

Semiconductor Manufacturing Technology



Michael Quirk
Julian Serda

SEMICONDUCTOR MANUFACTURING TECHNOLOGY

MICHAEL QUIRK

STAKTEK CORPORATION

JULIAN SERDA

ADVANCED MICRO DEVICES



Upper Saddle River, New Jersey
Columbus, Ohio

Library of Congress Cataloging-in-Publication Data

Quirk, Michael

Semiconductor manufacturing technology / Michael Quirk, Julian Serda.

p. cm.

ISBN 0-13-081520-9

1. Electronic industries. 2. Semiconductors—Design and construction. I. Serda, Julian.

II. Title.

TK7836.Q27 2001

00-041653

Vice President and Publisher: Dave Garza

Editor in Chief: Stephen Helba

Assistant Vice President and Publisher: Charles E. Stewart, Jr.

Production Editor: Alexandrina Benedicto Wolf

Production Coordination: Lithokraft II

Design Coordinator: Robin G. Chukes

Cover Designer: Linda Fares

Cover Image: John Foxx

Production Manager: Matthew Ottenweller

Marketing Manager: Barbara Rose

This book was set in Times Roman by Lithokraft II. It was printed and bound by R. R. Donnelley & Sons Company. The cover was printed by Phoenix Color Corp.

Copyright © 2001 by Prentice-Hall, Inc., Upper Saddle River, New Jersey 07458. All rights reserved. Printed in the United States of America. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permission(s), write to: Rights and Permissions Department.

Prentice
Hall

10 9 8 7 6 5 4 3
ISBN 0-13-081520-9

PREFACE

This text started with a simple premise: as instructors, we need to teach relevant microchip technology to students and employees in semiconductor manufacturing. Unfortunately, in the semiconductor industry, changes in technology are measured in months, not years. Our challenge was to write a relevant book that would not be outdated by the time it was published. With that in mind, we researched the material and applied ourselves to writing the chapters and creating the artwork. Following the aggressive pace of Moore's law, the technical material in our book is at most only 18 to 24 months old. This permits us to keep abreast of the changing technology nodes swirling through the semiconductor industry.

This text is written for students in two-year and four-year technology programs at community colleges and universities. The text will also be a practical reference as well as a standard text in corporate and technical training classes. Students are expected to have an understanding of high school chemistry, physics, and math. Chapters are organized around the broad technologies applicable to semiconductor manufacturing.

ORGANIZATION OF THE TEXT

Our goal is to accomplish three objectives:

1. Help technology students grasp the fundamental technologies used in manufacturing semiconductor devices.
2. Present some of the many challenges in microchip fabrication.
3. Instill in the reader an appreciation of the conceptual simplicity of semiconductor manufacturing.

All fundamental technical information relevant to semiconductor manufacturing is first presented in Chapters 1 to 8. Chapter 9 presents a process model overview with a general flowchart that links the major areas in a wafer fab. Chapters 10 to 19 cover each of the major processes in the fab. Finally, Chapter 20 provides an overview of the back-end process for IC assembly and packaging. The content in the process chapters (Chapters 10 to 20) addresses critical process technology, followed by the various equipment designs needed to support this technology. Each process chapter concludes with a summary of quality measures and troubleshooting issues to familiarize the student with the practical, day-to-day challenges encountered during wafer fabrication.

The latest technologies for sub-0.25 μm processing are covered in detail. This includes chemical mechanical planarization (CMP), shallow trench isolation (STI), chemically amplified deep UV photoresists, step-and-scan systems, copper metallization with dual damascene, and the widespread move to process integration with cluster tools. Throughout the text, we explain all process and equipment technology in light of the long history of change in the industry. Early tools and processes are described to clarify the development of current technology. In some cases, the linkage between the latest equipment and earlier tools is obvious, while in other instances the change is dramatic.

Professors, students, and other readers of this book can send comments or questions about this text to the authors at the following website: <http://www.smtbook.com>. We look forward to any exchange of information that can help advance semiconductor manufacturing education.



Vacuum Tubes

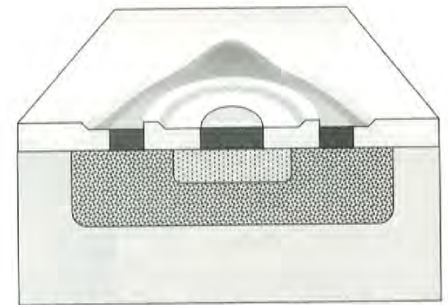


FIGURE 1.2 The First Planar Transistor

now known as Silicon Valley. The first commercial planar transistor was fabricated at Fairchild Semiconductor in Palo Alto, California, in 1957. It had a layer of aluminum interconnect material that was deposited on top of the silicon wafer to connect the different parts of the transistor (see Figure 1.2). A natural oxide layer thermally grown from the silicon was used to insulate the aluminum conductor. The use of layers was an important development in semiconductors and is the reason for the term *planar technology*.

CIRCUIT INTEGRATION

The semiconductor name derives from the material's ability to sometimes be a conductor to electricity and other times act as a nonconductor to electricity. The earliest semiconductor material was germanium, built into a single chip with one function (referred to as a discrete component). Nowadays, greater than 85% of all microchips are made from silicon semiconductor material. For this reason, we will emphasize silicon in this book.

An important step forward in the semiconductor industry was the integration of multiple electronic components on one silicon substrate. Referred to as an *integrated circuit*, or *IC*, it was coined independently by Robert Noyce at Fairchild Semiconductor and Jack Kilby at Texas Instruments in 1959. On the silicon surface of an IC are manufactured the many different semiconductor *devices*, such as transistors, diodes, resistors, and capacitors, which are connected into a circuit to define how the chip will function. The expression *IC* is often meant to describe the chip and all its components.

The first IC, developed in July 1958 by Jack Kilby at a Texas Instruments facility in Dallas, Texas, was made using a slice of germanium semiconductor material as the substrate.⁶ The invention combined a transistor and other components on the germanium while also using the natural resistance of the germanium as a resistor. The devices were connected with individual wires.

Robert Noyce at Fairchild Semiconductor also invented the concept of an IC by extending the idea of how to interconnect the different components on the planar silicon material. His idea was to use an aluminum metal conductor on the silicon surface to interconnect the different transistors while using a layer of oxide grown from the silicon as an insulator to separate the metal conductor from the silicon devices. This was the first practical structure of an IC as a single-structure silicon chip.⁷

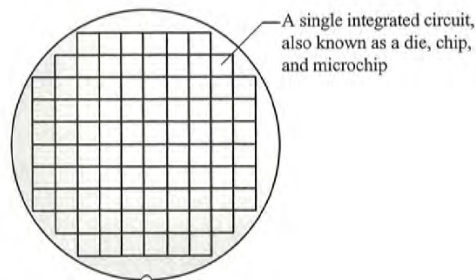
Circuit integration has increased dramatically since the initial IC. Because all the components are integrated on the silicon substrate, the IC has developed into a cost-effective and reliable way to produce and interconnect many components. The ability to integrate many different components on an IC has spurred engineers to design ever more complex electronic circuits to meet new customer needs.



The First Transistor from Bell Labs
(Photo courtesy of Lucent Technologies, Bell Labs Innovations)



Jack Kilby's First Integrated Circuit
(Photo courtesy of Texas Instruments, Inc.)



A single integrated circuit,
also known as a die, chip,
and microchip

FIGURE 1.3 Top View of Wafer with Chips

Integration Eras

We can roughly categorize the integration era by the number of components integrated on a chip. This is a useful method of organizing the development of the semiconductor industry since 1960 to the present (see Table 1.1).

TABLE 1.1 Circuit Integration of Semiconductors

Circuit Integration	Semiconductor Industry Time Period	Number of Components per Chip
No integration (discrete components)	Prior to 1960	1
Small scale integration (SSI)	Early 1960s	2 to 50
Medium scale integration (MSI)	1960s to Early 1970s	50 to 5,000
Large scale integration (LSI)	Early 1970s to Late 1970s	5,000 to 100,000
Very large scale integration (VLSI)	Late 1970s to Late 1980s	100,000 to 1,000,000
Ultra large scale integration (ULSI)	1990s to present	> 1,000,000

Circuit integration continues today as the number of devices on a chip is continually increased. A significant challenge for circuit integration is the capability of the semiconductor manufacturing process to improve the fabrication technology to produce highly integrated ULSI chips at acceptable costs. To this end, the semiconductor industry has become highly standardized, with most manufacturers using similar manufacturing process and equipment technology. The key to market success is the ability of a company to deliver the right product at the right time. Industry

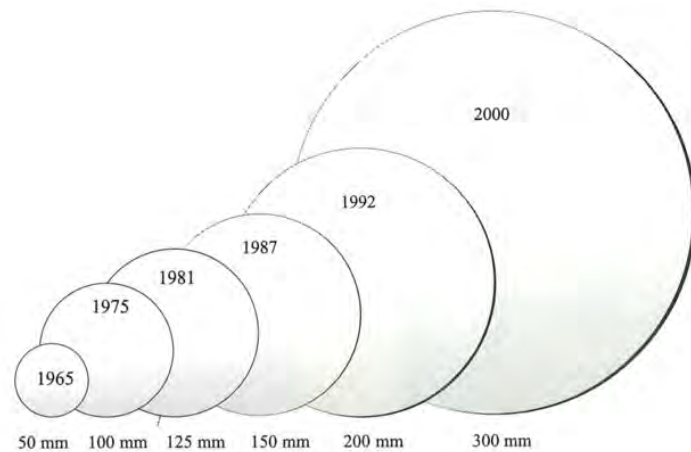


FIGURE 1.4 Evolution of Wafer Size

consortiums, such as International SEMATECH and the Semiconductor Industry Association (SIA), have been created to assist chip producers to be competitive in the world market of high-performance ICs.

IC FABRICATION

Tens or hundreds of identical chips are fabricated simultaneously on a silicon *wafer* (see Figure 1.3 on page 5). The number of chips on a wafer varies depending on the type of product and the size of each chip. The chip size changes depending on the level of integration on the chip.

The chip is also referred to as a *die* (singular and plural for chips or ICs), while the silicon wafer is often called the *substrate*. The wafer diameter has evolved over the years from an initial diameter of less than 1" to the present common diameter of 8" (200 mm), which is undergoing a change to 12", or 300 mm (see Figure 1.4). The cost of fabricating ICs drops dramatically if there are more chips on a wafer due to benefits from the economies of scale (more chips are produced from the same effort).

Fabrication of the semiconductor devices occurs only in the first few microns of the silicon near the wafer surface. The bulk of the silicon wafer thickness is to give the wafer adequate rigidity during processing. Once the devices are fabricated in the silicon, then layers of metal circuitry are placed on the silicon to interconnect between the devices and the various electronic signals outside the chip (see Figure 1.5). The concept and materials for interconnecting a modern day IC are remarkably similar to the initial planar transistor first commercialized at Fairchild Semiconductor in 1957. The major difference is that today's chips are much more complex.

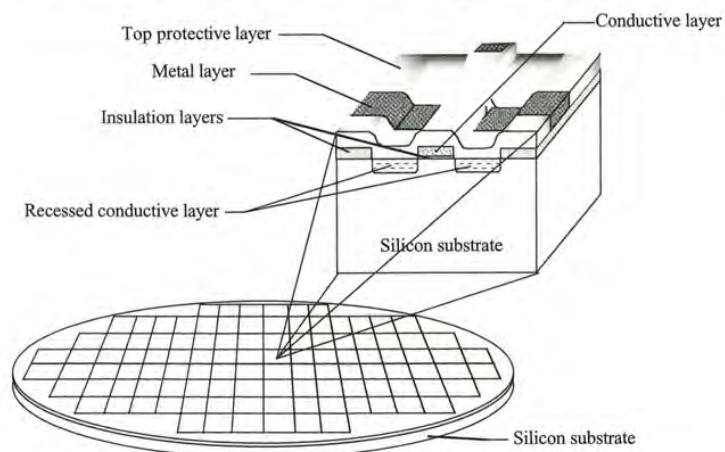


FIGURE 1.5 Devices and Layers from a Silicon Chip

Wafer Fab

Early wafer fabs were simple, with operators manually processing wafers through the different operations. An essential aspect of wafer fabs is that as wafers become more densely integrated, the permissible level of contamination decreases dramatically. Contamination that can damage wafers and cause them not to function properly evolves from many sources: humans, materials, water, air, and equipment. Modern wafer fabs have become specialized facilities that provide the clean manufacturing environment and specialized equipment to produce wafers with minimal defects from contamination. This includes limited human exposure, ultrapure chemicals and utilities, and special wafer handling procedures needed to fabricate ICs in the ULSI era.

A wafer typically requires two to three months of processing to complete the 450 or more process steps in a wafer fab. At the end of the fabrication process, the individual chips are separated from the entire wafer and then prepared for packaging into the final product.

Stages of IC Fabrication

Fabrication of microchips involves five general stages of manufacturing (see Figure 1.6):

- ◆ Wafer preparation
- ◆ Wafer fabrication
- ◆ Wafer test/sort
- ◆ Assembly and packaging
- ◆ Final test

These five stages are interdependent, with a large infrastructure within the semiconductor company and a support network of industries that provide specialized chemicals and equipment. Companies that operate only in an individual stage (such as a chip company that only fabricates chips) have to meet industrywide standards to ensure that the final microchip meets performance objectives.

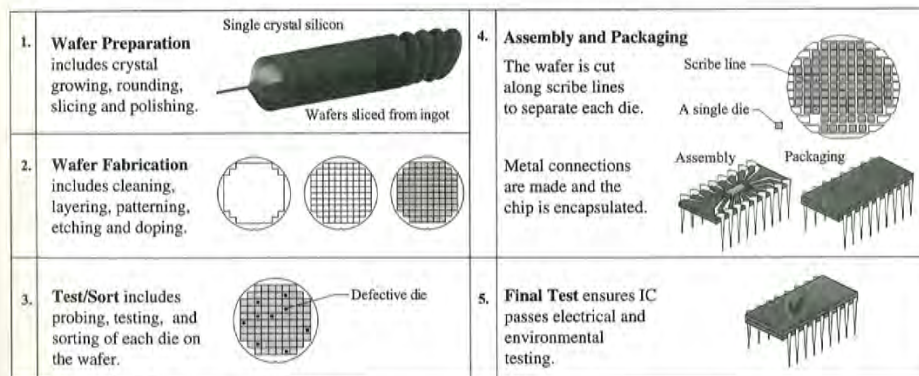


FIGURE 1.6 Stages of IC Fabrication

Wafer Preparation ■ In the first stage, *silicon* is mined from sand and purified. It undergoes specialized processing to create a silicon ingot with the appropriate diameter (see Figure 1.7 on page 8). The silicon ingot is then sliced into thin wafers that are used to fabricate the microchips. Wafers are prepared to special specifications for criteria such as flatness and contamination.

Since the 1980s, most companies that fabricate microchips purchase their wafers from suppliers who specialize in crystal growth and wafer preparation. The industry also produces wafers made from germanium or compound semiconductor materials. These are specialized applications, with the majority of all semiconductor wafers made from silicon.

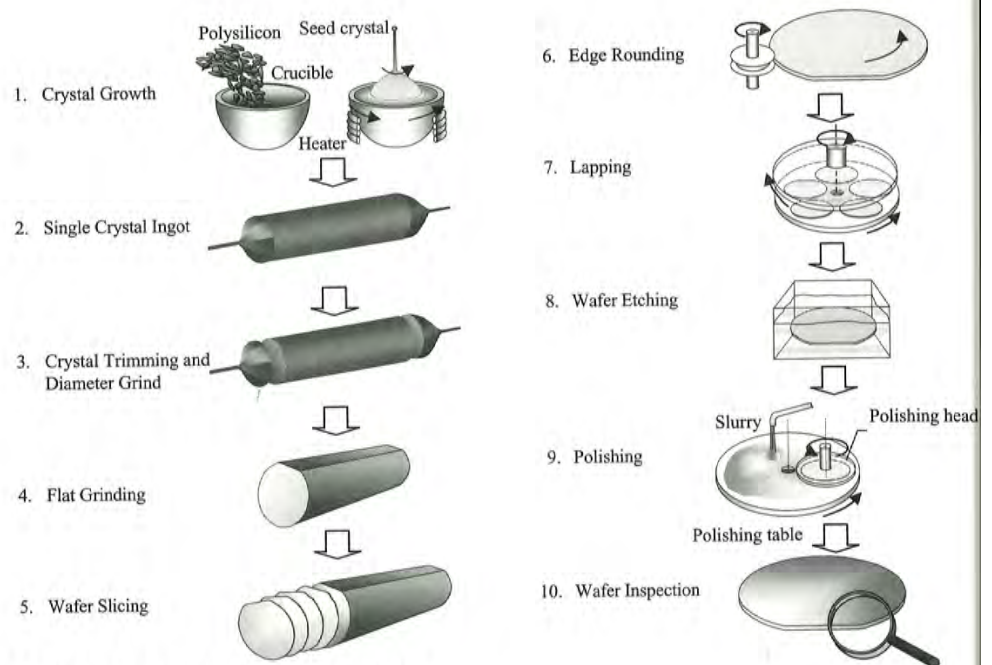
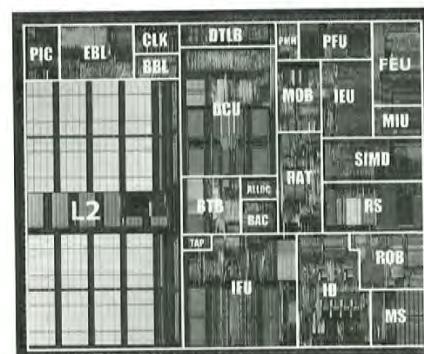


FIGURE 1.7 Preparation of Silicon Wafers
(Note: Terms in Figure 1.7 are explained in Chapter 4.)

Wafer Fabrication ■ The fabrication of microchips from wafers is the second stage, referred to as *wafer fabrication*. The bare silicon wafers arrive in the wafer fab, then they go through a variety of cleaning, layering, patterning, etching, and doping steps. The completed wafers have a full complement of integrated circuits permanently etched into each silicon wafer. Other names for wafer fabrication are *microchip fabrication*, *chip fabrication*, and *chip fab*.

Companies that fabricate chips are both merchant and captive producers. *Merchant chip suppliers* fabricate chips to sell on the open market, such as a chip manufacturer that produces memory chips for customers. *Captive chip producers* fabricate chips to be used in-house in their company's products. An example of a captive producer is a company that makes computers and also fabricates chips for their computers. Some chip makers fabricate chips for in-house use and also sell chips on the open market, while others will fabricate specialty chips and purchase other chips on the open market.

Another type of chip maker is the *fabless company*. This type of company designs chips for a particular market, such as a video microchip, while another chip maker fabricates the chips. Finally,



ULSI Chip
(Photo courtesy of Intel Corporation)

another type of semiconductor manufacturer, the *foundry*, produces chips only for other companies. Semiconductor foundries have become more common since the 1980s, with about 10% of all chips now fabricated at foundries. A major reason for the increase in fabless chip producers and foundries is the high cost of constructing and maintaining a wafer fab. Currently, a wafer fab for high-performance ICs costs around \$1.5 to \$3.0 billion, with about 75% of the total cost for equipment.

Wafer fabrication of ICs involves the interaction of many complicated process steps using automated equipment to produce the hundreds of millions of devices on one ULSI wafer. Because of the complexity associated with fabricating high-performance ICs, the semiconductor industry has always been at the leading edge of equipment design and manufacturing technology. This innovation has fostered continual improvements in wafer fabrication.

Wafer Test/Sort ■ At the completion of wafer fabrication, the wafer is then sent to the test/sort area where individual die are probed and electrically tested. Defective die are marked for sorting later into acceptable and unacceptable chips. Die that fail the wafer test will not be sent to customers, while die that pass the wafer test will continue through the process.

Assembly and Packaging ■ After wafer test/sort, the wafer enters assembly and packaging to enclose the individual chips in a protective package. The backside of the wafer undergoes a backgrind to reduce the thickness of the substrate. A thick membrane of acetate tape is attached to the backside of each wafer, then the front of the wafer is sliced along the scribe lines with a diamond-tip saw blade to separate the die from each wafer. The sticky acetate holds the silicon chips to keep them from falling off. At the assembly plant, the good die have wire bonds added or have bumps applied to form the attachment to the assembly package. Soon after, the die is hermetically sealed in a plastic or ceramic package. The actual form of the final package varies depending on the type of chip and its application (see Figure 1.8 on page 10). See Chapter 20 for more information on assembly and packaging.

Final Test ■ To ensure a chip's functionality, each packaged integrated circuit is tested to meet the manufacturer's electrical and environmental specifications. After the final test, the chips are delivered to the customer for assembly into the specific application, such as mounting memory components on a circuit board for a personal computer.

SEMICONDUCTOR TRENDS

The rapid technology changes needed to design and fabricate ICs with ULSI integration leads to the continual introduction of new equipment and processes. The semiconductor industry introduces new fabrication technology into the wafer fab every 18 to 24 months. Wafer fabrication technology changes are driven by the customers' needs. Customers require faster, more reliable, and lower cost chips. To achieve this, chip manufacturers have learned to reduce the size of components on a chip. This increases chip speed and reduces power consumption. Chips undergo extensive testing and analysis to verify long-term reliability. The number of chips on a wafer are increased to reduce cost.

There are three major trends associated with improvements in microchip technology:

- ◆ Increase in chip performance
- ◆ Increase in chip reliability
- ◆ Reduction in chip cost

As the different process areas of wafer fabrication are studied in this text, the introduction of new technologies will be discussed based on their contribution toward these trends.

Increase in Chip Performance

The performance of semiconductor microchips has dramatically improved since the SSI era of the early 1960s. A common way to judge chip performance is speed. *Chip speed* will improve if devices are made smaller and packed closer together on the chip because the electrical signal moving through the circuit has less distance to travel. Another way to increase speed is to use materials that

CHAPTER 12

METALLIZATION

Metallization in wafer fabrication is the process of depositing metal film over a dielectric film and later patterning it to form the interconnecting metal lines and plugs of integrated circuits. This is similar to using insulated copper wire in an automobile to interconnect all the electrical components to make a fully-functional electrical system. The metal lines are sandwiched between dielectric layers for electrical integrity. High-performance microprocessors use metal lines to interconnect tens of millions of devices on one chip. Transistor densities are projected to reach 1 billion transistors per chip by the year 2010, with a corresponding increase in interconnect complexity.

Microchip interconnect technology has become a critical challenge for future IC performance due to the need to decrease signal propagation delay. Because the density of ULSI circuit elements is increasing, interconnect resistance and parasitic capacitance increase, thereby slowing the signal propagation.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Explain the terminology for metallization.
2. List and describe the six categories of metals used in wafer fabrication. Discuss the performance requirements and give applications for each metal category.
3. Explain the benefits for using copper metallization in wafer fabrication. Describe the challenges for implementing copper.
4. State the advantages and disadvantages to sputtering.
5. Describe the physics of sputtering and discuss different sputtering tools and applications.
6. Describe the benefits and applications for metal CVD.
7. Explain the fundamentals of copper electroplating.
8. Describe a process flow for dual-damascene processing.

A significant change currently underway in wafer fabrication technology is the reduction in interconnect metal resistivity, ρ . This reduction is achieved by the replacement of aluminum alloy with copper as the primary conducting metal. With deep submicron linewidths, there is also a need for a low- k interlayer dielectric (ILD). Lowering the dielectric constant will reduce parasitic capacitance, which contributes to signal delay.

Traditionally, the methods used to deposit metals on a wafer have been physical processes. On the other hand, processes used to deposit insulating and semiconducting layers usually involved the CVD chemical reactions studied in Chapter 11. This separation into physical and chemical processes is less defined with the introduction of new IC metallization technology.

TYPES OF METALS

The types of metals and metal alloys used for chip interconnects are evolving based on the performance requirements. Requirements for a successful metal material are:^{3, 4}

1. **Conductivity:** Must be highly conductive and capable of handling high current densities while maintaining electrical integrity.
2. **Adhesion:** Ability to adhere to the underlying substrate and easily connect to external connections. Low contact resistance to semiconductor and metal interfaces.
3. **Deposition:** Readily deposited with a uniform structure and composition (for alloys) by a relatively low-temperature process. Deposition into high-aspect ratio gaps for damascene metallization technique.
4. **Patterning/Planarization:** High-resolution patterning for traditional Al-based metallization without etching the underlying dielectric. Ease of planarization for damascene metallization.
5. **Reliability:** The metal is relatively soft and ductile in order to withstand cyclic temperature variations during processing and service.
6. **Corrosion:** High resistance to corrosion with minimal chemical interactions with the adjacent layers and underlying device regions.
7. **Stress:** Resistance to mechanical stress to reduce wafer distortion and material failures for reasons such as cracking, void formation, and stressed-induced corrosion.

The melting temperature and resistivities of different metals found in wafer fabrication and silicon are shown in Table 12.1.⁵

TABLE 12.1 Silicon and Select Wafer Fab Metals (at 20°C)

Material	Melting Temperature (°C)	Resistivity ($\mu\Omega\text{-cm}$)
Silicon (Si)	1412	$\approx 10^9$
Doped Polysilicon (Doped Poly)	1412	≈ 500 to 525
Aluminum (Al)	660	2.65
Copper (Cu)	1083	1.678
Tungsten (W)	3417	8
Titanium (Ti)	1670	60
Tantalum (Ta)	2996	13 to 16
Molybdenum (Mo)	2620	5
Platinum (Pt)	1772	10

The various metals and metal alloys used in wafer fabrication can be grouped into the following categories:

- ◆ Aluminum
- ◆ Aluminum-copper alloys
- ◆ Copper
- ◆ Barrier metals
- ◆ Silicides
- ◆ Metal plugs

Aluminum

The earliest interconnect metal in semiconductor fabrication was aluminum, and it is still the most common interconnect metal in wafer fabrication. Aluminum is projected to continue as the dominant fab interconnect metal for several more years.⁶ Copper interconnect metal is expected to replace aluminum for high-performance IC fabrication during the early 2000s. Nevertheless, it is beneficial to review the background into the selection of aluminum for metallization because many of the same basic technical challenges exist today.

Aluminum-Copper Alloys

Aluminum metal was selected as the primary interconnect material for ICs because of its low resistivity and its compatibility with silicon and the wafer fabrication process. However, aluminum suffers from a reliability problem known as *electromigration*. This is the movement of aluminum atoms in the conductor due to momentum transfer from the electrons carrying the current.⁹ Under high current-density conditions, electrons collide with aluminum atoms, causing the atoms to gradually move. This movement of atoms leads to a depletion of atoms at the negative end of the conductor. In conductor regions where depletion occurs, this action leads to voids, thinning of lines, and a potential open circuit. In other regions of the conductor that have an accumulation of metal atoms, metal atoms pile up and form hillocks (see Figure 12.6).¹⁰ *Hillocks* are protrusions on the surface of metal films due to electromigration. If excessive or large hillocks form, adjacent lines or lines on two levels can short together. With advanced circuitry designs in ULSI technology, there is increased current density and chip temperature, both of which make the aluminum chip metallization more prone to electromigration.

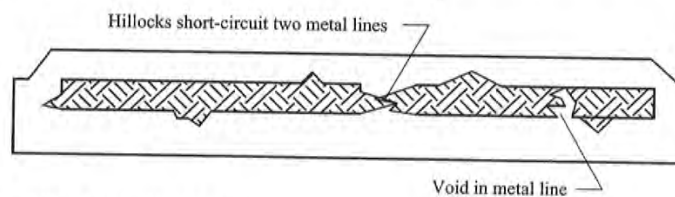


FIGURE 12.6 Hillock on a Metal Line

Electromigration in interconnects is controlled by alloying the aluminum with between 0.5 to 4% copper. This alloying essentially increases the conductor's current-carrying capacity by reducing the grain boundary diffusion effects in the aluminum. At the same time, it has been shown that values of copper in aluminum greater than 8% will actually increase electromigration.¹¹ The exact reason why the addition of copper reduces the chance of electromigration failure in Al-Cu alloys is not well understood.¹² If copper is alloyed with aluminum, more care must be given to plasma etching this alloyed aluminum (etching is discussed in Chapter 16). Copper is difficult to etch, and any residual copper remaining after etching the aluminum alloy interconnect can promote corrosion.¹³ It is also possible to have electromigration in contacts and vias. Electromigration failure in contacts is resolved by the use of barrier metallization.

Electromigration is probably one of the most widely studied failure mechanisms of integrated circuits. Studies of this failure began in the 1950s with the development of the semiconductor and intensified in the late 1960s with the discovery of crack formation in the aluminum conductor leading to failures.¹⁴ The significance of electromigration is that it usually occurs after the chip has been in service for an amount of time, which means that failure occurs catastrophically during customer usage.

Copper

A major transformation is underway in IC design and fabrication by means of the introduction of *copper interconnect* technology. Copper interconnects will replace aluminum metallization with copper to achieve significant benefits in chip performance. A legitimate question that someone may pose is why introduce copper if aluminum has performed so admirably over the years.

Need for Copper ■ The benefits for introducing copper for IC interconnect metallization are:¹⁵

1. **Reduction in resistivity.** The interconnect wiring resistivity reduces from 2.65 $\mu\Omega\text{-cm}$ for aluminum to 1.678 $\mu\Omega\text{-cm}$ at 20°C for copper, reducing RC signal delay and increasing chip speed.
2. **Reduction in power consumption.** Narrower lines consume less power.

3. **Tighter packing density.** Narrower lines permit tighter circuitry packing, which means that fewer levels of metal are needed.
4. **Superior resistance to electromigration.** Copper does not have a concern for electromigration.
5. **Fewer process steps.** Potential for 20 to 30% fewer processing steps with damascene processing of copper.

As wafer fabrication design rules reduce to 0.15- μm linewidths and below, the increased packing density of devices on the chip permits more electrical signal speed from device to device (the transistors are closer; therefore, the signal has less distance to travel). This density leads to improved chip performance. However, this improved chip performance is only possible if the interconnect system between the devices is optimized. Narrower linewidths lead to increased line resistance. Tightly spaced conductor lines with a dielectric material between them act as capacitors, leading to a degradation in performance from an increase in resistance (R) and capacitance (C). This condition is the signal delay, or interconnect delay discussed in Chapter 11. If either or both of these two parameters are reduced, then the signal delay reduces, leading to increased chip performance.

One possible method reducing interconnect resistance is to increase the conductor cross section. However, this contradicts the goal of increased packing density since wider conductors will take substantially more space. Smaller IC feature sizes are not achievable with larger linewidths. This is why the semiconductor industry is placing emphasis on copper—to lower the interconnect resistance, which lowers R and therefore the overall signal delay. The optimum improvement to RC signal delay is gained when resistance (R) is reduced and the capacitance (C) is lowered by using a low- k dielectric along with thinner barrier metals (see Table 12.2).¹⁶

TABLE 12.2 Change in Interconnect Delay Compared to 0.25- μm Device Generation

Technology	0.25 μm	0.18 μm	0.13 μm
Conventional Interconnect Technology: • Al/Cu interconnect alloy and TiN barrier metal	0	+21%	+93%
New Technology Introduced by Generation: • Reduced barrier thickness • Low- k (3.0) dielectric • Dual damascene Cu interconnect and plugs	-10%	-27%	-16%

Another benefit from the implementation of copper is that smaller linewidths can carry the same amount of current, permitting a tighter packing density on each metal level. This condition produces fewer overall levels of metal on a chip, leading to significantly reduced manufacturing costs.¹⁷

Copper is a relatively soft metal. It has superior resistance to electromigration, a common reliability problem with aluminum.¹⁸ This means that chips fabricated with copper can handle higher electrical power densities, which permits the development of new product applications. Table 12.3 compares Cu and Al for different properties and processes common in wafer fabrication.¹⁹

TABLE 12.3 Comparison of Properties/Processes Between Al and Cu

Property/Process	Al	Cu
Resistivity ($\mu\Omega\text{-cm}$)	2.65 (3.2 for Al—0.5% Cu)	1.678
Electromigration resistance	Low	High
Corrosion resistance (in air)	High	Low
Etch processing	Yes	No
CMP (chemical mechanical planarization) processing	Yes	Yes

Copper Challenges ■ There are three major challenges involved with using copper as semiconductor interconnects that distinguish it from traditional aluminum interconnects. These three challenges differ significantly from aluminum technology and must be resolved before the implementation of copper for IC interconnects.²⁰

1. Copper diffuses quickly into oxides and silicon. This is a concern if copper diffuses into the active region of the silicon (i.e., source/drain/gate region of the transistor) because it will damage the device by creating junction or oxide leakage.
2. Copper cannot be easily patterned using regular plasma etching techniques (see Chapter 16 for a discussion of etching). Copper dry etching does not produce a volatile by-product during the chemical reaction that is necessary for economical dry etching.
3. Copper oxidizes quickly in air at low temperatures ($<200^{\circ}\text{C}$) and does not form a protective layer to stop further oxidation.

These challenges are addressed by converting to dual damascene processing with special barrier metals optimized for copper. Damascene processing eliminates the need to etch copper. Furthermore, tungsten plugs are expected to be used as the first level of metal to contact the source, drain, and gate regions.²¹ This use of tungsten overcomes the concern of copper contaminating the silicon (referred to as copper poisoning). The tungsten may even be patterned for wiring of the local interconnect (LI). All other metal lines and vias for the multiple metal layers are expected to be copper.

Another major hindrance to converting to copper in semiconductor manufacturing is the natural reluctance to introducing new materials into production. New materials means there are new sources of contamination, new equipment and procedures, unpredictable results, and so on. Change always brings unforeseen problems that pose increased risks to the semiconductor manufacturer. However, these risks are balanced by the benefits from changing the interconnect material in order to improve IC performance.

Barrier Metals

A more effective way to make reliable ohmic contacts to shallow junctions without material diffusion or problems such as junction spiking is to use barrier metallization. A *barrier metal* is a thin layer of deposited metal or metals that is designed to prevent intermixing of the materials above and below the barrier (see Figure 12.7). The thickness of barrier metals has typically been around 100 nm for features sizes in the 0.25- μm generation, which is substantially thinner than the 400 to 600 nm thickness required for 0.35- μm generation of devices. Thickness of barrier metals is projected to decrease to 23 nm or less for the 0.18- μm technology node and beyond.²²

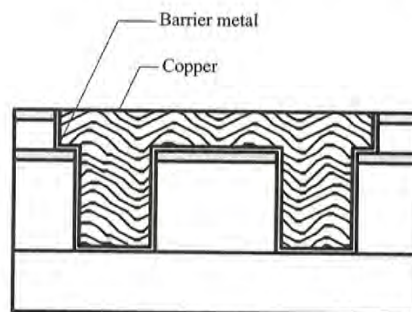


FIGURE 12.7 Barrier Layer for Copper Interconnect Structure

Barrier metals are widely used in semiconductor manufacturing. For contacts between the tungsten (W) plug that connects the aluminum interconnect wiring and the silicon source/drains, the barrier metal prevents the silicon and W from coming into contact with each other. This barrier prevents diffusion between the W and silicon and any junction spiking.

The essential properties of an acceptable barrier layer metal are:

1. Good diffusion barrier properties so that the diffusivity of the two interface materials (e.g., tungsten and silicon) is low at the sintering temperature (*sintering* refers to joining the materials by thermal means).
2. High electrical conductivity with low ohmic contact resistance.
3. Good adhesion between the semiconductor and the metal.

interconnect technology for an ILD-2 and metal-2 (one complete level) are described in Steps 11 and 12 of the process flow described in Chapter 9. This traditional process can be viewed as an *aluminum subtractive process* because the aluminum is deposited as a blanket film and then etched away (subtracted) to form the circuitry. The critical step that defines the width and spacing of the aluminum interconnect lines is metal etch, while SiO_2 is deposited into the narrow gaps between the lines.

New materials and processes are being rapidly introduced into advanced wafer fabs to support the 0.18- μm technology node with traditional aluminum alloy metallization. This is a transitional period until copper metallization is eventually implemented. As a rule, most fabs will change only one major microchip material at a time (e.g., metal or insulator) to minimize the inherent risk to production. With these new fabrication materials and their stringent processing requirements, the general trend is for increased use of high-density plasma CVD (HDPCVD) for dielectrics, a shift from sputtered metals to metal CVD, and increased use of chemical mechanical planarization (CMP). The goals are to increase product performance and manufacturing productivity.

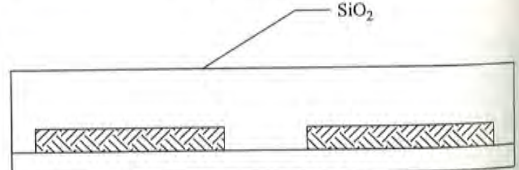
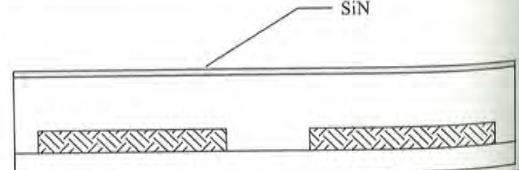
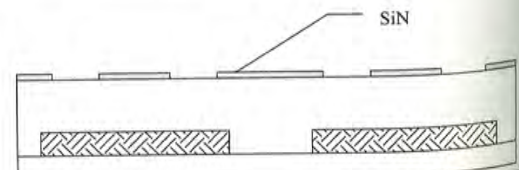
Copper Damascene Structure

The semiconductor industry is implementing copper as the interconnect material for microchip fabrication. The traditional process flow will not be used for copper metallurgy because copper is not suitable for dry-etching. Copper also has special requirements to minimize its diffusion into silicon. To form copper interconnect wiring, a dual-damascene process is used to avoid copper etching. In the damascene process, it is no longer metal etch that defines the critical line width and spacing, but rather dielectric etch.

Dual Damascene Process ■ Since copper is difficult to etch, early developers of copper technology were forced to consider alternative approaches to creating metal lines. The *dual-damascene process* became the consensus process for copper metallurgy. It creates both the vias and lines for each metal layer by etching holes and trenches in the ILD, depositing copper in the etched features and using chemical mechanical planarization to remove the excess copper (referred to as *copper overburden*).

There are many possible scenarios for a dual-damascene process flow. Table 12.5 explains a basic process flow with the essential technologies that may or may not include all the steps performed by a particular manufacturer.

TABLE 12.5* Copper Metallization using Dual Damascene

Process Step	Description	Structure
1. SiO_2 deposition	ILD oxide deposition with PECVD to desired thickness for via. There is no critical gap fill; therefore, PECVD is acceptable.	
2. SiN etch stop deposition	Thin (250 Å) SiN etch stop is deposited on ILD oxide. The SiN needs to be dense and pinhole-free; therefore, HDPCVD is used.	
3. Via patterning and etch	Photolithography to pattern and dry etch via openings into SiN. Strip photoresist after completion of etch.	

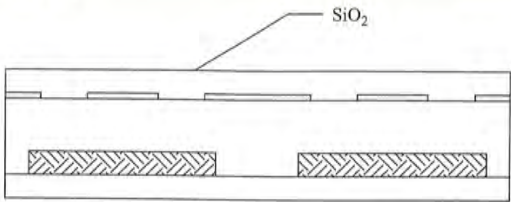
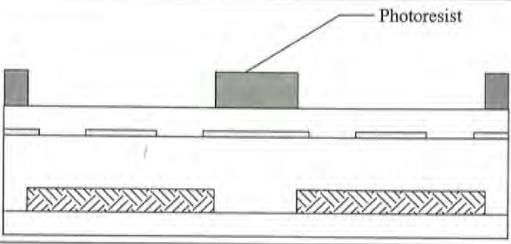
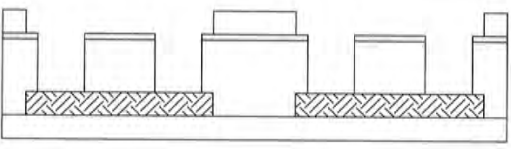
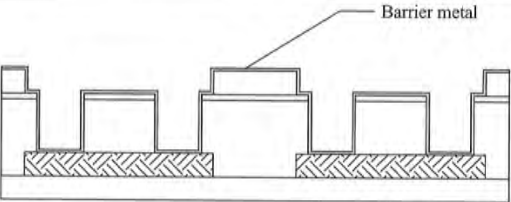
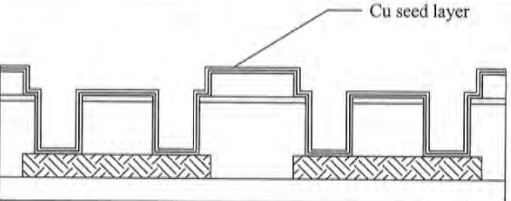
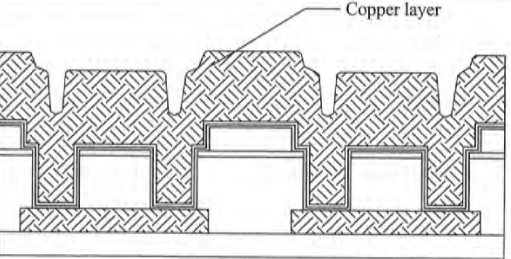
Steps 11
s an *alu-*
n etched
ig of the
between

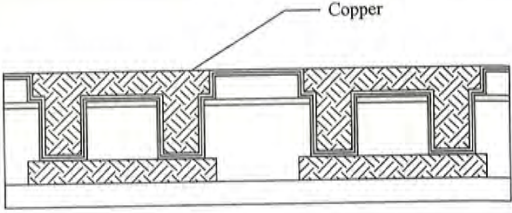
s to sup-
a transi-
ll change
rent risk
ents, the
s, a shift
arization

microchip
copper is
ision into
d copper
width and

per tech-
amascene
l lines for
d features
as *copper*

explains a
steps per-

Process Step	Description	Structure
4. Deposit remaining SiO_2	PECVD oxide deposition for remaining ILD oxide.	 A cross-sectional diagram of a semiconductor structure. It shows a base layer with a patterned layer on top. A new layer of SiO_2 is being deposited over the patterned layer, filling the gaps between the features.
5. Interconnect patterning	Photolithography to pattern SiO_2 trench with resist. Previously patterned via openings are located in trench.	 A cross-sectional diagram showing a photoresist layer being patterned into a trench. The trench is filled with a material, and the photoresist is being removed from the top surface.
6. Etch trench for interconnect and hole for via	Dry etch trench in ILD oxide, stopping on the SiN layer. Etch continues to form via opening by passing through opening in SiN.	 A cross-sectional diagram showing the etching of a trench and a via hole. The trench is being etched into the ILD oxide layer, and the via hole is being etched through the SiN layer.
7. Deposit barrier metal	Deposit Ta or TaN diffusion layer with ionized PVD on bottom and sidewalls of trench and via.	 A cross-sectional diagram showing the deposition of a barrier metal layer. The barrier metal is being deposited on the bottom and sidewalls of the trench and via hole.
8. Deposit Cu seed layer	Deposit continuous Cu seed layer with CVD. The layer must be uniform and free of pinholes.	 A cross-sectional diagram showing the deposition of a Cu seed layer. The seed layer is being deposited on the bottom and sidewalls of the trench and via hole.
9. Deposit Cu fill	Deposit Cu fill with electrochemical deposition (ECD). Fill both via opening and trench.	 A cross-sectional diagram showing the electrochemical deposition of Cu fill. The Cu fill is being deposited into the trench and via hole, filling the space.

Process Step	Description	Structure
10. Remove excess Cu with CMP	Remove excess Cu using chemical mechanical planarization (CMP). This planarizes the surface and prepares for the next level. The resulting surface is a planar structure with metal inlays in the dielectric to form the circuitry.	

* P. Singer, "Making the Move to Dual Damascene Processing," *Semiconductor International* (August 1997): p. 79.

For dual-damascene processing, a two-tiered metal inlay is created that includes a series of holes (for contacts or vias) in addition to the interconnect trench for metal lines. The vias and metal line levels undergo copper fill concurrently, which saves a process step and eliminates the interface between the via and metal line.

There are different approaches to performing dual-damascene processing, each with their own set of challenges. Most of the manufacturing methods integrate the use of etch-stop materials to control how deep etching occurs for the via and trench. Etch stop is typically done by using a barrier metal such as silicon nitride (Si_3N_4 or Si_xN_y) as a hard mask at the bottom of the via or trench. Si_3N_4 etches at a much slower rate than the dielectric material, thus effectively stopping the etch process. Some approaches may even have two separate Si_3N_4 layers, which adds process complexity. In addition, Si_3N_4 has a high- k and increases the interlayer capacitance of the interconnect stack, which is why the Si_3N_4 layer needs to be thin.

The dual-damascene flow shown above has the via etch done in the same etch step as the trench, with the SiN etch stop to define the bottom of the trench. In other variations of dual-damascene, the via is etched first, then the trench is patterned and then etched. Trench structures created over topography from a previous via etch pose a significant challenge for photolithography.

The most important reason for using a damascene approach to copper metallization is to avoid metal etch. A second advantage to damascene processing is that there is no further need for dielectric gap fill between etched metal lines, since the dielectric is applied as a blanket and then etched.

The dual-damascene process in wafer fabrication potentially uses 20 to 30% fewer steps than traditional aluminum metal interconnect. Not only are there fewer manufacturing steps, but dual damascene also eliminates or reduces some of the most difficult steps of traditional aluminum metallization, including aluminum etch and many of the tungsten and dielectric chemical-mechanical polish steps.⁵⁴ Reducing process steps in wafer fabrication is important for improving the yield of the process, since fewer process steps translate to fewer sources of error that reduce assembly yield.

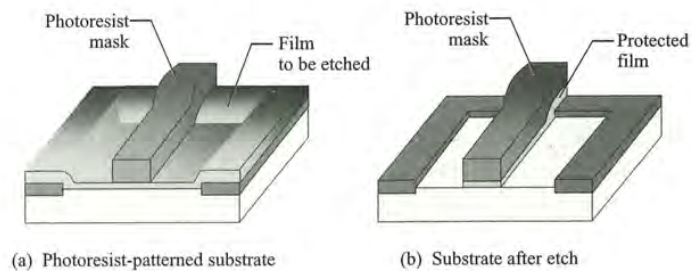
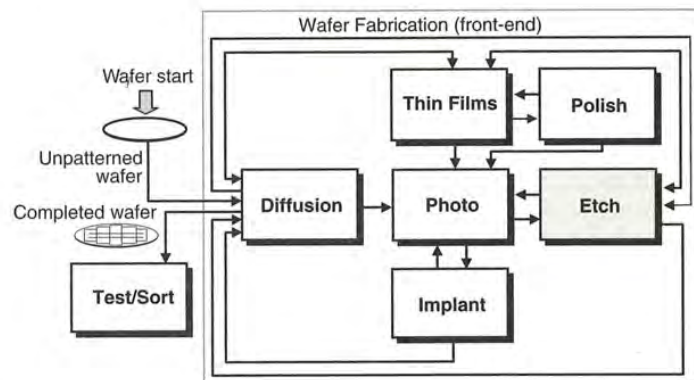


FIGURE 16.1 Applications for Wafer Etch in CMOS Technology

FIGURE 16.2 General CMOS Process Flow
(Used with permission from Advanced Micro Devices)

Etch Processes

Two basic types of etch processes are used in semiconductor manufacturing: dry etch and wet etch. *Dry etch* exposes the wafer surface to a plasma created in the gaseous state. The plasma passes through the openings in the patterned resist and interacts physically or chemically (or both) with the wafer to remove the surface material. Dry etching is the primary method used for etching devices with submicron geometries. Thus, dry etching will receive the most coverage in this text. In *wet etch*, liquid chemicals such as acids, bases, and solvents are used to chemically remove wafer surface material. Wet etch is generally applicable only for larger geometries ($>3\ \mu\text{m}$). It is still used to remove some layers from the wafer and to remove dry etch residues.

Dry etch can also be categorized by the type of material removed. The three major material categories for etching are metal etch, dielectric etch, and silicon etch. *Dielectric etch* is used for applications with dielectric material, such as silicon dioxide. The formation of contact and via structures require a dielectric etch to form openings in the ILD. This is a challenge for etching high-aspect ratio openings (ratio of height to width for an opening). *Silicon etch* (including polysilicon) is used for applications requiring silicon removal, such as etching polysilicon transistor gates and silicon trench capacitors. *Metal etch* is primarily used for removal of aluminum alloy stacks to form interconnect wiring on metal layers. At the time of this writing, there is no acceptable method for etching copper metal with submicron feature sizes (which is an important reason for the introduction of dual-damascene processing into the wafer fab).

Etch can also be categorized as either patterned or unpatterned etching. *Patterned etching* uses a masking layer (patterned photoresist) to define areas of the surface material that are to be etched. Only selected portions of this wafer surface layer are removed during the etch process. Patterned etching is used to form the many different features on the wafer surface, including gates, metal interconnects, vias, contact holes, and trenches. *Unpatterned etching*, *etchback*, or *stripping*, occurs when the entire wafer surface is etched with no mask present. This etching process is used to strip masking layers (e.g., STI nitride strip and the titanium strip after the salicide process used

contaminants in the etched film, improperly chosen etch chemistry (e.g., etching too fast), contaminants in the chamber, and nonuniform dopant distribution in the film. There are different names for the residues left after etch, including stringers, veils, crowns, and fences. Stringers are small residues of the etched material that are not totally removed, are electrically active, and can form an undesirable connecting short between features. Residues are a source of wafer contamination for IC fabrication and can cause problems during resist stripping. An overetch is sometimes done at the end of the etch process to remove residues. In some cases the etch residues can be removed by a resist strip process or by wet chemical etching.

Polymer Formation

A *polymer formation* is sometimes intentionally deposited on the sidewalls of the etch feature to form an etch-resistant film that prevents lateral etching (see Figure 16.10). This produces highly anisotropic features because it blocks the etching of the sidewall, thus increasing the etch directionality. The result is better control of critical dimensions of patterned structures. The polymers come from photoresist carbon converted into polymers during etching and combines with etching gases (i.e., C_2F_4) and etch by-products to form this sidewall polymer.² The need for sidewall polymers depends on the type of etching gas used.

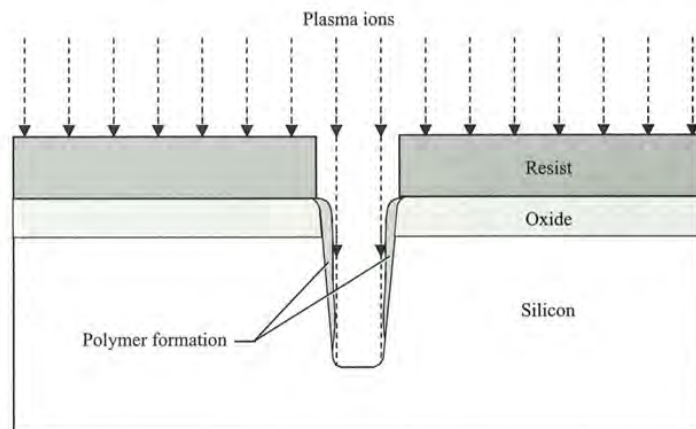


FIGURE 16.10 Polymer Sidewall Passivation for Increased Anisotropy

These sidewall polymers are complex, containing etchant elements and reaction by-products such as aluminum, titanium from barrier layers, oxides, and other inorganic materials. The polymer chains have strong carbon-fluorine bonds that are difficult to oxidize and remove.³ However, the polymers must be removed after the etch process or device yield and reliability is affected. The cleaning of the sidewall often requires special gas chemistry for a plasma stripping process or possibly a wet clean using strong solvents followed by DI-water rinsing (see the following section).

Unfortunately, an undesirable side effect of polymer deposition is that the internal parts of the process chamber also become coated with polymer. Etch process chambers require periodic cleaning to remove the polymer and to replace parts that can not be cleaned.

Plasma-Induced Damage

Plasma consists of energized ions, electrons, and excited molecules, which can cause *plasma-induced damage* to sensitive devices on the wafer surface. A major type of damage is nonuniform plasma that creates trapped charges on the gate electrode of a transistor, causing a breakdown of the thin gate oxide.⁴ Plasma becomes nonuniform due to poor equipment design or from operating the plasma reactor outside of the optimum process window. Another type of device damage is from energetic ion bombardment to gate oxide that is directly exposed to the ions. This damage could occur at the edge of a gate electrode during etching. Plasma damage is sometimes removed from the wafer through anneals or chemical wet etching.

Particle Contamination

Wafer damage from plasma can also come from particle contamination generated by the plasma near the wafer surface. Studies have shown that particles are trapped near the plasma/sheath interface because of the electrical potential difference.⁵ When the plasma is turned off, these particles fall onto the wafer surface. Plasma based on fluorine gas chemistries produces fewer particles than chlorine or bromine gas chemistries because fluorine generates etch by-products with a higher vapor pressure. Control of particle contamination is done by an optimized tool design, proper tool operation and shutdown, and use of the appropriate gas chemistry for the film being etched.

DRY ETCH

Dry etch is the primary etching method used to remove surface material in semiconductor manufacturing. The goal of dry etch is to reproduce the image of a mask on the wafer surface with a high degree of integrity. The advantages to using dry etch over wet etch are listed in Table 16.2.

TABLE 16.2 Advantages of Dry Etch Over Wet Etch

Description of Advantages
1. Etch profile is anisotropic with excellent control of sidewall profiles.
2. Good CD control.
3. Minimal resist lifting or adhesion problems.
4. Good etch uniformity within wafer, from wafer-to-wafer, and from lot-to-lot.
5. Lower chemical costs for usage and disposal.

There are disadvantages to using dry etch. The primary disadvantages are poor selectivity to the underlying layer, risk for device damage from plasma, and expensive equipment.

In the dry etching process, a low-pressure plasma discharge is used to remove material in small feature sizes of integrated circuits (the creation of a plasma was discussed in Chapter 8). The plasma interacts with the wafer surface to cause etching action and the subsequent removal of the surface material. The major actions that occur for plasma etching of a substrate material are shown in Figure 16.11.⁶

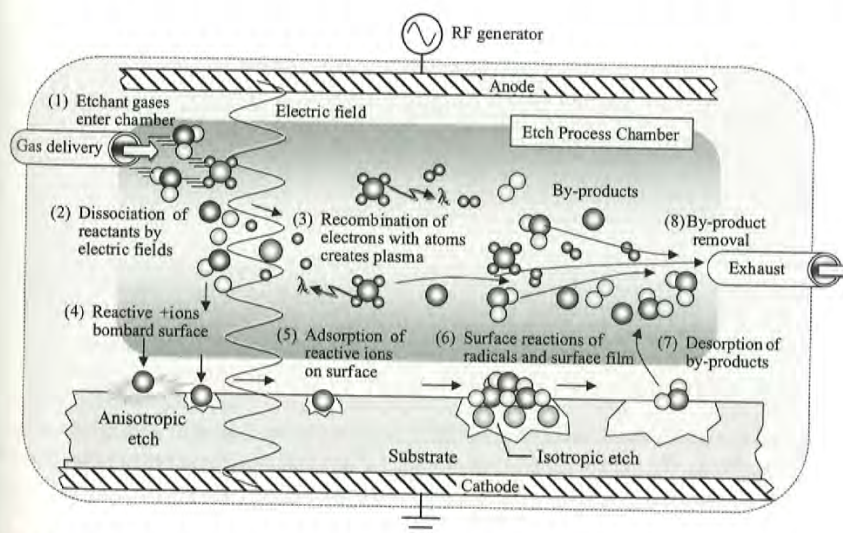


FIGURE 16.11 Plasma Etch Process of a Silicon Wafer

Reactive Ion Etch (RIE)

Reactive ion etch (RIE) is a technique for removing material from the wafer surface with both a reactive chemical process and a physical process using ion bombardment. RIE is similar to the standard parallel-plate plasma etcher, except the wafers are placed on the RF-powered electrode (cathode) and the powered electrode size is greatly reduced relative to the grounded electrode size (see Figure 16.19). In this manner a DC self-bias develops on the cathode and the wafers acquire a large voltage difference with respect to the plasma. This condition creates directionality for the ionized species moving toward the wafer, creating improved anisotropic sidewalls for features. There is no sputtering on the anode. The pressure is relatively low at <0.1 torr.

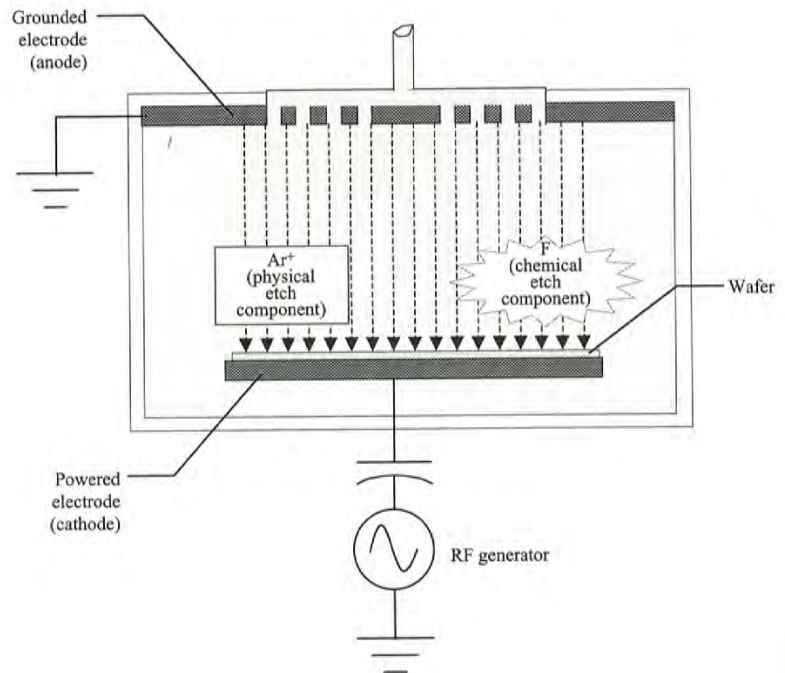


FIGURE 16.19 Parallel-Plate RIE Reactor

High-Density Plasma Etchers

The most predominant dry etching methodology in use for critical layers of advanced ICs is *high-density plasma etch* with single-wafer processing in a cluster tool.¹² Standard plasma systems previously used in wafer fabrication operated in a vacuum regime of a few hundred mtorr, where it is relatively straightforward to create a plasma (such as by applying an RF signal across two parallel plates). However, for $0.25\text{-}\mu\text{m}$ geometries and below, it has become too difficult to get etchant ions into and etch by-products out of high-aspect ratio features. The etching action slows down and will actually stop altogether at the bottom and lower sides of the feature profile.

The solution is to lower the system pressure to about 1 to 10 mtorr to increase the mean free path lengths of gas molecules and ions. This condition effectively reduces collisions that cause loss of profile control. However, a drawback to reduced pressure is the lower etch rate due to the rapid reduction in ion density as pressure decreases. To overcome this concern, a high-density plasma is needed to generate enough ions for an acceptable etch rate with a lower pressure. High-density refers to the number of active species in the plasma relative to conventional plasma at the same process pressure. In conventional plasma, the degree of ionization is typically on the order of 0.01% to 0.1%. High-density plasma more efficiently couples input power with the plasma, resulting in greater dissociation of etch species to achieve a degree of ionization as high as 10%. This technique yields highly directional, low-energy ions that produce anisotropic etching in high aspect ratio openings.

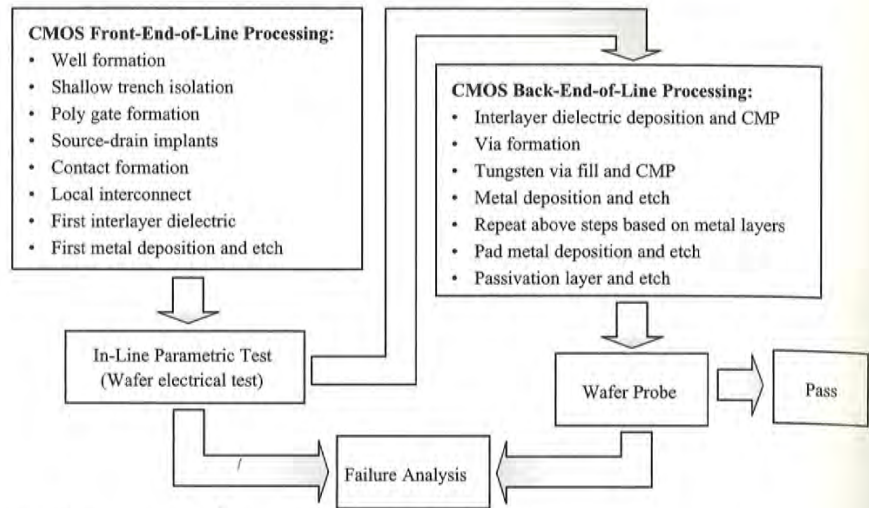


FIGURE 19.2 Location of Wafer Fabrication Electrical Tests

The reasons for the in-line parametric test are:

1. **Identify process problems:** Early identification of process problems during wafer fabrication (instead of waiting until an entire wafer is fabricated before finding out if there is a problem).
2. **Pass/fail criteria:** Determine whether wafers should continue the fabrication process based on pass/fail criteria.
3. **Data collection:** Collect wafer data to assess process trends in order to make process improvement (e.g., channel length variation).
4. **Special tests:** Assess specific performance parameters when required (e.g., special customer requests).
5. **Wafer level reliability:** Perform optional wafer level reliability tests on lots when needed to assure reliability concerns of process conditions.

In-line parametric tests are done as early as possible in the fabrication process. No power supply and signal voltages are applied to the devices on the wafer. Instead current, voltage, and capacitance measurements are made using special parametric test structures to determine the processes' capability. The test is important because it is the first time the wafer undergoes a complete set of tests to verify it was fabricated correctly. The pass/fail data establishes a close link between process conditions and device performance.

With increasing IC complexity, testing early in the process is critical. If the process conditions for fabricating the wafer are unacceptable, then parametric tests will detect this by demonstrating the product fails the electrical limits of the test. The defect reduction team is alerted to the problem and can immediately take corrective action. The data provides ongoing information to monitor and optimize the front-end semiconductor processes.²

Wafer Test Structures ■ Parametric testing is not performed on the individual wafer devices. Rather, it is done on special *test structures* (also referred to as *process control monitors*, or *PCMs*) arranged at specific locations on the wafer. Test structures are used because they prevent damage to the actual production die during testing. For early design verification tests, the test structures can be a test pattern on an entire special chip because of the need for many structures and test data. For production wafers, space is a premium; therefore, test structures are commonly located in the scribe line region between the individual die. These are sometimes referred to as *scribe line monitors* (*SLMs*) (see Figure 19.3). The width of the scribe line that the SLM must fit in might be 100 to 150 μm ; thus the SLMs are limited in size.

PCM test structures are used to test a wide range of parameters. There is a growing trend in the industry toward standardized test structures for various measurements, mainly due to the



FIGURE 19.3 Scribe Line Monitor Test Structure

structure design complexity and limited development time for new products.³ Examples of some typical test structures are provided in Table 19.2.⁴

TABLE 19.2 Examples of Test Structures

Test Structure	Fault Measurement
Discrete transistors	Leakage current, breakdown voltage, threshold voltage, and effective channel length
Various linewidths	Critical dimensions
Box in a box	Critical dimensions and overlay registration
Serpentine structure over oxide steps	Continuity and bridging
Resistivity structure	Film thickness
Capacitor array structure	Insulator materials and oxide integrity
Contact or via string	Contact resistance and connections

Figure 19.4 shows a representative PCM test structure used to assess the film thickness of the first metal contacts as a function of sheet resistance. The in-line parametric test measures the series resistance of the structure. A problem with the contacts, such as an improper barrier film thickness, would cause this test structure to electrically fail the parametric test. Since the test structure was fabricated at the same time and under the same process conditions as the wafer, it highlights a problem existing with the actual contacts on individual wafer die. This information is valuable to the production team for defect reduction.

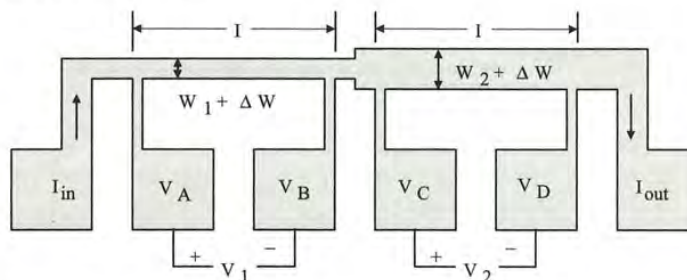


FIGURE 19.4 Test Structure for First Metal Contacts

Redrawn from *Microelectronics Manufacturing Diagnostics Handbook*, ed. A. Lanzberg (New York: Van Nostrand Reinhold, 1993).

CHAPTER 20

ASSEMBLY AND PACKAGING

The use of wafers is a cost-effective way to simultaneously fabricate many chips. Once all fabrication and tests are complete, chips are separated from the wafer and assembled into the final IC package. The assembly and packaging process takes good electrical devices, places them in a package, and interconnects the device bonding pads to the package leads. Packaging provides a means of protecting the chip and attaching it to a higher level of assembly.

In the early days of semiconductor manufacturing, the cumbersome interconnection of the chip to the metal enclosure surrounding the chip was a manual process. U.S. chipmakers viewed it as highly labor intensive and transferred this process technology to offshore facilities or contractors. In the modern era, IC assembly and packaging has become highly automated and is an increasingly important part of the product price and performance for the end user. High-volume IC assembly is still dominated by Asia-Pacific region companies.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Describe the general trends and design constraints of assembly and packaging.
2. State and discuss the traditional assembly methods.

There are many new chip assembly and packaging designs that are in development for ICs. New designs are continually benchmarked for cost and performance against the existing approaches that have been proven over the past 20 to 30 years. As customer demand improves chip performance, new requirements are placed on IC packaging. Critical packaging parameters are the number of input and output (I/O) pins, electrical performance, thermal dissipation, and size.

The general trend is the merging of wafer fabrication technology with assembly and packaging. The size of the package is continually reducing to reflect the chip size. Some chipmakers are packaging ICs at the wafer level. Multiple chips are placed in one package. Ultimately, it is inevitable that the wafer fabrication process and the IC assembly and packaging eventually become part of the same process.

3. Describe the different traditional packaging options.
4. Discuss the benefits and limitations of seven advanced assembly and packaging techniques.

INTRODUCTION

At the completion of the fab process, wafers that pass electrical test are ready for assembly and packaging of individual chips. This is done during *final assembly and packaging*, referred to as the *back-end* of the IC manufacturing process. IC final assembly and packaging is a big business. In 1998, 62 billion IC package units were sold worldwide.¹

Final assembly and packaging are two distinct processes in the back-end of IC manufacturing. Each has their specific processes and tools. In the traditional process, *IC final assembly* separates each good die from the wafer and attaches the die to a metal leadframe or substrate. For leadframe assembly, the metal bond pads on the chip surface are interconnected with fine diameter wire to an internal leadframe to provide electrical access to the chip. After final assembly, *IC packaging* encloses the die in a protective package. The most common package in use today is the encapsulation

of the chip with a plastic molding compound. This plastic package provides environmental protection and formed leads for connection to the next higher level of assembly (e.g., soldered to a circuit board). An outline of the traditional final assembly and packaging process is shown in Figure 20.1.

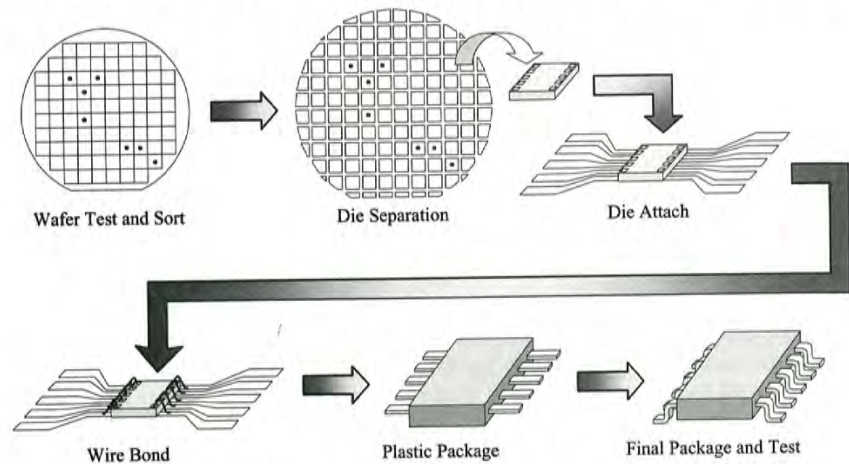


FIGURE 20.1 Traditional Assembly and Packaging

IC packaging has four important functions for all chips:

1. Protection from the environment and handling damage.
2. Interconnections for signals into and out of the chip.
3. Physical support of the chip.
4. Heat dissipation.

Many packaging variations exist in the industry, with some of the most common packages shown in Figure 20.2 (additional information on packages is provided later in this chapter). The package is selected so that the four functions are optimized to meet certain design constraints: performance, size, weight, reliability, and cost objectives (see Table 20.1). Packaging the IC permits the microchip to function in a wide range of customer environments, such as in a notebook computer, in the engine compartment of an automobile, and sandwiched between the plastic layers of a credit card. The various environmental conditions of contamination, moisture, temperature, mechanical vibration, and physical abuse all must be taken into account when the design engineers select the IC package.

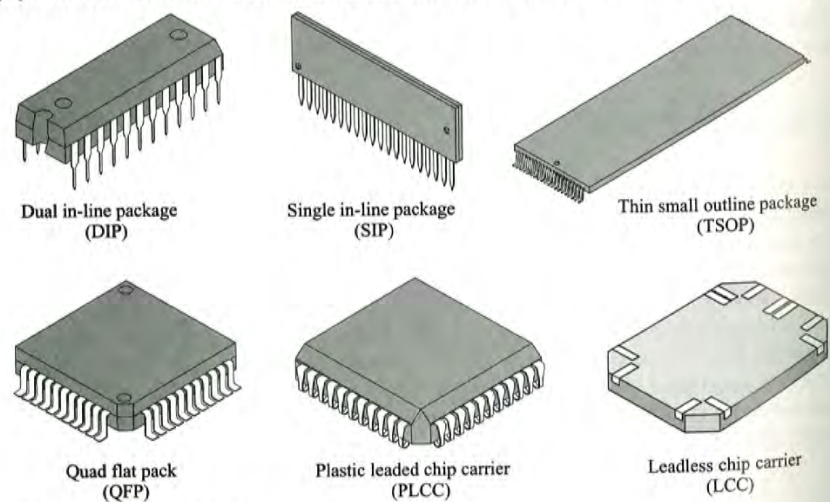


FIGURE 20.2 Typical IC Packages

TABLE 20.1 Design Constraints for IC Packaging

Design Parameters	Design Constraints
Performance	• RC time delay • Number of signal inputs/outputs (IOs) • Wirebond versus bump attachment • Signal rise time • Switching transients • Thermal • Power dissipation • Input and output impedance • Frequency response
Size/weight/form	• Chip size • Package size • Bond pads size and pitch • Package leads size and pitch • Substrate carrier pads size and pitch • Design of heat sink
Materials	• Chip substrate (plastic, ceramic, or metal) • Carrier (organic, ceramic) • Thermal expansion mismatch • Lead metallurgy
Cost	• Integration into existing process • Package materials • Yield
Assembly	• Method of die attach • Package attach (through hole, surface mount, or bumped) • Heat sink assembly • Encapsulation

For high-performance chip applications, such as high-end computers, performance and reliability are critical. For most consumer applications, cost is an important criteria, along with size and weight.

Packaging Levels

Two different packaging levels exist for electronic components (see Figure 20.3 on page 574). The chip assembly and packaging, which is the topic of this chapter, is referred to as *1st level packaging*. Once the chip is packaged into an IC component, the package I/O terminals connect the chip to the next level of assembly. The *2nd level packaging* assembles the IC component into a system with many components and connectors.² In most instances of 2nd level packaging, the IC component is assembled onto a printed circuit board using eutectic Sn/Pb solder (melting temperature of 183°C). A *printed circuit board (PCB)*, also referred to as a *substrate* or *carrier*, has the circuit interconnections to attach the chip-carrying IC components onto the board with solder, while using connectors to interface into the remaining electronic subsystems of the product. Assembled circuit boards are then put into the final product (sometimes referred to as box-line assembly, such as for the assembly of computers).

This chapter covers the traditional method of IC assembly and packaging and then uses this information as a foundation to introduce the advanced assembly and packaging approaches. Many of the traditional methods are still in use in a wide range of products, primarily because of their low cost and proven reliability. Advanced packaging techniques are generally used for high-performance microchip applications.

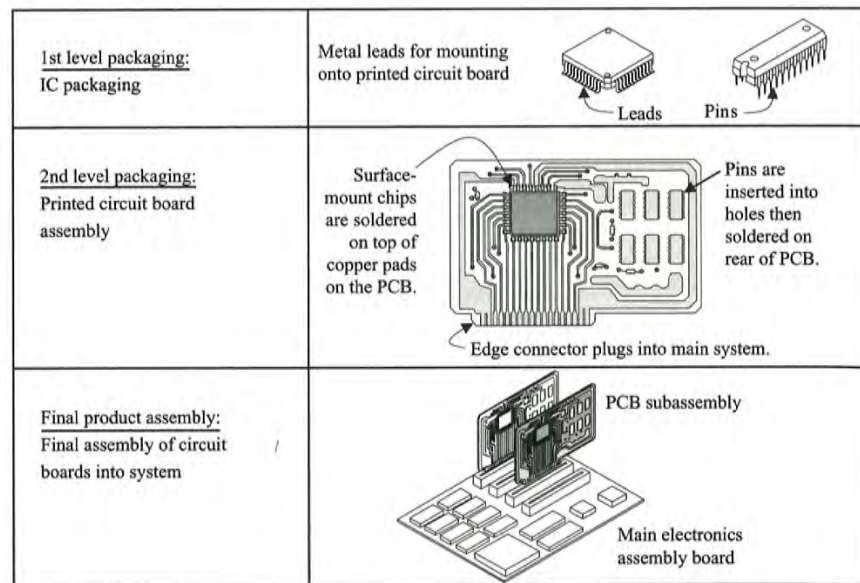


FIGURE 20.3 Levels of IC Packaging

TRADITIONAL ASSEMBLY

Final assembly consists of the operations required to attach the chip to the IC packaging. Manufacturing yield is critical during final assembly, since the majority of the fabrication cost is already in the chip. Traditional final assembly, which has been used for an estimated 95% of all ICs assembled during the latter 1990s,³ consists of the following four steps:

- ◆ Backgrind
- ◆ Die separation
- ◆ Die attach
- ◆ Wire bonding

Backgrind

The first operation in final assembly is *backgrind* (sometimes this is done right after wafer sort before shipment to final assembly). Large diameter wafers are relatively thick to minimize breakage during front-end fabrication (a 300-mm wafer is 775 μm thick, which is roughly $1/32$ of an inch). However, wafers must be thinned before assembly begins. Wafers are usually thinned to a thickness of 200 to 500 μm .⁴ Thinner wafers are easier to dice into chips and have improved thermal dissipation, which is beneficial for reducing thermal stresses in thin ULSI packages. Thinner chips also reduce the size profile and weight of the final IC package.

The backgrind operation is done with fully automated machines (see Figure 20.4). Backgrind is carefully controlled to minimize the introduction of stress into the wafer.⁵ Stress can cause the wafer to warp, making it more likely to break during dicing and more difficult to mount into the package. In some cases, backside metallization is deposited after backgrinding. The metal is typically a gold film applied to improve electrical conductivity to the substrate and for eutectic die bonding (see the following section).

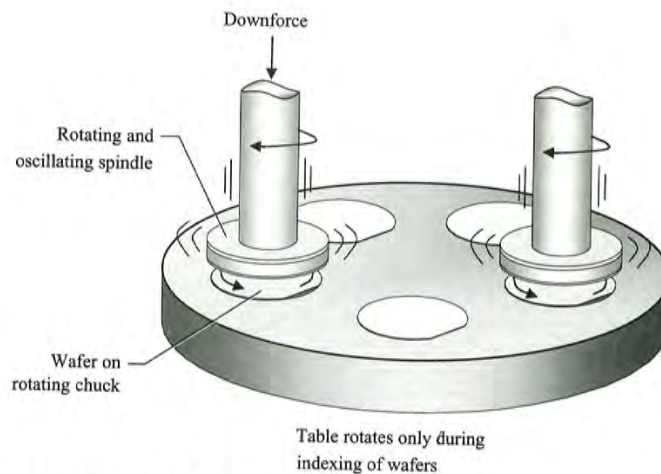


FIGURE 20.4 Schematic of the Backgrind Process

Die Separation

Die separation (also known as *die singulation*) cuts each die from the wafer using a diamond-blade dicing saw. Before dicing, wafers are removed from a cassette and properly oriented and placed onto an adhesive film secured to a rigid frame. The adhesive film holds the wafer intact once all chips are diced. The wafer is transferred into the saw, sprayed with DI water, and then sliced in the X and Y direction by a 25- μm thick diamond saw blade rotating up to 20,000 RPM (see Figure 20.5). The wafer is rinsed with DI water to remove the silicon slurry residue from the cutting action while each individual die is held in place by the adhesive backing. The saw usually cuts the wafer with 90 to 100% saw-through of the scribe lines. Wafer dicing is done with fully automated equipment that uses alignment systems and integrated wafer cleaning. Older dicing methods with 50% saw-through followed by a breaking process are usually not acceptable for ULSI devices.⁶

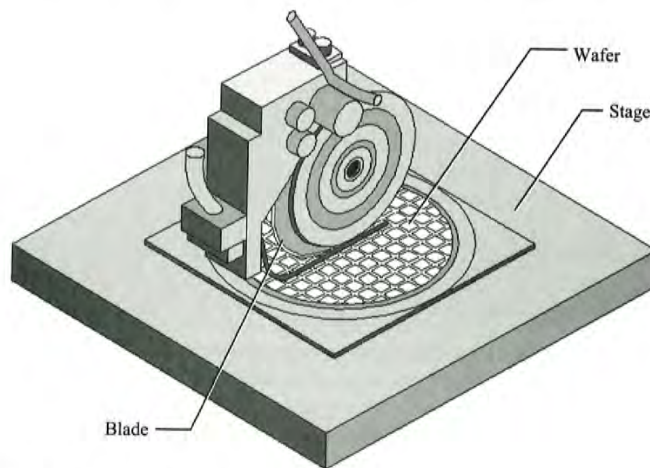


FIGURE 20.5 Wafer Saw and Sliced Wafer

Die Attach

After die separation, the wafers are moved to the *die attach* operation. At die attach, each good die is individually picked from the adhesive backing and physically attached to the substrate or leadframe. An example of a leadframe is shown in Figure 20.6 on page 576, with the leads fanning out

from the internal chip-bonding region to the larger lead pitch needed for higher-level assembly (pitch is the center-line to center-line spacing of leads). Leadframe magazines are small handling racks used to efficiently handle the leadframes and move them from tool to tool. The automated die bonder is a high-speed tool that uses a special gripper, referred to as a collet, to pick up die by their edge (to avoid chip damage) and place them on the substrate or leadframe for assembly. Die bonder tools require flexibility to attach die to a variety of applications, including leadframes, ceramic substrates, and circuit boards. The good die are selected based on a detector that recognizes the absence of an ink mark or by using the computerized wafer mapping data supplied from wafer sort.

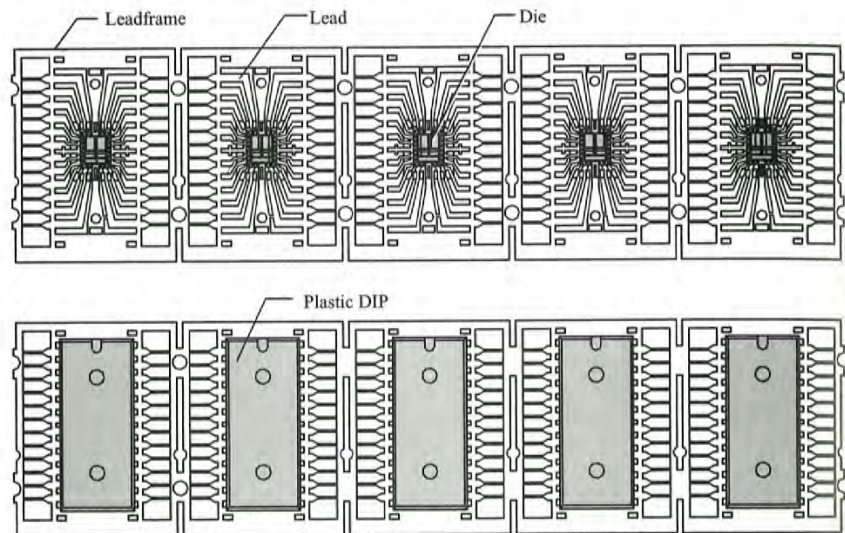


FIGURE 20.6 Typical Leadframe for Die Attach

Chip Bonding ■ The die is physically attached to the leadframe or substrate using one of the following techniques:

- ◆ Epoxy attach
- ◆ Eutectic attach
- ◆ Glass frit attach

Epoxy Attach. *Epoxy attach* is the most common method for bonding the chip to the leadframe or substrate. An epoxy is dispensed in the center region of the leadframe or substrate. The die bonder tool places the back of the chip onto the epoxy (see Figure 20.7), followed by a thermal heat cycle to cure the epoxy (e.g., 125°C for one hour). Most MOS applications use straight epoxy. However, the epoxy can be formulated with silver flakes to make a thermally conductive epoxy if there is a need for heat dissipation between the chip and the rest of the package.

Eutectic Attach. To attach a chip using *eutectic attach*, a thin film of gold (Au) is deposited on the backside of the wafer after backgrind (see Figure 20.8). Recall that eutectic defines a solution composition in order to minimize its melting point. This gold is then joined by alloying to the substrate, which is usually either a metal leadframe (e.g., alloy 42, a Ni-Fe alloy) or a ceramic substrate (e.g., 90 to 99.5% Al_2O_3). Typically, the substrate has a metallized surface of Au or silver (Ag). When heated to 420°C for approximately six seconds, which is slightly above the Au-Si eutectic temperature, this method forms a eutectic alloy interconnection between the chip and leadframe. Eutectic attach provides a good thermal path and mechanical strength. The eutectic attach technique is more common for bipolar ICs.

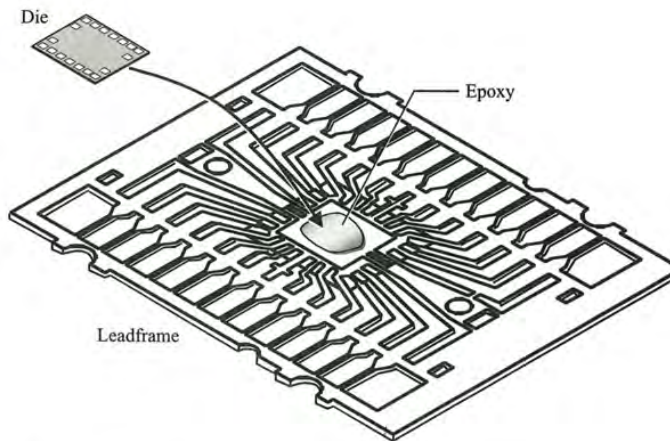


FIGURE 20.7 Epoxy Die Attach

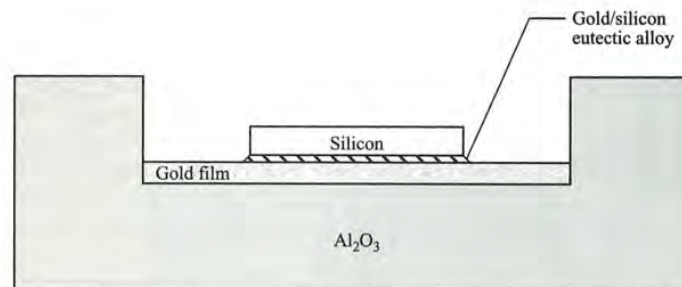


FIGURE 20.8 Au-Si Eutectic Attach

Glass Frit Attach. The *glass frit attach* consists of a mixture of silver and glass particles suspended in an organic medium. It is used to attach chips directly to Al_2O_3 ceramic packages without metallization to attain a hermetic seal. *Hermetic sealing* refers to protecting the silicon device from the external environment, specifically moisture and contamination. The silver and glass used in glass frit soften during cure and form a bond to the ceramic with good thermal conduction. It requires a relatively high temperature to cure the silver-filled glass.

Wirebonding

Wirebonding is the most common method for electrically connecting the aluminum bonding pads on the chip surface to the package inner lead terminals (sometimes referred to as *posts*) on the leadframe or substrate (see Figure 20.9 on page 578). This high-speed operation spools and bonds a fine diameter wire from the chip bond pads to leadframe inner lead pads, with the ability to form multiple bonds per second (e.g., a common speed is ten wire bonds per second). The tool bonds the wire at each die bonding pad or leadframe pad and steps to the next location. Wirebond placement accuracy is often $\pm 5 \mu\text{m}$.⁷ The bond wire is either Au or Al wire because it bonds well to both chip pads and leadframe inner lead pads, with a common wire diameter between 25 to 75 μm . A standard wire diameter of 25 μm is used on a die bonding pad pitch of 70 μm .⁸

The three basic types of wirebonding derive their names from the type of energy used during the wire termination process. The three methods of wirebonding are:

- ◆ Thermocompression bonding
- ◆ Ultrasonic bonding
- ◆ Thermosonic ball bonding