Digital Bipolar Integrated Circuits

MOHAMED I. ELMASRY

Digital Bipolar Integrated Circuits

MOHAMED I. ELMASRY
Department of Electrical Engineering
University of Waterloo, Ontario

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York • Chichester • Brisbane • Toronto • Singapore

Library of Congress Cataloging in Publication Data:

Elmasry, Mohamed I., 1949-

Digital bipolar integrated circuits.

"A Wiley-Interscience publication."

Includes bibliographical references and index.

1. Digital integrated circuits. 2. Bipolar transistors. I. Title.

TK7874.E5 1983 621.381'73 82-21868

ISBN 0-471-05571-9

Printed in the United States of America

Preface

One of the most exciting and challenging engineering careers in recent years has been integrated circuit (IC) engineering. Many reasons contribute to this fact. In integrated circuits an engineer is given a rare opportunity to be innovative within the constraints imposed by the rules of physics. This has led to an exponential growth in many areas, for example, circuit density and performance. This in turn has widened the area of applications of ICs, and reduced the manufacturing cost. Much electronic equipment can now be built economically for the masses and the word IC is a familiar word to the public. This has no doubt enhanced IC engineers' sense of achievement and attracted many talented people to the field.

This book addresses one important aspect of IC engineering, that is, circuit design and analysis of ICs—more specifically, *circuit design and analysis of digital bipolar ICs*.

It is assumed that the reader is familiar with the basics of semiconductor physics, particularly *pn* junctions. Thus the book can be used as a text or a reference for senior undergraduate or graduate courses.

The emphasis of this book is circuit analysis and design, and this is covered in Chapters 4, 5, and 6. Because circuit designers have to know the process they design with, the model of the device they are using, and the system constraints they design within, these topics are discussed in Chapters 2, 3, and 7, respectively. In the introductory first chapter, the link between the different aspects of IC design is explained. Appendixes A, B, C, and D deal with processing steps, ion implantation, design projects, and symbolic-design layout rules for bipolar circuits.

Chapter 1 discusses the trade-offs involved in the design of bipolar digital integrated circuits among such factors as power dissipation, speed, and area. A design example is given which demonstrates that different bipolar logic families can be used on the same chip to optimize the design. It is suggested that the reader go through this example twice, once at the beginning of the course and a second time at the end, after becoming familiar with the two families under consideration, I^2L and EFL.

Chapter 2 discusses some basic bipolar processing steps and technologies. This is to serve as background information to circuit designers. Thus, no details of processing steps are given. Both junction- and dielectric-isolation

bipolar technologies are discussed. Integration of transistors, resistors, and Schottky diodes is included.

Chapter 3 deals with the characterization and modeling of the bipolar transistor. The approach is aimed toward computer-aided design (CAD) and new material is included. Examples are given. The course instructor can generate similar problems using available circuit-analysis computer programs.

Chapter 4 first discusses some basic general definitions for digital circuit design. This is followed by the analysis of the transistor–transistor logic (T^2L) family. Different circuit configurations are considered. Schottky T^2L (ST^2L) is also studied. A circuit application is given which uses PLA (programmable logic arrays).

Chapter 5 deals with the analysis of integrated injection logic (I^2L). The upward mode of operating a bipolar transistor is studied and the application of I^2L in logic and memory realizations is discussed. Typical layouts are given and I^2L-related families, for example, ISL, are also discussed.

Chapter 6 first discusses emitter-coupled logic (ECL). This is followed by the analysis of emitter-function logic (EFL). The applications of EFL in logic and storage element design and typical layouts are given.

Chapter 7 deals with the system/subsystem aspects of IC design. A design example is discussed in some detail.

Appendix A discusses in general the different steps involved in the design of an IC. Appendix B discusses in some detail the use of ion implantation in bipolar ICs. Appendix C includes a list of 30 design projects for the student reader. Finally, Appendix D offers universal stick diagram/layout design rules that can be used in association with the projects of Appendix C.

It is my hope that this book offers the reader an in-depth study of digital bipolar integrated circuits and conveys some of the know-how regarding the design of these circuits.

MOHAMED I. ELMASRY

Waterloo, Ontario
April 1983

Contents

Chapter 1

Introduction to Digital Bipolar Integrated Circuits

1.1 INTRODUCTION

The purpose of this chapter is to introduce the general topic of digital bipolar integrated circuits and to set the stage for the chapters that follow. The approach taken throughout this text is one which builds upward in complexity from the basics of integrated circuit (IC) processing technology. That is, bipolar devices and circuit blocks will be discussed always with the IC technology aspects clearly visible. The intent of this approach is to bridge the gap between the device-physics and device-technology disciplines and the computer-system designers. Moreover, this volume will provide the digital-system designer with a background vocabulary and understanding of IC techniques for bipolar digital design. As will become clear from the organization of the text, there are many facets to bipolar digital ICs (BIDICs). Two distinctly different aspects of the bipolar world are high-speed circuits, such as those used in large mainframe computers, and micropower circuits, such as those used for battery-operated electronics (i.e., calculators and watches). A third facet of BIDICs are the ubiquitous medium scale integrated (MSI) circuits used to create custom functions and interface circuitry not available in large scale integrated (LSI) circuits. The text includes chapters that address each of these families of BIDICs separately. However, there are common themes that cut across the device-technology lines of all digital families, and these fundamentals are discussed next. Specifically, the system level problem of functional partitioning and energy balance of power versus speed are now considered.

1.2 DIGITAL IC DEFINITION

The definition of a digital IC involves the confluence of a computer-system architecture and a specific IC technology. The computer-system viewpoint of digital ICs may well be characterized as a top-down approach. Starting with algorithms and high-level system descriptions, the process of design is one of

expanding downward toward finer subdivisions of the system until functional blocks at all levels are defined. IC technology, on the other hand, is a medium. In this sense it is a bottom-up solution. New technology improvements evolve from understanding the physics and materials science; however, the maturation process for a new IC technology is heavily dependent on applicability for LSI. It is this *necessary* relationship between technology and system design that will be discussed.

The story begins at a system definition and partitioning level. Given a system architecture—either for general or special purpose digital information or signal processing—a number of hardware realizations can be created. In the process of generating any particular realization the system and technology constraints shape the nature of the solution. For example, ambient conditions, available power, and packaging methodology quickly become first-order design constraints. These system-level problems apply for all systems whether they involve a mainframe computer, a portable instrument, or a robot to be launched into space. The combination of packaging technology and device technology dictate the details of subsystem partitioning. Specifically, the device technology imposes constraints on the maximum number of components per package. The package technology imposes another set of constraints—sometimes at odds with the device constraints—as to power-handling capabilities and maximum number of interconnections to other subsystem blocks. On an overall system-packaging level there are constraints as to subblock interconnections and power dissipation. It is the job of the digital LSI circuit designer to resolve all these constraints in a way that optimizes the system performance. This optimum may mean any number of things:

1. Maximum speed and/or functional complexity.
2. Minimum power and/or space.

However, given these objective functions for optimization and a set of system and subblock packaging technologies the problem reduces to a set of IC design problems at the lowest level. The system partitioning finally reaches the point where the overall design allocates a set of subblock interconnects that also imply requirements concerning signal levels, power supplies, and total power dissipation. The designer is then able to attack the problem of designing an IC chip to meet the goals and requirements of his or her particular subsystem. The phenomenal growth in complexity of digital ICs over the past decade suggests that very substantial pieces of digital architecture can now be realized as single chips. Hence, the word "subsystem" may seem restrictive as design specification. Nonetheless, the overall approach of designing a monolithic IC based on terminal and functional specifications is general and changes only in scale. To illustrate the definition problem at the IC subblock level an example is now discussed.

1.3 IC SUBBLOCK DEFINITION AND SPECIFICATION

In a digital-system design environment registers are essential building blocks. Throughout the subsequent chapters a variety of bipolar circuit technology realizations of registers will be used as examples. By way of introduction and viewing the problem top-down, Fig. 1.1*a* shows a four-bit register implemented with JK flip-flops (FF). Figure 1.1*b* shows the gate-level representation of each JK/FF using NOR and AND gates. The IC designer views the problem at even a more primitive level of abstraction as discussed later. At the device level one can postulate that each input and output reflects the need for one transistor. For MSI realizations using conventional transistor–transistor logic (T^2L) gates the number may be twice the nominal estimate, whereas, for circuit technologies such as integrated injection logic (I^2L) gates where a wired-AND function is used extensively, the actual transistor count may in fact be much lower than the nominal estimate. However, using the device count of one per input–output (I/O) as a reference, each JK/FF as shown in Fig. 1.1*b* contains nominally 26 transistors and the 4-bit register requires 104 devices. To gain some insight as to what these numbers imply from an area perspective and to correctly interpret the power–speed balance several back-of-the-envelope calculations are appropriate.

There exists a tight loop of reasoning as one seeks to resolve the questions of area–speed–power. Stated briefly the arguments go as follows. Power is proportional to the product of current times power supply voltage:

$$P = V_{CC}I \tag{1.1}$$

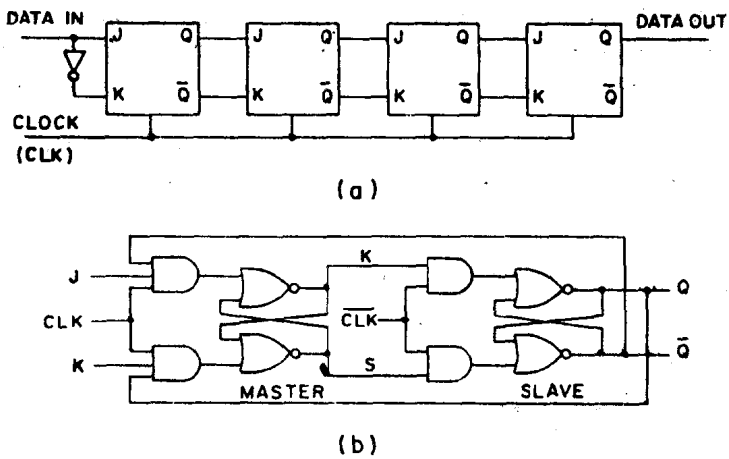

Figure 1.1 (*a*) Four-bit shift register implemented using JK flip-flops; (*b*) JK master–slave flip-flop at gate level.

The signal delay can be represented as the time needed to store or remove a packet of charge by means of a source current:

$$\tau_D \propto \frac{Q}{I} \tag{1.2a}$$

and if the charge represents only that stored on a parallel plate capacitance:

$$\tau_D \propto \frac{CV_l}{I} \tag{1.2b}$$

where V_l is the logical swing of voltage across C. Combining the above equations the power delay product, which represents energy required per switching event, is given as

$$P\tau_D \propto CV_{CC}V_l \tag{1.3}$$

As yet we have not introduced the area constraint. IC chip area is bounded by two factors: packaging and fabrication yield. The primary constraint in packaging is the amount of power a package can dissipate. Current state of the art is 1–10 W/package. Hence, given the number of devices per chip, each dissipating a given amount of power, one can determine the chip area that bounds the capacity of package. The fabrication yield problem imposes a different constraint on area. Random as well as technology-specific defects occur during IC fabrication and are measured in terms of number/cm^2. As the number of defects and die per square centimeter become comparable, the yield goes to zero. Thus one always seeks to minimize defects and keep die size within an acceptable range to produce a reasonable yield. Figure 1.2 shows a typical plot of expected yield versus chip size for a bipolar process circa 1976. Now applying each of these area factors to the power-delay equation, the generic extremes of maximum speed and maximum functional density can be defined.

Consider a 1-W package and two different IC chips—one designed to maximize speed and the other designed to maximize transistor count. For the

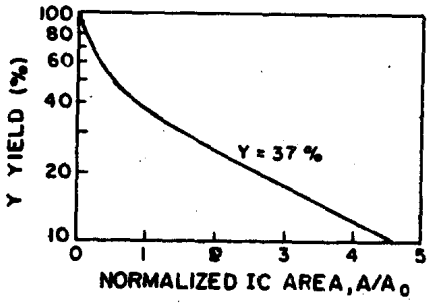

Figure 1.2 Yield logarithm versus normalized area for a bipolar process circa 1976.

case of maximum speed the following assumptions and calculations represent state-of-the-art:

$$\tau_D(\text{min}) \approx 20 \text{ ps} = 20 \times 10^{-12} \text{ s}$$

$$\text{assume } V_l \approx 0.5 \text{ V and } V_{CC} = 3 \text{ V}^*$$

there is now a trade-off between P and C. Nominally C is set by circuit fan-out and wire capacitance to other gates. For state-of-the-art minimum geometry devices and typical wiring, C ranges from 10 to 100 fF (fF = 10^{-15} farads). Using Eq. (1.3) the average power per gate becomes 0.75–7.5 mW so that given the package constraint on total power, the high-speed chip can contain on the order of 130–1300 gates. These numbers have the obvious flaw involved in assuming that *all* gates are active and the minimum delay is required everywhere on the chip. Nonetheless, the small gate counts indeed reflect the worst-case cost of high speed.

Turning to the other extreme case of maximum functional complexity, Eq. (1.3) can be used in reverse. Namely, assume

$$P = 1 \ \mu\text{W per gate}$$

$$V_l = 0.5 \text{ V}, \quad V_{CC} = 3 \text{ V}$$

so that the total gate count now becomes 1×10^6 gates/chip. Again using the capacitance bounds of 10–100 fF, the delays range between 15 and 150 ns. Clearly for the extreme of micropower to maximize component count the gate delays have been compromised substantially. It is clear that choices between 1000 gates with subnanosecond delays and 10^6 gates with hundreds of nanoseconds delays are the rule rather than the exception. Moreover, for most systems only a portion of the circuitry must run at maximum speed. Hence, it is possible to have designs with both high-speed sections as well as micropower sections. Such an example is presented shortly. First, however, it is useful to define the range of BIDIC technologies that presently span the performance space between nanosecond delays and very high density.

1.4 BIPOLAR LOGIC FAMILIES

In Chapter 2 a review of technology specifics for bipolar circuitry is presented. It is assumed that the reader has a basic familiarity with the planar process and the general meaning of dielectric isolation and Schottky-clamped circuitry. The purpose of this section is to introduce the overall framework of bipolar circuit technologies so that some of the system level trade-offs concerning speed and

*Whereas these choices are somewhat arbitrary, it is likely that for bipolar devices the limits of $V_l \approx 4\text{–}5 \ kT/q$, i.e., 100–125 mV and $V_{CC} \approx 5\text{–}10 V_l$ i.e., $\approx$ 0.5–1.5 V will be close to fundamental.

power can be viewed generally. The subsequent chapters will explore the various circuit families in much greater depth.

Figure 1.3 shows a plot of gate delay τ_D versus power per gate. The solid diagonal lines represent constant energy per switching event which is in units of picojoules. In essence the lines represent the fact that the right-hand side of Eq. (1.3) can remain constant at a given energy level and the P and τ_D terms on the left-hand side can be traded off. Also shown on the figure are experimental points for various bipolar circuit technologies and levels of circuit integration ranging from small scale and medium scale integration (SSI and MSI) to large scale integration (LSI) of circuit functions. Although Fig. 1.3 shows values of τ_D and P corresponding to a given state-of-the-art (1976), the general features of the circuitry are discussed briefly.

Transistor–transistor logic (T^2L) is the first bipolar digital integrated circuitry to achieve MSI functional complexity, primarily because it became government standardized in the early 1960s. Figure 1.4*a* shows the basic MSI T^2L NAND gate with the multiple-emitter (Q_1) input transistor and the so-called totem-pole output driver pair (Q_3-Q_4). The name reflects the fact that both input and output devices are transistors as compared with earlier circuit families such as RTL (resistor–transistor) or DTL (diode–transistor) logic circuits. The delays for T^2L are typically less than 10 ns and the power delay product ranges from 10 to 100 pJ. The corresponding power levels reflect the fact that functional densities greater than 100 gates are possible within 1-W package. From the basic circuit one can see that four active devices as well as several passive components are needed. The use of several devices increases power consumption but gives no added functional advantage other than

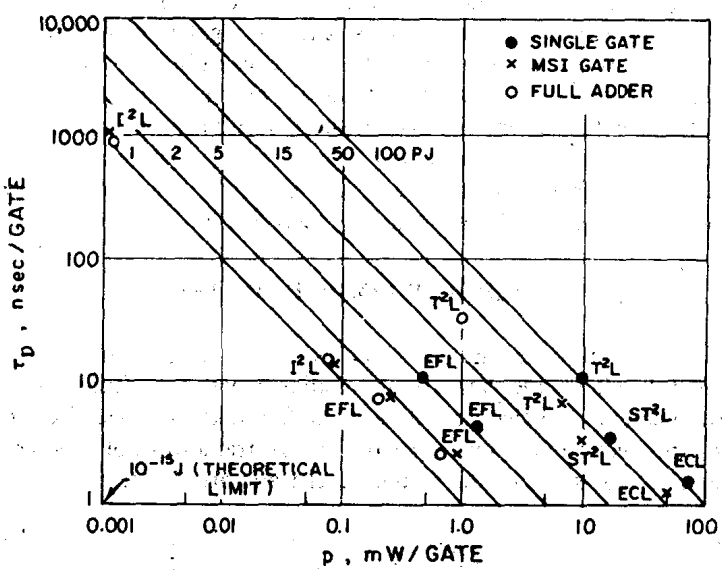

Figure 1.3 Gate delay τ_D versus power per gate for bipolar logic families.

off-chip current driving capability. Subsequent developments of digital IC families in the 1970s have both reduced component count per gate and increased speed.

One set of circuit changes for T^2L came with the development of Schottky clamping and the use of smaller on-chip voltage supplies. For example, the gate component count can be cut in half and a 3-V supply used with a logic swing of 1.5 V. These changes result in more than a tenfold increase in components per chip. Figure 1.4*b* shows one such gate configuration where the conventional Schottky notation is used. Chapters 3 and 4 present further details concerning the device and circuit impact of Schottky clamping. The primary point to be emphasized here is the fact that reduced voltage levels and capacitance both, as reflected through Eq. (1.3), result in a reduced power-delay product.

Integrated injection logic (I^2L) is the successor to T^2L with a substantial potential for LSI applications, both digital and as an analog-compatible circuitry. In the early 1970s the technology was reported with special emphasis on micropower. The measured results shown in Fig. 1.3 clearly reflect this micropower emphasis with power delay products of 1 pJ or less. The power levels of 0.001–0.1 mW reflect the potential for LSI and even VLSI since gate counts could reach 10^6 for a 1-W package. The obvious deficiency is the gate delays, which typically are greater than 5–10 ns. Details of the I^2L circuitry

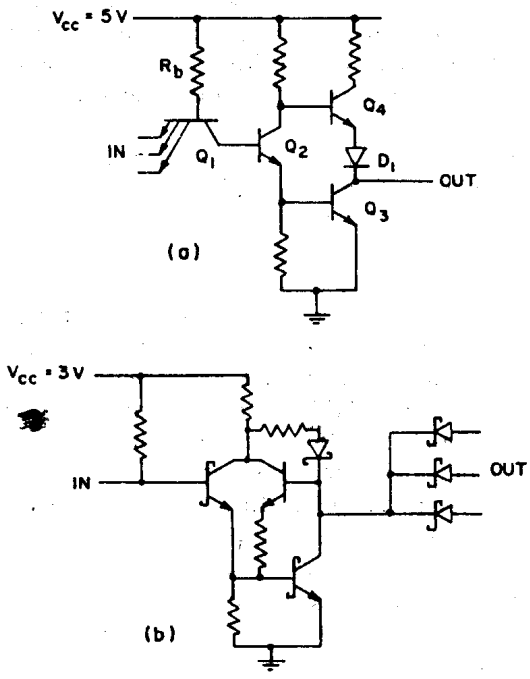

Figure 1.4 (*a*) Basic T^2L NAND gate with totem-pole output stage. (*b*) A Schottky T^2L gate.

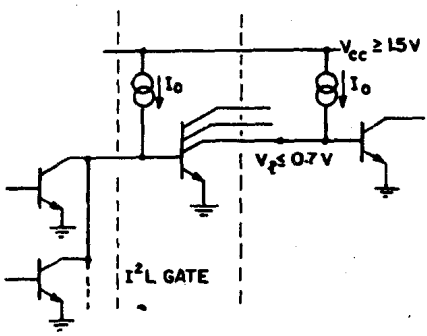

Figure 1.5 Basic I^2L gate with driving gates and a loading gate.

are discussed in Chapter 5. Figure 1.5 shows an overly simplified gate representation, which illustrates two key points:

1. The gate is based on a single transistor-current source pair with multiple inverting outputs and a wired-AND input node.
2. The scaled-down V_{CC} and V_I values make the micropower operation possible, although the ultimate speed limitations as predicted by Eq. (1.3) hinge on the node capacitance values.

Comparing the I^2L and T^2L gates two dominant differences are apparent: a major reduction in both subcircuit components and voltage levels. The I^2L technology has set the stage for further LSI/VLSI innovations in digital bipolar circuits that emphasize both decreased gate size and power consumption.

Emitter-coupled logic (ECL) is the highest-speed bipolar technology that has been developed to date. Compared with the T^2L and I^2L bipolar families, ECL avoids saturated-device effects by limiting the logic swing and steering a fixed current through the load resistors. Figure 1.6 shows the basic NOR/OR gate. The sources I_{EE} and V_{EE} set the power level; resistors R_{C1} and R_{C2} set the logic swing. The circuit shown in Fig. 1.6 is only the skeleton—a schematic—since the voltage reference, current source, and output emitter-followers have been omitted. These extra components, like for T^2L, consume area and power but add only MSI off-chip advantages. As will be shown in Chapter 6, there are on-chip ECL-type circuits such as emitter-function logic (EFL) that offer more logic functions per unit area and per unit source dissipation. Nonetheless, neither ECL/EFL nor T^2L bipolar logic families can match the density advantages of I^2L. Especially for ECL/EFL, the need for both a current source and voltage reference are major drawbacks in terms of density. However, since the objective with ECL/EFL is clearly its speed advantages, the sacrifice in density is most often masked by the dominance of package power limitations.

As a final comment for this section, combinations of the above three families on a single chip are common and have major functional advantages. In