Design Strategies and Circuit Techniques for Low-Power and Low-Supply CMOS Circuits Used in Low-Cost Optical Communication Systems

by

Bangli Liang

A Thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of
Master of Applied Science

Ottawa-Carleton Institute for
Electrical and Computer Engineering

Department of Electronics
Carleton University
Ottawa, Ontario, Canada
November 2008

Copyright ©
2008 - Bangli Liang
NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.
The undersigned recommend to
the Faculty of Graduate Studies and Research
acceptance of the Thesis

Design Strategies and Circuit Techniques for Low-Power and
Low-Supply CMOS Circuits Used in Low-Cost Optical
Communication Systems

Submitted by **Bangli Liang**
in partial fulfilment of the requirements for the degree of
**Master of Applied Science**

_________________________
Tadeusz Kwasniewski, Supervisor

_________________________
Zhigong Wang, Co-Supervisor
Southeast University, Nanjing, P.R. China

_________________________
Langis Roy, Department Chair

Carleton University
2008
Important Information

The information used in this thesis comes in part from the research programs of Dr. Tadeusz Kwasniewski and Dr. Zhigong Wang, and their associates. The research results appearing in this thesis represent an integral part of the ongoing research program. All research results in this thesis including tables, graphs, and figures, but excluding the narrative portions of the thesis are effectively incorporated into the research program and can be used by Dr. Tadeusz Kwasniewski, Dr. Zhigong Wang, and their associates for educational and research purposes, including publication in open literature with appropriate credits. The matters of intellectual property may be pursued cooperatively with Carleton University as well as Southeast University (through Dr. Tadeusz Kwasniewski and Dr. Zhigong Wang) when and where appropriate.
Abstract

This thesis begins with the discussion of the background and motivation of this thesis, CMOS technology and circuit logic, and then low-power high-speed circuit techniques and design bottlenecks are reviewed.

Based on the comprehensive overview, a number of advanced design strategies and key circuit techniques such as DC bias optimizing, device sizing, split-resister (S-R) low capacitive load, compact active inductor, MOS-based capacitor and resistor, active negative feedback topology are identified and analyzed to realize low-power high-speed circuits with reduced silicon area.

To validate the selected strategies and techniques, a low-jitter wide tuning range CDR, three low-power 1.25-Gb/s to 10-Gb/s limiting amplifiers, a high modulation efficiency LDD/MD were designed and fabricated; 1V-supply circuits such as 1:2 DE-MUX, 2:1 MUX, data decision circuit, and high input sensitivity 2:1 static frequency divider were designed and simulated.

Finally, the measurement data and post-layout simulation data confirmed that the selected circuit techniques and design strategies indeed results in circuits with lower power dissipation or lower supply voltage, and higher operating speed.
Acknowledgments

I would like to express my gratitude to everyone that has enabled me to complete this thesis.

I am deeply grateful to my supervisor, Dr. Tadeusz Kwasniewski, and my co-supervisor, Dr. Zhigong Wang. I would like to thank them for letting me explore different fields, for always making me understand and interpret the true nature of things and great technical guidance.

I would also like to thank Dr. Shoujun Wang for his support and trust in helping me to enroll at Carleton University.

I would also like to thank my course instructors, Dr. Garry Tarr, Dr. Jim Wight, Dr. Calvin Plett, Dr. Ralph Mason and Dr. John W. M. Rogers for their instructions in the courses I took. I wish to thank all the staff of DOE for their great assistances. I would like to thank Carleton University and Department of Electronics for the financial support and for providing a great environment for me to do research and interact with outstanding people.

I am very grateful to Dianyong Chen, Bo Wang, John Cheng, Lei Zhang, M. Usama, and other group members for being great colleagues and friends. I wish to thank all the people in VLSI team for all the help and assistance.

I wish to thank my wife Qian Gao, my parents, and parents-in-law for giving me the reason to be happy, and strength to work and live. Each in their own way gave me the strength to stay on the course, and for that I dedicate this thesis to them.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Important Information</td>
<td>iii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>List of Symbols</td>
<td>xvi</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xvii</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Chapter Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Thesis Goal</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Document Organization</td>
<td>3</td>
</tr>
<tr>
<td><strong>2 CMOS Technology and Circuit Logic</strong></td>
<td>5</td>
</tr>
<tr>
<td>2.1 Chapter Overview</td>
<td>5</td>
</tr>
</tbody>
</table>
2.2 IC Technologies and Technology Choice ........................................... 5
  2.2.1 Technology Requirements ...................................................... 5
  2.2.2 Transistor Performance ......................................................... 6
  2.2.3 Passive Devices: Inductors and Varactors ............................. 8
  2.2.4 Building Blocks of Optical Communication System .................. 9
  2.2.5 Technology Choices ............................................................ 11

2.3 Circuit Logic .............................................................................. 12
  2.3.1 CMOS Rail-to-Rail Logic ......................................................... 14
  2.3.2 CMOS Current Mode Logic ..................................................... 16
  2.3.3 Circuit Logic Choice ............................................................... 18

2.4 Chapter Summary ...................................................................... 18

3 Low-Power High-Speed Techniques ................................................. 19
  3.1 Chapter Overview ................................................................. 19
  3.2 Overview of High-Speed Amplifiers ....................................... 19
    3.2.1 Resistor-Loaded CS Amplifier .............................................. 19
    3.2.2 NMOS-Load CS Amplifier ..................................................... 21
    3.2.3 CS-CG Cascode Amplifier (Reduction of $C_{gd}$ Impact) .......... 22
    3.2.4 CS Amplifier with Neutralization ......................................... 23
    3.2.5 Shunt-peaked CS Amplifier .................................................. 24
    3.2.6 Zero-peaked CS Amplifier .................................................... 27
    3.2.7 Bandwidth Enhancement with $f_T$ Doublers ........................... 28
    3.2.8 Broadband Amplifiers Using Cascading Gain Stages ............ 29
    3.2.9 Distributed Amplifier (DA) .................................................... 32

  3.3 Overview of High-Speed Analog Circuits .................................. 33
    3.3.1 Broadband Output Buffers .................................................... 33
    3.3.2 CMOS Driver for Laser Diode or Optical Modulator ............ 35

vii
3.3.3 High-Speed Equalizer Filter ........................................ 37
3.4 Overview of High-Speed Digital Circuits ............................ 38
  3.4.1 Monolithic Transformer Coupled 2 : 1 MUX ..................... 38
  3.4.2 High-Speed Frequency Divider ................................. 41
3.5 Chapter Summary ..................................................... 52

4 Advanced Design Strategies and Circuit Techniques for Low-Power
High-Speed Circuit Implementations .................................... 53
  4.1 Chapter Overview ................................................... 53
  4.2 Design Fundamentals ............................................... 53
    4.2.1 Current-Mode Logic ......................................... 54
    4.2.2 Advanced CMOS Components ................................. 55
  4.3 Design Strategies and Circuit Techniques ........................ 57
    4.3.1 DC Bias Level Optimizing ................................... 57
    4.3.2 Device Aspect Ratio Optimizing ............................. 58
    4.3.3 Split-Resistor (S-R) Loads ................................ 60
    4.3.4 Active Inductors, Capacitors, Resistors ................... 61
    4.3.5 Other Useful Circuit Techniques ............................ 67
  4.4 Chapter Summary ................................................... 71

5 Low-Power and Low-Supply ICs Design ............................... 72
  5.1 Chapter Overview ................................................... 72
  5.2 Low-power Analog Circuits ....................................... 72
    5.2.1 High Modulation Amplitude LDD/MD ......................... 73
    5.2.2 Low-Power Limiting Amplifiers ............................ 79
    5.2.3 Monolithic CDR Circuit .................................... 87
  5.3 Low-Supply Digital Circuits ..................................... 97
    5.3.1 Low-Supply Circuits Design ................................. 98
List of Tables

2.1 Performance of State-of-the-Art Semiconductor Technologies .................. 6
3.1 Gain, BW, and GBW of Cascading Gain Stages ................................. 32
5.1 Measurement Summary of LD/MZM driver ........................................ 78
5.2 Measurement Summary of LAs .......................................................... 87
5.3 Measurement Summary of 0.6μm CMOS CDR .................................... 96
5.4 Performance Comparison of CMOS DEMUX ...................................... 103
5.5 Performance Comparison of CMOS MUX .......................................... 105
5.6 Performance Comparison of CMOS Frequency Divider ....................... 109
5.7 Performance Comparison of CMOS Decision Circuit ......................... 111
List of Figures

2.1 Simulated CMOS inductors: (a) Q vs. Frequency, (b) L vs. Frequency. 8
2.2 Simulated CMOS varactors: (a) Q vs. Frequency; (b) C vs. Frequency. 9
2.3 Simulated CMOS varactors: (a) Q vs. $V_{\text{control}}$; (b) C vs. $V_{\text{control}}$. 10
2.4 Comparison of power consumption for CMOS tail-to-tail logic and CMOS current mode logic versus frequency. 12
2.5 Two cascading inverters and the equivalent circuit of the first stage. 14
2.6 CMOS CML differential pair. 16
3.1 Resistor-loaded CS amplifier and its gain-frequency characteristic. 20
3.2 Resistor-loaded CS differential amplifier. 21
3.3 NMOS-loaded CS amplifier and its gain-frequency characteristic. 21
3.4 CS-CG cascode amplifier. 22
3.5 Simulation of cascode amplifier and resistor-loaded common-source amplifier: (a) AC response; (b) Transient response. 23
3.6 CS amplifier with neutralization. 23
3.7 Shunt peaking amplifier. 24
3.8 Spiral inductor model and active inductor model. 24
3.9 Bandwidth extension and normalized gain versus m. 26
3.10 Zero-peaked common source amplifier. 26
3.11 Comparison between zero-peaked and resistor-loaded common-source amplifier: (a) AC response; (b) Transient response. 27
3.12 Creating a single-ended output. ............................................. 28
3.13 Creating the bias for $M_2$. .................................................. 28
3.14 Block diagram of cascading gain stages. .............................. 29
3.15 Normalized transfer function and achievable bandwidth. ........ 31
3.16 Cascading gain stages and AC response. ............................. 31
3.17 Typical Distributed amplifier. ............................................. 32
3.18 Output buffer: (a) Circuit schematic. (b) Operating region of the input pair transistors. (c) Simplified equivalent circuit. ................. 34
3.19 Open-drain driver for distributed select circuit. .................... 35
3.20 Quasi push-pull source follower and its equivalent circuit. ........ 35
3.21 Schematic of equalizer filter. ............................................ 37
3.22 Top-level diagram of MUX and schematic diagram of the MS-FF. ... 39
3.23 Schematic of the MUX stage with the monolithic transformer. .... 39
3.24 (a) Divide-by-2 SCL; (b) Logic diagram of toggle switch. ......... 42
3.25 (a) Dynamic DFF implementation; (b) TSPC frequency divider. ... 44
3.26 (a) Modified 2:1 frequency divider; (b) Conventional frequency divider;
(c) Timing chart of the proposed 2:1 frequency divider. ............... 46
3.27 TSPC ÷2 implementation and the third stage showing capacitances. 48
3.28 Dynamic CMOS logic divider operating at 6.5GHz and 14GHz. ..... 50
3.29 Dynamic CMOS logic divider operating at 10GHz. .................. 51
3.30 TSPC dynamic frequency divider operating at 5GHz and 18GHz. .... 51
3.31 TSPC frequency divider operating at 10GHz. ....................... 52
4.1 $f_T$ versus $V_{GS}$, $V_{DS}$, and W/L: (a) $(2\mu m \times 10)/0.12\mu m$ LVT and RVT devices; (b) $(5\mu m \times 10)/0.12\mu m$ LVT device under various $V_{DS}$ .... 55
4.2 Proposed static frequency divider. (a) Divider core; (b) Output buffer. 58
4.3 Common source amplifiers with (a) Split-resistor; (b) Shunt peaking;
(c) Split-resistor, shunt peaking, and series peaking. ................. 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Active inductors and small signal model.</td>
</tr>
<tr>
<td>4.5</td>
<td>AC response of shunt-peaked CS amplifier: (a) Trimmable active inductor with a gate resistor $R_g$; (b) Comparison of shunt peaking inductors.</td>
</tr>
<tr>
<td>4.6</td>
<td>Low voltage-drop active inductor and capacitive voltage converter.</td>
</tr>
<tr>
<td>4.7</td>
<td>2.5Gb/s Transimpedance Amplifier in 0.35µm CMOS technology.</td>
</tr>
<tr>
<td>4.8</td>
<td>(a) Schematic and equivalent circuit of the regulated cascode active inductor; (b) Complete circuit schematic of the fully integrated quadrature hybrid.</td>
</tr>
<tr>
<td>4.9</td>
<td>Current sources. (a) Schematic. (b) $I_D$ versus $V_{DS}$.</td>
</tr>
<tr>
<td>4.10</td>
<td>CSA with BW extension: (a) AC response; (b) Transient response.</td>
</tr>
<tr>
<td>4.11</td>
<td>Output buffers. (a) For DEMUX; (b) For data decision circuit.</td>
</tr>
<tr>
<td>5.1</td>
<td>Schematic of the proposed LD/MZM driver circuit.</td>
</tr>
<tr>
<td>5.2</td>
<td>Microphotograph of the fabricated LD/MZM driver (Total area is 0.6mm×0.65mm including 0.1mm×0.1mm bonding pads).</td>
</tr>
<tr>
<td>5.3</td>
<td>Block diagrams of evaluation setup for (a) LD driver; (b) MZM driver.</td>
</tr>
<tr>
<td>5.4</td>
<td>Optical output of the fabricated LD/MZM driver: (a) 625-Mb/s from LD; (b) 1.25-Gb/s from LD; (c) 625-Mb/s from MZM; (b) 1.25-Gb/s from MZM.</td>
</tr>
<tr>
<td>5.5</td>
<td>Limiting amplifier architecture.</td>
</tr>
<tr>
<td>5.6</td>
<td>Schematic of the proposed circuits: (a) Gain cell with active inductors. (b) Active feedback gain cell with folded active inductors for low supply LA.</td>
</tr>
<tr>
<td>5.7</td>
<td>Schematic of the proposed output buffer.</td>
</tr>
<tr>
<td>5.8</td>
<td>Schematic of the used offset cancelation feedback circuit.</td>
</tr>
<tr>
<td>5.9</td>
<td>Schematic of the used on-chip signal loss detection and alarm circuit.</td>
</tr>
</tbody>
</table>
5.10 Microphotographs of the fabricated circuits: (a) 0.6μm CMOS LA (0.5mm×0.4mm); (b) 0.25μm CMOS LA (0.7mm×0.5mm); (c) 0.18μm CMOS LA (1.0mm×0.7mm).

5.11 Measurement results at 5mVpp input of the fabricated LAs: (a) 0.6μm CMOS LA at 1.25-Gb/s; (b) 0.25μm CMOS LA at 6-Gb/s; (c) 0.18μm CMOS LA at 10-Gb/s.

5.12 Block diagram of the fabricated CDR.

5.13 The proposed I/Q VCO: (a) Block diagram; (b) Schematic of the proposed variable delay gain stage; (c) Differential control circuit.

5.14 Schematic of the proposed (a) PD/QPD, and (b) FD.

5.15 Timing diagrams of (a) PD/PFD; (b) PD, QPD, and FD.

5.16 Schematic of the realized loop filter (Q1 from PD, Q3 from FD, and Vctrl to control circuit).

5.17 Microphotograph of the fabricated monolithic 0.6μm CMOS CDR.

5.18 Measurement results: (a) Frequency control curve of differential tuning VCO; (b) Measured spectrum of the in-locked VCO.

5.19 Measurement results of 0.6μm CMOS CDR: (a) Eye diagram of the recovered data at 622Mb/s; (b) Jitter histogram of the locked VCO at 622MHz.

5.20 Proposed circuits. (a) Latch in DEMUX; (b) MUX with a buffer; (c) Data decision circuit core.

5.21 Layout of (a) 1:2 DEMUX; (b) Data decision Circuit.

5.22 20-Gb/s DEMUX output eye-diagrams: (a) DEMUX core without peaking; (b) Buffered DEMUX without peaking; (c) DEMUX core with peaking; (d) Buffered DEMUX with peaking.

5.23 PVT simulations of DEMUX (40-Gb/s data, 20GHz clock).

5.24 Post-layout PVT simulations of DEMUX (40-Gb/s data, 20GHz clock).
5.25 Simulations of the proposed MUX (40-Gb/s data output) 103
5.26 Transient waveforms of the proposed MUX (40-Gb/s operating) 104
5.27 PVT simulations of the proposed MUX (40-Gb/s data output) 104
5.28 Voltage effect on self-resonant frequency \( f_{SR} \) of the proposed frequency divider: (a) \( f_{SR} \) versus \( V_{DD} \); (b) \( f_{SR} \) versus \( V_{CM} \) 105
5.29 Simulated input sensitivity curves of 3 frequency dividers: V1-Traditional static divider; V2-Static divider with SP; V3-Static divider with SP and S-R 106
5.30 Transient waveforms of frequency divider: (a) V1; (b) V2; (c) V3 107
5.31 Transient powers of the proposed frequency dividers 107
5.32 Starting self-oscillations of the frequency divider with SP and S-R under different supply voltages: (a) \( V_{DD}=1V \); (b) \( V_{DD}=1.2V \); (c) \( V_{DD}=1.5V \) 108
5.33 Transient waveforms of the proposed frequency divider (V3) at 42GHz 108
5.34 Data decision circuit operating at 40-Gb/s: (a) Input data; (b) Core circuit output; (c) Full circuit (with buffer) output 109
5.35 S-R ratio optimization to a trade-off between amplitude and jitter 110
5.36 Operating at 42-Gb/s: (a) \( V_{DD}=1.0V \); (b) \( V_{DD}=1.2V \); (c) \( V_{DD}=1.5V \) 110
5.37 PVT simulations of data decision circuit at 40-Gb/s: (a) "FF", 1.5V, -55°C; (b) "SS", 1.0V, 125°C 111
List of Symbols

NMOS
PMOS
VDD/GND
Inverter/Buffer
NAND/AND Gate
NOR/OR Gate
Transmission Gate
Resistor/Inductor/Capacitor/Diode
D-Latch/DFF
Current Source
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Alternative Current</td>
</tr>
<tr>
<td>APD</td>
<td>Avalanche Photodiode</td>
</tr>
<tr>
<td>BER</td>
<td>Bit-Error Rate</td>
</tr>
<tr>
<td>BiCMOS</td>
<td>Bipolar and CMOS Transistors</td>
</tr>
<tr>
<td>BW</td>
<td>Bandwidth</td>
</tr>
<tr>
<td>C²MOS</td>
<td>Clocked CMOS</td>
</tr>
<tr>
<td>CDR</td>
<td>Clock and Data Recovery</td>
</tr>
<tr>
<td>CM</td>
<td>Common-Mode</td>
</tr>
<tr>
<td>CML</td>
<td>Current Mode Logic</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary Metal-Oxide-Semiconductor</td>
</tr>
<tr>
<td>CS-CG</td>
<td>Common Source/Common Gate</td>
</tr>
<tr>
<td>DA</td>
<td>Distributed Amplifier</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>DEMUX</td>
<td>Demultiplexer</td>
</tr>
<tr>
<td>DFF</td>
<td>D Flip-Flop</td>
</tr>
<tr>
<td>EOI</td>
<td>Electrical/Optical Interface</td>
</tr>
<tr>
<td>ESD</td>
<td>Electrostatic Discharge</td>
</tr>
<tr>
<td>FD</td>
<td>Frequency Detector</td>
</tr>
<tr>
<td>$f_{MAX}$</td>
<td>Maximum Oscillation Frequency</td>
</tr>
<tr>
<td>$f_T$</td>
<td>Transit Frequency of Transistors</td>
</tr>
<tr>
<td>GBW</td>
<td>Gain-Bandwidth Product</td>
</tr>
<tr>
<td>HBT</td>
<td>Heterojunction Bipolar Transistor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>HEMT</td>
<td>High Electron Mobility Transistor</td>
</tr>
<tr>
<td>I/O</td>
<td>Input/Output</td>
</tr>
<tr>
<td>IC</td>
<td>Integrated Circuit</td>
</tr>
<tr>
<td>ILFD</td>
<td>Injection-Locked Frequency Divider</td>
</tr>
<tr>
<td>LA</td>
<td>Limiting Amplifier</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Networks</td>
</tr>
<tr>
<td>LDD</td>
<td>Laser Diode Driver</td>
</tr>
<tr>
<td>LSI</td>
<td>Large Scale Integration</td>
</tr>
<tr>
<td>LVT/RVT</td>
<td>Low-$V_T$/Regular-$V_T$</td>
</tr>
<tr>
<td>MAN</td>
<td>Metropolitan Area Networks</td>
</tr>
<tr>
<td>MD</td>
<td>Modulator Driver</td>
</tr>
<tr>
<td>MIM</td>
<td>Metal-Isolator-Metal</td>
</tr>
<tr>
<td>MMIC</td>
<td>Monolithic Microwave Integrated Circuits</td>
</tr>
<tr>
<td>MAN</td>
<td>Metropolitan Area Networks</td>
</tr>
<tr>
<td>MOSFET</td>
<td>Metal Oxide Semiconductor Field Effect Transistor</td>
</tr>
<tr>
<td>MS-FF</td>
<td>Master-Slave Flip-Flop</td>
</tr>
<tr>
<td>MUX</td>
<td>Multiplexer</td>
</tr>
<tr>
<td>MZM</td>
<td>Mach-Zehnder $LiNbO_3$ External Modulator</td>
</tr>
<tr>
<td>NRFD</td>
<td>Narrowband Regenerative Frequency Divider</td>
</tr>
<tr>
<td>NRZ</td>
<td>Non-Return-to-Zero</td>
</tr>
<tr>
<td>OAN</td>
<td>Optical Access Networks</td>
</tr>
<tr>
<td>OC</td>
<td>Optical Carrier</td>
</tr>
<tr>
<td>PD</td>
<td>Phase Detector</td>
</tr>
<tr>
<td>PDN</td>
<td>Pull-Down Network</td>
</tr>
<tr>
<td>PFLL</td>
<td>Phase/Frequency-Locked Loop</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical Layer</td>
</tr>
<tr>
<td>PLL</td>
<td>Phase-Locked Loop</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>PRBS</td>
<td>Pseudo-Random Bit Sequences</td>
</tr>
<tr>
<td>PUN</td>
<td>Pull-Up Network</td>
</tr>
<tr>
<td>PVT</td>
<td>Process, Voltage, Temperature</td>
</tr>
<tr>
<td>Q</td>
<td>Quality Factor</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RMS</td>
<td>Root Mean Square</td>
</tr>
<tr>
<td>SAN</td>
<td>Storage Area Networks</td>
</tr>
<tr>
<td>SDH</td>
<td>Synchronous Digital Hierarchy</td>
</tr>
<tr>
<td>SerDes</td>
<td>Serializer/Deserializer</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SOC</td>
<td>System-On-a-Chip</td>
</tr>
<tr>
<td>SONET</td>
<td>Synchronous Optical Networking</td>
</tr>
<tr>
<td>SP</td>
<td>Shunt Peaking</td>
</tr>
<tr>
<td>SR</td>
<td>Split Resistor</td>
</tr>
<tr>
<td>STM</td>
<td>Synchronous Transport Module</td>
</tr>
<tr>
<td>THL</td>
<td>Transparent High Latch</td>
</tr>
<tr>
<td>TIA</td>
<td>Transimpedance Amplifier</td>
</tr>
<tr>
<td>TLL</td>
<td>Transparent Low Latch</td>
</tr>
<tr>
<td>TSPC</td>
<td>True Single-Phase Clocked</td>
</tr>
<tr>
<td>TWA</td>
<td>Traveling Wave Amplifier</td>
</tr>
<tr>
<td>TX/RX</td>
<td>Transmitter /Receiver</td>
</tr>
<tr>
<td>VCO</td>
<td>Voltage-Controlled Oscillator</td>
</tr>
<tr>
<td>VSR</td>
<td>Very Short Reach</td>
</tr>
<tr>
<td>VT</td>
<td>Threshold Voltage</td>
</tr>
<tr>
<td>WAN</td>
<td>Wide Area Networks</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Networks</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Chapter Overview

This chapter describes the background and what motivated the research described in this thesis, thesis goal, and also organization of this document.

1.2 Introduction

The demand for wide-bandwidth networks such as local area networks (LAN), wide area networks (WAN), metropolitan area networks (MAN), very short reach (VSR), storage area networks (SAN) and wireless local area networks (WLAN) is increasing because of the rapid growth of broadband internet access and the performance of computer and storage systems. A wider bandwidth network has also been required in home electronics appliances such as full-motion video, multimedia, e-commerce, and advanced digital services to transfer large-volume image streams. This increase in speed is much higher than the speed increase that corresponds to the famous Moore’s law prediction of LSI growth. The way to meet the wide-bandwidth network requirement and then improve performance through low power and low cost is to develop a high-speed CMOS interface and implement it in a system-on-a-chip (SOC).
Current generation Ethernet LANs are being deployed with 10/100-Mb/s connections at the desk and 1-Gb/s on the LAN backbones. 10-Gb/s physical layer (PHY) IC offerings used in long haul applications exist but power and cost issues limit the sustainability of the end product for VSR, SAN, LAN, and MAN applications.

Since the mainstream CMOS technology has successfully invaded into the low GHz applications, it will be interesting to investigate the feasibility of CMOS technology for the emerging 10 to 40-GHz applications. To meet the demands mentioned above, optical communication systems should be implemented through low-cost low-power high-integration technologies. Among Silicon-based technologies, GaAs- and InP-based III/V technologies, and SiGe-based technologies, CMOS should be the best candidate to realize low-cost commercial applications and more and more researches are focused on low-power high-speed CMOS circuits.

Today's optical fiber communication systems operate at bit rates between 155-Mb/s and 40-Gb/s. Current high-end communication ICs are mainly implemented in GaAs, InP, or SiGe bipolar technologies. Several high-speed chips in CMOS technologies are reported in [1-9], which confirm CMOS to be a viable alternative for high-speed circuit design because advanced circuit techniques and a state-of-the-art fabrication process can be combined to extend speed limits. This approach is very economical due to the lower production costs, higher yield, and integration density. On the other hand, it is desirable to implement low-end chips for 10/100-Mb/s connections at the desk and 1-Gb/s on the LAN backbones using low power CMOS and employing very compact circuit topologies to reduce manufacturing cost.

As key blocks in optical communication systems, low power high modulation current/voltage amplitude laser diode/modulator driver (LDD/MD), broadband amplifier, high bitrate demultiplexer (DEMUX), multiplexer (MUX), high operating frequency divider and data decision circuit (data retimer) used in clock and data recovery (CDR) circuit are widely used and have attracted more and more attention.
In this thesis, major research will focus on design strategies and circuit techniques to result in a trade-off among speed, power dissipation, and silicon area.

1.3 Thesis Goal

The thesis goal is to seek suitable design strategies for power and speed optimization and to select a number of effective and feasible circuit techniques and topologies for experimental verification. The goal and the rule to select techniques are:

1. Choose some circuit techniques to improve effective $f_T$, to lower supply voltage and power consumption.

2. Select one or more effective circuit methods to reduce RC time constant and device mismatch, to increase input sensitivity, to lower supply voltage and power consumption.

3. Introduce feasible topologies to improve gain-bandwidth product (GBW) by reducing critical node capacitance and internal signal swing.

4. Use area-saving devices and compact topologies to boost circuit bandwidth, to reduce silicon area, to improve circuit performance, process tolerance, integration density, and compatibility with digital circuits in CMOS logic.

1.4 Document Organization

This document is organized as follows. Chapter 2 reviews CMOS technology and circuit logic. Chapter 3 reviews low-power high-speed circuit techniques and design bottlenecks. In Chapter 4, advanced design strategies and key circuit techniques are identified and analyzed based on CMOS current-mode logic (CML) topology. In Chapter 5, low-power analog circuits using compact active inductors are fabricated
and evaluated, and low-supply digital circuits using spiral inductors and low-voltage circuit techniques are designed and simulated, which are used to validate the identified design strategies and selected circuit techniques. Previous chapters are summarized and conclusions are drawn in the last chapter. References cited in this thesis are listed at the end.
Chapter 2

CMOS Technology and Circuit Logic

2.1 Chapter Overview

The rapid progress of semiconductor technologies makes it possible to implement high performance chips for super high speed data communication systems. However, it is still a big challenge to realize low-power lower speed data communication systems (10-Gb/s per channel and below) and low-supply higher speed data communication systems (40-Gb/s per channel and above). To design competitive chips, it is required to understand well the relationship among semiconductor technology, circuit logic and circuit, block, even system specifications. In this chapter, CMOS technology and circuit logic will be reviewed.

2.2 IC Technologies and Technology Choice

2.2.1 Technology Requirements

The first generation OC-192 systems have relied exclusively on GaAs-based technologies with typical $f_T$ (Transit frequency) and $f_{MAX}$ (Maximum oscillation frequency) values of 60-70GHz. This situation is now being tipped in favor of SiGe npn and
Table 2.1: Performance of State-of-the-Art Semiconductor Technologies

<table>
<thead>
<tr>
<th>Parameter</th>
<th>InP DHBT</th>
<th>GaAs HBT</th>
<th>SiGe HBT</th>
<th>RVT NMOS</th>
<th>RVT PMOS</th>
<th>LVT NMOS</th>
<th>LVT PMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size ((\mu)m)</td>
<td>1</td>
<td>2</td>
<td>0.25</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>(f_T) (GHz)</td>
<td>160</td>
<td>70</td>
<td>70</td>
<td>100</td>
<td>52</td>
<td>115</td>
<td>58</td>
</tr>
<tr>
<td>(f_{MAX}) (GHz)</td>
<td>200</td>
<td>70</td>
<td>90</td>
<td>110</td>
<td>54</td>
<td>120</td>
<td>60</td>
</tr>
<tr>
<td>(J_{cpfT}) (mA/(\mu)m²)</td>
<td>1.2</td>
<td>0.8</td>
<td>2</td>
<td>2.6</td>
<td>1.3</td>
<td>2.6</td>
<td>1.3</td>
</tr>
<tr>
<td>(I_{pfT}) (ps/V)</td>
<td>1</td>
<td>0.5</td>
<td>1.4</td>
<td>1.4</td>
<td>0.7</td>
<td>1.4</td>
<td>0.7</td>
</tr>
<tr>
<td>(BV_{CEO}) (V)</td>
<td>6</td>
<td>&gt;10</td>
<td>2.8</td>
<td>&gt;1.5</td>
<td>&gt;1.5</td>
<td>&gt;1.5</td>
<td>&gt;1.5</td>
</tr>
<tr>
<td>(V_{BE/VT}) (V)</td>
<td>0.8</td>
<td>1.4</td>
<td>0.9</td>
<td>0.385</td>
<td>-0.400</td>
<td>0.300</td>
<td>-0.290</td>
</tr>
<tr>
<td>Reference</td>
<td>[10]</td>
<td>[10]</td>
<td>[10]</td>
<td>This</td>
<td>This</td>
<td>This</td>
<td>This</td>
</tr>
</tbody>
</table>

CMOS transistors as their \(f_T\) and \(f_{MAX}\) exceed 40GHz. Below is a list of active and passive device requirements for highly integrated IC’s [10].

- Transistor speed: \(f_T > 4\) bitrate, \(f_{MAX} > 5\) bitrate, which must be reached at \(V_{CE}=V_{CC}/3\) or \(V_{DS}=V_{DD}/2\) over all process corners and entire temperature range

- Large Q (Quality Factor) high resonant frequency inductor, MIM (Metal-Insulator-Metal) capacitor and varactor diode

- Back-end with large number of metal layers which helps integration levels

- Low-K moderate thickness dielectrics for good isolation and low-parasitic capacitance

### 2.2.2 Transistor Performance

The most relevant transistor performance for fiber-optic IC’s is adequately captured by the following figures of merit [10]:

- Transistor speed: \(f_T > 4\) bitrate, \(f_{MAX} > 5\) bitrate, which must be reached at \(V_{CE}=V_{CC}/3\) or \(V_{DS}=V_{DD}/2\) over all process corners and entire temperature range

- Large Q (Quality Factor) high resonant frequency inductor, MIM (Metal-Insulator-Metal) capacitor and varactor diode

- Back-end with large number of metal layers which helps integration levels

- Low-K moderate thickness dielectrics for good isolation and low-parasitic capacitance
• Peak $f_T$ and $f_{MAX}$ values (Determine circuit speed)

• Peak $f_T$ current density $J_{cp}f_T$ (Determines power dissipation)

• $I_{cp}f_T/C_{BC}(C_{GD})$ the ratio of the peak $f_T$ current and output capacitance (Intrinsic slew-rate)

• $BV_{CEO}/BV_{DG}$ (Dictates output swing and is at least 2 times larger for GaAs and InP devices than for Si-based devices operating at the same speed)

• Threshold voltage: $V_{BE}/V_T$ (Limits the minimum value of the power supply voltage and favors CMOS over bipolar devices, and InP and SiGe HBTs over GaAs HBTs)

• Minimum feature size (Affects power dissipation, and integration levels and favors Si CMOS and SiGe BiCMOS due to the more mature processing techniques)

• Thermal resistance (Is related to the semiconductor material and to the minimum feature size and affects indirectly the integration levels and the transistor speed that can be achieved under reliable operating conditions)

For 40-Gb/s applications, devices with over 160GHz $f_T$ and $f_{MAX}$ values are required and such performance has so far been reached only by InP HBTs and HEMTs. For the same generation of lithography, NMOS display half the speed of the SiGe HBT with p-channel device performance yet another factor of 2 behind. SiGe HBTs have been aggressively scaled in the last two generations in order to retain their $\times 2$ and $\times 4$ speed advantage over NMOS. In general, for the same $f_T$, the NMOS feature size is two generation ahead of the bipolar one. However, as a rule of thumb, for CMOS to truly exhibit its low power advantage over bipolar devices, the $f_T$ of the PMOS must be sufficiently high to meet the application requirements, i.e. 40GHz
for 10-Gb/s. Otherwise, power-hungry CML-like implementations using NMOS and resistive loads must be employed in order to achieve a bitrate of 10-Gb/s and beyond. $f_T$ values of up to 250GHz have been recently reported for SiGe HBTs. Such performance has only been reached at very high current densities, exceeding $6mA/\mu m^2$, and with breakdown voltages of 1.5-1.8V. These high cutoff frequency figures have so far failed to be concomitantly accompanied by similar $f_{MAX}$ values. The intrinsic slew rate is an important figure of merit for bipolar digital circuits and output drivers. In the case of MOSFETs, the slew rate is very high but the device performance is offset by the low output resistance. Table 2.1 summarizes the transistor performance for the main semiconductor technologies. Because of the large turn on voltage (1.4V), GaAs HBTs can only be used with supply voltages of 5V or higher.

![Figure 2.1: Simulated CMOS inductors: (a) Q vs. Frequency, (b) L vs. Frequency.](image)

### 2.2.3 Passive Devices: Inductors and Varactors

Recent results indicate that inductors with Q values larger than 10 and resonant frequencies beyond 50GHz can be realized on Silicon substrate. Measured characteristics of four octagonal inductors are shown in Figure 2.1 as a function of frequency. For smaller inductors (200pH and below), the peak Q is above 15 in the frequency range
Figure 2.2: Simulated CMOS varactors: (a) $Q$ vs. Frequency; (b) $C$ vs. Frequency.

from 10GHz to 40GHz, about 40% lower than that of III-V inductors of comparable size and operating in the same frequency range [10]; However, the self-resonance frequency goes beyond 100GHz.

Si varactor diode $Q(f)$ and $C(f)$ characteristics, obtained from S-parameter simulations under 0V are shown in Figure 2.2. As shown in Figure 2.2 (a) and Figure 2.3 (a), the $Q$ remains around 10 up to 40GHz even at a varactor voltage of 0V while the capacitance ratio is higher than 2.5 illustrated in Figure 2.3 (b). In general, such a high varactor capacitance ratio is difficult to achieve in III-V technologies where varactor diodes are rarely integrated in a fast HBT process due to conflicting epitaxial layer requirements.

Based on these simulation data, passive devices offered by current main-stream CMOS technologies can be used to design circuits operating at 10-Gb/s to 40-Gb/s.

2.2.4 Building Blocks of Optical Communication System

The building blocks making up a wideband data communication system have different requirements in terms of transistor and passive device performance. The requirements for various digital and analog blocks are summarized below [10]:

...
Digital blocks: a) high $f_T/f_{MAX}$ (speed); b) low peak $f_T$ current density (power dissipation); c) low $V_{GS}$ (power supply and power consumption); d) small device size (power dissipation); e) fine metal pitch (integration)

VCO requirements: a) high Q inductor (low noise), b) high Q varactor with large capacitance ratio (process spread); c) high Q, low parasitic capacitance MIM capacitor (low noise and large oscillation frequency and tuning range); d) high $f_{MAX}$ transistor (large power and low-noise); e) low 1/f noise transistor.

50Ω driver: a) large intrinsic slew-rate for bandwidth and $S_{22}$ matching; b) large breakdown voltage for voltage swing; c) high $f_{MAX}$ (bandwidth)

Transimpedance and Post Amplifier: a) high $f_{MAX}$ for (bandwidth) and b) low noise figure for good sensitivity.

For Si technologies, $f_T$ values have been stuck at around 100-GHz for the most advanced technologies. Generally, the analog functions place more demanding requirements on the speed of transistor technologies than do digital functions. One exception is the master-slave DFF (D flip-flop) in the data decision circuit which is typically clocked at a frequency equal to the data rate. At such high speeds, both bipolar and
FET digital circuits use CML topologies buffered by emitter/source follower stages. Such a topology does not favor low supply voltage and power dissipation. The sensitivity of HEMT or CMOS digital circuits is systematically lower than that of the corresponding bipolar implementations due to the poor $V_T$ match.

To allow low-supply CMOS to play a more important role in high-speed data communications, it is necessary to find new topologies or modify existing topologies and techniques to remove power-hungry emitter/source followers, and to lower supply voltage and power dissipation further, which will be addressed in Chapter 4 and employed to design low-power high-speed circuits described in Chapter 5.

### 2.2.5 Technology Choices

For 10-Gb/s (and below) SerDes (Serializer and Deserializer) functions operating at or below 5V supply, 0.18μm CMOS ($f_T=49\,\text{GHz}$) and well-proven sub-micron (0.6μm, 0.35μm, and 0.25μm) CMOS will most likely become the technologies of choice. 10-Gb/s short to medium reach applications require the most cost effective yet performance technology available. SiGe BiCMOS is a technology that has been in high volume production, driven by consumer wireless applications, for the past years. It is perfectly suited for the EOI (Electro/Optical Interface) family at 10-Gb/s because of the high speed (SiGe) bipolar transistors and CMOS for some control implementation. Technology needs for SerDes are similar to EOI in terms of speed performance. From the point of view of implementing large amounts of digital logic, the CMOS is crucial. However, such a CMOS process is higher cost than conventional "digital" CMOS processes as it will have to integrate high quality varactor, MIM capacitor and thick top metal inductors. GaAs HBTs and p-HEMTs will retain the 5V modulator driver markets in long-haul 10-Gb/s SONET/SDH applications. For requirements at 40-Gb/s serial and above, no technology is currently mainstream enough to be cost effective and manufacturable in high volume.
In fact, it is feasible and economical to realize monolithic laser diode driver and modulator driver without on-chip spiral inductor and external component by using 5V-supply sub-micron (0.6μm and 0.35μm) CMOS technologies for low-end applications (155-Mb/s to 2.5-Gb/s). At the same time, it is possible and promising to realize 40-Gb/s digital blocks such as MUX, DEMUX, frequency divider, and data decision circuits in most advanced CMOS technologies using low supply circuit topologies. In this thesis, both analog circuits for 622-Mb/s to 10-Gb/s systems and digital circuits for 40-Gb/s application are proposed and verified in CMOS technologies, which confirms that CMOS should be the best candidate for low-cost low-power designs.

![Comparison of power consumption for CMOS tail-to-tail logic and CMOS current mode logic versus frequency.](image)

**Figure 2.4:** Comparison of power consumption for CMOS tail-to-tail logic and CMOS current mode logic versus frequency.

### 2.3 Circuit Logic

Since each circuit style has its pros and cons in terms of low-power or high-performance, choosing an appropriate circuit style for a given application requires a good understanding of the characteristics of circuit styles. In this part, the types of logic discussed will be CMOS rail-to-rail logic and CMOS current mode logic (CML). At low frequencies, CMOS rail-to-rail logic is preferred for its simplicity and low static power dissipation, while, at higher frequencies, CML is used, as it can operate faster with lower power because of the reduced signal swing. As shown in Figure 2.4, when
it is not switching, CMOS rail-to-rail logic does not consume any current, while CML does. CMOS rail-to-rail logic consumes current only during transitions, and its power consumption is proportional to the operation frequency. CML bias current must rise as the speed of switching increases, just as CMOS rail-to-rail logic does, but it does so at a slower rate. Thus, above some frequency, CML becomes a lower-power solution. Furthermore, CML logic is differential and, therefore, has good power-supply rejection, which is preferred in many high speed low power applications [11].

The main limitation of circuit speed is the capacitive load and parasitic capacitive load [12,13]. The voltage change on a capacitor can be written as

\[
\Delta V = \frac{\Delta Q}{C_{tot}} = \frac{1}{C_{tot}} \cdot \int_{t_0}^{t_0+\Delta t} i(t) \cdot dt = \frac{I_{ave} \cdot \Delta t}{C_{tot}}
\]  

(2.1)

Where \( C_{tot} \) is the total load capacitance, \( \Delta t \) is the time it takes to change the voltage on the load capacitor with a voltage amplitude \( \Delta V \), and \( I_{ave} \) is the average current to charge or discharge the load capacitor. To make a circuit fast, the time \( \Delta t \) must be sufficiently small.

\[
\Delta t = \frac{\Delta V \cdot C_{tot}}{I_{ave}}
\]  

(2.2)

It is quite straightforward how to make a circuit faster: to reduce the voltage swing \( \Delta V \), or/and make the load capacitance \( C_{tot} \) smaller, or/and increase the charging current \( I_{ave} \). A large logic family exploits these fundamental methods to make digital circuit faster. For example, Pseudo-NMOS and Domino-Logic exclude PMOS capacitance from the input, because PMOS input capacitance is usually 2 ~ 3 times as large as NMOS input capacitance if they provide the same current. Technology scaling down reduces capacitance as well. For the consideration of performance and power, only CML logic and CMOS rail-to-rail logic are discussed here.
2.3.1 CMOS Rail-to-Rail Logic

Static CMOS rail-to-rail logic is by far the most commonly used circuit logic. Despite the very high speed, CMOS rail-to-rail logic is still extensively used in some high-speed transceivers. The reasons are technology scaling down, reduced power supply voltage, and simplicity, maturity, and robustness of static CMOS logic. Static CMOS logic has a pull-up network and a pull-down network. At any time except transitions, either pull-up network is turned on to pull the output to the power supply voltage or pull-down network is turn on to pull down the output to ground. Since pull-up and pull-down are unable to be turned on simultaneously except during transitions, in principle static CMOS logic consumes zero static power. Therefore, static CMOS logic exhibits extremely low power consumption at low frequency applications.

The speed and power consumption of static CMOS logic in high speed applications can be roughly estimated by using two inverters connected in series as shown in Figure 2.5.

Assume the initial input signal \( V_{IN} \) is high, thus PMOS transistor \( M_{P1} \) is turned off, NMOS transistor \( M_{N1} \) is turned on, and the voltage \( V_{O1} \) is low. Let’s further assume that the input signal has very sharp edges and sufficiently large driving capability, then when \( V_{IN} \) jumps from high to low, the time it takes to turn on \( M_{P1} \) and
turn off $M_{N1}$ is negligible. The voltage $V_{O1}$ is pulled up to $V_{DD}$ through transistor $M_{P1}$. However, it cannot change abruptly since it has to drive the gate capacitance of $M_{P2}$ and $M_{N2}$, as well as parasitic capacitances from the four transistors. When $V_{O1}$ increases, $V_{DS}$ of transistor $M_{P1}$ reduces until the channel is not pinched off. Transistor, $M_{P1}$ falls into triode region and the charging current reduces. When $V_{O1}$ reaches $V_{DD}$, the energy ($\epsilon_{store}$) stored in the gate capacitors and parasitic capacitors ($C_{tot}$) is

$$\epsilon_{store} = \frac{1}{2} \cdot C_{tot} \cdot V_{DD}^2 = \frac{1}{2} \cdot \frac{I_{ave}^2 \cdot \Delta T^2}{C_{tot}}$$

(2.3)

Since $V_{DD}$ charges the capacitors through the channel of $M_{P1}$, some energy is consumed by the channel resistance. When the input $V_{IN}$ goes from low to high, PMOS transistor $M_{P1}$ is switched off, and NMOS transistor $M_{N1}$ is turned on. Assume this process is sufficiently fast, then power supply $V_{DD}$ is cut off from the capacitors abruptly so that the power consumption caused by short-circuit effect is negligible, and it provides no energy during the process of discharging the capacitors. However, the stored energy $\epsilon_{store}$ is completely consumed by the channel resistance of transistor $M_{N1}$ when the capacitors are discharged to the ground. The energy consumption for an input cycle is the sum of $\epsilon_{store}$ and the energy dissipated in charging process. The average power consumption of an inverter can be estimated as

$$P_{ave} = V_{DD} \cdot I_{ave} \cdot \Delta T \cdot f = V_{DD}^2 \cdot C_{tot} \cdot f$$

(2.4)

Some conclusions can be drawn on the basis of the simple analysis. Firstly, static CMOS logic has to drive the input capacitance of the pull-up network and the input capacitance of the pull-down network simultaneously. The pull-up network is composed with PMOS transistors and has larger capacitance. From Eq.2.2, this slows down the circuit because of the large capacitance; Second, static CMOS realizes a rail-to-rail output. According to Eq.2.2, this also slows down the circuit because of
the large signal swing. Based on Eq.2.4, this greatly increases the power consumption because of the large swing; Thirdly on the basis of Eq.2.4, static CMOS logic consumes much power at high frequencies because the power consumption is proportional to switching frequency; Lastly, static CMOS logic has poor immunity against common mode noise due to the single-ended operation. Therefore, high-speed CMOS digital design favors current mode logic (CML) rather than static CMOS logic.

An important observation from the two inverters connected in series is that the output of the first inverter has finite slew rate. This is different from our previously assumption that the input voltage signal to an inverter has very sharp edges and infinite driving capability. Therefore, the second inverter will not switch instantaneously, and additional delay is added. This realistic consideration applies to all digital circuits.

![CMOS CML differential pair](image)

**Figure 2.6:** CMOS CML differential pair.

### 2.3.2 CMOS Current Mode Logic

CMOS CML is based on differential pair as shown in Figure 2.6. Since PMOS transistors are removed from the gain stage, it should operate faster than static CMOS logic. The fully balanced differential topology offers excellent immunity against common mode noise. When the input voltage $V_{in}$ is sufficiently large, one of the two
branches can be switched off, while the other takes all the tail current $I_0$. The minimum input voltage can be derived using the following equations.

\[ I_{1,2} = \frac{\mu \cdot C_{ox}}{2} \cdot \left(\frac{W}{L}\right)^2 \cdot (V_{gs1,2} - V_{th})^2 \]  

(2.5)

\[ I_1 + I_2 = I_0 \]  

(2.6)

\[ V_{in} = V_{gs1} - V_{gs2} \]  

(2.7)

Solving Eq.2.5-Eq.2.7 leads to an expression of $I_1$ (or $I_2$). The minimum voltage that can fully switch the differential pairs is given when this current is equal to $I_0$. It can be written as:

\[ \text{Min}(V_{in}) = \sqrt{\frac{2 \cdot I_0}{\mu \cdot C_{ox} \cdot (W/L)}} \]  

(2.8)

The voltage swing is

\[ \Delta V = V(i = 0) - V(i = I_0) = R \cdot I_0 \]  

(2.9)

The voltage swing is the product of the load resistance and the tail current. Therefore, it is possible to reduce the voltage swing to improve the speed of the circuit. However, excessive reduction of voltage swing will reduce the noise margin. In addition, it may not be able to fully switch the following differential pairs.

Similar to static CMOS logic, the speed and power consumption of CMOS CML can be estimated by using two serially connected inverters. We still assume the input to the first inverter has very sharp edges and sufficient driving capability, then using the first order approximation, the output of the first inverter is essentially a step response of charging or discharging a capacitor with a current source of finite internal
impedance. The change of the output voltage in one branch can be written as below:

$$\Delta V_{O1,2}(t) = \pm (R \cdot I_0) \cdot (1 - e^{-t/RC})$$  \hspace{1cm} (2.10)

Where $R$ is the load resistance and $C$ is the load capacitance of the first inverter. Fast switching only relies on small $RC$. However, to maintain the required voltage swing, the tail current $I_0$ has to increase. Additionally, the speed of differential pairs can be enhanced by using inductive peaking technique.

The power consumption of a CMOS CML inverter to the first order can be estimated easily.

$$P_{\text{diss}} = V_{DD} \cdot I_0$$  \hspace{1cm} (2.11)

Obviously CML inverter consumes static power. However, on the basis of the first order estimation, the power consumption is independent on frequency. Therefore, CML is suitable for high frequency applications in terms of speed and power consumption.

### 2.3.3 Circuit Logic Choice

Based on above discussions, CML should be the best choice for low-power high-speed circuit design, which will allow us using various bandwidth-boosting techniques and power-saving topologies to realize low-power high-speed circuits.

### 2.4 Chapter Summary

This chapter reviewed semiconductor technology background and circuit logic, and identified CMOS CML as the best candidate to realize low-power high-speed circuits. So the literature overview in Chapter 3 will focus on CMOS CML based circuit topologies and design techniques.
Chapter 3

Low-Power High-Speed Techniques

3.1 Chapter Overview

To identify feasible design strategies and select suitable circuit techniques for low-power high-speed IC design, existing low-power high-speed circuit topologies, techniques, and design bottlenecks will be reviewed, some of them will be verified by transistor-level simulations based on previously reported circuits and techniques in this chapter.

3.2 Overview of High-Speed Amplifiers

In this part, various high-speed amplifiers and bandwidth boosting techniques are reviewed and compared to select those most effective design strategies and circuit techniques for low-power high-speed circuit implementations.

3.2.1 Resistor-Loaded CS Amplifier

As shown in Figure 3.1, a common source (CS) amplifier with Unsilicided Poly resistor load is the fastest non-enhanced amplifier with the following advantages: Unsilicided poly is a pretty efficient current provider (good current to capacitance ratio); Output
swing can go all the way up to $V_{DD}$; Allows following stage to achieve high $f_T$; Linear settling behavior (in contrast to NMOS load). On the other hand, it has several undesired limitations, which can be interpreted by the following equations:

$$g_{m1} = \frac{dI_d}{dV_{gs}} = \frac{2 \cdot I_d}{V_{GS} - V_T} \quad (3.1)$$

$$A_v = g_{m1} \cdot R_L = \frac{2 \cdot I_d \cdot R_L}{V_{GS} - V_T} = \frac{2 \cdot V_{RL}}{V_{GS} - V_T} \quad (3.2)$$

$$A_{MAX} = \frac{2 \cdot V_{DD}}{V_{GS} - V_T} \quad (3.3)$$

$$f_{-3dB} = \frac{1}{2\pi \cdot R_L \cdot C_{tot}} \quad (3.4)$$

$$GBW = \frac{g_{m1}}{2\pi \cdot C_{tot}} \quad (3.5)$$

$$C_{tot} = C_{db1} + \frac{C_{RL}}{2} + C_{gs2} + K \cdot C_{OV2} + C_{fixed} \quad (3.6)$$

Where $K = 1 + |A_v|$ is Miller multiplication factor. Obviously, high $V_{GS} - V_T$ is required for high bandwidth, but this reduces gain. With low $V_{DD}$, the gain is very limited. Typically implementation is fully differential pair benefiting from self-biased topology and good common-mode rejection but consumes more power than single-ended version, as shown in Figure 3.2.
Figure 3.2: Resistor-loaded CS differential amplifier.

Figure 3.3: NMOS-loaded CS amplifier and its gain-frequency characteristic.

3.2.2 NMOS-Load CS Amplifier

As illustrated in Figure 3.3, NMOS-loaded CS amplifier is very simple and its NMOS active load does not suffer from serious process fluctuations, while Poly resistor loaded CS amplifier does. Its performance can be expressed by the following equations:

\[ g_{m1} = \frac{dI_d}{dV_{gs}} = \frac{2 \cdot I_d}{V_{in} - V_T} \]  (3.7)

\[ g_{m2} = \frac{dI_d}{dV_{gs}} = \frac{2 \cdot I_d}{V_{dd} - V_{out} - V_T} \]  (3.8)

\[ A_v = \frac{g_{m1}}{g_{m1}} = \sqrt{\frac{W_1/L_1}{W_2/L_2}} \]  (3.9)

\[ f_{-3dB} = \frac{g_{m2}}{2\pi \cdot C_{tot}} \]  (3.10)
\[ GBW = \frac{g_m}{2\pi \cdot C_{tot}} \]  
\[ C_{tot} = C_{db1} + C_{sb2} + C_{gs2} + K \cdot C_{OV3} + C_{fixed} \]

Where K is Miller multiplication factor. From Eq.3.7-3.12, we can choose the minimum channel length (L) for maximum speed while choose ratio of \( W_1/W_2 \) to achieve appropriate gain. But the problems are that \( V_T \) of \( M_2 \) lowers the bias voltage of the next stage (thus lowering its achievable \( f_T \)) and the performance of this amplifier will severely hampered when it is cascaded.

![Diagram](image)

**Figure 3.4:** CS-CG cascode amplifier.

### 3.2.3 CS-CG Cascode Amplifier (Reduction of \( C_{gd} \) Impact)

Main performance limitations of the amplifier illustrated in Figure 3.4 are:

- The cascode device lowers the gain seen by \( C_{gd} \) of \( M_1 \)
- Cascoding lowers achievable voltage swing

Based on the simulation data in Figure 3.5, CS-CG cascode amplifier has a higher DC gain and smaller -3dB bandwidth than resistor loaded amplifier; Transient simulation shows that CS-CG cascode amplifier has a reduced voltage amplitude and worse rise, fall edges. So we can say that this type of amplifier is ill-suited for high speed applications.
3.2.4 CS Amplifier with Neutralization

CS amplifier with neutralization is proposed to cancel the effect of $C_{gd}$ as shown in Figure 3.6. Choosing $C_N = C_{gd}$, so the charging of $C_{gd}$ is provided by $C_N$. The benefit is that the impact of $C_{gd}$ was removed, so $Z_{in} = 1/(s \cdot C_{gs})$. In fact, the impact of $C_{gd}$ cannot be completely removed if $C_N$ is not precisely matched to $C_{gd}$:

$$C_{in} = C_{gs} + (1 + |A_v|) \cdot C_{gd} + (1 - |A_v|) \cdot C_N \tag{3.13}$$

Since the neutralization does not completely remove the effect of $C_{gd}$, we can make $C_N$ slightly larger than $C_{gd}$ to "over neutralize". However, over neutralization can
reduce the effect of $C_{gs}$, but if $C_N$ is too large, the input capacitance is negative and can compromise stability. In addition, at high frequencies, this can lead to inductive input impedance. In practice, leverage differential signals can be used to create an inverted signal and $C_N$ should be matched to $C_{gd}$, which can be implemented by using lateral metal caps or CMOS transistor. If $C_N$ is too low, residual influence of $C_{gd}$ still functions. If $C_N$ is too high, input impedance has inductive component, which will cause undesired peaking in frequency response. Acceptable level of peaking is often evaluated using eye diagrams.

### 3.2.5 Shunt-peaked CS Amplifier

As shown in Figure 3.7, an inductor is used in load to extend bandwidth, which is often implemented as a spiral inductor or an active inductor shown in Figure 3.8.

**Figure 3.7:** Shunt peaking amplifier.

**Figure 3.8:** Spiral inductor model and active inductor model.
We can view the impact of inductor in both time and frequency domains: peaking of frequency response and delay of charging current in $R_L$. To determine the principles of parameter optimization, the expression for gain can be written as:

$$A_v = g_m Z_{out} = g_m \left[ (s \cdot L_d + R_L) \| \left( \frac{1}{s \cdot C_{tot}} \right) \right] = \frac{(g_m R_L)[s \cdot (L_d/R_L) + 1]}{s^2 \cdot L_d C_{tot} + s \cdot R_L C_{tot} + 1} \quad (3.14)$$

Let $m = R_L \cdot C_{tot}/\tau$, and $\tau = L_d/R_L$, which corresponds to ratio of RC to LR time constants, and the parameterized gain can be expressed as:

$$A_v = g_m \cdot R_L \cdot \frac{s \cdot \tau + 1}{s^2 \cdot \tau^2 \cdot m + s \cdot \tau \cdot m + 1} \quad (3.15)$$

Compare new and old -3dB frequencies and set $s = j\omega$:

$$\omega_1 = \frac{1}{R \cdot C} \quad (3.16)$$

$$\tau = \frac{1}{m \cdot \omega_1} \quad (3.17)$$

$$|A_v| = g_m \cdot R_L \cdot \frac{j\omega/m\omega_1 + 1}{-(\omega/m\omega_1)^2 \cdot m + (j\omega/m\omega_1) \cdot m + 1} \quad (3.18)$$

Define $\omega_2$ as new -3dB frequency, $\omega_1$ note that is old one. Thus,

$$\frac{j\omega/m\omega_1 + 1}{-(\omega/m\omega_1)^2 \cdot m + (j\omega/m\omega_1) \cdot m + 1} = \frac{1}{\sqrt{2}} \quad (3.19)$$

$\omega_2/\omega_1$ was solved after much algebra as:

$$\frac{\omega_2}{\omega_1} = \sqrt{\left(-\frac{m^2}{2} + m + 1\right)} + \sqrt{\left(-\frac{m^2}{2} + m + 1\right)^2 + m^2} \quad (3.20)$$

We can see that $m$ directly sets the amount of bandwidth extension. Once $m$ is
chosen, inductor value is determined as below:

\[
L_d = \frac{R_L^2 \cdot C_{tot}}{m}
\]  \hspace{1cm} (3.21)

From Figure 3.9, we can see that the highest extension is \(\omega_2/\omega_1 = 1.85\) at \(m = 1.414\) (However, peaking occurs!), the maximally flat response occurs at \(m = 2.41\) (extension=1.72), the best phase response appears at \(m = 3.1\) (extension=1.6), and no peaking occurs at \(m = \infty\) only. Generally, eye diagrams are often used to evaluate best \(m\).

![Figure 3.9: Bandwidth extension and normalized gain versus m.](image)

![Figure 3.10: Zero-peaked common source amplifier.](image)
3.2.6 Zero-peaked CS Amplifier

Based on Figure 3.10, the performance of Zero-peaked common source amplifier can be expressed by the following equations:

\[ A_v = \frac{g_m \cdot R_L}{1 + g_m \cdot R_s} \quad (3.22) \]

\[ \omega_0 = \frac{1}{2\pi \cdot C_s \cdot R_s} \quad (3.23) \]

\[ \omega_p = \frac{1}{2\pi \cdot C_{tot} \cdot R_L} \quad (3.24) \]

Basically, inductors are expensive with respect to layout area. So we can instead achieve bandwidth extension with capacitor according to the idea of degenerating gain at low frequencies while removing degeneration at higher frequencies (i.e., create a zero). Unfortunately, we have to increase \( R_L \) to keep same gain (lowers pole), which results in lower achievable gate voltage bias (lowers device \( f_T \)). Based on the simulation results in Figure 3.11, there is just a little bit improvement of gain-bandwidth product in high frequency end with the price of lower DC gain, which means that this type of gain cells are not suitable for wideband circuit design.

![Figure 3.11](image.png)

**Figure 3.11:** Comparison between zero-peaked and resistor-loaded common-source amplifier: (a) AC response; (b) Transient response.
3.2.7 Bandwidth Enhancement with $f_T$ Doublers

A MOS transistor has $f_T$ [14] calculated as:

$$2\pi \cdot f_T = \frac{g_m}{C_{gs} + C_{gd}} \approx \frac{g_m}{C_{gs}}$$  \hspace{1cm} (3.25)

The operating principle of a $f_T$ doubler amplifier is to double effective $f_T$ by increasing the ratio of transconductance to capacitance. So we can make the argument that differential amplifiers are $f_T$ doublers as shown in Figure 3.12. And capacitance seen by $V_{in}$ for single-ended input is $C_{gs}/2$. A MOS transistor has $f_T$ calculated as:

$$i_1 - i_2 = \frac{V_{in}}{2} \cdot g_m - \left( -\frac{V_{in}}{2} \right) \cdot g_m = V_{in} \cdot g_m$$  \hspace{1cm} (3.26)
According to Eq. 3.26, transconductance to capacitance ratio is doubled as $2g_m/C_{gs}$. Input voltage is again dropped across two transistors. The ratio given by voltage divider in capacitance ideally is $1/2$ of input voltage on $C_{gs}$ of each device; Input voltage source sees the series combination of the capacitances of each device ideally $1/2$ of the $C_{gs}$ of M1 and currents of each device add to ideally yield ratio: $2g_m/C_{gs}$.

As shown in Figure 3.13, current mirror can be used for bias, which is inspired by bipolar circuits. $V_{bias}$ should be properly set such that current through input transistor has the desired current of $I_{bias}$. Thus, the current through current source will ideally match that of input transistor. However, achievable bias voltage across above transistors is severely reduced (thereby reducing effective $f_T$ of device), which means it does not favor low supply technologies.

\[ \frac{V_{out}}{V_{in}} = \left( \frac{A}{1 + s \cdot \omega_0^{-1}} \right)^n = A^n \cdot \frac{1}{(1 + s \cdot \omega_0^{-1})^n} \]  

Figure 3.14: Block diagram of cascading gain stages.

### 3.2.8 Broadband Amplifiers Using Cascading Gain Stages

In general, we can significantly increase the gain of an amplifier by cascading $n$ stages shown in Figure 3.14, and the total gain [14] can be written as:
As we all known, the bandwidth of each stage will degrade, so we need to figure out how much bandwidth of each stage will shrink. A -3dB is denoted by \( \frac{1}{\sqrt{2}} \) amplitude drop for \( \frac{V_{out}}{V_{in}} \) in Eq.3.27, so now let

\[
\frac{V_{out}}{V_{in}} = \left| \frac{A}{1 + s \cdot \omega_0^{-1}} \right|^n = \frac{A^n}{\sqrt{2}}
\]

\[
\left[ \frac{A}{\sqrt{1 + (\omega_1 \cdot \omega_0^{-1})^2}} \right]^n = \frac{A^n}{\sqrt{2}}
\]

\[
\left[ 1 + (\omega_1 \cdot \omega_0^{-1})^2 \right]^n = 2
\]

\[
\omega_1 = \omega_0 \cdot \sqrt{2^{1/n}} - 1 \quad \text{(i.e. } n = 2, \omega_1 = \omega_0 \cdot \sqrt{2} - 1 \approx 0.63 \cdot \omega_0)\]

Where, \( \omega_1 \) is the overall bandwidth, \( A \) and \( \omega_0 \) are the gain and bandwidth of each section. From Eq.3.31, bandwidth decreases much slower than gain increases, and overall gain bandwidth product (GBW) of amplifier can be increased. The transfer function for cascaded sections is given in Eq.3.32 and plotted in Figure 3.15.

\[
H(f) = \left| \frac{1}{1 + j \cdot 2\pi \cdot f} \right|^n
\]

So far, the key issue is to choose the optimal number of stages for specific design. To first order, there is a constant gain-bandwidth product for each stage:

\[
\omega_0 = \frac{\omega_T}{A}
\]

On one hand, increasing the bandwidth of each stage requires that we lower its gain; On the other hand, we can make up for lost gain by cascading more stages. We found
that the overall bandwidth is calculated as:

$$\omega_1 = \omega_0 \cdot \sqrt{2^{1/n} - 1} = \frac{\omega_T}{A} \cdot \sqrt{2^{1/n} - 1}$$  \hspace{1cm} (3.34)

Assume that we want to achieve $G$ with $n$ stages, so $A = G^{1/n}$, $\omega_1 = \frac{\omega_T}{G^{1/n}} \sqrt{2^{1/n} - 1}$. From this, Tom Lee finds optimum gain $\approx 1.65$ [15]. Figure 3.15 shows the curve of achievable bandwidth versus $G$ and $n$, and we can see that gain per stage derived from plot is $A = G^{1/n}$ and maximum is fairly soft. Thus, we can dramatically lower power dissipation (and improve noise) by using larger gain per stage.

**Figure 3.15:** Normalized transfer function and achievable bandwidth.

Assume that we want to achieve $G$ with $n$ stages, so $A = G^{1/n}$, $\omega_1 = \frac{\omega_T}{G^{1/n}} \sqrt{2^{1/n} - 1}$. From this, Tom Lee finds optimum gain $\approx 1.65$ [15]. Figure 3.15 shows the curve of achievable bandwidth versus $G$ and $n$, and we can see that gain per stage derived from plot is $A = G^{1/n}$ and maximum is fairly soft. Thus, we can dramatically lower power dissipation (and improve noise) by using larger gain per stage.

**Figure 3.16:** Cascading gain stages and AC response.
**Table 3.1:** Gain, BW, and GBW of Cascading Gain Stages

<table>
<thead>
<tr>
<th>R-load CS amplifier</th>
<th>DC Gain</th>
<th>-3dB BW</th>
<th>GBW</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-stage</td>
<td>17dB</td>
<td>3.98GHz</td>
<td>66.7dB-GHz</td>
</tr>
<tr>
<td>4-stage</td>
<td>34dB</td>
<td>2.13GHz</td>
<td>96.1dB-GHz</td>
</tr>
<tr>
<td>6-stage</td>
<td>51.3dB</td>
<td>1.59GHz</td>
<td>112.5dB-GHz</td>
</tr>
</tbody>
</table>

The simulation results shown in Figure 3.16 (b) and Table 3.1 confirmed that we can design broadband amplifiers with a GBW much larger than the $f_T$ of used CMOS technologies.

![Typical Distributed amplifier](image)

**Figure 3.17:** Typical Distributed amplifier.

### 3.2.9 Distributed Amplifier (DA)

Distributed amplification is widely used as a circuit topology for achieving a flat gain and good input and output matching over a very large bandwidth. Some broadband amplifiers are implemented in CMOS for high data rate optical communication systems.

For a typical distributed amplifier shown in Figure 3.17, there is a basic trade-off between gain and bandwidth. We can achieve higher gain for a given load resistance by increasing the device size (i.e., increase $g_m$), but the increased capacitance lowers bandwidth. We therefore get a relatively constant gain-bandwidth product. On
the other hand, by lumping input capacitance into LC network corresponding to a transmission line, so that signal ideally sees a real impedance rather than an RC lowpass filter at the input. Based on such a lumped networks (i.e. T-coils), we can now trade delay (rather than bandwidth) for gain, which is why the operating speed of DA is not limited by unity-gain frequency of the individual stages.

Distributed amplifiers have very high bandwidth, but the negatives are high power, poor noise performance, and expensive in terms of chip area.

3.3 Overview of High-Speed Analog Circuits

Broadband amplifiers are widely used in data communication systems as gain stage, delay cell, buffer, and driver for various external loads such as laser diodes, optical modulators, sensors or alarm devices. In this part, some typical applications of broadband amplifiers will be reviewed to select feasible design strategies.

3.3.1 Broadband Output Buffers

Figure 3.18 (a) shows the 3-stage cascaded CML output buffer with inductor peaking. The inductor-peaking topology is previously presented and results in a \( \pi \)-network composed of \( C_1 \), \( C_2 \), and \( L_C \) between the stages in the small-signal equivalent circuit shown in Figure 3.18 (c). The fanout, \( m \), is 1 and \( C_1 \) and \( C_2 \) are equal. The resonance of the \( \pi \)-network leads to 1.8dB gain peak and 3.5 times larger bandwidth than that of CML without inductors. However, when the fan-out, \( m \), is 2, the increased capacitance of \( C_2 \) decreases the voltage gain of the circuit and thus, the bandwidth is also decreased. Large input/output swing requirement is another challenge in designing the buffer. The required voltage swing 0.6V\( _{PP} \) causes the input transistor pair to operate partially in the triode region, shown in Figure 3.18(b), introducing non-linearity. Therefore, the waveforms on the two output nodes become asymmetric.
Figure 3.18: Output buffer: (a) Circuit schematic. (b) Operating region of the input pair transistors. (c) Simplified equivalent circuit.

The frequency response for both the rising-step output and the falling-step output had to be optimized [1].

Another example reported in [16] is an open-drain driver for distributed select circuit as shown in Figure 3.19, which is used to drive the clock and data to the distributed select circuit. As compared to a conventional back-terminated driver, for the same voltage swing and device current density, this would require half the DC current and device sizes. This reduces the power consumption and improves the high-speed performance since the output loading conductance is halved. The only issue is the reflection from the source. This is not a problem for the data and clock input lines since the signals travel only in one direction—from the source to the load. The signal traveling to the right is absorbed by the matched on-chip termination. The situation is different for the output line, however. Since the input signals are applied at various locations along the output line, these signals will propagate in both directions. Thus,
the output transmission line must be doubly terminated.

It is easy to see that the reviewed buffers can be used in very high speed applications, but the design complexity and silicon area are not acceptable for low-cost implementations.

![Open-drain driver for distributed select circuit](image)

**Figure 3.19:** Open-drain driver for distributed select circuit.

![Quasi push-pull source follower and its equivalent circuit](image)

**Figure 3.20:** Quasi push-pull source follower and its equivalent circuit.

### 3.3.2 CMOS Driver for Laser Diode or Optical Modulator

A versatile driver for laser diode or optical modulator using a 0.35-μm CMOS is reported in [17], in which a quasi push-pull source follower is introduced.
Other drivers usually use conventional source followers for level shifting and impedance transformation. But due to the lack of sinking current, the conventional source follower causes an asymmetry of the rise edge and fall edge of the signal, and also lowers the driver’s slew rate. During the signal transition of falling edge, the sinking current decreases because of the mechanism of channel-length modulation. When the signal amplitude is very large, the amount of current is reduced considerably by channel-length modulation. It is the assignable cause of signal quality worsening. To solve this problem, a quasi push-pull source follower as depicted in Figure 3.20 is introduced, in which two cross-coupled transistors \( M_1, M_2 \) are added, and the small signal model of this follower is shown in the Figure 3.20, where the \( g_m' \) and \( C' \) are introduced by added transistor. \( C_L \) represents the total capacitance seen at the output except for \( C' \). (Here the body effect and source resistance are ignored, but we can still get some design insights). Then we have:

\[
\frac{V_o}{V_i} = \frac{g_m \cdot \left( 1 + s \cdot \frac{C_L}{g_m} \right)}{2 \cdot (g_m - g_m') \cdot \left( 1 + s \cdot \frac{C_g + C_L + C'}{g_m - g_m} \right)}
\]

(3.35)

We can see that this quasi push-pull source follower contributes some gain with a reduction of the bandwidth, which has important implication for the driver optimization. Since the follower’s bandwidth quite large, an appropriate bandwidth reduction will have no strong effect on the driver’s overall performance. With the gain contributed by the source follower, we can relax the requirement for the differential amplifiers’ gain, and thus have more freedom to balance overall performance of the driver. Combined with a dynamic amplification technique, the driver’s slew rate and output swing are increased; Meanwhile the overshoot is efficiently reduced. The driver works under 5V supply to give high enough voltage swing across the 50Ω load. However, the used source follower consumes much current, which is undesired for the use of laser diode driver, in which device cooling and automatic power control are
3.3.3 High-Speed Equalizer Filter

Shown in Figure 3.21, this equalizer intersperses three peaking stages with two gain stages to provide a boost factor of about 22 dB at 5 GHz while exhibiting a low-frequency loss of less than 3dB. The design exploits the modified reverse scaling technique to allow optimization for high-frequency peaking and low-frequency loss.

As mentioned above, simple resistively-loaded differential pairs cannot yield the required bandwidth. Thus, inductive peaking and negative Miller capacitances have been added to improve the speed without sacrificing the voltage headroom. To save area, only three of the stages incorporate inductive peaking.

The peaking stages in the equalizer path employ a variable degeneration resistance along with MOS varactors $M_4$ and $M_5$ to provide a wide boost range. As the control voltage rises, the on-resistance of $M_3$ falls and so does the capacitance of $M_4$ and $M_5$, raising the magnitude of the zero. Note that the simultaneous change of the

![Figure 3.21: Schematic of equalizer filter.](image)
resistance and capacitance greatly simplifies the adaptation loop.

The cascade employs capacitive coupling between some stages to isolate common-mode (CM) levels. This mitigates the voltage headroom issue and, more importantly, avoids variability in the CM level seen by - due to the preceding stage, thus maintaining a constant tuning range. The capacitors 0.25pF are realized using multi-finger fringe structures having a parasitic component of about 3%. The CM level is generated using a resistive divider. The corner frequency associated with this capacitive coupling is around 3MHz, resulting in negligible droop with encoded data.

This equalizer described does not incorporate a passive network at the input, providing greater sensitivity but consuming a higher power, which means that some modifications are required for low power applications [18].

3.4 Overview of High-Speed Digital Circuits

3.4.1 Monolithic Transformer Coupled 2:1 MUX

The monolithic transformer coupled 2:1 MUX IC (Figure 3.22) consists of a master-slave flip-flop (MS-FF), a master-slave-master flip-flop (MSM-FF) and the 2:1 multiplexer (MUX) [19]. The inputs of the MUX IC are two in-phase differential 15-Gb/s data signals, D1 and D2. Full-rate data acquisition is done by the MS-FF and MSM-FF (f_{CLK}=15GHz). The desired phase shift of 90° between the MUX 2:1 inputs is achieved by adding an extra latch in series to one path (MSM-FF). Finally, the data streams D1 and D2 are multiplexed by the MUX 2:1 to a 30-Gb/s data stream. The MUX uses no output buffer.

The MS-FF (Figure 3.22) consists of two latches connected in series. All NMOS transistor in the core are low-V_T 0.12µm NMOS devices. The latches use series gating between clock and data inputs. All data path transistors are of the same size and
Figure 3.22: Top-level diagram of MUX and schematic diagram of the MS-FF.

are 3/5 the width of the clock transistors. Poly-silicon resistors are used as load. One latch consumes 4mA at 1V supply. Clock input matching is realized with 100Ω on-chip resistors, which are connected to a DC level shifter ($V_{DD}/2$).

Figure 3.23: Schematic of the MUX stage with the monolithic transformer.

Figure 3.23 shows the schematic diagram of the monolithic transformer coupled MUX stage. The transformer splits the conventional CML-MUX design into the MUX-Core and the MUX-Clock section. The clock is the only signal in the MUX which is no broadband signal. We can take advantage of this fact and use a monolithic transformer to couple the clock signal from the MUX-Clock to the MUX-Core (Figure
There are outstanding advantages due to the on-chip transformer. Because of the missing DC path between primary and secondary side of the transformer the MUX Core and the MUX Clock section can use the full supply voltage. The effective supply voltage for the MUX 2:1 circuit is doubled. The supply voltage of MUX-Core and MUX-Clock is connected to the center taps of the monolithic transformer. In this circuit are only two gates in series while using the full supply voltage.

The input-transformer XI is connected as a parallel resonant device. The MOS capacitors C are connected in parallel to the primary windings of the transformer. The resonant tuning increases the current transfer ratio of the transformer. The cascode transistors T5 and T6 provide isolation between the data-path transistors T1-T4 and the transformers parasitic capacitances. All transistors of the MUX stage are NMOS devices because of their higher speed compared to PMOS transistors. Except for the current-source transistors low-$V_T$ devices with gate lengths of 120μm are used everywhere. The MUX stage uses 70Ω poly-silicon load resistors, which is a compromise between high voltage swing and reasonable output matching. The tail current is set to 5mA. The DC level of the sinusoidal clock signal is $V_{DD}/2$.

As described above, the advantages of this MUX stage are summarized as below:

- Using 70Ω poly-silicon load resistors which is a compromise between high voltage swing and reasonable output matching allow to drive 50Ω external load without output buffer, so much power are saved

- Using on-chip transformer allows sub-1V supply because only two transistors are stacked vertically. On the other hand, both data path and clock path transistors operate under relatively larger $V_{DS}$, which means higher effective $V_T$ can be achieved even if using a low supply

- Smaller clock and data device pairs can be selected to reduce undesired parasitic
capacitance because high $V_{DS}$ and tail-to-tail input clock signal are available. As a result, the clock pair can switch faster and the MUX core can operate at a bit rate as high as 30-Gb/s just consuming a power of 28mW.

However, this MUX also have the following drawbacks:

- The effective output signal swing is not high due to the lack of output buffer.
- Since there is a resonant network in the MUX block, broadband application is impractical.
- The immunity to PVT variation could be poor due to the use of low supply and on-chip transformer which required accurate models.

3.4.2 High-Speed Frequency Divider

A high-speed frequency divider is one of the key elements in digital communication systems. In fact, a multiplexer (MUX)/demultiplexer (DEMUX) is indispensable in such systems, and the operating speed of MUX and DEMUX is mainly limited by the operating frequency of the 2:1 divider because the divider operates at the highest speed in MUX and DEMUX circuits. The frequency divider is also widely used to generate a precision I/Q signal if the input signal has a 50% duty cycle, for the modern in-phase and quadrature (I/Q) modulator or demodulator. For the signal with duty cycle other than 50%, an additional divide by-2 can be used to generate the 50% duty cycle [20].

Using CMOS technology for high-speed frequency dividers is an effective way to reduce the power consumption. When such high-speed circuits, like MUX and DEMUX, are made using CMOS technology, these circuits can be integrated with large CMOS logic LSIs in the same chip. This leads to lower power consumption,
smaller modules, higher mounting density in communication equipment, and lower cost.

A frequency divider circuit takes a periodic input signal and generates a periodic output signal at a frequency that is a fraction of the input signal. The input waveform can be either analog or digital. Analog divide-by-two frequency dividers operating on the principle of regenerative feedback have been demonstrated for frequencies well into the millimeter-wave range. Other analog divider implementations use the mechanism of injection locking. Most digital frequency dividers can also operate with analog input signals. Some of the more mature CMOS digital dividers use cross-coupled toggle flip-flops in a feedback configuration implemented with D-latches. These are called "static" frequency dividers. With this approach, a division ratio of 2 is obtained. A diverse number of dividers use the D-latch approach have been suggested and the differences between them relate to how the latches are actually implemented. In contrast to the static dividers, circuits using dynamic CMOS logic have also been suggested. A different method for achieving frequency division in the analog/digital domain consists of a chain of inverters. This mechanism is used to implement a basic divide-by-two and a divide-by-four system.

**Figure 3.24:** (a) Divide-by-2 SCL; (b) Logic diagram of toggle switch.
A. Digital Logic Approaches for Frequency Divider

Generally speaking, it is much easier to understand and analyze the digital approach in the digital domain and the analog approach in the analog domain. Within the digital domain, the design strategy can be further divided into two categories: static logic and dynamic logic.

The static implementation is the most popular approach. The memory cell is a true bistable circuit, unlike the parasitic capacitor used in the dynamic approach. One standard design is the divider-by-2 cell shown in Figure 3.24.

As shown in Figure 3.24, it is a logic gate for the toggle switch, which is essentially an edge-triggered master/slave D flip-flop (DFF). The inverted output is fed back to the input port D. The same clock is used to drive both level-triggered DFF with opposite logic. The reason for the inverter is to make an edge-triggered DFF out of the two level-triggered DFF. The first DFF is commonly called the master DFF and the second one is normally referred to as the slave DFF. Either master DFF or slave DFF is activated in each clock cycle, not both DFF at the same time (because of the inverter between them). In Figure 3.24, each positive input clock cycle is loaded into the DFF. On the next cycle, inverted output again is fed back to the input, which causes the output to toggle. It is why toggle DFF is a more descriptive name for this circuit. The same event repeats for every two input clock cycle. Thus, output frequency is half of the input frequency.

B. Dynamic DFF Implementation

In dynamic DFF implementation, there is no dedicated bistable circuitry. The parasitic cap between the node acts as the storage element. It is called Clocked CMOS ($C^2MOS$). One such circuit is shown in Figure 3.25 (a).

This circuit uses a far less number of transistors. The theory of operation is simple.
Figure 3.25: (a) Dynamic DFF implementation; (b) TSPC frequency divider.

Transistors $M_{N1}$ to $M_{P1}$ complete essentially a tri-state inverter (INV1). Transistors $M_{N3-5}$ and $M_{P3-4}$ form another inverter (INV2). The capacitor in the middle is a model for the parasitic capacitance between the gates. The capacitor’s responsibility is to store the signal. $M_{N5}$ and $M_{P6}$ make a simple CMOS inverter to complete the feedback path needed for the toggle latch. In the positive clock cycle ($f_{in+}$), INV1 is on and INV2 is off, so the signal is clocked into the storage capacitor. In the negative clock cycle ($f_{in-}$), INV1 is off and INV2 is on, the signal is clocked out. There are many different flavors using a similar theory of operation.

The circuit in Figure 3.25 (a) requires the complementary clock input ($f_{in+}$ and $f_{in-}$). Sometimes, it is more desirable to drive the frequency divider single ended. The type of logic is called true single-phase clocked (TSPC) logic. It is built on the basic $C^2MOS$. It eliminates the differential drive requirement at the expense of more transistors. One such circuit is shown in Figure 3.25 (b).

Transistors $M_{N1-4}$ and $M_{P1-2}$ form the master latches. Transistors $M_{N5-6}$ and $M_{P3-6}$ form the slave latches. $M_{N7}$ and $M_{P7}$ are simply the inverter to complete the feedback path needed in the toggle latch. The master latch is sometimes called double $CN^2MOS$ because two NFETs are needed in the PDN. Similarly, the slave latch is sometimes referred to as the double $PC^2MOS$ because two PFETs are needed in the PUN. When the clock is high, the master latch is activated while the slave latch is
off. The opposite is true when the clock is low. The circuit shown is Figure 3.25 (b) is the most basic form. There are more variations and clever designs used to reduce the number of the transistors.

Comparing the static and dynamic implementation, static logic is more reliable. The dynamic logic uses a far fewer number of transistors and is easier to implement. However, it could consume more power because full CMOS swing is needed in certain applications.

In summary, frequency dividers are an interesting and indispensable building blocks in various communication systems. There are two main design approaches: digital logic and analog using injection-lock techniques. The digital logic approach can be further divided as the static DFF and dynamic DFF. Like any other receiver design, design trade-offs required depend on the application. In applications from around 40GHz to 100GHz, an ILFD is a suitable approach because today's standard CMOS process still doesn't have the bandwidth in that range. From 10GHz to 40GHz, successful CML-based frequency divides have been reported.

**C. Reported CMOS Frequency Dividers**

To gain more insights about frequency dividers, CMOS logic static 2:1 frequency divider, and TSPC 2:1 frequency divider will be reviewed and transistor-level simulations based on IBM 0.13µm CMOS will be carried out in this part.

\[\text{§1 Dynamic 2 : 1 Frequency Divider in CMOS Logic}\]

Figure 3.26 (a) shows the circuit diagram of a typical dynamic 2:1 frequency divider. The core of the divider (shaded area) consists of three CMOS inverters and one CMOS transmission gate. The modified divider has circuit topology in which one transmission gate (TG2) is removed from a conventional one, which is shown in Figure 3.26 (b) [21]. This leads to improved speed performance in two ways.
First, the delay of the critical path is reduced. The operating frequency of the conventional divider is limited by the delay from node “c” through nodes “d”, “e”, “a”, and “b” to node “c”, which is composed of the delays of the three inverters and two transmission gates. In the modified divider, the critical-path delay, which is from node “a” through nodes “b”, “c”, and “d” to node “a”, is reduced by the delay of one transmission gate. This effect essentially increases the speed performance of the modified divider by about 20% compared with the conventional one, because the critical paths of the modified version and the conventional one are composed of four...
gates and five gates, respectively.

Second, the load of the input clock in the modified divider is reduced to half that in the conventional one. The input clock drives only one transmission gate (TG1) in the modified divider. Thus, the slope of the input clock waveform becomes steep. This enables the clock to operate at higher frequencies and decreases the delay of the transmission gate because gate delay decreases as the slope of the input waveform steepens. When the slope steepens double, the gate delay decreases by 10% to 20%. As a result, due to the first and second effects, the total speed improvement of over 30% can be estimated. No DC current, except for off-leak current, flows in the modified divider, in the same way as in the conventional one in Figure 3.26 (a).

Input clock $CK_{in}$ is input to an inverter for inverted clock $CKN$ and a transmission gate for noninverted clock $CK$. $CK$ is delayed by the transmission gate in order to reduce the skew between $CK$ and $CKN$. This divider operates like a ring oscillator and has the maximum and minimum input frequencies. The output frequency is controlled by turning the transmission gate (TG1) on and off by the input clock. The timing chart of the proposed 2:1 divider is illustrated in Figure 3.26 (c). When $CK$ becomes low, TG1 turns on and then the voltage level of node "b" changes to the same voltage of node "a" with the propagation delay of one transmission gate ($t_{tg}$). The transition in node "b" is propagated through nodes "c" and "d", and returns to node "a" continuously, with the delay of three inverters ($t_{iv} \times 3$). For successful 2:1 divider operation, the voltage level of node "a" has to change while TG1 is off. This determines the maximum and minimum input frequencies of the proposed divider.

After the voltage level of node "a" changes, the transition does not propagate to next node "b" until TG1 becomes on. When $CK$ becomes low and TG1 turns on again, the propagation of the transition through nodes "b", "c", and "d" to node "a" is repeated. Consequently, triggered by the falling edge of $CK$, the input clock is divided by two. Of course, the divider triggered by the rising edge can be designed
easily.

Figure 3.27: TSPC +2 implementation and the third stage showing capacitances.

§2 TSPC Dynamic Frequency Divider

The divide-by-2 circuit realized in the TSPC logic is shown in Figure 3.27. The salient feature of the TSPC clocking technique is that there is only one clock signal needed to trigger the flip-flops and no extra clock phase is required whatsoever. This technique is mainly used in dynamic CMOS circuits and helps to simplify the design.

The circuit consists of three parts. The first part is a gated inverter that consists of $M_{P1}$, $M_{P4}$ and $M_{N1}$, which passes the divider output to the following stage when “$f_{in}$” goes low. The second part is a latch stage that consists of $M_{P2}$, $M_{P3}$, $M_{N2}$, $M_{N3}$, $M_{N4}$ and $M_{N5}$. This circuit will be activated and store the output of the gated inverter when “$f_{in}$” is high. The PMOS transistors $M_{P1}$ and $M_{P2}$ are used to pre-charge the internal nodes to increase the speed of the circuit. The output of the flip-flop is directly connected back to the D-input to obtain the divide-by-2 function because the TSPC circuit can completely isolate the sense and latch stage at different phases of the clock signal. The static power of the circuit is zero because no direct path from supply to ground exists and it only consumes dynamic power. One of the advantages of the TSPC divider is its simplicity. The circuit consists of
only nine transistors. In some cases where the inverted output is required an inverter is included which adds two more transistors. The circuit requires an input signal of large amplitude, and it is very sensitive to the slope of the signal. Therefore, a high-frequency input buffer may sometimes be needed in front of the divider to drive it. The speed of the circuit greatly depends on the voltage supply. The circuit will be slow if a low-voltage supply is used. In order to operate at higher frequency, larger sizes of the transistors are needed to increase the $g_m$ and thus make it operate faster. However, increasing the size also increases loading for the previous stage and thus the trade-off should be considered during design. Moreover, using larger transistor sizes will increase the degree of charge leakage and charge sharing at the output nodes and thus will affect the minimum operation frequency of the circuit.

Consider that this circuit is divided into three stages, the first made up of transistors $M_{P1}$, $M_{P4}$ and $M_{N1}$, the second made up of $M_{P2}$, $M_{N2}$ and $M_{N3}$ and the third made up of $M_{P3}$, $M_{N4}$ and $M_{N5}$. The output of this divider has to drive both the input of the following stage and an inverter in its own first stage, both of which appear to the output simply as capacitive loads. Besides these capacitive loads there are other capacitances that should be considered like the interconnect capacitance and the parasitic capacitances of the output transistors.

To begin designing, a frequency of operation $f_{out}$ is assumed for the circuit. From this, the approximate switching time is obtained as, $\tau_{sw} = 1/f_{out}$. Switching time is made up of rise time $\tau_{LH}$ and fall time $\tau_{HL}$, and $\tau_{sw} = \tau_{LH} + \tau_{HL}$ can also be written as $\tau = R \cdot C$.

Consider the third stage of the divide-by-2 circuit made up of transistors $M_{P3}$, $M_{N4}$ and $M_{N5}$ as shown in Figure 3.27. The figure also shows the various capacitances to be considered while designing this stage. The sum of parasitic capacitances associated with the output transistors $M_{P3}$ and $M_{N4}$ is denoted as $C_0$. The equations for $C_0$, $C_L$ and $C_x$ are written as: $C_0 = C_{gd,M_{P3}} + C_{db,M_{P3}} + C_{gd,M_{N4}} + C_{db,M_{N4}}$. 
where \( C_{gd} \) is the gate-to-drain capacitance and \( C_{db} \) is the drain-to-bulk capacitance; 
\[
C_L = C_{L1} + C_{L2} + C_{\text{int}},
\]
where \( C_L \) represents total load capacitance, \( C_{L1} \) is the input capacitance looking into the next stage which is given as 
\[
C_{L1} = C_{ox} \cdot W \cdot L
\]
(here \( W \) and \( L \) denote the gate width and length respectively, of the transistor being driven in the next stage, \( C_{ox} \) \( (C_{ox} = \varepsilon_{ox}/t_{ox}) \) is the gate oxide capacitance, \( C_{L2} \) is the capacitance looking into the inverter in the first stage of this same circuit (refer Figure 3.27), 
\[
C_{L2} = (C_{ox} \cdot W \cdot L)_{M_{P4}} + (C_{ox} \cdot W \cdot L)_{M_{N1}}
\]
and \( C_{\text{int}} \) is the interconnect capacitance (the parasitic capacitance associated with the metal layer forming the connection and the substrate); 
\[
C_x = C_{gd,M_{N5}} + C_{gs,M_{N5}} + C_{gs,M_{N4}} + C_{sb,M_{N4}}
\]
where \( C_{gs} \) is the gate-to-source capacitance and \( C_{sb} \) is the source-to-bulk capacitance.

\[\text{Figure 3.28: Dynamic CMOS logic divider operating at 6.5GHz and 14GHz.}\]

\[\text{D. Transistor-Level Simulations}\]

Based on IBM 0.13\( \mu \)m CMOS technology, both dynamic CMOS logic frequency divider and TSPC frequency divider are simulated. As shown in Figure 3.28 and 3.29, the modified dynamic frequency divider in CMOS logic can operate from 6.5GHz to 14GHz with a power consumption of about 120\( \mu \)W. The duty cycle is almost 50% in the entire operating frequency range.

As shown in Figure 3.30 and 3.31, the overviewed TSPC dynamic frequency divider
can operate from 5GHz to 18GHz with a power consumption of about 200μW. The duty cycle is not 50% and moves farther away from 50% with the increase of input signal frequency.

Obviously, both dynamic frequency divider in CMOS logic and TSPC dynamic frequency divider are suitable for super low-power applications, but neither of them is able to operate at 40GHz and above.
E. Summary of CMOS Frequency Divider

In this part, the application specification, operating principle, implementation approach of CMOS frequency divider were reviewed, previously reported divider topologies were analyzed, and transistor-level simulations were also carried out. It turned out that it is very difficult to design very high speed frequency divider with low power, low complexity, and high reliability. Undoubtedly, advanced design strategies and techniques have to be identified and used to result in a trade-off among speed, power, complexity, and reliability.

3.5 Chapter Summary

This chapter reviewed low-power high-speed circuit topologies, techniques, and design bottlenecks, which is the foundation and motivation to identify advanced design strategies and circuit techniques for low-power high-speed circuit implementations in next chapter.
Chapter 4

Advanced Design Strategies and Circuit Techniques for Low-Power High-Speed Circuit Implementations

4.1 Chapter Overview

Based on the extensive review of low-power high-speed circuit techniques and design bottlenecks in Chapter 3, this chapter will put emphasis on advanced design fundamentals, strategies and circuit techniques. Systematic analysis and comparison of the trade-offs for the techniques reported in [1-36] result in the selection of most feasible and effective design strategies and techniques such as active inductors, capacitors, resistors, negative feedback topology, split-resistor low capacitive loads, device sizing and DC bias level optimizing for low-power low-supply high-speed circuit implementations.

4.2 Design Fundamentals

In this section, the natural advantages of CML and the features of available advanced active and passive devices are addressed to show the feasibility of the design strategies
and circuit techniques to be selected.

4.2.1 Current-Mode Logic

Based on Chapter 2, CML rather than CMOS rail-ro-rail logic was chosen for low-power low-supply high-speed ICs design. Moreover, CML-based circuits also can benefit from the following advantages:

- Well-controlled signal swing (Same current, much smaller signal swing and RC time constant, higher speed)

- Low-voltage requirement (Allow low supply design benefiting from available low-$V_T$ transistor without suffering from relative larger leakage)

- Lower power dissipation (Better for higher speed operation since $P = I \cdot V_{DD}$ rather than $P = k \cdot C \cdot f \cdot V_{DD}^2$)

- Superior signal integrity (Canceled common-mode noise, higher immunity to supply fluctuation and ground bounce, lower switching and substrate noise due to smaller signal amplitude and nearly constant current)

- Small propagation delay (Lower critical node capacitance, smaller signal amplitude)

- Improved current source (Higher PVT-fluctuation immunity and better reliability)

More importantly, CML is very friendly to various circuit techniques such as cascading multiple-stage topology, active feedback, active inductor and spiral inductor shunt peaking, low capacitive load, feasibility for device sizing and DC bias optimizing. So, on the basis of CML, low-cost high performance ICs can be realized using the
advanced design strategies and combining multiple circuit techniques to lower supply voltage, power consumption, and silicon area, simultaneously.

![Graph](image_url)

**Figure 4.1:** $f_T$ versus $V_{GS}$, $V_{DS}$, and W/L: (a) (2$\mu$m×10)/0.12$\mu$m LVT and RVT devices; (b) (5$\mu$m×10)/0.12$\mu$m LVT device under various $V_{DS}$.

### 4.2.2 Advanced CMOS Components

#### A. NMOS Transistors

As addressed in Chapter 2, effective $f_T$, $V_{GS} - V_T$, output capacitance, and transistor feature size are major parameters to determine the circuit speed, supply voltage requirement, and power dissipation. Therefore, it is useful to gain more insights of the available CMOS devices.

As shown in Figure 4.1 (a), the used low-$V_T$ device (LVT, $V_T \approx 300$mV) has a higher $f_T$ than the used regular-$V_T$ (RVT, $V_T \approx 385$mV) device under the same bias, which means that low-$V_T$ device is a better candidate for low supply high speed design. Compared Figure 4.1 (a) and (b), it is clear that transistor with a larger aspect ratio (W/L) can approach higher effective $f_T$. In another word, sufficient effective $f_T$ can be maintained by increasing the transistor size while lowering the DC bias level. Therefore, it is very cheerful to make full use of the available multi-$V_T$
CMOS technology and feasible device sizing methodology for low-supply high-speed IC design.

Conservatively, a NMOS transistor can operate at the frequency of $f_T/4$. Thus the available 0.13μm Si CMOS technology with a maximum $f_T$ of about 120-GHz should be able to use in the ICs with upper limit frequencies of 30GHz. Further, 1-Hz frequency band can carry approximately 2-bit data, resulting in the highest bitrates of 60-Gb/s. However, the affordable effective $f_T$ is about 80GHz for low voltage design, which is illustrated in Figure 4.1 (b). At high speeds (40-Gb/s and above), both bipolar and FET digital circuits have to use CML topologies with multiple bandwidth boosting techniques. In addition, both the aspect ratio (W/L) and DC bias level have to be carefully chosen according to specific optimizing strategies.

**B. Topmost Metal Spiral Inductors**

Although on-chip inductors have lower quality factor (Q) than bonding wire inductors due to the limited conductivity of the metal, substrate loss, and parasitic capacitance, a high Q-value of the spiral inductors is not indispensable for most wide-bandwidth circuits because the effective Q is determined by the poly-silicon load resistors that are connected in series to the inductors. However, high self-resonance frequency is necessary for peaking inductors, which can be achieved by using only the two topmost metal layers to reduce capacitance to the substrate. Spacing between adjacent turns of the inductor is made wider than required by the design rules to reduce turn-to-turn capacitance. As shown in Figure 2.1, the available inductors have both sufficiently high self-resonance frequency and high quality factor at operating frequency, which is a guarantee for high speed circuit implementations.
4.3 Design Strategies and Circuit Techniques

Since it is possible and necessary to design economical high performance CMOS circuits, feasible design strategies and achievable design goals are determined as below:

- Further lower supply and power (Using lower supply voltage and both LVT and RVT NMOS to design low power dissipation ICs which is compatible with digital circuits in CMOS logic for control, calibration, and signal processing)
- Maximize effective $f_T$ (To meet the speed specification while reduce supply and power)
- Minimize silicon area (Using area-saving components and compact topologies to save silicon area and improve integration density)
- Improve circuit performance (Using optimized circuit architectures and multiple circuit techniques to boost bandwidth, improve input sensitivity and switching speed, and to achieve higher immunity against PVT fluctuations, fabrication reliability, and less complexity)

According to above design strategies, the following circuit techniques are selected, analyzed here, and will be used to design high-speed ICs addressed in Chapter 5.

4.3.1 DC Bias Level Optimizing

The proposed static frequency divider as shown in Figure 4.2 are used here as an example to illustrate the principles of DC bias level optimizing.

Peak $f_T$ values have been stuck at around 100GHz for the most advanced Si technologies. In fact, the achievable effective $f_T$ is much lower than 100GHz for D-latch-based circuits with a 1V-supply because there is no device in latches operating under $V_{GS} > 0.6V$ and $V_{DS} > 0.8V$. From Figure 4.1 (b), it is necessary to keep
$V_{GS} > 0.45V$ and $V_{DS} > 0.25V$ for devices in data and clock paths so as to maximize the effective $f_T$. From Figure 4.1 (a), using LVT devices in both data and clock paths would be much better than employing RVT transistors for low supply (1V) operations. Therefore, the DC bias level of data and clock transistor pairs can be optimized to meet corresponding requirements. In this example, $V_{GS} \approx 0.5V$ and $V_{DS} \approx 0.5V$ is chosen for devices in data path to achieve enough transconductance, $V_{GS} \approx 0.5V$ and $V_{DS} \approx 0.25V$ is selected for clock pair to achieve enough switching speed and to save voltage room for other transistors, $V_{GS} = 0.45V \sim 0.65V$ and $V_{DS} = 0.25V \sim 0.65V$ are adopted for differential pairs in output buffers.

![Figure 4.2: Proposed static frequency divider. (a) Divider core; (b) Output buffer.](image)

### 4.3.2 Device Aspect Ratio Optimizing

For high speed switch circuits, there is a basic trade-off between high input sensitivity and high switch speed limited by output capacitance and external loading capacitance. Transistors in data path of latches use much smaller width to reduce output node capacitance for fast switching. Larger clock devices are used to increase the input sensitivity while the relatively larger input capacitances are partly tuned out by input matching networks. LVT NMOS devices are used in high frequency parts, because of their higher speed compared to RVT NMOS transistors under same
gate-source voltage $V_{GS}$ as shown in Figure 4.1 (a). As to DEMUX, decision circuit and frequency divider, the cross coupled transistor pair used in the latch is one of the largest contributors to the output capacitance. In other words, smaller size in hold branch than that in sampling branches would help to reduce parasitic capacitance further. However, the device size in hold branch should be large enough to maintain a particular logic state at the appropriate clock phase and to avoid large duty cycle distortion. For MUX, there is no cross coupled transistor pair in the latch, so the devices in data path should have the same size.

Finally, the aspect ratio of $W_{Data}/W_{Clock}$ are optimized as: $W_{Data}/W_{Clock}=3/5$ for DEMUX to offer enough gain for 20-Gb/s deserialized signal; $W_{Data}/W_{Clock}=1/3$ for MUX to provide an intermediate gain and sufficient bandwidth for 40-Gb/s serialized data stream; $W_{Data}/W_{Clock}=2/15$ for frequency divider and decision circuit to increase the input sensitivity for 40GHz input clock.

As shown in Figure 4.2 (b), both the transistor sizes and bias currents of the used output buffer were scaling up by a factor of 2 to drive heavy external loads since the driving capability was improved via transforming the output impedance step by step.

The device aspect ratios of other proposed circuits described in Chapter 5 were also optimized according to the same strategies addressed here.

![Figure 4.3: Common source amplifiers with (a) Split-resistor; (b) Shunt peaking; (c) Split-resistor, shunt peaking, and series peaking.](image-url)
4.3.3 Split-Resistor (S-R) Loads

Poly-silicon resistors with S-R topology ($R_1$ and $R_2$ are separated by output node) shown in Figure 4.3(a) are used as loads, which is the fastest non-enhanced amplifier. The advantages of this amplifier can be seen that: Unsilicided poly is a pretty efficient current provider (i.e., has a good current to capacitance ratio); Output swing can go all the way up to $V_{DD}$; Allows following stage to achieve high $f_T$; Linear settling behavior (in contrast to NMOS load). On the other hand, the bandwidth and gain limitations can be estimated below:

\[
g_m = \frac{dI_d}{dV_{gs}} = \frac{2I_d}{V_{GS} - V_T} \quad (4.1)
\]

\[
A_v = g_m R_1 = \frac{2I_d R_1}{V_{GS} - V_T} = \frac{2V_{R_1}}{V_{GS} - V_T} \quad (4.2)
\]

\[
f_{-3dB} = \frac{1}{2\pi \cdot R_1 C_{tot}} \quad (4.3)
\]

\[
C_{tot} = C_{db1} + \frac{C_{R_1}}{2} + C_{gs2} + K \cdot C_{OV_2} + C_{fixed} \quad (4.4)
\]

Where, $g_m$ is the transconductance of the transistor; $A_v$ is the DC gain of the amplifier; $f_{-3dB}$ is the -3dB bandwidth of the amplifier; $C_{tot}$ is the total capacitance at the output node; $K = 1 + |A_v|$ can be defined as Equivalent Miller multiplication factor which can be adjusted by optimizing the ratio of $R_1/R_2$ for given constant load resistance $R = R_1 + R_2$. Based on above equations, poly-silicon resistors with S-R topology can be used as loads to reduce critical node parasitic capacitances and Miller Effect further, which is a compromise between high internal signal swing (DC gain) and small RC time constant (bandwidth or speed). It is important for high speed digital circuits to introduce such an additional design parameter, the ratio of $R_1/R_2$. 
4.3.4 Active Inductors, Capacitors, Resistors

A. Operating Principle of Active Inductor

Various versions of the simplest active inductor and the small-signal model are shown in Figure 4.4. From the small-signal analysis, the nodal and loop equations can be expressed as

\[ s \cdot C_{gs} V_{gs} + g_m V_{gs} = -I_x \]  
(4.5)

\[ s \cdot C_{gs} V_{gs} R_g + V_{gs} = -V_x \]  
(4.6)

The equivalent impedance is

\[ Z_{DS} = \frac{V_x}{I_x} = \frac{1 + s \cdot C_{gs} R_g}{g_m + s \cdot G_{gs}} \]  
(4.7)

\[ Z_{DS} = s \cdot \frac{1}{\omega_T} \cdot \frac{R_g - g_m^{-1}}{1 + (\frac{\omega}{\omega_T})^2} + \frac{g_m^{-1} + R_g (\frac{\omega}{\omega_T})^2}{1 + (\frac{\omega}{\omega_T})^2} \]  
(4.8)

\[ Z_{DS} = s \cdot L_{eff} + R_{eff} \]  
(4.9)

Thus, the equivalent inductance is

\[ L_{eff} = \frac{1}{\omega_T} \cdot \frac{R_g - g_m^{-1}}{1 + (\frac{\omega}{\omega_T})^2} \]  
(4.10)
And the equivalent resistance is

\[ R_{\text{eff}} = \frac{g_m^{-1} + R_g \left( \frac{\omega}{\omega_T} \right)^2}{1 + \left( \frac{\omega}{\omega_T} \right)^2} \]  

(4.11)

And the effective quality factor is

\[ Q_{\text{eff}} = \omega \cdot \frac{L_{\text{eff}}}{R_{\text{eff}}} = \frac{\omega}{\omega_T} \cdot \frac{R_g - g_m^{-1}}{g_m^{-1} + R_g \left( \frac{\omega}{\omega_T} \right)^2} \]  

(4.12)

At low frequencies, the active inductor has an impedance of about \( g_m^{-1} \). At high frequencies, the low-pass filter formed by \( R_g \) and \( C_{gs} \), cuts the gate-drain connection, causing the impedance of the circuit to increase similar to that of an inductor. It turns out that for \( R_g > g_m^{-1} \), the impedance becomes inductive in a certain frequency range.

Fortunately, this inductance can be conveniently trimmed with the gate resistor \( R_g \), while keeping the series resistance \( g_m^{-1} \) approximately constant for operating frequency much lower than the transit frequency \( (f_T) \). The active inductor can be used for frequencies up to about \( f_T/2 \) [22]. Compared with spiral inductors, active inductors are much smaller and amenable to monolithic integration. For a common source amplifier with active inductor loads, its voltage gain is about

\[ A_v = \frac{g_{m\text{MA}}}{g_{m\text{ML}}} = \frac{\sqrt{W_A}}{\sqrt{W_L}} \quad L_A = L_L = L_{\text{min}} \]  

(4.13)

It means that the voltage gain of this gain stage just depends on the aspect ratio between the input transistors and the load transistors and offers excellent immunity against the fluctuations of process, supply voltage, and temperature.

To verify the derived equations, Spice simulations based on 0.18\( \mu \)m CMOS are carried out and the simulated AC response are illustrated in Figure 4.5. From Figure 4.5 (a), with the gate resistor increasing, the undesired peak goes higher corresponding to
the increasing effective inductance, which confirms that this type of active inductor can be trimmed conveniently. Moreover, in the case shown in Figure 4.5 (a), the optimal value of the gate resistor should be around 2KΩ in this typical case, which is not a problem to realize such a resistor in modern CMOS technologies. Based on Figure 4.5 (b), an active inductor with gate NMOS has stronger shunt peaking effect and is easier to trim the inductance value than spiral inductor, but its low frequency gain is much lower when the gate NMOS has a too large equivalent resistance (bandpass).

According to the simulated data, using active inductors with gate resistors will result in an optimal trade-off for wide-bandwidth circuits, which is not limited by finite self-resonance frequency, while on-chip spiral inductors with large inductance often suffer from low self-resonance frequency due to undesired parasitic.

**B. Applications of Active Inductor**

Some circuits which take the advantage of simple active inductors are reported in the past years [22–24]. In [22], a 3GHz, 32dB CMOS limiting amplifier for SONET OC-48 receivers was implemented using active inductor to boost amplifier bandwidth. In this design, to solve the problem of the large DC voltage drop across the conventional active inductor at low supply voltages, a low voltage-drop active inductor topology
Figure 4.6: Low voltage-drop active inductor and capacitive voltage converter. was employed as shown in Figure 4.6 (a). The solution is to bias the resistors of the active inductors one NMOS threshold voltage (note that this threshold voltage is increased by the back-gate effect) above $V_{DD}$ reducing the voltage drop across the inductor by about half shown in Figure 4.6 (a). Since no current is drawn from this bias voltage ($V_{BH}$), it can be generated on-chip with a capacitive voltage converter as shown in Figure 4.6 (b). However, this active inductor topology is not useful for high supply voltage (5V) applications and unfeasible for lower supply voltage (1V to 1.8V) circuits.

In [23], a 2.5-Gb/s CMOS transimpedance amplifier using inductive load was realized, in which novel active inductor architecture was proposed. Figure 4.4 shows the circuit with the proposed active inductor load. In Figure 4.7, PMOS $M_{11}$ and $M_{12}$ are used for the current bleeding purposes in order to reduce the voltage drop across load resistor $R_3$ and $R_4$. For the same available voltage drop, much larger load resistor can be used and the DC voltage can be specified for the direct coupling with next stage. Thus, the amplifier in Figure 4.7 using the proposed novel active inductor
Figure 4.7: 2.5Gb/s Transimpedance Amplifier in 0.35\(\mu\)m CMOS technology.

load effectively becomes the same architecture like the typical shunt-peaked amplifier shown in Figure 3.7. The notable differences between them are that the inductance \(L_{eq}\) can be much larger than \(L_d\), and the resistor \(R_{3,4}\) can also be larger value than \(R_L\). Thus, the proposed active inductor can be useful for low-frequency inductive peaking purposes (below 3 GHz for 0.35\(\mu\)m CMOS technology).

In fact, active inductors also can be used in MMIC. In [24], using active inductors in the lumped equivalent circuit, a significant area reduction was achieved while maintaining enhanced circuit performance at multi-gigahertz frequencies as shown in Figure 4.8. In addition, the tunable inductance of the active inductors enables the control of the center frequency by the bias currents, which can be utilized for the realization of reconfigurable RF front-ends in multi-standard wireless systems.

In summary, active inductor is a useful alternative for on-chip spiral inductor in shunt-peaked amplifiers and can be tactfully used to design compact trimmable high speed circuits, which is verified by many realized chips as addressed above and will be confirmed further by the fabricated chips described in Chapter 5.
Figure 4.8: (a) Schematic and equivalent circuit of the regulated cascode active inductor; (b) Complete circuit schematic of the fully integrated quadrature hybrid.

C. Active Capacitors and Resistors

Some important sub-circuits such as DC offset cancelation feedback networks and signal loss and detection circuits require low-pass filters (LPF) with lower corner $f_c$, which can be realized by active resistors and capacitors with large values to save layout area, to improve reliability and immunity against PVT fluctuations. Here active resistors and capacitors are just briefly introduced due to the limited length of this thesis and will be mentioned again for specific circuits in Chapter 5.
**4.3.5 Other Useful Circuit Techniques**

![Circuit Diagram](image)

**Figure 4.9:** Current sources. (a) Schematic. (b) $I_D$ versus $V_{DS}$.

### A. Stacked Current Source

In Figure 4.9 (a), the used current source 4 consists of two stacked NMOS transistors. The upper transistor is a LVT device and the bottom is a RVT device, and they are connected in series. This configuration increases the output resistance of the current source. Ignoring the body effect, the output resistance of the stacked current source [2] approximately is:

$$R_{out} = r_o + r_o (1 + g_m r_o) = 2r_o + g_m r_o^2 \approx g_m r_o^2$$  \hspace{1cm} (4.14)

Where, $r_o$ is output resistance of LVT and RVT devices, assuming that they have the same output resistance. This results in a flat current source characteristic as shown in Figure 4.9 (b). Even though the stacked current source is just a little bit better than the conventional cascode current source, the number of used transistors is reduced and no additional DC bias is required. The main disadvantage of stacked current sources is higher operating voltage to keep the devices in saturation. However, the
minimum operating voltages of four current sources are almost same as 250mV based on Figure 4.9 (b), which permits larger PVT fluctuations.

In addition, stacked LVT and RVT NMOS transistors with a channel length of 180nm are used as current source to reduce short channel effects and geometric mismatches. Both latches and buffers in the high speed digital circuits use stacked current sources to achieve excellent immunity against PVT variations.

![Figure 4.10: CSA with BW extension: (a) AC response; (b) Transient response.](image)

**B. Inductive Shunt Peaking**

The maximum speed of a differential pair is limited mainly by the parasitic capacitances of the transistors and the layout. Therefore, spiral inductors are connected in series with the load resistors as shown in Figure 4.3 (b). This inductive peaking enhances the operating bandwidth without deteriorating the low-frequency response. The physics of shunt peaking is quite straightforward. Since the voltage on capacitors is unable to change abruptly, at the very beginning the additional current is almost provided by capacitors only and the output voltage changes quickly. With the output voltage going down, the current provided by the load resistor will increase, but the current provided by capacitors will decrease. The addition of the inductor in series with the drain resistor delays the current flow through the branch containing the
resistor, making more current available for charging the device capacitors, and reducing the rise and fall times. From another perspective, the addition of an inductance in series with the load capacitance introduces a zero in the transfer function of the common source stage which helps offset the roll-off due to parasitic capacitances [13]. In perfect world, inductive peaking can increase the bandwidth to about 1.72 times larger than the unpeaked case. Inductance values are scaled with the same factor as the drain resistors are.

Both spiral inductors and bonding wire inductors are often used as shunt peaking inductors. However, additional pads and poor accuracy control for bond inductors are hardly acceptable for the required chip area and integration density. For 40-Gb/s operation, it makes sense to use inductive peaking inductors only at the fastest part of the system. Shunt peaking can improve the bandwidth by approximately 50%, assuming the use of on-chip inductors.

Another technique to enhance the bandwidth of CML circuits is series peaking. An inductor is connected in series to the output of the CML circuit as shown in Figure 4.3 (c). As illustrated in Figure 4.10, the output network acts as a filter which consists of various parasitic capacitances, load resistors, bond inductances, and on-chip inductors. Series peaking can additionally improve the bandwidth by approximately 45% when combined with shunt peaking. Series peaking makes sense if it is used in combination with shunt peaking. If only series peaking is used, then the enhancement in bandwidth is low. Using series peaking and shunt peaking can nearly double the bandwidth of a CML circuit. However, series peaking is seldom used in compact chips for low-cost applications.

C. Input Matching Networks

On-chip input matching for clock path is realized with low resistance (50Ω to 250Ω) poly-silicon resistors shown in Figure 4.2 (a), which also acts as a DC level shifter
and ESD protection to allow AC-coupling of the clock input. This input DC level was optimized for fast switching. The used DC level shifters provide a DC level of $0.5V_{DD}$ for MUX, DEMUX, decision circuit but a $0.6V_{DD}$ for clock divider.

**Figure 4.11:** Output buffers. (a) For DEMUX; (b) For data decision circuit.

**D. Output Buffers**

Limited bandwidth is a big challenge for high speed buffer design, which can be relaxed by the inductor-peaking network. Large input/output swing requirement is another challenge in designing such a buffer. The required large voltage swing causes the input transistor pair to operate partially in the triode region, introducing non-linearity. Therefore, the waveforms on the two output nodes become asymmetric. The frequency response for both the rising-step output and the falling-step output had to be optimized. This problem can be partly relieved by optimizing the ratio of $R_1/R_2$ in split resistor topology mentioned above.

The proposed buffers in Figure 4.11 consist of multiple-stage differential amplifiers in series which are not for voltage gain, but are required for driving the external 50Ω load. In each stage the tail current is twice the current of the previous stage. The first stage offers a high-voltage swing, which drives the second stage. For Figure 4.11
(b), the second stage works as a limiting amplifier to provide appropriate amplitude, proper common-mode level. The last differential amplifier is designed to provide enough driving capability and good matching to external 50Ω loads.

4.4 Chapter Summary

Based on Chapter 2 and 3, this chapter addressed the design fundamentals, strategies, and identified four key circuit techniques, which will be used for low-power low-supply high-speed circuit implementations in the following chapter. Other useful circuit techniques were also discussed at the end of this chapter, which will be used together with the selected strategies and key techniques for specific circuits described in Chapter 5.
Chapter 5

Low-Power and Low-Supply ICs Design

5.1 Chapter Overview

To validate the identified design strategies and circuit technologies in Chapter 4, the fabrications and evaluations of low-power analog circuits using compact active inductors, and the design and simulations of low-supply digital circuits using spiral inductors and other low-voltage techniques will be addressed in this chapter.

5.2 Low-power Analog Circuits

Optical transceivers operating at 622-Mb/s to 10-Gb/s are widely used in optical access networks (OANs), backbone telecommunication networks and Ethernet fiber optic links which used for local area networks (LANs). Some high speed CMOS circuits with spiral inductor loads or resistive loads were realized, which consume large power or large chip size [25–30]. However, low power dissipation compact chips are the first choice for commercial communication systems.

Shunt peaking (SP) technology is widely used to extend the bandwidth of gain stages, which can be realized by topmost metal spiral inductors or on-chip active inductors. Although active inductors introduce a little bit larger noise and have
smaller quality factors, they consume much smaller silicon area, which is still a good choice for low-cost applications.

In this section, several CMOS circuits are presented, which use area-saving active inductors as loads of gain cells to boost the bandwidth, and to reduce the power consumption and silicon area.

5.2.1 High Modulation Amplitude LDD/MD

With the rapid development of optical access networks, high performance, low cost optical transmitters (TXs) are needed. In the past years, other laser diode (LD) drivers with higher costs were realized using resistive-load or transistor active-load, which usually have a lower power efficiency [25], more chip area [25–27], much smaller modulation current [26,27]. In general, large voltage drop over LD makes it difficult to reduce the power supply voltage of the LD driver down to 3.3V or below. Furthermore, for reliability and application reasons, high power supply CMOS drivers providing large modulation current for laser diodes or high modulation voltage for optical modulators are indispensable. In this part, a LD/MZM driver using on-chip shunt peaking active inductors and direct-coupled topology is realized in 0.6μm CMOS for low-cost 1.25-Gb/s optical transmitters.
A. Circuit Design

The schematic of the proposed LD/MZM driver is shown in Figure 5.1. The direct-coupled, fully differential circuit consists of an input level-shifter, two pre-drivers and an output stage. The whole circuit is fully balanced using differential amplifiers to maximize its operating speed and to minimize undesired common-mode noises.

The upper NMOS transistor pair in each pre-driver is a differential current amplifier. The lower transistor forms the current source for the NMOS transistor pair. The active inductors pair composed of NMOS pair and resistor pair is used as loads. Active inductors with low Q value are used here to boost the bandwidth due to the following reasons: Firstly, high Q value, high inductance on-chip spiral inductors are not easily obtained; Secondly, on-chip spiral inductors inevitably consumed very large chip area; Thirdly, high Q value is not indispensable for wide band amplifiers and the Q value is determined by the poly-silicon load resistors that are connected in series to the inductors; Additionally, active inductors can be used for frequencies up to $f_T/2$ while the application of spiral inductors are limited by definite self-resonance frequency due to undesired capacitance to the substrate; Lastly, active inductors are very compact, easily fabricated and have very good immunity against PVT variations. A differential source follower pair to realize the functions of level-shifting and impedance transforming is omitted here to reduce power dissipation and chip dimension. Thus, direct coupled topology is employed in this design. The input level shifter implemented through a group of 5KΩ poly-silicon resistors and a pair of 50Ω poly-silicon resistors provides the function of DC level shifting and input impedance matching. In this work, identical 5KΩ resistors in the level-shifter are employed to obtain an accurate input common-mode level and a more accurate DC bias voltage for current sources. Because DC level offset can be eliminated using identical resistors with an equal resistance tolerance resulting from inevitable process variation rather
than different resistors with different resistance tolerances. In addition, parasitic inductors of bonding wires for output stage can be properly used to reduce the jitter, rise time of output signal pulse and to enlarge the bandwidth of this LD/MZM driver.

Due to very large output current/voltage swing of the LD/MZM driver, most of the devices have huge aspect ratios (W/L). This will necessarily result in increased parasitic capacitance, especially in critical nodes, which strongly impacts the high speed operation of the LD/MZM driver. Multiple-finger structure offering nearly 50% reduced source and drain area is employed for all large transistors to minimize parasitic capacitance. Special attention has been given to keep routes as short as possible while using large line width to meet the requirement of current density and to minimize parasitic inductance. Great care also has been taken as to virtually avoid large area overlap between supply and signal lines. For the purpose of effective biasing and minimization of substrate bouncing, substrate contact arrays have been extensively used. In addition, double metals which composed ground lines, supply lines and other DC bias lines are used in large area as MIM (Metal-Isolator-Metal) capacitors to cancel undesired noises. Therefore, the performance of this circuit was improved greatly due to all design techniques mentioned above.
**Figure 5.3:** Block diagrams of evaluation setup for (a) LD driver; (b) MZM driver.

**B. Circuit Fabrication and On-Chip Evaluation**

The proposed LD/MZM driver was fabricated in 0.6μm double-poly double Metal N-well CMOS technology. The microphotograph of the die is shown in Figure 5.2. The chip dimensions including bonding pads (0.1mm×0.1mm) are 0.6mm×0.65mm. On the chip, only one tenth of the total chip area in the middle region is used for the active part.

The performance of the fabricated LD/MZM driver is evaluated via on-wafer probing on uncut wafers employing a CASCADE MICROTECH probe station, an ADVANTEST D3186 Pulse Pattern Generator, an ADVANTEST R6142 Programmable DC Voltage/Current Generator, a ROHDE & SCHWARZ SMP04 Signal Generator (10MHz-40GHz), an Agilent 83430A Lightwave Transmitter, an Agilent Lightwave Multimeter and an Agilent Infinium DCA 86100A Wide-bandwidth Oscilloscope. A low threshold 1.55μm wavelength InAsP/InGaAsP strained multi-quantum well laser diode and a Mach-Zehnder $LiNbO_3$ external modulator (MZM) are used to observe
Figure 5.4: Optical output of the fabricated LD/MZM driver: (a) 625-Mb/s from LD; (b) 1.25-Gb/s from LD; (c) 625-Mb/s from MZM; (b) 1.25-Gb/s from MZM.

output optical signals.

The block diagrams of evaluation setup of the realized LD/MZM driver are illustrated in Figure 5.3. The LD driver provides modulation current and bias current for the used laser diode. The driver is used to drive a MZM which has a single-end RF input. If a differential MZM is used, the measured result of the differential MZM driver should be much better.

The DC current of the LD/MZM driver under a single supply of 5V is less than 94mA, corresponding to a power dissipation of 470mW. The circuit has been tested using an input voltage of $500mV_{pp}$ at different bit rates. The measured eye-diagrams at the bit rates of 625-Mb/s and 1.25-Gb/s from one single-ended output of the
Table 5.1: Measurement Summary of LD/MZM driver

<table>
<thead>
<tr>
<th>Parameters</th>
<th>LD</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply</td>
<td>5V</td>
<td>5V</td>
</tr>
<tr>
<td>Power</td>
<td>470mW</td>
<td>470mW</td>
</tr>
<tr>
<td>Bitrate</td>
<td>1.25-Gb/s / 0.625-Gb/s</td>
<td>1.25-Gb/s / 0.625-Gb/s</td>
</tr>
<tr>
<td>SNR</td>
<td>10.35 / 13.06</td>
<td>9.83 / 9.65</td>
</tr>
<tr>
<td>RMS Jitter</td>
<td>29ps / 18ps</td>
<td>31ps / 38ps</td>
</tr>
<tr>
<td>Output Intensity</td>
<td>1.223mW / 1.239mW</td>
<td>0.417mW / 0.425mW</td>
</tr>
<tr>
<td>Extinction Ratio</td>
<td>5.80dB / 5.91dB</td>
<td>5.51dB / 5.67dB</td>
</tr>
</tbody>
</table>

LD/MZM driver are shown in Figure 5.4. The modulation current range at each single-end output of the LD driver is 0-80mA corresponding to an optical modulation amplitude of 1.24mW.

The maximum modulation voltage for MZM is over $4V_{pp}$ corresponding to a 425µW optical modulation amplitude with a 2.15V internal DC bias for the output stage. This driver circuit with a wide input common-mode level range from 1.4V to 3.8V can operate well with a controllable modulation current/voltage output under a single supply voltage ranging from 4.5V to 5.5V. Moreover, the temperature coefficient of this driver is only $0.06mA/°C$ for LD or $5mV/°C$ for MZM with an operating temperature ranging from $-40°C$ to $+85°C$.

In Table 5.1, optical measurement data shows the performance of the implemented LD/MZM driver. The RMS jitter of output optical signal from the driven laser diode and the modulated Mach-Zehnder $LiNbO_3$ external modulator is no more than 38ps, the extinction ratio is larger than 5.5dB, and the SNR ranges from 18.5dB to 27.5dB. From tested eye-diagrams, we can come to a conclusion that the performance of the fabricated chip verges on the anticipant result and can operate at a bit rate up to 1.25-Gb/s.
C. Design Summary

The realized 0.6μm CMOS LD/MZM driver uses two pre-drivers with on-chip shunt peaking active inductors to boost the bandwidth and to remove the power-hungry source followers without consuming large silicon area. Supplied by this LD/MZM driver, the modulation voltage of $4V_{pp}$ for MZ modulator, and the modulation current of $80mA_{pp}$ for laser diode, result in fully-open optical eye diagrams at the bit rate of 1.25-Gb/s. This driver has a much higher efficiency factor of $V_{out}^2/P_{total}$ or $I_{out}^2/P_{total}$ than its counterparts due to the used low-power techniques such as DC bias level optimizing, device sizing and matching, active inductor peaking, and direct-coupling topology.

5.2.2 Low-Power Limiting Amplifiers

Synchronous optical network (SONET) is an industry standard for broadband optical fiber networks and is defined for various transmission speeds, such as OC-3, OC-12, OC-24, OC-48, OC-96, OC-192, up to OC-768. In general, a distributed feedback (DFB) laser launches a signal at a power of -2dBm into an optical fiber. After passing through about 100Km of a single-mode fiber, the signal is attenuated to -28dBm. To detect this signal with a safety margin, we require the receiver to have a sensitivity of -31dBm, corresponding to about 1μW average optical power [22].

As the name SONET implies, the transmission is synchronous and continuous. Non-return-to-zero (NRZ) coding is used and the bit stream is scrambled to make the probability for long runs of zeros or ones small. The SONET OC-3 $\times$ $2^N$ ($N=0, 1, 2, \cdots$) standard further prescribes system bit-error rates (BER). For OC-48, the required BER should be $10^{-10}$ or better, which corresponds to less than one error every four seconds.

For a typical SONET OC-48 receiver front-end, under worst-case conditions, a
signal of -31dBm is received and converted by an avalanche photodiode (APD) into a current signal of 5.3μA_{pp}. A transimpedance amplifier (TIA) with a typical transimpedance of 1.5KΩ converts this current into a voltage signal of 8mV_{pp}. The task of the limiting amplifier (LA) is to amplify this small voltage signal to a voltage level sufficient for the reliable operation of the clock and data recovery (CDR) circuit (≥250mV_{pp}). It follows that the LA must have a small-signal gain of 30dB. If the received signal is stronger than -31dBm, for example, because the optical link is shorter than 100Km, the LA will receive a stronger signal. If the amplitude rises above about 30mV_{pp}, the output amplitude of the amplifier will saturate, hence the name “limiting amplifier”.

Amplifier parameters are chosen such that the TIA determines the system performance (e.g., sensitivity) and the LA does not degrade this system performance. A typical bandwidth chosen for the TIA is 70% of the clock frequency (For a bitrate of 2.5Gb/s, the clock frequency is 2.5GHz and the required bandwidth should be around 1.75GHz), a good trade-off between inter-symbol interference (ISI) and noise. The bandwidth of the LA is chosen to be 70% of the clock frequency or larger such that it adds very little ISI. A typical input-referred noise current of the TIA is 400nA. The noise figure of the LA is chosen to be 16dB, which adds only about 3.5% to the input-referred noise current of the TIA (corresponding to a sensitivity penalty of 0.15 dB). In summary, the design goal for the LA is a differential gain of 30dB, a -3dB bandwidth of 0.7 clock frequency, and a noise figure of 16dB.

In addition, the LA offset may also impact the receiver performance [29]. Vertical shift of the signal with respect to the decision threshold reduces the peak signal level, degrading the receiver sensitivity. In addition, the LA offset leads to pulse width distortion complicating the design of the CDR circuit. Continuous-time offset cancelation circuits introduce a lower cut-off frequency in the transfer function and “droop” in the time domain after long runs (also known as baseline wander). At
the end of the droop period, the signal is again shifted with respect to the decision threshold. To minimize this effect, the lower cut-off frequency must be sufficiently small, typically on the order of a few tens of kilohertz.

\[ \text{Figure 5.5: Limiting amplifier architecture.} \]

**A. Circuit Design**

**§1 LA Architecture**

Shown in Figure 5.5, the architecture of the LA consists of a broadband input-matching network, five identical gain stages shown in Figure 5.6 (a) comprising the 1.25-Gb/s and 6-Gb/s LA cores while three identical gain stages shown in Figure 5.6 (b) comprising the 10-Gb/s LA core, an offset cancelation feedback loop, and an output buffer. The LA core must provide sufficient gain and bandwidth. It is, therefore, desirable to employ various bandwidth boosting techniques without introducing too much noise. Designed to operate as a stand-alone module, the LA must deliver large voltage swings to 50Ω loads, requiring a high-current output buffer. The core itself must provide a relatively large driving capability for the large input capacitance of the buffer.
Figure 5.6: Schematic of the proposed circuits: (a) Gain cell with active inductors.
(b) Active feedback gain cell with folded active inductors for low supply LA.

§2 Amplifier Core

A cascade of identical gain cells is used as the amplifier core to achieve enough voltage gain and -3dB bandwidth. However, a critical difficulty stems from the relationship between the number of gain cells, m, and the overall input-referred noise. For a larger m, the lower gain per stage leads to rapid accumulation of noise. For the input-referred noise levels targeted in this design, m must fall below approximately 6.

As shown in Figure 5.6 (b), this work introduces active negative feedback as a means of improving the GBW of amplifiers. Based on the discussion in [29], active
feedback increases the GBW beyond the technology $f_T$ by a factor equal to the ratio of $f_T$ and the cell bandwidth.

In addition to active feedback, all proposed broadband amplifiers employ gain cells with active inductors shown in Figure 5.6 (a) and the modified version shown in Figure 5.6 (b). Because conventional active inductors consume rather large voltage headroom, folded active inductors have to be used in low supply circuits (i.e. 1.8V 0.18µm CMOS LA). In general, consuming larger voltage headroom is one of main drawbacks of active inductors. However for high supply circuits, large voltage headroom consumption permits directly coupled multiple gain stages without level-shifting source followers (SFs), so there is no SF-induced gain drop but nearly half power is saved.

![Bonding Wire Inductor(Off-Chip)](image)

**Figure 5.7:** Schematic of the proposed output buffer.

§3 Output Buffer

Buffers driving off-chip loads typically present a bandwidth bottleneck resulting from the large input transistors that are necessary for high current drive capability. In broadband applications, the buffer must drive an on-chip back termination resistor of about 50Ω in addition to an off-chip load of 50Ω. To deliver a single-ended voltage swing of 0.2V to the equivalent resistance of 25Ω, the buffer must steer 8 mA, requiring
Figure 5.8: Schematic of the used offset cancelation feedback circuit.

Figure 5.9: Schematic of the used on-chip signal loss detection and alarm circuit.

A tail current of 10 to 12 mA when the incomplete switching of the stage is taken into account. Consequently, the input devices must be wide and bonding wire inductors can be used to partly tune out the relatively large output capacitance, which is shown in Figure 5.7.

§4 Offset Cancelation

The principal difficulty in the design of the offset cancelation loop relates to the required corner frequency $f_c$ of the resulting high-pass filter. In order to ensure negligible droop in the output in the presence of long runs, $f_c$ must fall in the range of
a few tens of kilohertz. As shown in Figure 5.8, $R_F \cdot C_F$ must reach a few milliseconds for to be equal to a few tens of kilohertz. In this design, a 50MΩ active resistor serves as $R_F$ and a 40-pF MOS capacitor as $C_F$, which are used to realize large value resistors and capacitors without chip area penalty.

Another issue stems from the low load resistance seen by the feedback amplifier at the input of the LA. To compensate for an input-referred offset voltage of roughly 20mV, and the differential pair of input stage must steer about 1 mA to their loads while sensing an output offset of less than 10 mV, a value determined by pulse width distortion requirements. Thus, these transistors must be sufficiently wide.

§5 Signal Loss Detection and Alarm

To obtain signal loss information of the incoming data stream, an on-chip simple loss detection circuit is included as shown in Fig.5.9. Firstly the amplified signal was low-pass (LP) filtered by the proposed active low-pass filter and then the filtered signal was send to the input of the used voltage comparator and compared with the reference voltage. The output of voltage comparator is converted to rail-to-rail digital signal by cascading inverters to be used as the switching signal for the alarm LED (light emitting diode). Generally, there is no signal loss, LP filter outputs a higher voltage level than the reference and the output of the comparator is low. If signal loss occurs, the output signal level of LP filter will be lower than the reference voltage and the comparator will generate a “high” voltage level to turn on the alarm LED.

B. 1.25-Gb/s, 6-Gb/s, and 10-Gb/s LA Fabrications

In this section, LAs operating at 1.25-Gb/s to 10-Gb/s have been designed and fabricated using 0.6μm, 0.25μm, 0.18μm CMOS technologies. The microphotographs of the dies are shown in Figure 5.10. The chip dimensions including bonding pads are 0.5mm×0.4mm for 1.25-Gb/s LA, 0.7mm×0.5mm for 6-Gb/s LA, and 1mm×0.7mm
for 10-Gb/s LA, respectively. The dimension of pads is 0.1mm×0.1mm and the core area of the fabricated chips is rather small.

![Figure 5.10](image1.jpg)

**Figure 5.10:** Microphotographs of the fabricated circuits: (a) 0.6µm CMOS LA (0.5mm×0.4mm); (b) 0.25µm CMOS LA (0.7mm×0.5mm); (c) 0.18µm CMOS LA (1.0mm×0.7mm).

![Figure 5.11](image2.jpg)

**Figure 5.11:** Measurement results at 5mV_{pp} input of the fabricated LAs: (a) 0.6µm CMOS LA at 1.25-Gb/s; (b) 0.25µm CMOS LA at 6-Gb/s; (c) 0.18µm CMOS LA at 10-Gb/s.

### C. Circuit Measurement

The performance of the fabricated chips has been evaluated via on-wafer probing on uncut wafers employing a CASCADE MICROTECH probe station, an ADVANTEST D3186 Pulse Pattern Generator, an ADVANTEST R6142 Programmable DC Voltage/Current Generator, a ROHDE & SCHWARZ SMP04 Signal Generator (10MHz-40GHz), and an Agilent Infinium DCA 86100A Wide-bandwidth Oscilloscope.
Table 5.2: Measurement Summary of LAs

<table>
<thead>
<tr>
<th>Parameters</th>
<th>0.6μm LA</th>
<th>0.25μm LA</th>
<th>0.18μm LA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (mW)</td>
<td>108</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>Supply (V)</td>
<td>5</td>
<td>3.3</td>
<td>1.8</td>
</tr>
<tr>
<td>Gain (dB)</td>
<td>46</td>
<td>56</td>
<td>46</td>
</tr>
<tr>
<td>Rate (Gb/s)</td>
<td>1.25</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Jitter (ps)</td>
<td>6</td>
<td>7.4</td>
<td>2</td>
</tr>
<tr>
<td>Bandwidth (GHz)</td>
<td>0.88</td>
<td>4.3</td>
<td>7.1</td>
</tr>
<tr>
<td>Input Amplitude (mV)</td>
<td>5-500</td>
<td>3-500</td>
<td>4-400</td>
</tr>
<tr>
<td>Output Amplitude (mV)</td>
<td>200</td>
<td>178</td>
<td>125</td>
</tr>
</tbody>
</table>

The testing results of realized LA circuits are illustrated in Figure 5.11, and a summary of all proposed circuits is listed in Table 5.2.

On the basis of above measured data, we can say that the realized LAs with active inductors can reach the comparable performance as its counterparts realized with spiral inductors [28, 29]. However, the power dissipation is much lower and the chip size is smaller, which means they are better for low cost applications.

**D. Design Summary**

LAs operating at OC-24, OC-96, and OC-192 have been fabricated in 0.6μm, 0.25μm, and 0.18μm CMOS technologies. The observed performances of the fabricated chips confirmed that the used design strategies and circuit techniques such as active inductors, capacitors, resistors, and negative feedback indeed result in a feasible trade-off among gain-bandwidth product, power consumption, and silicon area.

**5.2.3 Monolithic CDR Circuit**

As a vital component, CDR plays an increasingly important role in high-capacity networks. In modern high-speed data communication systems, NRZ data stream is
Figure 5.12: Block diagram of the fabricated CDR.

normally transmitted and the clock must be extracted from the data. Monolithic full-rate CDR without any external components is desired for low-cost applications. In comparison with narrow-band regenerative frequency divider (NRFD), phase-locked loop (PLL) is a more attractive choice for the clock recovery since it is easier to be implemented and has high integration, reliability and flexibility [31].

Generally PLL-based clock recovery (CR) has a very narrow loop bandwidth for good jitter attenuation. However, the tuning range of an on-chip VCO must be wide enough to cover process, voltage and temperature (PVT) variations. Therefore, frequency acquisition aid is indispensable for CR implementation. In a phase-and frequency-locked loop (PFLL), the phase detector (PD) and frequency detector (FD) can be designed independently to achieve wide capture range, small jitter, and low power dissipation [32].

For most applications, the time jitter and the phase noise are two important
design criteria of a PLL. Unfortunately, the switching activity of digital modules in mixed-signal systems introduces power-supply or substrate noise, which greatly disturbs those noise-sensitive blocks in a PLL. In particular, noises injected onto the voltage-controlled oscillator (VCO) pose the dominant jitter source of a PLL.

In this work, based on the negative conductance (NC) configuration, novel differential delay cells with active inductor loads are used to lower phase noise and to improve linearity of the proposed ring VCO.

**Figure 5.13:** The proposed I/Q VCO: (a) Block diagram; (b) Schematic of the proposed variable delay gain stage; (c) Differential control circuit.
A. Circuit Design

As shown in Figure 5.12, the proposed CDR circuit is composed of an I/Q VCO, a PFD, a loop filter and a data retimer which is based on a master-slave flip-flop. This CDR does not need any edge-detection circuits to preprocess the incoming NRZ data stream. In the VCO, 0° in-phase (I) and 90° quadrature-phase (Q) clocks are generated. The PFD has three functional blocks: phase detector (FD), quadrature phase detector (QPD) and frequency detector (FD). All of the main blocks in the CDR are fully differential. A differential VCO structure reduces the effects of common-mode noise, the magnitude of current spikes injected to power supply and substrate, and ultimately the clock jitter generation. Similarly, differential architectures adopted in the PFD and the loop filter can improve the performance of the CR with a noisy supply and substrate. Input, output buffers and inter-stage buffers are used to realize input matching, DC level shift, impedance transforming and decouple the CDR core from external 50Ω environment. Subcircuits are described below.

§1 Voltage Controlled Ring Oscillator

As a pivotal building block in PLL, high frequency and RF VCOs can be implemented monolithically as LC oscillators or ring oscillators. In comparison, monolithic high-Q LC oscillators have lower phase noises but ring VCOs offer wider tuning ranges and consume smaller chip areas. The realized I/Q ring VCO comprises a differential control circuit for VCO tuning, four-stage delay cells and two buffering amplifiers as shown in Figure 5.13 (a). The adopted novel differential variable-delay gain stage with active inductor loads and the employed control circuit to generate differential control voltages $V_{con+}$ and $V_{con-}$ for the gain stages of VCO are shown in Figure 5.13 (b) and (c), respectively.

Since a high performance ring VCO can be easily obtained using the negative
conductance [33], the used VCO is also developed in this technology. As shown in Figure 5.13 (b), the variable-delay gain stage is constituted by transistor pair $M_{AP}$, $M_{AN}$, cross-coupled pair $M_{FP}$, $M_{FN}$, active inductor load pair $M_{LP}/R_g$, $M_{LN}/R_g$ and other transistors used as current sources, source followers and biasing blocks. Firstly, a gain stage with MOS loads is difficult to operate at high data rate due to the large time constant of load capacitance, but a gain stage with inductive loads can provide much larger gain-bandwidth product. Using inductive loads, the capacitive loading can be partly tuned out, and then the pole of each gain stage can be pushed toward high frequency end, which is the so-called shunt-peaking technique. In general, inductive loads can be implemented with on-chip spiral inductors or active inductors. It is very difficult to realize a high-inductance and high-Q on-chip spiral inductor with a small die size. Contrarily, active inductors are compact and offer adequately high operating speed [34]. Thereupon, active inductor pair $M_{LP}/R_g$ and $M_{LN}/R_g$ is introduced here as the loads of transistor pair $M_{AP}$, $M_{AN}$ to maximize the operating frequency of the proposed VCO. Secondly, the cross-coupled pair $M_{FP}$, $M_{FN}$ introduces a negative average conductance that reduces the overall output conductance and equivalently increases the output impedance and hence the delay. Thus, this VCO can operate at the expected frequency range with less gain stages and the phase noise is enormously lowered. Thirdly, to keep the output voltage swing a constant, a differential control circuit shown in Figure 5.13 (c) is employed, in which differential pair $M_{cl}$, $M_{c2}$ is used to steer $I_{SS}$ to $M_{AP}$, $M_{AN}$ and $M_{FP}$, $M_{FN}$. Moreover, in the control circuit, $V_{con+}$ and $V_{con-}$ can be viewed as differential control lines and thus provide higher noise immunity for the control input. Finally, differential signals are obtained through a pair of source followers, which offers two advantages: an easy direct-connection with the subsequent differential PD and a low-noise output signal due to common-mode noises suppression.

For a clock recovery application, choosing appropriate transistors $M_{AP}$, $M_{AN}$,
\( M_{FP}, M_{FN} \) and \( M_{c1}, M_{c2} \) is the key point. To avoid latch-up, the transconductance of \( M_{FP}, M_{FN} \) must be less than that of \( M_{AP}, M_{AN} \). Additionally, the dimension of \( M_{c1}, M_{c2} \) should be adjusted so carefully that proper VCO gain, loop gain and consequently well-balanced tuning range, linearity, and noise performance can be optimized simultaneously.

![Diagram](image)

**Figure 5.14:** Schematic of the proposed (a) PD/QPD, and (b) FD.
§2 Phase Detector and Frequency Detector

Compared to a conventional PLL with PD only, PFLL could significantly increase acquisition range and reduce locking time. To optimize the operating speed and avoid problem caused by internal crosstalk, the proposed subcircuits are all based on differential current mode logic (CML). In Figure 5.14 (a), a CMOS version of Pottbäcker PD is proposed [35]. At every transition of the input data, I and Q clocks are sampled by the input NRZ data directly without preprocessing circuit. This operation generates beat notes with 50% duty cycle at PD/QPD outputs when the VCO frequency $f_{osc}$ and bit-rate frequency (data rate) $f_b$ are different. As shown in Figure 5.14 (b), this FD is a differential logic circuit that receives inputs from PD/QPD and generates frequency difference signal at the output $Q_3$. As shown in Figure 5.15 (a), when $f_{osc} < f_b$, PD output $Q_1$ lags QPD output $Q_2$ and the superposition of $Q_1$ and $Q_3$, is positive. On the other hand, when $f_{osc} > f_b$, $Q_1$ leads $Q_2$ and the superposition of $Q_1$, and $Q_3$ is negative, as shown in Figure 5.15 (b). The superposition of $Q_1$ and $Q_3$ indicates a clear DC component driving the loop towards lock.

![Figure 5.15: Timing diagrams of (a) PD/PFD; (b) PD, QPD, and FD.](image-url)
Figure 5.16: Schematic of the realized loop filter (Q1 from PD, Q3 from FD, and $V_{ctrl}$ to control circuit).

§3 Loop Filter

Figure 5.16 shows the schematic of the employed loop filter. Unlike other design [36], this loop filter integrated on chip without any off-chip component. $Q_1$ and $Q_3$ are first added up and then low-pass filtered. The DC component drives the loop towards lock. The transfer function of the loop filter is dominated by $C_0$, $2R_1$, and $R_2$.

§4 Loop Design

The bandwidth of PLL affects the stability, the suppression of phase noise from VCO, and the oppression of spurious modulation and pull-in time. Since the amount of long-term jitter that will result depends on the sensitivity of the VCO to noise, low-Q VCOs based on RC oscillators, such as relaxation or ring oscillators, are very sensitive to noise. Thus, low-Q VCOs can only obtain low long-term jitter by maximizing the loop bandwidth and tracking the input frequency as close as possible. In this design, it is set to 50MHz. By analyzing the closed and open loop responses, the phase margin is found to be 65 degree.
Figure 5.17: Microphotograph of the fabricated monolithic 0.6μm CMOS CDR.

B. Circuit Fabrication and Measurement

The designed CDR circuit was fabricated in 0.6μm CMOS. The chip microphotograph of the die is shown in Figure 5.17. The chip dimension including bonding pads is 1.3mm×1.3mm. The dimension of pads is 0.1mm×0.1mm. Obviously, bonding pads and on-chip capacitors are major consumers of layout area.

The performance of the fabricated chips has been evaluated via on-wafer probing on uncut wafers employing a CASCADE MICROTECH probe station, an ADVANTEST D3186 Pulse Pattern Generator, an ADVANTEST R6142 Programmable DC Voltage/Current Generator, a ROHDE & SCHWARZ SMP04 Signal Generator (10MHz-40GHz), an Agilent Infinium DCA 86100A Wide-bandwidth Oscilloscope, and a HP 8593A spectrum analyzer. Firstly, the performance of the VCO was evaluated in open loop mode. The tuning range of the VCO is from 360MHz to 1060MHz as displayed in Figure 5.18(a), which is wide enough to cover large PVT variations. Then the loop was closed, and differential 622Mb/s 231-1 PRBS data streams were used as input. As shown in Figure 5.18(b), the spectrum of the recovered
Figure 5.18: Measurement results: (a) Frequency control curve of differential tuning VCO; (b) Measured spectrum of the in-locked VCO.

clock signal was measured from the in-locked VCO, which illustrates a phase noise of -92.95dBc/Hz at 10-kHz offset. Figure 5.19 (a) gives the eye-diagram of the retimed 622Mb/s NRZ data, and the measured jitter histogram of the in-locked VCO at 622MHz is shown in Figure 5.19 (b). The circuit is able to acquire lock in a frequency range between 398MHz and 960MHz. The testing results of the realized monolithic CDR were summarized in Table 5.3.

Table 5.3: Measurement Summary of 0.6μm CMOS CDR

<table>
<thead>
<tr>
<th>Technology</th>
<th>0.6μm CMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply</td>
<td>5V</td>
</tr>
<tr>
<td>Power</td>
<td>363mW</td>
</tr>
<tr>
<td>Frequency Range</td>
<td>398MHz to 960MHz</td>
</tr>
<tr>
<td>Phase Noise</td>
<td>-92.95dBc/Hz at 10kHz Offset</td>
</tr>
<tr>
<td>Clock Jitter</td>
<td>12.7ps (RMS) / 82.2ps (Peak-to-peak)</td>
</tr>
<tr>
<td>Bitrate</td>
<td>622-Mb/s</td>
</tr>
<tr>
<td>Application</td>
<td>SONET OC-12 / SDH STM-4</td>
</tr>
</tbody>
</table>
C. Design Summary

A low-power truly monolithic CDR circuit has been fabricated in 0.6µm CMOS using design strategies and circuit techniques such as on-chip active inductors, fully differential topology, device sizing, and DC bias optimizing. The validity of the adopted design strategies and techniques are confirmed by the measurement data.

5.3 Low-Supply Digital Circuits

Current 40-Gb/s optical fiber communication ICs are mainly implemented in GaAs, InP, or SiGe bipolar technologies. Several high-speed chips in standard CMOS technologies were reported in [1–9], which confirm CMOS to be a viable alternative and a very economical approach for broadband circuit design.

As basic building blocks in optical fiber communication systems, current CMOS MUXs, DEMUXs, data decision circuits and frequency dividers have already achieved bitrates higher than 20-Gb/s or operating frequency higher than 20GHz [1–9]. In this section, a 50-Gb/s 1:2 DEMUX, a 50-Gb/s 2:1 MUX, a 42-Gb/s decision circuit and a 43GHz 2:1 frequency divider in 0.13µm CMOS are designed and simulated to validate
the identified design strategies and circuit techniques.

5.3.1 Low-Supply Circuits Design

In this section, four key building blocks in a 0.13μm CMOS technology are described. Since such a high speed is well beyond the reach of conventional CMOS designs, circuit techniques identified in Chapter 4 are adopted to achieve the bandwidth and timing required for transmitting a 40-Gb/s NRZ bit stream. The schematic of proposed circuits are shown Figure 5.20, Figure 4.2, and Figure 4.11.

The 1:2 DEMUX consists of two MS-FFs and output buffers. Each MS-FF includes two latches connected in series. When a 40-Gb/s data stream is applied, the MS-FFs are clocked at 20GHz. To sample every bit of the 40-Gb/s input data, the clock of one MS-FF is in phase while the other one is inverted. A separate buffer for each output decouples the MSFFs from the 50Ω environment. All transistors in the core are LVT 120nm NMOS devices for low supply (1V) operation. The latches showed in Figure 5.20 (a) use series gating between clock and data inputs. Poly-silicon resistors are used as loads which is a compromise between high voltage swing and reasonable RC time constant. Clock input matching is realized with 50Ω on-chip resistors, which are connected to a DC level shifter (V_{DD}/2). The tail current of all latches is set to 7mA for 1.2V supply.

The MUX in Figure 5.20(b) uses low resistance poly-silicon resistors with peaking inductors in series as loads to achieve enough bandwidth. The tail current of the latch is set to 10mA for 1.2V supply. Due to the small load resistors and large tail current, only a single-stage output buffer is employed. To achieve sufficient input sensitivity, both devices in data path of the latch and the differential pair of buffer amplifier are large-size LVT NMOS. The adopted 10mA tail current and peaking inductors will compensate the excessive parasitic capacitance.

The decision circuit consists of a MS-FF in Figure 5.20 (c) and a 3-stage output
Figure 5.20: Proposed circuits. (a) Latch in DEMUX; (b) MUX with a buffer; (c) Data decision circuit core.

buffer in Figure 4.11 (b). The MS-FF is clocked at 40-GHz. The tail current of both latches is set to 6mA for 1.2V supply. To achieve full voltage swing, at least 20-GHz bandwidth is needed for the latches and output buffer. To enhance the bandwidth, SP and S-R are implemented using on-chip spiral inductors and poly-silicon resistors with an optimal resistance ratio ($R_1/R_2 = 3/2$), respectively.

The internal dividing function of this frequency divider is based on a MS-FF by connecting the inverted slave outputs to the master inputs. The tail current of both latches is set to 7mA for 1.2V supply. To achieve very high operating frequency and low clock jitter, S-R with a ratio of $R_1/R_2 = 1/1$ as shown Figure 4.2 (a) is used and SP are implemented using on-chip spiral inductors with high Q-value and high
Figure 5.21: Layout of (a) 1:2 DEMUX; (b) Data decision Circuit.

Since fully differential topology is applied, the layout is devised to be maximally symmetrical to keep the circuit as balance as possible for high immunity against common-mode disturbances. Furthermore, all interconnects are kept as short as possible. Especially the lines between slave outputs and master inputs of clock divider are affecting the maximum operation frequency, because of their propagation delay and capacitive load. For DEMUX shown in Figure 5.21 (a), input signal path and output signal paths are routed perpendicularly to minimize possible crosstalk; Data pads and clock pads are separated as far as possible to lower possible interference. Conversely, data signal path and clock signal paths are routed perpendicularly, input pads and output pads are placed at the left-end and the right-end respectively considering layout symmetry and low crosstalk for data decision circuit shown in Figure 5.21 (b). Additionally, “SGS” and “PSGSP” pad patterns are used for high symmetry, low disturbance and easy on-chip test, where “P”, “G”, “S” represent “power supply”, “ground” and “signal” terminals, respectively.
Figure 5.22: 20-Gb/s DEMUX output eye-diagrams: (a) DEMUX core without peaking; (b) Buffered DEMUX without peaking; (c) DEMUX core with peaking; (d) Buffered DEMUX with peaking.

5.3.3 Circuit Simulations

To verify the proposed digital blocks, circuit simulations are carried out using the simulator, Cadence Spectre, and BSIM4 model based on IBM 0.13 μm CMOS technology. Differential pseudo-random bit sequences (PRBS) of $2^{31}-1$ is used as data signal and differential sinusoidal waveform is employed as clock signal for transient analyses.
A. DEMUX

As shown in Figure 2.1 (a), the employed on-chip inductors have a quality factor above 14 in a frequency range from 8GHz to 20GHz and are implemented by topmost metal with pattern ground shield to achieve a higher quality factor. The used inductors can improve the bandwidth by approximately 50% in practice rather than 72% in the ideal case, which confirmed by the simulated data given in Figure 5.22 (c).

Figure 5.22 shows the simulated eye-diagrams of the differential output signal at a data rate of 20-Gb/s. Based on Figure 5.22, we can see that the used SP inductors along with the employed 2-stage buffers greatly boosted the bandwidth and improved the output waveform, which resulted in an increase of effective signal amplitude from $2 \times 140mV_{pp}$ to $2 \times 400mV_{pp}$ and a decrease of peak to peak (PP) jitter from 11.8ps to 7.2ps. Further simulations are done for higher data rates and PVT variations,
The proposed DEMUX can operate up to 50-Gb/s under supply voltages from 1.0V to 1.5V (Simulated with process corners) and can operate well at 40-Gb/s under various PVT conditions as demonstrated in Figure 5.23 and 5.24. It can be seen from Figure 5.23 and 5.24 that the employed stacked current source offer very good immunity against PVT fluctuations and the circuit bandwidth are greatly improved by the used on-chip inductors. Combined above circuit techniques together, the designed DEMUX can operate at 40-Gb/s and beyond under a lower supply and consume less power. The simulation data are summarized in Table 5.4.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Bitrate (Gb/s)</th>
<th>Ratio</th>
<th>$P_{\text{total}}$ (mW)</th>
<th>Supply (V)</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>40</td>
<td>1 : 4</td>
<td>62</td>
<td>1.2</td>
<td>90nm</td>
</tr>
<tr>
<td>[2]</td>
<td>40</td>
<td>1 : 2</td>
<td>108</td>
<td>1.5</td>
<td>120nm</td>
</tr>
<tr>
<td>This work</td>
<td>50</td>
<td>1 : 2</td>
<td>25-87</td>
<td>1.1-1.5</td>
<td>0.13μm</td>
</tr>
</tbody>
</table>

**B. MUX**

As illustrated in Figure 5.25, the proposed MUX can achieve a bit rate of 40-Gb/s with an output amplitude of $2 \times 500mV_{pp}$ and a PP jitter of 3.8ps.
Figure 5.26: Transient waveforms of the proposed MUX (40-Gb/s operating).

Figure 5.27: PVT simulations of the proposed MUX (40-Gb/s data output).

The proposed MUX can operate up to 50-Gb/s under supply voltages from 1.0V to 1.5V (Simulated with process corners) and can operate well at 40-Gb/s under various PVT conditions as demonstrated in Figure 5.26 and 5.27. These figures show that the designed MUX has very good immunity against PVT fluctuations and enough bandwidth for 40-Gb/s and beyond operating. The main performance of this MUX is listed in Table 5.5.
Figure 5.28: Voltage effect on self-resonant frequency $f_{SR}$ of the proposed frequency divider: (a) $f_{SR}$ versus $V_{DD}$; (b) $f_{SR}$ versus $V_{CM}$.

Table 5.5: Performance Comparison of CMOS MUX

<table>
<thead>
<tr>
<th>Reference</th>
<th>Bitrate (Gb/s)</th>
<th>Ratio</th>
<th>$P_{total}$ (mW)</th>
<th>Supply (V)</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>40</td>
<td>4:1</td>
<td>132</td>
<td>1.2</td>
<td>90nm</td>
</tr>
<tr>
<td>[2]</td>
<td>40</td>
<td>2:1</td>
<td>100</td>
<td>1.5</td>
<td>120nm</td>
</tr>
<tr>
<td>This work</td>
<td>50</td>
<td>2:1</td>
<td>15-42</td>
<td>1-1.5</td>
<td>0.13μm</td>
</tr>
</tbody>
</table>

C. Frequency Divider

As shown in Figure 5.28 (a), the operating frequency of a conventional CML static frequency divider with resistor loads decreases very slightly with the increase of its supply voltage, $V_{DD}$. Since with the increase of supply voltage both the charging current and signal amplitude go up, the required charging time almost keep a constant, which means the operating frequency just varies a little bit from $V_{DD} = 1$V to 1.6V. However, higher supply voltage will result in larger power dissipation. So it is possible to implement high speed frequency dividers with much lower power consumption. Based on Figure 5.28 (b), the operating frequency decreases very slightly with the increase of its input common-mode voltage, $V_{CM}$, which means that the proposed CML static frequency divider has a wide range of input common-mode voltage and very good immunity against input common-mode voltage and supply voltage variations. Based on Figure 5.28, CML static frequency divider with reduced internal
signal swing can work at much higher frequency than dynamic CMOS logic frequency divider with rail-to-rail signal amplitude shown in Figure 3.26.

Figure 5.29: Simulated input sensitivity curves of 3 frequency dividers: V1-Traditional static divider; V2-Static divider with SP; V3-Static divider with SP and S-R.

Simulated input sensitivity curves of three frequency divider topologies operating under 1.2V supply are given in Figure 5.29. The frequency divider with SP and S-R (V3) operates at the highest frequency of 43.2GHz, which confirms that the employed on-chip spiral SP inductors and S-R loads can improve the operating speed of CML circuits up to 50%. On one hand, shunt peaking inductors and split-resistor loads can greatly boost the circuit bandwidth by partly tuning out the parasitic capacitances at output nodes, further reducing internal signal swing and weakening undesired Miller Effect as verified by simulation data in Figure 5.29. On the other hand, we have to pay more layout area (more passive devices) and higher complexity for the improvement of operating speed.

The transient waveforms of traditional CML static frequency divider, CML static frequency divider with SP, and CML static frequency divider with SP and S-R are illustrated in Figure 5.30. All the proposed frequency dividers have very good waveform and enough output signal amplitude, but different operating frequencies due to the use of different bandwidth boosting techniques.
As shown in Figure 5.31, frequency divider V3 has a little higher power dissipation due to its higher self-resonant frequency predicted by Figure 2.4. However, it is worth paying 1.5mW power penalty for the 8GHz improvement of self-resonant frequency.

Figure 5.32 shows the transient self-oscillation waveforms of the frequency divider (V3) under different supply voltages. Unlike conventional static CML frequency divider, the proposed frequency divider with SP and S-R operates at different frequencies with different output voltage amplitudes because the used on-chip inductor appears high AC resistance under high frequency and high current density. As a result, higher supply voltage is desired for the frequency divider V3 to operate at very
high speed.

Figure 5.32: Starting self-oscillations of the frequency divider with SP and S-R under different supply voltages: (a) $V_{DD}=1\text{V}$; (b) $V_{DD}=1.2\text{V}$; (c) $V_{DD}=1.5\text{V}$.

Figure 5.33: Transient waveforms of the proposed frequency divider (V3) at 42GHz.

Figure 5.33 illustrates that the frequency divider with SP and S-R can operate at 42GHz with very good voltage waveform and enough signal amplitude under a 1.2V supply. Simulation results of the proposed frequency divider (V3) and the comparison to previous work are given in Table 5.6. The designed frequency divider (V3) can operate under 1V supply in a wider frequency range.

**D. Data Decision Circuit**

Figure 5.34 shows the simulated eye-diagrams of the differential output signal at a data rate of 40-Gb/s. The effective signal amplitude is increased from $2\times300mV_{pp}$ to $2\times550mV_{pp}$ and the peak-to-peak (PP) jitter is decreased from 4.2ps to 1.7ps simultaneously using the output buffer shown in Figure 4.11 (b). Figure 5.35 illustrates the optimization of the resistance ratio $R_1/R_2$ of the employed S-R loads. Choosing
Table 5.6: Performance Comparison of CMOS Frequency Divider

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Supply (V)</th>
<th>Frequency Range (GHz)</th>
<th>Input Sensitivity</th>
<th>Power (mW)</th>
<th>Type</th>
<th>CMOS Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5]</td>
<td>1.1</td>
<td>31-41</td>
<td>390mV@40GHz</td>
<td>4</td>
<td>Dynamic</td>
<td>80nm</td>
</tr>
<tr>
<td>[6]</td>
<td>2.5</td>
<td>38-40.6</td>
<td>840mV@40GHz</td>
<td>4</td>
<td>Resonance</td>
<td>180nm</td>
</tr>
<tr>
<td>[7]</td>
<td>1.8</td>
<td>34.5-38</td>
<td>720mV@38GHz</td>
<td>2</td>
<td>Ring-Oscillator</td>
<td>120nm</td>
</tr>
<tr>
<td>[8]</td>
<td>1.5</td>
<td>5-27</td>
<td>630mV@25GHz</td>
<td>2</td>
<td>Static</td>
<td>120nm</td>
</tr>
<tr>
<td>[9]</td>
<td>1.2-1.5</td>
<td>5-25</td>
<td>1.50V@25GHz</td>
<td>2</td>
<td>Static</td>
<td>120nm</td>
</tr>
<tr>
<td>This</td>
<td>1.0-1.5</td>
<td>28-44</td>
<td>300mV@40GHz</td>
<td>2</td>
<td>Static</td>
<td>120nm</td>
</tr>
</tbody>
</table>

**Figure 5.34:** Data decision circuit operating at 40-Gb/s: (a) Input data; (b) Core circuit output; (c) Full circuit (with buffer) output.

$R_1/R_2=60/40$, a PP jitter of 2.2ps and an effective signal amplitude of $2 \times 580mV_{pp}$ are achieved. In the same way, a minimum PP jitter of 1.7ps can be obtained with a slightly reduced signal swing of $2 \times 550mV_{pp}$ through the optimization of the S-R loads in the second stage of the used output buffer.

Further simulations are done for higher bitrates and PVT variations, respectively. The proposed decision circuit can operate up to 42-Gb/s under supply voltage ranging from 1.0V to 1.5V shown in Figure 5.36 (Simulated with “TT” process corner and room temperature) and can operate well at 40-Gb/s under various PVT conditions shown in Figure 5.37. It can be seen from Figure 5.36 and Figure 5.37 that the employed stacked current sources offer very good immunity against PVT fluctuations.
5.3.4 Design Summary

In this section, low-supply digital circuits in 0.13μm CMOS are designed and simulated. All of them use fully-balanced CML topologies for super-high speed operation. Optimized DC bias level and transistor aspect ratio, on-chip shunt peaking coils, and split-resistor loads are the major contributors of bandwidth extension, and supply voltage reduction, which is confirmed by post-layout simulation data. Stacked current sources are employed to ensure that the proposed circuits have excellent immunity.
against PVT fluctuations, high reliability and manufactureability. Moreover, all the proposed circuits can work well under a lower supply (1V), which is compatible with digital circuits in CMOS logic realized in low-supply nano-scale CMOS technologies.

Table 5.7: Performance Comparison of CMOS Decision Circuit

<table>
<thead>
<tr>
<th>Reference</th>
<th>Bitrate (Gb/s)</th>
<th>( P_{\text{tach}} ) (mW)</th>
<th>( P_{\text{total}} ) (mW)</th>
<th>Supply (V)</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>[4]</td>
<td>37-40</td>
<td>10.8-20</td>
<td>13-240</td>
<td>1.2-1.5</td>
<td>90nm</td>
</tr>
<tr>
<td>This work</td>
<td>42</td>
<td>3.9-9.3</td>
<td>22-45</td>
<td>1.0-1.5</td>
<td>0.13( \mu )m</td>
</tr>
</tbody>
</table>

5.4 Chapter Summary

Using the identified design strategies and techniques, low-power monolithic analog circuits were fabricated and evaluated, and low-supply digital circuits were designed and simulated in this chapter. The measurement data and post-layout simulation data confirmed the validity of the adopted design strategies and techniques.
Chapter 6

Conclusions

6.1 Summary of Work in This Thesis

In this thesis, the relationship among semiconductor technology, circuit logic and circuit, block, even system specifications was reviewed. Low-power high-speed circuit topologies, techniques, and design bottlenecks were overviewed. Most importantly, advanced design strategies and key circuit techniques were identified to realize high-speed circuits with lower supply and power, smaller silicon area. Furthermore, CMOS low-power analog circuits were fabricated and measured, and low-supply CMOS digital circuits were designed and simulated which validated the adopted design strategies and used circuit techniques, topologies.

6.2 Summary of Thesis Contribution

Thesis contributions are in the area of design strategy to a trade-off among speed, power dissipation, and silicon area. Based on and extensive literature review in Chapter 3, key circuit techniques were identified as below:

1. **DC bias optimizing**: To improve effective $f_T$, to lower supply voltage and power consumption.
2. **Device sizing**: To reduce RC time constant and device mismatch, to increase input sensitivity, to lower supply voltage and power consumption.

3. **Split-resistor (S-R) load**: To improve gain-bandwidth product (GBW) by reducing critical node capacitance and internal signal swing.

4. **Compact L, R, C, and feedback topology**: To boost circuit bandwidth, to reduce silicon area, to improve circuit performance, process tolerance, integration density, and compatibility with digital circuits in CMOS logic.

Experimental verification allowed to confirm that a systematic application of these circuit techniques and optimization strategies indeed results in circuits with lower power dissipation or lower supply voltage, and higher operating speed.

### 6.3 Future Work

Low power high speed pre-emphasis and equalization circuits, low phase noise VCO and low supply CDR circuits for serial data communications can be the potential research interests, which will take the advantage of this thesis and give more design insights in this area.

### 6.4 Summary of Publications

During the period of study and research in Carleton University, 12 papers have been published and are listed below, which is the summary of my major work and contributions in this area.

1. **Bangli Liang**, Dianyong Chen, Bo Wang, Dezhong Cheng, Tad Kwasniewski, "A 43-GHz Static Frequency Divider in 0.13μm Standard CMOS,"


List of References


