INTERCONNECT DESIGN WITH LARGE TRANSISTOR CONSTRAINTS FOR MULTI-CHIP MODULES AND LARGE DIE SOI/SOS

By

JIE WEN

Bachelor of Science

Huazhong University of Science and Technology

WuHan, China

1989

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 1998

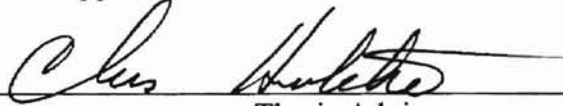
INTERCONNECT DESIGN WITH LARGE

TRANSISTOR CONSTRAINTS FOR

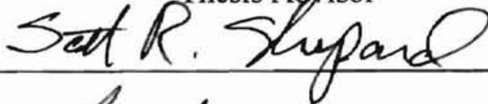
MULTI-CHIP MODULES AND

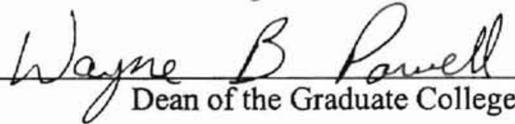
LARGE DIE SOI/SOS

Thesis Approved:



Thesis Advisor





Dean of the Graduate College

ACKNOWLEDGEMENTS

This study was conducted to develop a set of general design guidelines and performance estimation equations of long interconnects for high-speed multi-chip modules and large die SOI/SOS applications. Analysis and simulation results are presented. A novel layout for transistors with large width is also developed.

I wish to express my sincere appreciation to my major advisor, Dr. Chriswell Hutchens for his intelligent supervision, constructive guidance and inspiration. My sincere appreciation extends to my other committee members Dr. Louis G. Johnson and Dr. Scott Shepard. I would like to thank the Naval Ocean Surveillance Command for providing the research opportunity and financial support to this project.

Moreover, I wish to express my sincere gratitude to my colleagues who provided suggestions and assistance for this study.

I would also like to give my special appreciation to my wife, Li Zhou, for her strong encouragement at times of difficulty, love and understanding throughout this whole process. Thanks also go to my parents for their support and encouragement.

Finally, I would like to thank the Department of Electrical and Computer Engineering for supporting during these years of study.

TABLE OF CONTENTS

Chapter Page

1. INTRODUCTION.. 1

1.1 Objective.. 3

1.2 Organization.. 5

2. EXTRACTION OF R, L, C AND Z_0.. 7

2.1 Transmission Line.. 7

2.2 Selection of Material for the Substrate.. 9

2.2.1 Effective Dielectric Constant of Substrate.. 9

2.2.2 Selection of Material of Substrate.. 10

2.3 Extraction of Self Capacitance and Inductance of Microstrip.. 11

2.3.1 Calculation of Self Capacitance (*pF/cm*) and inductance Microstrip(*nH/cm*).. 11

2.3.2 Properties of Self Capacitance and Inductance of Microstrip.. 12

2.3.3 Comparison of Self Capacitance on Insulating Substrate and on Oxide-Passivated Silicon Substrate.. 14

2.4 Extraction of Self Resistance (Ω/cm) of Microstrip.. 15

2.4.1 Calculation of Sheet Resistance.. 15

2.4.2 Skin Effect.. 16

2.4.2.1 Calculation of Resistance with Skin Effect.. 16

2.4.2.2 Cutoff Frequency.. 17

2.5 Computation of Characteristic Impedance Z_0 (Ω).. 18

2.5.1 Characteristic Impedance Z_0 of Transmission Line........................ 18

2.5.2 A Comparison of the Microstrip Characteristic Impedance on Different Substrates.. 19

2.6 Points of Design.. 21

3. DESIGN OF TRANSMISSION LINES.. 23

3.1 Lossless, Lossy and Fully Lossy Transmission Line.............................. 24

3.1.1 Definition of Lossless, Lossy and Fully Lossy Transmission Line...... 24

3.1.2 Characteristic of Three Types of Transmission Line..................... 25

3.2 Optimal Delay Design of A Transmission Line.................................... 27

3.2.1 First Incidence Voltage.. 28

3.2.2 Lumped RLC Model for a Lossy Transmission Line..................... 29

3.2.3 The Delay Estimation for a transmission line.............................. 32

3.2.4 Delay Factor Sensitivity Analysis... 34

3.2.5 Oscillation of Series *RLC* Circuits... 35

3.2.6 Optimizing the Delay... 37

3.3 Peripheral Parts' Effect on Interconnection Performance........................ 39

3.4 Loss of Transmission Line.. 41

3.5 PSPICE Simulation of Transmission Line Signal Delay.......................... 43

3.6 Points of Design.. 46

4. DESIGN CONSIDERATIONS TO ELIMINATE SELF GENERATED NOISE... 48

4.1 Reflection Noise.. 49

4.1.1 Series Termination.. 49

4.1.2 Optimal Resistor Value of Termination..................................... 51

4.1.3 Signal Termination Simulation.. 54

4.2 Driver Size Design... 56

4.2.1 Driver Size Estimation Model.. 56

4.2.2 Simulation and Discussion... 58

4.3 Crosstalk.. 62
4.3.1 Calculation of Crosstalk... 63
4.3.2 Simulation of Cross-Talk.. 67
4.3.2.1 The Accuracy of Mathematical Model....................................... 68
4.3.2.2 Cross-Talk Vs. Spacing Between Two Signal Lines.................... 69
4.3.2.3 Cross-Talk Vs. Termination Resistance....................................... 71
4.3.2.4 Cross-Talk Vs. Transmission Line Driver Size............................ 72
4.3.2.5 Cross-Talk Vs. Transmission Line Length................................... 73
4.3.2.6 Comparison of Coplanar Waveguides and Adjacent Microstrips.... 74
4.3.3 Discussion and Reduction of Cross-Talk.. 76
4.4 Simultaneous Switching Noise (Ground Bounce)................................. 78
4.4.1 Analysis of Simultaneous Switching Noise..................................... 80
4.4.2 Forming Damped *RLC* Equivalent Circuit of Power Trace................. 83
4.4.3 Simulation of Ground Bounce.. 85
4.4.4 Summary of Ground Bounce Reduction... 91
4.5 Points of Design.. 92

5. NOVEL LAYOUT DESIGN FOR LARGE TRANSISTOR.................................. 96

5.1 Overview of problems... 97
5.2 The Effect of R_g on f_T, f_{max}, Thermal Noise and Transient Response.............. 98
5.2.1 Cutoff Frequency f_T.. 99
5.2.2 Maximum Frequency of Oscillation f_{max}.. 100
5.2.3 Thermal Noise.. 102
5.2.4 Device Delay.. 103
5.3 The Effect of R_s on g_m, g_{ds} and Bandwidth.. 104
5.4 Novel Structure for Large Transistor.. 106
5.4.1 Selection of Basic Cell.. 107
5.4.2 Novel Layout Based on Triangle and Diamond Cell............................ 108
5.5 Comparison of Triangle and Finger Structure...................................... 113
5.6 Conclusion.. 115

6. CONCLUSION AND SUGGESTION 116

6.1 Conclusion 116

6.2 Test Structure 120

6.3 General Design Procedure and Suggestions 122

BIBLIOGRAPHY 128

LIST OF TABLES

Table Page

2.1 (a)(b) Example cal. Parameter data of a lossy transmission........................... 21

3.1 Definition of lossless, lossy and fully lossy transmission line.......................... 25

3.2 Sensitivity of line delay to various parameters 34

3.3 Calculation and simulation data of signal delay for a transmission line.............. 44

4.1 Series termination simulation data.. 53

4.2 The effect of driver width on signal risetime.. 59

4.3 Risetime calculation and simulation data with different line lengths................. 60

4.4 Calculation and simulation data for mathematical model accuracy.................... 68

4.5 Simulation and calculation data of crosstalk vs. Spacing................................ 70

4.6 Simulation data of crosstalk vs. termination resistance R_t.............................. 72

4.7 Simulation data of crosstalk vs. driver size.. 73

4.8 Simulation and calculation data of crosstalk vs. different line lengths................ 74

4.9 Crosstalk simulation data for waveguides with separating GND trace
and adjacent microstrip lines with separating GND trace.............................. 75

4.10 Ground bounce simulation data.. 88

5.1 The extracted data comparison between triangle and finger structure.............. 114

LIST OF FIGURES

Figure Page

1.1 A simple diagram of interconnection........ 4

2.1 Cross-sections of typical transmission line structure........ 8

2.2 Cross-sections of microstrips on insulating substrate (a) and oxide-passivated silicon (b)........ 11

2.3 Microstrip self capacitance and inductance versus line width (w) and substrate height (h)........ 13

2.4 Self capacitance on oxide-passivated silicon, sapphire and GaAs........ 14

2.5 Effect of parallel loads on characteristic impedance........ 19

2.6 Properties of characteristic impedance........ 20

3.1 The basic lumped *RLC* segment for a lossy transmission line........ 30

3.2 (a) "Tee" type symmetric lumps; (b) "Pi" type symmetric lumps; (c) "Gamma" type lumps........ 31

3.3 Lumped and distributed model for interconnection........ 33

3.4 Typical series *RLC* circuits and typical response........ 36

3.5 Delay versus line width........ 38

3.6 Graphical illustration of bond wire delay........ 39

3.7 Interconnect equivalent circuit with bond wire and pads........ 40

3.8 Equivalent *RLC* circuit of input and output ends........ 41

3.9 (a) Resistance attenuation constant versus line width; (b) Resistance attenuation versus operating frequency........ 43

3.10 Delay simulation schematics and segments........ 44

3.11 Delay simulation results (a), (b), (c) and (d)........ 45

3.12 Comparison of calculation and simulation data for signal delay........ 45

4.1 (a) Parallel termination; (b) Series termination........ 49

4.2 (a) Second-incidence switching problem of series termination with multiple loads; (b) Parallel fan-out with multiple series termination........ 51

4.3 The function of termination resistance, R_t, vs. (a) load capacitance C_L; (b) line resistance, R_{line}; (c) Driver channel resistance, R_{on}........ 53

4.4 Simulation waveform of series termination
(a) Optimal termination with different transmission line length;
(b) Different termination resistance with fixed line length i.e. line resistance........ 55

4.5 Estimation model of signal risetime........ 57

4.6 The simulation results of signal risetime (a) risetime versus transmission line length and (b) risetime versus driver size........ 59

4.7 Risetime simulation schematics and segment........ 59

4.8 Risetime comparison between calculation and simulation (a) $Z_0=50\Omega$ and (b) $Z_0=75\Omega$ with different load capacitance, termination resistance and line width 61

4.9 Risetime comparison between cal. and sim. with different driver size.............. 61

4.10 Cross-sections of (a) coplanar waveguides; (b) microstrips........................... 64

4.11 Circuit model for crosstalk due to capacitance coupling............................ 65

4.12 Capacitance crosstalk of coplanar waveguides versus ground line width and signal frequency.. 66

4.13 Capacitance crosstalk simulation (a) Simulation schematics; (b) One sample simulation waveform.. 67

4.14 Crosstalk comparison between simulation and calculation data..................... 69

4.15 Capacitive crosstalk versus spacing of two signal lines............................ 70

4.16 Crosstalk versus termination resistance... 72

4.17 Crosstalk versus driver size.. 73

4.18 Crosstalk versus line width... 74

4.19 Effect of GND trace on crosstalk... 75

4.20 Equivalent circuit and switching current model....................................... 80

4.21 Equivalent *RLC* circuit for power supply system...................................... 83

4.22 Ground bounce simulation circuit and typical response............................. 85

4.23 Ground bounce simulation results (a), (b), (c), (d) and (e)........................... 87

4.24 Ground bounce vs. Decoupling capacitance to load capacitance ratio.............. 88

5.1 Equivalent circuit models used in analysis of R_g effect on transistor performance. (a) lumped model; (b) small-signal model of a MOSFET.......................... 98

5.2 Calculated f_T and f_{max} data versus test data.. 102

5.3 R_s effect on $g_{m,}$ g_{ds} and bandwidth.. 105

5.4 (a) g_m versus transistor width; (b) g_{ds} versus transistor width......................... 106

5.5 (a) Polygon with n identical edges; (b) Polygon perimeter versus number of edges.. 107

5.6 Triangle and diamond basic cell and layout structure. (a) triangle cell; (b) layout structure based on triangle cell; (c) diamond cell; (d) layout structure based on diamond cell.. 109

5.7 Comparison of triangle and diamond structure (a), (b), (c), (d) and (e)............ 113

5.8 Side by side comparison of layouts with triangle and finger structure.............. 113

6.1 Test structure for transmission line and large transistor.............................. 121

NOMENCLATURE

c_0	Speed of Light in Vacuum
C	Line Capacitance per Unit
$C_{coupling}$	Mutual Capacitance for Crosstalk
C_{gs}	Gate to Source Capacitance of Transistor
C_{db}	Drain to Body Capacitance of Transistor
C_{line}	Total Line Capacitance
C_{lump}	Capacitance of Lumped *RLC* segment
C_{source}	Capacitance of Signal Source
C_{total}	Total Capacitance of Interconnection
C_D	Decoupling Capacitance
C_L	Load Capacitance
d	Interconnect Line Length
d_b	Bond Wire Length
d_{max}	Maximum Interconnect Line Length
f	Signal Operating Frequency
f_c	Cutoff Frequency of Skin Effect
f_T	Cutoff Frequency of Transistor
f_{max}	Maximum frequency of transistor
g_m	Transconductance Parameter
g_{ds}	Drain to Source Transconductance Parameter
h	Height of Substrate
L	Line Inductance per Unit
	Channel Length for Transistor
L_b	Bond Wire Inductance

L_{lump} Inductance of Lumped *RLC* segment

L_{line} Total Line Inductance

L_{total} Total Inductance of Interconnection

N Number of Lumped *RLC* segments

Q Quality Factor

r Radius of round Wire

R Line Resistance per Unit

R_g Gate Resistance of Transistor

R_s Source Resistance of Transistor

$R_{critical}$ Resistance set to "critical damping" for Series *RLC* circuit

R_{lump} Resistance of Lumped *RLC* segment

R_{line} Total Line Resistance

R_{sheet} Sheet Resistance of Metal Trace

R_{on} Turn-on Resistance of Transistor

R_{out} Output Impedance of Driver

R_S Source Output Resistance

R_t Termination Resistance

R_{teff} Effective Termination Resistance

R_{total} Total Resistance of Interconnection

S Spacing Between two Parallel Traces

t Thickness of Interconnection Line

t_b Bond Wire Induced Delay

t_d Delay of Interconnect including Buffer Chain

t_{d1} Delay of interconnect

t_{db} First Stage Delay of Buffer Chain

t_{ds} Delay of One Lumped *RLC* Segment

t_f Time-of-flight

t_{ox} Thickness of Oxide Layer

t_r Signal Risetime

V_{12} Crosstalk Voltage

V_{first} First Incidence Voltage

V_{in} Input Voltage

V_A Fast Rising Portion of Unit Step Signal Response for a Transmission Line

V_{dd} Positive Power Supply

V_{GS} Gate to Source Voltage

V_{IH} Minimum Input High Voltage

V_{IL} Maximum Input Low Voltage

V_R Slow Rising Portion of Unit Step Signal Response for a Transmission Line

V_T Threshold Voltage

w Interconnection Width

w_{eff} Effective Width of Interconnection

w_g GND trace width

W Channel Width of Transistor

W_{eff} Effective Channel width of Transistor

Z_0 Characteristic Impedance of Transmission line

Z_{load} Effective Load Impedance

α Scaling Factor of Buffer Chain

α_D Dielectric Attenuation Constant

α_R Resistance Attenuation Constant with Skin Effect

α_S Resistance Attenuation Constant without Skin Effect

β Electrical "Length" of Lumped *RLC* segment

ε Dielectric Constant of Material

ε_0 Permittivity of Free Space

ε_{eff} Effective Dielectric Constant

ε_r Relative Dielectric Constant

μ Magnetic Permeability of Free Space

Transistor Self Gain

μ_r Relative Magnetic Permeability of Material

ρ Resistivity of Material

δ_s Skin Depth

ω Signal Frequency

ν	Signal Propagation Speed
τ	Signal Delay of per Unit
ΔC	Mutual Capacitance per Unit
Δv	Inductance Induced Voltage

CHAPTER 1

INTRODUCTION

The development of integrated circuit has gone through the stages of small-scale integration, medium scale integration and very large scale integration (VLSI). This technology revolution increases complexity and density of semiconductor devices allowing electronic systems to reach even higher performance. To keep pace with the advances in IC technology, a higher performance packaging and interconnection have been developed [5][7]. Large die SOI/SOS and multiple chip modules (MCMs) are developed as advanced system applications with complex functions and high speed. In large size single die and MCMs applications, the interconnection density and length increases dramatically. For example, the die size has grown up to $20mm \times 20mm$ [6] in a recent report. In MCMs, the wiring capacity is greater than $400cm/cm^2$, and the size reaches $127.5mm \times 127.5mm$ *(IBM ES900)* [1] with the maximum length reaching to tens of *cm*. Simultaneously, as the fabrication technology has improved, the device geometry has scaled down to $L_{eff} = 0.09\mu m$, with a highest unity-current-gain frequency of *100GHz* for *NMOS,* (f_T) [55]. Due to the greater line length, finer metal1/metal2 etc., small dielectric thickness and higher operating frequency (a few *GHz*), the interconnect can no longer be modeled as a lumped circuit and becomes a critical factor in large high-speed system design [2][3].

The electrical high performance constraints of an electronic system are the system working frequency (i.e. system bandwidth, settling time or clock frequency), signal

integrity and noise level. For a given architecture, the system must be designed to run at a maximum operating frequency correctly and reliably. To maximize operating frequency, the delay in critical paths, i.e. the paths with maximum delay, should be minimized. The system reliability is limited by the system noise level. Electrical noise may induce inadvertent logic transition, i.e. errors in the system. To achieve high performance, both maximum operating frequency and acceptable noise level are the most important issues in the design exclusive of process yield.

In an electronic system, total delay is the sum of various delays in the system. Delays caused by devices and interconnections are the two dominant factors. Device delays are reduced by improvements in fabrication technology. Since the fabrication technology has made significant advances in recent years, interconnect delays are becoming the dominant problem replacing gate delay. Interconnect capacitance can become larger than gate capacitance hence determining the overview speed performance of the system. This large capacitance also increases power consumption and noise. For a properly designed interconnect, delay is totally determined by the conductor material and the line geometry.

Circuit density is the maximum number of electronic elements that can be contained in a chip or MCMs [1][2]. It is determined by the wiring capacity. Wiring capacity is affected by several factors, one of which is wiring density. Wiring density is a function of minimum line width, line separation and via size. The characteristic length of a line with minimum line width, separation and via size is specified by electrical effects. Too high a density can result in unexpected noise. If an interconnect is not properly designed, signal reflects and shape degrades significantly. In addition, two parallel signal

lines must be separated by a specific minimum spacing to avoid excessive crosstalk noise.

In high-speed digital design, in contrast to low frequency design, the parasitic circuit elements of an interconnect, such as resistance, capacitance and inductance can not be neglected. These design issues are focused on the behavior of passive circuit elements, such as the wires, circuit boards and integrated-circuit packages. At low frequency, the effects of passive elements are typically neglected as just parts of a product's packaging. However, in high frequency application, they play a direct and important rule in electrical performance [2].

Interconnection has a greater and greater effect on system electrical performance. The proper interconnect (including packaging) design distributes to the achievement of high speed, high reliability and lower power dissipation in high speed and high-density applications. The effects of parasitic parameters of an interconnect line on signal propagation (ringing and reflection, etc.), interaction between signals (crosstalk) and other interference (simultaneous noise, etc.) has become very critical issues in designing valuable high speed system.

1.1 Objective

The objective of this thesis is to develop a novel layout for large transistor based on triangle cell to mitigate the gate and source resistance limits on transistor performance. A set of general design guidelines and theoretical parameter and performance estimation equations are also developed for interconnects with long lengths or used in high speed

applications (a few *GHz*) on MCMs and large *SOI/SOS* dies. This research is based on the investigation of interconnect parasitic parameters extraction, review of transmission line properties, signal delay analysis, crosstalk estimation, simultaneous noise calculation, termination resistance selection, proper driver design and novel layout realization. The simulation and layout-extracted results are also presented to support this proposal.

Figure 1.1 shows a simple interconnection diagram. The main parts include interconnection wires, loads, signal drivers and power supply. The design issues covered in this thesis are outlined as follows.

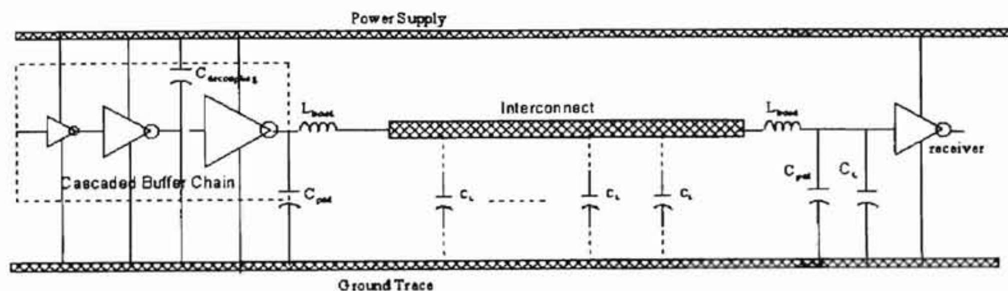


Fig. 1.1 A simple diagram of interconnection

1. Chapter 2. Extraction and calculation of electrical parasitic parameters of an interconnection.
2. Chapter 3. Interconnection types, lossless, lossy and fully lossy transmission line; and the application of each.
3. Chapter 3. Maximum line length and proper signal risetime applied.
4. Chapter 3. How to achieve the optimal delay; Line geometry design (width, selection of characteristic impedance).

5. Chapter 4. Proper termination methods (parallel or series) and the optimal termination resistance.
6. Chapter 4. Proper driver size and fanout topology design
7. Chapter 4. Crosstalk reduction and effect of distribution parameters.
8. Chapter 4. More issues on line geometry design (line spacing and termination)
9. Chapter 4. Controlling ground bounce, estimating decoupling capacitance and effect of distribution parameters.
10. Chapter 5. Limitations in the design of transistor with large effective width. A novel transistor layout and its advantages.

1.2 Organization

Chapter 1 introduces the background and proposes for this study.

Chapter 2 reviews the extraction of parasitic parameters, i.e. self capacitance, inductance, resistance and characteristic impedance of an interconnect. The effect of substrate material, line width, substrate height and operating frequency is also discussed.

Chapter 3 discusses basic issues of a transmission line design, especially in optimal delay design. Three different types of transmission lines and their characteristics are discussed. A distributed Lumped *RLC* model is used for lossy transmission line simulation. Optimal delay design issues, including delay estimation, peripheral part effect, i.e. bond wire and pads, and parameter sensitivity to delay, are discussed and simulation results are presented. The signal loss for very long interconnects is also included in this chapter.

Chapter 4 analyzes the control of reflection, crosstalk and simultaneous noise. The proper driver size design is also covered along with the proper selection of a termination type and resistance value. In section 4.1, signal reflection and termination method is studied and simulation results are presented as verification. In section 4.2, proper driver size design is discussed. In section 4.3, a crosstalk estimation model is reviewed and the effect of a GND separation, termination resistance, driver size and line length on crosstalk are discussed in detail. The simulation results confirm the selected model. In the last section, a simultaneous noise model is presented along with analysis and verification through simulations. In conclusion, simple guidelines for proper termination, GND trace width design for reduction of crosstalk and higher trace density, decoupling capacitance selection and proper driver size design are given.

Chapter 5 presents a novel layout for large transistor based on the analysis of gate and source resistance limitations on transistor performance, specially cutoff frequency f_T, maximum frequency of oscillation f_{max} [4], thermal noise, gate delay, transconductance g_m and output conductance g_{ds}. The extracted results and comparison with a conventional finger structure are presented.

Chapter 6 summarizes the conclusion and recommendations. A test structure and general design flow diagram are presented.

CHAPTER 2

EXTRACTION OF R, L, C AND Z_0

With the increase of circuit operating frequency and the advent of deep submicron technologies, the VLSI interconnects become one of the most important limiting factors in high-speed and high-density circuit performance. For system-level designs, such as large on-chip circuits or VLSI, multi-chip modules (MCMs) and printed circuit boards (PCBs), the interconnects can induce considerable delays and coupling noise due to transmission line effects. In general, the interconnect delay and coupled noise must be considered in the evaluation of total system performance.

To model the interconnect effects, the electrical parameters must be first extracted. The general description of the electrical parameters of an interconnect, assuming it is a transmission line, requires five electrical parameters: line capacitance, line inductance, line series resistance, line shunt conductance and line characteristic impedance [8][9]. In VLSI and MCMs applications, the shunt conductance can generally be neglected without losing generality. These parameters are determined by the interconnect geometry and play a very important role in the analysis of interconnect performance, such as generated noise and delay. The following parts of this chapter give out a general and relatively simple estimation of interconnect electrical parameters.

2.1 Transmission Line

Interconnects in VLSI and MCMs applications can be represented by several different models. The *RC* and transmission line models are the two most frequently used models. In the *RC* model, only resistance and capacitance effects are considered, but in the transmission line model, the effect of line inductance is included. The *RC* model is normally adequate for lower clock frequencies and short interconnects giving fairly accurate results for those applications. However, if the interconnect is sufficiently long or the clock frequency is sufficiently fast, a transmission line model must be used since the *RC* model is no longer sufficient. In Chapter 3, the definition for a *RC* and transmission line model will be discussed in detail. In MCMs applications, interconnects are normally considered as transmission lines.

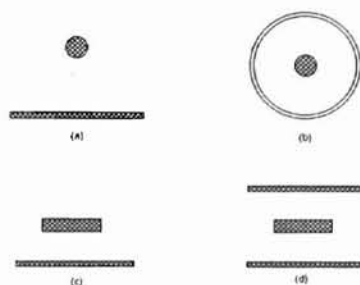


Figure 2.1 Cross-sections of typical transmission line Structure

Figure 2.1 presents the cross-sections of different transmission line structures which are used to model the different interconnects in different electrical applications. The wire above ground model (2.1a) can be used to represent a bond interconnect. Figure 2.1b is a coaxial model used for local area network or test probe connections. The on-chip interconnect is normally modeled as microstrip line (2.1c) and both microstrip and strip line (2.1d) are used in MCMs and PCBs. In recent years, the microstrip has been used extensively because it is easier to be fabricated and supplies a free and

accessible surface for solid-state device. A strip line is more expensive than microstrip. In this thesis, all transmission lines are microstrips.

2.2 Selection of Material for the Substrate

The selection of the substrate on which the die will be mounted is one of the considerations for MCMs application. In the following sections, it will be illustrated that some electrical parameters of interconnect like line capacitance C_{line} and characteristic impedance Z_0 are dependent on the dielectric constant of substrate. Different choices of a substrate material result in different signal delays and losses. Reference [10] gives out the advantages and limitations of two commonly used substrate materials, organic and ceramic. In this part, The guideline of MCMs substrate selection is illustrated.

2.2.1 Effective Dielectric Constant of Substrate

The relative dielectric constant is defined as $\varepsilon_r = \varepsilon / \varepsilon_0$ $(\varepsilon_0 = 8.854 \cdot 10^{-12} F/m)$. The effective dielectric constant depends on the selected material, line width and substrate height. J. Howard gives out the following formula for effective ε_r as in (2.2.1) [2],

$$\varepsilon_{eff}(w_{eff}) = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[\left(1 + \frac{12 \cdot h}{w_{eff}}\right)^{-0.5} + 0.04 \cdot \left(1 - \frac{w_{eff}}{h}\right)^2\right] \qquad (2.2.1)$$

$$w_{eff} = w + \frac{1.25 \cdot t}{\pi}\left(1 + \ln\left(\frac{4 \cdot \pi \cdot w}{t}\right)\right) \qquad (2.2.2)$$

where h is the height of substrate, w_{eff} is the effective width of signal microstrip, t is the thickness of signal microstrip and ε_r is dielectric constant of substrate material. Fig. 2.2 shows the typical cross-section of microstrip.

Dielectric constant has effect on signal delay and loss. More detail discussion will be given in Chap. 3. Normally, the dielectric loss can always be neglected in MCMs applications when maximum frequency is less than 10GHz [11][15]. But large magnetic loss should be considered for *Si* in high frequency application.

2.2.2 Selection of Material of Substrate[11]

- ε_r should remain constant over the frequency and temperature range of interest.
- ε_r often reduces slightly (*~5%*) as frequency increases from dc to high frequency. The value of ε_r around *100M~10GHz* should be used for simulating high-speed signal propagation, not the dc value.
- Dielectric resonance may exist at microwave frequency (*>>10GHz*), but is not general concern for MCMs applications.
- Low dissipation factor or loss tangent ($\tan\delta$).
- For Polymer dielectrics, ε_r increases with ambient humidity.
- High thermal conductivity.
- Low expansion in *X-Y* (substrate) plane.
- ε_r should be as small as feasible. In this thesis, $\varepsilon_r = 2.6 \sim 4.3$ for MCMs; $\varepsilon_r = 4.0$ for SiO_2; $\varepsilon_r = 10.5$ for sapphire.

2.3 Extract Self Capacitance and Inductance of Microstrip

Line capacitance and inductance are very important electrical parameters of interconnects. They determine the interconnection delay and coupling noise (mutual capacitance and inductance).

2.3.1 Calculation of Self Capacitance and Inductance of Microstrip (*pF/cm*)

Figure 2.2 shows the cross-sections of microstrips fabricated on insulating substrate and oxide-passivated silicon substrate.

Fig. 2.2 Cross-sections of microstrips on insulating substrate (a) and oxide-passivated silicon

In the above figure, w is microstrip width, h is the substrate height ($100\mu m \sim 250\mu m$ for SOS and Bulk applications; $10\mu m \sim 30\mu m$ for MCMs application [15]). t_{ox} is the oxide thickness (field oxide plus bulk oxide) for SOI ($5000A^o \sim 10000A^o$). Under such conditions, the formulas suitable for calculating capacitance and inductance of an microstrip which can be derived from characteristic impedance of a microstrip transmission line have been studied previously by Scheider [12] and Hassegawa *et al.* [13]. These formulas are:

$$C = \frac{2\pi\varepsilon_{eff}\varepsilon_0}{\ln\left(\frac{8h}{w_{eff}} + \frac{w_{eff}}{4h}\right)} \qquad r_0 = 1(\lambda \Delta V g_m \tag{2.3.1}$$

$$C = \varepsilon_r\varepsilon_0\left[\frac{w_{eff}}{t_{ox}} + 2.42 - 0.44\frac{t_{ox}}{w_{eff}} + \left(1 - \frac{t_{ox}}{w_{eff}}\right)^6\right] \quad w_{eff} >> t_{ox} \tag{2.3.2}$$

$$L = \frac{\mu_0}{2\pi}\ln\left(\frac{8h}{w_{eff}} + \frac{w_{eff}}{4h}\right) \tag{2.3.3}$$

Equation (2.3.1) is used for a microstrip on an insulating substrate (SOS and MCMs) and (2.3.2) is used for a microstrip on an oxide-passivated silicon substrate (SOI). The same equation (2.3.3) can be used to estimate the inductance for a microstrip for SOS/SOI and MCMs with proper ground plane [14].

2.3.2 Properties of Self Capacitance and Inductance of Microstrip

The plots in fig. 2.3 shows capacitance and inductance change versus line width and substrate height. The selected substrate materials are sapphire, gallium arsenide and oxide-passivated silicon. The height (*h*) used for calculation is $25\mu m$ for MCMs and $250\mu m$ for SOS. t_{ox} ($6000A^o$) is used for SOI. Induced capacitance reduces quickly as the height of substrate increases for SOS and MCMs applications. To keep capacitance to a small value, larger substrate height and narrow line width are preferred. When line width is much smaller than the height of substrate, i.e. h/w>>3, the capacitance reduces slowly. This implies that the capacitance can't be reduced more when the thickness of

substrate increases beyond $h/w >> 3$. For SOI, thicker isolating layer and narrower line induces smaller capacitance.

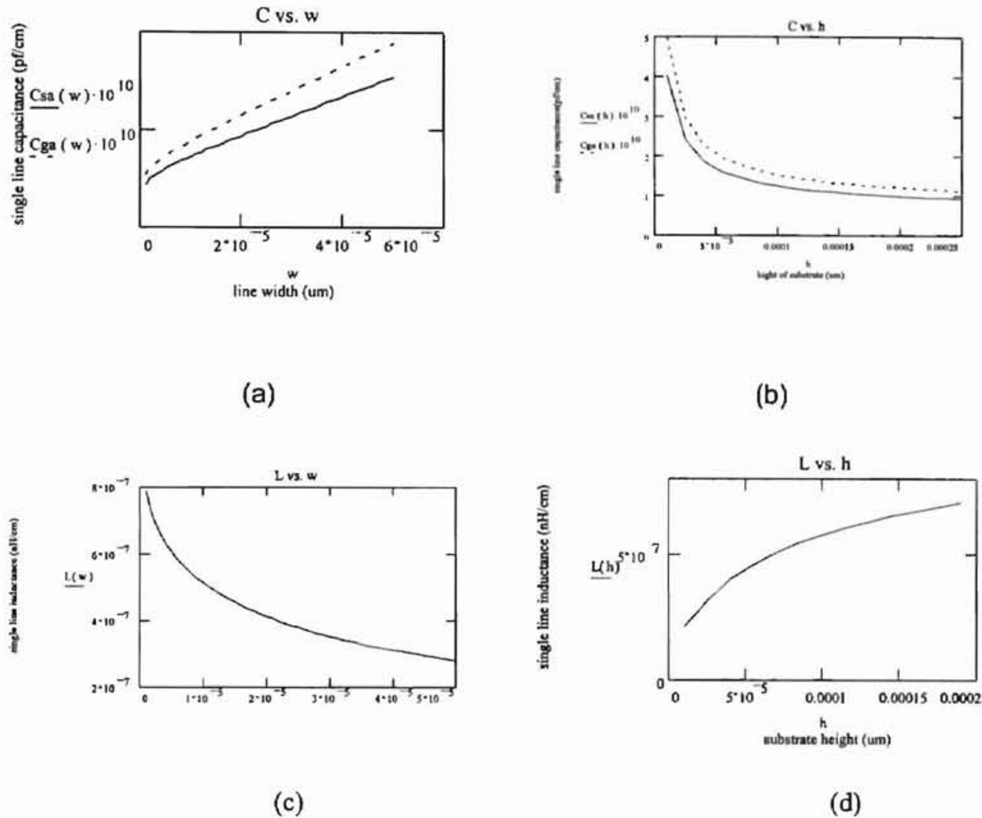


Fig 2.3 Microstrip self capacitance and inductance versus line width (w) and substrate height (h)

The inductance for SOS and MCMs reduces as the width of the transmission line increases. To keep the inductance small, wider lines are preferred. However this will increase the capacitance and results in decreased wiring capacity. The inductance for SOI wiring increases as line width increases and substrate height decreases ($L \propto \ln(w/4h)$)). A narrow line has both a smaller capacitance and inductance for a SOI system. In high-speed application, the reduction of inductance is a very important issue.

The proper selection of line width is determined by which parameter, capacitance or inductance, is dominant in system performance.

Besides the self-capacitance and inductance, there is also mutual capacitance and inductance between two or more adjacent parallel transmission lines. We will discuss the model and resulting design method in the "crosstalk" section in Chapter 4.

2.3.3 Compare the Self-capacitance on Insulating Substrate and on Oxide-passivated Silicon Substrate

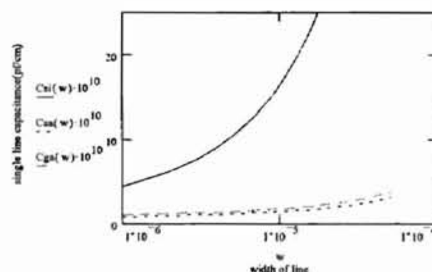


Fig. 2.4 Self capacitance on oxide-passivated silicon, sapphire and GaAs

Figure 2.4 shows that the line capacitance increases as the line width increases on all three substrates. The interconnection made on an insulating substrate (Sapphire, GaAs) induces considerably small capacitance (*0.5pF~5pF*). The capacitance on silicon substrate is larger than that on insulating substrate. However, this advantage diminishes at small line widths. The capacitances are scaled differently, being a logarithmic function on the insulator but a linear function of line width on the silicon substrate. The reason for larger capacitance on silicon is the thin oxide layer, t_{ox}, forming a large capacitor. For the purpose of capacitance estimation of on-chip interconnects, the thickness of field oxide (bulk) or thickness of field oxide plus buried oxide (*Simox)* should be used.

2.4 Extract Self Resistance of Microstrip (Ω/cm)

2.4.1 Calculation of Sheet Resistance

The resistance of conductors in MCMs and large die SOI/SOS has a significant effect on signal integrity, delay and termination. At low frequency, the line resistance is its dc resistance and is given by equation (2.4.1), where ρ is conductor material resistance, t is line thickness, R_{sheet} is the sheet resistance of metal trace, $R_{sheet} = \rho / t$, d *and* w_{eff} are line length and width respectively.

$$R_{line} = \rho \frac{d}{w_{eff} t} = R_{sheet} \frac{d}{w_{eff}} \qquad (2.4.1)$$

D.B. Tuckerman [15] points out the following issues should be noted in MCMs applications:

- Thin film resistivities are greater than bulk resistivities and depends on the grain size and impurities.
- Alloy resistivities are greater than pure metal resistivities.
- High temperature resistivities are greater than that at room temperature. For *Al* and *Cu*, it is normally *0.3%~0.4%/°C*. This resistance increase can be significant at high temperature applications.
- Conductors often have a non-rectangular cross-section. To estimate resistance more accurately, more complex models should be used.

Oklahoma Sta

- Line width and line thickness may vary from design value by up to *10%* because of process factors.

2.4.2 Skin Effect

At low frequencies, the current is distributed uniformly throughout the cross section of conductor through which it flows; however, at high frequency, the current distribution can be imagined as concentric tubes. The inner rings have more inductance than the fatter outer rings. The current follows the path with least inductance at high frequency, i.e. the current is concentrated on the surface of conductor. This is known as skin effect.

2.4.2.1 Calculation of Resistance With Skin Effect

A constant called skin depth δ_s is used to measure the skin effect. It is the distance at which current density becomes a fraction $1/e$ of its value at the surface. The skin depth δ_s is expressed as (2.4.2). It is inversely proportional to the square root of frequency. Here, f is the operating frequency and $\mu = 4\pi \times 10^{-7}$ Henrys/meter. For *Al* and *Cu*, $\rho = 2 \sim 3\mu\Omega \cdot cm$ [15]

$$\delta_s = \sqrt{\frac{\rho}{\pi \mu f}} \qquad (2.4.2)$$

So the effective resistance of transmission line with skin effect is

$$R = \rho \frac{d}{w_{eff}\delta_s} = \frac{d}{w_{eff}}\sqrt{\pi\mu f \rho} \qquad (2.4.3)$$

For example, for a $3.3\mu m$ plated *Cu* trace, $\rho = 2.8\mu\Omega \cdot cm$, when *f=1GHz*, $\delta_s = 2.6\mu m$; when *f=10GHz,* $\delta_s = 0.8\mu m$ and *25%* larger for *Al* on die. As the signal frequency increases, skin depth becomes thinner resulting in increased line resistance.

Equations (2.4.1) and (2.4.3) shows that at low frequencies, a conductor has a constant dc resistance; while at high frequency, resistance grows proportional to the square root of frequency.

2.4.2.2 Cutoff Frequency

There is a cutoff frequency [2][15] for any interconnect with a specific thickness, at which the conductor thickness is equal to the skin depth. It is defined as (2.4.4). For every interconnect, the cutoff frequency, f_c, offers a criterion to determine if the skin effect can be ignored.

$$f_c = \frac{\rho}{\pi\mu t^2} \qquad (2.4.4)$$

When $f < f_c$, skin effect can be ignored. The proper thickness of conductor is roughly equal the skin depth $t \approx \delta_s$ or $t \approx \sqrt{\rho/(\pi\mu f)}$ (equation 2.4.2) to achieve the maximum line length at the lowest resistance (see Section 3.2.1). Too thick a trace doesn't help to reduce resistance and increases material cost and signal distortion. The

Oklahoma State University Library

effect of skin effect on signal attenuation and selection of transmission line types is discussed in Chapter 3.

2.5 Computation of Characteristic Impedance Z_0

2.5.1 Characteristic Impedance Z_0 of Transmission Line

Characteristic impedance Z_0 is another basic parameter of transmission line. It is defined as

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} = \left(\frac{L}{C}\right)^{1/2} \sqrt{1 + \frac{R}{j\omega L}} \qquad (2.5.1)$$

where R is the resistance per unit along the line

L is the per unit inductance along the line

G is the per unit conductance shunting the line

C is the per unit capacitance shunting the line

When $R/j\omega L << 1$, i.e. $\omega >> R/L$

$$Z_0 \approx \sqrt{L/C} \qquad (2.5.2)$$

The characteristic impedance at lower frequencies is not real or constant. It has a substantial capacitive component. A "50Ω line" or "75Ω line" is only a meaningful concept at sufficient high frequency.

The different transmission line structures, such as microstrip, stripline, coaxial line, etc., have different approximate equations for Z_0 [2]. The microstrip Z_0 is approximately given by

Oklahoma State University Library

$$Z_0 = \frac{60}{\sqrt{\varepsilon_{eff}}} \ln\left(\frac{8h}{w_{eff}} + \frac{w_{eff}}{4h}\right) \quad (2.5.3)$$

Parallel capacitive loads and peripheral parts, such as bond wire and input/output pads, change the value of characteristic impedance. In the case shown in figure 2.5, the effective Z_0 is given by equation (2.5.4). Note, the non-transmission line parameters (L_b, C_{pad}, C_{source}) should be comparable to the transmission line per unit *L and C* in this case.

$$Z_{eff} = \sqrt{\frac{L_{line} + 2L_b}{C_{source} + C_{line} + NC_L + 2C_{pad}}} \quad (2.5.4)$$

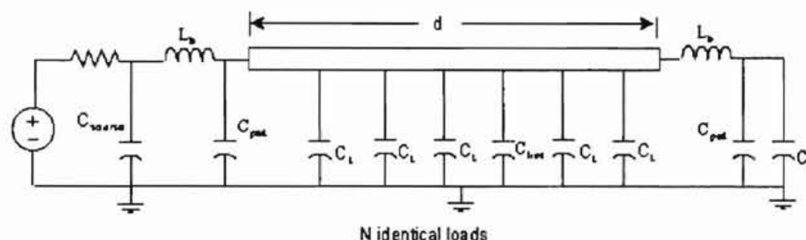


fig. 2.5 Effect of parallel loads on characteristic impedance

2.5.2 A comparison of the microstrip characteristic impedance on different substrates

Figure 2.6 gives out some information of characteristic impedance (Z_0). Microstrips with same geometry on different substrate have different Z_0 (Plot (a)). Plot (b) compares two Z_0 estimation models (equation (5.2.2) and (5.2.3)) and Plot (c) compares Z_0 and line resistance and gives important information for Z_0 selection.

- Plot (a) presents the change of Z_0 on different substrates with the change of line width. Z_{0si}, Z_{0sa} *and* Z_{0ga} are characteristic impedances of microstrips on *Si*, Sapphire and GaAs respectively. The plot implies Z_0 reduces and the change becomes slower as the line width increases. A line built on sapphire has lower characteristic

Oklahoma State University Library

impedance than that on a silicon substrate. To get a smaller Z_0, larger dielectric constant should be chosen.

- A comparison of Z_0 estimation using equations (2.5.2) and (2.5.3) is shown in plot (b). The equation (2.5.2) is simple and accurate enough and is also useful in line inductance estimation using TDR (See Chapter 3).

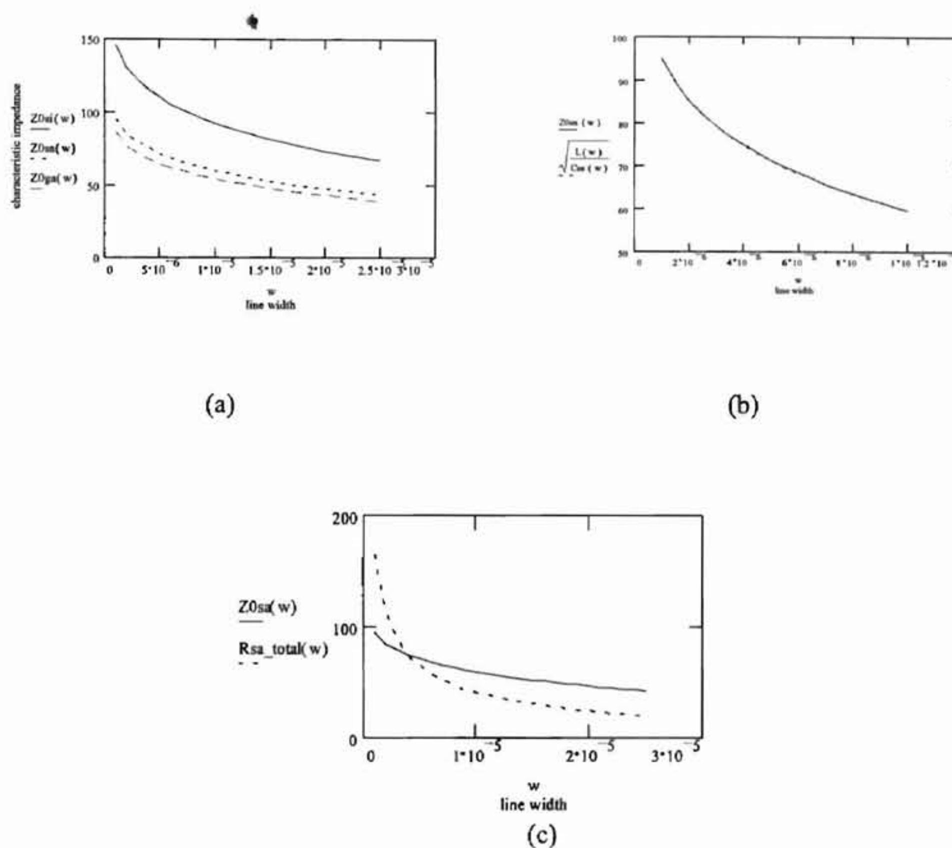


Fig. 2.6 Properties of Characteristic impedance

- From plot (c), when Z_0 has a large value, the line resistance is very large. That implies higher resistive attenuation. The high Z_0 (>100Ω) should be avoided in MCMs application due to high line resistance. But larger Z_0 helps to achieve smaller driver size (smaller power dissipation but lower speed) (see section 3.2.1). So there is

Oklahoma State University Library

a tradeoff in the selection of the proper Z_0 between signal loss and line length/driver design.

- Line resistance becomes smaller than Z_0 as the width increases (plot (c)). An additional series resistance should then be added to achieve the optimal delay and termination, $R_{total} \approx (1 \sim 3)Z_0$ (See section 3.2.6)

2.6 Points of Design

Table 2.1a Example cal. parameter data of a lossy trans. line. ($t = 5\mu m$, (with skin effect))

$W(\mu m)$	$C(Pf/cm)$	$L(nH/cm)$	$R(\Omega/cm)$	$Z_0(\Omega)$	$d_{min}(cm)$	$d_{max}(cm)$
28,83	2.233	3.568	3.29	40	1.216	11.4
16.85	1.749	4.36	5.08	50	0.98	10.1
3.915	1.128	6.336	13.91	75	0.54	6.1
0.707	0.828	8.273	36.77	100	0.272	3.3

Table 2.1b $t = 1\mu m$ (without skin effect)

$W(\mu m)$	$C(Pf/cm)$	$L(nH/cm)$	$R(\Omega/cm)$	$Z_0(\Omega)$	$d_{min}(cm)$	$d_{max}(cm)$
32.5	2.247	3.59	9.91	40	0.4	3.8
20.2	1.76	4.394	15.35	50	0.326	3.3
6.11	1.138	6.387	42.5	75	0.176	2.0
1.52	0.834	8.34	113.2	100	0.088	1.1

Table 2.1 gives out example calculation data of *C, L, R and* Z_0 with $h = 25\mu m$ (MCMs), $f = 1GHz$ ($\delta_s = 2.6\mu m$). Concluding the analysis above, the following design highlights are most important consideration for *R, L, C and* Z_0 extraction. The line geometry, signal frequency and material of conductor and substrate determine interconnect parameters.

- Use verified physical data for the technology in the intended design
    - Dielectric constant ε_r

Oklahoma State University Library

- Conductor sheet resistance
- Layer thickness (h) and conductor line width (w)
- Intended characteristic impedance (Z_0)
- Interested operating frequency

- Thickness of metal trace t has no influence on L, C.
- Line capacitance increases but inductance decreases as the line width increases or substrate height decreases for MCMs and SOS.
- The design of line width and substrate height is determined by either the self capacitance or inductance, which is dominant in system performance. Factors considered include signal delay, integrity and coupling noise
- If the maximum line length is desired, the thickness of metal trace should be set to the skin depth $t = \delta_s(f)$. Too thick a metal trace is not helpful in improving the performance of a transmission line. If $t > \delta_s(f)$, extraneous material is wasted; if $t < \delta_s(f)$, higher resistance will be resulted. Under the satisfaction of signal delay (risetime) and loss requirements, the thickness of line is preferred to be thinner with a maximum value $\delta_s(f)$.
- If thickness t is less than skin depth, the resistance of trace will increase. This characteristic helps to design short lossy transmission lines, for example, an on-chip transmission line.
- Too high a characteristic impedance should be avoided because of larger line resistance. The most popular used values of Z_0 are 50Ω (75Ω) or less ($30\Omega \sim 100\Omega$). The parallel loads may reduce the effective value of Z_0.

Oklahoma State University Library

CHAPTER 3

DESIGN OF TRANSMISSION LINES

As the signal frequency and circuit density increase, the performance of interconnect becomes increasing significant. In an electronic system, interconnects include: on-chip wires, wires on the MCM substrates, wires on PCBs, the package pins, lead frames, bonding wires and solder bumps, etc. Their electrical performance varies widely and depends on the lengths and cross-sections of interconnects and operating frequencies. The most popular used models are the *RC* and transmission line models.

As mentioned in part 2.1, the transmission line model is used for high frequencies and long interconnects. In this model, interconnect inductance has been considered. N. Sherwani and Q. Yu [1] offer a simple merit to determine whether the *RC* or transmission line model should be used for given geometry. In brief, when the signal risetime, t_r, is much less than signal time-of-flight, t_f, i.e., $t_r < 2.5t_f$, a transmission line model should be used. On the other hand, when t_r is much larger than t_f, i.e., $t_r > 5t_f$, lumped *RC* model is sufficient. When t_r is between $2.5t_f$ and $5t_f$, either model will suffice. t_f is the time-of-flight of the signal and is given by:

$$t_f = \frac{d}{v} \qquad v = \frac{c_0}{\sqrt{\varepsilon_r \mu_r}} = \frac{c_0}{\sqrt{\varepsilon_r}} \tag{3.1}$$

where c_0 is speed of light in vacuum ($3 \cdot 10^8 m/s$); ε_r and μ_r are the relative permittivity and permeability of propagation medium. μ_r is approximately 1 for nonmagnetic material.

Oklahoma State University Library

For MCMs applications, interconnects are typically modeled as transmission lines due to short risetime and long lengths. In figure 2.1, four typical transmission line cross-sections were presented. Compared to ordinary point-to-point wiring, a transmission line has less signal distortion, less radiation and less crosstalk due to the reduction of line inductance and the short distance to the ground plane [2].

3.1 Lossless, Lossy and Fully Lossy Transmission Line

An ideal transmission line consists of two perfect conductors. There are: balanced (twisted pair) and unbalanced (coax, microstrip and stripline) transmission lines. An unbalanced transmission line is also called a single-ended line. In this type of transmission line, signal current flows through one signal trace and returns back along the ground trace. The ground trace is usually wider than signal trace and can be shared among many signal traces [2].

Every transmission line goes through lossless (LC) $\rightarrow$ lossy (RLC) $\rightarrow$ fully lossy (distributed RC) transition as operating frequency or length increases. The key factor is the total line resistance (R_{line}). The ratio of $Z_0 = \sqrt{L/C}$ (Ω) to line resistance per unit R (Ω/cm) determines the critical useful length.

3.1.1 Definition of Lossless, Lossy and Fully Lossy Transmission Line

Recall the formula (2.5.1), if the conductance G is ignored, the transmission line characteristic impedance is a function of frequency.

$$Z_0 = \sqrt{\frac{R + j\omega L}{j\omega C}} = \left(\frac{L}{C}\right)^{\frac{1}{2}} + \left(1 + \frac{R}{j\omega L}\right) \qquad (3.1.1)$$

The following Table gives out a definition of the three type transmission lines

Table 3.1 Definition of lossless, lossy and fully lossy transmission line

types	Definition(I)	Definition(II)
Lossless	$\frac{R}{\omega L} << 1$ or $\omega >> \frac{R}{L}$	$R_{total} << Z_0$
Lossy	$\frac{R}{\omega L} \approx 1$ or $\omega \approx \frac{R}{L}$	$R_{total} \approx Z_0$
Fully lossy	$\frac{R}{\omega L} >> 1$ or $\omega << \frac{R}{L}$	$R_{total} >> Z_0$

The critical length of each transmission line type is given by:

$$d = K\sqrt{L/C}/R \qquad (3.1.2)$$

where L, C *and* R are the inductance, capacitance and resistance per unit of a transmission line. For lossless line, $K \leq 0.1$; for lossy line, $0.1 \leq K \leq 2$; and for fully lossy line, $K \geq 2$. The critical length increases from lossless to lossy and fully lossy. The skin effect in high frequency applications will reduce the critical length [15].

For example, if $R = 10\Omega/cm$ (recall table 2.1) and $Z_0 = 50\Omega$, when $d \leq 0.5cm$; $0.5cm \leq d \leq 10cm$ and $d \geq 10cm$, we have lossless, lossy and fully lossy transmission lines respectively.

3.1.2 Characteristics of Three Types of Transmission Line

Normally, the interconnects, such as wires on the PCB, package pins, lead frames, bonding wires and solder bumps, all have low resistance due to their large cross-sections.

As a result, these interconnects can frequently be treated as lossless transmission lines or inductors.

For MCMs wiring, because of their small dimension and long length, the resistance is usually significant and they are treated as lossy transmission line. This case will be more common for large die on-chip interconnects in the future.

A brief description of the properties of lossless, lossy and fully lossy transmission line is presented as follows [15]:

a. lossless transmission line

- R is neglected and Z_0 is a real constant in the frequency of interest. $Z_0 = \sqrt{L/C}$
- No energy loss occurs in the line.
- Propagation delay is linear with length and the signal propagation speed is given by formula (3.1).
- The signal is undistorted along the line but should be terminated.

b. lossy transmission line

- R, L and C can not be ignored and Z_0 is complex and frequency dependent. In most cases, $Z_0 \approx \sqrt{L/C}$.
- Significant energy loss exists in line and series termination or non-termination is useful because of significant line resistance (DC losses) (more discussion in section 4.1)
- Signal delay ranges from linear to quadratic
- Signal is somewhat distorted

c. fully loss transmission line

Oklahoma State University Library

- L is neglected. Z_0 is complex, frequency dependent and is not an useful concept at any frequency.
- High line energy dissipation and no signal reflections.
- Signal propagation is proportional to RC, and quadratic with length. $\tau \sim RC \sim length^2$.

If there are multiple loads along the line, attempts should be made to use a lossless transmission line since the signal has no distortion and only delay at different points along the line. Use of a lossy transmission line can be difficult for multiple loads due to the signal distortion. If there is only one load on the end, both lossless and lossy transmission line can be applied. When a lossy transmission line must be applied to connect multiple loads, parallel fanout is preferred.

The selection of transmission line type for an application depends on loading, operating frequency and line length. For long interconnects on large single die or MCMs, they are usually lossy transmission lines. The resistance R_{line} is limited by

$$0.1\left(\frac{L}{C}\right)^{1/2} \le R_{line} \le 2\left(\frac{L}{C}\right)^{1/2} \qquad (3.1.3)$$

3.2 Optimal Delay Design of a Transmission Line

All electrical devices from basic gates to large chips have their specific delay. The cumulative effects of inductance and capacitance result this delay. The gate delay reduces faster due to the scaled down device geometry and the interconnect delay becomes one of the most important limiting factors in today's high-speed and high-density circuit performances. As a result, it is vital to minimize interconnect delay to

Oklahoma State University Library

achieve the highest performance. The interconnect delay is determined by many factors, such as line length, conductor geometry and material, etc. In this section, the design of optimal delay will be discussed.

3.2.1 First Incidence Voltage

Assuming a unit step signal is applied in the input end of a transmission line, the response of this line at a distance x from the input end is consisted of two parts: the fast rising portion, V_A, and the slow rising portion, V_R. $V_A = e^{-\alpha x}$ is an attenuated function exponentially dependent on attenuation constant α and the distance x; V_R represents a RC-like behavior [1][16].

To achieve minimum propagation delay on a transmission line, the fast rising portion, V_A, must have sufficient amplitude. In another words, the first signal (first incidence voltage) to arrive at the end of the line must have sufficient voltage to switch the receiver, e.g. its value should exceed V_{IH} on $1 \rightarrow 0$ transition or V_{IL} on a $0 \rightarrow 1$ transition. V_{IH} is minimum voltage that is required to be input of receiving device for a logic 1; V_{IL} is the maximum voltage that is applied to the input of a receiving device for logic 0. This situation is referred to as first incidence switching [16].

The first incidence voltage at the receiving end is given by

$$\frac{V_{first}}{V_{in}} = \frac{Z_0}{R_{out} + Z_0} \cdot e^{\left(-Rd/2Z_0\right)} \qquad (3.2.1)$$

$$d_{\max} = \frac{-2Z_0}{R} \ln\left(\frac{0.5 \cdot (Z_0 + R_{out})}{Z_0}\right) \qquad (3.2.2)$$

Oklahoma State University Library

where V_{in} is the open circuit output of driver, R is the per unit resistance of transmission line, d is transmission line length and d_{max} is the maximum length for fixed driver size, characteristic impedance Z_0. When $Z_0 >> R_{out}$ and V_{in} equals $V_{dd}/2$:

$$d_{\max} \approx 1.38 Z_0 / R \quad (3.2.3)$$

To get enough first-incidence-switch value V_{first}, the resistance R should be small and Z_0 should be significantly greater than driver output impedance R_{out}. Because Z_0 of a transmission line will always be chosen between (30Ω~100Ω), the output impedance of a CMOS driver is designed around (5Ω~20Ω). If Z_0 is selected too small, a large width driver is required to achieve small output impedance. Too a large driver width results in higher power dissipation and device performance limitations due to the increased gate (R_g) and source resistance (R_s) (see Chapter 5). If Z_0 is too large, line resistance will be large (see section 2.5.2). This may in turn create problems in meeting the signal risetime and loss requirements resulting longer line unavailable The selection of Z_0 should be a tradeoff between signal speed/loss and driver size/power consumption.

3.2.2 Lumped *RLC* Model for A Lossy Transmission Line

As mentioned in the beginning of this chapter, the *RC* and transmission line models are the two most popular models being used to describe the properties of interconnects. A lossy transmission line can be modeled and simulated as a series lumped *RLC* segments using lossless transmission line sections separated by lumped resistance and conductance. Figure 3.1 presents the structure of this basic segment [17].

Oklahoma State University Library

The characteristic impedance ($Z_0 = \sqrt{L/C}$) and lossless transmission line delay ($\tau = d_s \cdot \sqrt{LC}$) for each segment are the same and the segment delay is the total delay divided by the number of segments. d_s is the line length of each segment.

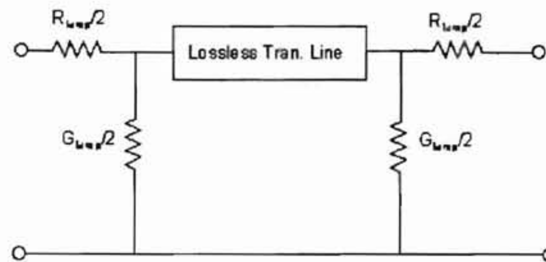


Fig.3.1 The basic lumped *RLC* segment for a lossy transmission line

The only consideration in basic segment design is the number of segments should offer a fine enough granularity such that for each lumped segment, the propagation delay is much smaller than signal risetime. In another words, the *RLC* lumped segment's propagation delay $d_s(LC)^{1/2}$ should be a small fraction β of the risetime. The quantity β can be thought as the electrical "length" of each lumped segment ($\beta = \frac{1}{3} \sim \frac{1}{10}$).

$$d_s(LC)^{1/2}/t_r = \beta, \quad (\beta = \tfrac{1}{3} \sim \tfrac{1}{10}) \tag{3.2.4}$$

Assuming $\beta=1/10$, the number of lumped sections, N, is given by:

$$N \geq 10\frac{d\sqrt{LC}}{t_r} \tag{3.2.5}$$

where d is total transmission line length; L and C are line inductance and capacitance per unit respectively. The lumped capacitance, inductance and resistance are given by:

$$R_{lump} = \frac{R \cdot d}{N} \tag{3.2.6}$$

Oklahoma State University Library

$$L_{lump} = \frac{L \cdot d}{N} \tag{3.2.7}$$

$$C_{lump} = \frac{C \cdot d}{N} \tag{3.2.8}$$

For example, assuming a lossy transmission line with $Z_0=50\Omega$, $d=3cm$, $C=1.76pF/cm$, $L=4.394nH/cm$, $R=15.35\Omega/cm$ (Table 2.1) and signal risetime $t_r=350ps$. with number of segments, N, equal 8 then $C_{lump}=0.66pF$, $L_{lump}=1.648nH$, and $R_{lump}=5.76\Omega$.

This series lumped segment model of a lossy transmission line can be used in PSPICE for transmission line simulation. The input should be "filtered" (i.e. using an inverter) to avoid introducing unphysically high frequencies, which may excite resonance of the individual lumped segment.

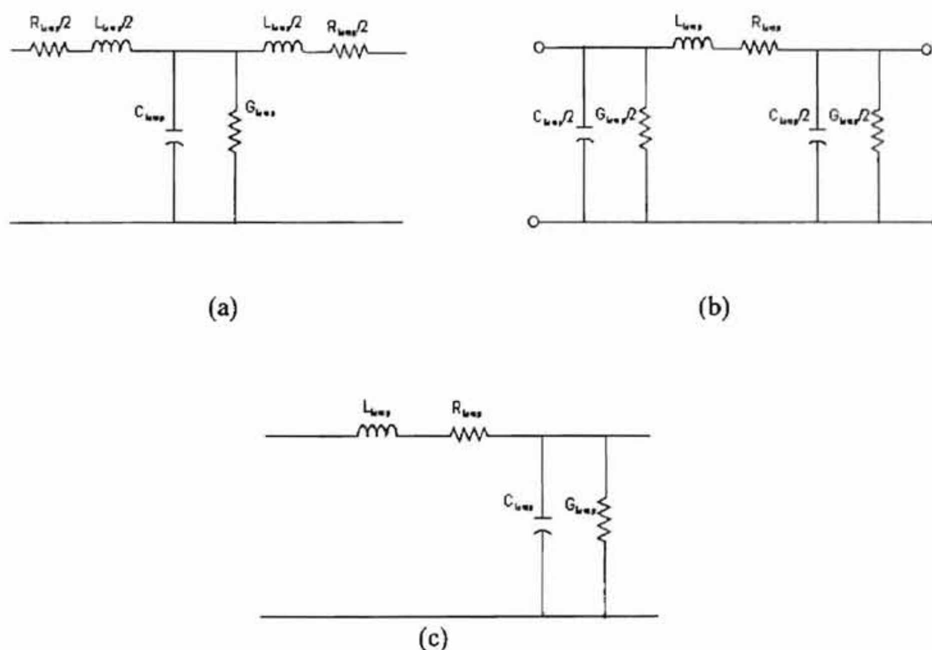


Fig. 3.2 (a) "Tee" type symmetric lumps; (b) "Pi" type symmetric lumps; (c) "Gamma" type lumps;

Oklahoma State University Library

Figure 3.2 gives out other three types but equivalent segments used for simulation where G is always neglected. In this thesis, a "Gamma" type lumped segment is used and recommended to reduce network size.

3.2.3 The Delay Estimation for A Transmission Line

The extraction formulas for capacitance, inductance and resistance of a lossy transmission line were illustrated in Chapter 2. It is known the line capacitance, inductance and resistance are functions of the line width w. Roughly the following equations are held:

$$R(w) \propto f_1(w) \qquad (3.2.9)$$

$$C(w) \propto f_2(w) \qquad (3.2.10)$$

$$L(w) \propto f_3(w) \qquad (3.2.11)$$

Figure 3.3 presents an interconnect driven by cascaded buffer chain (a) lumped model; (b) distributed model (series RLC model).

It is well known that the cascaded buffer chain with increasing width ratio, α, offers the minimum delay and it is easy to show $\alpha = e(2.718)$ for ideal case. Practically α always varies from 2.5~5 depending on the process.

For a lumped model, the delay of an interconnect with a capacitance load at the end is related to line width. Using the Elmore delay equation, it is given by:

$$t_{d1} = R_{out}C_L + R_{out}df_2(w) + C_L df_1(w) + \frac{d^2}{2} f_1(w) f_2(w) \qquad (3.2.12)$$

Where R_{out} is the output impedance of the last stage of buffer chain. Combining driver delay with (3.2.12), the total delay is

$$t_d = N\alpha t_{db} + R_s(C_L + Cd) + RdC_L + \frac{1}{2}RCd^2 \tag{3.2.13}$$

where t_{db} is the first stage delay in the buffer chain; R_s is sum of driver output impedance, R_{out}, and termination resistance, R_t ($R_s = R_{out} + R_t$).

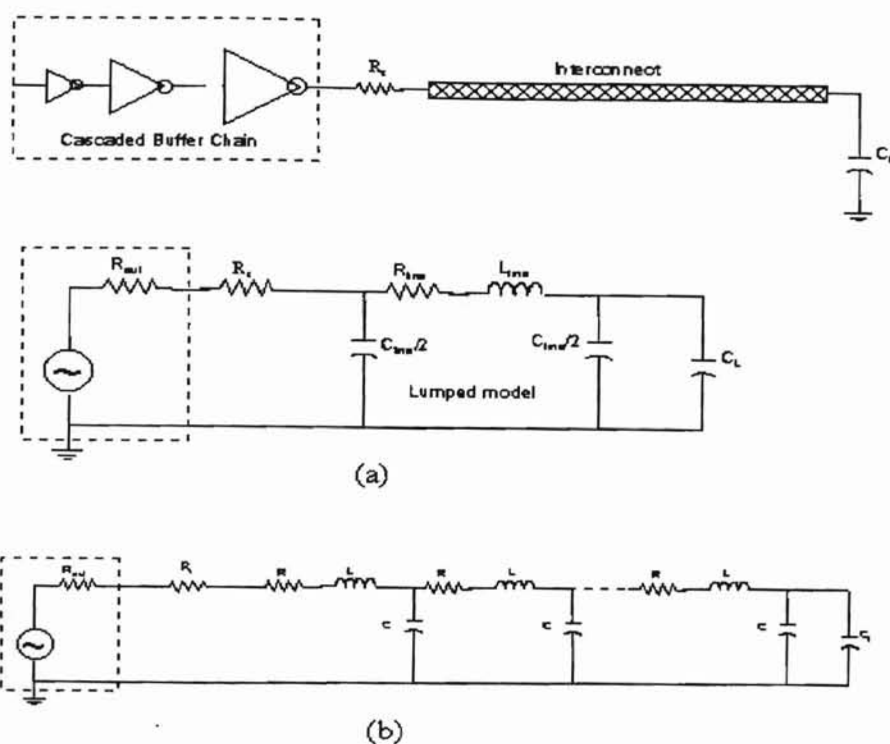


Fig. 3.3 Lumped and distributed model for interconnection

For lossless and lossy transmission lines, the signal propagation delay roughly equals to time-of-flight and is determined by the line inductance and capacitance. Except for the signal amplitude loss, a lossy transmission line has the same properties as a lossless transmission line. The delay of only the transmission line and the total delay including driving buffer and pad inductance are given by (3.2.14) and (3.2.15), where

L_{line} and C_{line} is total inductance and capacitance of the transmission line respectively. The present of series inductance (bond wire) increases delay slightly by approximatly L_b/Z_0 (see section 3.3).

$$t_{d1} = \sqrt{L_{line}C_{line}} \approx \frac{d\sqrt{\varepsilon_r}}{c_0} \tag{3.2.14}$$

$$t_d = N\alpha t_{db} + \sqrt{L_{line}C_{line}} + 2L_b/Z_0 \tag{3.2.15}$$

3.2.4 Delay Factor Sensitivity Analysis

The transmission line delay is a function of driver output impedance R_{out}, series termination resistance R_t, line resistance R_{line}, capacitance C_{line} and inductance L_{line}, supply voltage V_{dd}, load capacitance C_L, and driver end capacitance, C_D. It can be presented as:

$$t_d = f(R_{out}, R_{line}, C_{line}, L_{line}, R_t, C_L, C_D, V_{dd}) \tag{3.2.16}$$

Ayman & Karem [18] gave out the analysis for the sensitivity of delay versus these parameters. They pointed out the delay of lossless lines is a linear function of C_L/C_{line}; and for a lossy transmission line, the driver output impedance and line resistance can not be neglected. Ayman & Karem's analysis data are presented in table 3.2. The sensitivity of delay to parameters is briefly summarized as follows.

Table 3.2 Sensitivity of line delay to various parameters

Parameters	Sensitivity
Line length d	0.936
Line resistance per unit R	0.0493
Line capacitance per unit C	0.474
Load capacitance C_L	0.120
Driver output impedance R_{out}	-0.0254

- The sensitivity to line length is close to one, which means that line length increases by 5%, the delay increases almost 5%. Reference [15] confirms the same result.
- The sensitivity to line capacitance is about *0.5*, which can be derived directly from the lossy transmission line delay equation $t_d = d\sqrt{LC}$.
- The sensitivity to line resistance R_{line} is about 5%. Compared to the sensitivity of the line capacitance, the sensitivity to line resistance is much smaller.
- The sensitivity of driver output impedance is negative because delay decreases with stronger driver.
- The sensitivity to dielectric constant ε_r is dependent on different types of transmission lines. For a diffusion transmission line, $t_d \sim R_{total}C_{total}$, the sensitivity to ε_r is 100%; For lossless transmission line, $t_d \sim (L(C_{line} + C_{load}))^{1/2}$, if $C_{line} = C_{load}$, the sensitivity is 25%; For lossy and lossless transmission lines, the sensitivity is almost equal.

The sensitivity implies the priority of factors in optimizing delay. First, the possible shortest line length is preferred by optimal placement and wiring. Second, it is more advantageous to decrease line capacitance before trying to decrease resistance in order to decrease the delay. A narrower line is preferred if the risetime and signal loss specifications are satisfied.

3.2.5 Oscillation of Series RLC Circuits

When designing a transmission line, the resonation of the interconnect line must be considered. A series *RLC* circuit will resonate when loop resistance equals zero. The concern is excessive overshoot and undershoots due to under damping. Overshoot is not

a critical detriment to digital applications (this is not the case for ground bounce.), but undershoot can cause logic faults and increase the effective delay. The equivalent *RLC* circuit and its response are shown in figure 3.4. The extra delay induced by the resonation is illustrated in figure (b).

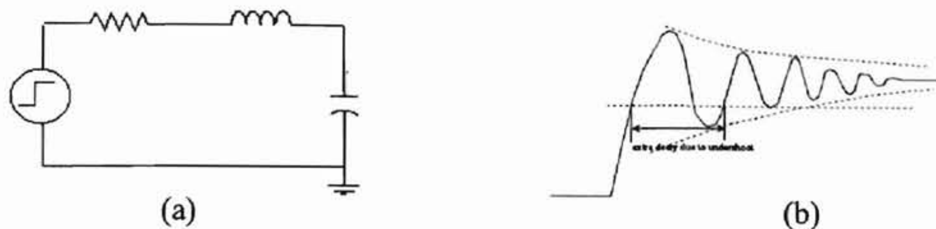


Fig 3.4 Typical series RLC circuits and typical response

The resonation frequency of a *RLC* circuit and Q factor are given by:

$$\omega = j\frac{R}{2L} + \frac{1}{2}\sqrt{\frac{R^2}{L^2} - \frac{4}{LC}} \qquad (3.2.17)$$

$$Q = (L/C)^{1/2}/R \qquad (3.2.18)$$

if $R = 2\sqrt{L/C}$, the real part of the frequency equals 0 and $Q = 0.5$. This is critical damping for a series *RLC* circuit. For a lossless transmission line. $R_{critical} \approx \sqrt{L/C} \approx Z_0$; for lossy transmission line, $R_{critical} \approx (1 \sim \pi)\sqrt{L/C} \approx (1 \sim \pi)Z_0$ [15].

The series resonation of a *RLC* equivalent circuit is only a problem with a low resistance line (lossless or low-lossy transmission line) and very wide CMOS driver (very low output impedance). In this case, the Q factor can be very high. The following solutions can be used to increase circuit damping toward the "critical damping" condition.

a. Reducing line width or decreasing conductor thickness (if this is possible without hurting signal risetime) to increase line resistance.

b. Using a larger termination resistance.

c. Reducing driver width to increase its output impedance, however it should be noted, this will increase the buffer delay causing an increase in the total delay.

d. If there is no room to increase the resistance, the inductance in system including the bond wire inductance and/or line inductance should be reduced by using a thinner substrate, wider conductor, parallel fanout topology, etc.

3.2.6 Optimizing the Delay

Optimizing the delay of a transmission line is to optimize the dimension of the line and the size of driver, especially the transmission line width. If the series resistance is too small, the circuit will ring; if it is too large, the delay will increase (for the *RC* model). In Chapter 4, it will be shown out that the series resistance also plays important role in proper termination of a transmission line. For this reason we will have a tradeoff between the resonance and delay. Usually the optimal delay can be achieved when $R_{total} = (1 \sim 3)Z_0$ [15], that implies

$$R_{total} = (1 \sim 3)\sqrt{L_{total} / C_{tatol}} \quad (3.2.19)$$

where $R_{total} = R_{out} + R_{line} + R_t$;

$L_{total} = L_b + L_{line}$; L_b is the inductance of bond wire;

$C_{total} = C_{line} + C_{load} + C_P$; C_P is capacitance of pad;

Now considering the two extreme cases:

a. For a very wide line. $R_{line} \to 0$; $L_{line} \to 0$; $C_{line} \to \infty$ and $\tau \sim R_{out} C_{line} \sim w$;

b. For a very narrow line. $R_{line} \to \infty$; $L_{line} \to$ constant (very high); $C \to$ constant (very low and $\tau \sim R_{line} C_L \sim 1/w$;

For wide line, delay increases as the C_{line} increases, e.g. w increases; For narrow line, delay increases as R_{line} increases e.g. w decreases. This implies there should be optimal line width to achieve the minimum delay without signal resonance. Figure 3.4 shows the curve of delay versus line width [15].

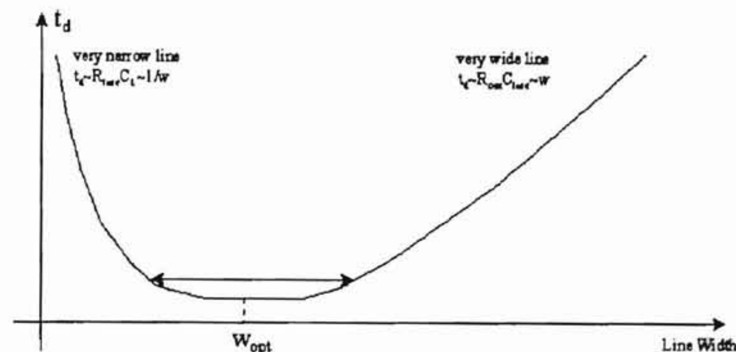


Fig. 3.5 Delay vs. line width

Equation 3.2.19 gives out the range of total resistance for a transmission line, which leads to optimal delay without signal resonation. R_{total} makes the circuit more or less critically dumped. Going to a lower resistance than required doesn't improve the system performance and in fact can introduce resonance and extra delay. From figure 3.5, it is shown that there is a very flat region around optimum line width. The choice of line width also has an effect on driver size design and termination resistance selection due to line resistance [19]. In practice, in the initial design stage, one or two line widths should be selected and then adjusted by other design considerations (signal termination, crosstalk, etc.).

3.3 Peripheral Parts' Effect on Interconnection Performance

Peripheral parts, such as bond wires and pads, also have a significant effect on interconnect performance. Bond wire induces extra inductance and pad capacitance maybe larger than interconnect line capacitance. When evaluating the overview performance of the system, those effects must be considered.

A bond wire is always modeled as an inductance. It is estimated by [15]:

$$L_b = 2(nH/cm) \cdot d_b \cdot \left(\ln(\frac{2d_b}{r}) - \frac{3}{4} \right) \quad (3.3.1)$$

where r is the radius and d_b is the length of round wire. This inductance induces extra delay ($t_{bd} = L_b / Z_0$) for signal propagation. For example, $d_b = 1.25mm$, $r = 50\mu m$, $Z_0 = 50\Omega$, $L_d = 0.79nH$ and $t_{bd} = 15.8ps$. Figure 3.6 presents a graphical analysis of this induced delay (assuming L/Z_0 << risetime).

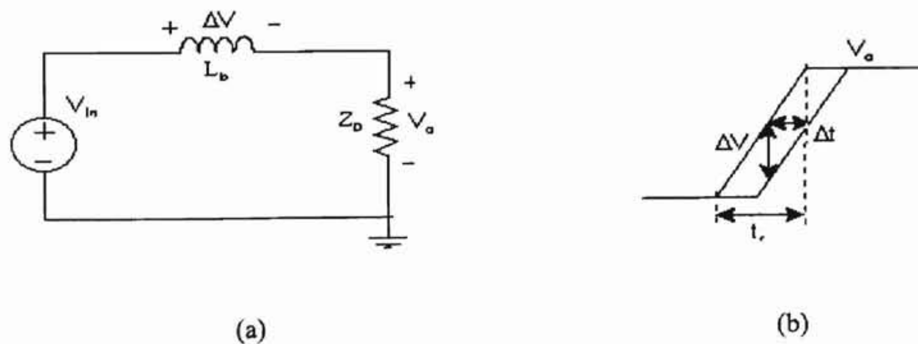


Figure 3.6 Graphical illustration of bond wire delay

From fig 3.6 (b),

$$\Delta v = L_b \frac{\partial I}{\partial t} = L_b \frac{\partial (V_{in}/Z_0)}{\partial t} = \frac{L_b}{Z_0}\frac{V_0}{t_r} = \frac{L_b}{Z_0}\frac{\Delta v}{\Delta t}$$

$$\Delta t = \frac{L_b}{Z_0} \qquad (3.3.2)$$

Equation (3.3.2) is used for lumped model analysis. The bond wire inductance also has a significant effect on the power distribution. This inductance would induce a greater possibility of high "simultaneous noise" and power distribution line oscillation. These issues will be further discussed in Chapter 4.

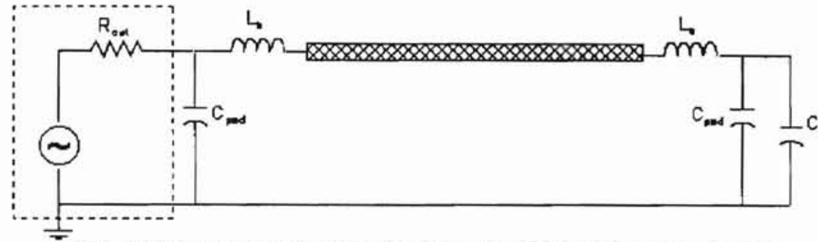
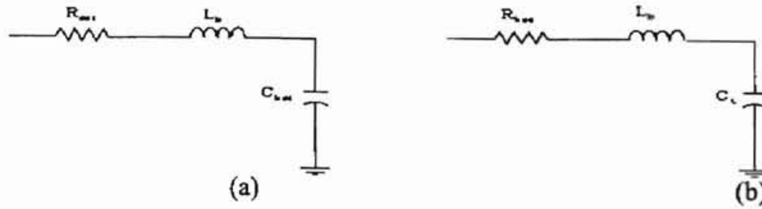


Fig. 3.7 Interconnect equivalent circuit with bond wire and pads

At both ends of transmission line, there are output and input pads. These pads add extra capacitance to interconnect. They have effects on signal propagation delay as well as characteristic impedance. Figure 3.7 presents the equivalent circuit. The effective characteristic impedance is given by equation (2.5.4) (see section 2.5.1) and effective delay is:

$$t_{delay} = \left(L_{total} C_{total}\right)^{1/2} \qquad (3.3.3)$$

$$Q \leq \sqrt{L_{total} / C_{total}} \Big/ R_{total} \qquad (3.3.4)$$



Fig, 3.8 Equivalent *RLC* circuit of input and output ends

When selecting bond wire, the resonation of series *RLC* should also be considered. The equivalent circuits at ends of output and receiver are shown as in figure 3.8. For fixed R_{out}, C_{line}, R_{line} *and* C_L, correct package technology (L_b) is required to achieve a low quality factor, Q, avoiding resonance (equation 3.3.4). In order to improve the system performance, the pad capacitance and bond wire should be made as small as possible. In simple terms, pad out must match the die process technology.

3.4 Loss in a Transmission Line

In this section, signal integrity along a transmission line is discussed. The signal attenuation constant of a microstrip model depends on line dimension, electronic properties of substrate and conductors and the operating frequency. Normally there are two types of loss in transmission line model, one is dielectric loss due to substrate and the other is ohmic skin loss in strip conductors and ground plane [1] [11]. The signal loss of a transmission line with length d can be expressed as

$$\frac{V(x=d,t=t_d)}{V(x=0,t=0)} = e^{-\alpha d} \tag{3.4.1}$$

where α is attenuation constant and given by:

$$\alpha = \max(\alpha_R, \alpha_S) + \alpha_D \tag{3.4.2}$$

where α_R is resistance attenuation constant of line without skin effect;

α_S is resistance attenuation constant of line with skin effect;

α_D is dielectric attenuation constant.

The signal conductor loss is caused by line resistance. Thickness of the line, t, is used in the resistance calculation at low frequency and skin depth δ_S for high frequency. α_R and α_S are given by:

$$\alpha_R = \frac{R}{2Z_0} = \frac{\rho}{2wtZ_0} \tag{3.4.3}$$

$$\alpha_S = \frac{\sqrt{\pi\mu f\rho}}{wZ_0} \tag{3.4.4}$$

The dielectric attenuation constant is defined as:

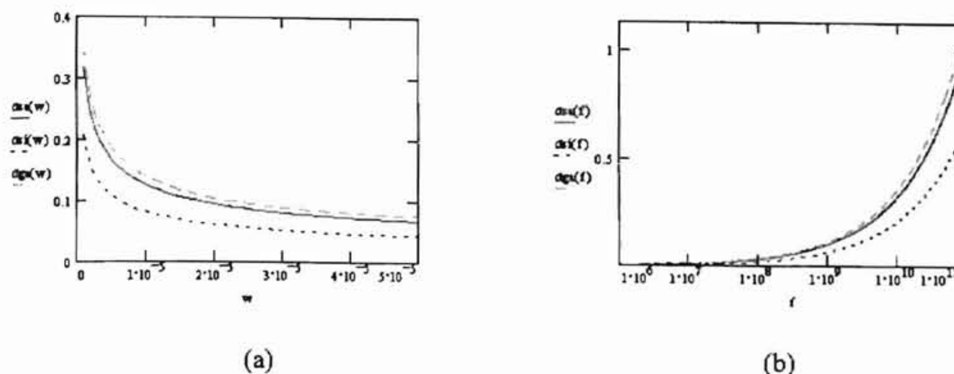
$$\alpha_D = \frac{\pi f\sqrt{\varepsilon_{eff}}}{c_0}\tan\theta \tag{3.4.5}$$

where c_0 is light speed in vacuum and $tan\theta$ is dielectric loss tangent and is restricted to lower than *0.001*. According to reference [1], $\tan\theta = \sigma/(\omega \cdot \varepsilon_r)$. Here, σ is the conductivity of the dielectric substrate and ω is the angular frequency. Simplifying equation (3.4.5),

$$\alpha_D = \frac{\sigma}{2c_0\sqrt{\varepsilon_r}} \tag{3.4.6}$$

From equation (3.4.6), it is observed that if the substrate conductivity is frequency independent, the dielectric attenuation factor is also frequency independent. And if the

dielectric constant and conductivity are chosen correctly, the dielectric loss are negligible at most application frequencies.



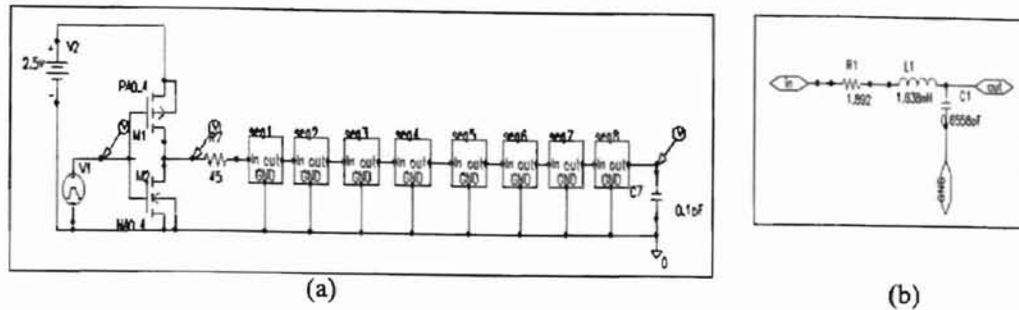
(a) (b)

Fig. 3.9 (a) Resistance Attenuation Constant Versus Line Width; (b) Resistance Attenuation Constant Versus Operating Frequency

Figure 3.9 presents the change of resistance attenuation constant versus line width w and operating frequency f. From the plot shown above, the conductor attenuation factor reduces as the line width increases due to the decrease of line resistance. On the other hand, the conductor attenuation factor goes up fast as the frequency increases due to the skin effect increasing the line resistance. Therefore as signal frequency goes high, the loss can not be neglected. Signal loss is another important limitation for transmission line length in high-speed application.

3.5 PSPICE Simulation of Transmission Line Signal Delay

The Figure 3.10 gives the delay simulation schematics and segments. Here simulation is valid for line length (d) and the number of segment (N) satisfying $d/t_r \leq N/(\beta\sqrt{LC})$ (see section 3.2.2).



(a) (b)

Figure 3.10 Delay simulation schematics and segment

Table 3.3 contains calculation and simulation data of signal delay for a transmission line with a characteristic impedance of *50Ω* and *75Ω* and Figure 3.11 is simulation result for different lengths, series termination resistances and different characteristic impedances. Figure 3.12 gives out the comparison of calculation and simulation data for signal delay.

a. *Z_0=50 ohm, C=1.749Pf/cm, L=4.368nH/cm, R=5.05Ω/cm, W=16.85um, t=3.0um*, (table 2.1(a)), N is number of segments

Table 3.3a, Calculation and simulation data of signal delay for transmission line

Length (cm)	N	C_{lump} (pF/cm)	L_{lump} (nH/cm)	R_{lump} (Ω)	T_{delay} (Cal)(ps)	T_{delay} (Sim) (ps)
0.5	2	0.4372	1.092	1.26	43.7	57.52
1.0	3	0.583	1.456	1.682	87.4	109.78
1.5	4	0.6558	1.638	1.892	131.1	157.2
2.0	5	0.6996	1.747	2.018	174.8	198.54
2.5	7	0.6246	1.56	1.802	218.5	252.46
3.0	8	0.6558	1.638	1.892	262.2	298.5

b. *Z_0=75 ohm, C=1.128Pf/cm, L=6.338nH/cm, R=13.82Ω/cm, W=3.91um, t=3.0um*

Table 3.3b

Length (cm)	N	C_{lump} (pF/cm)	L_{lump} (nH/cm)	R_{lump} (Ω)	T_{delay} (Cal)(ps)	T_{delay} (Sim) (ps)
0.5	2	0.2819	1.584	3.455	42.27	57.8
1.0	3	0.3759	2.113	4.607	84.54	108.6
1.5	4	0.4229	2.377	5.183	126.8	156.8
2.0	5	0.4511	2,535	5.528	169.1	208.5
2.5	7	0.4028	2.263	4.936	211.4	247.55
3.0	8	0.4229	2.377	5.183	253.6	295.7

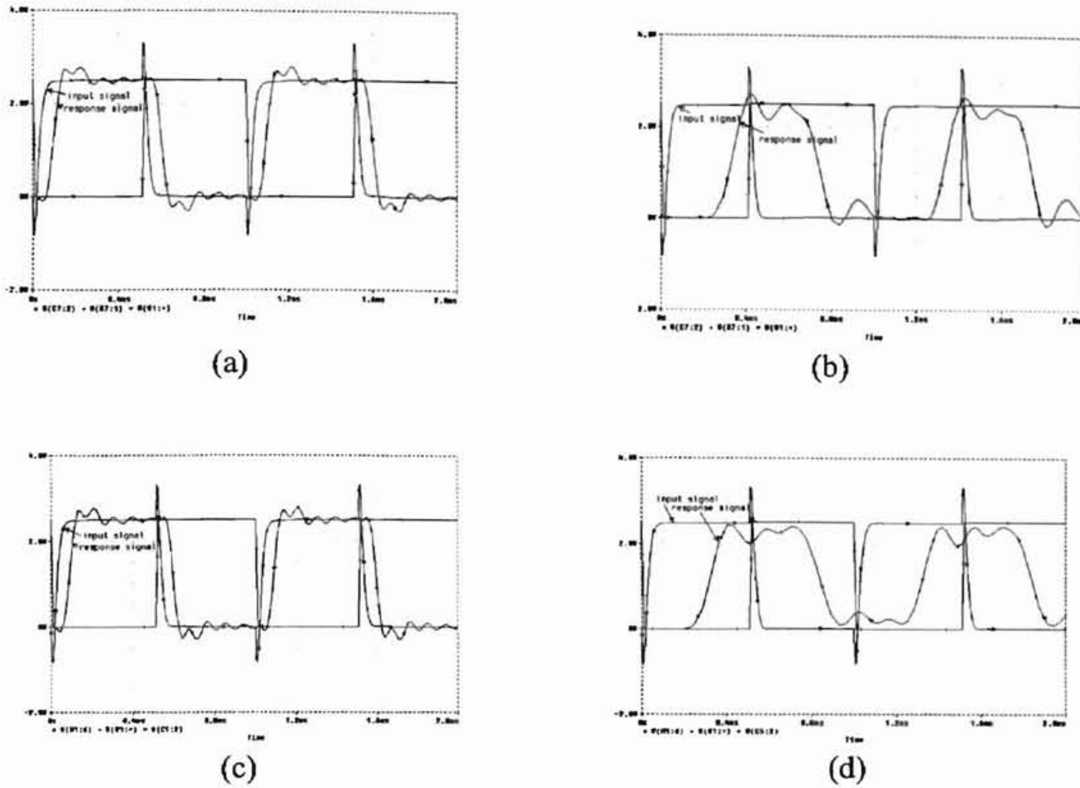


Fig. 3.11 Delay Simulation results. a. Z_0=50Ω, cm R_t =45Ω, R_{on}=3.53Ω, length =0.5; b. Z_0=50Ω, R_t=45Ω, R_{on}=3.53Ω, length = 3cm; c. Z_0= 75Ω, Rt = 70Ω, R_{on}=3.53Ω, length =0.5cm; d. Z_0= 75Ω, Rt = 70Ω, R_{on}=3.53Ω, length =3cm; where Rt is the series termination resistance of transmission line.

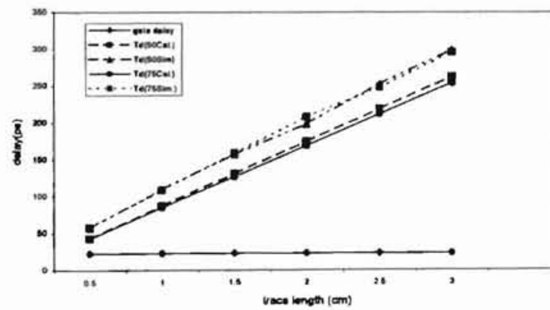


Fig 3.12 Comparison of calculation and simulation data for signal delay

Comparing the calculation and simulation data, the following conclusion are made:

1. The delay of a lossy transmission line is independent of characteristic impedance. It depends on line inductance and capacitance.

$$T_d = Length \cdot \sqrt{L_{line} C_{line}} \qquad (3.5.1)$$

   In practice, the total inductance and capacitance should include bond wire inductance and pad capacitance.

2. The delay is linear with line length.
3. The lumped *RLC* distributed model of the lossy transmission line is accurate enough to estimate the delay. If more segments are used (select bigger N), estimation will be more accurate.
4. Comparing gate delay with interconnection delay, as line length becomes larger, the interconnection delay becomes dominant (much larger than gate delay, >10 times).

3.6 **Points of Design**

- Check out the possible length range of interconnects and determine if transmission lines should be used by using risetime specification and possible line lengths (time-of-flight) ($t_r < 2.5t_f$). (see section 3.1)
- Select a characteristic impedance which is much larger than output impedance of line driver to increase signal incidence voltage but not too large avoiding high signal resistive loss (Z_0~(30Ω~100Ω)) (see section 3.2.1)
- Compare the estimated line resistance and chosen characteristic impedance. Determine the type of transmission line, lossless, lossy or fully lossy.

- For a lumped *RC* interconnect, the lumped delay model (equation 3.2.13) is used to estimate signal delay. For a transmission line, the delay is proportional to line length (equation 3.5.1) and a lumped distributed *RLC* model can be used for simulation.
- Optimal delay is obtained when setting $R_{total} = (1 \sim 3)Z_0$ (3.2.19). When selecting the total resistance, termination methods, driver size should be considered (see chapter 4).
- From the chosen characteristic impedance and delay specification, total line inductance and capacitance can be estimated. Based on equations (3.5.1) and (2.5.2)

$$C = t_d / Z_0 \qquad L = Z_0 t_d \qquad (3.6.1)$$

- When estimating the interconnection delay, the effect of bond wire and input/output pads must be considered. Make bond wire inductance and pad capacitance as small as possible. Check out if there is a resonance induced by including the bond wire inductance at both input and output ends.
- Make sure $C_{pad} << C_{line}$ to achieve parallel termination for lossless transmission line and open circuit for lossy transmission line (see section 4.1).
- Estimate line resistance and check the quality factor, *Q*, to make sure the interconnect being critically damped. If not, make the total resistance larger by increasing either output impedance of driver, line resistance or termination resistance (series termination).
- Check the signal risetime using equation (4.2.1). If the risetime doesn't satisfy the specification, modify line resistance and redesign the line geometry.

CHAPTER 4

DESIGN CONSIDERATIONS TO ELIMINATE SELF GENERATED NOISE

Interconnects not only cause signal delay but also generate noise. The interconnection noise becomes larger as a result of faster signal risetimes, larger switching current, longer interconnects and smaller spacing between interconnects. Certain types of noise, such as crosstalk, that previously were considered only at the board level become significant in high density and high frequency on chip circuits as well as in Multi-Chip Module systems.

Signal reflection, crosstalk and simultaneous switching noise are three important design issues regarding interconnects. In practice, a real transmission line is of finite length and the end of the transmission line introduces a discontinuity that may generate reflection noise. As the circuit density increases, smaller spacing between the interconnects results in higher crosstalk noise because the capacitive, inductive and resistive mutual coupling between become more significant. In addition, with large amounts of current switching simultaneously, the inductive noise associated with power line also increases, which is known as simultaneous switching noise.

In this chapter, design considerations for reflection noise, crosstalk and simultaneous switching noise are discussed. The content includes the factors that contribute to the generation of those noises and the effects of noise on the system

performance. The methods of controlling these noises and interconnect driver width design are also covered in the chapter.

4.1 Reflection Noise

Reflection noise is the noise caused by the discontinuities in a transmission line. Whenever a change in characteristic impedance occurs, part of the incident electromagnetic wave is reflected. The reflection coefficient, Γ, is given out by:

$$\Gamma = \frac{Z_{load} - Z_0}{Z_{load} + Z_0} \tag{4.1.1}$$

where Z_0 is the characteristic impedance of the transmission line on which the incident wave travels, and Z_{load} is the load impedance of the line. If the load is matched to the line impedance, e.g., $Z_{load} = Z_0$, no reflection occurs.

4.1.1 Series Termination

There are a few methods to mitigate the reflection noise. Parallel and series termination are two most usually used methods. The following figure presents the simplified circuit for both approaches.

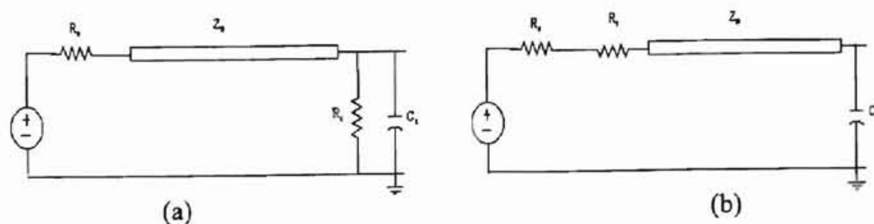


Fig. 4.1 (a). Parallel termination; (b). Series termination

When comparing these two methods, they have advantages and disadvantages respectively. Parallel termination can have several loads (C_i) without creating excessive reflections and has one half the rise-time of a comparable series-termination line with same load [2]. However the output with parallel termination depends on signal loss and reflection coefficient and its series dc losses costs noise margin. It also results in more DC power consumption and greater difficulty in selecting the correct terminating resistor because the parallel loading reduces the characteristic impedance (see section 2.5.2) [15].

For example, assuming $Z_0 = 50\Omega$, $R_{out} = 3.53\Omega$, $R_{line} = 7.575\Omega$ and $R_t = 49\Omega$,

$$V_0(DC) = V_S R_t / (R_{on} + R_{line} + R_t) = 0.815 V_S$$

There is around 20% loss in DC voltage. This consumes a significant amount of the noise margin. In order to compensate for the DC loss, R_t can be slightly larger than Z_0 or R_{out} and R_{line} should be very small compared to R_t or both. Making R_t a little larger than ideal termination resistance can obtain small positive reflection coefficient resulting in improvement of signal first incidence voltage but may limit the line to one load. Reducing R_{out} requires a very large transistor that can eventually contribute to other problems such as source degeneration and gate resistance limitations for reduced device operating bandwidth (see chap. 5) Normally, the series resistance is limited to *10%* of R_t for successful parallel termination due to DC loss problem. For a parallel termination transmission line, a lower characteristic impedance with lower line resistance is preferred. Only at the end of a parallel termination line can R_t be slightly larger than Z_0 to compensate for signal loss.

Series termination can overcome the DC loss problem of the parallel termination. It has no DC *IR* drop, V_o always get to the correct value "eventually". This also results in no DC power consumption. This characteristic is useful for CMOS and low power MCMs. The termination is less sensitive to exact used values of the resistance, the selection range is limited by the signal ring (resistance is too small) and excessive delay (resistance is too large). However, it is only suitable for use with single load. It is often the case on MCMs with CMOS drivers that design of the driver output impedance in combination with line resistance and termination resistance produces an effective series termination value. The disadvantage of a series termination includes: the slow risetime and "second-incidence switching" when multiple loads are on the line. In another word, far end load switches at $t = t_d$ as parallel termination but the near end load switches at $t = 2t_d$ rather than at $t = 0$ for parallel termination. This problem can be solved using the parallel fanout with multiple series termination. The number of multiple loads must be limited by (2~5) [15][1]. Figure 4.2 shows "second-incidence switching" problem and its solution.

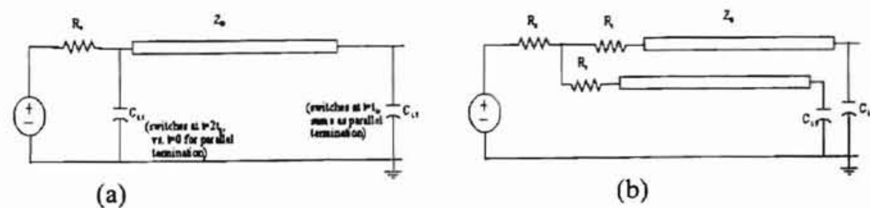


Fig. 4.2 (a). Second-incidence switching problem of series termination with multiple loads; (b). Parallel fan-out with multiple series termination.

4.1.2 Optimal Value of Termination

The selection of parallel or series termination is determined by the application. In brief, for an application with faster signal risetime and multiple loads along the line, such as critical clock paths, parallel termination is preferred. For an application with considerable line resistance or driver output impedance (DC loss can not be neglected), single load and low power consumption, the series termination has better performance. Reference [2] discusses the design of parallel terminations in detail. In MCMs and large SOI/SOS VLSI applications, series termination or non-termination is recommended for long interconnects.

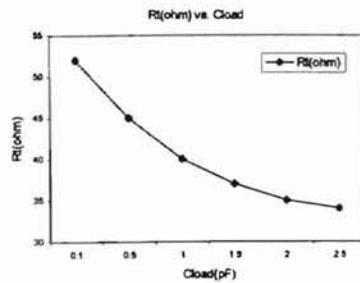
The series termination resistance is often selected to be equal to the characteristic impedance of transmission line. But Rihini Gupta [28] points out this design is only optimal for limited cases. R_t equal to Z_0 will result in over-damping of the signal and extra delay in the following cases. (1) The line driver has a significant rise/fall time, e.g., the channel resistance is considerable compared to Z_0, which is normal case for CMOS driver; (2) The line resistance is large, such as long transmission line; (3) The circuit has significant capacitive loading.

Variant factors, such as driver channel resistance (driver size), transmission line resistance, loading capacitance and fan-out topologies, have effects on the optimal termination. For the parallel fanout shown in 4.2(b), the signal reflection coupling by driver output impedance should be considered. In brief, $R_t = Z_0 - NR_{out} - R_{line}$. Here N is the number of parallel lines (limited 2~5) and every line must have same length and identical load [2].. Table 4.1 gives out the optimal termination data from simulations with different line resistance, load capacitance and driver size. Figure 4.3 (a ~ c) are

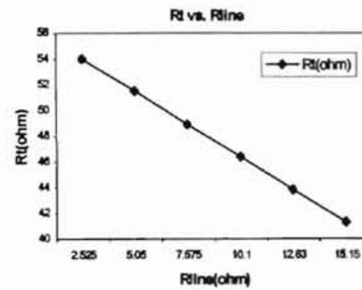
plots of termination resistance versus line resistance, width and loading capacitance respectively.

Table.4.1 Series termination simulation data

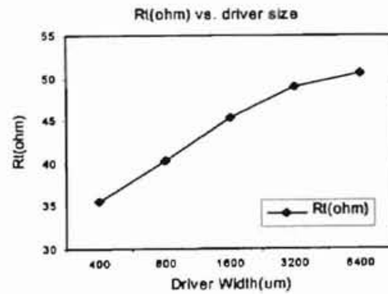
R_{line} (Ω)	R_t (Ω)	C_{load} (pF)	R_t (Ω)	R_{on} (Ω) (width(μm))	R_t (Ω)
2.525	54	0.1	52	28.24 (400)	35.5
5.05	51.5	0.5	45	14.12 (800)	40.3
7.575	48.9	1.0	40	7.06 (1600)	45.4
10.1	46.37	1.5	37	3.53 (3200)	49
12.625	43.8	2.0	35	1.765 (6400)	50.6
15.15	41.3	2.5	34		



(a)



(b)



(c)

Fig. 4.3 The function of termination resistance, R_t, vs. (a). load capacitance C_{load}; (b). line resistance, R_{line},; and (c). Driver channel resistance, R_{out}.

a. Figure 4.3a shows the optimal termination resistance ($R_{t\text{-}opt}$) as function of capacitive load (C_L). The required termination resistance (R_t) reduces as the load capacitance increases. Recall section 2.5.1, the load capacitance reduces the value of effective Z_0

(equation 2.5.4) and lower Z_0 requires smaller R_t for optimal termination. When $C_L < C_{line}$, optimal termination resistance is equal to characteristic impedance, i.g. $R_{t-opt} = Z_0$. In addition, as C_L increases, the quality factor, Q, reduces. For the same Q, a smaller R_t is required.

b. 4.3b is the function of R_{t-opt} versus line resistance (R_{line}). The optimal termination resistance for a lossy transmission line decreases with the increase of R_{line}. The interconnect for MCMs with high frequency signals has a significant line resistance due to its long length and the skin effect. This has a large effect on signal termination and delay.

c. 4.3c shows that the termination resistance (R_t) increases as driver size becomes larger. The driver size contributes to the signal termination by its channel resistance, which also contributes to signal risetime.

When selecting a series termination resistance, the effective characteristic impedance of line should be used. The line resistance and driver output impedance should be subtracted from the ideal series-termination value (Z_0).

4.1.3 Signal Termination Simulation

A lossy transmission line is divided into several segments as shown in figure 3.10 for simulation. Changing transmission line length, value of termination resistance, loading capacitance and driver size, we get the curves shown in figure 4.4 and data in table 4.1.

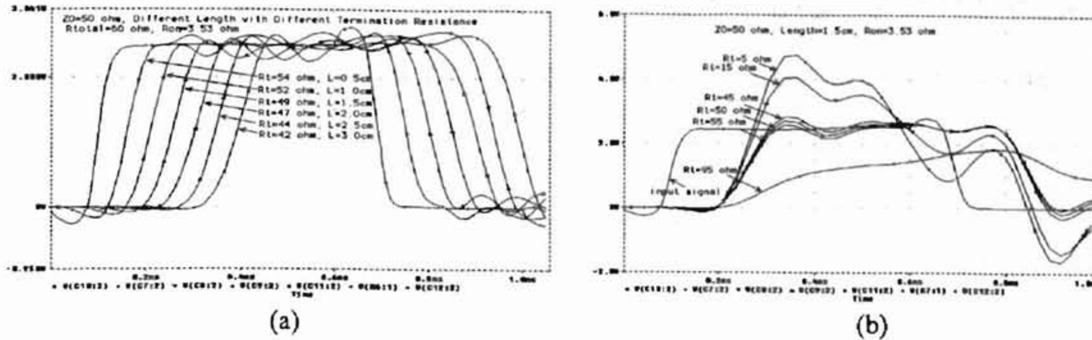


Fig. 4.4 Simulation waveforms of series termination. (a) Optimal termination with different transmission line length; (b). Different termination resistance with fixed line length i.e. line resistance.

From the simulation waveforms and data, the following conclusions are made:

- Series signal termination is determined not only by series termination resistance but also line resistance and driver channel resistance.

$$R_{teff} = R_{line} + R_{out} + R_t \tag{4.1.2}$$

- The total resistance R_{total} for optimal value for termination is roughly

$$R_{teff} \approx (0.9 \sim 1.5) Z_0 \tag{4.1.3}$$

Figure 4.4(b) shows the simulation results of optimal termination resistance that all the overshoot or undershoot are less than 10%. $R_{teff} = (1 \sim \pi) Z_0$ can also be used in R_{teff} estimation [28]. Here the effective characteristic impedance of line should be used for Z_0. Combining equations 4.1.2 and 4.1.3, optimal termination resistance is easy to be estimated.

- Too small a termination resistance causes excessive ringing. The quality factor should keep small (Figure 4.4b).

$$Q = \sqrt{L_{total} / C_{total}} \Big/ R_{teff} \qquad Q \approx 0.5 \sim 1.1 \tag{4.1.4}$$

- The different line geometries will have different termination resistance.

- The optimal series termination can be obtained by changing any of followings, line geometry (R_{line}), driver size (R_{on}) and termination resistance (R_t), separately or adjusting them together.
- The signal self-termination, $R_t=0$, can also be obtained by designing proper driver size and line geometry in some cases.
- There are also other several topologies of signal termination, such as split parallel termination, capacitive termination, etc. [2].

4.2 Driver Size Design

As noted previously, transmission line driver size is a very important parameter in interconnection design. Based on the analysis in Chapter 3 and the previous section, it is known that the driver width has significant effects on signal propagation delay of fully lossy transmission line and reflection noise control. It is also an important parameter in the estimation of maximum available line length (see section 3.2.1). The output impedance of driver is one of the most important factors in determine the simultaneous switching noise [1] (see section 4.4). In this section, we discuss how to estimate the driver size with a known risetime, optimal termination, fixed line width and length (R_{line}) and fixed load capacitance (C_L).

4.2.1 Drive Size Estimation Model

As we discussed in chapter 3, a transmission line can be modeled as lumped *RLC* segments in series and signal delay along line is independent on output impedance of driver. It is difficult and complex to estimate signal risetime t_r using the distributed *RLC* model. A simple lumped model is desired to estimate t_r and determine the output impedance of the driver, i.e. driver width, combined with the consideration of driver size effects on signal termination and reduction of simultaneous noise. Figure 4.5 shows the model which is used to estimated signal risetime.

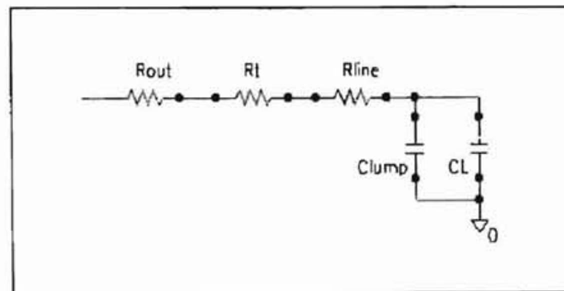


Fig. 4.5 Estimation model of signal risetime

From the above model, equation 4.2.1 is held:

$$T_{rise} = 2.2(R_{out} + R_t + R_{line})(C_{lump} + C_L) \qquad (4.2.1)$$

Here, R_t is the termination resistance

R_{out} is the average channel resistance (output impedance) of driver

R_{line} is the resistance of transmission line

C_{lump} is the capacitance of the last segment

C_L is the load capacitance

The average channel resistance is

$$R_{ohmic} < R_{out} < R_{on\max} \qquad (4.2.2)$$

where

$$R_{ohmic} = \frac{1}{K(V_{GS} - V_T)} \frac{L}{W} \quad (4.2.3)$$

$$K = \mu \frac{\varepsilon_{ox}}{t_{ox}} \quad (4.2.4)$$

R_{onmax} is more conservative estimation which is obtained from the current with maximum gate and source bias $(|V_{GS}|=V_{DD}, |V_{DS}|=V_{DD})$

$$R_{on\max} = \frac{V_{DD}}{\frac{\varepsilon_{ox}}{t_{ox}} \upsilon_{\max} (V_{DD} - V_T)} \frac{1}{W} \quad (4.2.5)$$

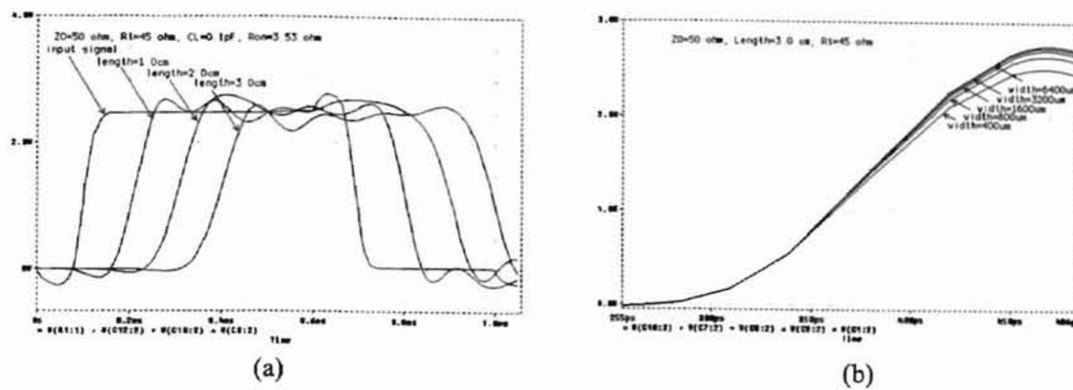
$$W = \frac{V_{DD}}{\frac{\varepsilon_{ox}}{t_{ox}} \upsilon_{\max} (V_{DD} - V_T)} \frac{1}{R_{opt} - R_{line} - R_t} \quad (4.2.6)$$

Where ν_{max} is maximum electron velocity in velocity saturation. Here, it is $1.5 \cdot 10^5 m/s$. L and W here are transistor channel length and width respectively and R_{opt} is total optimal termination resistance. In this section, R_{onmax} (4.2.5) is used to estimate the driver channel resistance.

4.2.2 Simulation and Discussion

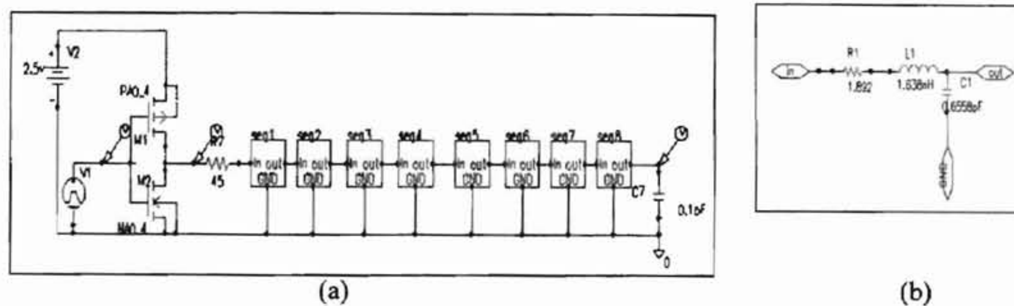
The circuit schematics shown in figure 4.7 are used to simulate the signal risetime. Changing the driver size, loading capacitance and transmission line length, several groups of data are obtained and shown in Table 4.2 and 4.3. Table 4.2 shows the simulation and calculation data for different drive sizes. Table 4.3 gives out simulation and calculation data with different line length and loading capacitance for 50Ω and 75Ω

characteristic impedance respectively. Figure 4.6 gives out the simulation results of signal risetime versus (a) transmission line length and (b) driver size.



(a) (b)

Fig. 4.6 The simulation results of signal risetime (a) Risetime vs. transmission line length and (b) Risetime vs. driver size.



(a) (b)

Figure 4.7 Risetime simulation schematics and segment

Table 4.2 The effect of driver width on signal risetime

Z_0=50Ω, Length=3cm, C_{load}=0.1pF

Width	400um	800um	1600um	3200um	6400um
$R_{out}(\Omega)$	28.24	14.12	7.06	3.53	1.765
T_{rise}(pS) (CAL)	131.97	118.49	111.75	105.88	102.95
T_{rise}(pS) (SIM)	116.47	108.26	104.79	102.89	102.58

Table 4.3 Risetime calculation and simulation data with different line lengths

(a) Z_0=75 ohm, C=1.128Pf/cm, L=6.338nH/cm, R=13.82Ω/cm, W=3.91um, t=3.0um
Termination Resistance R_t = 70Ω, C_L =0.01pF and C_L =0.1pF

Length (cm)	C_{lump} (pF/cm)	R_{line} (Ω)
0.5	0.2819	6.91
1.0	0.3759	13.82
1.5	0.4229	20.73
2.0	0.4511	27.64
2.5	0.4028	34.55
3.0	0.4229	41.46

T_{rise}(Cal)(ps) R_t = 70Ω, C_L =0.01pF	T_{rise}(Sim) (ps)	T_{rise}(Cal)(ps) R_t = 70Ω, C_L =0.1pF	T_{rise}(Sim) (ps)	T_{rise} (Cal)(ps) R_t = 35Ω, C_L =0.1pF	T_{rise}(Sim) (ps)
51.65	43.36	67.58	52.38	38.1	36.5
74.16	68.25	91.45	80.53	54.8	52.4
89.77	82.59	108.43	91.76	68.14	64.77
102.6	91.72	122.6	99.72	80.2	74.3
104.15	95.65	119.55	103.58	80.8	74.97
109.5	104.01	132.28	113.67	92.02	90.26

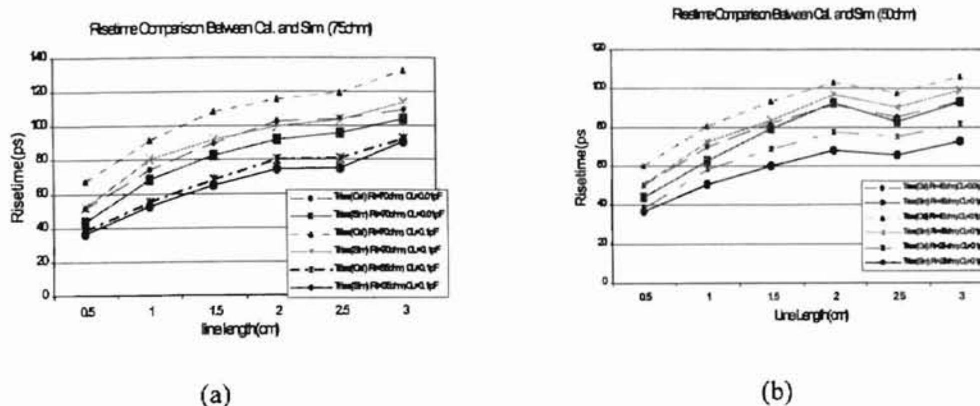
(b) Z_0=50 ohm, C=1.749Pf/cm, L=4.368nH/cm, R=5.05Ω/cm, W=16.85um, t=3.0um

Length (cm)	C_{lump} (pF/cm)	R_{line} (Ω)
0.5	0.4372	2.525
1.0	0.583	5.05
1.5	0.6558	7.575
2.0	0.6996	10.1
2.5	0.6246	12.625
3.0	0.6558	15.15

T_{rise} (Cal)(ps) R_t = 45Ω, C_L =0.01pF	T_{rise}(Sim) (ps)	T_{rise} (Cal)(ps) R_t = 45Ω, C_L =0.1pF	T_{rise}(Sim) (ps)	T_{rise}(Cal) (ps) R_t = 25Ω, C_L =0.1pF	T_{rise}(Sim)(ps)
50.23	43.9	60.33	50.45	38.28	36.7
69.9	62.62	80.5	72.05	58.4	50.4
82.18	79.24	93.29	83.57	68.83	60.03
91.52	92.33	103.1	96.68	77.37	67.95
85.38	82.89	97.49	90.25	75.28	65.6
93.27	92.93	105.88	98.96	81.71	72.6

Comparing the simulation and calculation data, figure 4.8 and 4.9 give insight into the parameter effects on risetime. Figure 4.8(a) and (b) are risetime versus line length curves for different load capacitance (C_L) and termination series resistance (R_t) for

50Ω and 75Ω characteristic impedances respectively. Figure 4.9 shows the calculation and simulation data curves of risetime versus driver size with $Z_0=50\Omega$, $C_L=0.1pF$ and $R_t=45\Omega$.



(a) (b)

Figure 4.8. Risetime comparison between calculation and simulation. (a)$Z_0=50\Omega$ and (b) $Z_0=75\Omega$, with different load capacitance, termination resistance and line width

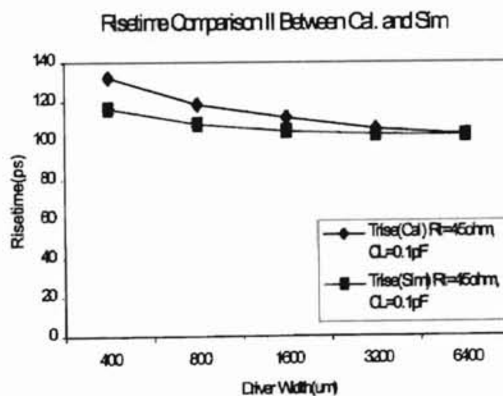


Fig. 4.9 Risetime comparison between calculation and simulation with different driver size

Based on the above analysis and simulation results, the following observations are obtained:

1. The model of Figure 4.5 is good enough to estimate the driver size with the largest error being less than 15%. The channel resistance of driver has effect on the signal risetime combined with line resistance and load capacitance. For fixed line geometry and load, model in figure 4.5 can be used for driver width estimation.
2. The error between calculation and simulation is due to the estimation of the R_{out}. R_{out} is estimated in the velocity saturation region. This is the more conservative estimation so that the calculation values are larger than those obtained from the simulation (see Dr. Johnson's notes).
3. The shorter the line, the larger the estimation error. For shorter line, line resistance is smaller so the R_{out} has more effect on the calculation result.
4. From the figure 4.9, the same conclusion can be obtained. The smaller the driver width, the larger the channel resistance, so the larger error of the.
5. The channel Resistance of the driver also has effect on the termination and the delay of signal (see section 3.2.3 and .41.2).
6. This model supplies a simple estimation of driver size for the engineer. Transmission line driver design is related to many complex factors. The driving ability and transient switching noise are two main considerations. With fixed line length, loading capacitance, optimal termination resistance and decoupling capacitance, the model in Figure 4.6 is a quick and easy method for estimating driver size.

4.3 Crosstalk

VLSI interconnects become one of the important limitation factors of today's high speed and high density circuits performances with the advent of deep-submicron technologies and sub-nano second switching circuits. As chip dimensions and clock frequency increase, the wavelength of signals become comparable to interconnection length and this makes interconnects better "antennas" [1]. Mutual capacitance and inductance induce unwanted electrical coupling known as crosstalk noise on neighboring wires. Whenever a signal edge travels along a signal trace, bond wire and connector lead, both a forward and a backward noise pulses are generated. Crosstalk is one of the most critical noises between interconnect wires of high-speed and high-density circuits. Several authors [9][14] offered experimental simulation results and concluded that crosstalk is dependent on substrate material, buffer size, termination resistance, interconnect line length and spacing between the lines. C.T. Chang and G.A. Garcia suggested that coplanar waveguides are much better than two parallel microstrips built on the same substrate [29]. In this chapter, capacitive crosstalk of both microstrips and coplanar waveguides are discussed and the simulation results of crosstalk versus line space, termination resistance, drive size and line length are given.

4.3.1 Calculation of Crosstalk

Microstrips and coplanar waveguides are the primary interconnect structures for very high-speed integrated circuits. Compared to other kinds of interconnections, coplanar waveguides are easier to connect to shunt circuit elements and maintains the same characteristic impedance when its dimensions are scaled. These advantages are due

to signal and ground conductors of coplanar waveguides being located on the same side of insulating substrate. Figure 4.10 shows the cross-sections of microstrips and coplanar waveguides.

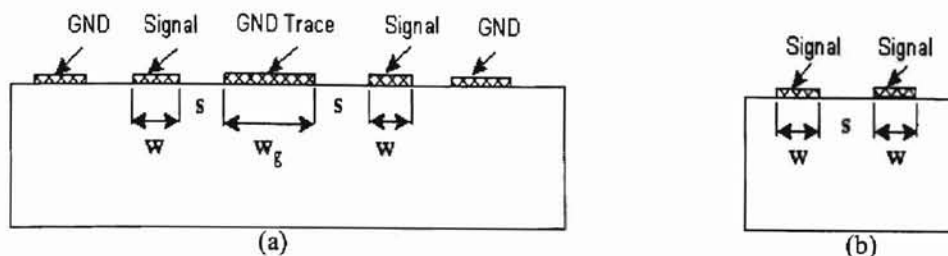


Fig. 4.10 Cross-sections of (a) coplanar waveguides; (b) microstrips

Where w is transmission line width, s is the space between the signal and ground trace. w_g is ground trace width. For microstrips, s is the space between two signal traces.

The method of conformal mapping and an elliptical transformation are used by C.T. Chang *et. al* [29] and Y.C. Lim *et. al* [30] to calculate the capacitive coupling between two coplanar waveguides. The method of conformal mapping converts the planar geometry of coplanar waveguides to the geometries of parallel plate transmission line. More physical insight than original planar geometry is provided and approximation of capacitive coupling is allowed by conformal mapping. The elliptic integrals are reduced to simplified forms. In brief, the elliptical integral transformation used here is defined as:

$$F(\alpha,\Phi)=\int_0^{\Phi}\frac{d\theta}{\sqrt{1-\sin^2\alpha\sin^2\theta}} \tag{4.3.1}$$

where α for this application is

$$\alpha=\arcsin(\frac{w}{w+2s}) \tag{4.3.2}$$

so the mapping parallel plate transmission line width is given by:

$$\Delta w = F\left(\alpha, \arcsin\frac{w+2s}{w+4s+2w_g}\right) - F\left(\alpha, \arcsin\frac{w+2s}{3w+4w+2w_g}\right) \quad (4.3.3)$$

The electrical field lines originate from one signal trace and terminate at another neighboring one by the paths through the underlying substrate and the above free space. The two capacitance in parallel are taken into account by using the dielectric constant $(\varepsilon_r + 1)\varepsilon_0$ which is equivalent to $2(\varepsilon_r + 1)\varepsilon_0 / 2$ being the effective dielectric constant and factor representing two capacitance in parallel [30]. The coupling capacitance per unit is given by:

$$\Delta C = 1.5(\varepsilon_{eff} + 1)\varepsilon_0 \frac{\Delta w}{\Delta d}$$

$$= 1.5(\varepsilon_{eff} + 1)\varepsilon_0 \frac{F(\alpha, \arcsin\frac{w+2s}{w+4s+2w_g}) - F(\alpha, \arcsin\frac{w+2s}{3w+4s+2w_g})}{F(\frac{\pi}{2} - \alpha, \frac{\pi}{2})}$$

(4.3.4)

$F(\frac{\pi}{2} - \alpha, \frac{\pi}{2})$ is the mapping spacing between two signal traces and 1.5 is the approximate correcting factor for the end effect at the edge of the signal trace. The total coupling capacitance is given by:

$$C_{coupling} = \Delta C \cdot d \quad (4.3.5)$$

Figure 4.11 shows a simple circuit model for capacitive crosstalk between two coplanar waveguides. Here coupling capacitance is treated as lumped capacitor between two terminated transmission lines with characteristic impedance Z_0.

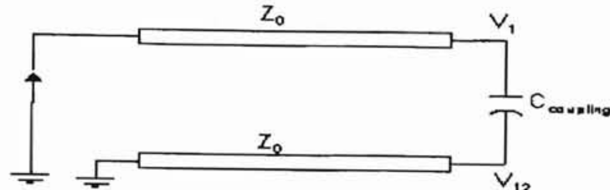


Fig 4.11 Circuit model for crosstalk due to capacitive coupling.

Based on the circuit model, the following formula is held:

$$j2\pi f\Delta Cd(V_{12}-V_1)Z_0 = V_{12} \tag{4.3.6}$$

Solving for crosstalk, and the real part of V_{12} is

$$V_{12} = \left(\frac{(2\pi f\Delta CdZ_0)^2}{1+(2\pi f\Delta CdZ_0)^2}\right)^{\frac{1}{2}} V_1 \tag{4.3.7}$$

and

$$T = 10\lg\left|\frac{V_{12}}{V_1}\right| = 10\lg\left(\frac{(2\pi f\Delta CdZ_0)^2}{1+(2\pi f\Delta CdZ_0)^2}\right) \tag{4.3.8}$$

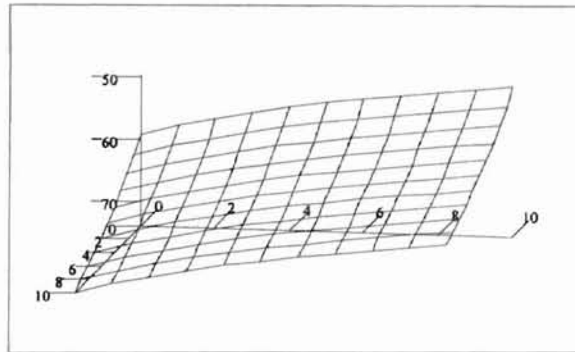


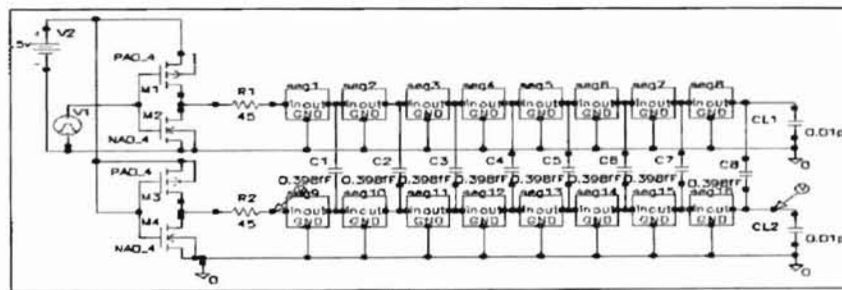
Fig. 4.12 Capacitive crosstalk of coplanar waveguides versus ground line width and signal frequency.

Figure 4.12 gives out the capacitive crosstalk changes versus ground trace width and signal frequency. Note the axis of ground line width and signal frequency are

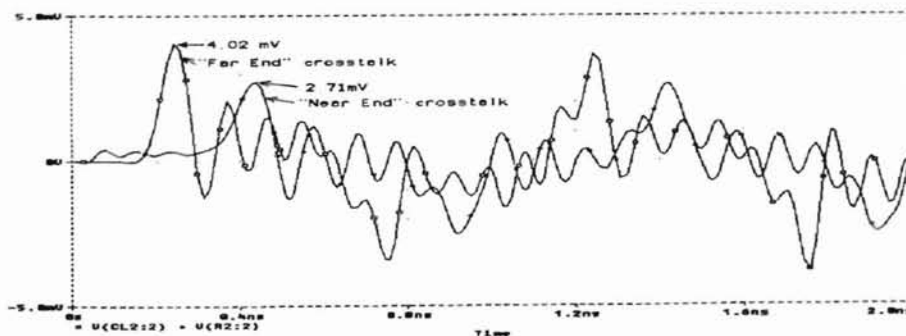
already normalized. It shows the coupling crosstalk reduces as ground separation trace width increases but increases as signal frequency increases. More detail discussion will be given in the next section.

4.3.2 Simulation of Crosstalk

The capacitive crosstalk simulation schematic is shown in figure 4.13(a). In this simulation, the transmission lines are modeled by series lumped *RLC* segments and each segment has same equivalent parameters. The values of *R, L, C* for different line lengths are different. The peak values of crosstalk at the "far end" and "near end" are measured respectively. Figure 4.13(b) shows one sample of simulation waveform and measured values.



(a)



(b)

Fig. 4.13 Capacitive Crosstalk Simulation; (a) Simulation Schematic; (b) One sample simulation waveform.

4.3.2.1 The Accuracy of Mathematical Model

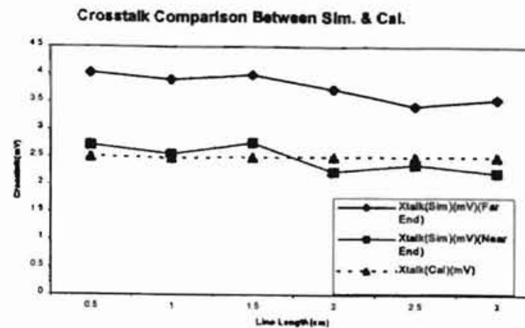
To check the accuracy of the mathematical model, transmission lines with different line lengths are used. Because of differences in line length the lumped *RLC* parameters of segments are different, however the total line capacitance, inductance and resistance are linearly scaled. For observation convenience, the coupling capacitance is kept the same for different line lengths by changing the spacing (inserted ground trace width, w_g, figure 4.10). The longer the line, the larger the spacing. Table 4.4 contains calculation and simulation data of transmission lines with six different line lengths.

Table 4.4 Calculation and simulation data for mathematical model accuracy

Length (cm)	N	w_g(um)	C(w_g) (pF/cm)	C_{ctotal}(w_g) (fF)	$C_{clu mp}$ (fF)	TN_{cross} (mV)(Cal)	TN_{cross}(mV) (Sim)(Near End)	TN_{cross}(mV) (Sim)(Far End)
0.5	2	110.8	0.6373	3.186	1.593	2.503	2.71	4.02
1.0	3	167.2	0.3169	3.169	1.056	2.489	2.55	3.9
1.5	4	210.0	0.2117	3.176	0.794	2.494	2.75	3.99
2.0	5	245.7	0.1595	3.19	0.638	2.505	2.24	3.73
2.5	7	278.0	0.1272	3.181	0.454	2.5	2.36	3.42
3.0	8	306.85	0.1061	3.182	0.398	2.499	2.21	3.55

Where w_g is the width of inserted ground trace (*GND* trace); $C(w_g)$ is self capacitance of GND trace; C_{ctotal} *and* C_{clump} are coupling capacitance for total line and per lump respectively. Note: The total coupling capacitance is roughly the same because the total crosstalk noise TN_{cross} is setup to be the same *(-60dB)*.

Here $Z_0=50ohm$, $C=1.749Pf/cm$, $L=4.368nH/cm$, $R=5.05\Omega/cm$, $w=16.85um$, $s=4.21um$, $t=3.0um$ and N is the number of segments.



4.14 Crosstalk comparison between simulation and calculation data

Figure 4.14 shows the comparison between the simulation and calculation data. The near end crosstalk data is much closer to the calculation data. As H.W. John and M. Graham [2] pointed out the forward crosstalk is proportional to the derivative of input signal and each coupling capacitance and inductance. Since the forward crosstalk arrives at the far end of line simultaneously, the total forward crosstalk is proportional to the total coupling capacitance and is cumulative. In our simulation, the crosstalk is doubled at far end because of open-circuit load. The reverse coupling is different. Even though the total coupling capacitance is same as forward crosstalk but the reverse crosstalk spread over the round trip transmission line delay and the value is not cumulative. After a round trip delay, the far end crosstalk reflects back to near end. So the far end crosstalk is larger due to open-circuit reflection. Based on the data in table 4.4 and figure 4.14, this model is accurate enough to estimate the crosstalk and only the far end crosstalk is valuable.

4.3.2.2 Crosstalk vs. Spacing between Two Signal Lines

Table 4.5 gives out the simulation and calculation data with different spacing between two signal lines. *3.0cm* and *1.5cm* long lines are used. The signal line geometries are same as those used in previous section (recall table 2.1). Figure 4.15 shows the comparison result of simulation and calculation data, where $Z_0=50\Omega$, $L=3.0cm$ *and* $1.5cm$, $Rt=42\Omega$, $C_L=0.1pF$

Table 4.5. Simulation and calculation data of crosstalk vs. spacing

Space(μm)	Crosstalk (mV, Cal., 3cm)	Crosstalk (mV, Sim., 3cm)	Crosstalk (mV, Cal., 1.5cm)	Crosstalk (mV, Sim., 1.5cm)
100	17.75	18.99	8,874	8.41
200	5.443	5.214	2.721	2.56
300	2.605	2.69	1.303	1.44
400	1.523	1.43	0.762	0.73
500	0.998	0.914	0.499	0.485

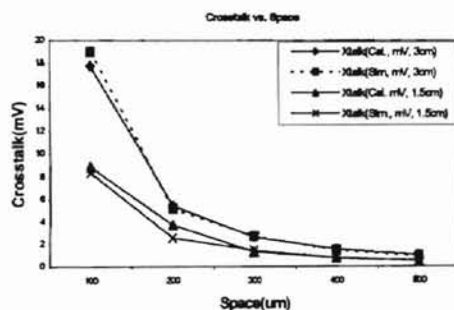


Fig. 4.15. Capacitive crosstalk vs. spacing of two signal lines.

It is easily observed that the calculation data is very close to simulation results. The model given by equations 4.3.4 and 4.3.6 is accurate enough to predict the capacitive coupling for coplanar waveguides. The following observation cab be made:

- The capacitive coupling crosstalk decreases fast as the spacing of two signal lines increases. Here the spacing is the ground trace width. The importance of ground separation trace will be further discussed in section 4.3.2.7. The ratio of the spacing (GND trace width) vs. signal line width depends on line length and required noise

margin. For example, in the above simulation case, a *3cm* long signal line with *16μm* width required *350μm* to obtain *–60dB* crosstalk (*2.5mv)*; a *1.5cm* length line requires only *220μm* to achieve the same crosstalk.

- Crosstalk reduction "saturates" as the spacing becomes larger. In the above simulation condition, when the space is larger than *500μm*, the crosstalk will not be reduced more and crosstalk difference between lines with various lengths becomes very small. This information implies we can set up a spacing point for different line lengths to achieve minimum crosstalk and here it is around *25~30* times of signal trace width. This point can be found through the PSPICE simulation.

4.3.2.3 Crosstalk vs. Termination Resistance

The series termination resistance has an effect on the crosstalk. S. Seki and H. Hasegawa [33] pointed out that the floating interconnection has larger crosstalk amplitude and as the termination resistance is reduced, the lumped *RLC* oscillation becomes dominant and determines the crosstalk amplitude (see section 4.1). As termination resistance R_t increases, the *RLC* oscillation reduces resulting reduced crosstalk. However too large a termination resistance will contribute to additional signal delay, more signal loss and larger risetime (see chapter 3 and section 4.1 and 4.2). Table 4.6 and figure 4.16 give the simulation data and curve of *1.5cm and 3.0cm* long transmission lines and confirm the crosstalk change trend discussed above.

Table 4.6 Simulation data of crosstalk vs. termination resistance R_t

Rt(Ω)	Xtalk(mV, 3cm)	Xtalk(mV, 1.5cm)
10	6.1	5.35
30	3.16	2.96
45	2.32	1.82
70	1.928	1.42
90	1.67	1.27

Where Z_0=50Ω, *Length=3.0cm and 1.5cm*, C_L=0.1pF

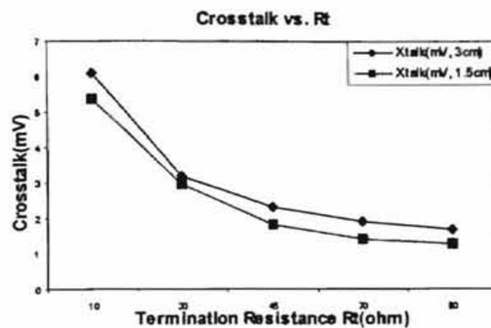


Fig 4.16 Crosstalk vs. termination resistance

As the termination resistance increases, capacitive coupling crosstalk reduces and the change becomes slowly for large termination resistance. There is a tradeoff in the selection of termination resistance between signal performance (signal delay, integrity and risetime) and noise control (reflection and crosstalk noise).

4.3.2.4 Crosstalk vs. Transmission Line Driver Size

By changing driver size and keeping the other parameters (signal line geometry, spacing and termination resistance, etc.) unchanged, the crosstalk simulation data is obtained. Table 4.7 and figure 4.16 shows the corresponding data and curves with line length equal *3.0cm* and *1.5cm* respectively. Here Z_0=50Ω, *Length=3.0cm and 1.5cm*, Rt=42Ω, C_L=0.1pF and driver size is changed from *400μm to 6400μm*.

Table 4.7 Simulation data of crosstalk vs. driver size

DriverSize(um)	Xtalk(mV, 3cm)	Xtalk(mV,1.5cm)
400	2.28	1.53
800	2.57	1.64
1600	2.61	1.71
3200	2.70	1.73
6400	2.68	1.8

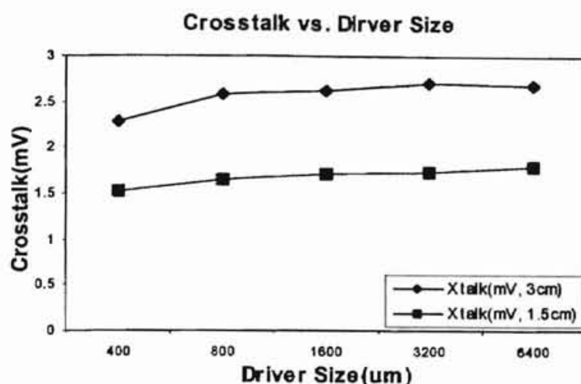


Fig. 4.17 Crosstalk vs. driver size

From the data in table 4.7 and the curves shown in figure 4.17, the following conclusions are made:

- Driver size has little effect on the crosstalk because the signal risetime is dominated by both termination resistance R_t and line resistance R_{line}. Recall the data of risetime model in section 4.2, we can find the risetime data ranges from *116.5 ps to 102.6 ps* as the driver size changes from *400μm to 6400 μm*, a variation less than 10%.
- Faster signal risetime increases crosstalk. From the Fig. 4.15, we can see that smaller drivers with larger risetime have reduced crosstalk. The signal risetime is preferred to be close to the required specification as slow as possible.

4.3.2.5 Crosstalk vs. Transmission Line Length

Changing the coplanar waveguide length and keep the inserted ground trace width fixed (*300μm*), Table 4.8 and figure 4.18 give the simulation and calculation data and comparison curves. Here $Z_0=50\Omega$, line length changed from *0.5cm to 3.0cm*, $R_t=42\Omega$, $C_L=0.1pF$ and driver width equals *3200μm (PFET)*.

Table 4.8 Simulation and calculation data of crosstalk vs. different line length

Length(cm)	Xtalk(mV) (Cal. Wg=300um)	Xtalk(dB) (Cal. Wg=300um)	Xtalk(mV) (Sim. wg=300um)	Xtalk(dB) (Sim. wg=300um)
0.5	0.434	75.2	0.75	70.46
1.0	0.868	69.19	1.18	66.52
1.5	1.303	65.66	1.77	63.00
2.0	1.737	63.16	2.14	61.35
2.5	2.17	61.23	2.52	59.93
3.0	2.605	59.64	2.76	59.14

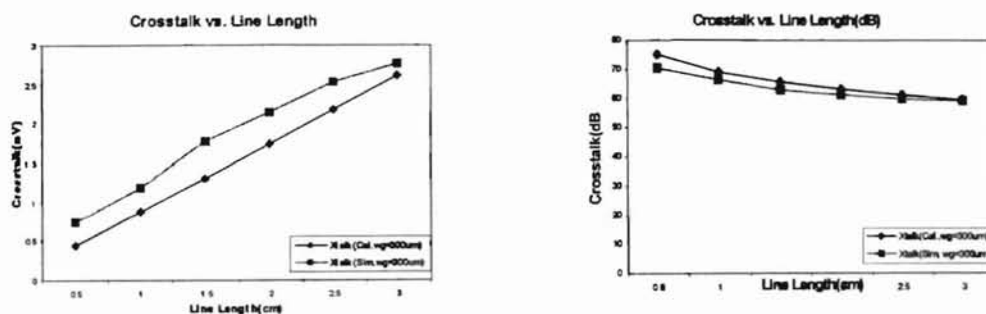


Fig. 4.18 Crosstalk vs. line length.

As the line length increases, the crosstalk and the coupling capacitance increases linearly as equation 4.3.4 indicates. For longer signal traces, larger spacing is required to maintain lower crosstalk.

4.3.2.6 Comparison of Coplanar Waveguides and Adjacent Microstrips

The ground between two signal lines plays very important role in the shielding of crosstalk noise (recall figure 4.10). Using the same model for coplanar waveguides, which has separating ground trace between two signal lines, and two adjacent microstrips without separating ground trace i.e. $w_g = 0$, the spacing data with same crosstalk in two cases are obtained and shown in table 4.9, and figure 4.19 shows the comparison results.

Table 4.9 Crosstalk simulation data for waveguides with separating GND trace and adjacent microstrip lines without separating GND trace.

Xtalk(-dB)	30	35	40	45	50	55	60	65	70	75	80
Space(μm, with GND)	43.72	63.01	88.95	123.7	170.2	232.3	315.2	425.8	573.3	770	1032
Space(μm, without GND)	76.23	141.8	258.4	466	835.2	1492	2659	4735			

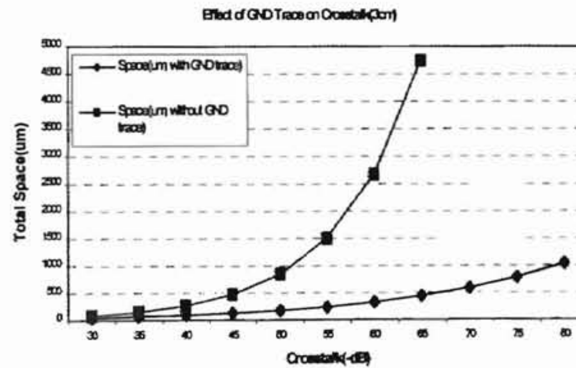


Fig. 4.19 Effect of GND trace on crosstalk

- GND separation is very helpful on crosstalk reduction. The coplanar waveguides have much less crosstalk than two parallel microstrips with same spacing. The coplanar waveguides can save wiring space over parallel microstrips

- Crosstalk reduces as the width of GND separation between two signal traces increases. The width of ground trace depends on line geometry and required crosstalk noise margin.
- Improvement in crosstalk approaches zero as the GND width increases, e.g. for 50 ohm, *3cm* coplanar waveguide, there is no further improvement when GND width *>500um* (a w_g/w ratio is about *25~30*). This point can be estimated by using the illustrated model for the worst case (maximum line length, minimum crosstalk margin or fastest signal risetime). More accurate estimation can be obtained by simulation.

4.3.3 Discussion and Reduction of Crosstalk

a. From the model used for analysis, the crosstalk is contributed by the mutual coupling capacitance. This model is more accurate for low frequency. For high frequency, the effect of magnetic coupling (the inductive coupling) will play a dominant role in determining the crosstalk. The inductive coupling decreases slowly as the spacing between two signal lines increases because the magnetic field reduces slowly along the substrate and it can not be shielded. H.T. Yuan, *et. al* [14] presents the capacitive and inductive coupling data which indicated that the mutual capacitance reduces by a factor five when the spacing between two lines increases eight times, but mutual inductance just reduces 30 percent. Therefore, inductive coupling will become more troublesome than capacitive coupling in high speed application.
b. The coupling capacitance and coplanar waveguide (CPW) characteristic impedance are not dependent on the absolute dimensions of CPW. They are dependent on the

relative dimensions, w, s and w_g. When these parameters are scaled, the coupling capacitance per unit length will remain constant [29].

c. The transmission line length, spacing between two lines, termination and output driver size (the risetime of the signal) have large and complex effects on crosstalk. In the above sections, we discussed these effects respectively. Brief conclusions are summarized as follows:

   1. Longer signal lines, larger the crosstalk.
   2. Larger termination resistance, less crosstalk.
   3. Higher frequency signals have larger crosstalk than lower frequency signals for identical geometries.
   4. Driver size has little effect on crosstalk.
   5. A GND separation trace reduces crosstalk effectively.
   6. Substrate thickness and line load also has measurable effects on crosstalk [33].
   7. *RLC* oscillation increases the crosstalk.

d. The crosstalk can be reduced by

   1. Using a ground trace between two adjacent signal traces. The ground trace width is determined by signal line geometry and crosstalk noise margin.
   2. Increasing the spacing between two adjacent signal lines.
   3. Avoiding routing parallel signal lines especially sensitive signals.
   4. Using largest the termination resistance as possible.
   5. Using the shortest line as possible.
   6. Full shielding can be used for some ultra-sensitive signal line [15].
   7. Choosing the smallest driver possible.

8. Using the transmission line with larger line resistance.

9. Reducing the substrate thickness to reduce inductive coupling.

10. Using a shielded multi interconnection scheme [33].

4.4 Simultaneous Switching Noise (Ground Bounce)

In large die SOI/SOS or MCMs system, as device geometry is scaled down, the number of devices (gates) integrated in one system has increased enormously. This results in increased switching current. In a high-speed digital system, large output drivers are used to improve the switch speed and drive ability. The increase of a driver's width increases the switching current that flows in and out the chips through power supply and ground pins. Since the drivers drive not only the chips but also the interconnects, the substrate to package wire bond and the package parasitics and all these components have inductance, current and ground bounce swing can be significant. The change of current, which is primarily due to the charging and discharging of parasitic capacitance, causes a voltage drop ($\Delta v = L\,dI/dt$) across the parasitic inductance. This kind of noise is referred as simultaneous switching noise (ground bounce). For constant field scaling, both switching current I and speed t_r scale down by factor α but parasitic inductance L doesn't scale down resulting Δv unchanged. Since V_{dd} scales down, therrefore ground bounce consumes more noise margin, i.e. $\Delta v/V_{dd}$ increases. To reduce Δv, parasitic inductance L should be scaled down but this is not effective because inductance changes as a logarithm function of line geometry. In SOI/SOS application, ground bounce becomes worse because the well to bulk capacitance, which works as temporary current

source when device switches, has been eliminated. If ground bounce is too large, it will cause logic errors.

Ground bounce is a very complex phenomenon in integrated circuits. The reasons are: (1) ground bounce is typically not significant unless large number of gates or drivers switch simultaneously; (2) ground bounce happening in different system sections can interact by sharing the same power/ground traces; (3) the enormous number of transistors in a system make it difficult to estimate switching capacitance accurately; (4) the short circuit current, as well as charging and discharging current, has an effect in determining ground bounce.

Controlling ground bounce becomes a more important design issue and has received much more attention in high speed and high pin count VLSI design. Previous work has used simple lumped element circuit models to discuss the properties of ground bounce [34][35]. A transmission line model is also used by Senthinathan [36] to describe the propagation of ground bounce along multichip module interconnects. In this study, a simple equivalent lumped circuit model is assumed.

Several techniques are used to control ground bounce. The decoupling capacitance is placed as close as possible to the chip (gates) in order to serve as temporary source of current to maintain a constant DC power supply. Damped power/ground traces due to proper trace geometry and driver width design are helpful in reducing ground bounce [9][37][38]. In addition, evenly distribution circuitry among many power/ground pins are preferred [34]. The following sections set up the model for simultaneous switching noise analysis and its equivalent *RLC* circuit. Finally the simulation results and methods to reduce the ground bounce are discussed.

4.4.1 Analysis of Simultaneous Switching Noise (Ground Bounce)

Figure 4.20 presents (a) an approximate model of the switching circuit with parasitic inductance, (b) a equivalent analysis circuit and (c) equivalent input voltage and switching current waveform. In order to simplify the analysis, we use a one-sided equivalent circuit with parasitic parameters of the power supply trace and consider the transient switch from $0 \rightarrow 1$. The behavior of the ground trace is identical to the power supply trace when the signal switches from $1 \rightarrow 0$. The switching current is modeled as triangle to approximate the real current waveform.

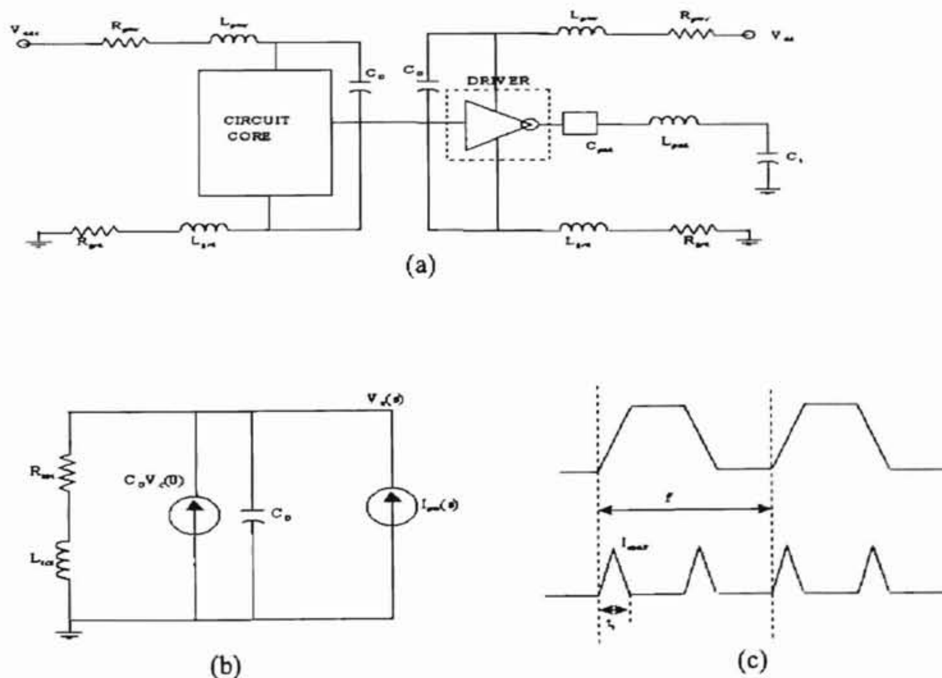


Fig. 4.20, Equivalent circuit and switching current model

In (a) an on-chip circuit core and off chip driver use separated power supplies and grounds to reduce interaction. The behavior of ground bounce in both sections is

identical and can be represented by the simplified circuit model in (b). In (b), L_{tot} is the total inductance including trace and package parasitic inductance (power and ground pins). R_{tot} is total resistance of power supply system. C_D is the inserted decoupling capacitance and $C_D V_c(0)$ represents its initial condition. When a signal switches from $0 \rightarrow 1$, the capacitor C_D's initial condition is V_{dd}. $I_{sw}(s)$ is the equivalent switching current shown in (c). The shape of switching current depends on driver ability (size) and the risetime /falltime of input signal. Here we assume signal risetime and falltime are equal. t_r is the switching time, f is the signal frequency and I_{max} is the maximum switching current. In the following analysis, only the rising edge is considered to simplify the analysis.

Writing the node equation of $V_x(s)$, the following equation is held:

$$\frac{V_x(s)}{L_{tot}s + R_{tot}} + V_x(s)C_D s - V_c(0)C_D - I_{sw}(s) = 0 \tag{4.4.1}$$

Solving for $V_x(s)$:

$$V_x(s) = \frac{V_{dd}\left(s + \dfrac{R_{tot}}{L_{tot}}\right)}{\left(s + \dfrac{R_{tot}}{2L_{tot}}\right)^2 + \left(\dfrac{1}{L_{tot}C_D} - \dfrac{R_{tot}^2}{4L_{tot}^2}\right)} + \frac{1}{C_D}\frac{\left(s + \dfrac{R_{tot}}{L_{tot}}\right)}{\left(s + \dfrac{R_{tot}}{2L_{tot}}\right)^2 + \left(\dfrac{1}{L_{tot}C_D} - \dfrac{R_{tot}^2}{4L_{tot}^2}\right)} I_{sw}(s) \tag{4.4.2}$$

let

$$\alpha = \frac{R_{tot}}{2L_{tot}}; \quad \beta = \left(\frac{1}{L_{tot}C_D} - \frac{R_{tot}^2}{4L_{tot}^2}\right)^{1/2}; \quad \omega_0 = \alpha^2 + \beta^2 = \frac{1}{\sqrt{L_{tot}C_D}} \tag{4.4.3}$$

substituting equation (4.4.3) into (4.4.2), (4.4.2) is simplified

$$V_x(s) = \frac{V_{dd}(s+2\alpha)}{(s+\alpha)^2+\beta^2} + \frac{1}{C_D}\frac{(s+2\alpha)}{(s+\alpha)^2+\beta^2} I_{sw}(s) \quad (4.4.4)$$

From (c) in fig. 4.20, the switching current at one rising edge can be expressed as:

$$I_{sw}(s) = \frac{2I_{max}}{t_r}\frac{1}{s^2}\left(1-2e^{-\frac{t_r}{2}s}+e^{-t_r s}\right) \quad (4.4.5)$$

Considering switching current, we have:

$$I = C_L\frac{dV}{dt}; \quad \text{and } dV = V_{dd}\ ,\ dt = t_r$$

$$\text{so } I_{max} = \frac{C_L V_{dd}}{t_r}. \quad (4.4.6)$$

Now Substituting it in (4.4.4)

$$V_x(s) = h_1(s) + h_2(s)$$

$$= \frac{V_{dd}(s+2\alpha)}{(s+\alpha)^2+\beta^2} + \frac{2C_L V_{dd}}{C_D t_r^2}\frac{(s+2\alpha)}{(s+\alpha)^2+\beta^2}\frac{1}{s^2}\left(1-2e^{-\frac{t_r}{2}s}+e^{-t_r s}\right) \quad (4.4.7)$$

The close form of $V_x(t)$ is unnecessarily complicated to solve. Here we use a Taylor series to approximate the second term in (4.4.7), so

$$\Delta v = V_{dd}\left[1-e^{-\alpha t}\left(\cos(\beta t)+\frac{\alpha}{\beta}\sin(\beta t)\right)-\frac{C_L t^2}{C_D t_r^2}\left(1-\frac{\omega_0^2}{12}t^2\right)\right] \quad (4.4.8)$$

Equation (4.4.8) presents an estimate of ground bounce. It can be readily observed that the first component of interests is an oscillation with a decaying magnitude. The oscillation frequency ω_0 and decay factor α are given by (4.4.3). For fast decay of oscillation, damped power traces are required. The second term of (4.4.7) can be thought as the initial offset of oscillation. Because a Taylor Series is used and only low order terms are kept to simplify the expression, the equation may introduce errors. Examining (4.4.8), it is desired this offset is very small and as a result the ratio of C_L/C_D should be

very large, i.e. $C_D >> C_L$. Note, because only signal edge is considered, t is limited, $t < 1/(2f)$. For a more accurate estimation, simulation is required.

4.4.2 Forming Damped *RLC* Equivalent Circuit of Power Trace

The inductance and resistance of power supply traces, load capacitance, and driver turn-on resistance form an equivalent *RLC* circuit. It may oscillate if it is not damped and the oscillation also induces simultaneous noise to the system. The equivalent *RLC* and simplified circuit are given by T. Gabara [37][38] and shown in figure 4.21. For more accurate analysis, further investigation is required for equivalent *RLC* circuit.

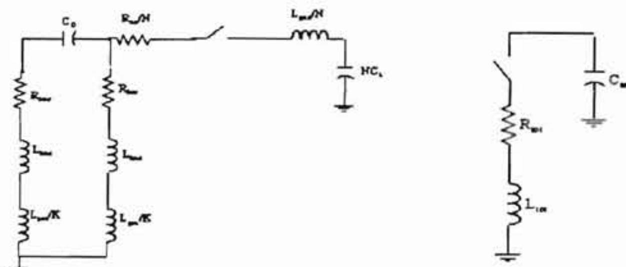


Figure 4.21 Equivalent *RLC* circuit for power supply system

In figure 4.21, L_{tot}, C_{tot} *and* R_{tot} are total effective inductance, capacitance and resistance respectively. In a MCMs or SOI/SOS large die application, we assume there are a total of K power supply and ground pins. Note, in order to simplify the analysis, we assume (a) every power or ground pin has an identical inductance and the inductance of all bond wires is much larger than total pad inductance allowing pad inductance to be

neglected; and (b) decoupling capacitance is much larger than load capacitance. The total resistance, capacitance and inductance are given:

$$L_{tot} = L_{line} + \frac{L_{grd}}{K} \quad (4.4.9)$$

$$C_{tot} = C_D \quad (4.4.10)$$

$$R_{tot} = R_{on} + R_{line} \quad (4.4.11)$$

In the above equations, L_{line} *and* R_{line} are total inductance and resistance of power supply and ground traces. C_D is the total decoupling capacitance. R_{on} is total effective turn-on resistance of the switched gates/drivers. The value of L_{tot}, C_{tot} *and* R_{tot} will determine whether an oscillation will be generated at the power/ground nodes. The critical or over damped condition is given by:

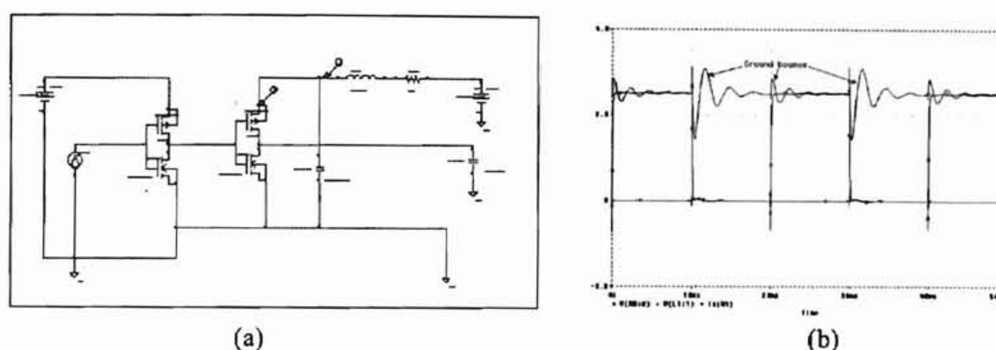
$$R_{tot} \geq 2\sqrt{L_{tot}/C_{tot}} \qquad \text{critically or over-damped} \quad (4.4.12a)$$

$$Q = \sqrt{L_{tot}/C_{tot}}\,/R_{tot} \quad (4.4.12b)$$

Resistive damping techniques to reduce the switching noise have been studied [37][38]. If the equivalent *RLC* circuit is either critically or over-damped, no oscillation happens and ground bounce decays faster. However, the response of signal becomes slower due to the increased resistance and capacitance and the design concern becomes the delay of switching. When total resistance or capacitance is reduced and condition (4.4.12a) is not satisfied, the equivalent *RLC* circuit exhibits oscillatory behavior resulting in ground bounce. This also can be observed in the simulation waveform (4.23d). The design issue becomes how not to tradeoff the switching delay or driver delay and avoid the switching noise. Over-damping of the equivalent circuit can be obtained by the proper design of power traces, driver size and proper number of power

pins. The power pin inductance plays the dominant inductance role in ground bounce. The limited number of physical *I/O* pins makes reducing pin inductance difficult. In many cases, power and ground leads can consume up to *30%~40%* of the total *I/O* pins to ensure a low inductance path. In summary, the total capacitance can be increased by inserting more decoupling capacitance and the smallest driver possible while satisfying the risetime specification. The Simulation result (4.23e) are in agreement with the analysis.

4.4.3 Simulation of Ground Bounce



(a) (b)

Fig. 4.22 Ground bounce simulation circuit and typical response

Figure 4.22 shows the simulation circuit and typical ground bounce response. In the simulation, we just observe the ground bounce on power supply. The observed bounce on ground trace is exactly similar as that on power supply. Two inverters are used here for simplicity. One (2nd) works as switching driver and the other buffers ideal pulse to ensure a worst case but realistic clock edge into 2nd inverter. Those two inverters have different power supplies to isolate switching current. Line inductance and

resistance data is taken from table 2.1. C_L represents the total equivalent switching capacitive loads.

The following simulations have been done to observe the effective changing current and ground bounce by changing parameters.

(a). Load capacitance is changed but the other parameters are kept constant. Here C_L=1pF, 10pF and 50pF; C_D=100pF; L_{line}=5nH; R_{line}=5Ω; Driver Width W=3200μm. The measured ground bounce and switching current are shown in 4.23 (a).

(b) Decoupling capacitance is changed but the others are kept constant. Here C_L=10pF; C_D=20pF, 100pF and 200pF; L_{line}=5nH; R_{line}=5Ω; Driver Width W=3200μm. 4.23(b)

(c) Inductance is changed but the others are kept constant. Here C_L=10pF; C_D=100pF; L_{line}=1nH, 5nH and 10nH; R_{line}=5Ω; Driver Width W=3200μm. 4.23(c)

(d) Resistance is changed but the others are kept constant. Here C_L=10pF; C_D=100pF; L_{line}=5nH; R_{line}=1Ω, 5Ω and 10Ω; Driver Width W=3200μm. 4.23(d)

(e) Driver width is changed but the others are kept constant. Here C_L=10pF; C_D=100pF; L_{line}=5nH; R_{line}=5Ω; Driver Width W=320μm, 1600μm and 3200μm. The effective output impedance are about 5Ω, 15Ω and 32Ω. 4.23(e)

(f) Selection of decoupling capacitance. Keeping resistance, inductance and driver width constant, the ratio of decoupling capacitance and load capacitance are varied to observe ground bounce. Table 4.4.1 gives out the simulation data. Here, C_L=1pF, 5pF, 10pF, 15pF, 20pF and 25pF; C_D=25pF, 50pF, 100pF, 200pF, 300pF, 400pF, 500pF, 600pF and 700pF. L=5nH, L=5Ω and Driver Width W=3200μm.

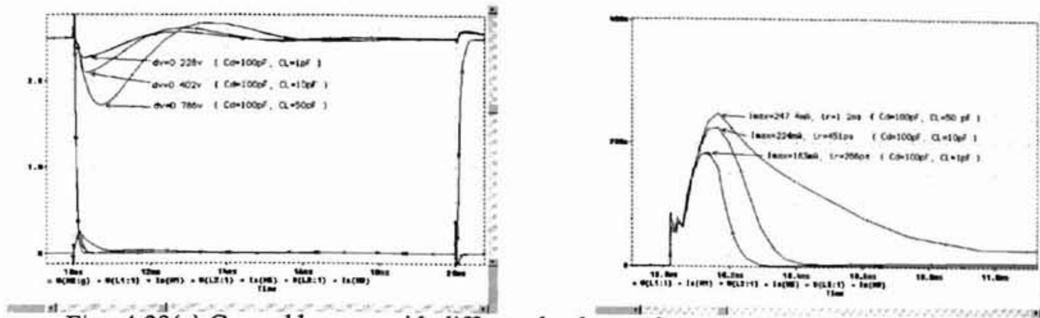


Fig. 4.23(a) Ground bounce with different load capacitance

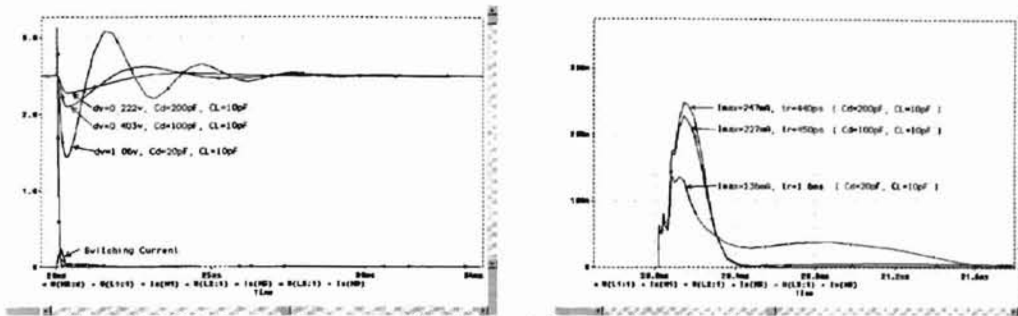


Fig. 4.23(b) Ground bounce with different decoupling capacitance

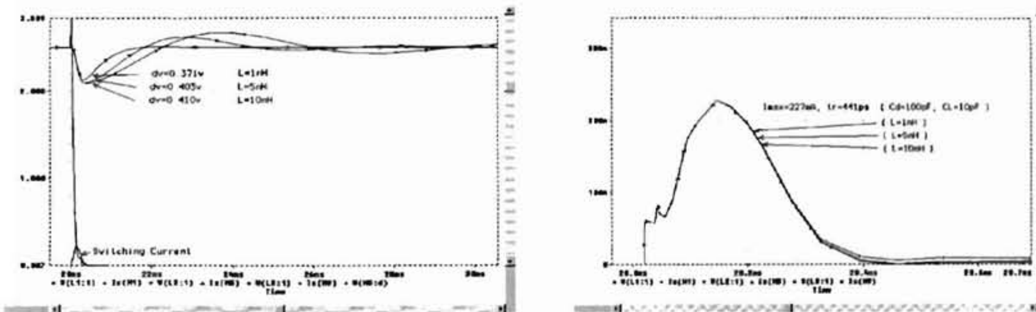


Fig.4.23(c) Ground bounce with different equivalent inductance

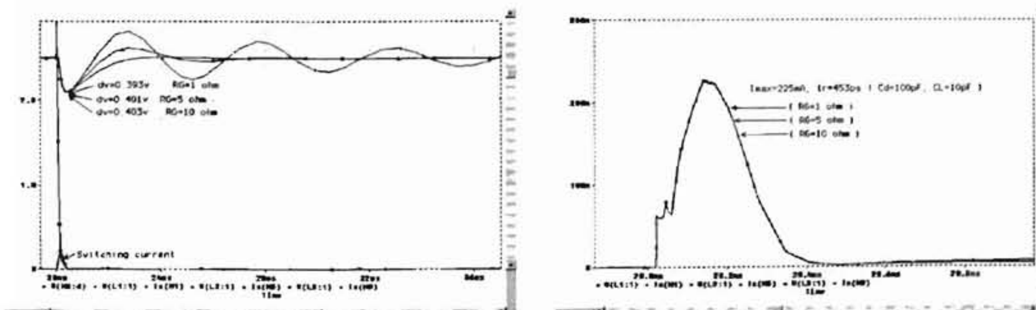


Fig. 4.23(d) Ground bounce with different line resistance

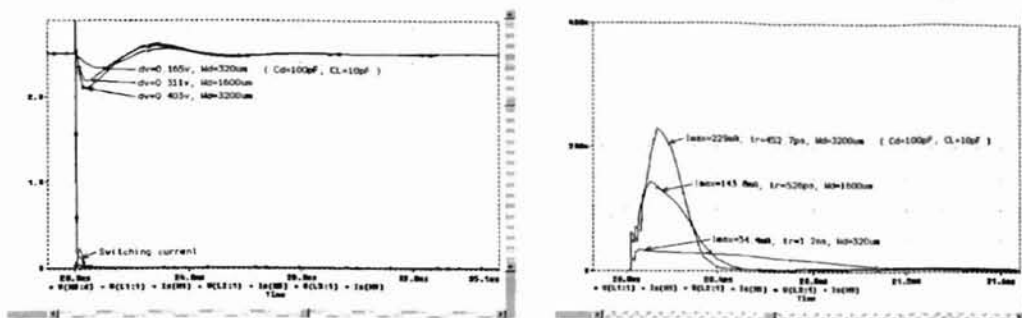
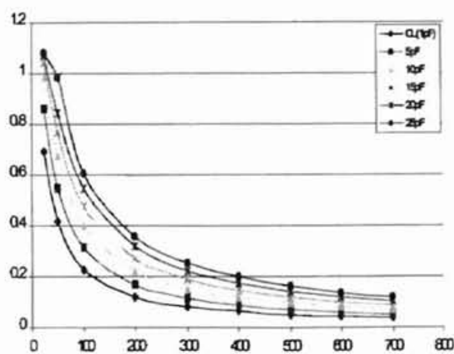


Fig. 4.23(e) Ground Bounce with different driver width

Table 4.10 Ground bounce simulation data

$C_D \backslash C_L$ (pF)	1	5	10	15	20	25
25	27.7%	34.4%	39.4%	41.8%	42.8%	43.2%
50	16.7%	21.8%	27.0%	30.7%	33.7%	35.8%
100	9.16%	12.5%	16.1%	19.1%	21.8%	24.1%
200	4.84%*	6.68%	8.88%	8.68%	12.7%	14.3%
300	3.32%	4.60%*	6.16%	7.60%	8.92%	10.2%
400	2.52%	3.52%	4.68%*	5.84%	6.92%	7.92%
500	2.04%	2.84%	3.84%	4.72%*	5.60%	6.44%
600	1.68%	2.36%	3.20%	3.96%	4.72%*	5.48%
700	1.44%	2.04%	2.76%	3.44%	4.08%	4.72%*



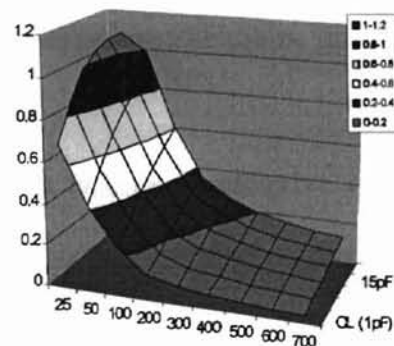


Fig. 4.24 Ground Bounce vs. decoupling capacitance to load capacitance ratio

From the simulation waveforms and measured results, the following phenomena can be observed.

(a) The ground bounce consists of two parts: the first undershoot and oscillatory response (equation 4.4.8 presents same results). Oscillation of the power system makes ground

bounce worse (4.23(d)). The first undershoot or is determined by total switching current (4.23(a), 4.23(b) and 4.23(e)) and the oscillatory response is determined by damping of the equivalent *RLC* circuit.

(b) The switching current depends on load capacitance C_L, decoupling capacitance C_D and the driver size. The larger the average (area under current waveform) switching current, the larger the first undershoot. Line resistance and effective has little effect on switching current (4.23(c), 4.23(d)). So resistance and inductance have a minimal effect on first undershoot.

(c) The first undershoot can be reduced by either increasing the ratio of decoupling capacitance against load capacitance (4.23(a)(b)) or reducing the driver width (4.23(e)). In an application, total switching capacitance is roughly constant and application dependent. The decoupling capacitance and driver size are two design issues to reduce the first undershoot. Decreasing driver width will slow down signal. Under the satisfaction of delay and risetime/falltime specifications, the smaller driver is preferred.

(d) The first undershoot is not very sensitive to a change in effective inductance and resistance. A change of inductance doesn't change switching current as much as a decoupling or load capacitance change does. In another words, it is more advantageous to reduce load capacitance if possible or increase decoupling capacitance before trying to reduce inductance or resistance in order to decrease the first undershoot.

(e) The total effective capacitance, inductance and resistance determine the power system damping. For a constant decoupling capacitance, a large inductance requires a large

resistance to ensure proper damping (4.4.12) or rapid decay (4.4.3). Increasing line resistance increases power dc loss. Inductance is kept as small as possible to ensure easily damping in power supply.

(f) Line resistance is used as compensation for a low turn-on resistance to ensure power system damped. The turn-on resistance is limited by signal risetime and/or clock rate. The geometry of power distribution lines can be designed to obtain the required resistance.

(g) The data in table 4.10 and figure 4.24 demonstrates that ground bounce decreases as the ratio of the decoupling to load capacitance increases. When C_D/C_L increases, the decrease of ground bounce begins to slow. This implies there is a critical point which we refer to as the minimum value of decoupling capacitance. First, let's check out the capacitance for damping. The output impedance of driver with *3200μm* width is about *5Ω* and line resistance is *5Ω*. So total resistance is *10Ω*. Substituting *L=5nH* and *R=10Ω* into (4.4.12), the damped capacitance is *200pF*. Checking data curves in Fig. 4.24, $C_D=200pF$ is at our critical point. The ground bounce with a decoupling capacitance greater than $C_D=200pF$ decreases much slower than that with a smaller decoupling capacitance. Based on experimental data in Table 4.10 and figure 4.24, we observed that the decoupling capacitance should be larger than the required capacitance for damping. From the simulation curves and experimental fitting, we have:

$$C_D = C_{damp} + KC_L = 4L_{tot}/R_{tot}^2 + KC_L \qquad (4.4.13)$$

where C_{damp} is the minimum capacitance to achieve a damping and K is a correcting factor depending on the required ground bounce noise margin. For a high ground

bounce tolerance, K can be chosen small. In the above case, for $\Delta v < 0.125V$, i.e. *5%* ground bounce, K is chosen as 20. The data marked by an asterisk in table 4.10 demonstrates the experimental fitting data for equation 4.4.13 which provides an estimate of the decoupling capacitance. If KC_L is very small compared to C_{damp}, the term of KC_L can be neglected and only the damped situation is considered. Normally K should be chosen as *10~20.*

4.4.4 Summary of Ground Bounce Reduction

- Ground bounce consists of two main parts. The first undershoot is determined by the switching current; and the oscillatory response is determined by power system damping. To reduce ground bounce, both components should be reduced.
- Larger C_L implies larger C_D.
- A reduction in switching current results in reduced ground bounce. Increasing the ratio of the decoupling capacitance to load capacitance and using as small a driver as possible are helpful in reducing the first undershoot.
- The equivalent *RLC* circuit must be critical or over-damped to avoid ground bounce. $Q = \sqrt{L_{tot}/C_D}/R_{tot}$ should be less than 1, normally Q=*0.1~0.7*. Increasing the total resistance helps system damping.
- Power pin inductance is dominant for off-chip drivers when the power line/bus is properly designed. The number of power pin can be increased to reduce inductance. Use as many total power/ground pins as economically feasible and keep bond wires short to reduce power pin inductance

- Proper design of built in decoupling capacitance helps to reduce the number of power pins while more power pins help to compensate the lack of decoupling capacitance resulting in power pin and decoupling capacitance design trade-off.
- Power/ground trace geometries should be designed carefully to ensure power system easily damping and avoid larger inductance.

4.5 Points of Design

Noise sources in the design of interconnects including signal reflection, crosstalk and ground bounce have been analyzed in this chapter. An estimation method for driver width is also presented. The design highlights of those noises control are summarized as follows.

(a) Reflection Noise

- The selection of termination type, i.e. parallel or series, is determined by the application. For faster signal risetime, multiple loads and lossless transmission line, parallel termination is preferred. For long interconnects (lossy transmission line), single load and lower power, series termination is preferred.
- Line resistance, driver output impedance and load capacitance affect the value of termination resistance. In brief, termination resistance reduces as line resistance or load capacitance increases but as driver width decreases. The effective termination resistance is composed of line resistance, termination resistance and driver output impedance (equation 4.1.2). The optimal value is $R_{teff} = (0.9 \sim 1.5)Z_0$.

- The termination resistance should be chosen to ensure the line being properly damped. The optimal series termination can be obtained by adjusting the line geometry (line resistance), driver size and termination resistance. Self-termination may be realized by proper design of line geometry and driver size.

(b) Driver Size Design

- Driver output impedance (R_{out}) make contributions to signal delay for lumped RC, signal risetime, signal termination, crosstalk and ground bounce control.
- To obtain maximum line length, driver output impedance should be much less than line characteristic impedance Z_0. Faster signal risetime and small signal delay also require smaller R_{out}. However, in the control of crosstalk and ground bounce, smaller driver width is preferred to slow down the signal risetime. The selection of driver size should be a compromise between the signal risetime and noise margin. Normally, for a CMOS interconnect/pad driver, the output impedance ranges from a few tens ohms to a few ohms.
- Driver output impedance is not constant and linear when signal switches. The estimated output impedance is calculated in the velocity saturated region (equation 4.2.5, 4.2.6).

(c) Crosstalk Noise

- Transmission line length, spacing between two lines, termination resistance, signal risetime have large and complex effects on crosstalk. In brief, crosstalk increases as line length or signal risetime increases; however it decreases as spacing or termination resistance increases.

- A ground separation trace between two parallel signal lines is very helpful in reducing crosstalk. Coplanar waveguides have much less crosstalk than two parallel microstrips with same spacing and save more wiring spacing. The width of a ground separation trace depends on line geometry and required crosstalk noise margin. Equation (4.3.4) is accurate for estimating crosstalk.
- Crosstalk reduction becomes slow as ground trace width is increased to a certain value. Beyond this value, crosstalk can not be effectively reduced further. This value can be estimated for worst case (maximum line length, minimum crosstalk margin and fastest signal risetime, etc.). Normally it roughly equal (25~30) times of signal line width. More accurate estimation can be obtained by simulation.
- All signal traces and ground separation traces should be properly damped.
- Use as small a driver as possible.

(d) Ground Bounce

- Ground bounce consists of two parts. One is the first undershoot, which is determined by average switching current. The other is oscillatory response, which is caused by under damping of the *RLC* equivalent circuit of the power system.
- The first undershoot can be reduced by decreasing switching current, which implies a smaller driver and a smaller load capacitance. The decoupling capacitance can reduce the first undershoot effectively and its value should be at least larger than the required capacitance to achieve damping. If the capacitance for damping is not sufficient, more capacitance should be added (KC_L). K is correct factor and ranges (10~20) (equation 4.4.13 and table 4.10).

- The equivalent *RLC* circuit of power/ground path should be critical/over damped, $Q=0.1\sim0.7$. A larger inductance requires larger decoupling capacitance for proper damping.
- Large line resistance contributes to rapid decay of oscillation. The power distribution line can be properly designed to obtain required resistance.
- The first undershoot is less sensitive to resistance and inductance than load and decoupling capacitance. It is more advantageous to reduce load capacitance or increase decoupling capacitance before trying reducing inductance or resistance in the control of ground bounce.
- To minimize ground bounce, first, the power supply system should be properly damped by increasing the number of power pins (minimizing inductance) and properly designing power trace geometries (lower inductance and larger resistance). Then calculate required capacitance to achieve damping as a basic starting value for decoupling capacitance. Increase decoupling capacitance according to simulation results and choose a correcting factor K for application.

CHAPTER 5

NOVEL LAYOUT DESIGN FOR LARGE TRANSISTOR

This chapter describes a novel layout for large width transistor using a triangle shaped sub-cell which is developed to mitigate the effect of both the gate resistance (R_g) and source resistance (R_s) on large width and short channel geometry transistors. Both a triangle and diamond structure are discussed with the triangle structure compared to the classical finger structure [48]. This chapter also reviews the effect of R_g on cutoff frequency, f_T, maximum frequency of oscillation, $f_{\max}$, thermal noise, device delay and Gain-Bandwidth Product (*GPB)* of large width transistors and the result shows that R_g is required to be much less than $1/g_m$ for minimum effect, i.e. $R_g \leq 1/((6 \sim 7)g_m)$. The effect of R_s on g_m, g_o and bandwidth of large transistor is also covered. The triangular cell has the largest perimeter along with the ability to be compacted resulting in a larger effective width and smaller cost of area compared to other right polygons in a fixed area. When compared to the conventional finger or tree structure with identical geometry W/L, the novel triangle structure reduces R_g by a factor of fifteen, R_s by a factor of twenty, improves $f_{\max}$ three times while increasing *GBP*, dynamic range *(DR)* and avoiding the additional gate delay. It also eliminates source degeneration and reduces area cost by half.

5.1 Overview of Problems

The operating frequency of circuits keeps going up as the device dimensions are scaled down, and as a result the relative influence of gate resistance R_g and source resistance R_s becomes increasing significant to the design of wide transistors [51][52][54]. In particular, in the design of the off-chip pad and long interconnection line drivers used in digital systems and high gain, low-noise amplifiers in analog applications, a transistor with large channel width must be employed to obtain the high current drive, transconductance and dynamic range required. More specifically large transistors must maintain large *GBPs* (Gain Bandwidth Product) and high *DR* (Dynamic Range) which can only be achieved by controlling both R_g and R_s. As the transistor width increases, the gate resistance typically introduces additional delay, reduces f_{max} and increases the thermal noise floor. The presence of a source resistance also degrades forward transconductance g_m (source degeneration). Both the effects degrade the transistor performance.

Previous methods and layout structures to reduce gate resistance include the use of many narrower transistor fingers connected in parallel to obtain an equivalent large width while lowering gate resistance further by silicidating the gate ploy [43][44] or using *Al* on the gate finger (T-gate) [45][52]. To reduce source resistance wider a metal trace is often applied. But these methods have certain limitations as will be demonstrated. Increasing the number of narrower transistor fingers increases both interconnect lengths and widths of source and drain leading to increased resistance and

capacitance and a reduction of circuit density. The gate silicide thickness must scale with channel length, thereby yielding a higher sheet resistivity for shorter device length [46].

Section 5.2 analyzes the effect of gate resistance on f_T, f_{max}, thermal noise and device delay. Section 5.3 describes the effect of R_s on g_m, g_o and bandwidth of large transistor. Section 5.4 develops the triangle structure and derives the effective total width, source and gate resistance. In this section, a diamond structure is also developed and compared to a triangle structure. Section 5.5 gives a direct comparison of triangle and finger structures. Finally section 5.6 summarizes the results.

5.2 The Effect of R_g on f_T, $f_{\max}$, Thermal noise and Transient Response

To analyze the effect of gate resistance R_g and source resistance R_s on large width transistor performance, the following ac equivalent lumped circuit and small-signal models of a MOSFET are used. In those figures, k is the loading factor, $C_{load} = kC_{gs}$, and W and L refer to transistor channel width and length respectively, where C_{gs} is the gate to source capacitance, g_m is the transconductance and g_{ds} is the output conductance.

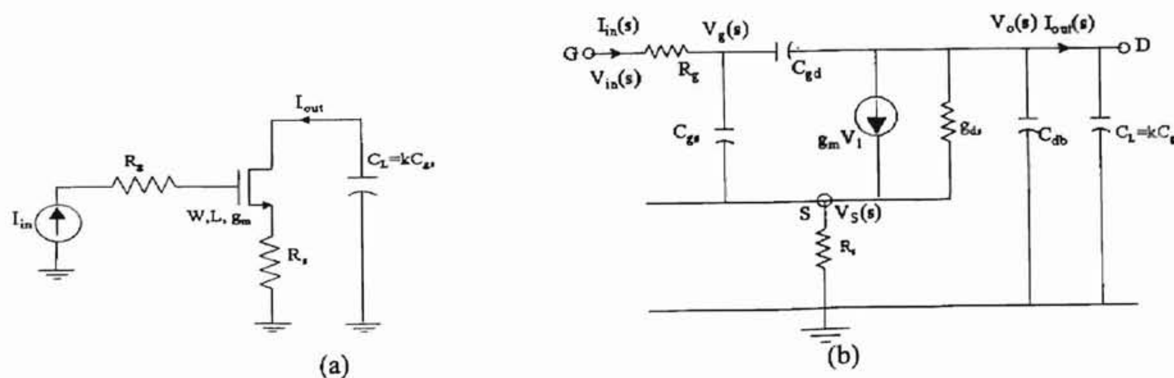


Fig 5.1 Equivalent circuit models used in analysis of R_g effect on transistor performance. a) lumped model; b) small-signal model of a MOSFET.

5.2.1 Cutoff Frequency f_T

The unit current gain bandwidth, defined as f_T [43], is a figure of merit for transistor speed exclusive of loading and gate resistance offering an insight into potential broadband utility. Figure 5.1(a) shows the gate resistance is modeled as lumped resistor in series with gate terminal and figure 5.1(b) presents the small signal model of a device including gate and source resistance (R_g and R_s) with bulk connected to ground. Based on the small signal model, the following set of node equations are held:

$$I_{in}(s) = (V_{in}(s) - V_g(s))/R_g \qquad (5.2.1)$$

$$(V_{in}(s) - V_g(s))/R_g - (V_g(s) - V_s(s)) \cdot sC_{gs} - (V_g(s) - V_o(s)) \cdot sC_{gd} = 0 \qquad (5.2.2)$$

$$I_{out}(s) = -g_m(V_g(s) - V_s(s)) - (V_o(s) - V_s(s)) \cdot g_{ds} - V_o(s) \cdot sC_{db} - (V_o(s) - V_g(s)) \cdot sC_{gd} \qquad (5.2.3)$$

$$V_s(s)/R_s = g_m(V_g(s) - V_s(s)) + (V_o(s) - V_s(s)) \cdot g_{ds} + (V_g(s) - V_s(s)) \cdot sC_{gs} \qquad (5.2.4)$$

Solving above set of equations for unit current gain $I_{out}(s)/I_{in}(s) = 1$, we get cutoff frequency is:

$$f_T = \frac{1}{2\pi}\frac{g_m}{C_{gs}} \qquad (5.2.5)$$

Taking into account the velocity saturation effect for short channel devices, where $g_m = WC_{ox}v_{sat}$, and v_{sat} is the saturation velocity, the cutoff frequency is given by:

$$f_T = \frac{g_m}{2\pi C_{gs}} = \frac{\upsilon_{Sat}}{2\pi L} \qquad (5.2.6)$$

Equation (5.2.5) and (5.2.6) indicate f_T is independent on R_g and transistor width.

In fact, for transistors with large width, the gate resistance is distributed along the gate width. A distributed model by decomposing the large transistor into n small devices is used for analysis by Behzad and Yan [43] with the resulting equation being the same as equation (5.2.6).

5.2.2 Maximum Frequency f_{max}

$f_{\max}$ is the maximum potential frequency of an oscillator employing that device [46] and it is determined by the frequency at which the unilateral power gain of a device drops to unity. Several authors have computed and discussed the effect of finite R_g on $f_{\max}$ using various approximations [44]-[47]. However, these approximations don't consider velocity saturation effects, substantial overlap capacitance, low output impedance and the distributed nature of gate resistance in large transistor deep submicron MOSFETs and may yield inaccurate results.

Behzad and Yan [43] gave out detail analysis for $f_{\max}$ without considering the effects of source resistance R_s and developed the following equation:

$$\omega_{\max}{}^2 \approx \frac{g_m{}^2 r_0}{4R_g(C_{gs}+C_{gd})[C_{gs}+(1+g_m r_0)C_{gd}]} \qquad (5.2.7)$$

Taking into account R_s and since solving is a simple matter of substitution we will present f_{max} in terms of a velocity saturated model. Using the small-signal model in Fig. 5.1(b) the f_{max} is given by

$$\omega_{\max}{}^2 \approx \frac{g_m{}^2 r_0}{4R_g(C_{gs}+C_{gd})[(1+\frac{R_s}{R_g})C_{gs}+(1+g_m r_0)C_{gd}]} \qquad (5.2.8)$$

Assuming $g_0 << g_m$, the simplified f_{max} expression is

$$f_{\max}{}^2 = \frac{f_T{}^2}{4R_g g_m\left(\frac{L_D}{L}+\frac{1}{\mu}\left(1+\frac{R_s}{R_g}\right)\right)} \quad (5.2.9)$$

where L_D is overlap of drain(source) diffusion and gate, μ is the self-gain of a MOSFET, $\mu = g_m / g_{ds}$.

If R_s is very small, i.e. $R_s << R_g$, its effect on $f_{\max}$ can be neglected and equation can be simplified as (5.2.10). Due to the nature of the R_s parasitic, it is more easily controlled and its reduction is a direct result of carefully reduction of R_g and the compact square symmetrical layout structure with parallel source traces (see section 5.4.2). However, in fact, R_s can be a several times larger than R_g (See the extracted data in *table 5.1*) so equation (5.2.9) will often give a more accurate estimation of f_{max}. Large R_s induces source degeneration in a MOSFET as demonstrated in section 5.3.

$$f_{\max}{}^2 = \frac{f_T{}^2}{4R_g g_m\left(\frac{L_D}{L}+\frac{1}{\mu}\right)} \quad (5.2.10)$$

Equation (5.2.7) and (5.2.8) reveals some trends as gate resistance or device width varies. If gate resistance varies and the other parameters remain constant, then $f_{\max} \propto 1/\sqrt{R_g}$; If we assume that g_m and R_g increase linearly to device width W and μ is constant, then $f_{\max} \propto 1/W$.

Fig. 5.2 compares measured and calculated f_T and f_{max} for three moderate geometry transistors fabricated in the NCCOSs' (San Diego, CA) *0.4μm* two metal single poly silicated SOS process [48]. Note: Measured data is supplied by Rob Johnson,

UCSD. Three transistors were laid out with width 50μm, 100μm and 200μm using an even number of finger pairs (W/L equal $50/0.5$). The measured result confirms equations (5.2.6) and (5.2.9) and $f_{\max} \propto 1/W$.

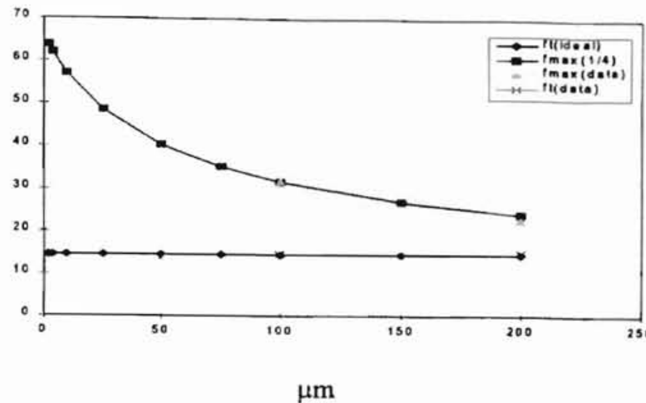


Fig. 5.2 Calculated f_T and $f_{\max}$ data versus test data

5.2.3 Thermal Noise

Thermal noise is the major noise source of transistors in high frequency applications. It is proportional to absolute temperature and places fundamental limits on the achievable dynamic range in electronic devices. It appears as white noise and can be modeled as a voltage source, $v^2(f)$. Because gate resistance R_g can not be neglected when compared to channel resistance for a large transistor, the total thermal noise is distributed by both gate and channel resistance. The thermal noise generated by gate resistance and channel resistance are presented by equations (5.2.11) and (5.2.12) [43][48]:

$$v_{Rg}{}^2 = 4KTBR_g \tag{5.2.11}$$

$$v_{ch}{}^2 = KTB\frac{8}{3}\frac{1}{g_m} \tag{5.2.12}$$

So the total thermal noise is sum of those two distributions:

$$v_{ntot}^2 = 4KTB(R_g + \frac{2}{3g_m}) \qquad (5.2.13)$$

where K is Boltzmann constant, T is absolute temperature and B is the bandwidth.

5.2.4 Device Delay

As device geometry are scaled down, the reduction of devices and logic gate delays increases the relative influence of interconnections in determining the total performance of VLSI circuits. In normal sized digital transistors, the channel resistance is so large and makes the large contribution to device delay and other parasitic resistances can be neglected. However this is not the case for large width transistor and small channel resistance (submicron channel length $L_{eff}=0.09\mu m$). The gate resistance becomes significant due to large length of gate poly, which causes large delays in the gate voltage propagation cross the transistor width. As the gate voltage propagates, the near end channel of transistor, which is close to the input, turns on or off earlier than the channel at further end. So the channel turns on partially until the gate voltage propagates across the whole poly line [51]. Using RC delay model the total device delay can be given by:

$$\begin{aligned} t_d &= \sqrt{t_{dg}^2 + t_{dch}^2} \\ &= \sqrt{(2.2R_g C_{gs})^2 + (2.2R_{ch} C_L)^2} \end{aligned} \qquad (5.2.14)$$

Substituting for the loading capacitance as $C_L = kC_{gs}$ and $R_{ch} = 1/g_m$ in velocity saturation region. Simplified (5.2.10), (5.2.11) follows:

$$t_d = 2.2C_{gs}\sqrt{R_g^{\ 2} + \left(\frac{k}{g_m}\right)^2} \tag{5.2.15}$$

Typically $k = (e \sim 5)$ for both pad and long interconnect drivers and is a lower bound for OTAs. The effect of R_s on f_T, f_{max}, thermal noise and delay has been neglected for the purpose of simplification.

From equations (5.2.6), (5.2.9), (5.2.10), (5.2.13) and (5.2.15), to avoid additional propagation delay in digital systems and degradation in potential *GPB*, $f_{\max}$, and *DR*($DR = V_{DD}^{\ 2}/V_{ntot}^{\ 2}$) in both broad and narrow bandwidth analog systems, the R_g term is required to be much less than the channel term, i.e. $R_g \le 1/((6 \sim 7)g_m)$. Considering the distributed nature of gate resistance for large transistor, B. Razavi, *et. al.* [43] pointed out that R_g should be scaled to one–third and lumped into a single resistor in series with the gate terminal for more accurate calculation.

5.3 The Effect of R_s on g_m, g_{ds} and Bandwidth

The drain and source regions are realized by means of diffusion and therefore represent a resistance which has a different effect on the operation of a MOSFET used in saturation region or as a switch.

When the MOSFET is used as a switch, the value of R_d and R_s simply appears in series with the ON resistance (R_{ch}) of that transistor. Their values are added to the value of R_{ch} so they can have a significant influence on device delay and signal risetime [53].

When the transistor is used as amplifier in saturation region in analog application, it can be modeled as a current source. In this case the source resistance plays a

significant role. It reduces the transconductance g_m and increases the output impedance r_o ($r_0 = 1/g_{ds}$), resulting the transistor self-gain μ reduces. Fig. 5.4 shows a lumped R_s included in a simple amplifying stage.

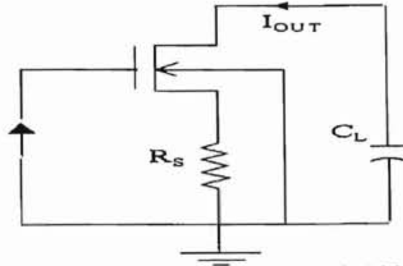


Fig. 5.3 R_s effect on g_m, g_o and bandwidth

From fig. 5.3, the following equation holds [48]:

$$V_{GS} = \Delta V + V_T + I_D R_s = \frac{I_D}{g_m} + V_T + I_D R_s \tag{5.3.1}$$

Solving I_D and deriving the effective g_m:

$$I_D = \frac{\Delta V_{eff}}{\frac{1}{g_m} + R_s} = \frac{\Delta V_{eff} g_m}{1 + R_s g_m} \tag{5.3.2}$$

$$g_{meff} = \frac{g_m}{1 + R_s g_m} \tag{5.3.3}$$

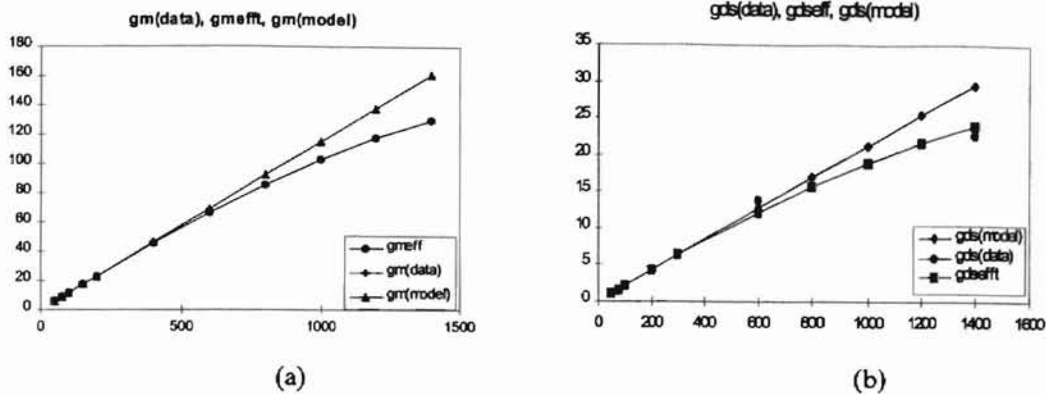
Substituting equation (5.3.3) into $r_0 = 1/(\lambda \Delta V g_m)$, the effective output impedance is:

$$g_{dseff} = \frac{1}{\frac{1}{g_{ds}} + \mu R_s} \tag{5.3.4}$$

$$r_{0eff} = r_0 + \mu R_s = r_0(1 + R_s g_m) \tag{5.3.5}$$

Figure 5.4 compares data of effective g_m *and* g_{ds} with R_s (g_{meff}, g_{dseff}), without R_s (g_{mmodel}, $g_{dsmodel}$) by equations (5.3.3) and (5.3.5) and measured data (g_{mdata}, g_{dsdata}) from

NCCOSs' (San Diego, CA) 0.4μm process. The curves give a clear picture of R_s effect on g_m and g_{ds}. Large R_s reduces g_m badly when it is comparable to $1/g_m$ and increases the output impedance introducing more difficulty in the design of transmission line driver which is preferred lower output impedance (See Chap. 3). Therefore, the source resistance must be made as small as possible ($R_s < 0.1g_m$).



(a) (b)

Fig 5.4 (a) g_m versus transistor width; (b) g_{ds} versus transistor width

For large transistors, the source metal trace becomes longer so interconnect resistance can not be negligible. Several methods are used to reduce R_s, for example, using as many as possible contacts along source-active region and making the source metal trace wider. However widening metal trace increases the total interconnect capacitance potentially slowing the device.

The following section presents a novel layout to reduce gate and source resistance and mitigates the negative effects of conventional methods and layout structures.

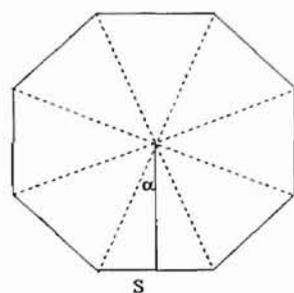
5.4 Novel Structure for Large Transistor

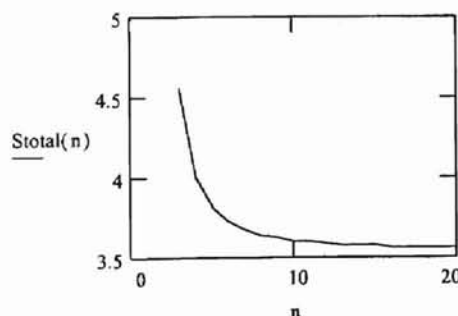
5.4.1 Selection of Basic Cell

Breaking the large transistor into small transistors (sub-cells) is an effective way to get an equivalent large width while maintaining small gate and source resistances. The shape of sub-cell plays most important role in the design. The following guidelines are basic considerations when designing a sub-cell: a) The cell should have the largest perimeter to achieve largest effective transistor width in a fixed area; b) The shape cell should be easy to compact together in a square area for drain and source resistance reduction; c) The cell shape should be symmetrical in order to route drain and source out in parallel.

Fig. 5.5(a) gives out one right polygon with *n* identical edges. It can be divided into *n* triangles. Assuming every edge length is *S* and the total area polygon is *A*, we can find that the total perimeter of the polygon is given by:

$$S_{total} = \sqrt{4nAtg\left(\frac{180}{n}\right)} \qquad (5.4.1)$$





(a) (b)

Fig. 5.5 (a) Polygon with *n* identical edges; (b). Polygon perimeter versus number of edges

Plot total perimeter S_{total} versus the number of edges in figure 5.6(b) assuming $A=1$. Fig. 5.6(b) shows that triangle (*n=3*) and diamond (*n=4*) have largest perimeter in fixed area and they are also symmetrical and easy to be compacted together.

5.4.2 Novel Layout Based on Triangle and Diamond Cell

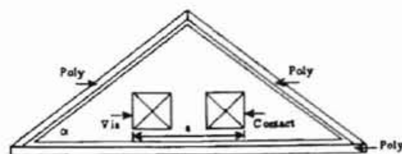
We have compared two novel structures for large width transistors to reduce its gate, source and drain resistance. One is based on a triangle cell and the other is based on diamond cell. See Figure 5.6 which shows the basic cells and structures for both layouts.

From the layout in figure 5.6, the effective width for the triangle and diamond cell are given by:

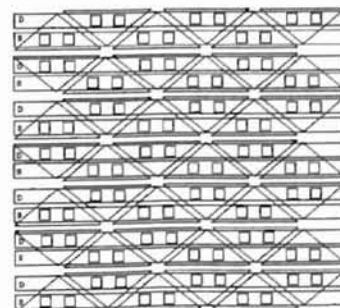
$$W_{efft}(\alpha) = \left(a + a\frac{1}{tg\alpha}\right)\left(1 + \frac{1}{\cos\alpha}\right) \tag{5.4.2}$$

$$W_{effd}(\beta) = 2a\left(\frac{1}{\cos\beta} + \frac{1}{2\sin\beta}\right) \tag{5.4.3}$$

where α, β are angles shown in fig. 5.6 and a is defined as the minimum active contact plus minimum via size plus minimum space between an active contact and via.



(a)



(b)

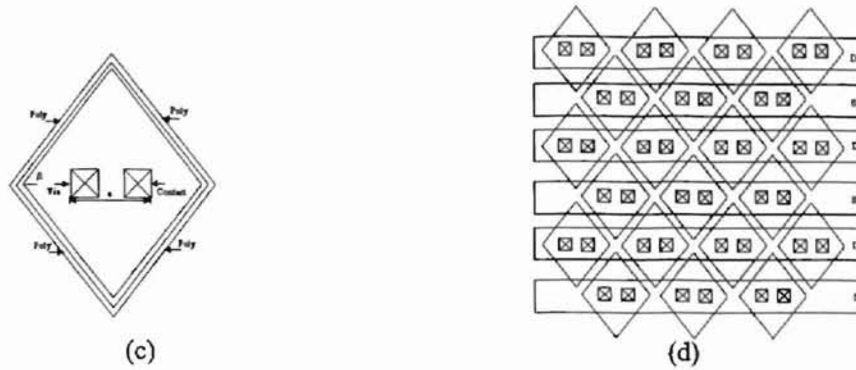


Fig.5.6 Triangle and diamond basic cell and layout structures. a). Triangle cell; b). Layout structure based on triangle cell; c). Diamond cell; d). Layout structure based on diamond cell

Assuming we use these cells to construct large transistors of fixed area $b \times b$ while taking into account the minimum spaces between diffusion islands and poly line to active field, the number of rows and columns using triangle cell are:

$$N_{rt}(\alpha) = \frac{b}{a}\frac{2\cos\alpha}{\sin\alpha + \cos\alpha} \qquad (5.4.4)$$

$$N_{ct}(\alpha) = \frac{b}{a}\frac{2\sin\alpha}{\sin\alpha + \cos\alpha} \qquad (5.4.5)$$

with the total effective width using a triangle cell in a fixed area $b \times b$ being

$$TW_{efft}(\alpha) = \frac{b^2}{a}\frac{2(\cos\alpha + 1)}{\sin\alpha + \cos\alpha} \qquad (5.4.6)$$

In the similar way, the number of rows and columns and total effective width using a diamond cell in same $b \times b$ area are derived and given by formula (5.4.7), (5.4.8) and (5.4.9).

$$N_{rd}(\beta) = \frac{b}{a}\frac{1}{tg\beta + 0.5} \qquad (5.4.7)$$

$$N_{cd}(\beta) = \frac{b}{a}\frac{tg\beta}{tg\beta + 0.5} \qquad (5.4.8)$$

$$TW_{effd}(\beta) = \frac{b^2}{a}\frac{2}{2\sin\beta + \cos\beta} \qquad (5.4.9)$$

Subscript t and d refer to the triangle and diamond devices respectively.

When we estimate the source resistance R_s, we need to take into account the minimum width of a metal trace (m_2), the minimum space between two metal traces and the number of vias and contacts which will scale inversely proportional to the number of cells. The maximum widths of the m_2 in the triangle and diamond structure are:

$$L_{\max t}(\alpha) = \frac{a}{4}(tg\alpha + 1) \qquad (5.4.10)$$

$$L_{\max d}(\beta) = \frac{a}{2}(tg\beta + 1) \qquad (5.4.11)$$

From the Figure 5.6(b) and 5.6(d), we find that all source (drain) traces are in parallel, i.e. the metal trace resistance of source (drain) is in parallel. So the added R_s and R_d due to cell interconnect are approximated by formula (5.4.12) and (5.4.13):

$$R_{st}(\alpha) = \frac{R_m b}{N_{rt}(\alpha)L_{\max t}(\alpha)} = 2R_m \qquad (5.4.12)$$

$$R_{sd}(\alpha) = \frac{R_m b}{N_{rd}(\alpha)L_{\max d}(\alpha)} = 2R_m \qquad (5.4.13)$$

where R_m is the sheet resistance of m_2 and the resistance of the contacts and vias are neglected. Here source resistance equals drain resistance although in general this need not be true.

Using a similar method, we can derive the gate resistance as:

$$R_{gt}(\alpha) = \frac{aR_p}{L_p}\frac{\sin\alpha + \cos\alpha}{\cos\alpha(\cos\alpha + 1)} \quad (5.4.14)$$

$$R_{gd}(\beta) = \frac{aR_p}{L_p}\frac{2\sin\beta + \cos\beta}{4\cos^2\beta} \quad (5.4.15)$$

where R_p is the sheet resistance and L_p is the minimum width of a poly line.

Making substitutions and plotting the above formula (5.4.2)-(5.4.15), we get the plots of Figure 5.7. Plot (a) is the total effective width of a triangle (TW_{efft}) and a diamond (TW_{effd}) structure versus base angle. . Plot (b) and (c) give out source resistance (R_{st} vs. R_{sd}) and gate resistance (R_{gt} vs. R_{gd}) respectively. Plot (d) shows the ratio of total effective width to source resistance (TW_{efft}/R_{st} vs. TW_{effd}/R_{sd}), an indicator of expected source degeneration and plot (d) shows the ratio of total effective width to gate resistance (TW_{efft}/R_{gt} vs. TW_{effd}/R_{gd}), an indicator of f_{max}, and DR performance. In all the plots, the values of Y axis is normalized.

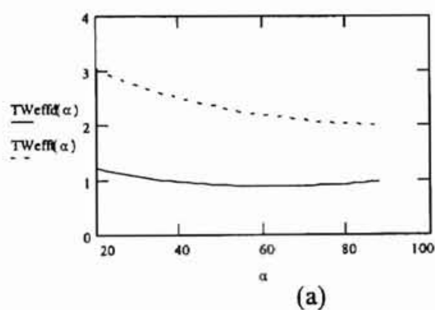
The following observations are made from the above plots:

a. A triangle structure always has a larger total effective width TW_{efft}, a higher ratio of TW_{eff} to R_s and TW_{eff} to R_g for identical contact size and fixed area $b \times b$. The triangle structure offers a more efficient performance than the diamond structure. Note that the waffle structure is a special case of the more general diamond structure.

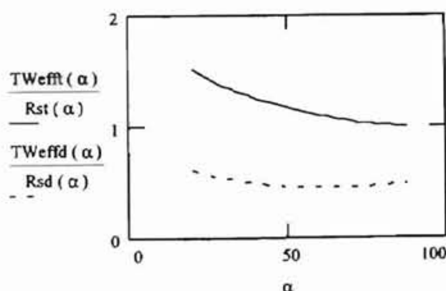
b. Source (drain) resistance is dependent on the effective transistor width and its value is small and constant. This is due to an assumed growth of a transistor in both X and Y dimensions at same rate. This characteristic keeps the total resistance constant for all

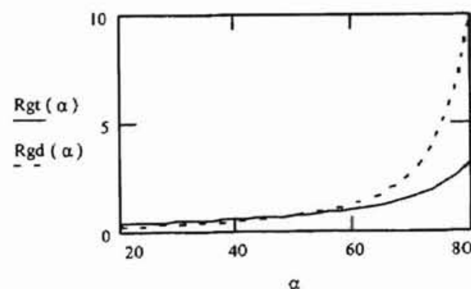
geometries with identical aspect ratios and is very helpful in reducing the source degeneration [48].

c. From fig. 5.7(a) and 5.7(c), a small angle is typically preferable since this results in a larger TW_{eff} and smaller R_g. However, the selected angle is limited by the size of the contacts and vias placed into the cell and for convenience we selected the angles $\alpha, \beta \propto (35^0 \sim 60^0)$. The plots in figure 5.7 are general useful in making design tradeoffs in drain, source and gate resistance versus circuit performance. To further reduce the gate resistance, a m_l layer placed over the poly layer (T-gate) can also be utilized [48].

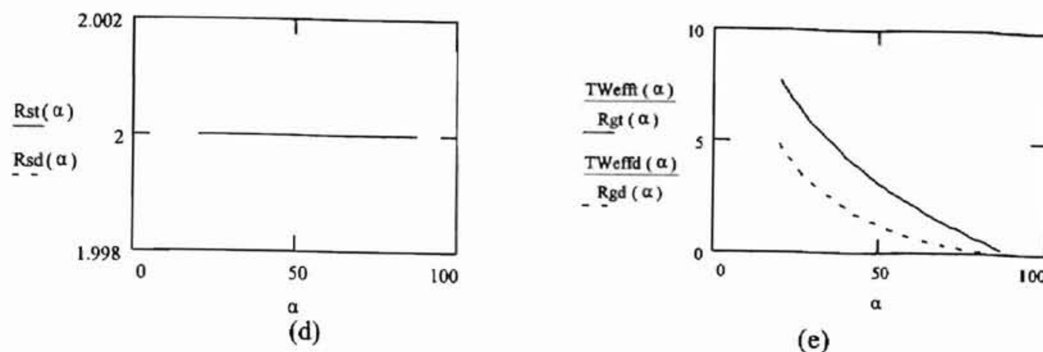


(a)





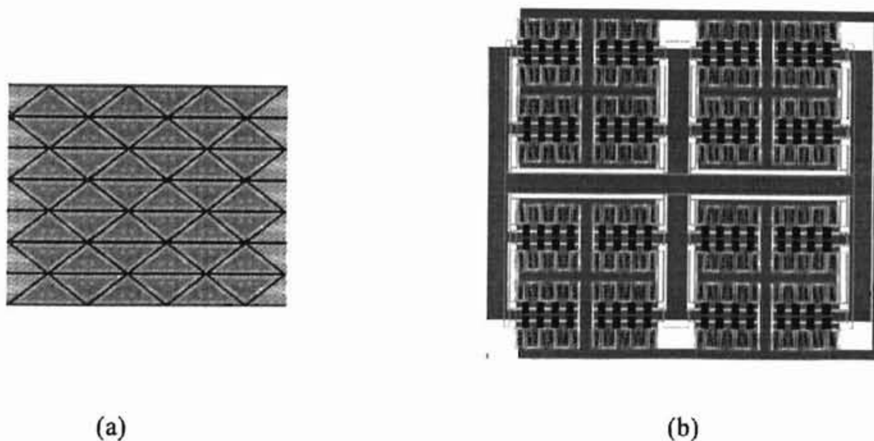
(b) (c)



(d) (e)

Fig. 5.7 Comparison of triangle and diamond structure. a). Total effective width (TW_{efft} vs. TW_{effd}); b). Source resistance (R_{st} vs. R_{sd}); b). c). Gate resistance (R_{gt} vs. R_{gd}); d). The ratio of total effective width to source resistance (TW_{efft}/R_{st} vs. TW_{effd}/R_{sd}). The ratio of total effective width to gate resistance (TW_{efft}/R_{gt} vs. TW_{effd}/R_{gd}).

5.5 Comparison of Triangle and Finger Structure



(a) (b)

Fig. 5.8 Side by side comparison of layouts with triangle and finger structure

The finger structure is one popular approach used to build equivalent large transistors. Figure 5.8(b) shows such a structure. As the effective width increases, the length of source and gate traces increase rapidly due to the branch growth resulting in

rapid increase in source (drain) and gate resistance. As a result, the source degeneration worsens and f_{max} reduces (see figure. 5.2, 5.4) while area costs go up (see figure.5.8).

Figure 5.8 shows the layouts for a finger and a triangle structures of nearly equal W/L. We used L-EDIT [49] to extract the gate and source resistance and computed effective g_m, g_{meff}/g_m (the effect of source degeneration), f_{max}, etc. from the extracted data. The following table provides a clear picture of the performance advantage of the triangle structure over classical multi-finger approach.

Table 5.1. The extracted data comparison between Triangle and Finger structure

Parameters	Triangle	Triangle	Finger
TW_{eff} (μm)	648μm	1000μm	614μm
R_g (Ω)	0.301	0.52	5.67
R_s (Ω)	0.668	0.629	15.533
$1/g_m$ (Ω)	8.57	5.56	9.05
f_{max} (GHz)	43.5	30.74	15.47
g_{meff}/g_m	0.926	0.898	0.368
$TW_{eff}/Area$	0.11	0.108	0.059
$Area$ (μm^2)	89.5×65.8	115×85	103.2×100.6

The triangle structure reduces R_g by a factor of fifteen and R_s by a factor of twenty, nearly eliminates source degeneration ($g_{meff}/g_m = 0.926$) while making twice as effective use of area and improving the f_{max} by a factor of three. It is clear from the analysis that the triangle structure potentially offers at the least a 2~4 times improvement over the classical finger structure. The R_s extracted data for *648μm and 1000μm* of triangle structure shows the R_s keeps constant while the effective width increases confirming conclusion (b) in section 5.4. This characteristic is helpful in overcoming the source degeneration effects in large transistors. Based on this improvement, one is

allowed to maintain a theoretical loading factor for digital pad drivers, theoretical gain-bandwidth product and DR in broadband analog applications and to consider CMOS devices for low noise figure, narrow band communication circuits with the noise floor limited only by the source generator.

5.6 Conclusion

The gate and source resistance have an increasing effect on transistor performance, such as f_{max}, thermal noise, device delay, transconductance and bandwidth, as CMOS circuits scale to submicron level in which the interconnection resistance becomes increasing significant and operating frequency goes high. A novel layout for large transistors based on a triangle cell was developed and the extracted data shows it reduces the R_g by one fifteenth and the R_S by one twentieth, increases f_{max} by three times when compared to the finger structure with same effective width. Source resistance maintains small and constant and only depends on the source interconnect sheet resistance. These improvements are due to the symmetric sub-cell layout organization and its resulting compactness which will thrust large CMOS transistors into communication systems.

The work was funded by NRaD, San Diego, CA and the authors would like to express their thanks to Dr. Isacc Lagnado and NRaD San Diego for their full support of this research effort and Rob Johnson for providing the s-parameter data for the finger transistor data.

CHAPTER 6

CONCLUSIONS AND SUGGESTIONS

The interconnection design has become a significant design issue in the applications of large die SOI/SOS and multiple chip modules as the device geometry has scaled down while signal frequency and die size increase. The signal delay of interconnects and noise level due to line length, signal reflection, crosstalk and ground bounce can easily dominate high-speed system performance. Basic interconnect design issues and a novel layout realization for large transistor have been discussed in this thesis. The conclusion and general design guideline are described in the following sections.

6.1 Conclusions

Interconnects refer to the medium used to connect any two or more circuit elements and differ widely in their electrical performance. RC and transmission line models are the two most frequently used models. When signal risetime t_r is much less than time of flight t_f, i.e. $t_r < 2.5t_f$, transmission line model should be used. When $t_r > 5t_f$, RC model is sufficient.

All transmission lines go through lossless, lossy and fully lossy transmission lines as line lengths and signal frequencies increase. The ratio of characteristic impedance Z_0

to line resistance R_{line} determines the type of transmission line. When $R_{line} << Z_0$, it is lossless transmission line; when $R_{line} \approx Z_0$, it is lossy transmission line; when $R_{line} >> Z_0$, fully lossy line should be used. For long MCMs interconnects, they are usually treated as lossy transmission lines due to their significant line resistance. This case will be more common for large die SOI/SOS applications (see section 3.1).

The parasitic parameters, such as capacitance, inductance and resistance determine the electrical behavior of interconnects. Small dielectric constant, narrow line width and thicker substrate are preferred to reduce the capacitance for both MCMs and SOI/SOS applications. The inductance increases in MCMs and SOS but decreases in SOI applications as the height of substrate increases. Thinner substrate is preferred in MCMs and SOS but thicker box layer is preferred in SOI for small inductance. The ideal substrate material should have smaller dielectric constant and lower dielectric loss. The proper selection of substrate material and line width is determined by which parameter, capacitance or inductance, is dominant in system performance. When $h/w >> 3$ for MCMs and SOS, the capacitance reduces very slowly. Skin effect must be considered with the optimal thickness of metal trace being the skin depth, i.e. $t = \delta_s(f)$ in high frequency applications The concept of characteristic impedance (Z_0) is just meaningful for both lossless and low lossy transmission line. Z_0 selection should satisfy the enough first incidence voltage ($Z_0 >> R_{out}$) and smaller signal resistive attenuation (see section 2.5.2). The most popular used values of Z_0 is 30Ω~100Ω.

The long interconnects in large die SOI/SOS and MCMs applications are modeled as lossy transmission lines which can be simulated using a distributed *RLC* lumped segment model. The delay of every lumped segment should be less than signal risetime

(equation 3.2.5). The signal delay depends on line geometry, i.e. the self-capacitance and inductance $t_d = length \cdot \sqrt{LC}$. Oscillation or ring of line and effect of peripheral parts like bond wire and pads should be considered in the design (see section 3.2.5, 3.3). For very long interconnection, signal loss will limits line length.

Noise control is very important in interconnect design. Considerations include reflection noise, crosstalk, ground bounce control and proper driver size design.

(a) Signal reflection

Parallel and series terminations are two most popular methods to eliminate the signal reflection. For multiple loads and faster risetime applications, parallel termination is better than series termination. But parallel termination has more dc power consumption and requires small positive reflection coefficient to compensate signal loss. For lossless transmission line, parallel termination is the first choice. For a lossy transmission line with single load and slower signal risetime, series termination is useful because of small resistive loss. The optimal termination resistance can be selected as $R_{t_{eff}} = (0.9 \sim 1.5)Z_0$. The termination resistance consists of the sum of driver output impedance, line resistance and termination resistance. Non-termination is commonly used for heavy lossy transmission line.

(b) Crosstalk

A ground separation trace is helpful in the reduction of crosstalk while coplanar waveguides are better than two parallel microstrips. Crosstalk increases linearly with the line length but reduces as the spacing between two signal lines increases. In order to reduce crosstalk, every signal and ground separation line should be "critical damping" or slightly over-damped to avoid oscillation. Larger termination resistance and slower signal

risetime (weaker driver) are preferred to reduce the crosstalk. Too wide ground separation trace width doesn't help reducing crosstalk further. The least spacing can be estimated in the worst case (maximum line length, minimum crosstalk margin or fastest signal risetime) and it is determined by signal line geometry. Roughly the ground separation trace width is around 25~30 times of signal trace width.

(c) Simultaneous noise (ground bounce)

Ground bounce consists of first undershoot and oscillatory response. The first undershoot is determined by the total switching current and oscillatory response is determined by equivalent *RLC* circuit damping of power system. Loading capacitance, decoupling capacitance and driver size has more contribution to total switching current than system inductance and resistance. The values of decoupling capacitance, inductance and resistance determine the system damping. To reduce the ground bounce, the power and ground trace should be properly damped ($Q = \sqrt{L_{tot}/C_D}/R_{tot}$, $Q = 0.1 \sim 0.7$) and a proper decoupling capacitance can be inserted between positive and negative power traces. The decoupling capacitance should be at least larger than the capacitance required for damping ($C_D = C_{damp} + KC_L$), here K is a correct factor and normally equal *(10~20)*. Minimum signal risetime (smaller diver) is preferred for reduction of ground bounce. Using as many power pins as possible to reduce the inductance and designing proper power/ground trace geometries ensuring power system easily damping and small line inductance also results small ground bounce.

(d) Driver size design

Driver size affects signal delay, risetime, termination, crosstalk and ground bounce. Larger driver is preferred for shorter delay and faster risetime but it introduces more

crosstalk and larger ground bounce. Driver size need not to be larger than required (risetime or delay). The driver width is estimated by output impedance in velocity saturation region.

For a large transistors used in the transmission line or pad drivers, the conventional finger structure has potential limitations due to excessive gate and source resistance. A triangle structure realizes twice the effective channel width per area while reducing gate resistance by factor fifteen and nearly eliminating the effect of the source degeneration. f_{max} is increased by a factor three. The triangle structure presented here is better than diamond structure and projects improved transistor performance of 2~4 times over the classical finger structure.

6.2 Test Structure

Figure 6.1 presents the test structure diagram of transmission line and large transistor. This test structure includes five parts. The oscillator generates the clock signal and clock signal is symmetrically distributed to four group 50Ω transmission lines. Four group transmission lines with different drivers, isolation ground traces and decoupling capacitances are set up to test transmission line behavior, crosstalk isolation and ground bounce reduction. For a performance comparison of large transistor with the triangle and finger structures, triangular cells are used in the last stage of buffer chain in part 3 and 1 with approximate $600\mu m$ *and* $1000\mu m$ effective width respectively. The same buffers with finger structure are used in part 4 and 2. The ground separation traces are used between oscillator and part 1 and 2. It is also setup in part 1 and 3 between

signal lines to investigate crosstalk isolation. Decoupling capacitances are inserted in all buffer chains to measure the reduced ground bounce.

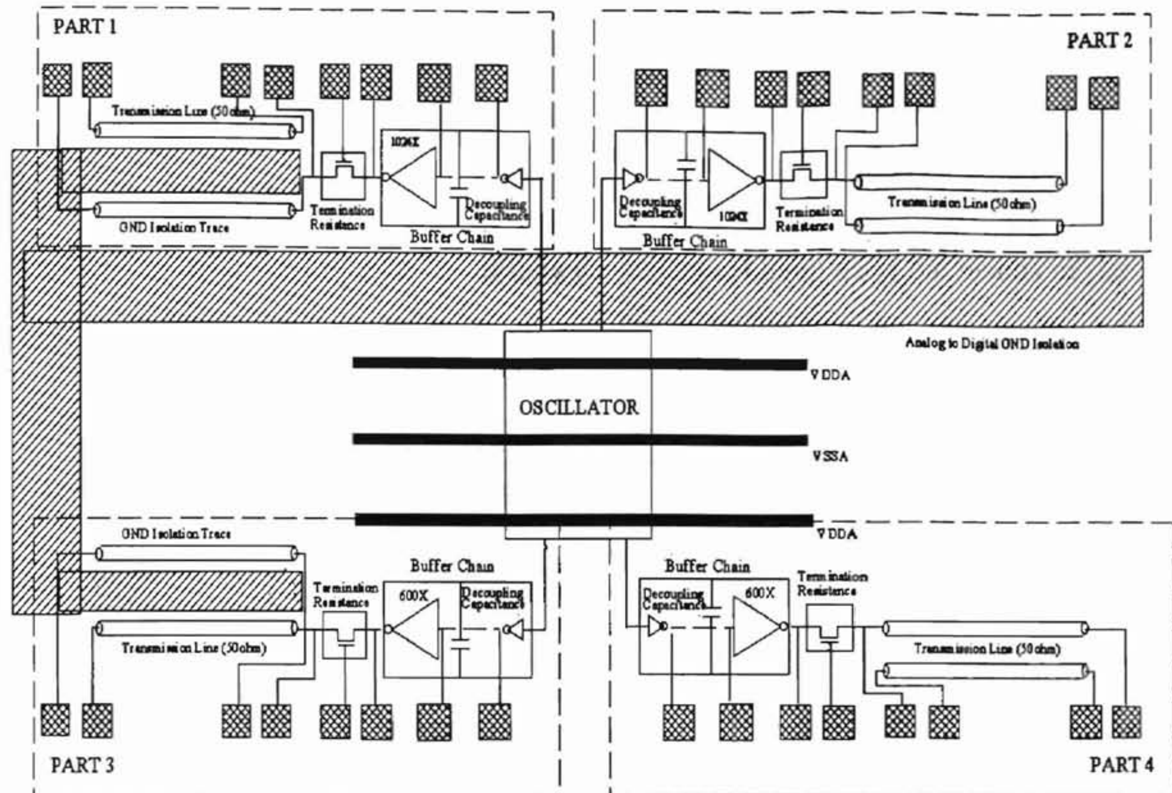


Fig. 6.1 Test structure for transmission line and large transistor

1. Analog/Digital Circuit Isolation

   Ground trace isolates part 1 and 2 from oscillator. Comparing the results of part 1,2 (with isolation) with that of part 3,4 (without isolation), we can get the information of analog/digital isolation.

2. Transmission Line Model

   Comparing the following measure results between part 1 and 3 or part 2 and 4

   a. delay

   b. risetime

c. driver size (1024X and 600X)

d. series termination(change the voltage to control the R_{on} of transistor)

3. Crosstalk

   Comparing the results of part 1 and 2 or part3 and 4. All of the four parts have same spacing between the transmission lines.

4. Decoupling Capacitance

   Every buffer chain has required decoupling capacitance. Comparing the results of part 1 and 3(1024X, 600X) or part 2 and 4

5. Large Transistor

   Triangle Structure is used for the last stage of the buffer chain in part 1 and 3 and finger structure is used in part 2 and 4. R_g and R_s effect will be observed through the measurement of the buffer output.

6. High Speed Application

   Change the oscillator frequency (>1GHz), repeat the experiments 1~5, characteristics of those parameters can be obtained for high speed application

6.3 General Design Procedure and Suggestions

General Design procedure

1. Verified the following data and specification at initial design stage
   - Available substrate ε_r
   - Available conductor, i.e. sheet resistance
   - Possible substrate thickness h

- Intended line width w
- Design frequency f
- Signal risetime and delay specification t_r *and* t_d of each signal. Use risetime no faster than requirement.
- Proposed loading situation, i.e. single or multiple loads
- Trace length range, d

2. Selection of substrate. Small dielectric constant is helpful to reduce capacitance (equation 2.3.1 and 2.3.2) and inductance (small ε_r requires thinner substrate for same capacitance resulting lower inductance (equation 2.3.3). Sapphire has larger ε_r than SiO_2 and Si resulting thicker substrate used in SOS and possible high inductance. The Si has significant magnetic loss when $f > 1GHz$. The ideal substrate for MCMs application has small dielectric constant and low magnetic loss at high-frequency. The selection of substrate thickness should be tradeoff between capacitance and inductance in MCMs application.
3. Determine whether a transmission line or lumped RC circuit should be used in your interconnects design by comparing signal risetime and time-of-flight through your possible line length range. If $t_r < 2.5t_f$, transmission line should be used; if $t_r > 5t_f$, lumped RC model should be used. If transmission lines will be used, go to the next step. Otherwise, use lumped RC analysis.
4. Check the loading. Lossless transmission lines are preferred for multiple loads and lossy transmission lines for a single load. If lossy transmission lines have to be used to drive multiple loads, parallel fanout (less than 2~5) should be used.

5. Optimize the placement to make the longest interconnect shorter and range most of long interconnects as lossless transmission lines as possible.
6. Calculate skin depth at the highest frequency of interest to select metal trace thickness (equations 2.4.2, 2.4.3). Estimate line resistance using selected width w and proposed longest length d_{max}. if d_{max} is less than your proposed largest line length, increase line width and repeat step 6.
7. Estimate characteristic impedance Z_0 (equations 3.2.2 and 3.2.3). Normally Z_0 is around $30\Omega\sim100\Omega$. If Z_0 is too large, go back to step 6 and increase line width to decrease the line resistance. If it is too small to obtain reasonable driver output impedance, go to step 6 and decrease line width. Z_0 is preferred much larger than R_{out} (normally $10R_{out}$) to ensure first incidence voltage for optimal delay (equation 3.2.2). Note in parallel fanout series termination structure, the driver output impedance is $1/N$ as the normal single fanout (see section 4.1.2).
8. Compare estimated resistance and characteristic impedance to determine transmission line type, lossless, lossy or fully lossy (table 3.1). Go back to step 4 to ensure lossless transmission line for multiple loads and lossy transmission line for single load. If necessary, repeat step 5.
9. Estimate line capacitance and inductance using chosen Z_0 and delay specification (equations 3.6.1). Based on the data of per unit resistance, capacitance, inductance and line length, verify line width and wire thickness. If not acceptable, go back to step 6, reselect line width and length, repeat steps 6~9.
10. Select termination method. Parallel termination is preferred for a lossless or low lossy transmission line with multiple loads and faster risetime. The termination resistance

could be a little larger than idea termination resistance ($R_t \approx 1.1Z_0$) for a small positive reflection coefficient to compensate signal loss. Series termination or non-termination is useful for a heavy lossy transmission line. Determine series termination resistance based on the estimated line resistance and driver output impedance (equation 4.1.2). The total optimal termination resistance is $(0.9 \sim 1.5)Z_0$ (equation 4.1.3).

11. Check the line quality factor, $Q \approx 0.5 \sim 1.1$ (equation 4.1.4), for critical damping. If not, modify termination resistance or driver width. If it is impossible for adjustment resistance due to signal risetime/falltime specification, line inductance should be reduced. If necessary, go back to 1 or 6, re-select substrate material or height and redesign interconnect geometry (see chapter 3).
12. Check the signal risetime. Adjust driver width or termination resistance (in series termination application) ensuring satisfaction of t_r/t_f (equation 4.2.1). If necessary, go back to 6 and redesign the line geometry (see chapter 4).
13. Build distributed lumped *RLC* segments (equations 3.2.4, 3.2.6~8) and simulate the delay performance. Delay estimation should include bond wire delay (equation 3.3.2). If delay doesn't satisfy specification, go back to 6 and redesign line geometry (see chapter 3).
14. Calculate crosstalk of two parallel adjacent signal lines (equations 4.3.4, 4.3.7~8). Use coplanar waveguides instead of parallel microstrips for sensitive signal traces. Based on noise margin requirement, determine the minimum spacing in worst case (maximum line length, minimum crosstalk or fastest signal risetime). This value is is around 25~30 times of signal trace width and is determined by line geometry and

crosstalk margin. Estimate the coupling capacitance and put in simulation. Larger termination resistance and smaller driver are preferred to get optimal results (see chapter 4).

15. Estimate the total switching capacitance, turn-on resistance and power system inductance and resistance. Use as many power pins as possible to reduce the power system inductance. Power/ground trace geometry should be designed carefully to ensure power system easily damping and avoid larger inductance (see chapter 4).
16. Select decoupling capacitance to ensure power supply trace properly damped. Check the quality factor, $Q \approx 0.1 \sim 0.7$ (equation 4.4.12). Use these estimation data (step 15) in equivalent switching circuit (figure 4.22) for simulation. Adjust decoupling capacitance to determine correct factor K for acceptable results (equation 4.4.13). K is normally around $(10 \sim 20)$(see chapter 4).
17. Based on the output impedance of driver, estimated the channel width (equation 4.2.5 and 4.2.6). Use buffer chain design to obtain minimum delay. For very large transistors, use a triangle structure to reduce the negative effects of gate and source resistance (figure 5.9) (see chapter 5).

The general design issues of interconnection delay and noise control are discussed in this study. It covers only lossless and lossy transmission line. More study is needed on fully lossy (distributed *RC*) transmission line for very long interconnection. More accurate models for various interconnects with different cross-sections are required to obtain more accurate extracted parasitic data and simulation results. In the discussion of crosstalk, only capacitive coupling are studied. In high frequency application, inductive

coupling will be dominant and it is more difficult to be reduced. So more studies on inductive coupling are very important in high-speed application. For ground bounce, more accurate analysis model including both power and ground parasitic parameters is desired. More direct/simple estimation of crosstalk and ground bounce is desired for engineer application. These estimation models and triangle structure for large transistors given in this study have to be verified by fabricated test results.

BIBLIOGRAPHY

[1] N. Sherwani, Q. Yu and S. Badida, Introduction to Multi-Chip Modules, 605 Third Avenue, New York, NY: John Wiley & sons, Inc. 1995

[2] Johnson, Howard W., High-speed Digital Design: a Handbook of black magic. Upper Saddle River, NJ: A simon & Schuster Company, 1993

[3] N.H.E. Weate and K. Eshraghian, Principles of CMOS VLSI Design, A System Perpective, Santa Clara, CA, Addison-Wesley Publisher Ltd. 1993

[4] K.F. Sander, Microwave Components and Systems, Santa Clara, CA, Addison-Wesley Publishing Company, 1987

[5] S. Kikuchi and H. Yamamoto et. al. "High Performance MCM-D Technology", in 1992 Sympo. VLSI Technol. Dig. Tech. Papers, pp9-14, 1992

[6] A. Masaki, "What Determines the Direction of Packaging Technology?" IEEE Trans. Components Hybrids, Manuf. Technol., vol. CHMT-10, no2, pp2-5, June, 1994

[7] D.J. Hathaway and R.R. Habra et. al. "Circuit Placement, Chip Optimization, and Wire Routing for IBM IC Technology", IBM J. Res. Develop, Vol. 40 No. 4, July 1996.

[8] S.Y. Kim, E. Tunner, et. al. "An efficient Methodology for extraction and Simulation of Transmission Lines for Application Specific Electronic Models," IEEE Trans. on Cir. and Sys. VOL. 41, pp58-64, 1993

[9] D.H. Cho, Y.S. Eo, et. al. "Interconnect Capacitance, Crosstalk and Signal Delay for 0.35μm CMOS Technology," IEEE Trans. On Cir. And Sys. Vol. 32, pp112-117, 1992

[10] "Material for Multichip Modules," Semiconductor International, pp209-212, Oct. 1996

[11] Samuel Y. Liao, Microwave Devices and Circuits. Eaglewood Cliffs, N.J. Prentice-Hall, INC, 1980

[12] M. V. Schneider, "Microstrip lines for microwave integrated circuits," Bell Syst. Tech. J. vol.48, pp1421-1444, May-June, 1969

[13] H. Hassegawa *et al.*, "Properties of microstrip line on Si-SiO_2 system," IEEE Trans. Microwave Theory Tech., vol. MTT-19. No.11, pp869-881, Nov. 1971.

[14] Han-Tzong Yuan *et al.*, " Properties of Interconnection on Silicon, Sapphire, and Semi-Insulating Gallium Arsenide Substrates," IEEE Trans. On Elec. Device. Vol. ED-29, No.4, April, 1982

[15] David Tuckerman, nChip Inc. Tutorial. MCM conference, 1994

[16] D.A. Doane and P.D. Franzon, Multichip Module Technologies and Alternatives: The Basic, Van Nostrand Reihold, 1993.

[17] Tanner Tools, Tutorial, 1995.

[18] A. I. Kayssi and K. A. Sakallah, "Delay Macromodels for Point-to-Point MCM interconnections," IEEE Trans. Cir. And Sys., Vol. 32 pp79-82, 1992

[19] R. Gupta, J. Willis, L.T. Pillage, "Wire Width Optimization of Transmission Lines for Low Power Design", In Prof. Of the Design Automation Conference, pp123-129, June, 1994

[20] Q. J. Yu, O. Wing, "Computational Models of Transmission Lines with Skin Effects and Dielectric Loss", IEEE Trans. On Cir. And Syst.. I: Fundamental Theory and Applications, Vol. 41, pp107-111, No. 2, Feb. 1994.

[21] T.A. Schreyer, Y, Nish and K.C. Saraswat, "A complete RLC Transmission Line Model of Interconnect Delay", IEEE Trans. On Solid-State Cir. Vol 36, pp95-96, 1993.

[22] T.V. Dinh and B. Cabon al et., "Modeling the Capacitance of Microstrip Line Using SPICE", IEEE Electronics Letters, vol.28, No. 2, pp194-196, Jan. 1992

[23] Y.L. Le and R. B. Iverson, "A High-Speed Multi-Dielectric Capacitance-Extraction Algorithm for MCM Interconnects". IEEE, Trans. On Solid-State Cir. Vol. 36, pp86-89, 1991

[24] S. Lum and M. S. Nakhla, "Sensitivity Analysis of Lossy Coupled Transmission Lines". IEEE, Trans. On Microwave Theo. And Tech. Vol. 39, No. 12 pp2089-2099, Dec. 1991

[25] H. Hasegawa and S. Seki, "Analysis of Interconnection Delay on Very High-Speed LSI/VLSI Chips Using an MIS Microstrip Line Model". IEEE, Trans. On Electron Device, Vol. ED-31, No. 12, pp1954-1960, Dec. 1984

[26] D.S. Gao and A.T. Yang et. al, "Modeling and Simulation of Interconnection Delays and Crosstalk in High-Speed Integrated Circuits". IEEE, Trans. On Circuit and System, Vol. 37, No. 1, pp1-9, Jan. 1990

[27] Xiaonan Zhang, "Coupling Effects on Wire Delay", IEEE, Trans. On Circuit and Devices, Vol. 26, No.8, pp12-18, Feb. 1996

[28] R.Gupta and L. T. Pillage, "OTTER: Optimal Termination of Transmission lines Excluding," IEEE J. of Solid-State Circuit, Vol. 31, pp640-645, Aug. 1994

[29] C. T. Chang and G. A. Garcia, " Crosstalk Between Two Coplanar Waveguides", Kurzberichte Letters, Band 43, pp55-58, 1989.

[30] Y. C. Lim and R. A. Moore, " Properties of Alternately Charged Coplanar Parallel Strips by Conformal Mapping," IEEE Trans. On Electron Devices, Vol. ED-15, No. 3, pp173-180, March, 1968

[31] T. Takken and D. Tuckerman, "Intergral Decoupling Capacitance Reduces Multichip Module Ground Bounce," IEEE Trans. On Electron Device. Vol. ED-32 No.12 pp79-84, 1993

[32] D. Shear, " Ground-bounce Tests," EDN Special Report. 1994

[33] S. Seki and H. Hasegawa, "Analysis of Crosstalk in Very High-Speed LSI/VLSI's Using a Coupled Multiconductor MIS Microstrip Line Model," IEEE Trans. On Electron Devices, Vol. ED-31, No. 12 pp1948-1953, Dec. 1984.

[34] Y.S. Chang, S.K. Gupta, *al et*, "Analysis of Ground Bounce in Deep Sub-Micron Circuit." IEEE Proceeding, 15th IEEE VLSI Test Symposium, pp110-116, may, 1997

[35] M.S. Haydt and R. Owens *al et*, "Modeling the Effect of Ground Bounce on Noise Margin." IEEE Proceedings of International Test Conference, pp279-285, Oct. 1994

[36] R. senthinathan, al et, "Modeling and simulation of Coupled Transmission Line Interconnects Over a Noisy Reference Plane", IEEE Trans. Comp. Hybrids, Manuf. Technol., Vol. 16, No. 7, pp705-713, Nov. 1993

[37] T. Gabara, "A Closed-Form Solution to the Damped RLC circuit with applications to CMOS Ground Bounce Estimation", IEEE Proceedings, Ninth Annual IEEE International ASIC Conference and Exhibit, pp73-78, Sept. 1996

[38] T. Gabara, al et, "Forming Damped LRC Parasitics Circuits in Simultaneously Switched CMOS Output Buffers". IEEE Journal of Solid-State Circuits, Vol. 32, No.3, pp407-417, Mar. 1997

[39] R.Downing et. al "Decoupling Capacitor Effects on Switching Noise", IEEE Trans. On Components, Hybrids, and Manuf. Technol. Vol. 16, No. 5, pp484-489, Aug. 1993

[40] J. Bandyopadhyay et al "Importance of Damping and Resonance in Thin-Film Integrated Decoupling Capacitor Design", Proc. Of 46th electronic components and Technology Conf. 1996, pp31-34.

[41] A.J. Rainal "Computing Inductive Noise of CMOS Drivers", IEEE Trans. On Component, Packaging and Manuf. Technol. –Part B, Vol. 19, No. 4, pp789-802, 1996

[42] R. Senthinnathan "Application Specific CMOS Output Driver Circuit Design Techniques to Reduce Simultaneous Switching Noise", IEEE J. of Solid-State Circuits, Vol. 28, No. 12, pp1383-1388, Dec. 1993

[43] Behzad Razavi, R. H. Yan, "Impact of Distributed Gate Resistance on the Performance of MOS Devices," IEEE Trans. On Cir. And Sys. VOL. 41, pp. 750-754, Nov. 1994

[44] Denis Deschacht, Christophe Dabrin, "Delay Propagation Effect in Transistor Gates," IEEE J. of Solid-State Circuit. Vol. 31, pp 1184-1189, Aug. 1996

[45] P. R. de la Houssaye et al. "Microwave performance of optically fabricated T-gate thin film silicon-on-saphire based MOSFET's," IEEE Electron Device Lett., vol. 16, no. 6 p.289, June 1995

[46] Y.Taur, D.A.Buchanan, et. al. "CMOS Scaling into the Nanometer Regime," Proceeding of the IEEE, Vol. 35, pp486-504, Apr. 1997

[47] S. Ohkawa, "Low noise GaAs field-effect transistor," Fujitsu Scient. Tech. J., pp151-173, Mar 1975

[48] Laker Sanaer, Design of Analog Integrated Circuits and Systems. Princeton Road, Hightstown, NJ: McGraw-Hill, Inc., 1994.

[49] Tanner Tools. Pasadena, CA: Tanner Research, Inc. 1996

[50] P.Larsson, "Parasitic Resistance in an MOS Transistor Used as On-Chip Decoupling Capacitance," IEEE J. of Solid-State Circuit, Vol. 32, pp 574-576, April 1997

[51] A. L. Caviglia and R.c. Potter et. al. "Microwave Performance of SOI n-MOSFET's and Coplanar Waveguides". IEEE Electron Device Letters, Vol. 12, No. 1, pp26-27, Jan. 1991.

[52] A.E. Schmitz, R. H. Walden, et. al "A Deep-Submicrometer Microwave/Digital CMOS/SOS Technology", IEEE Electron Device Letters, Vol. 12, No. 1, pp16-17, Jan. 1991

[53] V.V. Lee and S.A. Biellak, et. al "Series Resistance of Devices with Sunmicrometer Source/Drain", IEEE Electron Device Letters, Vol. 12, No. 12, pp664-666, Dec. 1991

[54] C. Raynaud and G. Guegan, et. al "High-Frequency Performance of Submicrometer Channel-Length Silicon MOSFET's", IEEE Electron Device Letters, Vol. 12, No. 12, pp667-669, Dec. 1991

[55] Yuan Taur et, al "CMOS Scaling into the Nanometer Regime", Proceeding of The IEEE, Vol.85, NO. 4, pp486-504, April 1997