will differ in the cases of n=2 and n=3, as illustrated in Fig. 5.7. When k=0.12, case n=3 provides a larger boost than n=2. On the other hand, when k=0.15, both n=2 and n=3 were insufficient in terms of delivering charge to boost up the supply, but the n=3 case is even

worse. As shown in the figure, by using the second method, the delay of the comparators remains roughly the same. The design of n=4 active decap was also simulated, and the results are similar to the n=2 basic active decap case. One noticeable difference is that as k increases above 0.12, the average V_{DD} voltage drops earlier for the n=4 case than the n=2 case. However, when k is at around 0.05, the average V_{DD} voltage for n=4 is slightly higher than both n=2 and n=3 cases.



Figure 5.7: Simulated V_{DD} voltage with extended active decap (n=3) on for two different k levels.



Figure 5.8: Average V_{DD} voltage as a function of k with different stack height n (from simulation).

Condition	Optimal <i>n</i> (simulated)	
<i>k</i> < 0.05	n=1 (use passive decap)	
0.05 < k < 0.07	n=4	
0.07 < <i>k</i> < 0.14	n=3	
0.14 < <i>k</i> < 0.16	n=2	
k > 0.16	n=1 (use passive decap)	

Table 5.2: Optimal stack height *n* selection based on the supply noise *k* (from simulation).

Using simulations, the average supply voltages per clock cycle for the n=2, 3 and 4 cases when the supply noise k varies from 0.02 to 0.2 were compared and plotted in Fig. 5.8. The corresponding optimal stack height n as a function of the supply noise k is summarized in Table 5.2. The crossover points of $k_{n1, n2}$ are similar between formula and simulation. The most important crossover point of the n=2 and n=3 cases from simulation is $k_{2,3}=0.14$, somewhat higher than the calculated value of 0.12. Above 0.14, no approach can raise the supply level back to 900mV, making the use of active decaps less valuable in this region. On the other hand, the lower bound of $k_{3,4}$ is at 0.07, slightly below the calculated level of 0.08. Therefore, a slightly wider k range of 0.07~0.14 for n=3 makes it superior to the basic active decap. Unlike the formula, when the k value is low, the active decaps do not switch due to the fixed triggering voltage of about 50mV, resulting in the active decaps producing slightly worse average supply voltage than the passive decap. However, the active decaps become worse then the passive decap when k < 0.05, and they should not to be used. Although the n=4 case is the best when k is in the range of 0.05~0.07, it only has limited value since this k range is small and the improvement over the n=3 active decap is marginal. Thus, it can be concluded from simulation that n=3provides the optimal level of the average supply voltage across a wide supply noise range of below 0.14. If the supply noise is above 0.14, a larger area is required to increase the average supply level to a satisfactory level above 900mV.

5.3 Charge-Borrowing Decap (CBD)

5.3.1 Charge-Borrowing Decap Concept

The main purpose of the active decap is to boost the voltage locally to reduce supply noise. Therefore, any technique that offers this type of improvement would also qualify as a viable alternative. For example, if charge is "borrowed" from a clean supply to boost up a noisy supply, it would help reduce the hot-spot *IR*-drop problem. That is the basic concept behind a charge-borrowing decap (CBD), which is a *novel* but relatively simple idea illustrated in Fig. 5.9. The key idea here is based on capacitive feedthrough. Assuming that the total area available is the same as before, the decoupling capacitance is $2C_{decap}$. In Fig. 5.9(a), when power supply noise kV_{DD} is present, a passive decap provides charge equal to $(2C_{decap})(kV_{DD}) = (2k)C_{decap}V_{DD}$. In the case of the CBD, as shown in Fig. 5.8(b), it can boost the local supply voltage to $2V_{DD}$ ideally, similar to the active decaps. From another perspective, the charge provided by the CBD circuit in one clock cycle can ideally be up to $(2C_{decap})\Delta V_{clk} = 2C_{decap}V_{DD}$, where ΔV_{clk} is the clock swing. Therefore, over one clock cycle, the charge-borrowing decap provides significantly more charge than a same-area passive decap.



Figure 5.9: Charge-borrowing decap concept shown in (b), compared to a passive decap in (a).

	Charge provided (ideal)	Boosted voltage (ideal)
Passive decap	$(2k)C_{decap}V_{DD}$	-
Active decap (n=2)	$(1/2)C_{decap}V_{DD}$	2V _{DD}
Active decap (n=3)	$(4/9)C_{decap}V_{DD}$	3V _{DD}
Active decap (n=4)	$(3/8)C_{decap}V_{DD}$	4V _{DD}
Charge-borrowing decap (CBD)	$2C_{decap}V_{DD}$	$2V_{DD}$

Table 5.3: High-level comparison among passive, active, and charge-borrowing decaps.

A high-level comparison between passive decap, active decaps, and CBD is highlighted in Table 5.3, where the differences in charge available and boosted voltage are shown. Note that the total decap area is fixed to $2C_{decap}$ in the comparison. In Table 5.3, the basic active decap (n=2) is better in charge provided than the passive decap if 2k is less than 0.5 (or, k<0.25), plus the active decap also provides a local boost in the supply voltage. The charge-borrowing decap is always superior to basic active decap since it supplies more charge while generating the same boost voltage level. The charge supplied from active decaps with n=3 or n=4 is about 11% or 25% less than the basic active decap (n=2), respectively, but their boosted voltages are higher. This intuitively explains the fact that the extended active decap is better than the basic case only in limited situations. From the concept, it is difficult to conclude the superiority between the extended active decaps and the CBD. More design details need to be studied before a conclusion can be made. From the table, the design that provides the most charge per clock cycle is the charge-borrowing decap so that it was named after this feature. If designed properly, the use of

charge-borrowing decaps can potentially remove hot spots as a drop-in replacement of passive decaps, similar to the active decaps. The rest of this chapter will provide results to support this argument.

In Fig. 5.9(b), there is a problem with the falling edge of the clock. When the Clk signal rises from 0 to V_{DD} , the supply is boosted to $2V_{DD}$ ideally due to capacitance feedthrough. Then, the user logic nearby switches, resulting in certain amount of charge to withdraw from the supply. The supply voltage will drop by ΔV . And then the Clk signal falls from V_{DD} back to 0, forcing the supply voltage to drop from $2V_{DD}$ - ΔV to V_{DD} - ΔV , which may not be acceptable if ΔV is excessively large. Therefore, the concept of charge-borrowing decap needs some additional circuitry to function properly, as shown in Fig. 5.10(a). Two diodes are inserted at node B1, one from the clean supply and the other from the noisy supply. Without the connection to the clean V_{DD} , when the clock is low, B1 stays at roughly V_{SS} since the current flow from V_{DD} to B1 is prevented by the diode, D2. When the clock goes high, B1 rises to V_{DD}. This boost in voltage will not trigger current flow from B1 to the supply because both B1 and the supply are at the same level of V_{DD} . Therefore, the voltage at B1 should be around V_{DD} when the clock is low. This ensures that the voltage at B1 can reach about 2V_{DD} when the clock goes high and the supply can be charged. To achieve that, access to a clean supply of V_{DD} is needed through D1. Therefore, the implementation of the charge-borrowing decap requires a clocking signal, two diodes, and a supply node that has less noise (clean).

Assuming there is one threshold voltage V_T drop across each diode, the operation of the CBD circuit can be illustrated in Fig. 5.10. When Clk is at 0, the noisy supply node is assumed to be at

 V_{DD} , while B1 is charged at V_{DD} - V_T . When Clk rises to V_{DD} , B1 rises to $2V_{DD}$ - V_T at the same time. As a result, the noisy supply node is increased to $2V_{DD}$ - $2V_T$. Before the clock falls, some drop occurs at the supply, causing it to drop by ΔV to $2V_{DD}$ - $2V_T$ - ΔV . Then, Clk falls back to 0, so that B1 also reduces to V_{DD} - V_T . However, D2 prevents charge from flowing back to B1 from the supply. Therefore, the noisy supply remains at $2V_{DD}$ - $2V_T$ - ΔV .



Figure 5.10: Diode inserted charge-borrowing decap showing the states of the clocking signal: (a) Clk at 0, (b) Clk rises to V_{DD} , and (c) Clk falls back to 0.



Figure 5.11: Possible implementations of charge-borrowing decap with (a) NMOS-formed diodes and (b) PMOS-formed diodes.

In a CMOS process, the diodes shown in Fig. 5.10 can generally be implemented in one of the two forms: NMOS and PMOS, as illustrated in Fig. 5.11(a) and 5.11(b), respectively [58][64]. In Fig. 5.11(a), the voltage at B1 when the clock signal is low is V_{DD} - V_{Tn1} , where V_{Tn1} is the

threshold voltage of Mdn1. When clock rises, B1 is increased to $2V_{DD}-V_{Tn1}$, causing current flow to charge up the noisy supply rail. Assuming that the supply is localized, then the boosted voltage is $2V_{DD}-V_{Tn1}-V_{Tn2}$ due to the threshold voltage degradation of Mdn2. Similarly, in Fig. 5.10(c), the boosted voltage is $2V_{DD}-|V_{Tp1}|-|V_{Tp2}|$ where $|V_{Tp1}|$ and $|V_{Tp2}|$ are the absolute threshold voltages of Mdp1 and Mdp2, respectively.

Both forms in Fig. 5.11(a) and 5.11(b) have practical limitations. In Fig. 5.11(a), when B1 reaches $2V_{DD}-V_{Tn1}$, the same voltage is applied at the gate of Mdn2. Such a high gate voltage may cause oxide breakdown in deep submicron processes, particularly 90nm and below [24]. A thick-oxide transistor should be used for Mdn2 to protect it from breakdown. In Fig. 5.11(b), when B1 is at $2V_{DD}-|V_{Tp1}|$, the drain of the transistor Mdp1 is at the same voltage, while the body of the transistor is tied to V_{DD} . This also creates a reliability concern that the pn junction of the transistor Mdp1 is forward-biased, resulting in the injection of a large amount of current back to the clean supply rail. Therefore, the PMOS implementation in Fig. 5.11(b) has to be modified. For example, if the gate voltage of Mdp1 were controlled separately using an appropriate voltage level and the bulk of Mdp1 were connected to B1, the forward biasing of the pn junction can be avoided.



Figure 5.12: Improved implementation of charge-borrowing decap with PMOS formed diodes.

The modified circuit shown in Fig. 5.12 resolves the issue of forward-biasing the transistor Mdp1. When the clock is low, B2 is set at V_{SS} , which allows B1 to be charged to V_{DD} instead of V_{DD} - $|V_{Tp1}|$. Forward-biasing the pn junction of Mdp1 here is no problem because the transistor behaves as a diode. When the clock rises, B1 is switched to $2V_{DD}$, and B2 is also raised to $2V_{DD}$. The high gate-source voltage of Mdp1 disables the current flow to back to the supply. The pn junction is now reverse-biased to prevent leakage. Note that in Fig. 5.12, the gate-body voltages of both transistors Mdp1 and Mdp2 are always within one V_{DD} of each other, ensuring gate-oxide reliability. Moreover, as a side effect, this circuit can generate a boosted voltage of $2V_{DD}$ - $|V_{Tp2}|$ instead of $2V_{DD}$ - $|V_{Tp1}|$ - $|V_{Tp2}|$, which improves the design.



Figure 5.13: Generation of the boosted voltage on node B2.

To generate the bootstrapped signal at node B2, the concept of a clock multiplier [58][65][66] can be used. The circuit for generating the bootstrapped signal at B2 is shown in Fig. 5.13. If one assumes that the clock has no previous activity, due to leakage through the transistors, it can be verified that both nodes B3 and B4 are at roughly V_{DD} while both transistors Mcn1 and Mcn2 are shut off. When the clock is turned low, B3 goes to $2V_{DD}$ and B4 is roughly at V_{DD} , which causes Mcn2 to turn on. The output node B2 is low, along with the clock signal. When the clock rises to

 V_{DD} , B3 is discharged from $2V_{DD}$ to V_{DD} , whereas B4 is charged up to $2V_{DD}$. This causes Mcn1 to turn on and Mcn2 to turn off. The rise of the clock signal also turns on Mip, allowing the output node B2 to follow B4. Because the gate of Mip is at V_{SS} , the voltage at B2 can rise to $2V_{DD}$ without any voltage loss. Similar to the previous design approach, the body of the transistor Mip is connected to B4 to ensure reverse-biased pn junctions. In Fig. 5.13, the two inverters are powered by the clean supply rail. The sizes of the PMOS-formed capacitors should be large enough to supply enough charge to the load capacitance on B2. For Mcn1 and Mcn2, thick-oxide transistors are used to reduce the risk of potential oxide breakdown because their gate-body voltages can be up to $2V_{DD}$.

5.3.2 "Clk" Signal Generation

A critical concern about the charge-borrowing decap design is the additional capacitive loading on the clock distribution system if the main clock of the chip is connected to the Clk input of the CBD. Since the CBD has a large capacitance value in the range of hundreds of picofarads, such a large capacitance loaded on the clock tree may cause an imbalance of the tree and introduce more clock skew and jitter [67]. In extreme cases, this extra loading may cause a functional failure in the clock distribution network. Therefore, the Clk input of the CBD should be generated from some other sources to keep the main clock tree unaffected.

The Clk input of the CBD simply requires a repetitive signal with enough buffer strength to drive a large capacitor. The frequency of the Clk signal should be roughly in the range of the chip's operating frequency to ensure sufficient charge pumped into the supply rail at every clock cycle. If the chip operates at a lower frequency, the extra charge provided from CBD will not harm the logic circuits connected to the local supply as long as the slew rate of the boosted voltages remains controlled within practical limitations. Another useful feature would be to implement an enable/disable function to the CBD block. When disabled, the block should behave like a regular passive decap. The CBD is turned on only when the *IR* drop exceeds certain predefined level or only during the period that the logic circuits connected to the local supply experience higher activities.



Figure 5.14: "Clk" signal generation using ring oscillator.

Satisfying the requirement above, a simple ring oscillator was selected to generate the Clk signal for the CBD, as illustrated in Fig. 5.14. A total of 39 inverter stages were used to provide an oscillation frequency of 1GHz, the upper limit of the targeted ASIC applications. A NAND gate replaces the regular inverter at the first stage to incorporate an *enable* signal. By using unit-sized inverters, the ring oscillator consumes about 65μ W of dynamic power. To provide enough current flow to charge and discharge the decoupling capacitance, in this case $2C_{decap}$, a chain of inverters was added and sized according to logical effort [7]. A fan-out factor of the inverter chain was selected to be about $3\sim4$ so that the delay through the chain was minimized. The number of stages required to generate the fan-out factor of 3 to 4 was then calculated. Of course, the circuit in Fig. 5.14 will cause additional supply noise on the "clean" supply, especially with a largely-sized last stage of the buffer chain. One has to ensure that the additional noise on the relatively clean supply node does not exceed certain noise budget when transferring charge from the clean supply to the noisy supply. As the size of the inverter chain increases, its dynamic power also increases, at a benefit of the improved slew rate (SR). This effect can be shown in Fig. 5.15. Note that in the figure the dynamic power includes the ring oscillator plus the entire buffer chain, in which each buffer is sized up properly to produce the minimum path delay.



Figure 5.15: Buffer size determines the edge delay while the buffer chain (with ring oscillator) consumes dynamic power.

The current through the last stage of the chain I_{last_stage} controls the decap charge and discharge, so it determines the slew rate. The edge delay can be defined here as the delay time for the positive side of the decap to rise a full swing, as follows:

Edge Delay =
$$\frac{1}{SR} \approx \frac{2C_{decap}}{I_{last_stage}}$$
 (5.2)

The edge delay is a better term in this scenario to illustrate the design tradeoffs in Fig. 5.15. For a targeting clock frequency of 1GHz, the corresponding clock period is 1ns. Having an edge delay in the range of 50 to 100ps (i.e., 1/20 - 1/10 of the clock period) is reasonable. Therefore, a buffer size of 300μ m/ 600μ m (NMOS/PMOS) for the last stage was selected to produce an edge delay of about 50ps, while the total dynamic power was around 3.8mW. In that case, the buffer chain was designed to have 7 stages sized up according to logical effort.

When determining the size of the buffer chain, the capacitance of $2C_{decap}$ is fixed at 0.68nF, consistent with the basic active decap design in Chapter 4. Clearly, assuming a fixed edge delay, the size of buffer required is proportional to the decap value. If a smaller decap is used, the buffer size can be made smaller to dissipate less power. This power drawn from the clean supply node is critical so that the supply noise caused by the ring oscillator and the buffer chain should not rise beyond the noise budget in its localized region. That is, the goal of generating the "Clk" signal from a clean V_{DD} is to provide charge from the clean supply that is not connected to the main system clock or any important circuitry. Designers must ensure that the clean supply itself does not become excessively noisy so that the local supply integrity is compromised.

5.3.3 Design of Charge-Borrowing Decap

With the considerations from the previous section, the complete charge-borrowing decap circuit is depicted in Fig 5.16. An *enable* signal is provided to turn off the CBD for test purposes. When the enable signal is low, the transistor Msp is off, preventing current flow from the clean supply. The decap is implemented using PMOS transistors, whose value is set to $2C_{decap}$ for comparison. When enabled, the voltage at B2 varies from 0 to $2V_{DD}$. The gate-source capacitance of Mdp1 creates certain noise on the clean supply node due to clock feedthrough. The existence of Msp provides shielding to the clock feedthrough to reduce the noise. With this circuit configuration, the practical boosted voltage is $2V_{DD}$ - $|V_{Tp2}|$, while the charge provided is at $2C_{decap}V_{DD}$. Note that the ESD concern on the thin-oxide decap in this circuit should be addressed by proper sizing of the two transistors, Mdp1 and Mdp2.



Figure 5.16: Complete circuit diagram of charge-borrowing decap.

As described earlier, after a local hot spot is identified, its nearby white space or passive decap area is occupied with an active decap or a CBD to reduce the supply noise. In the case of CBD, only the passive decap $2C_{decap}$ and the transistor Mdp2 need to be implemented locally. Other circuits showing in Fig. 5.16 can be located away from the hot spot but near a clean supply node, once such a clean supply is identified. Two global interconnects may be required to connect the two parts of circuits at node Clk and B1. The actual placement of the ring oscillator and the buffer chain will depend on the floorplan and location of power pins of the chip itself. Compared to the size of the passive decap, the size of Mdp2 is fairly small and even negligible to include in an area overhead. Since the clean supply node does not require a large area of decaps, the area

occupied by the circuits at the clean supply is relatively small. Thus, it is assumed that the CBD block requires a minimal area overhead, relative to the hot spot area where the CBD is placed.

5.3.4 Simulation Results

To validate the concept of charge-borrowing decap, HSPICE simulations were carried out. In the simulation setup, one charge-borrowing decap with enable signal is present, and no decoupling capacitor is connected between V_{DD} and V_{SS} . The load on the supply rail is a large buffer whose current demand can be controlled. As mentioned previously, the CBD can boost the supply voltage when the clock signal rises. The best scenario occurs when the current demand of the buffer also lines up with the rising edge of the clock. As a result, the dips on the supply voltage produced by the current demand and the peaks generated by the CBD cancel each other, causing a relatively low noise voltage profile. Such a case can be illustrated in Fig. 5.17.

In Fig. 5.17, the top part of the figure shows the clock signal, whereas the second part depicts the supply voltage when both passive decap and CBD are disconnected from the supply rail. The load buffer is designed to create voltage sags at the rising clock edges. In the third portion in Fig. 5.17, the load buffer is removed and only the CBD is connected and turned on. Clearly, the supply voltage is boosted at the rising edges. The dips near the falling edges can be considered as ripples since there is no decoupling capacitance connected. The last (fourth) part in the figure illustrates the voltage waveform when the CBD is turned on and the load is switching. The resulting supply voltage experiences a low level of noise because of the cancellation of peaks and dips.



Figure 5.17: Simulated V_{DD} voltage (on a 1GHz clock) with a CBD on and off (best case).

The above example can be considered as a best-case scenario because the current demand of the load and the clock rising edge are synchronized. On the other hand, the worst case would be that the current demand and the falling clock edge are lined up. The simulation result at a 1GHz external clock for the worst case is shown in Fig. 5.18. In the figure, the voltage sags created by the current demand are similar when the CBD circuit is turned on or off. The peaks created by the CBD are out of phase with the sags produced by the current demand. However, if the average V_{DD} value per clock cycle is used, similar to the previous approaches, it is clear that the CBD circuit produces a much higher average V_{DD} voltage because the voltage peaks in every clock cycle raise the average V_{DD} .



Figure 5.18: Simulated V_{DD} voltage (on a 1GHz clock) with a CBD on and off (worst case).

The simulations above are intended only to illustrate how the CBD circuit behaves under controlled conditions. Another set of simulations was used to compare the CBD with passive decap and active decap (both n=2 and n=3) for a clock frequency of 1GHz. As before, the size of the user logic buffer is changed to produce different supply noise levels k. The results are plotted in Fig. 5.19. At all k levels, the average V_{DD} voltage for the CBD is higher than the passive decap and the two active decaps. When k=0.15 for the passive decap, the average supply noise for the CBD is still at a satisfactory level of 100mV. Compared to the case of active decaps at the same k level, the average noise from the basic and extended active decaps fall close to that of the passive decap. This indicates that the CBD is more effective as a drop-in replacement than the other schemes.



Figure 5.19: Simulated average V_{DD} voltage as a function of k showing the case of CBD.

5.4 Test Chip Setup and Measurement Results

To validate this new approach to hot-spot removal, another test chip was fabricated in the same 90nm process. The degree of improvement is of interest as operation frequency increases when a basic active decap, an extended active decap, or a charge-borrowing decap is used as a drop-in replacement for a passive decap. The test chip setup is shown in Fig. 5.20. The decap circuits are individually controlled to be connected to the supply rail. In this case, the passive decap can be disconnected from the supply to observe the performance of the active decaps and the CBD by themselves.







Figure 5.21: Layout of charge-borrowing decap showing the relative size of the components.

The layout of the charge-borrowing decap for the test chip is shown in Fig. 5.21. Unlike the circuit diagram shown in Fig. 5.16, the ring oscillator and the buffer chain were not implemented, as the Clk signal was provided from the same external clock for synchronization purposes. The rest of the circuit, along with the passive decap $2C_{decap}$, was implemented on the test chip. The switching circuitry is located on the left, with the PMOS-formed decap on the right. The total layout area for the CBD is about $600\mu \text{m x} 150\mu \text{m} = 0.09\text{mm}^2$, in which the decap occupies an area of 0.083mm^2 . The switching circuitry, including the diode transistors and the bootstrapping circuit, accounts for 8% of the area. The gate-source voltage across the decap oxide is always less than or equal to one V_{DD} so that thin-oxide devices are acceptable in this case. Thus, thin-oxide PMOS transistors were used to implement the decap in the CBD.



Figure 5.22: (a) Annotated microphotograph and (b) layout of the second test chip.

The microphotograph and the layout of the second test chip are shown in Fig. 5.22. In the figure, the decap circuits are placed about 600µm away from the user logic, in which a large buffer with a large capacitive load was used to create supply noise, with the input controlled by an external signal. Similar to the first test chip, the size of the decaps was chosen to be only a few times

larger than the capacitive load to create a ~100mV voltage drop for the experiments. Note that the load cannot be dynamically changed to produce a variable k level. The clock frequency can be controlled externally by providing an input from a bit-rate analyzer. The probed pad of the supply node can be connected to an oscilloscope to measure the voltage waveforms. The test chip area is $1.1 \ge 0.86 = 0.95 \text{ mm}^2$ in total.

A collection of 9 sample chips was tested. The clock frequency was fixed at 1GHz for this test. The improvement across the sample space is shown in Fig. 5.23. The sample chips that are in the slow or normal process corners cannot be distinguished easily. However, the two sample chips that are in the fast process corner can be identified by correlating with simulation. As mentioned in Chapter 4, the average supply voltage varies significantly mainly due to the random process variation on R_{mesh} and L_{pack} , which determines the *IR* drop on the supply rails. Note that the supply noise reduction improvement by using the CBD is rather consistent under process variations across dies, which indicates the robustness of the design. The sample chip that provides the best improvement was used for further analysis.



Figure 5.23: Scatter plot comparing average V_{DD} for the tested sample chips.

The waveforms of a test chip with a 1GHz clock are depicted in Fig. 5.24. The dark gray curve is when the CBD is on, and the light gray (red in color) curve is when the passive decap is on. The two curves are superimposed. As expected, the supply voltage returns to high when extra charge is fed through from the clean supply on the rising edge of the clock, improving the average V_{DD} level per cycle.



Figure 5.24: Superimposed waveforms showing a CBD and a passive decap on a 1GHz clock. In the top panel of the figure, the upper trace is when the CBD is on, while the lower trace is when the passive decap is on.

It is also useful to obtain the CBD performance over the operating frequency range. Fig. 5.25 shows the impact of changing the external input frequency from 100MHz to 1.5GHz. Two solid trend lines are provided corresponding to the CBD and the passive decap. The gap between the two trend lines widens, indicating that the benefit of using the CBD increases as frequency increases. Even at 1.5GHz, the average V_{DD} level for the CBD is significantly higher than the passive decap, suggesting that the CBD is suitable for today's high-speed ASICs and medium-speed custom designs. At 1.5GHz, the test chip validated that the CBD has a maximum

improvement of 93 mV (or about 53% less noise) over the passive decap, 74 mV (48% less noise) over the basic active decap with n=2, or 46 mV (36% less noise) over the extended active decap with n=3. From 100MHz to 1.5GHz, compared to the passive decap, using the CBD reduces the supply noise from 42% to 55%. Note that the CBD outperforms both the basic and the extended active decaps across the operating frequency range.



Figure 5.25: Measured average V_{DD} voltages at different clock frequencies.

Another important observation is that, unlike active decaps, the CBD does not seem to have a specific frequency, at which the average V_{DD} voltages from the CBD and the passive decap crossover. In other words, there is always a gap in the average supply voltage between the CBD and the passive decap across the frequency range. This makes intuitive sense since the CBD

boosts the supply voltage at every clock cycle no matter the clock frequency. Although the noise increases as the frequency increases, the amount of charge provided by the CBD circuit remains roughly the same at every cycle (assuming that it is running at the clock frequency).

From the test results of the sample chips, it can be concluded that the charge-borrowing decap is capable of reducing the supply noise more efficiently than the passive and active decaps. In addition, there are other attractive features of the CBD, such as higher bandwidth, simplicity, and robustness. Moreover, the use of active decaps will increase the power consumption of the chip because of the internal switching circuitry. In the case of CBD, the local power overhead is fairly small since there is only very limited leakage current and most charge transferred from the clean supply to the noisy supply is not wasted. Specifically, the leakage power of the CBD is about 4μ W, about 0.1% of the total power consumption. However, the dynamic power of the Clk generation circuit is comparable to the active decap. Therefore, the equivalent power overhead with better performance makes the CBD circuit an appealing alternative to the active or passive decaps.

Although the charge-borrowing decap provides many advantages, it has a number of limitations. One important issue is that the supply voltage change is abrupt once the CBD is turned on. Highfrequency glitches on the supply may affect the logic circuit powered by it. To smooth out the supply voltages, large amount of passive decaps should be present in its vicinity. Although parasitic decoupling capacitance is inherently present on the supply, more decaps may be still needed. As a result, the existing area may not be completely replaceable by the CBD, as certain portion of the area should be reserved for the passive decaps. For a fixed area, the proportion of the CBD and the passive decap depends on the sensitivity of the local logic circuit to the supply glitches. Questions like the maximum allowed distance from the decap to the CBD and the minimum required size of the decap remain unanswered at this stage. However, since the local circuitry of the CBD is relatively simple (passive decap plus a diode-connected transistor), it is still an attractive alternative to the active decaps.

5.5 Summary

This chapter further extended the concept of active decap by increasing the stack height n to find an optimal value, depending on the supply noise k level presented at the local supply. It was found that the extended active decap with n=3 provided superior performance by delivering a higher average supply voltage than the basic active decap with n=2 when k is <14%. This chapter also introduced a novel design of charge-borrowing decap to provide better supply noise reduction than the basic and the extended active decaps. The charge-borrowing decap delivers more charge and an increased supply boost for a wide range of operating frequencies. Its relatively simplistic design and implementation ensures its robustness.

Chapter 6

Conclusions and Future Work

6.1 Summary and Conclusions

As technology scales further into the deep submicron regime, increasing clock frequency and decreasing supply voltage makes maintaining the quality of power supply a critical issue. Onchip power supply noise, due to *IR* drop and *Ldi/dt* effects, has a great impact on delay variation, and may even cause improper functionality. Power supply noise can be reduced by placing decoupling capacitors close to power pads and large drivers throughout the power distribution system. Decaps provide locally "instantaneous" current to the switching drivers and keep the power supply within certain noise budgets. Traditionally, a decap is made from an NMOS transistor outside the standard-cell blocks, or a pair of NMOS and PMOS transistors within the blocks. However, starting from 90nm technology, the reduction in oxide thickness of MOS transistors causes an increased ESD risk and more gate leakage. Standard decap designs, therefore, may no longer be appropriate for 90nm and beyond.

In this dissertation, the goal was to provide practical solutions to active and passive decap designs targeting ASIC applications in both white-space and standard-cell areas. The dissertation began with an overview of decap design basics, gate leakage phenomenon, ESD concerns, and standard-cell decap layout and placement. Some essential decap design issues were highlighted through the background section to motivate the topics for the rest of the dissertation. More importantly, the metric for power supply noise management was proposed and validated for decap performance comparisons used throughout the dissertation.

Next, the tradeoffs between high-frequency performance of decaps and ESD protection were investigated, and their impacts on the layout of standard-cell passive decaps were discussed. A design metric was introduced to determine the optimal number of fingers to use in the standard-cell layout to obtain a desired capacitance level over a target operating frequency. Models were developed to capture the frequency responses of R_{eff} and C_{eff} for any given technology with only a few parameters. For ESD protection, a cross-coupled design that had been previously proposed by cell library developers was shown to suffer from reduced frequency response and provided no savings in gate leakage. It was shown that more fingers than typically used were needed to provide the target resistance value for sufficient ESD protection. The layout with the smallest NMOS device and a multi-fingered PMOS device was described to deliver acceptable frequency response and ESD reliability, while providing the lowest leakage.

For white-space areas, the effectiveness of active decaps as a late-stage drop-in replacement for passive decaps was evaluated so that a completed chip layout need not be disrupted near the tapeout deadline. Improvements to the design of an active decoupling capacitor were described for removal of hot-spot *IR*-drop violations in ASIC designs running at up to 1GHz. The modified active decap using latch-based comparators in 90nm CMOS is able to switch in 0.5ns and consumes a relatively low power of 2.8mW, which is about 5X lower than a previous design

running at approximately the same speed. This reduced power makes it more suitable for use in ASIC designs. Measurement results from test chips indicate improvement over passive decaps of 10% - 20%, operating from 200MHz to 1GHz. The optimal active decap size to maximally remove hot-spot *IR* drop was identified. Placement analysis was also carried out and it was found that the active decap is most effective when placed in close proximity to the hot spot, as compared to the passive decap which is not as sensitive to the exact location. If sized and placed properly, active decaps can be up to 20% better when used as drop-in replacements of passive decaps for power supply noise reduction.

Further research on active decap design was explored. By increasing the stack height n to an optimal value, and depending on the supply noise k level presented at the local supply, a maximum supply boost can be achieved. It was found that the extended active decap with n=3 provided a superior performance by delivering a higher average supply voltage than the n=2 and n=4 cases when the supply noise k is in the range of 7-14%. When k is above 14%, n=2 must be used and beyond 16%, the area for the drop-in replacement of active decaps must be expanded to produce satisfactory improvement over the passive decap.

Finally, the novel design for charge-borrowing decap was proposed. This design provides better supply noise reduction than all other forms of active decaps. The charge-borrowing decap efficiently transfers charge from a clean supply rail to eliminate the hot spots on relatively noisy supply nodes. The CBD only requires a minimum power overhead and delivers a maximum supply boost for a wide range of operating frequencies. Test results indicate that the CBD outperforms both the basic and the extended active decaps by reducing the supply noise to a
lower level. The design and implementation of the CBD was kept relatively simple so that the robustness of the design can be maintained.

6.2 Contributions in this Dissertation

The following summarizes the major contributions in this dissertation:

- Developed standard-cell passive decap designs that properly trade off gate leakage, ESD reliability, and transient response. Provided simple and yet practical decap design metrics and guidelines for 90nm and 65nm CMOS technologies.
- Designed and implemented a white-space active decap using latch-based comparators that provides adequate supply noise reduction while consuming relatively low static power. Validated the design through a test chip. Explored the placement issues of the active decap.
- Extended the concept of active decap for an optimal design that produced the highest supply boost for the maximum supply noise reduction in a local area. Proposed a simple but novel charge-borrowing decap circuit that outperforms the basic and the extended active decaps. Validated the design through another test chip.

6.3 Future Work

The limitations on the charge-borrowing decap design needs to be evaluated further. Quantitative answers to questions like the optimal distance between the CBD and the logic, the accurate proportion between the CBD area and the nearby passive decap area, and the maximum allowable charge transferring from the clean supply node while still maintaining its supply integrity, should be investigated. A test chip with real industrial blocks that are placed as the user logic circuit is desired for the most accurate decap performance evaluation.

Another issue is the scalability of the active decaps and the CBD. The design and implementation of the active decap were only accomplished and validated in 90nm CMOS. It is desirable to make the active decap designs useful in future technologies, such as 65nm and 45nm CMOS. As technology scales, more design challenges will occur. If any part of the design needs to be modified to accommodate the more advanced technology, it should be investigated in future research.

Monitoring power supply fluctuations on-chip in real-time is also an emerging area of research [20]. The measured results of real-time power supply noise from a monitoring circuitry can be used as a validation of decap design and placement. Many techniques have been used to monitor power supply noise on-chip [68][69]. These techniques are not suitable for production environments as they either need a significant area or require complex data processing off-chip. To overcome these limitations, a simple monitoring based on an under-sampling technique [70] is worthy of being investigated. Under-sampling is used to capture a high-frequency periodic signal from a large number of cycles using a slower sampling signal to achieve an effective highspeed sampling rate. In the case of power supply noise, the dynamic voltage drop is not periodic in nature, but the same experiment could be repeated several times, and each time skew the sampling point by a small time shift, Δt , which represents the sampling period resulting in an equivalent sampling frequency of $f_{under-sampling} = 1/\Delta t$ [70]. Using this technique, the measurements may be repeated to average and cancel out the noise effect. This approach needs to be evaluated from concept to test chip in order to finally validate its advantages and disadvantages and serves as a promising area of future work.

APPENDIX: COMPARATOR DESIGN FUNDAMENTALS

Comparators are one of the most widely used components in analog integrated circuits. A voltage comparator is a circuit that compares the instantaneous value of an input signal with a reference signal and produces an output at logic level, depending on whether the input is greater or smaller than the reference level [57]. One important application for high-speed voltage comparators is in data converters, where the conversion speed is limited by the response time of the comparators [57]. Other issues related to comparator design include finite resolution, offset, power, and area. As technology scales, more advanced CMOS technologies allow comparators to be realized for higher speed and potentially smaller area and power. However, it is difficult to achieve high speed and high accuracy at the same time because of the existence of device mismatches [71].

A widely used comparator configuration is a high-gain differential input, single-ended output amplifier, whose symbol is shown in Fig. A.1. The output of the comparator V_{out} should have a large swing, ideally from V_{DD} to V_{SS} , as the input varies across a small swing, typically in the millivolt range [57]. In many applications, a comparator is used in open-loop operations, such that no frequency compensation is required [57][58][59]. However, in certain cases, due to the nature of AC coupling of the output and the input, a comparator may need frequency compensation to avoid oscillations [32].



Figure A.1: A differential input, single-ended output comparator symbol.



Figure A.2: DC transfer characteristics of (a) an ideal comparator and (b) a practical comparator with finite gain and offset voltage.

The DC characteristic curve of an ideal differential comparator is shown in Fig. A.2(a). When the positive input V_{in+} is greater than the negative input V_{in-} , the output is high (i.e., at V_{DD}). When V_{in+} is less than V_{in-} , the output is low (i.e., at V_{SS}). This ideal DC transfer curve corresponds to a differential gain of infinity. That is, an infinitely small polarity change in ($V_{in+} - V_{in-}$) will cause the output to switch. A more realistic DC transfer curve of a comparator is depicted in Fig. A.2(b). In practice, the differential gain is finite and equal to A_V . In Fig. A.2(b), the two voltages V_{IL} and V_{IH} are the *overdrive* voltages (also called the input excess voltages). The overdrive is the input level that drives the comparator from an initial saturated input condition to an input level that barely causes the output of the comparator to switch its level [57]. Another non-ideal effect of the comparator is the input referred DC offset voltage that is mainly caused by device mismatches. If no offset voltage is present, the comparator DC transfer curve will be symmetrical around the point where $V_{in+} = V_{in-}$. However, for a finite offset voltage of V_{OS} , the comparator output will switch at ($V_{in+} - V_{in-} + V_{OS}$). In general, the output voltage V_{out} of a comparator can be written as [57][58]:

$$V_{out} = \begin{cases} V_{DD} & if \quad \Delta V_{in} \ge V_{IH} \\ A_{V} \Delta V_{in} & if \quad V_{IL} \le \Delta V_{in} < V_{IH} \\ V_{SS} & if \quad \Delta V_{in} < V_{IL} \end{cases}$$
(A.1)

where $\Delta V_{in} = V_{in+} - V_{in-}$. Clearly, the finite gain and offset voltage affect the accuracy of the comparator. It is desired that a comparator is designed to have a high gain and a small offset voltage.

Three major challenges exist in any comparator design: high speed, high resolution, and low power [72]. High speed is achieved by having a fast response time. That is, following an input polarity change, the comparator should switch its output between V_{DD} and V_{SS} with fast rise and fall times. In order to achieve the highest comparison speed, the minimum channel length for a specific technology is often used in comparator designs [71]. In addition, low power consumption is always desired. As technology scales, due to V_{DD} scaling, the dynamic power of a comparator will scale accordingly. Meanwhile, the static power consumption of the comparator may increase due to larger leakage current in a more advanced deep-submicron technology. Overall, the total power consumption including both dynamic and static power may reduce as technology shrinks [72].

A high gain A_V is essential to achieve high resolution in a comparator design. For example, the input of the comparator needs to resolve 1mV of input variation, which requires the output to switch a full swing of 1V at V_{DD} , the voltage gain A_V is therefore 1V/1mV = 1000. It is difficult to achieve such a high gain within one stage of amplification. Hence, a multi-stage amplifier or a regenerative latch using positive feedback may be used as a comparator to achieve the high gain

requirement. A latch is normally faster then a multi-stage amplifier achieving the same gain [60][71][73]. Therefore, latch-based comparators are often used in practice.



Figure A.3: Circuit schematics showing positive feedback: (a) a basic differential pair with diodeconnected load, (b) a gain-enhanced differential pair, and (c) a differential pair using positive feedback to provide increased gain.

The concept of positive feedback used in the latch approach can be explained in the following example, as shown in Fig A.3. In Fig. A.3(a), a differential pair is loaded with two diode-connected PMOS devices. Its small-signal gain $A_{V(a)}$ can be given by [64][73]:

$$A_{V(a)} \approx \frac{g_{m1}}{g_{m3}} = \sqrt{\frac{\mu_n (W/L)_{M1}}{\mu_p (W/L)_{M3}}}$$
(A.2)

where g_{m1} and g_{m3} are the transconductance of M1 and M3, and μ_n and μ_p are the channel mobility of NMOS and PMOS devices, respectively. In Fig. A.3(b), a gain enhancement approach is added to the circuit with two additional transistors, M5 and M6. The small-signal gain of A_{V(b)} can be given by [54][57]:

$$A_{V(b)} \approx A_{V(a)} \cdot \sqrt{\frac{1}{1 - (2I_5 / I_{b1})}}$$
 (A.3)

where I_5 and I_{b1} are the current flow through M5 and Mb1, respectively. By properly choosing the size of M5, the small-signal gain can be improved. In Fig. A.3(c), the drain terminals of M5 and M6 are cross connected, creating a form of positive feedback. The positive feedback functions as follows: when V_{in+} is slightly larger than V_{in-} , it causes V_{out1} to be slightly smaller than V_{out2} . Larger V_{out2} forces M6 to deliver less current to the node V_{out1} . Similarly, smaller V_{out1} forces M5 to deliver more current to charge up V_{out2} . This positive loop reinforces V_{out2} to reach V_{DD} and V_{out1} to reach V_{SS} [54][57]. The small-signal gain of $A_{V(c)}$ can be given by [57]:

$$A_{V(c)} \approx A_{V(a)} \cdot \frac{1}{1 - \frac{(W/L)_{M5}}{(W/L)_{M3}}}$$
 (A.4)

Equation A.4 requires that the size ratio of $(W/L)_{M5} / (W/L)_{M3}$ is less than 1. When the size ratio is greater than unity, the small-signal gain will become infinity and the circuit will operate as a regenerative latch [57].

REFERENCES

- H. H. Chen and S. E. Schuster, "On-chip decoupling capacitor optimization for highperformance VLSI design," Symposium on VLSI Technology, Systems, and Applications, pp. 99-103, May-Jun. 1995.
- [2] S. Pant and E. Chiprout, "Power grid physics and implications for CAD," *IEEE/ACM Design Automation Conference*, pp. 199-204, Jul. 2006.
- [3] N. Srivastava, X. Qi, and K. Banerjee, "Impact of on-chip inductance on power distribution network design for nanometer scale integrated circuits," *International Symposium on Quality of Electronic Design (ISQED)*, pp. 346-351, Mar. 2005.
- [4] C. W. Fok and D. L. Pulfrey, "Full-chip power-supply noise: the effect of on-chip powerrail inductance," *International Journal of High Speed Electronics and Systems*, vol. 12, no. 2, pp. 573-582, Jun. 2002.
- [5] J. Kim, B. Choi, H. Kim, W. Ryu, Y. -H. Yun, S. -H. Hamm, S. -H. Kim, and Y. -H. Lee, "Separated role of on-chip and on-PCB decoupling capacitors for reduction of radiated emission on printed circuit board," *IEEE International Symposium on Electromagnetic Compatibility*, pp. 531-536, Aug. 2001.
- [6] H. H. Chen, J. S. Neely, M. F. Wang, and G. Co, "On-chip decoupling capacitor optimization for noise and leakage reduction," *Symposium on Integrated Circuits and Systems Design*, pp. 319-326, Sep. 2003.
- [7] D. A. Hodges, H. G. Jackson, and R. A. Saleh, *Analysis and Design of Digital Integrated Circuits in Deep Submicron Technology*, 3rd ed., New York: McGraw-Hill, 2004.
- [8] N. Na, T. Budell, C. Chiu, E. Tremble, and I. Wemple, "The effects of on-chip and package decoupling capacitors and an efficient ASIC decoupling methodology," *Electronic Components and Technology Conference (ECTC)*, pp. 556-567, Jun. 2004.
- [9] H. Su, S. S. Sapatnekar, and S. R. Nassif, "Optimal decoupling capacitor sizing and placement for standard-cell layout designs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 4, pp. 428-436, Apr. 2003.
- [10] M. Popovich and E. G. Friedman, "Decoupling capacitors for multi-voltage power distribution systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 14, no. 3, pp. 217-228, Mar. 2006.
- [11] M. Popovich, E. G. Friedman, M. Sotman, A. Kolodny, and R. M. Secareanu, "Maximum effective distance of on-chip decoupling capacitors in power distribution grids," *IEEE/ACM Great Lakes Symposium on VLSI*, pp. 173-179, May 2006.
- [12] P. Larsson, "Parasitic resistance in an MOS transistor used as on-chip decoupling capacitance," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 4, pp. 574-576, Apr. 1997.

- [13] P. Larsson, "Resonance and damping in CMOS circuits with on-chip decoupling capacitance," *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 45, no. 8, pp. 849–858, Aug. 1998.
- [14] M. D. Powell and T. N. Vijaykumar, "Exploiting resonant behavior to reduce inductive noise," *International Symposium on Computer Architecture*, pp. 288–299, Jun. 2004.
- [15] C. H. Ng, C. S. Ho, S. -F. S. Chu, and S. -C. Sun, "MIM capacitor integration for mixedsignal/RF applications," *IEEE Transactions on Electron Devices*, vol. 52, no. 7, pp. 1399-1409, Jul. 2005.
- [16] M. W. C. Goh, Q. Lim, R. A. Keating, A. V. Kordesch, and Y. Bin Mohd Yusof, "Design of radio frequency metal-insulator-metal (MIM) capacitors," *International Conference on Solid-State and Integrated Circuits Technology*, pp. 209-212, Oct. 2004.
- [17] C. H. Ng, C. S. Ho, N. G. Toledo, and S. -F. Chu, "Characterization and comparison of single and stacked MIMC in copper interconnect process for mixed-mode and RF applications," *IEEE Electron Device Letters*, vol. 25, no. 7, pp. 489-491, Jul. 2004.
- [18] C. H. Ng, C. S. Ho, S. -F. S. Chu, and S. -C. Sun, "MIM capacitor integration for mixedsignal/RF applications," *IEEE Transactions on Electron Devices*, vol. 52, no. 7, pp. 1399-1409, Jul. 2005.
- [19] H. Yamamoto and J. A. Davis, "Decreased effectiveness of on-chip decoupling capacitance in high-frequency operation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 6, pp. 649-659, Jun. 2007.
- [20] K. Arabi, R. Saleh, and X. Meng, "Power supply noise in SoCs: metrics, management, and measurement," *IEEE Design and Test of Computers*, vol. 24, no. 3, pp. 236-244, May-Jun. 2007.
- [21] TSMC 90nm CLN90G Process SAGE-X v3.0 Standard Cell Library Databook, Release 1.0, Artisan Components Inc., Sunnyvale, CA, 2004.
- [22] X. Meng, R. Saleh, and K. Arabi, "Layout of decoupling capacitors in IP blocks for 90-nm CMOS," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 11, pp. 1581-1588, Nov. 2008.
- [23] X. Meng, K. Arabi, and R. Saleh, "Novel decoupling capacitor designs for sub-90nm CMOS technology," *International Symposium on Quality Electronic Design (ISQED)*, pp. 266-272, Mar. 2006.
- [24] A. Amerasekera and C. Duvvury, *ESD in Silicon Integrated Circuits*, 2nd ed., Hoboken, NY: John Wiley & Sons, 2002.
- [25] J. Fu, Z. Luo, X. Hong, T. Cai, S. X. -D. Tan, and Z. Pan, "VLSI on-chip power/ground network optimization considering decap leakage currents," Asia and South Pacific Design Automation Conference, vol. 2, pp. 735-738, Jan. 2005.

- [26] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proceedings of IEEE*, vol. 91, no. 2, pp. 305-327, Feb. 2003.
- [27] F. Hamzaoglu and M. Stan, "Circuit-level techniques to control gate leakage for sub-100nm CMOS," International Symposium on Low Power Electronics and Design, pp. 60– 63, Aug. 2002.
- [28] R. S. Guindi and F. N. Najm, "Design techniques for gate-leakage reduction in CMOS circuits," *International Symposium on Quality Electronic Design (ISQED)*, pp. 61-65, Mar. 2003.
- [29] D. Lee, W. Kwong, D. Blaauw, and D. Sylvester, "Analysis and minimization techniques for total leakage considering gate oxide leakage," *IEEE/ACM Design Automation Conference*, pp 175-180, Jun. 2003.
- [30] L. Chang, K. J. Yang, Y. -C. Yeo, Y. -K. Choi, T. -J. King, and C. Hu, "Reduction of direct-tunneling gate leakage current in double-gate and ultra-thin body MOSFETs," *IEEE Transactions on Electron Devices*, vol. 49, no. 12, pp. 2288-2295, Dec. 2002.
- [31] X. Meng, K. Arabi, and R. Saleh, "A novel active decoupling capacitor design in 90nm CMOS," *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 657-660, May 2007. (Top 10 Honorable Mention Award)
- [32] X. Meng and R. Saleh, "An improved active decoupling capacitor for "hot-spot" supply noise reduction in ASIC designs," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 584-593, Feb. 2009.
- [33] R. Saleh, D. Overhauser, S. Taylor, "Full-chip verification of UDSM designs," *IEEE/ACM International Conference on Computer-Aided Design*, pp. 453-460, Nov. 1998.
- [34] S. Sapatnekar, "High-performance power grids for nanometer technologies," *IEEE International Conference on VLSI Design*, pp. 839-844, Jan. 2004.
- [35] G. Bai, S. Bobba, and I. N. Hajj, "Static timing analysis including power supply noise effect on propagation delay in VLSI circuit," *IEEE/ACM Design Automation Conference*, pp. 295-300, Jun. 2001.
- [36] H. Harizi, R. HauBler, M. Olbrich, and E. Barke, "Efficient modeling techniques for dynamic voltage drop analysis," *IEEE/ACM Design Automation Conference*, pp. 706-711, Jun. 2007.
- [37] M. Ang, R. Salem, and A. Taylor, "An on-chip voltage regulator using switched decoupling capacitors," *IEEE International Solid-State Circuits Conference*, pp. 438-439, Feb. 2000.
- [38] M. A. Ang, and A. D. Taylor, "Voltage regulating circuit for attenuating inductanceinduced on-chip supply variations," U.S. Patent 6509785, Jan. 21, 2003.

- [39] C. Giacomotto, R. P. Masleid, and A. Harada, "Four-state switched decoupling capacitor system for active power stabilizer," U.S. Patent 6744242 B1, Jun. 1, 2004.
- [40] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits: A Design Perspective*, 2nd ed., Upper Saddle River, NJ: Prentice Hall, 2004.
- [41] J. Gu, H. Eom, and C. Kim, "A switched decoupling capacitor circuit for on-chip supply resonance damping", *Symposium on VLSI Circuits*, pp. 126-127, Jun. 2007.
- [42] J. Gu, R. Harjani, and C. Kim, "Distributed active decoupling capacitors for on-chip supply noise cancellation in digital VLSI circuits", *Symposium on VLSI Circuits*, pp. 216-217, Jun. 2006.
- [43] W. C. Lee and C. Hu, "Modeling gate and substrate currents due to conduction- and valence-band electron and hole tunneling," *Symposium on VLSI Technology*, pp. 198-199, Jun. 2000.
- [44] K. Cao, W. -C. Lee, W. Liu, X. Jin, P. Su, S. K. H. Fung, J. X. An, B. Yu, and C. Hu, "BSIM4 gate leakage model including source drain partition," *International Electron Devices Meeting (IEDM)*, pp. 815-818, Dec. 2000.
- [45] X. Xi, M. Dunga, J. He, W. Liu, K. M. Cao, X. Jin, J. J. Ou, M. Chan, A. M. Niknejad, and C. Hu, "BSIM4.4.0 MOSFET model user's manual," University of California, Berkeley, 2004.
- [46] C. K. Alexander and M. N. O. Sadiku, *Fundamentals of Electric Circuits*, New York: McGraw-Hill, 2000.
- [47] X. W. Wang, Y. Shi, T. P. Ma, G. J. Cui, T. Tamagawa, J. W. Golz, B. L. Halpen, and J. J. Schmitt, "Extending gate dielectric scaling limit by use of nitride or oxynitride," Symposium on VLSI Technology, pp. 109-110, Jun. 1995.
- [48] T. P. Ma, "Opportunities and challenges for high-k gate dielectrics," International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), pp. 1-4, Jul. 2004.
- [49] T. P. Ma, "Electrical characterization of high-k gate dielectrics," International Conference on Solid-State and Integrated Circuits Technology, pp. 361-365, Oct. 2004.
- [50] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers, V. Naydenov, T. Khondker, S. Sarkar, and P. Singh, "Penryn: 45-nm next generation Intel[®] Core[™] 2 processor," *IEEE Asian Solid-State Circuits Conference (ASSCC)*, pp. 14-17, Nov. 2007.
- [51] M. T. Bohr, R. S. Chau, T. Ghani, and K. Mistry, "The high-k solution," *IEEE Spectrum*, vol. 40, no. 10, pp. 29-35, Oct. 2007.
- [52] J. Chia, "Design, layout and placement of on-chip decoupling capacitors in IP blocks," *M.A.Sc Thesis*, University of British Columbia, 2004.

- [53] S. Zhao, K. Roy and C. -K. Koh, "Decoupling capacitance allocation and its application to power-supply noise-aware floorplanning," *IEEE Transactions on Computer-Aided Design* of Integrated Circuits and Systems, vol. 21, no. 1, pp 81-92, Jan. 2002.
- [54] J. R. Hauser, "Bias sweep rate effects on quasi-static capacitance of MOS capacitors," *IEEE Transactions on Electron Devices*, vol. 44, no. 6, pp. 1009-1012, Jun. 1997.
- [55] W. Liu, MOSFET Models for SPICE Simulation Including BSIM3v3 and BSIM4, Wiley-IEEE Press, 2001.
- [56] H. Johnson and M. Graham, High-Speed Digital Design, Prentice-Hall, 1993.
- [57] R. Gregorian, Introduction to CMOS Op-Amps and Comparators, New York: John Wiley & Sons, 1999.
- [58] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*, 2nd ed., Piscataway, NJ: IEEE Press, 2005.
- [59] D. A. Jones and K. Martin, Analog Integrated Circuit Design, New York: John Wiley & Sons, 1997.
- [60] S. Sheikhaei, S. Mirabbasi, and A. Ivanov, "A 43mW single-channel 4GS/s 4-bit flash ADC in 0.18µm CMOS," *IEEE Custom Integrated Circuits Conference (CICC)*, pp. 353-356, Sep. 2007.
- [61] E. Alon and M. Horowitz, "Integrated regulation for energy-efficient digital circuits," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 8, pp. 1795-1807, Aug. 2008.
- [62] X. Meng and R. Saleh, "Active decap design considerations for optimal supply noise reduction," *International Symposium on Quality Electronic Design (ISQED)*, pp. 765-769, Mar. 2009.
- [63] X. Meng, R. Saleh, and S. Wilton, "Charge-borrowing decap: a novel circuit for removal of local supply noise violations," accepted to *IEEE Custom Integrated Circuits Conference* (CICC), Sep. 2009.
- [64] B. Razavi, Design of Analog CMOS Integrated Circuits, New York: McGraw Hill, 2001.
- [65] T. B. Cho and P. R. Gray, "A 10b, 20Msamples/s, 35mW pipeline A/D converter," IEEE Journal of Solid-State Circuits, vol. 30, no. 3, pp. 166-172, Mar. 1995.
- [66] A. M Abo and P. R. Gray, "A 1.5-V, 10-bits, 14.3-MS/s CMOS pipeline analog-to-digital converter," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 5, pp. 599-606, May 1999.
- [67] K. A. Jenkins, K. L. Shepard, and Z. Xu, "On-chip circuit for measuring period jitter and skew of clock distribution networks," *IEEE Custom Integrated Circuits Conference* (*CICC*), pp. 157-160, Sep. 2007.

- [68] E. Alon, V. Stojanovic, and M. A. Horowitz, "Circuits and techniques for high-resolution measurement of on-chip power supply noise," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 4, pp. 820-828, Apr. 2005.
- [69] T. Nakura, M. Ikeda, and K. Asada, "Design and measurement of on-chip di/dt detector circuit for power supply line," *IEEE Asia-Pacific Conference on Advanced System Integrated Circuits*, pp. 426-427, Aug. 2004.
- [70] B. Kaminska and K. Arabi, "Mixed signal DFT: a concise overview," *IEEE International Conference on Computer-Aided Design*, pp. 672-680, Nov. 2003.
- [71] B. Murmann, "A/D converter trends: power dissipation, scaling and digitally assisted architectures," *IEEE Custom Integrated Circuits Conference (CICC)*, pp. 105-112, Sep. 2008.
- [72] R. J. van de Plassche, J. H. Huijsing, and W. Sansen, Analog Circuit Design: High-Speed Analog-to-Digital Converters; Mixed-Signal Design; PLL's and Synthesizers, Boston: Kluwer Academic Publishers, 2000.
- [73] M. Gustavsson, J. J. Wikner, and N. Tan, CMOS Data Converters for Communications, Boston: Kluwer Academic Publishers, 2000.