# Adaptive Techniques for Dynamic Processor Optimization

## Theory and Practice

## Series on Integrated Circuits and Systems

Series Editor: Anantha Chandrakasan
Massachusetts Institute of Technology
Cambridge, Massachusetts

---

Adaptive Techniques for Dynamic Processor Optimization: Theory and Practice
Alice Wang and Samuel Naffziger (Eds.)
ISBN 978-0-387-76471-9

mm-Wave Silicon Technology: 60 GHz and Beyond
Ali M. Niknejad and Hossein Hashemi (Eds.)
ISBN 978-0-387-76558-7

Ultra Wideband: Circuits, Transceivers, and Systems
Ranjit Gharpurey and Peter Kinget (Eds.)
ISBN 978-0-387-37238-9

Creating Assertion-Based IP
Harry D. Foster and Adam C. Krolnik
ISBN 978-0-387-36641-8

Design for Manufacturability and Statistical Design: A Constructive Approach
Michael Orshansky, Sani R. Nassif, and Duane Boning
ISBN 978-0-387-30928-6

Low Power Methodology Manual: For System-on-Chip Design
Michael Keating, David Flynn, Rob Aitken, Alan Gibbons, and Kaijian Shi
ISBN 978-0-387-71818-7

Modern Circuit Placement: Best Practices and Results
Gi-Joon Nam and Jason Cong
ISBN 978-0-387-36837-5

CMOS Biotechnology
Hakho Lee, Donhee Ham and Robert M. Westervelt
ISBN 978-0-387-36836-8

SAT-Based Scalable Formal Verification Solutions
Malay Ganai and Aarti Gupta
ISBN 978-0-387-69166-4, 2007

Ultra-Low Voltage Nano-Scale Memories
Kiyoo Itoh, Masashi Horiguchi and Hitoshi Tanaka
ISBN 978-0-387-33398-4, 2007

Routing Congestion in VLSI Circuits: Estimation and Optimization
Prashant Saxena, Rupesh S. Shelar, Sachin Sapatnekar
ISBN 978-0-387-30037-5, 2007

Ultra-Low Power Wireless Technologies for Sensor Networks
Brian Otis and Jan Rabaey
ISBN 978-0-387-30930-9, 2007

Sub-Threshold Design for Ultra Low-Power Systems
Alice Wang, Benton H. Calhoun and Anantha Chandrakasan
ISBN 978-0-387-33515-5, 2006

High Performance Energy Efficient Microprocessor Design
Vojin Oklibdzija and Ram Krishnamurthy (Eds.)
ISBN 978-0-387-28594-8, 2006

*Continued after index*

Alice Wang • Samuel Naffziger
Editors

# Adaptive Techniques for Dynamic Processor Optimization

## Theory and Practice

*Editors*
Alice Wang
Texas Instruments, Inc.
Dallas, TX
USA
aliwang@ti.com

Samuel Naffziger
Advanced Micro Devices
Fort Collins, CO
USA
samuel.naffziger@amd.com

*Series Editor*
Anantha Chandrakasan
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139
USA

ISBN: 978-0-387-76471-9 e-ISBN: 978-0-387-76472-6
DOI: 10.1007/978-0-387-76472-6

Library of Congress Control Number: 2007943527

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.
While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

# Preface

The integrated circuit has evolved tremendously in recent years as Moore's Law has enabled exponentially more devices and functionality to be packed onto a single piece of silicon. In some ways however, these highly integrated circuits, of which microprocessors are the flagship example, have become victims of their own success. Despite dramatic reductions in the switching energy of the transistors, these reductions have kept pace neither with the increased integration levels nor with the higher switching frequencies. In addition, the atomic dimensions being utilized by these highly integrated processors have given rise to much higher levels of random and systematic variation which undercut the gains from process scaling that would otherwise be realized. So these factors—the increasing impact of variation and the struggle to control power consumption—have given rise to a tremendous amount of innovation in the area of adaptive techniques for dynamic processor optimization.

The fundamental premise behind adaptive processor design is the recognition that variations in manufacturing and environment cause a statically configured operating point to be far too inefficient. Inefficient designs waste power and performance and will quickly be surpassed by more adaptive designs, just as it happens in the biological realm. Organisms must adapt to survive, and a similar trend is seen with processors – those that are enabled to adapt to their environment, will be far more competitive. The adaptive processor needs to be made aware of its environment and operating conditions through the use of various sensors. It must then have some ability to usefully respond to the sensor stimulus. The focus of this book is not so much on a static configuration of each manufactured part that may be unique, but on *dynamic* adaptation, where the part optimizes itself on the fly.

Many different responses and adaptive approaches have been explored in recent years. These range from circuits that make voltage changes and set body biases to those that generate clock frequency adjustments on logic. New circuit techniques are needed to address the special challenges created by scaling embedded memories. Finally, system level techniques rely on self-correction in the processor logic or asynchronous techniques which remove the reliance on clocks. Each approach has unique challenges

and benefits, and it adds value in particular situations, but regardless of the method, the challenge of reliably testing these adaptive approaches looms as one of the largest. Hence the subtitle the book: Theory and Practice. Ideas (not necessarily good ones) on adaptive designs are easy to come by, but putting these in working silicon that demonstrates the benefits is much harder. The final level of achievement is actually productizing the capability in a high-volume manufacturing flow.

In order for the book to do justice to such a broad and relatively new topic, we invited authors who have already been pioneers in this area to present data on the approaches they have explored. Many of the authors presented at ISSCC2007, either in the Microprocessor Forum, or in the conference sessions. We are humbled to have collected contributions from such an impressive group of experts on the subject, many of whom have been pioneers in the field and produced results that will be impacting the processor design world for years to come. We believe this topic of adaptive design will continue to be a fertile area for research and integrated circuit improvements for the foreseeable future.

*Alice Wang* — *Texas Instruments, Inc.*
*Samuel Naffziger* — *Advanced Micro Devices, Inc.*

# Table of Contents

**Chapter 1 Technology Challenges Motivating Adaptive Techniques..... 1**
David Scott, Alice Wang
1.1 Introduction .......... 1
1.2 Motivation for Adaptive Techniques .......... 2
1.2.1 Components of Power .......... 2
1.2.2 Relation Between Frequency and Voltage .......... 2
1.2.3 Control Loop Implementation. .......... 4
1.2.4 Practical Considerations .......... 4
1.2.5 Impact of Temperature and Supply Voltage Variations .......... 7
1.3 Technology Issues Relating to Performance-Enhancing Techniques. .......... 9
1.3.1 Threshold Voltage Variation .......... 9
1.3.2 Random Dopant Fluctuations .......... 11
1.3.3 Design in the Presence of Threshold Voltage Variation .......... 13
1.4 Technology Issues Associated with Leakage Reduction Techniques .......... 14
1.4.1 Practical Considerations .......... 15
1.4.2 Sources of Leakage Current .......... 16
1.4.3 Transistor Design for Low Leakage .......... 20
1.5 Conclusion .......... 21
References .......... 21

**Chapter 2 Technological Boundaries of Voltage and Frequency Scaling for Power Performance Tuning .......... 25**
Maurice Meijer, José Pineda de Gyvez
2.1 Adaptive Power Performance Tuning of ICs .......... 25
2.2 AVS- and ABB-Scaling Operations .......... 28
2.3 Frequency Scaling and Tuning .......... 31
2.4 Power and Frequency Tuning .......... 33
2.5 Leakage Power Control .......... 37
2.6 Performance Compensation .......... 40
2.7 Conclusion .......... 44
References .......... 46

**Chapter 3 Adaptive Circuit Technique for Managing Power Consumption** ........ **49**
Tadahiro Kuroda, Takayasu Sakurai
3.1 Introduction ........ 49
3.2 Adaptive $V_{DD}$ Control ........ 50
3.2.1 Dynamic Voltage Scaling ........ 50
3.2.2 Frequency and Voltage Hopping ........ 51
3.3 Adaptive $V_{TH}$ Control ........ 55
3.3.1 Reverse Body Bias (VTCMOS) ........ 55
3.3.1.1 Self-Adjusting Threshold Voltage (SAT) Scheme ........ 55
3.3.1.2 Leakage Current Monitor ........ 56
3.3.1.3 $V_{\mathrm{TH}}$ Controllability ........ 57
3.3.1.4 Device Perspective ........ 59
3.3.2 Forward Body Bias ........ 60
3.3.3 Control Method and Granularity ........ 61
3.3.4 $V_{TH}$ Control Under Variations ........ 64
3.3.5 $V_{TH}$ Control vs. $V_{DD}$ Control ........ 66
3.4 Hardware and Software Cooperative Control ........ 68
3.4.1 Cooperation Between Hardware and Application Software ........ 68
3.4.2 Cooperation Between Hardware and Operating System ........ 70
3.5 Conclusion ........ 71
References ........ 71

**Chapter 4 Dynamic Adaptation Using Body Bias, Supply Voltage, and Frequency** ........ **75**
James Tschanz
4.1 Introduction ........ 75
4.2 Static Compensation with Body Bias and Supply Voltage ........ 76
4.2.1 Adaptive Body Bias ........ 77
4.2.2 Adaptive Supply Voltage ........ 82
4.3 Dynamic Variation Compensation ........ 84
4.3.1 Dynamic Body Bias ........ 84
4.3.2 Dynamic Supply Voltage, Body Bias, and Frequency ........ 87
4.3.2.1 Design Details ........ 87
4.3.2.2 Measurement Results ........ 89
4.4 Conclusion ........ 92
References ........ 92

**Chapter 5 Adaptive Supply Voltage Delivery for Ultra-Dynamic Voltage Scaled Systems** ........ **95**
Yogesh K. Ramadass, Joyce Kwong, Naveen Verma, Anantha Chandrakasan
5.1 Logic Design for U-DVS Systems ........ 97

5.1.1 Device Sizing ... 98
5.1.2 Timing Analysis ... 100
5.2 SRAM Design for Ultra Scalable Supply Voltages ... 101
5.2.1 Low-Voltage Bit-Cell Design ... 104
5.2.2 Periphery Design ... 105
5.3 Intelligent Power Delivery ... 107
5.3.1 Deriving $V_{DD}$ for Given Speed Requirement ... 107
5.3.2 DC-DC Converter Topologies for U-DVS ... 109
5.3.2.1 Linear Regulators ... 109
5.3.2.2 Inductor Based DC-DC Converter ... 109
5.3.2.3 Switched Capacitor Based DC-DC Converter ... 110
5.3.3 DC-DC Converter Design and Reference Voltage Selection for Highly Energy-Constrained Applications ... 112
5.3.3.1 Minimum Energy Tracking Loop ... 113
5.4 Conclusion ... 119
References ... 120

**Chapter 6 Dynamic Voltage Scaling with the XScale Embedded Microprocessor ... 123**
Lawrence T. Clark, Franco Ricci, William E. Brown
6.1 The XScale Microprocessor ... 123
6.1.1 Chapter Overview ... 124
6.1.2 XScale Micro-Architecture Overview ... 125
6.1.3 Dynamic Voltage Scaling ... 126
6.1.4 The Performance Measurement Unit ... 127
6.2 Dynamic Voltage Scaling on the XScale Microprocessor ... 129
6.2.1 Running DVS ... 130
6.3 Impact of DVS on Memory Blocks ... 134
6.3.1 Guaranteeing SRAM Stability with DVS ... 134
6.4 PLL and Clock Generation Considerations ... 138
6.4.1 Clock Generation for DVS on the 180 nm 80200 XScale Microprocessor ... 138
6.4.2 Clock Generation 90 nm XScale Microprocessor ... 139
6.5 Conclusion ... 142
References ... 142

**Chapter 7 Sensors for Critical Path Monitoring ... 145**
Alan Drake
7.1 Variability and its Impact on Timing ... 145
7.2 What Is a Critical Path ... 147
7.3 Sources of Path Delay Variability ... 148
7.3.1 Process Variation ... 149

7.3.2 Environmental Variation ........................................................ 149
7.4 Timing Sensitivity of Path Delay ................................................ 151
7.5 Critical Path Monitors ................................................................ 158
7.5.1 Synchronizer .......................................................................... 158
7.5.2 Delay Path Configuration ..................................................... 159
7.5.3 Time-to-Digital Conversion .................................................. 163
7.5.3.1 Sensitivity ........................................................................ 167
7.5.4 Control and Calibration ........................................................ 168
7.6 Conclusion ................................................................................. 169
Acknowledgements ......................................................................... 171
References ....................................................................................... 171

**Chapter 8 Architectural Techniques for Adaptive Computing......... 175**
Shidhartha Das, David Roberts, David Blaauw, David Bull, Trevor Mudge
8.1 Introduction ................................................................................ 175
8.1.1 Spatial Reach ......................................................................... 177
8.1.2 Temporal Rate of Change ...................................................... 177
8.2 “Always Correct” Techniques ..................................................... 179
8.2.1 Look-up Table-Based Approach ............................................ 179
8.2.2 Canary Circuits-Based Approach ........................................... 180
8.2.3 In situ Triple-Latch Monitor .................................................. 181
8.2.4 Micro-architectural Techniques ............................................. 182
8.3 Error Detection and Correction Approaches ............................... 183
8.3.1 Techniques for Communication and Signal Processing ........ 184
8.3.2 Techniques for General-Purpose Computing ........................ 186
8.4 Introduction to Razor .................................................................. 187
8.4.1 Razor Error Detection and Recovery Scheme ....................... 188
8.4.2 Micro-architectural Recovery ................................................ 190
8.4.2.1 Recovery Using Clock-Gating .......................................... 190
8.4.2.2 Recovery Using Counter-Flow Pipelining ........................ 191
8.4.3 Short-Path Constraints ........................................................... 192
8.4.4 Circuit-Level Implementation Issues ..................................... 192
8.5 Silicon Implementation and Evaluation of Razor ........................ 195
8.5.1 Measurement Results ............................................................. 196
8.5.2 Total Energy Savings with Razor ........................................... 197
8.5.3 Razor Voltage Control Response ........................................... 199
8.6 Ongoing Razor Research ............................................................. 200
8.7 Conclusion .................................................................................. 202
References ....................................................................................... 203

**Chapter 9 Variability-Aware Frequency Scaling in Multi-Clock Processors ........ 207**
Sebastian Herbert, Diana Marculescu
9.1 Introduction ........ 207
9.2 Addressing Process Variability ........ 209
9.2.1 Approach ........ 209
9.2.2 Combinational Logic Variability Modeling ........ 212
9.2.3 Array Structure Variability Modeling ........ 213
9.2.4 Application to the Frequency Island Processor ........ 215
9.3 Addressing Thermal Variability ........ 217
9.4 Experimental Setup ........ 218
9.4.1 Baseline Simulator ........ 218
9.4.2 Frequency Island Simulator ........ 219
9.4.3 Benchmarks Simulated ........ 219
9.5 Results ........ 220
9.5.1 Frequency Island Baseline ........ 220
9.5.2 Frequency Island with Critical Path Information ........ 221
9.5.3 Frequency Island with Thermally Aware Frequency Scaling ........ 222
9.5.4 Frequency Island with Critical Path Information and Thermally Aware Frequency Scaling ........ 224
9.6 Conclusion ........ 224
Acknowledgements ........ 225
References ........ 225

**Chapter 10 Temporal Adaptation – Asynchronicity in Processor Design ........ 229**
Steve Furber, Jim Garside
10.1 Introduction ........ 229
10.2 Asynchronous Design Styles ........ 230
10.3 Asynchronous Adaptation to Workload ........ 232
10.4 Data Dependent Timing ........ 234
10.5 Architectural Variation in Asynchronous Systems ........ 237
10.5.1 Adapting the Latch Style ........ 237
10.5.2 Controlling the Pipeline Occupancy ........ 240
10.5.3 Reconfiguring the Microarchitecture ........ 241
10.6 Benefits of Asynchronous Design ........ 244
10.7 Conclusion ........ 245
References ........ 245

**Chapter 11 Dynamic and Adaptive Techniques in SRAM Design...... 249**
John J. Wuu
11.1 Introduction .......... 249
11.2 Read and Write Margins .......... 250
11.2.1 Voltage Optimization Techniques .......... 251
11.2.1.1 Column Voltage Optimization .......... 252
11.2.1.2 Row Voltage Optimization .......... 255
11.2.2 Timing Control .......... 257
11.3 Array Power Reduction .......... 259
11.3.1 Sleep Types .......... 259
11.3.1.1 Active Sleep .......... 260
11.3.1.2 Passive Sleep .......... 261
11.3.2 P Versus N Sleep .......... 263
11.3.3 Entering and Exiting Sleep .......... 264
11.3.4 Dynamic Cache Power Down .......... 266
11.3.5 Data Bus Encoding .......... 266
11.4 Reliability .......... 267
11.4.1 Soft Errors .......... 267
11.4.2 Hard Errors .......... 267
11.4.2.1 Cache Line Disable .......... 268
11.4.2.2 Cache Line Remap .......... 268
11.4.2.3 Defect Correction .......... 268
11.5 Conclusion .......... 269
References .......... 270

**Chapter 12 The Challenges of Testing Adaptive Designs .......... 273**
Eric Fetzer, Jason Stinson, Brian Cherkauer, Steve Poehlman
12.1 The Adaptive Features of the Itanium 2 9000 Series .......... 273
12.1.1 Active De-skew .......... 273
12.1.2 Cache Safe Technology .......... 277
12.1.3 Foxton Technology .......... 278
12.2 The Path to Production .......... 281
12.2.1 Fundamentals of Testing with Automated Test Equipment (ATE) .......... 281
12.2.2 Manufacturing Test .......... 281
12.2.3 Class or Package Testing .......... 283
12.2.4 System Testing .......... 285
12.3 The Impact of Adaptive Techniques on Determinism and Repeatability .......... 286
12.3.1 Validation of Active De-skew .......... 287
12.3.2 Testing of Active De-skew .......... 290
12.3.3 Testing of Power Measurement .......... 291

12.3.4 Power Measurement Impacts on Other Testing .................. 294
12.3.5 Test Limitations and Guard-Banding .................................. 296
12.4 Guard-Band Concerns of Adaptive Power Management ............ 297
12.5 Conclusion ........................................................................ 300
References ............................................................................ 300

**Index ..................................................................................... 303**

# List of Contributors


| | |
|---|---|
| Alan Drake | IBM |
| Alice Wang | Texas Instruments |
| Anantha Chandrakasan | Massachusetts Institute of Technology |
| Brian Cherkauer | Intel Corporation |
| David Blaauw | University of Michigan |
| David Bull | ARM Ltd. |
| David Roberts | University of Michigan |
| David Scott | Taiwan Semiconductor Manufacturing Company Ltd. |
| Diana Marculescu | Carnegie Mellon University |
| Eric Fetzer | Intel Corporation |
| Franco Ricci | Marvell Semiconductor Inc. |
| James Tschanz | Intel Corporation |
| Jason Stinson | Intel Corporation |
| Jim Garside | The University of Manchester |
| John J. Wuu | Advanced Micro Devices, Inc. |
| José Pineda de Gyvez | NXP Semiconductors,<br>Eindhoven University of Technology |
| Joyce Kwong | Massachusetts Institute of Technology |
| Lawrence T. Clark | Arizona State University |

| | |
|---|---|
| Maurice Meijer | NXP Semiconductors |
| Naveen Verma | Massachusetts Institute of Technology |
| Sebastian Herbert | Carnegie Mellon University |
| Shidhartha Das | ARM Ltd.,<br>University of Michigan |
| Steve Furber | The University of Manchester |
| Steve Poehlman | Intel Corporation |
| Tadahiro Kuroda | Keio University |
| Takayasu Sakurai | University of Tokyo |
| Trevor Mudge | University of Michigan |
| William E. Brown | Ellutions, LLC |
| Yogesh K. Ramadass | Massachusetts Institute of Technology |

# Chapter 1 Technology Challenges Motivating Adaptive Techniques

David Scott,[1] Alice Wang[2]

[1]Taiwan Semiconductor Manufacturing Company Ltd., [2]Texas Instruments, Inc.

## 1.1 Introduction

In the design of an integrated circuit, the designer is faced with the challenge of having circuits and systems function over multiple operating points. From the point of view of performance, the circuit must meet its speed requirements over a range of voltages and temperatures that reflect the environment that the circuit is operating in. Also while the performance requirement must be met at a set of worst-case conditions for speed, the power requirement must be simultaneously met at another set of worst-case conditions for power.

Although each design is unique, the resulting instances of fabricated integrated circuits will number potentially in the billions. In addition, the number of components for each of the integrated circuits will also potentially number in the billions. Every single one of the billions of transistors in every one of the billions of circuits is unique. The success of an integrated circuit design is simply measured by the percentage of the fabricated integrated circuits with the transistors, as well as interconnections, meeting all the requirements.

The use of adaptive techniques allows for an integrated circuit to adapt for variations in the environment as reflected by both voltage and temperature and also for variations in the fabricated transistors. Adaptive techniques are intended to allow minimization of both dynamic and leakage power and also to increase the frequency of operation of the integrated circuit.

A. Wang, S. Naffziger (eds.), *Adaptive Techniques for Dynamic Processor Optimization*,
DOI: 10.1007/978-0-387-76472-6_1, © Springer Science+Business Media, LLC 2008

## 1.2 Motivation for Adaptive Techniques

### 1.2.1 Components of Power

The total power dissipation of an integrated circuit can be simply represented by the power equation below. The power is divided into three major components: the dynamic component, the subthreshold leakage component, and the parasitic leakage components. The dynamic component depends on the overall capacitance of the integrated circuit and the charge that must be displaced for each clock cycle. This is the power that is actually doing work to implement the function of the integrated circuit. Techniques such as clock gating [1] reduce power by gating the clock in unused parts of the integrated circuit, thereby reducing the effective capacitance of the integrated circuit:

$$P = CV^2f + V\,N\,(I_0 \exp(-nV_{th}/(kT/q)) + I_{tox} + I_{gedl}) \quad (1.1)$$

The subthreshold leakage current is simply tied to the threshold voltage of the transistors in the integrated circuit. As the threshold voltage ($V_{th}$) increases, the subthreshold current decreases exponentially. However, while significant leakage savings can be achieved by increasing the threshold voltage, a high threshold voltage tends to force designers to operate the circuits at higher voltages in order to achieve the performance goals. Gate oxide leakage ($I_{tox}$) and gate edge diode leakage ($I_{gedl}$) relate to the characteristics of both the gate oxide and the silicon. These parasitic components of leakage will be discussed in a later section.

### 1.2.2 Relation Between Frequency and Voltage

As shown in Figure 1.1, operating frequency increases as the supply voltage of the integrated circuit increases [2]. Note that the straight line in this plot does not extrapolate back to zero but rather a larger value that depends on the threshold voltage of the transistors in the circuit. Hence, at a given supply voltage, the frequency of the integrated circuit can be changed by changing the threshold voltage. Threshold voltage can be controlled dynamically by changing the transistor body bias. Hence, supply voltage and body bias provide two degrees of freedom over which to control both frequency and power.

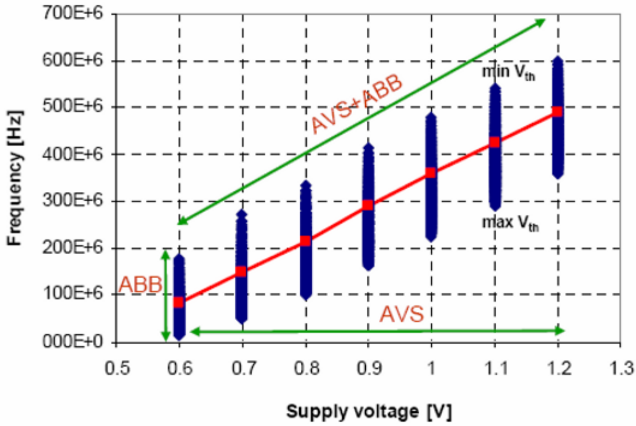



**Figure 1.1** Frequency versus voltage [2]. (© 2005 IEEE)

A conceptual plot for the case where the frequency is constrained to be a constant is shown in Figure 1.2. This illustrates the tradeoffs that can be made between the choice of supply voltage and threshold voltage. If the transistors have a low threshold voltage, the leakage power is very high and the dynamic power is quite low. That is because the operating frequency can be achieved at a relatively lower supply voltage and yet the low threshold voltage results in a high leakage current. As the threshold voltage is increased, the supply voltage to maintain the operating frequency is also increased and hence dynamic power increases. At the same time, the increasing threshold voltage results in a lower leakage power. For a given integrated circuit, there is an optimum point where the power is minimized. This is the point where the increase in dynamic power is offset by the decrease in leakage power.

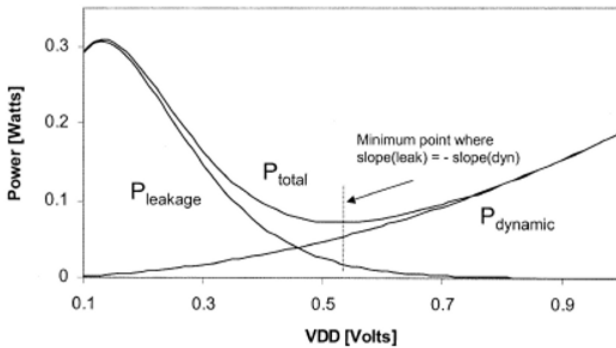



**Figure 1.2** As VDD is increased, the body bias is adjusted to keep operating frequency constant [3]. (© 2002 IEEE)

### 1.2.3 Control Loop Implementation

An example control loop to control body bias is shown in Figure 1.3 [3]. A clock signal is input into a replica circuit and into a phase detector at the same time. The purpose of the phase detector is to detect whether the signal edge is able to pass through the replica circuit in a single clock cycle. Based on whether the signal edge precedes or follows a single clock cycle, the output of the phase detector increases or lowers the body bias accordingly. For this scheme to work, the replica circuit must be representative of the other circuits within the chip that are being controlled by the control loop. A similar scheme can be implemented to control the supply voltage of the replica line where in this case the supply voltage is either incremented or decremented in order to control the speed of the replica circuit.

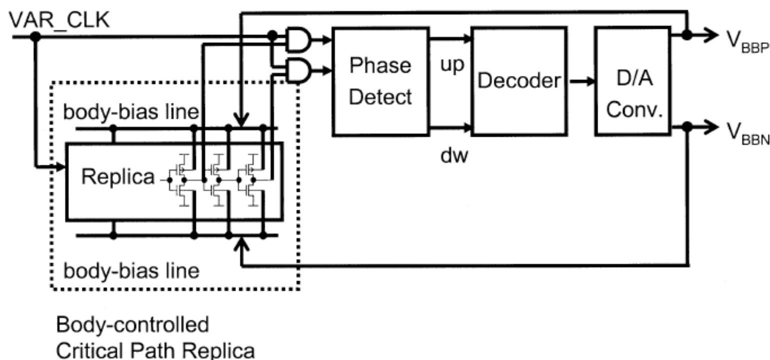



**Figure 1.3** Illustration of replica path [3]. (© 2005 IEEE)

### 1.2.4 Practical Considerations

The key limitation of implementing an adaptive technique is the extent to which the replica circuit represents the integrated circuit. The replica is just one circuit while an integrated circuit has literally thousands of delay paths. This oversimplification is often resolved, assuming that the replica circuit represents the worst-case delay path.

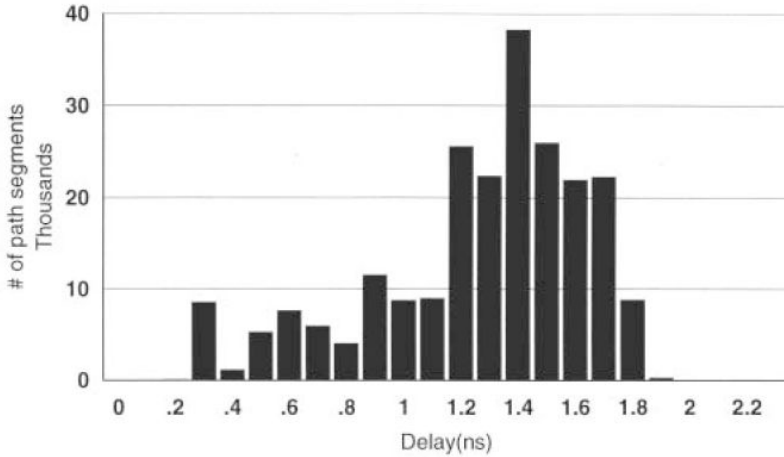
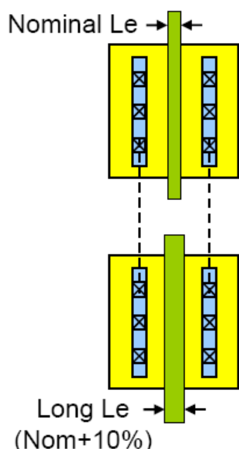



**Figure 1.4** An illustration of critical paths in a design [4]. (© 2004 IEEE)

A typical histogram of delay path segments is shown in Figure 1.4 [4]. As seen from observing this histogram, many of the paths are much faster than the slowest path, and this variation represents a further opportunity to reduce power. The transistors in the faster paths can be substituted with transistors with lower leakage. One way to do this is by selective use of transistors with longer channel length. Due to the longer channel length, these transistors will be slower but they will also have reduced leakage.

An example of this has already been implemented in an integrated circuit [5] through the use of a library of circuits that were implemented with both long and short gate lengths. A slight area penalty was incurred to make each circuit in the library footprint and layout compatible as in Figure 1.5. Hence, these circuits can be freely interchanged at any point in the design cycle to minimize power at the expense of path delay. This algorithm can be similarly implemented using multiple threshold voltage transistors.

The use of the above algorithm for substitution of longer gate length transistors to reduce leakage can occur on a massive scale as is shown in Figure 1.6. One result of implementing this type of algorithm is that all delay path segments become more critical as the extra slack in the design is harvested in order to reduce leakage current. Making all these paths more critical will tend to make the design less tolerant of circuit variations or circuit modeling inaccuracies.




- All transistors can be either nominal or long-Le
- Most library cells are available in both flavors
- Long-Le transistors are about 10% slower, but have 3x lower leakage
- All paths with timing slack use long-Le transistors
- Initial design uses only long channel devices

**Figure 1.5** Two transistors with the same layout footprint. Layout area efficiency is sacrificed in order to make the shorter channel transistor replaceable by the longer channel transistor [5]. (© 2006 IEEE)

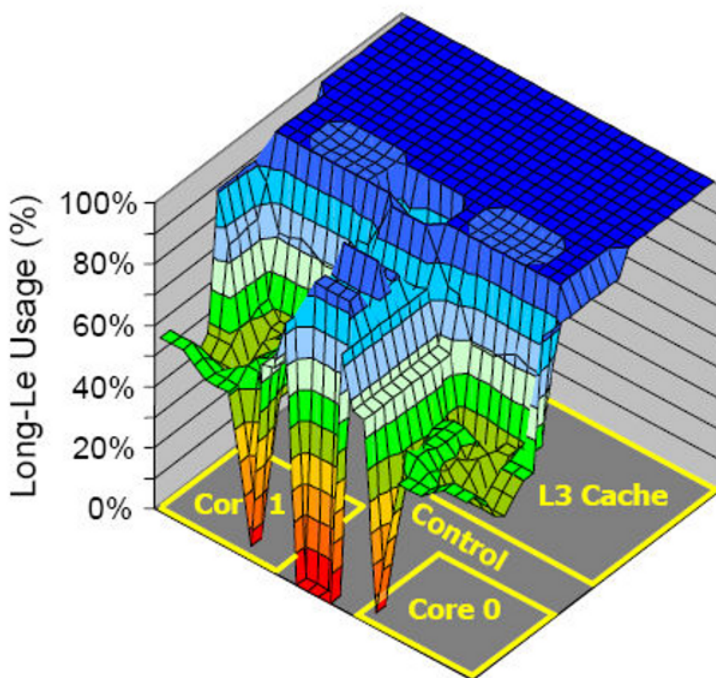



**Figure 1.6** Usage of long channel transistors in a design showing that a shorter channel is required for a small fraction of the transistors in this design [5]. (© 2006 IEEE)

### 1.2.5 Impact of Temperature and Supply Voltage Variations

In the last section, we showed how the operating frequency varies with supply voltage. In this section, we also factor in the temperature dependence as well as across chip variations. The operating frequency as a function of supply voltage is shown for two different temperatures in Figure 1.7. At low temperatures, the mobility of the carriers is lower and hence the operating frequency is lower when the supply voltage is high. At high temperature, the lower threshold voltage favors low voltage operation. These two curves cross at what is normally considered the nominal operating voltage. Hence in the absence of adaptive techniques, modern integrated circuits show very little sensitivity of operating frequency to temperature.

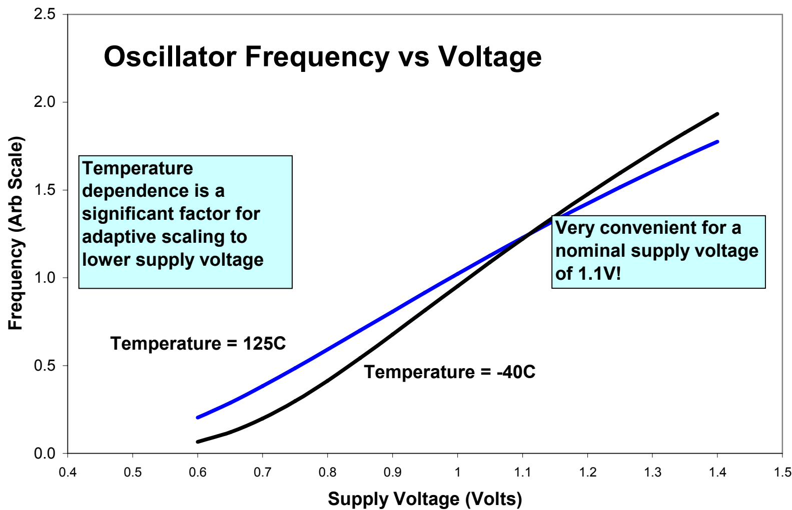



**Figure 1.7** A plot of frequency versus voltage for a circuit at two different temperatures [6]. (© 2007 IEEE)

At low voltages, the type of behavior described in Figure 1.7 greatly favors high-temperature operation, and hence there is a great deal of sensitivity of operating frequency to temperature at low voltages. In general, there is a great deal of temperature variation across a chip [7]. At low voltages, the coldest parts of the chip will have the most problem, while at high temperature, the hottest parts of the chip will be slowest. Any adaptive scheme must account for the changing sensitivity to temperature at the low and high operating voltages.

An analysis of the impact of supply voltage variation on the chip has been previously done for two different cases [8]. This analysis was

facilitated by recognizing that in steady state, the supply voltage across the chip must satisfy the following equation:

$$\nabla^2 V = R_S J_O \tag{1.2}$$

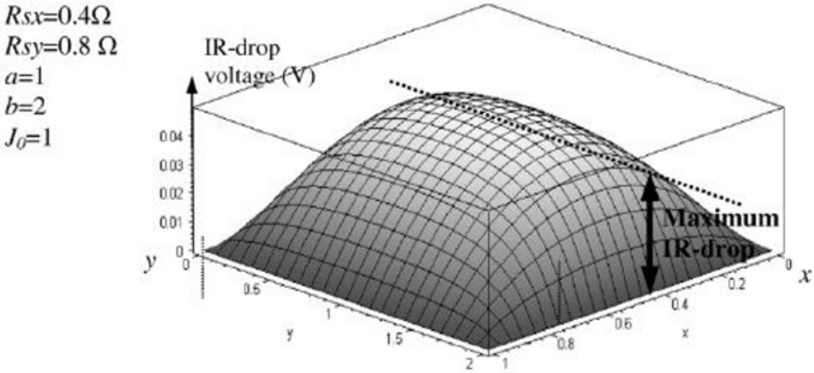



**Figure 1.8a** (Case 1) Supply voltage drop across a chip that has been wire bonded with supply pads on the edge of the chip [8]. (© 2005 IEEE)

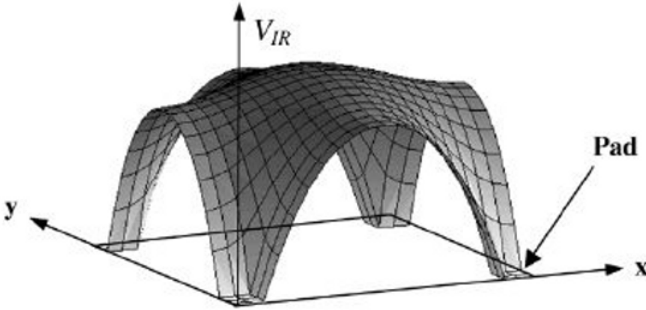



**Figure 1.8b** (Case 2) Supply voltage variations for flip chip where the power supply pads are arrayed over the complete chip area [8]. (© 2005 IEEE)

Case 1 is the wire bond case where the perimeter of the chip is pinned to the supply voltage and case 2, is the flip chip case where the supply and grounds pins are placed in a mesh across the chip. In Figure 1.8a, the wire-bound case is shown where clearly the maximum voltage supply loss is at the center of the chip. For the flip chip case shown in Figure 1.8b, each small part of the chip has the supply voltage pinned in the corners only, and this pattern is arrayed over the entire chip with the amplitude dependent

on the local power supply current. The maximum operating frequency of each of the path segments will depend on the local supply voltage.

## 1.3 Technology Issues Relating to Performance-Enhancing Techniques

### 1.3.1 Threshold Voltage Variation

While in the previous section we dealt with variations in the supply voltage caused by on-chip voltage drops, in this section we discuss the variation in threshold voltage. The impact of threshold voltage variation is shown in Figure 1.9 involving circuits with two different threshold voltages. The sensitivity of the frequency to threshold voltage has the impact of shifting the curve to the right as the threshold voltage increases.

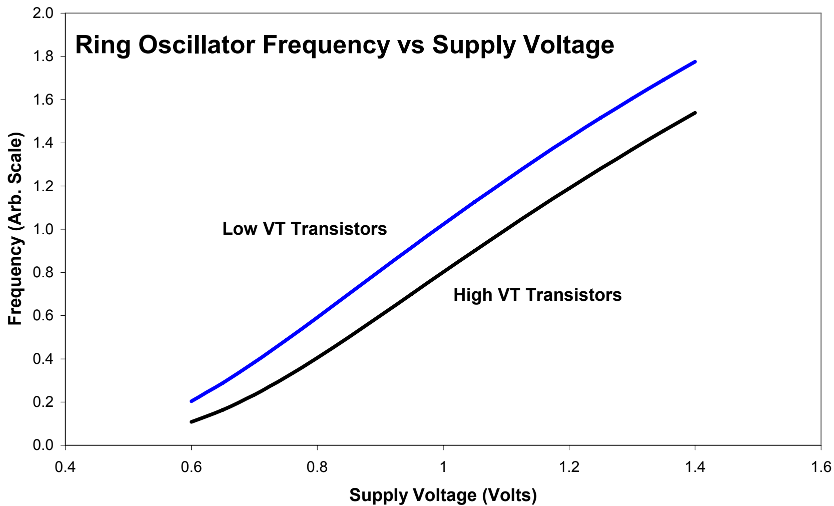



**Figure 1.9** As the threshold voltage is increased, the frequency versus supply voltage curve is shifted to the right [6]. (© 2005 IEEE)

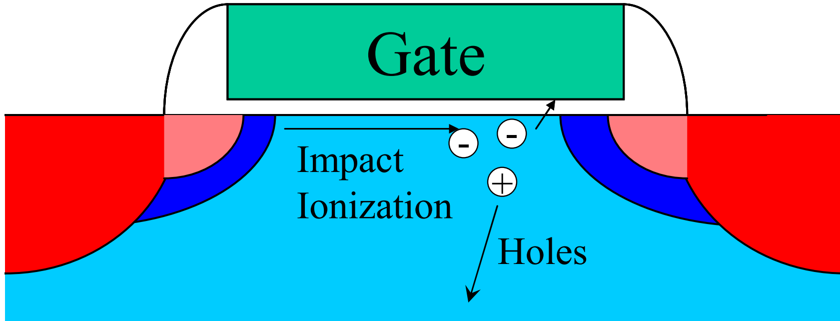



**Figure 1.10a** The mechanism of impact ionization is illustrated, with electron–hole pairs being generated at the drain end of the channel. Some of these generated carriers end up being trapped in the gate oxide [6]. (© 2005 IEEE)

The impact of hot carrier-induced threshold voltage shift is shown in Figure 1.10a. At high electric fields, carriers generated from impact ionization are trapped in the gate oxide. In general, the transistor lifetime decreases with the cube of the substrate current, and the gate voltage dependence of the substrate current is illustrated in Figure 1.10b. Also, there are separate degradation characteristics from both ac- and dc-related stress currents. It has also been found that the threshold voltage and other device parameters can shift over the lifetime of the product and not always in the same direction [8]. For example, the threshold voltage can recover after the stress is removed [9].

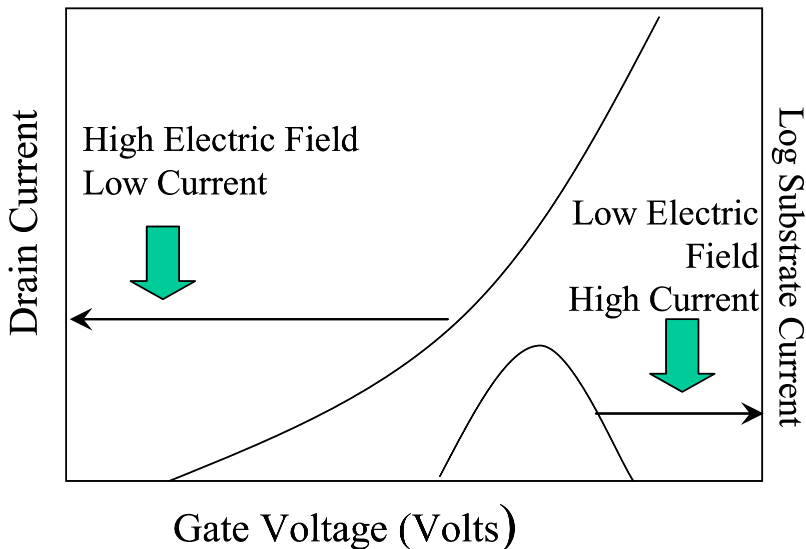



**Figure 1.10b** A peak in the substrate current occurs when high current flow and high electric field occur both at the same time [6]. (© 2005 IEEE)

Another source of threshold variation is negative bias temperature instability (NBTI). This phenomenon is commonly associated with p-channel transistors and is caused by the movement of charge in the gate oxide and at the interface. Of course in an integrated circuit, each transistor has a unique set of bias conditions over its lifetime and hence each transistor degrades differently. For a typical stress condition of negative bias, the variation of threshold voltage with time is given by the following equation [10]:

$$\Delta V_{th} = C \exp(-E_A / kT) \exp(\beta \mid V_G \mid) t^n \tag{1.3}$$

where $E_A$ is the activation energy, $V_G$ is the gate voltage, and $t$ is the stress time, and the other parameters are constants.

NBTI degradation has the biggest impact at lower supply voltage. This is due to the loss of headroom as the p-channel threshold voltage increases in absolute value. Previous work [10] has shown that as a ring oscillator is stressed, the low voltage frequency of operation is degraded due to the increase of p-channel threshold voltage. The impact of threshold voltage on low frequency operation can be observed in Figure 1.7. The circuit using transistors with the higher threshold voltage has a lower frequency of operation.

### 1.3.2 Random Dopant Fluctuations

As dimensions continue to shrink, the number of dopants in the channel has become discrete and measurable in discrete quantities The small number doping atoms in the channel means that the threshold voltage will be highly variable and will vary for transistors with otherwise identical characteristics. The variation in threshold voltage can be related to the average doping by the following equation [11]:

$$\sigma_{V_{th}} = \left( \frac{\sqrt[4]{4q^3 \varepsilon_{Si} \phi_B}}{2} \right) \bullet \frac{T_{ox}}{\varepsilon_{ox}} \bullet \frac{\sqrt[4]{N}}{\sqrt{W_{eff} L_{eff}}} \tag{1.4}$$

Measured data shown in Figure 1.11 shows that the standard deviation for typical state of the art dimensions is quite significant, with standard deviation in the range of 30 mV–50 mV being quite feasible. For a chip with many millions of gates, transistors with threshold voltages more than 5 standard deviations from the mean are relatively common.

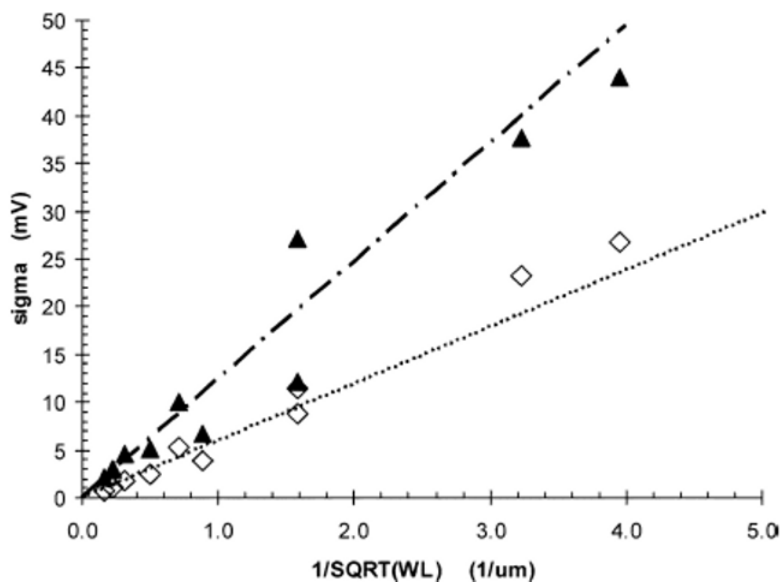



**Figure 1.11** Measured data showing the increase in threshold voltage variation as the area of the transistor is decreased. Diamonds are for strong inversion while triangles are for subthreshold region [11]. (© 2005 IEEE)

In terms of applying adaptive techniques, the difficulty that the circuit designer faces is compounded. Identical transistors placed in different parts of the circuit tend to have randomly different values of threshold voltage as described by Equation 1.4 and Figure 1.11.

In addition, as body bias is applied, the randomness of the threshold voltage will tend to increase [12]. When body bias is increased, more dopants are incorporated into the depletion region and hence the randomness of these additional dopants is also incorporated into the transistor.

New transistor design techniques are continuously under development, and these scaled transistors offer new challenges to the designer. Taking advantage of these new transistor design techniques can be of great value to the circuit design. As shown in Figure 1.12, Yasuda [12] found that as body bias is applied to a collection of transistors, their threshold voltage distribution has a tendency to reorder. That is to say different transistors have different responses to body bias and hence the transistor with the lowest threshold voltage in a distribution may no longer be the lowest when body bias is applied. Certainly, this is a concern to a designer who is using body bias to control transistor performance or leakage.

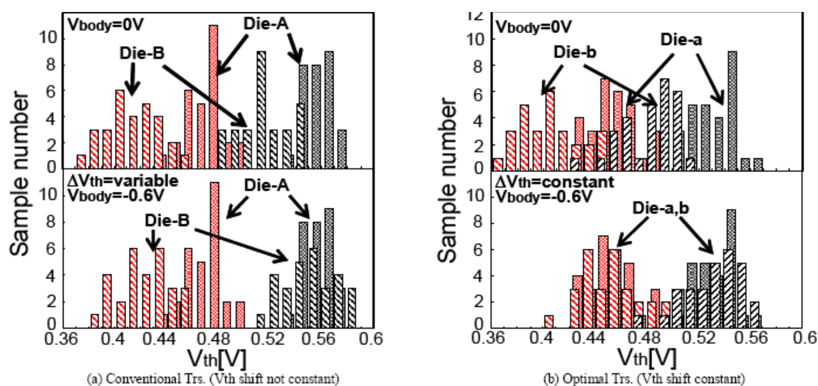



**Figure 1.12** The benefit of an optimized transistor design is shown where the threshold voltage shift with body bias is constant [12]. (© 2005 IEEE)

It has been shown that new transistor design techniques, taking advantage of Fermi level pinning present in [13], offer an advantage in the transistor design. If the channel of the transistor is optimally designed, the response of the transistor to applied body bias can be made much more predictable.

### 1.3.3 Design in the Presence of Threshold Voltage Variation

When designing an adaptive system, the designers must contend with a number of sources of threshold voltage variation that are not under their direct control. Transistor characteristics and in particular threshold voltage can vary from wafer to wafer and also from die to die within a wafer. These variations are generally known as global variations. In addition, the designers must contend with the local variations that occur within a die.

Local variations can be due to random dopant fluctuations including the transistors having different sensitivities to back gate bias, line edge roughness of the gate material, and systematic changes in device behavior such as temperature and temperature gradients across the die.

Transistor characteristics can also change over the lifetime of the integrated circuit as a result of hot carrier effects or negative bias temperature-induced (NBTI) changes in the threshold voltage. In addition, new techniques to improve transistor performance by using mechanical stress [14] also will bring additional sources of variation.

As a result of all of the above, designers are generally not looking at a single line on the frequency versus voltage curve that can be modulated with back gate bias. They must think of this line as having considerable

width, with the width of this line being defined by the sum total of all the variations that can occur in the transistor through fabrication, during its operation, and over its lifetime.

## 1.4 Technology Issues Associated with Leakage Reduction Techniques

A common technique to reduce subthreshold leakage is to merely increase the threshold voltage by applying back gate bias [15, 16]. In Figure 1.13, the waveforms show how the leakage of an integrated circuit can be reduced when the system is going into a lower power mode of operation. As discussed earlier, this type of scheme requires substrate terminals of the transistors to be available globally, and hence these two extra supply lines must be available to be globally routed. Also the substrate pump requires additional area and consumes current in order to operate. The operating current associated with the substrate pump offsets the leakage gains. The well bias generator function is often conveniently provided by the same supply as is used for the IO circuits, and hence an extra penalty for providing this extra supply is normally not incurred.

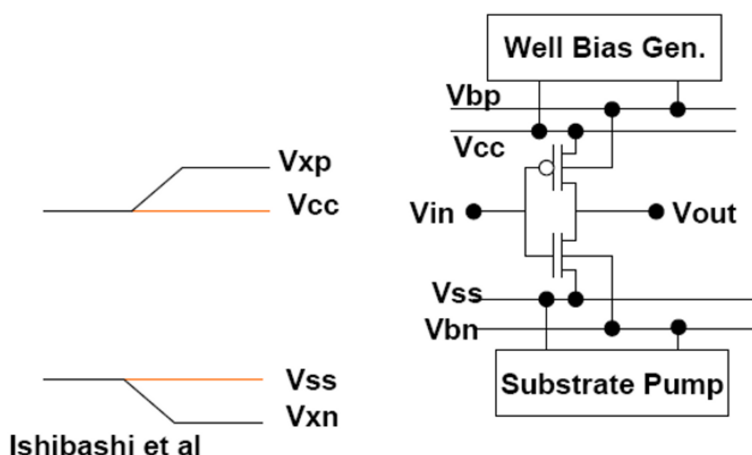



**Figure 1.13** One scheme to reduce leakage is to merely apply back gate bias to all transistors [6]. (© 2005 IEEE)

The need for a substrate pump has been avoided by some designers by raising Vss rather than having Vbn negative [1]. This type of scheme has the added benefit of current being reduced due to both back gate bias and

the lower operating voltage. However, switching of the Vss supply does require footer transistors that are large enough to conduct the entire active current of the circuit without having an excessive voltage drop. In addition, once added the footer transistors themselves become leakage sources.

### 1.4.1 Practical Considerations

The ability of a technology to support state-of-the-art integrated circuits and systems is conventionally judged by its leakage versus "on" current. Figure 1.14 shows such plots for three different technologies [14]. The individual data points for each different technology are achieved as a result of measuring transistors with different gate lengths and threshold voltage implants. In addition the normal process variations, discussed in the previous section, can play a role in smoothing out these curves. One important aspect of this tradeoff to note is that the "Ion" scale is linear while the leakage current scale is exponential.

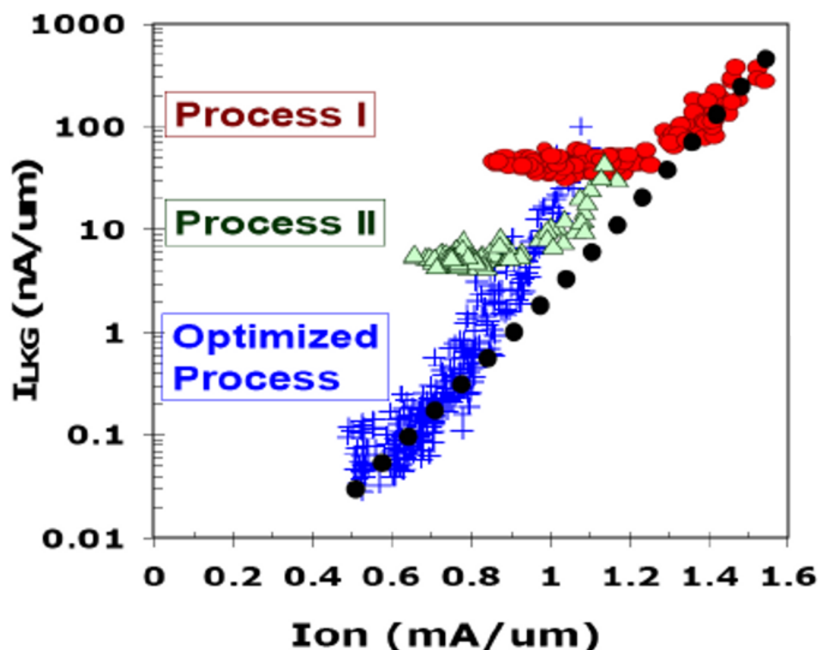



**Figure 1.14** Leakage versus ion for three different technologies [14]. (© 2005 IEEE)

Although Process 1 has a higher "on" current and higher "off" current than the process optimized for mobile applications, both of the plots are