FUNDAMENTALS OF MICROSYSTEMS PACKAGING

Rao R. Tummala
Georgia Institute of Technology
Editor

Angie Hughes
Nancy Trent
Mahesh Varadarajan
Managing Staff

Olatunji Cunningham
Travis Gary
Maurice Novembre
Alyssa Poole
Salley Theodore
Illustrators

Kristy D'Arcy
Melissa Sherrer
Administrative Assistance

McGRAW-HILL

New York Chicago San Francisco Lisbon London
Madrid Mexico City Milan New Delhi San Juan
Seoul Singapore Sydney Toronto

Samsung, EX1028, p. 1

Cataloging-in-Publication Data is on file with the Library of Congress.

McGraw-Hill

A Division of The McGraw-Hill Companies



Copyright © 2001 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

3 4 5 6 7 8 9 0 DOC/DOC 0 7 6 5

ISBN 0-07-137169-9

The sponsoring editor for this book was Steve Chapman and the production supervisor was Sherri Souffrance. It was set in Times Roman by Pro-Image Corporation.

Printed and bound by R. R. Donnelley & Sons.

This book was printed on acid-free paper.

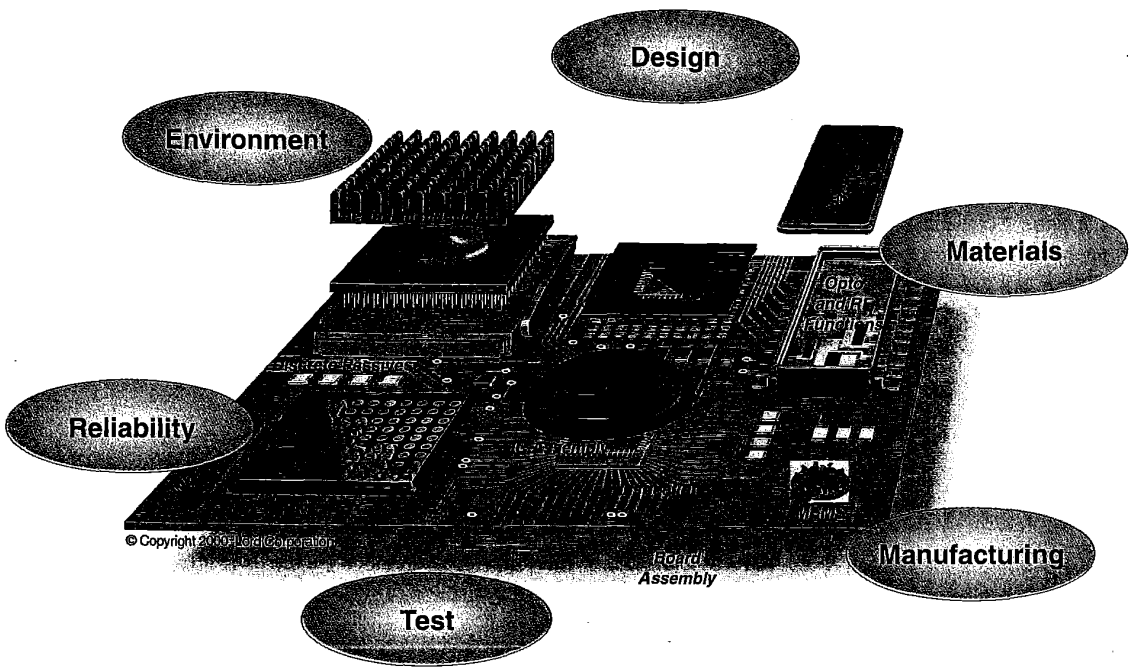
McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please write to the Director of Special Sales, McGraw-Hill Professional, Two Penn Plaza, New York, NY 10121-2298. Or contact your local bookstore.

Information contained in this work has been obtained by The McGraw-Hill Companies, Inc. ("McGraw-Hill") from sources believed to be reliable. However, neither McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein and neither McGraw-Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw-Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

CHAPTER 1

INTRODUCTION TO MICROSYSTEMS PACKAGING

Prof. Rao R. Tummala
Georgia Institute of Technology



1.1 What Are Microsystems?
1.2 Microsystem Technologies
1.3 What Is Microsystems Packaging (MSP)?
1.4 Why Is Microsystems Packaging Important?
1.5 System-Level Microsystems Technologies
1.6 What is Expected of You as a Microsystems Engineer?
1.7 Summary and Future Trends
1.8 Who Invented Microsystems and Packaging Technologies?
1.9 Homework Problems
1.10 Suggested Reading

CHAPTER OBJECTIVES

- Introduce the concept of Microsystems
- Present Microsystems building block technologies to contain Microelectronics, Photonics, MEMS, and RF/Wireless
- Describe the role of Packaging as IC and Device Packaging, and Microsystems Packaging
- Describe Systems Packaging to go from wafer to complete system
- Describe the role of electrical, mechanical and materials in Systems Packaging

CHAPTER INTRODUCTION

Microsystems and the technologies they constitute are the building blocks of information technology. These systems require a set of fundamental technologies that include not only microelectronics but also photonics, MEMS, RF and wireless. For these functions to be integrated into systems, they have to be designed, fabricated, tested, cooled and reliability assured. In other words, they have to be system-packaged. This book is about Systems Packaging.

Samsung, EX1028, p. 4

1.1 WHAT ARE MICROSYSTEMS?

Imagine a world without personal computers, cell phones, fax machines, camcorders, stereos, microwave ovens, calculators and all the other electronic products. Now imagine what is coming. Microsystems will impact many areas of life. These systems include voice controlled computers, electronic notepads and work surfaces, electronic newspapers on flat panel displays which can be rolled or folded, small mobile x-ray and diagnostic tools, micromedical implants, videophones in watches and wireless Internet access anytime anywhere. The technologies behind all these and millions of other electronic products in automotive, consumer, telecommunication, computer, aerospace, and medical industries are all based on microdevices and packaging technologies. They touch every aspect of human life with the potential to bring everyone around the globe into the digital age.

1.1.1 Microsystem Products

Microsystems are microminiaturized and integrated systems based on microelectronics, photonics, RF, micro-electro-mechanical systems (MEMS) and packaging technologies. These new systems, and technologies as illustrated in Figure 1.1 and Figure 1.2, provide a variety of integrated functions that include consumer, computing, communications, automobile, sensing and micromechanical functions to serve a variety of human needs.

Visions for future products cover all areas of life such as smart watches with integrated phone and video, wearable computers, multifunctional global phones, and micro miniaturized medical implants. Many futuristic devices are becoming feasible, or have already been partly realized, based on the miniaturization in microsystem technologies, leading to electronics with larger memory capacities, higher computing speeds with lower energy consumption batteries, and finally, to more powerful data networks and improved data

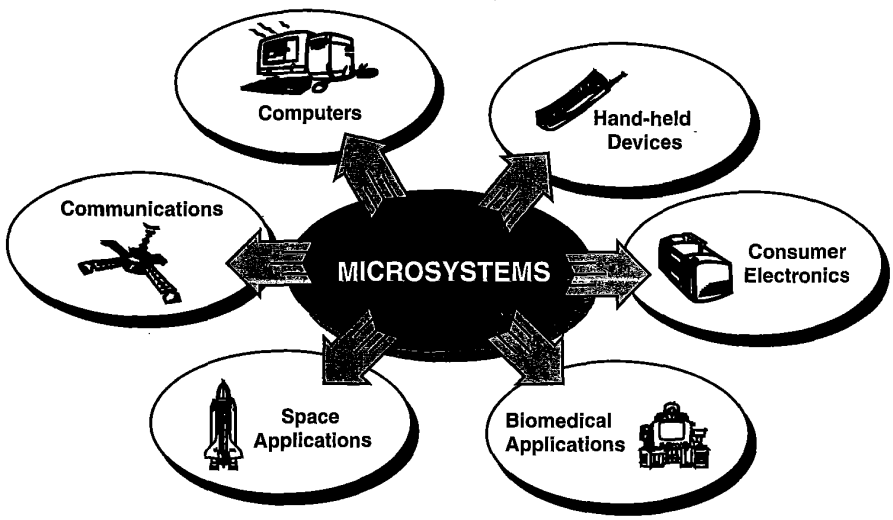


FIGURE 1.1 Microsystem products.

Samsung, EX1028, p. 5

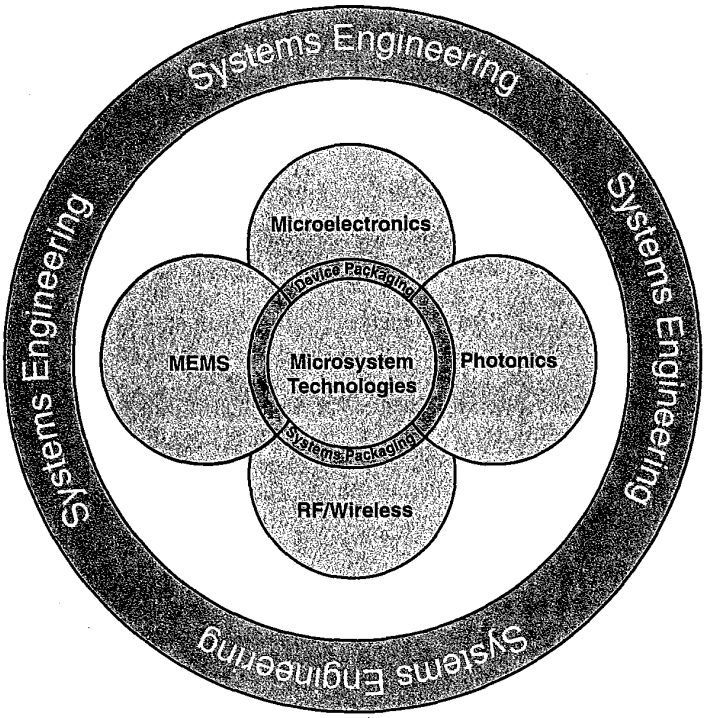


FIGURE 1.2 Microsystem technologies.

compression processes. Microsystem components also play an increasingly important role. Table 1.1 gives some of the trends in microsystems products.

1.2 MICROSYSTEM TECHNOLOGIES

The fundamental building block technologies behind all the electronic products, whether that product is a PC, a DVD player, a cell phone or an airbag in your car, are four technology waves: Microelectronics, RF/wireless, Photonics and MEMS. A fifth wave, consisting of Systems Packaging that integrates and engineers all these into products, is depicted in Figure 1.3.

1.2.1 Microelectronics: The First Technology Wave

Microelectronics is the first and most important technology wave. It started with the invention of the transistor. The three discoveries that made this possible were:

1. The invention of the transistor in 1949 by Brattain, Bardeen and Shockley at Bell Labs
2. The development of planar transistor technology by Bob Noyce in 1959

Samsung, EX1028, p. 6

TABLE 1.1 Trends in microsystems equipment.

Electronic Product	Trend
Consumer	Analog to digital
	TV/PC merge
	Network of music and video
	Interactive communication (Merge of PC with phone)
	Analog to digital audio (CD-ROM, MD, MP-3, etc.)
	Video to digital disk (DVD)
	LAN (Electrical→Optical→Wireless)
	Smart watch with integrated video and phone
	Smart clothes
	House sensors
	Wearable computer
Communication	From voice to data exchange
	Mobile phone (analog to digital)
	Electrical cable to optical fiber
	Multi-functional phone, global phone
Automobile	Navigation system
	Auto pilot
	Integrated electronics control
	(Air bag, anti-skid brake, automatic transmission, etc.)
	Collision warning system
	Information and traffic control systems
Computers	PC to networking system
	Multi-functional mobile equipment from Desktop
	→Notebook→Palmtop
Infrastructure	Big capacity, high speed server
	High bandwidth, wireless Internet access
	Analog to digital equipment
Medical	Mobile X-ray
	Diagnostic tools
	Miniaturized implants
	Pumping and injection systems
	Drug-screener
	Microinstruments for endoscopic neurosurgery

3. The first integrated circuit (IC), which incorporated two transistors and a resistor, (Figure 1.4) developed by Jack Kilby in 1959. Their combined discoveries earned them Nobel Prizes in 1972 and in 2000. The transistor is the single most important fundamental building block of all modern electronics. Microelectronics acts as the fundamental base of more than 90% of all microsystems products. Figure 1.5 illustrates the famous Moore's Law. In 1965, three years before he co-founded Intel with Bob Noyce, Gordon Moore published an article in *Electronics* magazine that turned out to be uncannily prophetic. Moore wrote that the number of circuits on a silicon chip would keep doubling every year. He

Samsung, EX1028, p. 7

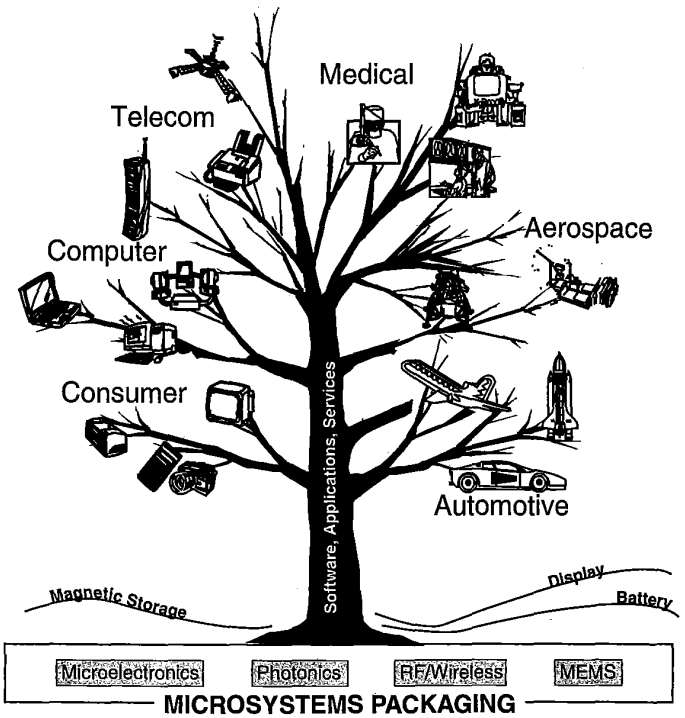


FIGURE 1.3 Building block technologies of the information age.

later revised this to every 18–24 months, a forecast that has held up remarkably well over several decades and countless product cycles.

The secret behind Moore's Law is that every 18 months or so chipmakers double the number of transistors that can be crammed onto a silicon wafer the size of a fingernail. They do this by etching microscopic grooves onto crystalline silicon with beams of ultraviolet radiation. A typical wire in a Pentium chip is now 1/500 the width of a human hair; the insulating layer is only 25 atoms thick.

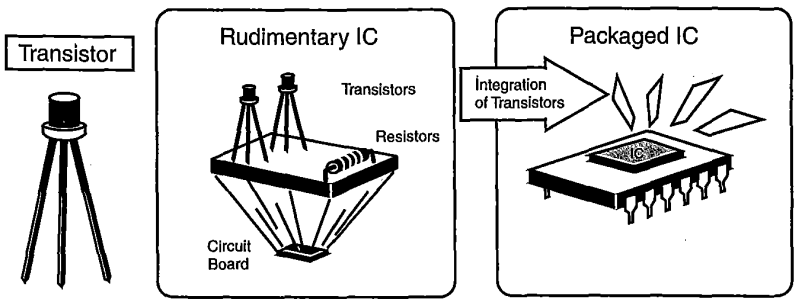


FIGURE 1.4 The invention of the first integrated circuit.

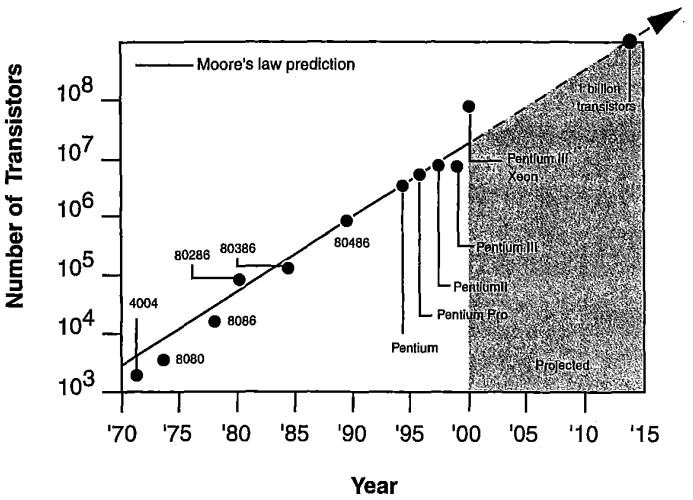


FIGURE 1.5 Moore's Law predicts the IC integration to double every 18 months.

But the laws of physics suggest that this doubling cannot be sustained forever. Eventually, transistors will become so tiny that their silicon components will approach the size of molecules. At these incredibly tiny distances, the bizarre rules of quantum mechanics take over, permitting electrons to jump from one place to another without passing through the space between. Like water from a leaky fire hose, electrons will spurt across atom-size wires and insulators, causing fatal short circuits.

Transistor components are fast approaching the dreaded point-one limit—when the width of transistor components reaches 0.1 micron and their insulating layers are only a few atoms thick. Last year, Intel engineer Paul Pakan publicly sounded the alarm in *Science* magazine, warning that Moore's Law could collapse. He wrote, "There are currently no known solutions to these problems."

They key word is known. The search for a successor to silicon has become a kind of crusade; it is the Holy Grail of computation. Among physicists, the race to create the Silicon Valley for the next century has already begun.

The economic destiny and prosperity of entire nations may rest on one question: can silicon-based computer technology sustain Moore's Law beyond 2020? Moore's Law is the engine pulling a trillion-dollar industry. It's the reason kids assume that it's their birthright to get a video-game system each Christmas that's almost twice as powerful as the one they got last Christmas. It's the reason you can receive (and later throw away) a musical birthday card that contains more processing power than the combined computers of the Allied Forces in World War II.

But microsystems are more than microelectronics. The microelectronics based on the transistor building block technology is one aspect of today's electronic systems, such as personal computers. But other systems, such as modern fiberoptic telecommunications, are based on photons, the fundamental properties of which are more superior in some respects, providing the needed higher bandwidth for today's Internet traffic. We call this the Photonic Wave. There are other systems that are not based on the building block

Samsung, EX1028, p. 9

transistors that are beginning to play a major role in modern electronics. These are RF and MEMS waves.

1.2.2 RF and Wireless: The Second Technology Wave

The world is going portable and wireless. The radio and wireless revolution started with Marconi in 1901. In December 1901, Gulielmo Marconi, in St. John's, Newfoundland received the first wireless message to cross the Atlantic. Sent from Poldhu, Cornwall, in England, his message was the letter S—three dots in Morse code. The demonstration of the transatlantic reception over 2900 km helped Marconi establish the business of wireless telegraphy. The originator of numerous innovations, including a method of continuous-wave transmission, as well as grounded antennas, improved receivers, and receiver relays, Marconi was also remarkable for his skills at marketing and promoting. In 1897, he had established a company that soon offered radio communications services, notably to shipping lines, though the transmission range was initially limited to some 240 km. By World War I, Marconi Companies in Britain and elsewhere were providing radio communications worldwide. This earned Marconi the Nobel Prize for Physics. The fundamental technology behind today's mobile phone is the same. A whole new industry has emerged with applications that span AM and FM radio to cellular phone to satellite to microwave communications, as illustrated in Figure 1.6, across the entire electromagnetic spectrum.

The main advantage of wireless is the fact that it cuts the cables, thus liberating the user from the tether to the network. It allows communications anywhere anytime. Of course, that is only realistic if the wireless equipment is small enough that it can actually be carried around everywhere. This is where Systems packaging applies.

Wireless technology is also increasingly used for non-communications functions. Top-of-the-line Mercedes cars, for example, are now equipped with a collision avoidance system that is based on radar. Navigational *global positioning systems* (GPS) are being

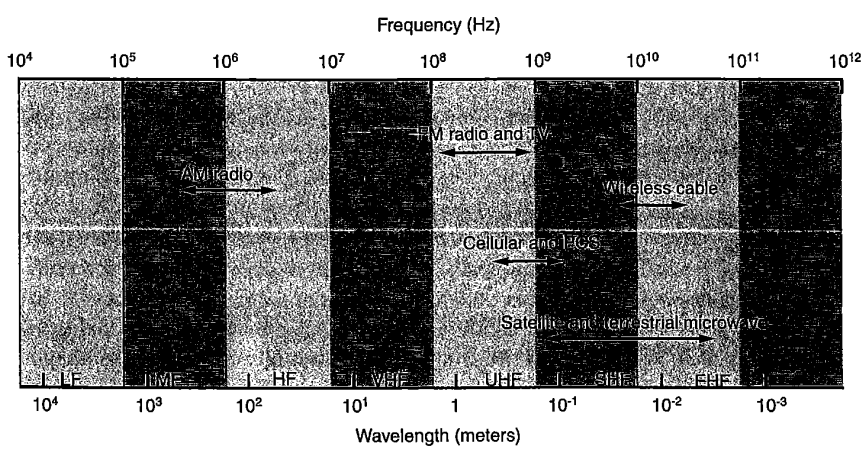


FIGURE 1.6 RF/Wireless applications.

Samsung, EX1028, p. 10

integrated into even more consumer items, where size and cost are paramount. For all of these products, a small, low-cost RF module is required. Another RF/wireless application is electronic toll taking on bridges and roads.

1.2.3 Photonics: The Third Technology Wave

In 1970, Corning Glass Works demonstrated highly transparent fibers, and Bell Laboratories demonstrated semiconductor lasers that could operate at room temperature. These demonstrations helped establish the feasibility of fiberoptic communications. These discoveries are the fundamental building block technologies of today's Internet networks.

As we enter the new millennium with exploding Internet volume and limitless business opportunities, it is natural to pause and think about how we are going to meet this challenge. Needless to say, there is no better known physical medium than fiber, and no signal source better than light to meet these new requirements. Therefore, as optical networking technology evolves from megabits per second to gigabits per second, we should become quite comfortable with the power of the exponents and what it means at the service levels. Fortunately, we have a ways to go to reach the data transport limit of fiber. For example, with the current state of device technology, a good laser source can emit 10^{16} photons/s, and a good detector—which can detect a bit with 10 photons—can detect 1 Pb (10^{15} b/s) on a single fiber. Nevertheless, the device technology is going to get better with time, further pushing the limit of optical fiber capacity. Thus, fiber optics is a future-proof technology. With wavelength-division multiplexing (WDM), it is now possible to transmit different colors of light over the same fiber, which has provided another dimension to increasing bandwidth capacity and channeling raw data capacity into smaller chunks of bandwidth. This advance is akin to a self-expanding highway where you open another channel when the traffic load increases without laying a new fiber. Thus, WDM optical networks offer, among other capabilities, flexibility, scalability, and capacity. Current systems are capable of delivering more than one Gb/s/channel on an over-100-channel system. By 2010, a 100 channel capability, each at 10 Gb/s as illustrated in Figure 1.7 optical interconnections, is expected to provide terabit capacities. In addition to bandwidth increase, new technologies also enable more functionality in optics. If history is any indicator, these capabilities will only get better and richer in features. Optical networks are already being deployed, not only in the backbone of networks, but also in regional, metropolitan, and access networks. Thus, optics will play a key role in next-generation network modes and eventually at customers' premises.

1.2.4 Micro-Electro-Mechanical Systems (MEMS) Technology: The Fourth Technology Wave

Imagine a machine so small that it is imperceptible to the human eye. Imagine working machines with gears no bigger than a grain of sand. Imagine these machines being batch fabricated, tens of thousands at a time, at a cost of only a few pennies each. Imagine a realm where the world of design is turned upside down and the seemingly impossible suddenly becomes easy. Welcome to the microdomain—a world occupied by an explosive new technology known as MEMS. MEMS are the next logical step in the silicon revolution. We believe that the next step in the silicon revolution will be different and more important than simply packaging more transistors onto the silicon.

Samsung, EX1028, p. 11

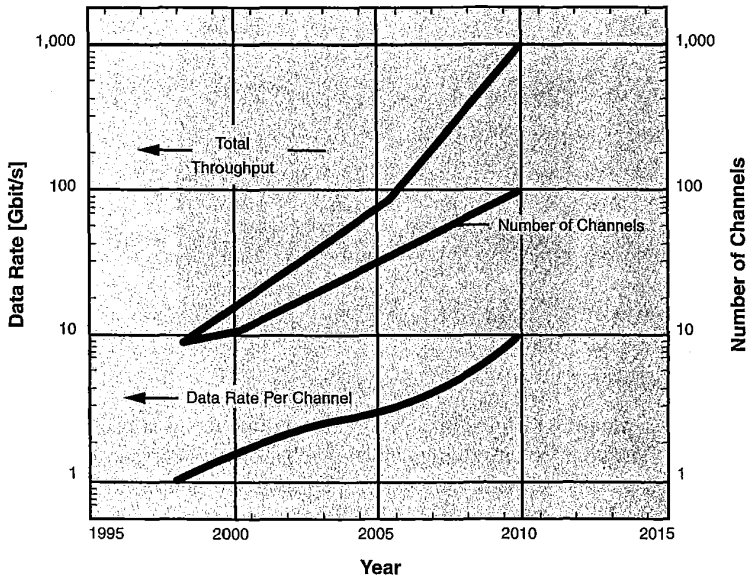


FIGURE 1.7 Potential of optoelectronics technology to terabits per second.

To the scientist or technologist, MEMS is like a dream come true. There is something magical about flicking on the electron microscope, zooming in and wandering through a micro-mechanical landscape that you conceived, created and understood, a world where the laws of nature don't behave as the layman expects. The micro-machinist's building site is the size of a grain of rice, the architecture the size of a human hair, the building elements smaller than a red blood cell, but our lithographic backhoes are the size of a city, so we can construct a thousand sites at once. We have mastered the art of building the impossible, from gyroscopes, micro-motors, gear trains and transmissions, to fluid pumps, x-y tables and entire self-assembling optical bench erector sets.

Enamored by this wondrous new frontier and its parallels with microelectronics, we technologists have declared MEMS as the fourth technology wave and the second semiconductor revolution. We have claimed unconditional applicability of Moore's Law and economic hockey stick curves, and many of us believe that MEMS will soon become as pervasive in all aspects of every day life as microprocessors are today.

Two factors drive Moore's Law in microelectronics: "smaller is better" and the "building blocks are universal across applications."

However, neither of these is particularly valid for MEMS. The second factor, about universal building blocks, is especially problematic. At the highest level of abstraction, the real power of microelectronics is not even its massively parallel fabrication paradigm. It is the existence of a generic element, called a transistor, which allows us to build extremely diverse functionality simply by implementing appropriate interconnection patterns within large collections of the generic elements. This is what makes semiconductor economics so vastly different from anything seen before in history. The impact of pushing the generic components along Moore's curve is therefore universal across all imaginable

Samsung, EX1028, p. 12

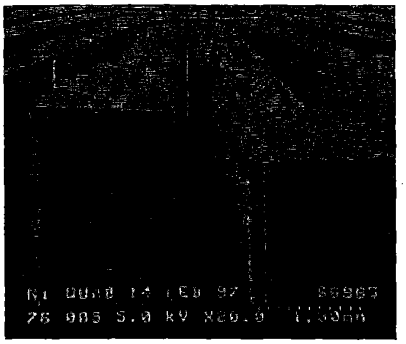


FIGURE 1.8 Example of a MEMS product.

application areas, which in turn justifies massive spending aimed at pushing even further along the curve. The gain factor in the financial feedback loop is greater than one because of the generic element paradigm.

In contrast, MEMS, by its very nature, does not have a set of generic elements. There is no MEMS transistor. MEMS "touches and participates" in the physical world of the mixed bag of applications and therefore needs to be much more application specific and less generic in every aspect of design, modeling, manufacturing, packaging, etc. Thus, it is much more challenging to keep the gain factor that drives Moore's Law greater than one. This is where many of the economic parallels with microelectronics break down, and economics is evidently what makes the difference between a possible future and a likely future for a technology. Figure 1.8 illustrates an example of MEMS.

1.3 WHAT IS MICROSYSTEMS PACKAGING (MSP)?

As the name implies, it includes three major technologies:

1. Microelectronics, Photonics, MEMS and RF Devices
2. Systems Engineering
3. Systems Packaging

Microelectronics typically refers to those micro devices, such as integrated circuits, which are fabricated in sub-micron dimensions and which form the basis of all electronic products. "IC" is an abbreviation for "Integrated Circuit" and is defined as a miniature or microelectronic device that integrates such elements as transistors, resistors, dielectrics, and capacitors into an electrical circuit possessing a specific function. "Systems" refers to all electronic products. "Packaging" is defined as the bridge that interconnects the ICs and other components into a system-level board to form electronic products. This view of microelectronic systems is depicted in Figure 1.9. The overlap of ICs and Packaging is referred to as *Packaged Devices* or *IC Packaging*.

An example of "packaged device" technology is today's microprocessors in your PC. The overlap of *Packaging* and *Systems* refers to incomplete or unintelligent system-level *Boards*, since these *Boards* do not contain the "brains" [the devices]. Finally, the overlap of *ICs* and *Systems* can be referred to as *Sub-Products*. These are considered sub-products because they perform a partial function of a system, limited by the magnitude of inte-

Samsung, EX1028, p. 13

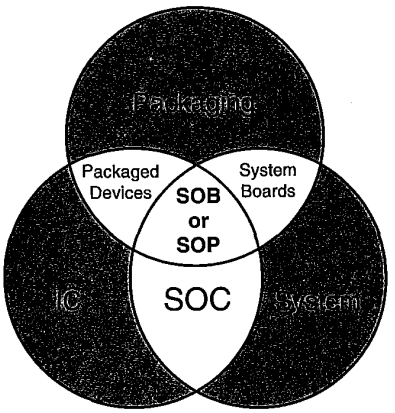


FIGURE 1.9 Integration of IC, packaging and system.

gration at the IC level and yet they typically don't involve extensive packaging. These "sub" or complete products depend heavily on the high integration of ICs without a dependency on packaging in order to meet a variety of product functions. In the future evolution of systems technology, this approach is predicted to evolve into a *system-on-chip* (SOC). A single chip radio is perhaps the best example of this. Most, if not all, products, however, are based on a number of packaged ICs and other components assembled onto a system-level board. This is referred to as *system-on-board* (SOB). A new paradigm called *system-on-package* (SOP), or *system-in-package* (SIP) is analogous to SOC, in that it is a single component, multi-function, multi-chip package providing all the needed system-level functions. These functions include analog, digital, optical, RF and MEMS. Both SOC and SOP are expected to be the wave of the future.

1.3.1 Electronics Systems Are Similar to Humans

The electronic product is like a human body. Electronic products have "brains" or microprocessors, and their packaging provides the "nervous" and "skeletal" systems. Therefore, note that without *packaging*, an electronic system is useless. It needs its packaging in order to be interconnected, powered or "fed," cooled via its "circulatory" system, and protected via its "skeletal" system. This is precisely what packaging is all about, as illustrated in Figure 1.10. There are other similarities such as the electrical and me-

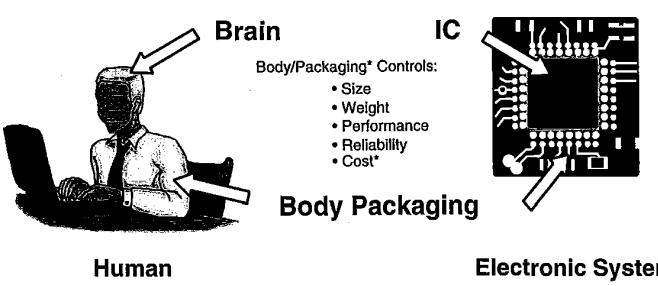


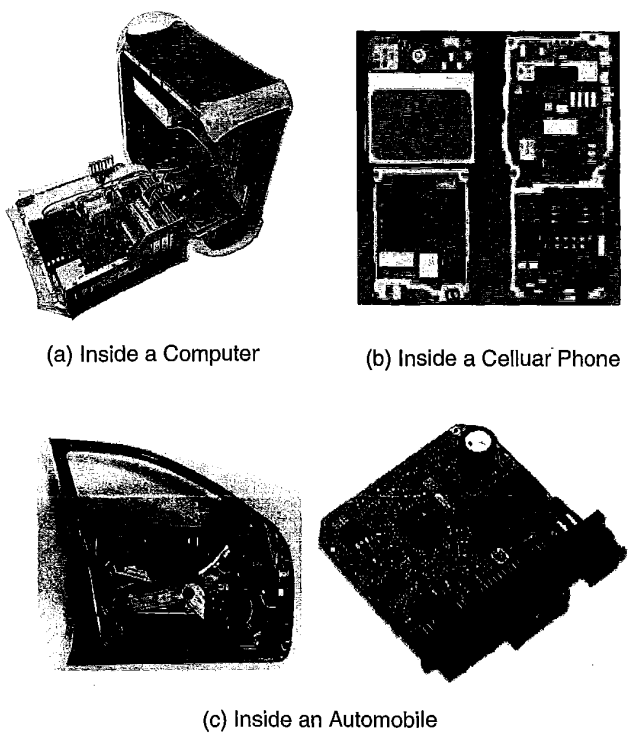
FIGURE 1.10 Electronic products are similar to humans.

Samsung, EX1028, p. 14

chanical transducers to those in the body, and the photodiode functions as the human eye. At the current IC integration rate following Moore's Law, it will be the year 2020 before the computer memory and processing technology catches up with the human brain.

1.3.2 Examples of Microsystems Packaging

Essentially, every electronic product contains: (1) semiconductor devices such as ICs, (2) packaging to integrate these ICs and other devices into components and (3) system-level boards which integrate these components to form the system-level assemblies that provide all functions required of the system. These functions are typically electrical, such as digital and analog, but the components providing these functions must, in turn, provide the needed mechanical and chemical functions. The microelectronics and packaging are integral parts of all these products. Figure 1.11 illustrates the system-level board packaging containing processor, memory and other ICs in personal computers, cell phones and automotives. Other products, not illustrated in this figure, include consumer, telecommunications, office automation, home appliances, entertainment, medical, and aerospace systems. Most of these products may contain other technologies such as:



(a) Inside a Computer

(b) Inside a Celluar Phone

(c) Inside an Automobile

FIGURE 1.11 Examples of systems packaging.

Samsung, EX1028, p. 15

- Magnetic and optical storage for storing information processed by the packaged ICs
- Displays (liquid crystal, flat panel, cathode ray tube, thin film transistor) for displaying information processed by packaged ICs
- Printers (laser, ink jet) for printing information processed by packaged ICs
- Fiber optics (silica fiber for telecommunications) for transferring information processed by packaged ICs

The central and fundamental building block of all electronic products therefore is packaged devices. Forming electronic products, however, requires a number of these packaged devices together with other components. This is referred to in this book as system-level packaging.

1.3.3 What Is the Relation between Information Technology (IT), Microsystems and Packaging?

The worldwide information technology (IT) and electronics market was at 1.2 trillion dollars in 2000, making it the largest industry and surpassing worldwide the agriculture, steel, transportation and automotive industries (Figure 1.12). Information technology, as illustrated, includes: 1) hardware such as microelectronics, photonics, RF/wireless and MEMS packaging, 2) software, 3) applications of hardware and software and 4) services such as electronic commerce. The evolution and growth of hardware and software (Figure 1.13) indicates the continual need for advancement of hardware on which to grow the software to make up the needed end products. The fundamental building blocks of these products are microelectronics, photonics, RF/wireless, MEMS and systems packaging involving all of these.

Of this 1.2 trillion-dollar market, the microelectronic and packaging market is approximately $250 billion (B)—about 25% of all IT and electronics, which includes software applications, hardware and Internet commerce.

Of this $250B, the microsystems packaging market is about 40% or $100B at system-level, and includes all the packaging technologies such as IC packages, printed wiring

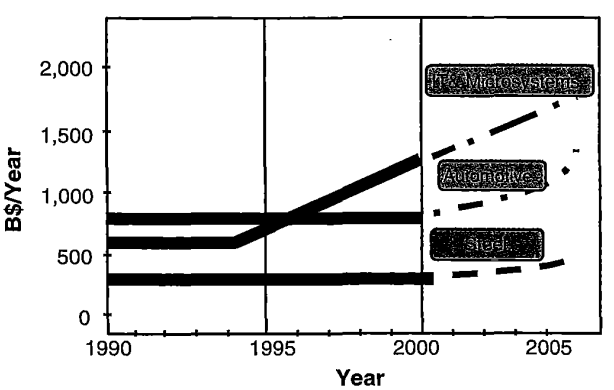


FIGURE 1.12 IT and Microsystems are the largest industry.

Samsung, EX1028, p. 16

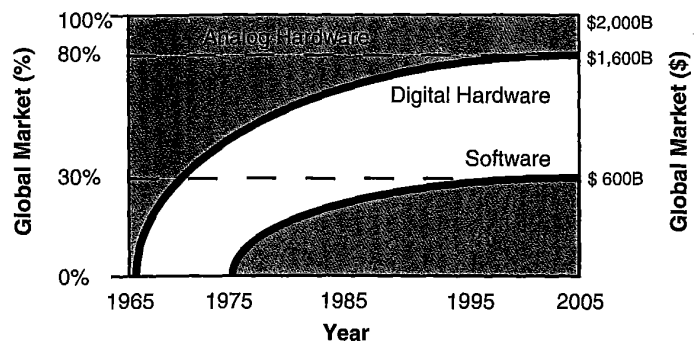


FIGURE 1.13 The evolution and growth of hardware and software segments.

boards, connectors, cables, optical, MEMS and RF packaging, heat sinks and others that constitute the total packaging solution. These markets are included in Table 1.2.

Today, the microsystems packaging technology, together with software, programming and system applications, is acting as the driving engine for leading-edge science, technology, advanced manufacturing and the overall economy of every country that participates in these. Together, these technologies are responsible for the largest industry and provide some of the highest-paying jobs in every corner of the world. Microsystems and Information Technology are considered by many to be the third industrial wave after agriculture and steel. The number of technologists employed in IT in the U.S. and worldwide is approximately 3 and 10 million, respectively.

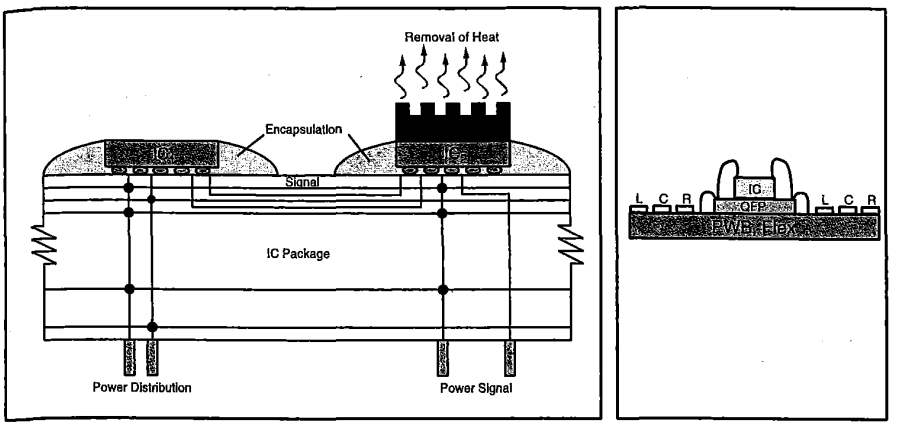
1.3.4 What Is IC and Systems Packaging?

Microsystems packaging involves two major functions: one at the IC or device level, and the other at the system-level, as shown in Figures 1.14a and 1.14b. At the IC level, it involves interconnecting, powering, cooling and protecting ICs. At this level, typically referred to as *Level 1*, the packaging acts as an IC "carrier." The IC carrier, also called Packaged IC, allows ICs to be shipped "certified or qualified" by IC manufacturers after

TABLE 1.2 Worldwide systems packaging market ($ billions).

Printed Wiring Boards	30
Flex Circuits	3.2
Assembly Equipment	3.3
Materials	9.9
Connectors	23.4
Optoelectronics Packaging	10
RF Packaging	1.2
Passive Components	25
Thermal and Heat Sinks	3
Total	**109**

Samsung, EX1028, p. 17



(a) IC Packaging (b) Systems Packaging

FIGURE 1.14 (a) IC packaging. (b) Systems packaging.

"burn-in" and electrical test to be "ready" for assembly onto a system-level board by end product or contract manufacturers.

Packaging a single IC does not generally lead to a complete system since a typical system requires a number of different active and passive devices. System-level packaging involves interconnection of all these components to be assembled on the system-level board, regardless of the type of component being assembled. The system-level board, also called "motherboard," not only carries these components on top and below, but also interconnects every component with conductor wiring so as to form one interconnected system. This system-level board is typically referred to as *Level 2* in the *Packaging Hierarchy*.

In forming an electrically-wired system-level board with assembled components, there are two additional interconnections that need to be made. First, interconnection must occur at the *IC level* where the input/output (I/O) pads on the IC are connected to the first level of the packaging. This is typically done by wire bonding the components to a lead frame that has been fabricated to a specific shape in order to make it ready for interconnection to the next level of packaging. This is referred to as IC assembly. The second interconnection is typically achieved by means of solder bonding between the lead frame of the first-level package and electrically conductive pads on the second-level package, which is typically a "card" or "board." This is referred to as board assembly. The system-level board, with components assembled on either or both sides, typically completes the system.

There are products, such as mainframes and supercomputers, that require a very large number of ICs. By today's standards, a single system-level board may not carry all the components necessary to form that total system, since some of these require several processors to provide the extremely high transactional throughput. These types of systems might be used to manage large amounts of data such as an airline reservation system or a corporate mainframe network, or process high-resolution imagery such as with certain types of medical equipment. In this case, connectors and cables typically connect the

Samsung, EX1028, p. 18

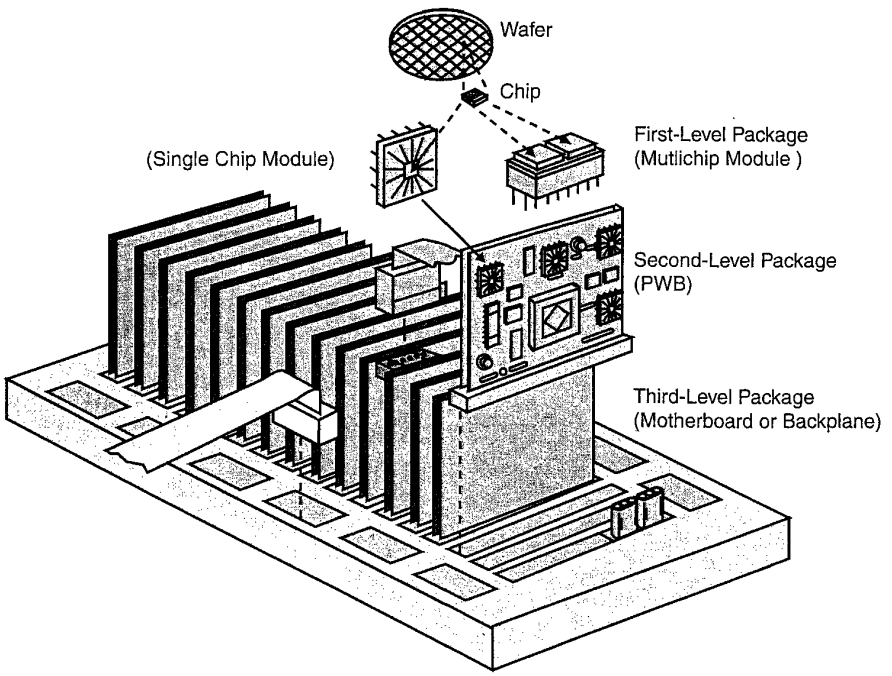


FIGURE 1.15 Packaging hierarchy.

several boards necessary to make the entire system. This is referred to as *Level 3* of Packaging. The *Three-Level Packaging Hierarchy* is illustrated in Figure 1.15.

1.3.5 Systems Packaging Involves Electrical, Mechanical and Materials Technologies

It should be recognized that in this three-level hierarchy, a transistor on an IC might communicate by means of an electrical or optical signal to another IC. This signal communication poses a whole set of electrical, mechanical, thermal, chemical and environmental challenges which, if not properly engineered and manufactured, may result in either poor communication or no communication at all. This is illustrated in Figure 1.16.

Electrical Packaging Technology

Electrical problems relate to both signal propagation between the transistors and to power distribution required to operate these transistors. The electrical parameters such as resistance, capacitance and inductance are always present and cause signal delays and signal distortions. Signal degradation is another problem that is due primarily to line resistance. Line resistance causes a voltage drop, thus increasing transition time. The power distribution problems stem from simultaneous switching of all the driving transistors in a given circuit, resulting in drawing a huge amount of current. This is referred to as "switching noise."

Samsung, EX1028, p. 19

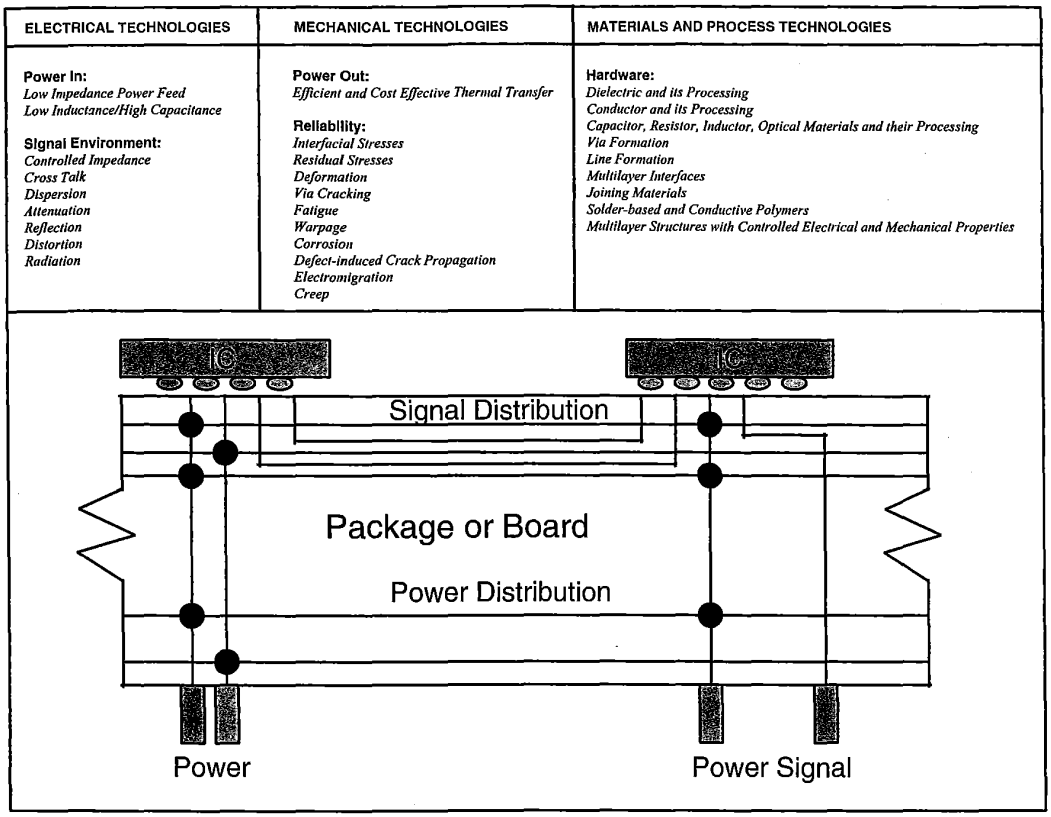


FIGURE 1.16 Systems packaging involves electrical, mechanical and materials technologies.

Since an electronic system involves more than one IC, effective communication between one transistor on one IC and another transistor on another IC, all the way through system-level board with the required signal quality, is required. Signal communication, however, does not start until an appropriate power is supplied to each and every transistor. Power distribution, however, poses a whole set of challenges that include voltage drop as a result of long and high resistive wiring from the power supply to the transistor through all the levels of packaging. Simultaneous switching of millions of transistors poses yet another challenge, in that the current drawn from the power supply results in what is referred to as delta-I noise. Signal distribution poses a different set of problems such as "cross-talk" between lines, as well as distortion, reflection and alternation of signals. Electromagnetic radiation as a result of all this radiated energy is another electrical challenge.

Materials Packaging Technology

The signal and power distribution requires appropriate use of materials to form the system-level packaging hierarchy. Power distribution, for example, requires metals of

Samsung, EX1028, p. 20

highest electrical conductivity for least voltage drop. Heat transfer requires materials of highest thermal conductivity. The minimized delta-I noise requires low inductance and high capacitance power distribution. High performance computers require high-speed signal propagation, which requires the use of lowest dielectric constant dielectrics in which to embed the best electrical conductors. Materials are also required to join ICs to packages to form IC packages as well as to join materials to form precise electrical structures with the required impedance, capacitance, resistance and inductance.

Mechanical Packaging Technology

The combination of power distribution through all levels of system packaging, and the use and fabrication of materials with the above diversity of properties, invariably lead to the development of thermomechanical stresses at every interface. These stresses, which develop not only during fabrication of IC and system-level packages, but also during shipment of product in hot and cold climates and during actual product usage, could lead to electrical failure of interconnections. Effective heat transfer, so as to keep the IC and the system-level packaging "cool," is one way to address the challenge.

The mechanical problems typically relate to reliability of the packaging structure that supports the electrical function. This occurs particularly at solder-to-chip interface and package-to-board interface during processing and fabrication of the IC packages and system-level boards. It also occurs during electrical operation of final electronic products. In both cases, stresses are developed due to the combined effect of mismatch in thermal expansion coefficients between various interfaces and temperature.

1.4 WHY IS MICROSYSTEMS PACKAGING IMPORTANT?

IC is not a microsystem, and no microsystem is complete without systems packaging. The importance of packaging, however, differs from one type of microsystem to another. The following is a summary of the importance of systems packaging.

1.4.1 Every IC and Device Has to Be Packaged

There are currently 60 billion ICs and devices that are manufactured worldwide, and all of these have to be packaged at the IC-level to form IC packages, and at system-level to form, system-level boards. Packaging at both levels is often considered the biggest bottleneck, because it controls the system's **electrical performance, cost, size** and **reliability.**

1.4.2 Controls Performance of Computers

The number of ICs and their interconnections required to form a processor or central processing unit (CPU) determine the cycle-determining path from IC through the package interconnections, and thus controls the speed or clock frequency of the CPU.

1.4.3 Controls Size of Consumer Electronics

The number and size of ICs in a given system, such as a cellular phone, tend to be small. However, it is the two levels of IC package and system-level hierarchy, involving passive,

Samsung, EX1028, p. 21

microwave, switch, relay and other components that form the bulk of the cell phone's size.

1.4.4 Controls Reliability of Electronics

Solid-state devices such as ICs are extremely reliable, with failure rates in parts-per-million (ppm). Since the majority of interconnections in a system are within packaging at IC package and system board levels, the failure rate tends to be more highly attributed to the packaging of the devices rather than to the devices themselves.

1.4.5 Controls Cost of Electronic Products

The cost of producing today's ICs and MEMS devices is low due to a variety of factors such as large-scale and high throughput wafer starts-per-day and automation. The approximate IC fabrication cost, excluding design cost, is about $\$4/cm^2$ at mature production levels. On the other hand, system-level packaging cost, with all the packaging components to form system-level boards, is much higher.

1.4.6 Required in Nearly Everything

Electronics are now a part of nearly all industries such as automotive, telecommunication, computer, consumer, medical, aerospace and military.

1.5 SYSTEM-LEVEL MICROSYSTEMS TECHNOLOGIES

Figure 1.17 and Table 1.3 illustrate all the critical microsystems and packaging technologies required to form the system-level board for most electronic products. These technologies form the basis of this book.

The master figure illustrated here in Figure 1.18 and in every chapter includes the integration of all the technologies to form system-level boards. This master figure includes IC and packaging, and the relationship between IC and packaging in chapter 2; microsystems and packaging, and the relationship between the two in chapter 3. The electrical design and design for reliability are presented in chapters 4 and 5. Thermal management is presented in chapter 6. The rest of the book is divided into three parts: 1) the technologies required to assemble the microelectronic, photonic, RF/wireless and MEMS devices in chapters 7 to 10; 2) the individual device technologies in chapters 11 through 14; and 3) system board technologies in the remaining chapters.

1.5.1 Science and Engineering Disciplines in Microsystems Packaging

Microelectronic packaging is perhaps the most cross-disciplined of all technologies, not only in the Information Technology industry, but also across all industries. It includes:

Physics: Physicists deal with fundamentals of signal speed, whether by electrons or photons, and their combinations called optoelectronics.

Chemistry: Chemists deal with such fundamental issues as synthesis, structure and properties of polymers and other materials.

Samsung, EX1028, p. 22

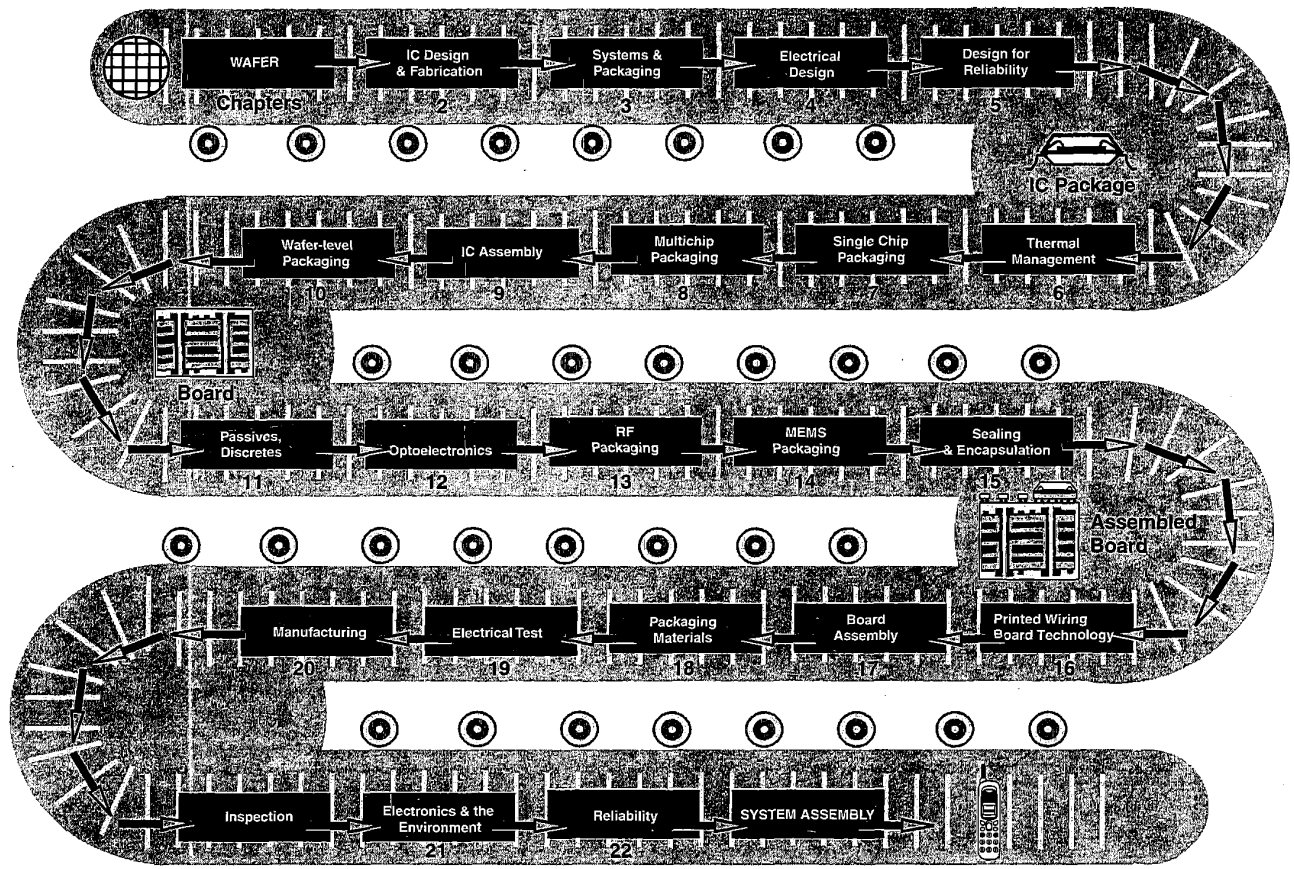


FIGURE 1.17 Critical microsystem packaging technologies from wafer to system-level board.

Samsung, EX1028, p. 23

TABLE 1.3 Microsystem technologies in a product and book chapters representing these technologies.

Semiconductor Technology (Chapter 2)	Optoelectronics Technology (Chapter 12)
Systems Technology (Chapter 3)	RF Packaging Technology (Chapter 13)
Electrical Design Technology (Chapter 4)	MEMS Packaging Technology (Chapter 14)
Design for Reliability (Chapter 5)	Sealing and Encapsulation Technologies (Chapter 15)
Thermal Management Technology (Chapter 6)	Printed Wiring Board Technology (Chapter 16)
Single Chip Packaging Technology (Chapter 7)	Board Assembly Technology (Chapter 17)
Multichip Packaging Technology (Chapter 8)	Materials, Processes and Properties (Chapter 18)
IC Assembly Technology (Chapter 9)	Electrical Test Technologies (Chapter 19)
Wafer-Level Packaging Technology (Chapter 10)	Manufacturing Technologies (Chapter 20)
Discrete, Integrated and Embedded Passive Technologies (Chapter 11)	Environmental Technologies (Chapter 21)
	Reliability Technologies (Chapter 22)

Electrical Engineering: Electrical engineers deal with signal and power distribution issues.

Computer Engineering: Computer engineers deal with design tools as well as system technologies.

Mechanical Engineering: Mechanical engineers deal with thermomechanical design, mechanical integrity, heat transfer and thermal management, MEMS and manufacturing.

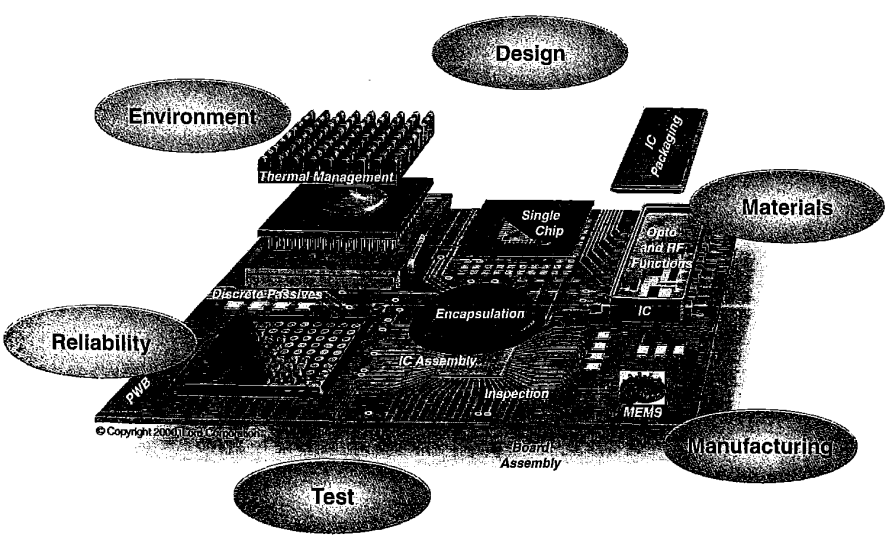


FIGURE 1.18 Master figure illustrating all microsystem technologies.

Samsung, EX1028, p. 24

Materials Science and Engineering: Materials scientists deal with material design, synthesis and characterization of metals, alloys, ceramics and polymers.

Chemical Engineering: Chemical engineers deal with chemical processing of metals, polymers and ceramics.

Manufacturing Engineering: Manufacturing engineers deal with cost effective manufacturing processes.

Business, Economics and Management: These professionals deal with business plans, resources and management of manufacturing facilities for cost-effective products.

Environmental Engineering: These engineers deal with the impact of microsystems design, chemical processing, usage and disposal of the product in the environment.

1.6 WHAT IS EXPECTED OF YOU AS A MICROSYSTEMS ENGINEER?

1.6.1 Engineering Professionalism and Ethics

Society has placed engineers in a position of trust, since products designed and manufactured under an engineer's supervision have the potential to do great harm if they are not designed, manufactured and used properly. For this reason, an engineer is expected to adhere to high ethical standards. Various branches of engineering have developed codes of ethics to address fundamental issues. The preamble to the code of ethics of the Institute of Electrical and Electronics Engineers (IEEE) is reprinted here.

1.6.2 IEEE Code of Ethics

Preamble

Engineers, scientists and technologies affect the quality of life for all people in our complex technological society. In the pursuit of their profession, therefore, it is vital that IEEE members conduct their work in an ethical manner so that they merit the confidences of colleagues, employers, clients and the public. This IEEE Code of Ethics represents such a standard of professional conduct for IEEE members in the discharge of their responsibilities to employers, to clients, to the community and to their colleagues in this Institute and other professional societies.

1.7 SUMMARY AND FUTURE TRENDS

Figure 1.19 summarizes the microsystems packaging as starting with a wafer and ending up with a finished system like a cellular phone. This is a very good example of technologies and systems in the 20th century.

1.7.1 So What Next?

Albert Einstein had defined time and space into a single variable at the turn of the century. The resulting energy mass equivalence led to profound changes in physics, which in turn gave mankind a better understanding of sub-atomic phenomena, as well as those on the

Samsung, EX1028, p. 25

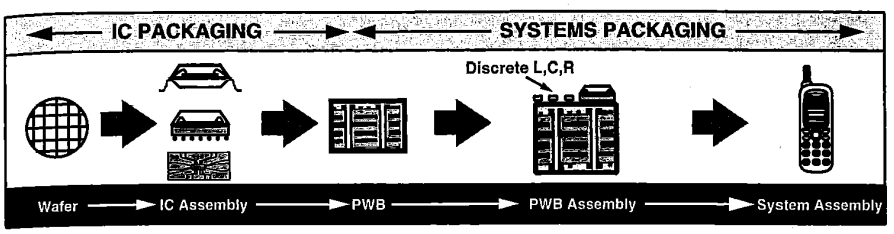


FIGURE 1.19 Summary of microsystems packaging.

cosmological scale. The technology arising out of this science has been largely used for the benefit of society.

As we get ready to enter the next century, yet another fundamental redefinition of time and space called the Internet promises to bring about unprecedented changes in society. Commonly known as the Net, it is an interconnection of computer and communication networks spanning the entire globe, crossing all geographical boundaries. Touching lifestyles in every sphere, the Net has redefined methods of communication, work, study, education, interaction, leisure, entertainment, health, trade and commerce. There now is a telecommuting global work force in redefined time and space. The Net is changing everything. From the way we conduct commerce, to the way we distribute information. Being an interactive two-way medium, the Net, through innumerable websites, enables participation by individuals in business-to-business, and business-to-consumer commerce, visits to shopping malls, bookstores, entertainment sites, and so on, in cyberspace.

What Are the Drivers of This Information Age? The Primary Drivers Are Microsystems Technologies and Markets

Labor and capital, which were the paramount assets of the industrial age, stand replaced by knowledge as the most important asset to be managed by businesses. The key business characteristics of industrial age and information age businesses are included in Table 1.4.

The Information Age is, thus, a *knowledge-based industrial revolution.* Information technology is used and companies are networked. Discovery and innovation are perceived to be more important to competitiveness than simply manufacturing.

TABLE 1.4 Characteristics of information age.

20th Century Industrial Age	21st Century Information Age
1. Mass production	Mass customization
2. Labor serves tools	Tools serve labor
3. Labor performs repetitive tasks	Labor applies knowledge
4. Command and control structure	Common control structure
5. Capital intensive	Knowledge intensive
6. Capitalists own production means	Labor owns production
7. Capital is primary driver	Knowledge is primary driver

Samsung, EX1028, p. 26

So What Are the Fundamental Building Block Technologies Behind this Revolution?

They are giga scale microelectronics, giga Hertz RF and wireless, terabit optoelectronics and micro-sized motors, actuators, sensors and medical implants, and most importantly the integration of all these into microsystems by microsystems packaging.

These are the subjects of this book.

1.8 WHO INVENTED MICROSYSTEMS AND PACKAGING TECHNOLOGIES? (Excerpted and adapted with permission from IEEE Spectrum, June 2000)

The microsystems products that we see today are a result of generations of discoveries, one built over the other, starting with Edison's light bulb. Some of the most important discoveries in microelectronics, photonics, RF/wireless, MEMS and systems packaging are outlined below. Those that are closely related to microsystems and packaging are in color.

1879 Thomas Alva Edison first displayed his iridescent **electric light bulb,** a glass tube with a conducting filament mounted in a vacuum. In 1883, Edison detected electrons flowing through the vacuum, but it was an English physicist, John Ambrose Fleming, who, in 1904, used this information to develop a diode vacuum tube.

1890 The population was booming in the U.S., and the Census Bureau realized it could not manage its headcounting duties. A contest was arranged to encourage inventors to propose a solution. Herman Hollerith won top prize with his **punch-card machine.** This is considered to be the grandfather of today's computers. Hollerith later formed the Tabulating Machine Company, which later became IBM.

1897 British physicist, Joseph John Thomas, declared that cathode rays were made up of negatively charged particles, which he called corpuscles. Once his proclamation was proved true, he received credit for **discovering the electron,** and was awarded a Nobel Prize in 1906.

1898 Danish scientist, Valdemar Poulsen, invented the telegraphone and early telephone-answering machine. This was the **first magnetic recording device.** The public waited another century for the answering machine to be augmented with the next innovation—caller ID.

1899 The commissioner of the U.S. Patent Office declared, "Everything that can be invented has been invented."

1900 Urged on by Charles P. Steinmetz, a pioneer in the scientific understanding of electric power, **General Electric established a research laboratory** in Schenectady, N.Y. Willis R. Whitney, its director until 1928, was credited with **key improvements in the incandescent bulb.**

1901 In December, Gulielmo Marconi, in St. John's, Newfoundland received the **first wireless message to cross the Atlantic.** Sent from Poldhu, Cornwall, in England, his message was the letter S—three dots in Morse code. The demonstration of the transatlantic reception over 2900 km helped Marconi establish the business of wireless telegraphy. The originator of the numerous innovations, including a method of continuous-wave transmission, as well as grounded antennas,

Samsung, EX1028, p. 27

improved receivers, and receiver relays, Marconi was also remarkable for his skills at marketing and promoting. In 1897, he had established a company that soon offered radio communications services, notably to shipping lines, though the transmission range was initially limited—some 240-km in 1900. By World War I, Marconi Companies in Britain and elsewhere were providing radio communications worldwide. This earned Marconi the Nobel Prize for Physics.

1904 Christian Hûlsmeyer, a German inventor fascinated by hertzian waves, was ahead of his time. One of many contributors to the development of electromagnetic waves for wireless communications, he got an idea for a different application: seeing ships through fog and darkness by transmitting waves and detecting the echoes. On April 30th, he applied for a German patent on a "means for reporting distant metallic bodies to an observer by use of electric waves." Though he demonstrated a range of 3000 meters for his Telemmobiloscope, neither naval nor shipping leaders were interested.

Around 1930, Hûlsmeyer's idea was taken up again or independently arrived at. At least eight countries developed **radar systems,** though for warning of air attack rather than for ship navigation. The term radar, an acronym for radio detection and ranging, was not proposed until 1940.

Also in 1904, English engineer, John Ambrose Fleming, **invented the diode,** a two-electrode vacuum tube used to rectify a wireless signal, so it could be detected by a galvanometer or telephone receiver. The diode was comprised of a heated electrode—and electron emitter—and a cold electrode that received the electrons in an evacuated glass tube.

Two years later, in the United States, Lee de Forest made a crucial improvement: interposing a cold grid-like electrode between the two others. This allowed control of the flow of electrons from the heated electrode. Calling it an **audion** (later **amplifier**), de Forest referred to it as a "device for amplifying feeble electrical currents" but until 1912 used it only for detecting radio waves.

1911 A mega-merger of three businesses resulted in the Computing Tabulating Recording Company. Thirteen years later, the name was changed to **International Business Machines,** signifying a focus away from coffee grinders and other food processing devices. The company then focused on selling calculation machines around the globe.

In February of 1911, Charles Kettering, an electrical engineer who had earlier electrified the cash register, demonstrated a self-starter on a Cadillac. Until then, gasoline engines had been started by hand cranking, a taxing method. Because of their simpler starting, battery-powered cars had developed a niche market for themselves, particularly among city women. But Kettering's **self-starting motor** now made gas engines simple to start, and with the engine running, operated as a generator to recharge the battery and power the headlamps. In November, Cadillac ordered 12,000 systems for its 1912 model, marketing them as "ladies' aid." But electric vehicles held onto another niche market, as delivery trucks for city use, well into the 1920s.

1912 Engineers were coming to realize that the three-electrode vacuum tube had other uses besides detecting radio waves. Fritz Lowenstein and de Forest, in the United States, as well as Robert von Leiben and Otto von Bronk, in Germany, saw that it could amplify weak signals and work as an **oscillator.**

Samsung, EX1028, p. 28

These functions were soon put to use. The triode was designed into telephone repeaters in several countries; a Western Electric repeater went into service between New York City and Philadelphia in October 1913. Through World War I, triode oscillators generated signals for radio transmitters and were used in radio receivers for heterodyne reception.

One of the most fundamental circuit inventions of the 20th century is the **regenerative circuit,** in which some of the output of an electron tube is returned to the input. In 1912, Edwin Howard Armstrong, a student at Columbia University in New York City, found he could obtain much higher application from a triode by transferring a portion of the current from the plate to the signal going to the grid. He also found that increasing this feedback beyond a certain level made the tube into an oscillator, a generator of continuous waves.

1918 Armstrong's invention of the **superheterodyne receiver** was a great advance. Essentially, the incoming high-frequency signal is converted to a fixed intermediate-frequency by heterodyning, or mixing, it with an oscillation generated by an electron tube in the receiver. Next, the intermediate-frequency signal, which always falls in the same frequency band, is amplified before it is subjected to the usual detection and amplification that produces the audio signal. The scheme offered improved sensitivity, as well as tuning by turning a single knob. The superheterodyne soon became, and remains today, the standard type of **radio receiver.** RCA marketed the first one in 1924. Actually, the heterodyne principle was introduced into radio, then called wireless, by Reginald Fessenden in 1901.

The phenomenon was well known and exploited by piano tuners: if two tones of frequencies A and B were combined, the listener hears A minus B. Fessenden suggested that it be employed in a radio receiver: the incoming radio frequency wave would be mixed with a locally generated wave of slightly different frequency, the combined wave then driving the diaphragm of an ear-piece at radio frequencies. Lacking an effective and inexpensive local oscillator, Fessenden had been unable to make a practical heterodyne receiver.

1920 On November 2nd, **KDKA, the first station licensed for general broadcasting** service, began operations, reporting election returns in the U.S. presidential race between Warren Harding and James Cox. Its 833-kHz, 100-W transmitter sat in a shack atop the roof of a Westinghouse Electric building in East Pittsburgh, PA. The station was the brainchild of Frank Conrad, a Westinghouse engineer and amateur radio operator.

He had supervised the manufacture of portable transmitters and equipment for the U.S. Army during World War I. After the war, he began playing phonograph music over his radio transmitter once a week, and then more often. The interest this aroused gave Westinghouse the idea of promoting such broadcasts to stimulate the sale of its radio receivers. One New York City radio station, WEAF, credits the first sale of airtime for a commercial message. On August 28, 1922, it broadcast a message for a real-estate developer.

1922 Although there were various levels of success in developing television, Philo Farnsworth is credited with an electronic design that was licensed to RCA. At long last, **commercial TV** was finally demonstrated to the public in 1939 at the New York World Fair. TV commercials were not long behind.

Samsung, EX1028, p. 29

1926 Western Electric developed a **sound-on-disk system for motion pictures** called **Vitaphone,** debuted by Warner Brothers on August 6th with *Don Juan,* followed on October 6th, with *The Jazz Singer* starring Al Jolson. Their reception was so enthusiastic that in the next year and a half the U.S. film business converted to sound movies, though sound-on-film systems eventually prevailed.

1927 Long-distance telephone service often required several stages of amplification, which introduced distortion. On his way to work on August 2nd, Bell Telephone Laboratories' engineer Harold S. Black had the idea that he could reduce distortion by feeding back some of the amplifier's output, in negative phase, to the input.

This idea was counter-intuitive and almost the opposite of Armstrong's idea of the regenerative circuit with its positive feedback. Although negative feedback lowers gain, it improves other amplifier characteristics, including flatness of response, and the invention is widely used in communications and controls systems.

Mervin Kelly, president of Bell Labs, wrote in 1957: Black's **negative feedback amplifier** "ranks . . . with de Forest's invention of the audion as one of the two inventions of broadest scope and significance in electronics and communications in the past 50 years."

1930 The Galvin Manufacturing Corporation offered the **first practical automobile radio,** sold as an accessory from car dealerships. The company's name is later changed to Motorola in an effort to link "motion" and "radio."

1931 On November 21st, AT&T inaugurated its **Teletypewriter Exchange Service** (TWX), which provided central switching, so subscribers could **communicate** with each other **by Teletype.** Introduced in the mid-20s, the teletypewriter required no skill in Morse code, so anyone could send and receive messages.

By 1937, TWX connected 11,000 stations. To compete, Western Union introduced its Telex Service in 1958. Some years later AT&T sold TWX to Western Union.

1932 Karl Jansky of Bell Telephone Laboratories was asked to investigate the sources of noise interfering with transatlantic radio transmissions. Writing in the December Proceedings of the IRE, he distinguished three types: noise from local thunderstorms, steadier and weaker static from distant storms, and a weak hiss of unknown origin. In a Proceedings article in October 1933, Jansky presented evidence that this last type of static came from outside the solar system. Bell Labs rejected his suggestion that a **30-meter dish-shaped antenna** be built for further investigation of this source. Only after World War II did **radio astronomy** became a recognized field of study.

1935 On February 12, Robert Watson-Watt of the British National Physical Laboratory sent a memorandum to the Air Ministry entitled "Detection of aircraft by radio methods" and later referred to it as **"the birth certificate of radar."** Two weeks later Watson-Watt demonstrated that a distant aircraft reflected radio waves transmitted toward it, which led to a major development effort. By the outbreak of war in September 1939, the British had a 25-station radar network, called Chain Home.

The Germans, too, had radar. It appears that the first one used in combat was their Freya radar, which detected 22 Wellington bombers approaching Wilhelmshaven on December 18, 1939 and guided defending Luftwaffe aircraft to

Samsung, EX1028, p. 30

them. Radar came to play many key roles in the war, including aircraft detection, ship and submarine detection, fire control, bombing, and navigation.

1936 In November, the British Broadcasting Corporation (BBC) tested two television systems: John Logie Baird's partly mechanical system, which used a spinning disk and EMI's all electronic one. The later proved far superior. Earlier in the year, Telefunken provided **TV coverage of the Berlin Olympics,** mainly to public television-viewing rooms in German cities.

In the United States, from the grounds of the **New York World's Fair,** RCA began experimental **TV broadcasting** on February 26, 1939. On April 30, 1941, the Federal Communications Commission approved broadcast TV standards recommended by the National Television Systems Committee—525 lines, 30 frames per second. World War II, however, put regular broadcasting on hold.

1937 The centimetric, or **short-wavelength, radar** of World War II depended upon two technological marvels: klystrons and cavity magnetrons, electron tubes capable of producing high-frequency oscillations. On June 5, Russell Varian conceived of **the klystron,** which achieves amplification by "velocity modulation." A stream of electrons are made to group themselves into bunches; these bunches then constitute a driving current for a resonant cavity, in which they are amplified.

Russell and his brother Sigurd, with William Hansen, built a tube which oscillated at 2.3 GHz, a wavelength of 13 cm. The klystron gave Englishmen Henry Boot and J.T. Randall the idea of introducing a resonant cavity into a magnetron; they built the **first cavity magnetron.** In most wartime radar sets, magnetrons produced the outgoing signal, but klystrons were the local oscillator for mixing with the reflected, incoming signal, since they were easier to tune.

1938 On October 22, Chester Carlson, an inventor living in the Astoria section of New York City, produced the **first eletrophotographic image.** It read "Astoria, 10-22-38." Commercialization proved difficult. In 1947 Haloid, a small maker of photographic paper in Rochester, N.Y., bought the rights and developed the invention, naming it xerography from xeros, the Greek word for dry. Early machines, which used special paper, did not sell well. But the first plain-paper copier, the Model 914 introduced in 1959, achieved rapid success. Such machines have since changed office practices everywhere.

1939 Siemens and Halske began selling electron microscopes in Germany. In the early 1930s, the Germans Ernst Ruska and, independently, Reinhold Rudenberg, invented the device, in which an electron beam achieves higher resolution than possible with light waves. Siemens and Halske hired Ruska for its development effort. Other developers of commercial machines included Philips, General Electric, and RCA.

Also in 1939, Otto Hahn, working in Berlin, and Lise Meitner, a refugee in Scandinavia from Nazi Germany, established that the atom can be split. Within two years, programs were established in Germany, Great Britain, the United States, and Russia to explore the possibility of building an atomic bomb.

1940 Motorola developed the **first hand-held two-way radio** for the U.S. Army Signal Corps. The portable "Handie-Talkie" AM radio became a World War II symbol.

1943 A **practical means of making printed circuits** was patented by electrical engineer Paul Eisler on February 3rd. He had fled to England in 1936, from anti-

Samsung, EX1028, p. 31

Semitism in Austria. He realized that electronic devices of all kinds were vital to the war effort and so renewed his earlier efforts to emulate printing technology in making circuits. The wiring pattern was printed onto a conducting sheet using an acid-resistant "ink;" the unwanted conductor was then dissolved. During the war, the U.S. military, not the British, took up the technology, using it extensively in making proximity fuses.

1945 Eniac, the electronic numerical integrator and computer, became operational in November. Designed mainly by Presper Eckert and John Mauchly, it was, with some 18,000 electron tubes, much more complex than any previous electronic devices. The U.S. military had sponsored the design and construction of the computer, intended to calculate ballistic tables for the paths of artillery shells. But Eniac could be wired, through hundreds of switches and patch cords, to perform many other tasks.

In a memorandum dated November 8, John Von Neumann presented the basic design of the digital stored-program computer, which carries out different algorithms without having to be re-wired. One of the most influential figures in 20th-century mathematics and computing, von Neumann was born in Budapest in 1903 and attended universities in Hungary, Germany, and Switzerland, obtaining his Ph.D. in 1926 at the age of 22. Four years later he moved to the United States.

In the '20s and '30s, he made important contributions to set theory, algebra, and quantum mechanics. During World War II, he played large roles in several projects, including the atom bomb project and the Eniac computer. After the war, von Neumann focused mainly on developing electronic computing, notably through the computer project he directed at the Institute for Advanced Study in Princeton, NJ.

1947 On December 23, John Bardeen and Walter Brattain demonstrated their invention of a **solid-state amplifier** at Bell Telephone Laboratories. It was the point-contact transistor, whose active part was the interface between metal leads and a germanium semiconducting crystal.

In 1947, Bell Labs' employees, John Bardeen and Walter Brattain, demonstrated a **point-contact transistor,** which did the work of a vacuum tube with less heat and without the tube, the vacuum with or without high voltage. As is sometimes the case for monumental events, their manager, William Shockley, is often given credit for this first transistor, although his junction transistor came a few weeks later. It took six months for Bardeen and Brattain to file for the patent; the same length of time it took for the public relations department to arrange for the press release. The Nobel Prize for their work on transistors came much later to all three researchers, in 1972.

In 1949, their project leader William Shockley proposed a new type of transistor that exploited semiconducting properties in the bulk of the crystal. Such a transistor required fabricating a crystal with a sandwich structure: two layers of n-type semiconductor, in which conduction occurs by movement of excess electrons, separated by a layer of p-type semiconductor, in which conduction occurs through the movement of vacancies in the electron structure of the crystal, called holes. Shockley, Morgan Sparks, and Gordon Teal demonstrated such a **germanium transistor** in April of 1950. Within a year or so, Bell Labs engineers had turned it into a practical, reliable, and manufacturable device, and on September 25, 1951 AT&T began offering manufacturing licenses for a nominal fee.

Samsung, EX1028, p. 32

1950 William Papian, under the direction of Jay Forrester, both at the Massachusetts Institute of Technology, built the **first magnetic-core memory,** a two-to-two array, in October 1950. On August 8, 1953 magnetic-core memory was installed for the first time in the university's Whirlwind computer, and the first commercial system to use the memory was the IBM 705 in 1955. Core memory was the standard random-access memory in computers until superseded by IC memory in the mid-70s.

1951 Presper Eckert and John Mauchly delivered the first **Univac** to the Bureau of the Census on June 14. After working on the Eniac and a stored-program computer, the Edvac, they had set up their own computer company and began work on the more powerful Univac **(Universal Automatic Calculator).** Remington Rand acquired the company in 1950. The machine attained fame in November 1952 when the Columbia Broadcasting System's TV network used it to predict the outcome of the presidential election on the basis of early returns.

1954 ASEA of Sweden completed a **landmark high-voltage direct-current power line for transmitting electric power,** a technology with advantages over AC transmission in certain situations, where underground or underwater cables are required, for instance, or where power systems of different frequencies are to be interconnected. The project connected the island of Gotland with the Swedish mainland by a 98-km submarine cable.

1955 Sony introduced the **first transistor radio** in Japan, marking the beginning of yet another reason for kids to get grounded—playing music too loud.

1956 On November 30, the first on-the-air use of a videotape recorder was to broadcast "Douglas Edwards and the News" on CBS. The TV networks had wanted an easy way to rebroadcast previously transmitted or recorded programs, especially because of time-zone differences. A film process called kinescope, then in use, was costly in time and labor, and the picture quality was often poor. Charles P. Ginsburg and Ray Dolby, working for Ampex, developed the **prototype videotape recorder** that used magnetic tape.

Motorola introduces a new radio-communication product. A small radio receiver delivers a radio message to an individual carrying the device. Doctors are the first to get beeped since **pagers** were initially used in hospitals.

IBM introduces the **first magnetic hard-disk drive,** the RAMAC (random access method of accounting and control). Disks are two feet in diameter and 50 of them are used to store 5 megabytes of data.

1957 The Soviet Union launched the **first artificial satellite, Sputnik I,** on October 4. It broadcast a beeping sound at 20 MHz and 40 MHz. A much heavier Sputnik II, launched just a month later, placed some six tons in orbit; a payload of 508.5 kg, which included the dog Laika, and the spent upper stage which remained attached. The first U.S. satellite, Explorer I, launched on February 1, 1958, weighed just 4.7 kg.

A pressurized light-water reactor (LWR), built by Westinghouse Electric, but derived from Admiral Hyman Richover's submarine-propulsion reactor, went critical in Shippingport, PA. The same year, a small boiling-water LWR built by General Electric started operation in California. LWR types were to dominate commercial production of nuclear electricity in the United States, Europe, and Asia for the next

Samsung, EX1028, p. 33

generation. Thermocompression bonding and ball bonding was developed at Bell Labs starting in 1957.

1958 The **integrated circuit** was developed by Jack Kilby of Texas Instruments. He had conceived of creating components in silicon by diffusing it with impurities to make p-n junctions. On September 12, he built a complete oscillator on a chip. Soon after, Robert Noyce and Jean Hoerni of Fairchild Semiconductor developed the planar process that was to commercialize the IC. Jack Kilby received the Nobel Prize more than 40 years later in the year 2000.

Texas Instruments' "Employee of the Year" was Jack Kilby, who developed the integrated circuit within a year of joining the company.

Charles Townes of Columbia University teamed with Bell Labs scientist, Arthur Schawlow, to find ways to extend the frequency range of the maser, a device that amplified microwaves by employing the simulated emission of photons, so as to create a **laser.** The article on their work, **"Infrared and Optical Masers,"** described the conditions required to make masers operate in the infrared, optical, and ultraviolet regions. Earlier, in 1954, Townes had designed and built the first maser.

1959 Leo Esaki of IBM invented the **tunnel diode** for his work at Sony, a feat viewed as the most important discovery since the transistor. He was awarded the Nobel Prize in 1972 for the device, which caused a lot of heartburn for manufacturers trying to move into eagerly awaited production mode. When the processing kinks were finally worked out, the market had fizzled, and only a few tunnel diodes found their way into microwave systems.

Resistors based on carbon, nichrome and capacitors based on anodized aluminum, paper, mylar, mica, ceramic, tantalum were developed. Sprague Electric and Kemet were one of the first ones to have commercialized some of these technologies.

1960 The U.S. Navy demonstrated the feasibility of using satellites as navigational aids with Transit-1B, launched on April 13. (Transit-1A, launched on September 17, 1959, failed to reach orbit.) A Transit receiver on a ship used the measured Doppler shift of the satellite's radio signal, together with known characteristics of the satellite's orbit, to calculate the ship's position. **Navigational satellites** are today well known because of the widely used Global Positioning System.

Building upon the ideas of Townes and Schawlow, and his own work on the solid-state maser, Theodore H. Maiman at Hughes Research Laboratories demonstrated the **first laser** on May 16. Since then, there have appeared a wide variety of types of lasers and an even wider variety of applications, including communications, as in optical fibers, holography, measuring distances and speeds, surgery, micromachining, computer printing, and optical recording systems.

Sony introduced the **first** fully transistorized, **portable black and white TV** in Japan.

1962 In the bipolar transistor, the usual type of discrete device of the '50s and in early ICs, electron action takes place within the body of the semiconductor. With the metal-oxide silicon field-effect transistor **(MOSFET),** electron action occurs at the surface. Steven Holstein and Frederick Heiman of the RCA Electronic Research Laboratory demonstrated an MOS IC in 1962. With MOS ICs, it proved possible to

Samsung, EX1028, p. 34

put more components onto a piece of silicon, so that the number of elements roughly doubled every 18 months, now known as Moore's Law, after Gordon Moore of Intel who first noted it.

Telstar, the **first communications satellite** (unless one counts Echo, an aluminized balloon, launched in 1960, that reflected radio signals passively), was launched on July 10. Telstar was an AT&T project, with NASA reimbursed for launch costs and some tracking and telemetry. It demonstrated the feasibility of an active broadband repeater in earth orbit, permitting live TV exchange between Europe and North America.

The first commercial communications satellite was Intelsat I, also called Early Bird, launched on April 6, 1965. It carried one TV and 240 voice channels.

1963 IBM began the first multichip multilayer ceramic substrate technology, building upon the invention of "via" by Howard Stetson of 3M and multilayer greensheet formation at RCA in Somerville New Jersey to connect one layer of metal wiring to another layer and thick film capacitor technology. Drs. Bernie Schwartz, David Wilcox, Rao Tummala and their teams were later credited for the industry's first development of **multilayer multichip module (MCM).**

1964 The Japanese pioneered **high-speed electrified trains** with Shinkansen (bullet trains), which at first had a maximum speed at 210 km/h. Another major advance in high-speed trains did not occur until 1981, when the French National Railroads began running the Train a Grande Vitesse between Paris and Lyon. These electrified trains, each a permanent combination of passenger cars between two locomotives, attained 270 km/h and increased to 300 km/h at the end of the decade.

Robert Moog offered an **electronic music synthesizer.** Five years later, jazz pianist Paul Bley gave a live performance on one.

IBM announced the System 360, the **first compatible "family" of computers.** Customers could choose from five processors and 19 combinations of power, speed, and memory.

IBM introduced **flip chip technology to replace wirebonding.** Lew Miller patented the process in 1969 and Kyoto ceramics (now Kyocera) in Japan quickly followed.

1965 The PDP-8, a product that reshaped the computer industry, was introduced by Digital Equipment Corp. (DEC). As **the first computer to take advantage of ICs,** it was smaller than earlier machines (about the size of a refrigerator) and less expensive (only U.S. $18,000). DEC sold some 40,000 PDP-8s in the next decade—the machine virtually defining the product known as the **minicomputer.**

DEC, formed in 1957 by Kenneth Olsen and Harland Anderson, had announced its first computer, the PDP-1 (PDP standing for programmed data processor) in 1960. Olsen intended the PDP-1 and subsequent machines for science and engineering, rather than business and communities.

Gordon Moore noticed that chip capacities double each year, an observation that becomes known as **"Moore's Law."** This early pace has slowed somewhat over the past few years, so the law has been amended to show a doubling every 18 months. However, the pace of announcements from 1997 by Intel (2-bit flash memory) and IBM (copper circuitry) may indicate that the pace of development is now up to Moore's original observation.

Samsung, EX1028, p. 35

Bryant Rogers of Fairchild oversaw the invention of **Dual-In-Line package (DIP)** with 14 leads, which TI implemented in 1965.

1966 IBM's Robert H. Dennard invented the **dynamic memory cell** by using only one transistor per bit of information. These memory cells called **DRAM** for dynamic random access memory substantially increased computer memory density and were rapidly adopted throughout the industry.

Frances Hugle, a silicon Valley Engineer, was credited with the patented invention of **Tape Automated Bonding (TAB).** GE was the first company to use the technology commercially.

1968 Intel is founded. The company's original focus was memory chips, but more recently has been on microprocessors. The company then went on to become the world's largest semiconductor manufacturer.

DARPA wrote a plan to develop what was then called "ARRANT." The execution of this plan produced what is now known as the **Internet.**

1969 On July 20, attention everywhere was riveted on **Neil Armstrong and Buzz Aldrin** as they **set foot on the moon.** Electronic technologies contributed in countless ways: in the design and building of the Apollo spacecraft and the lunar excursion module—which lowered them to the moon's surface—in control systems, for communications, and for navigation. And it was, of course, electronic communications that provided real-time coverage of the event worldwide.

The Advanced Research Projects Agency of the U.S. Department of Defense established **Arpanet,** a data communications network that beginning in late September linked four academic institutions (the University of Utah, Stanford Research Institute, and the Universities of California at Los Angeles and at Santa Barbara). It relied on packet switching, breaking messages into smaller packets before transmission, routing them as individual units, and reassembling them at the receiver. Arpanet grew into **today's Internet and World Wide Web.**

Packed into a "tiny" 6 by 4.5-inch format, the world's **first handheld calculator** was introduced by Sharp Electronics Corporation. For just $495, one could add, subtract, multiply, and divide, all on battery power.

1970 Corning Glass Works demonstrated highly transparent fibers, and Bell Laboratories demonstrated semiconductor lasers that could operate at room temperature. These demonstrations helped establish the feasibility of **fiber-optic communications.**

The Hamilton Watch Company introduced **the first digital watch,** the Pulsar, with digital images on a Light Emitting Diode display.

1971 Marcian E. ("Ted") Hoff of Intel developed the **first microprocessor, a single IC that performed all the elementary functions of a computer.** Announced on November 15, the 4004 was a 4-bit processor (that is, it operated on words four bits long) that Intel designed for a Japanese company for use in desktop scientific calculators. Within several years, there were microprocessors with vastly greater capabilities, such as the 8-bit Intel 8080 and the 16-bit PACE microprocessor by National Semiconductor, both announced in 1974.

The 4004 microprocessor (the early term for microprocessor) packed 2,300 transistors and executed 60,000 operations in one second using 10-micron

Samsung, EX1028, p. 36

technology. It sold for $200, about the same price as today's microprocessors with a transistor count more than three orders of magnitude greater.

James Fergason was awarded a patent on a "liquid crystalline material." This was the first of 100 patents he filed which formed the basis of the **liquid crystal display** industry.

1972 The **CT (computerized tomography) scanner** for producing cross-sectional images of the body was announced at the British Institute of Radiology Congress. Designed mainly by Godfrey Hounsfield and manufactured by EMI, it had been first demonstrated the year before at Atkinson Morley's Hospital in London. The scanner proved adept at diagnosing tumors and other lesions. Within a few years, hundreds of CT machines were in use throughout the world.

The slide rule had a prized place in every engineer's pocket. William Oughtred invented it in 1621, and after 350 years, Hewlett-Packard replaced it with **the first 35-key handheld scientific calculator.** Market analysts concluded that the HP-35, which was priced at just under $400, would not sell; they were wrong. It became as popular with engineers as pocket protectors did, and Hewlett-Packard couldn't produce them fast enough. As Hewlett-Packard stated: "This amazing new device weighs only nine ounces and fits easily into a shirt pocket, with speed and accuracy that surpasses that of the slide rule."

This year also saw the introduction of a revolutionary method to increase yield and wiring density in printed wiring boards. Rather than print-and-etch processing, Killmorgen Corporation's Photocircuits Division routed 4-mil insulated wire along precise paths. The technique handles dense wires so efficiently that it is still used today in military applications long after conventional printed wiring processing techniques have improved.

1973 On May 22, Robert Metcalfe, a research engineer at Xerox Corp.'s Palo Alto (CA) Research Center, typed a memo that outlined the design of a **local-area network** he **called Ethernet.** Metcalfe left Xerox in 1979 and founded 3Com, which began marketing a PC-version of Ethernet in 1982. By the mid '90s, there were five million Ethernets with some 50 million networked computers.

Gary Boone of Texas Instruments was awarded the U.S. patent number 3,757,305 for the **single-chip microprocessor architecture.** Texas Instruments also introduced the first commercial 4-kbit DRAM with a single-transistor storage cell.

Also in this year, Phillips and Sony announced the standard for **laser-read disks, CD-ROMs** that stored 600 megabytes of data.

1974 The first microprocessor with memory and input/output on the same chip was announced by Texas Instruments. With the prices of microchips dropping, the home version of the video game Pong was introduced. Nolan Bushnell of Atari invented the original in 1972.

Zhores I. Alferov of Physico Technical Institute, Russia and Herbert Kroemer of University of California, Santa Barbara, invented and developed **fast opto- and microelectronic components,** based on layered semiconductor structures termed semiconductor heterostructures. Fast transistors built using heterostructure technology are used in radio link satellites and the base stations of today's mobile phones. They both received Nobel Prizes in the year 2000.

1975 Videotape recording was developed by a U.S. company for use by television broadcasters. But Japan's Sony and JVC took the lead in turning videotape

Samsung, EX1028, p. 37

recorders into a consumer product. Sony introduced its **Betamax videocassettes** and recorders in the early '70s. JVC launched its **VHS system** in 1975, and this format came to dominate the market.

One of the first "portables" was introduced by IBM, the **5100 laptop,** with a memory storage capacity of 48 Kbits. By IBM's own admission, the system was more "luggable" than portable, since it weighed 50 pounds. It sold for between $9000 and $20,000.

Rao R. Tummala was the first to propose what he then called "glass-ceramic-copper multilayer technology" with dielectric constant half of alumina and electrical conductivity three times better than molybdenum or tungsten previously used and thermal expansion exactly matching silicon. This technology that IBM commercialized for all its high performance systems since the 1990s is now called LTCC, **low-temperature cofired ceramic.**

1976 Together with the venture capitalist Mark Markulla (who provided the $92,000 operating capital), **Steve Jobs and Steve Wozniak incorporated Apple Computer on January 3** to make and sell the Apple II personal computer. The machine traced its roots to the altair 8800, announced in January 1975. This microprocessor-based computer, which used the Intel 8-bit 8080 chip, was meant to be assembled by the hobbyist. And though it had to be programmed through switches on the front panel and had only 256 bytes of memory, it attracted much attention.

Early in 1976, Jobs and Wozniak designed a much more capable computer based on a $25 processor: the 8 bit 6502 from MOS Technology. They sold about 200 of these first Apple Computers (assembled by the two in the garage of Jobs' parents), but were unable to interest manufacturers in licensing the product.

Wozniak then designed a much-improved Apple II, which came with the Basic language and a video monitor that could display text and graphics in color.

Radio Shack followed closely with its version, the TRS-80. Personal computers had been on the market since 1972, but primarily in kit form that required major assembly.

Panasonic and JVC introduced Betamax's competitor, the Video Home System, better known as VHS.

1977 Hewlett-Packard introduced the world's first **watch with a built-in calculator.** Engineers the world over quickly realized the importance of user-friendliness when they misplaced the stylus required to punch the tiny keys. The initial, and final, production was sold to HP employees at a deep discount.

1978 The first commercial product using a **speech-synthesis chip** was introduced by Texas Instruments. The Speak & Spell educational toy was a marketing phenomenon.

1979 Sony introduced the **Walkman personal stereo,** changing the way many listen to music.

Alan Heeger of University of California, Santa Barbara, Alan Macdiarmid of University of Pennsylvania, Philadelphia and Hideki Shirakawa of University of Tsukuba, Japan won Nobel Prizes for their discovery and development of **electrically-conducting polymers.**

1980 Professor Lawrence Nagel and his colleagues at University of California, Berkeley have been credited for their contribution to the well known **SPICE modeling and simulation** used in all modern microelectronic designs.

Samsung, EX1028, p. 38

1981 IBM introduced a personal computer on August 12. Uncharacteristically, the company relied heavily on outside contractors for the microprocessor and most other parts, as well as the operating system. The IBM Personal Computer used the 8-bit Intel 8088 processor, came with 16 KB of Memory and a floppy-disk drive (the disk able to hold 160 KB), and sold for $2495. (The PC-DOS operating system, from a small company called Microsoft, cost $40 more.) Sales surpassed expectations, and other manufacturers (including Compaq, Tandy, Commodore, and Zenith) moved quickly to produce IBM-compatible computers. Software makers produced thousands of programs for these machines, and almost all computer manufacturers adopted the IBM standard; Apple being the only important exception.

A caricature of Charlie Chaplin introduced IBM's first PC, marking the company's foray into consumer electronics. Intel's microprocessor was inside from the start. In fact, most of the components were sourced outside of IBM, including the Disk Operating System purchased from a 32-person company called Microsoft. The PC had 16 KB of memory and one or two floppy drives. The basic price started at $1,565 and 136,000 units were sold in the first 18 months.

The year 1981 was officially recognized as the introductory year of **surface-mount packaging technology.** In reality, 50 mil pitch flat packs had been used in the Apollo missions. Through-hole DIPs were developed, because the flat packs were so difficult to assemble. Also this year, Sony introduced the **first 3.5-inch floppy disk drive,** which became the standard for data storage on personal computers.

1982 Sony unveiled the **Watchman personal TV** and the world's **first CD player.**

Time Magazine named the computer as the man of the year. Never before had an inanimate object been chosen for this honor.

1983 Philips and Sony began selling compact disk players after the rival companies combined forces to develop the novel recording medium. The CD is digital, with each second of sound initially encoded as 705 600 bits 16 bits—for each of 44 100 samples—and finally encoded with about three times as many bits to allow for error correction and other requirements. A CD player performs a sequence of signal-processing tasks to produce sound much higher in quality than earlier recording technologies: greater bandwidth, flatter frequency-response, greater dynamic range, and better signal-to-noise ratio. The CD quickly replaced the phonograph record and contributed to the development of other media, such as the CD-ROM, by Philips and Sony in 1984 as a computer memory, and the DVD for video and audio, programs, in the late '90s.

Radio Shack introduced the TRS-80 Model 100, the first laptop with an operating system in ROM and 29 K bytes of memory.

1984 AT&T and the Bell System as a regulated monopoly providing telephone service in the United States came to an end on January 1. A consent decree between AT&T and the U.S. Justice Department inaugurated full competition in long-distance telephony and, after the Telecommunications Act of 1996, local service as well.

1985 Toshiba launched its T1100, the first laptop computer generally accepted by the market. The system was powered by an 8088 processor and was based on a 3.5 floppy drive. It weighed 2.9 kg.

Samsung, EX1028, p. 39

1986 Yukata Tsukada, an engineer at IBM Japan, proposed what he called Surface Laminar Circuitry (SLC), now generally known as build-up technology. This is the first thin film technology, similar to MCM, applied to printed wiring boards. This build-up technology is a $2 billion industry and is at the heart of most microprocessor packaging.

1987 The second generation of personal computers from IBM became available. 100 patents protected the PS2 design.

Yukata Tsukada, now an IBM Fellow, was the first to recognize the effect of underfill encapsulant between IC and organic package in improving reliability of IC connection that is attached to printed wiring board. Hitachi was the first to use **underfill** with ceramic **packages.**

1988 Sony created the world's first 5.25 inch magneto optical disk.

1989 The number of universities offering electronic packaging courses in the U.S. were four. Dr. Rao R. Tummala's *Microelectronics Packaging Handbook* was credited as the first step leading to the 40+ universities offering packaging courses today.

1990 The first version of the World Wide Web software, created by Tim Berners-Lee and others, began operating within GERN (the Geneva-based European Organization for Nuclear Research) in December. The Internet took a second step toward near-universal use with the invention of **Mosaic, the first Web browser, by a team at the University of Illinois, Urbana-Champaign.**

The **Hubble Space Telescope was launched** and, despite a hugely embarrassing error in fabricating the mirror that had to be corrected, soon produced images of the cosmos of unprecedented clarity and depth. NASA's Compton Gamma Ray Observatory launched in 1991, and the Chandra X-ray Observatory joined the Hubble in space in 1999. Together, the three instruments proved a horn of plenty, making the last decade of the 20th century a golden age for astronomers, astrophysicists, and cosmologists.

1992 For 20 years, Gil Hyatt fought with the U.S. Patent Office for recognition of his "single-chip integrated-circuit computer architecture." In 1992, six Japanese electronics manufacturers agreed to pay him licensing fees anyway.

1993 Intel launched the Pentium processor in 60 or 66-MHz speeds. Its 3.1.M transistors were built with 0.8-micron technology and packaged in a 273 I/O pin grid array.

1994 The **Moving Picture Experts Group (MPEG)** completed the main elements of the MPEG-2 standard for **video compression and decompression,** in effect establishing the core of a world standard for digital television, both standard and high definition.

The Department of Defense announced a plan to pump $600M into the U.S. flat-panel display effort in an attempt to catch up to, and hopefully surpass, Japanese technology.

1995 The **gigachip era began.** At an IEEE circuits conference, Hitachi and NEC announced the first 1.Gb dynamic RAMs. The first giga-instruction-per-second microprocessor was announced by DEC.

Samsung, EX1028, p. 40

The **DVD** (also referred to as the **digital video disk** and the digital versatile disk) was introduced, promising to consign videotape to the dustbin of history, much as audio disks had wiped out long-playing records a decade earlier. The DVD relied on infrared lasers and strong focusing to achieve a capacity of 4.6 GB.

Intel's Pentium chips moved to a 0.35-micron process and clock speeds increased to 166 MHz. Taiwanese companies made plans for 12 inch wafers and deep submicron processes.

Hand-held computing hit the market with the introduction of the **Pilot 1000** by Palm Computing subsequently bought by 3Com Corporation.

1996 A **250-KW fuel cell** stack, manufactured by Energy Research and the first to be connected to the electric power grid, began operation in Santa Clara, California.

1997 On May 11, IBM's **Big Blue computer system defeated world chess champion Garry Kasparov** in a tournament. One possible element in its victory: computers do not tire under pressure. In Garry Kasparov vs. Deep Blue, IBM designers did their best to explain that the chess match between man and machine was just a demonstration of the technology's capability but the red-faced Kasparov could not be consoled after his loss.

Moore's Law vs. The Law of Physics: at an Intel developer forum, Gordon Moore stated that the industry's ability to shrink a microprocessor through improved manufacturing processes was going to start butting against the finite size of atomic particles. One study estimated that the brick wall would be hit in 2017.

Researchers at Georgia Tech, under the leadership of Prof. Rao Tummala and with NSF funding, proposed a way of dealing with Moore's Law by a new concept they called **System-On-Package** or **SOP.** With this concept, the complete integration onto one chip known as SOC is no longer necessary and yet a single component would provide digital analog, optical and MEMS functions.

IBM introduced a **matchbook-size hard drive** that measured just over an inch square (to Compact Flash Type II standards) and packed 340 megabytes of data. Inside are four chips that must be flip chip attached to fit in the 0.5-mm height. Critics claimed that it was just a reincarnation of Hewlett-Packard's Kittyhawk drive, the 1.3 inches introduced in 1992 that never took off.

Copper, one of the best conductors of electricity on earth, moved into chips with IBM's announcement of a process to fabricate copper interconnects to replace aluminum.

1998 IBM produced the **first mainframe server to exceed 1000 million instructions per second.** And IBM and Motorola independently introduced techniques permitting aluminum to be replaced by copper wiring in ICs, which has less resistance.

In September, Microsoft became the world's most valuable company, surpassing General Electric. The following year, Cisco Systems, the supplier of Internet backbone equipment, surpassed Microsoft. The longest record in terms of number of quarters as the most valuable company still belonged to IBM as of 2000.

Sony announced its next-generation 3.5-inch floppy disk, the HiFD that stores 200 Mbytes of data. The company is now in competition with many others trying to replace Sony's original floppy format.

1999 Construction began on the **International Space Station,** intended as the first permanently manned space facility.

Samsung, EX1028, p. 41

Developments are progressing on **quantum computing,** a method based on quantum bit or qbit, to process information much faster than we realize today. NEC has developed a **super-conductive circuit** fabricated on a silicon substrate, a solid state device that could function as a quantum bit. Bell Labs and Michigan State University are pursuing an alternative approach that uses electrons floating on liquid-helium in a vacuum.

Other leading-edge research involves using proteins and polymers as the basis of computer processing. The materials act as bi-stable devices since their configuration change is a result of outside stimulation. Studies are underway to determine how to apply their molecular characteristics to form massive computing power and data storage. It is likely that one or both of these approaches will form the basis of the nano-computers of the future and will play a key role in the electronics industry well into the 21st century.

2000 IBM announced the **fastest super computer** that is so powerful that weather prediction can now be extended to 7 days in advance. This speed of **3 billion operations per second** is attributed to materials science and system design.

1.9 HOMEWORK PROBLEMS

1. What is a microsystem? Compare and contrast with microelectronics.
2. Why integrate microsystem technologies into single products?
3. What is the role of packaging in microsystems?
4. What is the fundamental building block of IC? What is the fundamental building block of MEMS, Optoelectronics and RF?
5. Why is packaging important?
6. What do you do in systems packaging if you are a physicist, metallurgist, ceramist, industrial engineer, computer engineer, chemical engineer, electrical engineer or mechanical engineer?

1.10 SUGGESTED READING

1. W. Brown, *Advanced Packaging*. IEEE Press, 1999.
2. L. W. Couch. *Digital and Analog Communication Systems*. Fourth Edition. New York: Macmillan, 1993.
3. D. Doane and P. Franzen, eds. *Multichip Module Technologies and Alternatives*. New York: Van Nostrand Reinhold, 1993.
4. K. Feher. *Wireless Digital Communications*. Englewood Cliffs: Prentice Hall, 1995.
5. T. L. Floyd, *Electronics Devices*. Prentice Hall, 1998.
6. S. Harsany. *Principles of Microwave Technology*. New Jersey: Prentice Hall, 1997.
7. J. Hecht, *Understanding Fiber Optics*. Prentice Hall, 1999.
8. W. Johnson, et al., eds. *Multichip Modules*. New Jersey: IEEE Press, 1990.
9. H. L Kraus, C. W. Bostonian, and F. H. Raab. *Solid State Radio Engineering*. New York: John Wiley & Sons, 1980.
10. R. Lasky, et al., eds. *Principles of Electronic Packaging*. New York: McGraw-Hill, 1989.
11. J. H. Lau. *Chip Scale Package Design, Materials, Processes, Reliability, and Applications*. New York: McGraw-Hill, 1999.
12. J. H. Lau. *Flip Chip Technologies*. New York: McGraw-Hill, 1996, 565.
13. J. Lau. *Handbook of Fine Pitch Surface Mount Technology*. New York: Van Nostrand Reinhold, 1993.

Samsung, EX1028, p. 42

14. J. H. Lau. *Low Cost Flip Chip Technologies: for DCA, WLCSP, and PBGA Assemblies.* New York: McGraw-Hill, 2000.
15. J. D. Lenk. *Optimizing Wireless/RF Circuits.* New York: McGraw-Hill, 1999.
16. M. J. Madou. *Fundamentals of Microfabrication.* Boca Raton: CRC Press, 1997, 589.
17. D. Messner, et al., eds. *Thin Film Multichip Modules.* New Jersey: IEEE Press, 1993.
18. E. Peeters. "Challenges in Commercializing MEMS." *IEEE Computer Science Engineering* Jan–Mar 1997.
19. K. Peterson. "Silicon as a Mechanical Material." *Proceedings, IEEE.* Vol. 70, No. 5, 1982, 420–457.
20. B. Razavi, *RF Microelectronics.* Prentice Hall, 1998.
21. D. E. Ricken, and W. Gessner. *Advanced Microsystems for Automotive Applications '98.* New York: Springer, 1998.
22. B. F. Romanowicz. *Methodology for the Modeling and Simulation of Microsystems.* New York: Kluwer Academic Publishers, 1998.
23. D. Rutledge, *The Electronics of Radio.* Cambridge University Press, 1999.
24. R. Schneiderman. "GaAs Continues to Gain in Wireless Applications." *Wireless Systems Design.* March 1997, 14–16.
25. B. Sheu. *Microsystems Technology for Multimedia Applications: An Introduction.* New York Institute of Electrical and Electronics Engineers, 1995. 713.
26. B. Sklar. *Digital Communications.* Englewood Cliffs: Prentice Hall, 1998.
27. R. Steele, ed. *Mobile Radio Communications.* Piscataway: IEEE Press, 1992.
28. F. E. Terman. *Radio Engineers Handbook.* New York: McGraw-Hill, 1943.
29. R. Tummala, E. Rymasweski, and A. Klopfenstein, eds. *Microelectronics Packaging Handbook.* Second Edition. New York: Van Nostrand Reinhold, 1989.
30. J. Vardaman, ed. *Surface Mount Technology—Recent Japanese Developments.* New Jersey: IEEE Press, 1993.
31. C. J. Weisman. *Essential Guide to RF and Wireless.* New Jersey: Prentice Hall, 2000.
32. W. Wolf, *Modern VLSI Design.* Prentice Hall, 1998.
33. Lau, Wong, et al. *Electronic Packaging: Design, Materials, Process, and Reliability.* New York: McGraw-Hill, 1998.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contribution of Dr. Don Estreich of Agilent Technologies for critically reviewing this chapter.

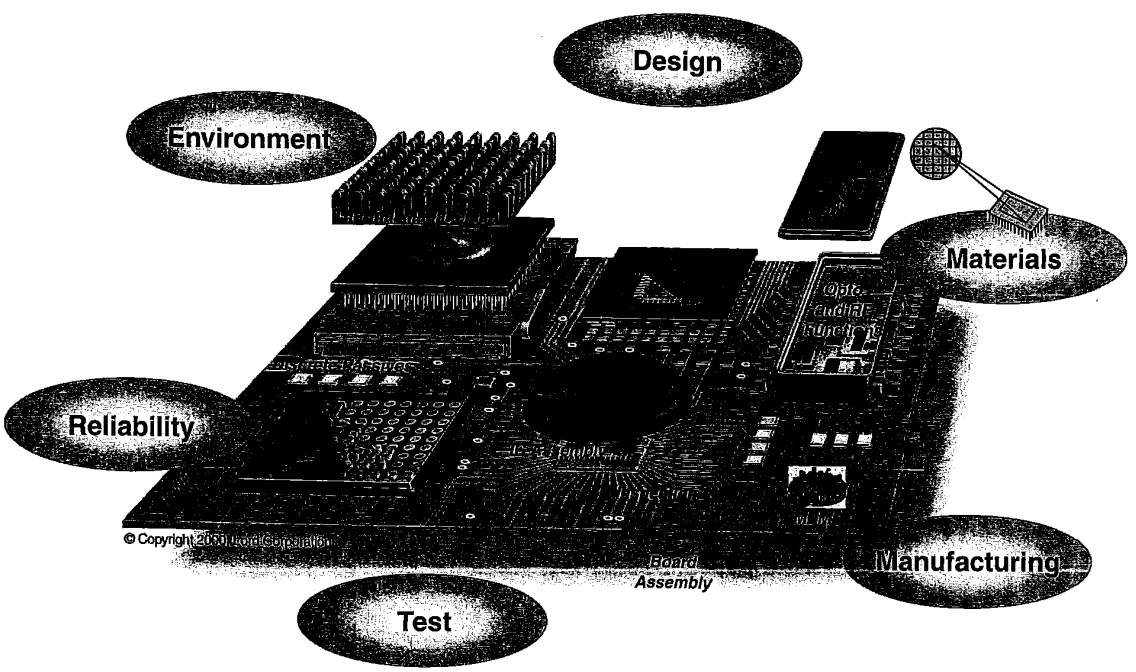
Samsung, EX1028, p. 43

CHAPTER 2

THE ROLE OF PACKAGING IN MICROELECTRONICS

Prof. James E. Morris
State University of New York at Binghamton

Prof. Rao R. Tummala
Georgia Institute of Technology



2.1 What Is Microelectronics?
2.2 Characteristics of Semiconductors
2.3 Microelectronic Devices
2.4 Integrated Circuits
2.5 IC Packaging
2.6 Semiconductor Roadmap
2.7 IC Packaging Challenges
2.8 Summary and Future Trends
2.9 Homework Problems
2.10 Suggested Reading

CHAPTER OBJECTIVES

- Define microelectronics
- Describe characteristics of semiconductors
- Define and describe integrated circuits and their evolution
- Describe IC processes
- Introduce the System-on-Chip (SOC) concept
- Describe the role of IC packaging
- Translate the semiconductor roadmap into IC packaging challenges

CHAPTER INTRODUCTION

Integrated circuits (ICs) based on microelectronic devices form the basis of all modern electronic products. Since the invention of the transistor in 1947, electronic products began shifting from vacuum tubes to transistors in the 1950s, and to integrated circuits in the 1960s. These pioneering technology developments formed the basis of large scale integrated (LSI) devices in the 1970s and very large scale integrated devices (VLSI) in the 1980s, paving the way for gigascale integration and the System-on-Chip (SOC) in the 21st century. Every IC has to be packaged before it can be used. Packaging starts where the IC stops. The typical parameters important for IC packaging include I/Os, power and size of the chip. This chapter sets the stage for packaging ICs, eventually leading to all microsystems.

Samsung, EX1028, p. 45

2.1 WHAT IS MICROELECTRONICS?

The invention of the transistor in 1947 revolutionized the semiconductor industry as it replaced the usage of vacuum tubes in electronic products and enabled shrinkage in the size of the products considerably. In contrast to the large dimensions of vacuum tubes, the transistor's size was measured in micrometers (μm) and it became the dominating device in microelectronics. To achieve high functionality and high-performance products, *integrated circuit* (IC) technology was developed in the early 1960s to integrate hundreds of transistors on a single semiconductor chip. Integrated circuits (ICs), based on microelectronic devices, form the basis of all modern electronic products. Continuous advances in reducing the size of the transistors allowed the progressive integration of tens, hundreds, then thousands of transistors on a single IC in technologies called *small, medium, and large scale integration* (SSI, MSI, and LSI). This led to the integration of up to a million transistors on an IC, called *very large* or *ultra scale integration* (VLSI or ULSI). Driven by the functional and performance requirements of modern and future electronics, the technology drivers in the semiconductor industry envision the era of *gigascale integration* (GSI) and terascale integration in the immediate and far futures, where many billions and trillions of transistors may be integrated on a single semiconductor chip. Table 2.1 indicates approximate ranges of transistors, gates, *input/output* (I/O) connections, and the system-level functions that can be achieved. The idea of *system-on-chip* (SOC) calls for an IC where not only transistors, but other crucial components of an electronic system, are integrated on a chip. For example, in a modern cellular phone, many discrete components such as capacitors, resistors, power transistors, semiconductor chips, etc., are assembled on a *printed wiring board* (PWB) which provides electrical interconnections among these components. The SOC concept is expected to integrate all of these components on a single semiconductor chip and facilitate making electronic products very small and highly functional. This evolution of microelectronics technology from a single transistor to many billions of transistors on a chip is summarized in Figure 2.1.

TABLE 2.1 Advances in IC integration.

Integration Level	Number of Transistors	Equivalent Gates	Typical Functions of Systems	Typical Number I/Os
SSI	1–40	1–10	Single Circuit Function (e.g., Transistors)	14
MSI	40–400	10–100	Functional Network	24
LSI	400–4,500	100–1,000	Hand Calculator or Digital Watch	48
VLSI	4,500–300,000	1,000–80,000	Microprocessor	64–300
ULSI	Over 300,000	Over 80,000	Small Computer on a chip	300 +
GSI	1 Billion	Over 100 Million	Supercomputer	10,000 +

Samsung, EX1028, p. 46

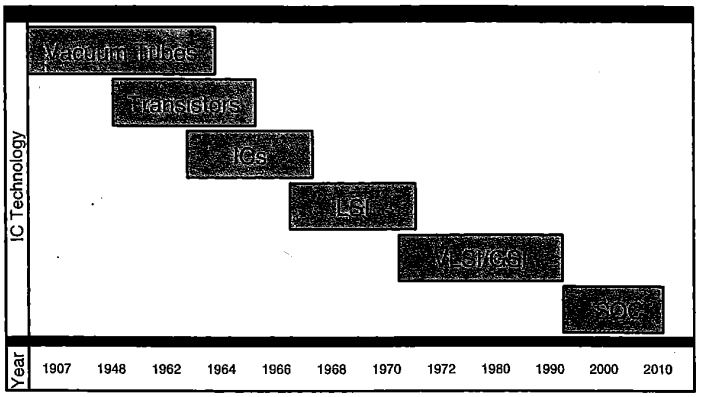


FIGURE 2.1 Evolution of microelectronics.

2.1.1 What Is Microelectronics Packaging?

The functions of an electronic package are to protect, power, and cool the microelectronic chips or components and provide electrical and mechanical connection between the microelectronic part and the outside world. Whether a single transistor or a GSI chip, they have to be packaged. As such, the electronic package is an integral part of the microelectronic system. The challenge for the package is to provide all crucial functions required by the microelectronic part without limiting the performance of the part. To meet this challenge, the package technology has also evolved from a simple metal can to very complex multilayer ceramic and organic structures that are described in detail throughout this book. But as the semiconductor technology progresses towards higher levels of integration, high performance and increasing functionality, the design and fabrication of the package that will meet the requirements of modern and future microelectronic systems becomes increasingly complex and challenging. To establish a better understanding of the relationship between microelectronics and packaging, IC packaging technologies are succinctly described in the following sections. The projections of the International Technology Roadmap for Semiconductors (ITRS) are also discussed and their impact on the packaging technology is described.

2.2 CHARACTERISTICS OF SEMICONDUCTORS

All materials are classified into three categories: insulator, semiconductor and conductor. Insulators have a maximum of eight electrons in their valence band. A tremendous amount of energy is required to remove them from this orbit. At the other extreme, the valence electrons of an ideal conductor require virtually no energy to be removed. Most good conductors have one or two valence electrons. Semiconductors have four valence electrons, which require a moderate amount of external energy in order to be removed.

2.2.1 Energy Bands

The valence shell of an atom represents a band of energy levels and the valence electrons are confined to that band. When an electron acquires enough additional energy from an

Samsung, EX1028, p. 47

external source, it can leave the valence shell and become a free electron and exist what is known as the conduction band. The difference in energy between the valen band and the conduction band is called an energy gap or bandgap. This is the amou of energy that a valence electron must have in order to jump from the valence band the conduction band. Once in the conduction band, the electron is free to move throug out the material and is not tied to any given atom. Figure 2.2 shows energy diagran for bandgaps of insulators, semiconductors, and conductors. Notice in section (a) th insulators have a very wide energy gap. Valence electrons do not jump into the conducti band except under breakdown conditions where extremely high voltages are appli across the material. In section (b), semiconductors have a much narrower energy ga This gap permits some valence electrons to jump into the conduction band and becon free electrons. By contrast, in section (c), the energy bands in conductors overlap. In conductive material, there is always a large number of free electrons.

2.2.2 Conductors

A conductor is a material that easily conducts electrical current. The best conductors a single-element materials, such as copper, silver, gold, and aluminum, which are chara terized by atoms with only one valence electron very loosely bound to the atom. The loosely bound valence electrons can easily break away from their atoms and become fr electrons. Therefore, a conductive material has many free electrons that, when movir in the same direction, make up the current.

2.2.3 Insulators

An insulator is a material that does not conduct electrical current under normal condition Most good insulators are compounds rather than single-element materials. Valence ele

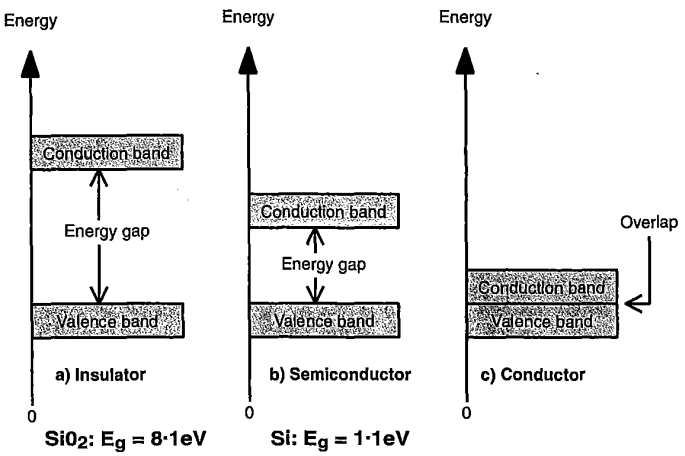


FIGURE 2.2 Bandgaps of conductor, semiconductor, and insulator materials.

trons are tightly bound to the atoms; therefore, there are very few free electrons in an insulator.

2.2.4 Semiconductors

A semiconductor material falls between conductors and insulators in its ability to conduct electrical current. A semiconductor in its pure (intrinsic) state is neither a good conductor nor a good insulator. The most common single-element semiconductors are silicon and germanium. Compound semiconductors, such as gallium arsenide, are also commonly used. The single-element semiconductors are characterized by atoms with four valence electrons.

Current flow in semiconductor devices can be controlled by an external force—mainly an electric potential. When appropriately designed and engineered, such a device can function like a switch. If a sufficient amount of current is flowing in the device, then the switch is said to be "closed," otherwise it is "open." These states of the switch can then be interpreted as binary "0" or binary "1," which form the basis of modern computing. As a result, these semiconductor devices are major building blocks of microelectronics.

2.2.5 Types of Semiconductors

The semiconductor materials mostly used in microelectronics can be classified into three categories—intrinsic, extrinsic, and compound—as shown in Figure 2.3.

Intrinsic

Intrinsic semiconductors are perfect semiconductor crystals that have no impurities or lattice defects. At low temperatures, there are no free electrons in the intrinsic semiconductor and the semiconductor behaves like an insulator. When the semiconductor is heated, some of its valence electrons acquire enough energy to break away from their parent atom and are excited to the conduction band, thereby increasing the conductivity

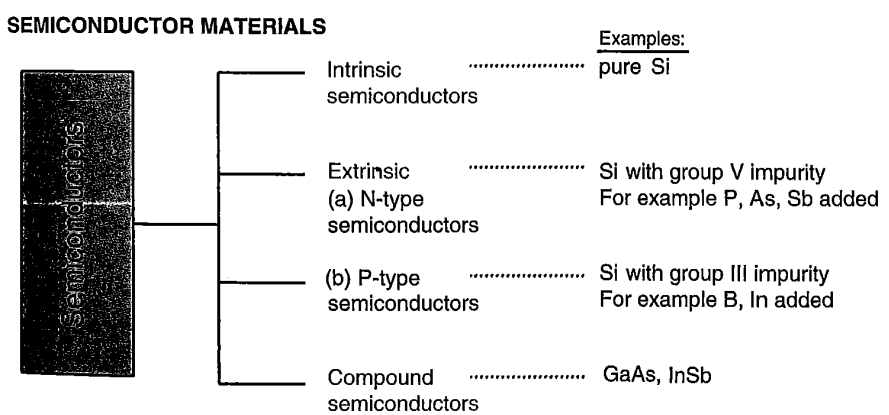


FIGURE 2.3 Three categories of semiconductor materials.

Samsung, EX1028, p. 49

of the material. In an intrinsic material, the number of electrons in the conduction band equals the number of holes in the valence band.

Extrinsic

The most common semiconductor materials, such as silicon (Si) and germanium (Ge), have four valence electrons in the outer shell. These elements are found in Group IV of the periodic table. When an element having five valence electrons (from Group V of the periodic table), such as arsenic (As) or phosphorus (P), is added to a Group IV element, such as Si, there are more valence electrons available than are required to form a covalent bond with silicon. One of the five electrons is, therefore, free to move and becomes a free charge carrier. This method of creating free carriers in intrinsic material by adding different elements or introducing impurities into a semiconductor crystal is called doping. Once a semiconductor is doped, it is called an extrinsic semiconductor. If the extrinsic semiconductor has an extra electron, such as silicon doped with As or P, it is called N-type material. The electrical conductivity of silicon is raised from 4×10^{-4} to 2.24×10^{5} $(\text{ohm-m})^{-1}$ by 2×10^{-4} atomic % of phosphorus.

In contrast to an N-type semiconductor, when a Group III element, like boron (B) or indium (In) is added to the silicon (Group IV), the impurity atom is one electron short of the four valence electrons required to form a covalent bond with silicon. This deficiency of an electron creates an extra hole in the material. In this case, the hole is a free carrier and the semiconductor is said to be a P-type material. Extrinsic semiconductors may contain doping concentrations as high as 5×10^{18} atoms/cm^3 and significantly improve the conduction properties of intrinsic semiconductors.

2.2.6 Why Silicon?

Silicon is the most widely used semiconductor material in microelectronics because of its material properties, and also because it can easily be extracted from naturally abundant silica (SiO_2). The Si atom consists of a nucleus with 14 electrons revolving around it in a system, which may be compared to the sun and its planets moving in their orbits. Electrons have negative charge and the Si nucleus has a positive charge, which offsets the charge on these 14 electrons, so the Si atom is electrically neutral. The electrons move in orbits around the nucleus. An atom which has eight electrons in its outermost orbit, or the outermost shell, is relatively stable, which it acquires by sharing one from each of four adjacent atoms in covalent bonds. Silicon has four electrons in its outermost shell, and it requires four additional electrons to become stable. Silicon can be easily doped to increase its electrical conductivity to create N-type or P-type material, which is then used to make various microelectronic devices.

The atomic structures of silicon and germanium are shown in Figure 2.4. Silicon is the most widely used material in diodes, transistors, integrated circuits, and other semiconductor devices. Both Si and Ge have the characteristic four valence electrons. The valence electrons in Ge are in the fourth shell while those in silicon are in the third shell, closer to the nucleus. This means that the Ge valence electrons are at higher energy levels than those in Si and, therefore, require a smaller amount of additional energy to escape

Samsung, EX1028, p. 50

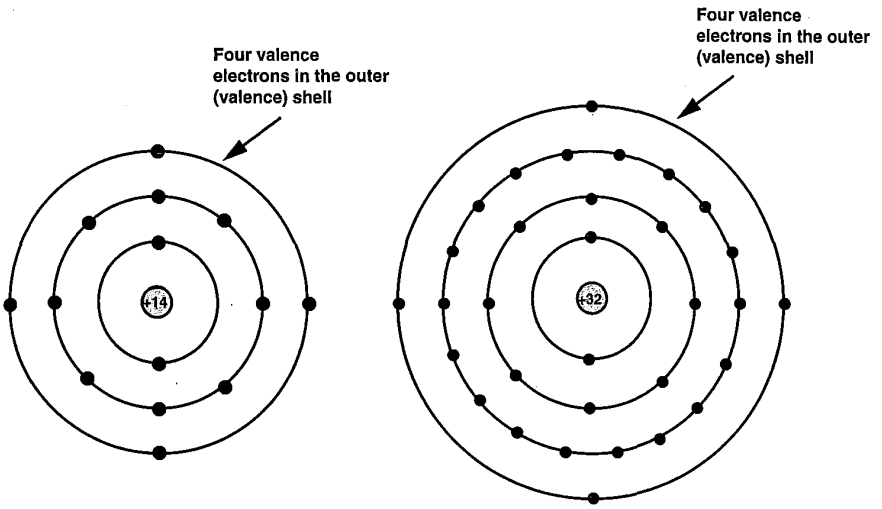


FIGURE 2.4 Atomic structure of (a) silicon and (b) germanium.

from the atom. This property makes Ge conductive at high temperatures, and this is a basic reason why Si is the most widely used semiconductor material.

2.2.7 Compound Semiconductors

Two or more elements can be combined on a single crystalline semiconductor material. These are called compound semiconductors. The compound semiconductors can be a mixture of two, three, or four elements and are called binary, ternary, or quaternary compound semiconductors, respectively. Some examples of compound semiconductors used in microelectronics are: binary (InSb, GaAs), ternary (GaAsP, AlGaAs), and quaternary (InGaAsP). Compound semiconductors are widely used in high speed and optoelectronics applications. The speeds of electrons in various semiconductor materials are shown in Figure 2.5 where it is observed that the speed in some binary compounds such as SiGe and GaAs can be much higher than in intrinsic or extrinsic semiconductor materials such as Ge or Si. However, the cost involved in manufacturing the compound semiconductors is much higher than for silicon. As such, the efforts of the semiconductor industry have been primarily focused on developing the silicon technology since the invention of the transistor. The GaAs technology has superior material properties that lead to higher performance devices than silicon, but the cost of manufacturing has deprived it of the astounding infrastructure and resources spent in developing the silicon technology. Silicon has traditionally led the way in defining the roadmap for the semiconductor industry and GaAs and other compound semiconductor technologies have always lagged behind. As the modern generation of electronic products rapidly moves into

Samsung, EX1028, p. 51

(a)

Semiconductor Material	Electron Mobility (cm²/V-s)	Hole Mobility (cm²/V-s)
silicon (Si)	1500	600
germanium (Ge)	3900	1900
gallium arsenide (GaAs)	8500	400
gallium antimonide (GaSb)	4000	1400
indium phosphide (InP)	4600	150
indium arsenide (InAs)	33,00	460
indium antimonide (InSb)	78,000	750

(b)

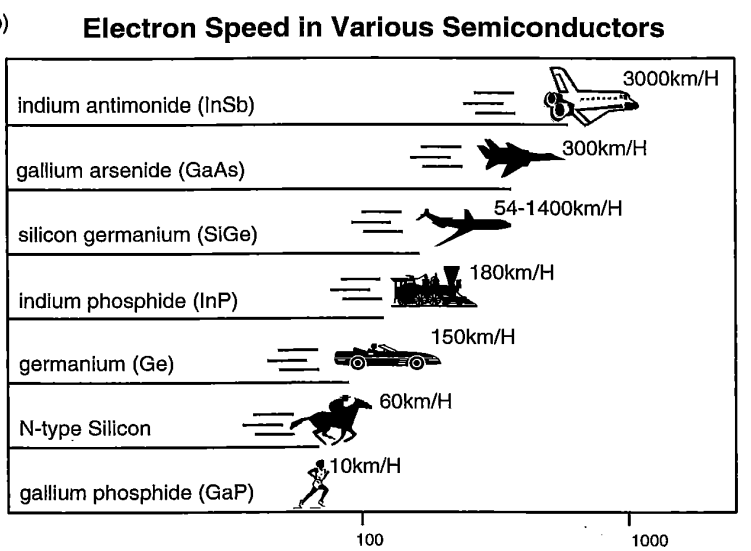


FIGURE 2.5 (a) Mobility and (b) electron speed of semiconductor materials.

information technology and the wireless world, the compound semiconductor technologies, such as SiGe and GaAs, may play an important role in microelectronics.

2.2.8 Silicon-Germanium

SiGe devices are emerging as an important technology in the broadband communications arena where high-frequency operation is required. The SiGe operating frequencies, in terms of process and transmission, can be significantly higher compared to silicon based devices. These high-frequency devices, which operate at lower power than silicon, are poised for great advances in wireless products for a number of reasons. They will help

Samsung, EX1028, p. 52

extend the battery life of wireless phones and will play a significant role in the development of other entertainment and information technologies like digital set-top boxes, Direct Broadcast Satellite (DBS), automobile radar systems, and *personal digital assistants* (PDA). SiGe promises to make these popular communications technologies less expensive, smaller and more durable.

SiGe devices are based on bandgap engineering, a way of dealing with the energy levels of silicon so as to accelerate the flow of electrons through the transistors. Products using SiGe technology are capable of dramatically greater functions using fewer chips. In the future, SiGe technology may be the force behind the creation of powerful new microcommunication devices like single-chip watch-sized phones and products that combine the capabilities of cellular phones, global positioning systems and internet access, all in one package. These multifunction, low-cost, mobile client devices capable of communicating over voice and data networks are considered to be key elements in the future of computing and communications.

2.2.9 Gallium Arsenide

At present, GaAs technology is primarily used in applications where emission or absorption of light is required. Laser sources emitting light over a wide spectrum of visible wavelengths, photodetectors, and diodes define a few applications where GaAs is used. At present, optical interconnects are being investigated as a technology that may replace electrical interconnects on semiconductor chips. Interconnects are wires that are used to communicate between devices on a chip. Electrical wires have inherent speed limitations where optical wires can operate at the speed of light. A crucial component for such an optical interconnect scheme is to be able to transmit and receive optical signals at high efficiencies. GaAs based technologies are mostly used as transmitters and detectors.

The microelectronics products, therefore, are comprised of many different types of semiconductor materials that are used to build application specific devices and circuits. Each technology has its own set of performance and functional requirements to contribute to the final product, and each component therefore requires its own packaging solution.

2.3 MICROELECTRONIC DEVICES

Semiconductor materials can be engineered and designed to build devices whose characteristics under controlled conditions can be used to construct microelectronic products. One of the basic devices is called a PN junction diode.

2.3.1 PN Junction Diode

When P-type and N-type extrinsic semiconductor materials are formed together, they form a PN junction. The current flow in the device is controlled by the applied potential across the junction which is shown in Figure 2.6. The PN junction carries large current when positive voltage is applied to the P material and negative voltage is applied to the N material. This is referred to as forward biasing the junction. When the applied potential is 0 volts, no current flows in the device. As the potential increases to a certain threshold value, the current increases by a minimum amount. Once the potential passes the threshold point, a small change in the applied voltage results in large current flow through the

Samsung, EX1028, p. 53

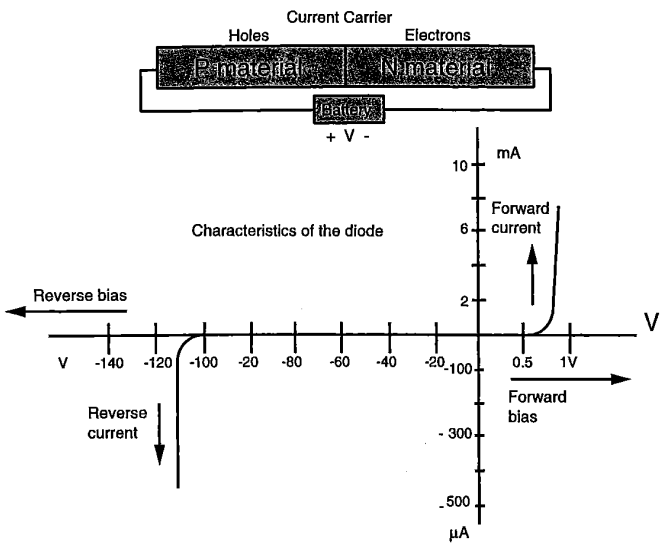


FIGURE 2.6 PN junction diode characteristics.

device. When the junction is reverse biased, positive voltage is applied to the N material and negative voltage applied to the P material, the device blocks the current flow. Thus, the PN junction has rectifying properties. Such a rectifier is called a junction diode. If the potential across the reverse biased junction is increased beyond a certain value, approximately −110 volts in the example of Figure 2.6, the device enters into a breakdown region.

2.3.2 Bipolar Junction Transistors (BJTs)

When N-type and P-type materials are placed as shown in Figures 2.7a and 2.7b, they form a *bipolar junction transistor* (BJT). Since both the electrons and the holes are involved in charge transfer, it is called a bipolar device. The middle material in the BJT structure is called the base region, and materials on either side of the base are called the collector and the emitter since their function is to emit and collect charge carriers.

There are two types of bipolar transistors: PNP and NPN. In the PNP transistor, holes from the emitter are accelerated into the collector if the base region is very thin. In the NPN transistor, it is the electrons from the emitter that cross the base to reach the collector. If the base-emitter junction is forward biased and a base current flows, a large current, tens to hundreds times the base current, flows between the collector and emitter. In other words, a small base current controls a large collector current and the transistor acts as a "current amplifier." The ratio of the collector current to the base current is called the current gain of the transistor, or the current amplifier in this case.

The device characteristics of a BJT are shown in Figure 2.8, where the *collector current* (I_C) is controlled by the *base current* (I_B). There are three regions of operations: cut-off, linear, and saturation. The linear region is used for amplifier applications, and

Samsung, EX1028, p. 54

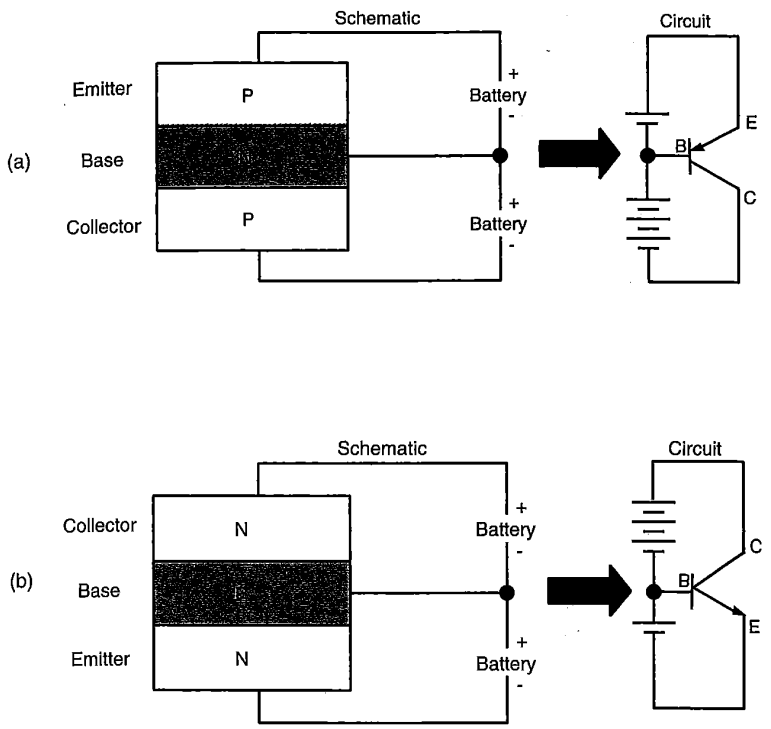


FIGURE 2.7 Bipolar transistors (a) PNP and (b) NPN.

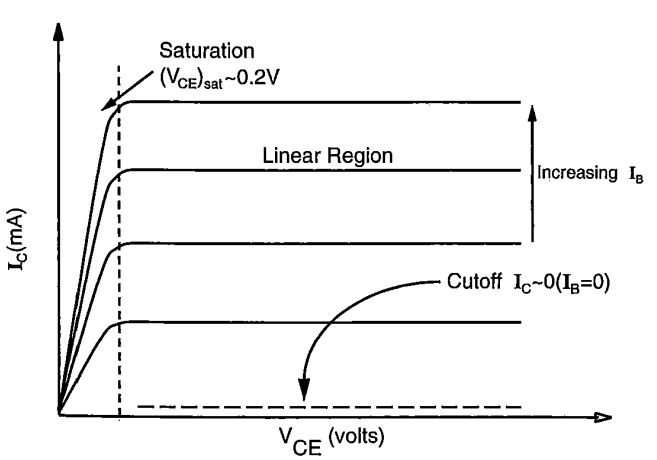


FIGURE 2.8 Bipolar junction transistor (BJT) characteristic curves.

the cut-off and saturation regions are used in digital systems to represent binary "1" and binary "0."

2.3.3 Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

The *metal oxide semiconductor field effect transistor* (MOSFET) is a counterpart to the BJT. Unlike the BJT, the current in the MOSFET is controlled by an input voltage rather than a current. MOSFETs are unipolar devices because the current flow only involves one carrier type, either holes or electrons. The MOSFET is a three terminal device with the structure illustrated in Figures 2.9a and 2.9b. The NMOS (or N-channel) device in Figure 2.9a has three electrical contacts: 1) two are to the N-type regions, termed the source and drain, which are formed by the diffusion of Group V impurities into the P-type substrate; 2) the third contact, called the gate, is formed by deposition of metal (or polysilicon) material on a thin oxide layer grown on the surface of the substrate. The PMOS (or P-channel) device has complementary structure to the NMOS.

The operation of the MOSFET can also be categorized in three regions: cut-off, triode, and saturation. However, it is important to note that the saturation region of the MOSFET is not the same as the saturation region of the BJT. The MOS transistor characteristics are illustrated in Figure 2.10. The drain current, I_D, is controlled by the potential applied between the gate and the source V_{GS}. When V_{GS} is less than a certain threshold voltage V_T, there is no current flow in the device and it is said to be in the cut-off region. As V_{GS} increases, the current starts flowing from the drain to the source. Each curve in the I_D vs. V_{DS} plane of Figure 2.10 corresponds to the current flow in the device at constant V_{GS}. For a particular value of V_{GS}, the current is a function of V_{DS}. An initial increase in V_{DS} results in a significant increase in the drain current, I_D, which is called the triode region or the amplification region. However, a saturation point is reached where the drain current can no longer be increased by increasing the V_{DS} potential. An increase in the drain current can only be accomplished by increasing V_{GS} and therefore by moving to higher curves in the graph. The MOSFET currents are given by:

$$I_{DS} = K[2(V_{GS} - V_T) - V_{DS}]V_{DS} \quad \text{in the triode region, and}$$

$$I_{DS} = K[V_{GS} - V_T]^2 \quad \text{in the saturation region,}$$

where $K = C_{ox}(\mu W/2L)$, μ is the electron or hole mobility, W and L are the channel width and length, and C_{ox} is the gate capacitance per unit area.

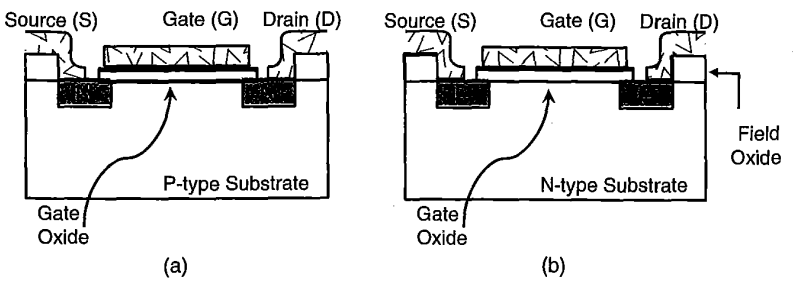


FIGURE 2.9 Metal oxide semiconductor field effect transistors: (a) N-channel and (b) P-channel.

Samsung, EX1028, p. 56

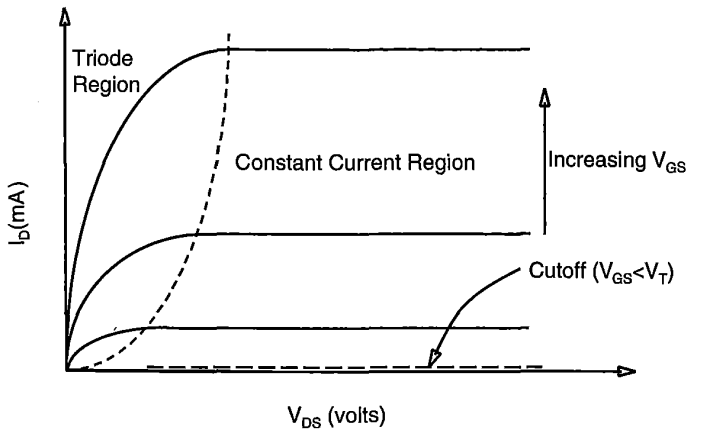


FIGURE 2.10 MOSFET characteristic curves.

2.3.4 CMOS Circuits: the CMOS Inverter

The currently prevalent electronic technology combines NMOS and PMOS devices in the *complementary* MOS (CMOS) family of circuits. The basic element is the CMOS inverter, illustrated in Figure 2.11a, where the common gate and drain connections serve as input and output terminals respectively. If V_{IN} is either +V or 0 volts (ground), the output voltage is the opposite: 0 or +V respectively, as shown in the "transfer characteristic" of Figure 2.11b, where V_{TN} and V_{TP} are the threshold voltages required before the devices start to turn on. These two input/output conditions are used for digital logic representations of "1" and "0," and lead to the representation of the two transistors as switches. In Figure 2.11c, the +V logic 1 input establishes a low resistance conducting channel in the bottom NMOS device, which is therefore represented as a closed ("on") switch, and "opens" the top PMOS "switch," or turns it "off." For many purposes, this view is an adequate one for the initial analysis of more complicated CMOS circuits. Note the salient feature that NO current flows in EITHER logic state, because one of the switches is always OFF.

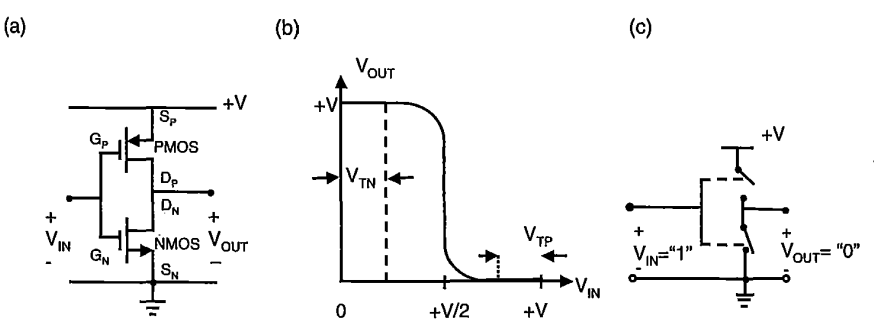


FIGURE 2.11 The CMOS inverter: (a) circuit, (b) transfer characteristic and (c) equivalent switch circuit.

Samsung, EX1028, p. 57

As the CMOS inverter "switches" between states, the input exceeds the thresholds of both devices for a short time, and both transistor "switches" are partly ON, so current flows. This switching behavior is shown in Figure 2.12 for a ramp-step input. The current pulse height is $K(V/2 - V_T)^2$. The gate capacitances take a short while to charge and discharge, and to establish the channel charges, so the output response has an additional small delay of typically less than a nanosecond (ns = 10^{-9} second), in addition to that shown. The current pulse shown accompanies every transition, whether 1→0 or 0→1.

On-chip interconnection wiring also has enormous significance to the high-frequency limitations of CMOS logic systems. If the distance from one circuit's output to the next input is short, the interconnection might be represented by the wiring resistance (R_2) in Figure 2.13a, which limits the current available to charge the input gate capacitance (C_{G2}) of the following circuit, and slows the input voltage transition. (CMOS technology is currently changing from aluminum interconnection wiring to copper, which will reduce this effect.) Signals also take a finite time to travel along these interconnection lines, even if they do so at close to the speed of light. For longer lines, if this time (T) is greater than the input or output signal rise and fall times, then the interconnections must be represented by transmission lines, as in Figure 2.13b. Similar effects are seen at both the package and PWB levels, and will be explored more fully in the next chapter.

Whether the metal used on the chip is Al or Cu, one of the requirements of the package wiring is to provide a stable, low resistance, and mechanically strong bond to the chip input and output contact pads. The metallurgical issues involved in satisfying these three requirements will be explored in a later chapter.

Unipolar MOSFETs may operate slower than BJTs due to gate capacitances; however, they dissipate small amounts of power, since zero current flows in the quiescent state, and also occupy significantly less area compared to the BJTs. The small area of the

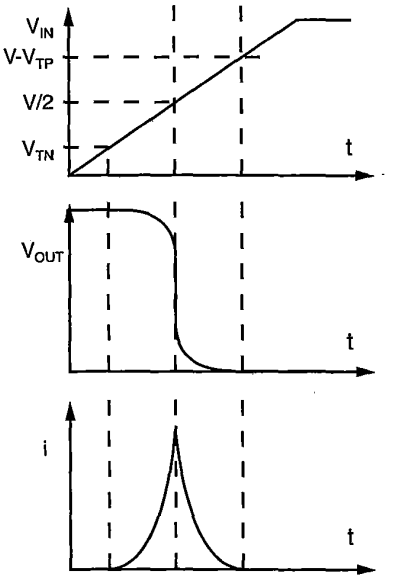


FIGURE 2.12 CMOS inverter switching waveforms.

Samsung, EX1028, p. 58

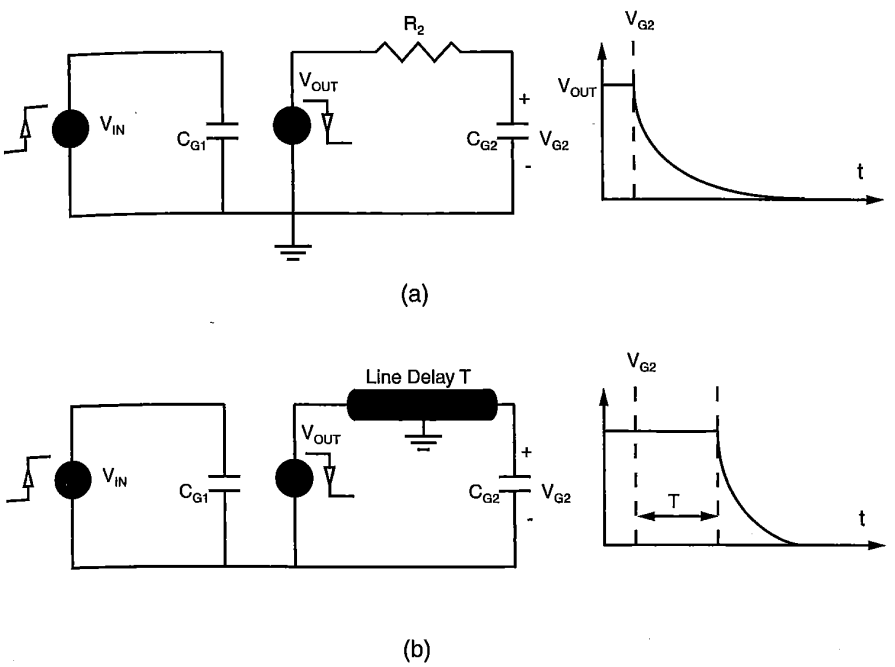


FIGURE 2.13 Transient responses of the CMOS inverter with short interconnections (a) and long interconnections (b).

MOSFETs helps in increasing the density of transistors on a chip. MOSFETs are also cost-effective in terms of fabrication. In microelectronics, BJTs are mostly used in applications where speed is more important than the area or power consumption. An integrated circuit may contain varieties of devices in addition to those discussed here.

2.4 INTEGRATED CIRCUITS

2.4.1 What Is an Integrated Circuit?

The semiconductor materials constitute the devices, and the devices are interconnected to form functional circuits. An integration of many such circuits or components on a single chip is called an integrated circuit (IC). The advances and the functionality of modern electronics would not have been possible without the development of ICs. Before the advent of ICs, discrete components such as transistors, diodes, capacitors, resistors, inductors, etc. were mounted on a *printed wiring board* (PWB) to form a circuit block, which was then connected to other circuit blocks to build a completely functional unit.

ICs enabled one to monolithically integrate all of these blocks on a single chip resulting in miniaturized products that became low in cost and high in reliability. Typical sizes of an IC can be as small as 1 mm to 30 mm on a side. In functionality, an IC can be a single component such as a power amplifier or a power transistor to a fully integrated microprocessor used in modern PCs and high performance servers and workstations. Theoretically, there is no limit to integrating multiple functions on a single chip, giving

Samsung, EX1028, p. 59

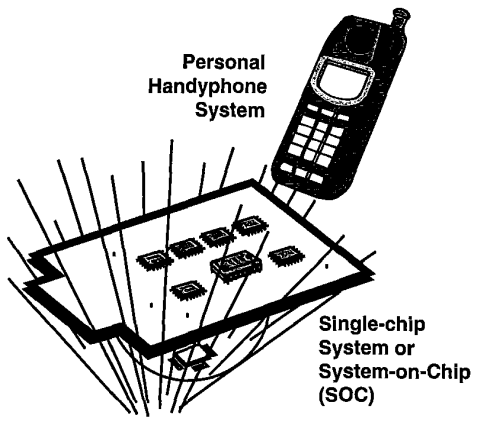


FIGURE 2.14 SOC design concept.

rise to concepts such as system-on-chip (SOC). The SOC is a concept where electrical, optical, mechanical, chemical, and biological devices will be integrated together on a single chip. Figure 2.14 shows a personal handyphone system design by the SOC concept, where various components shown in the figure are integrated on a single chip.

2.4.2 Types of Integrated Circuits

There are many types of integrated circuits. They are classified by their material and composition, degree of integration or number of transistor elements, principles of operation, manufacturing method, device type, etc. Figure 2.15 illustrates the many types of ICs based on different applications. Active components such as transistors and diodes,

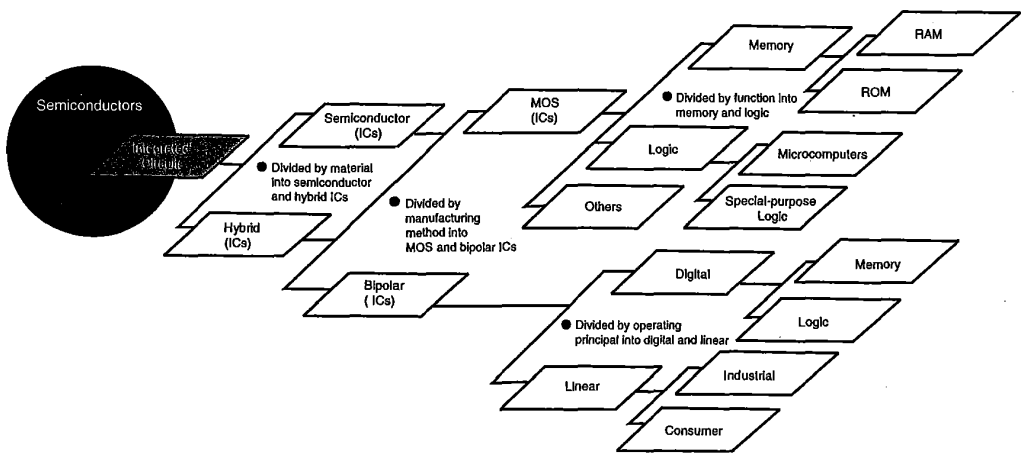


FIGURE 2.15 Types of integrated circuits.

Samsung, EX1028, p. 60

and passive components such as resistors, capacitors, and inductors are fabricated on a small substrate and interconnected to form a functional circuit. ICs are roughly divided into two categories: (1) hybrid and (2) semiconductor.

A *hybrid* IC (HIC) contains interconnected diodes, capacitors and resistors fabricated on a single plastic, ceramic or insulated aluminum substrate and has the capability of performing a complete electronic circuit function. It is called hybrid since it involves two or more material types. Semiconductor ICs only contain devices fabricated from semiconductor materials. Except for discrete components, virtually all microelectronic products can be addressed as ICs as seen in Figure 2.15.

2.4.3 Integrated Circuit Evolution

The number of elements incorporated into a chip is called the degree of integration. The first IC in the world in the 1960s had only about 10 elements. Since then, the degree of integration has progressed at a startling rate, roughly quadrupling every three years. As ICs grow larger in scale, they graduate from LSI to VLSI. Today, the degree of integration has reached the era of *ultra large scale integration* (ULSI), where the degree of integration is on the level of 20 to 100 million transistors per chip. Within a decade this integration is projected to reach a billion transistors on a single chip, as illustrated in Figure 2.16.

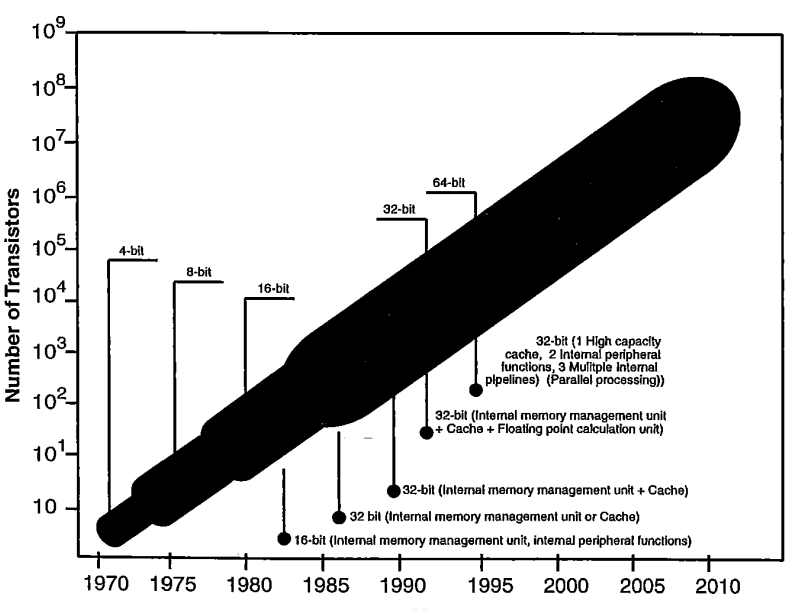


FIGURE 2.16 Evolution of the microprocessor.

Samsung, EX1028, p. 61

2.4.4 Why Does Integration Progress?

As ICs achieve higher densities, the number of elements increases, thereby raising their information processing capacity, improving performance, decreasing size and weight, and lowering cost. Furthermore, automation of the manufacturing process drastically reduces the defect rate and raises reliability. For these reasons, the integration of ICs continues to progress. Over the past ten years, IC cost has sharply declined while performance has greatly improved. In other words, the cost-performance ratio has shown a huge improvement. No other product of today's industry has achieved such a triumph. A key element of this remarkable progress is due to lithography, a process by which the devices are fabricated on a chip. State of the art lithographic technology can produce MOSFETs as small as 180 nanometers (nm).

How will lithographic technology continue to evolve in the future? At present, circuit patterns are created on silicon surfaces by the use of light. The light most commonly used for this purpose is ultraviolet (UV) with a wavelength of 356 nm. In the future, however, light of even smaller wavelengths will be needed to achieve even finer device geometries projected by the semiconductor industry. This will produce circuit feature sizes below optical wavelengths as indicated in Figure 2.17.

2.4.5 IC Materials and Processes

Current ICs use aluminum (Al) conductors and silica (SiO_2) dielectrics, as practiced since the days of the first IC. However, new materials have been introduced recently. The

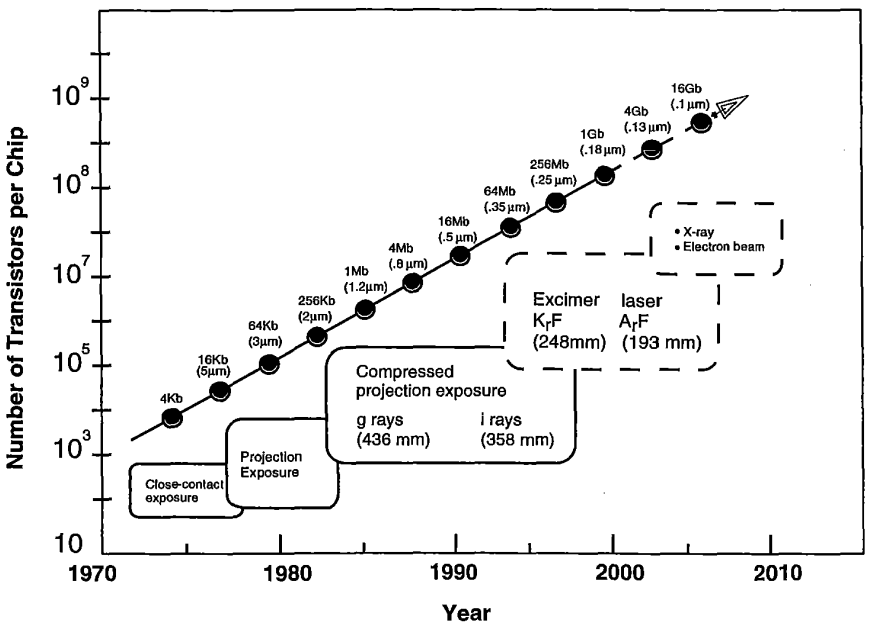


FIGURE 2.17 Logic and memory IC evolution.

Samsung, EX1028, p. 62

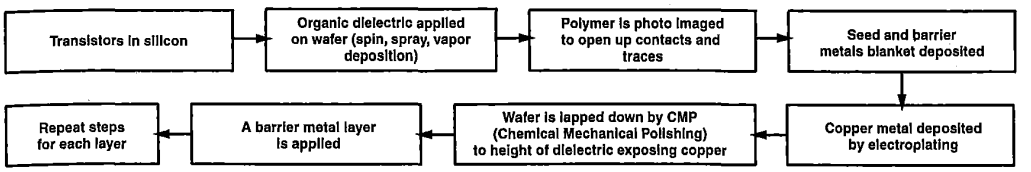


FIGURE 2.18 IC fabrication by chemical mechanical polishing.

driving force for these new materials is improved electrical performance at higher operating frequencies. Specifically, it is the reduced gate delay resulting from the higher electrical conductivity of copper vs. aluminum and the lower capacitances achievable with lower inner level dielectric constants (ε_r) of 2.5 vs. 4.0. Other advantages of copper and low dielectric constant dielectrics include less crosstalk, lower failure rates from electromigration, and the potential for much lower processing costs based on reduced process steps, higher yield, and less capital equipment. The technique for depositing copper and the entire imaging process is expected to change radically from today's Al fabrication process. The new process is called "Copper Damascene" from the Syrian city Damascus where pottery was decorated in 2000 B.C. by making trenches in moist clay, sprinkling gold dust and colored minerals into the trenches, kiln drying, and polishing down the surface. The analogous process on silicon is generically described in Figure 2.18.

The transition to low ε_r and copper traces began progressively. The inner layer dielectric materials and the process for depositing them are evolving with the first low inner level dielectrics being fluorinated silicate glass with organic materials soon to follow as indicated in Table 2.2.

Figure 2.19 illustrates one of the first copper Damascene ICs fabricated, based on the materials and processes described above.

TABLE 2.2 Lower dielectric constant materials for IC fabrication.

	Dielectric Constant						
	4.0	3.5	3.0	2.5	2.0	1.5	1.0
Chemical Vapor Deposition	SiO_2	Silicon Oxyfluoride	Methyl-doped SiO_2	Nanoporous Silica-based	PTFE		
Spin-on			Hydrogen Silsesquioxane (HSQ)	BCB, Silk	Porous Organosilicate		
				Methyl Silsesquioxane (MSQ)	Nanoglass		
			Silicon-based Polymer		PTFE		

Samsung, EX1028, p. 63

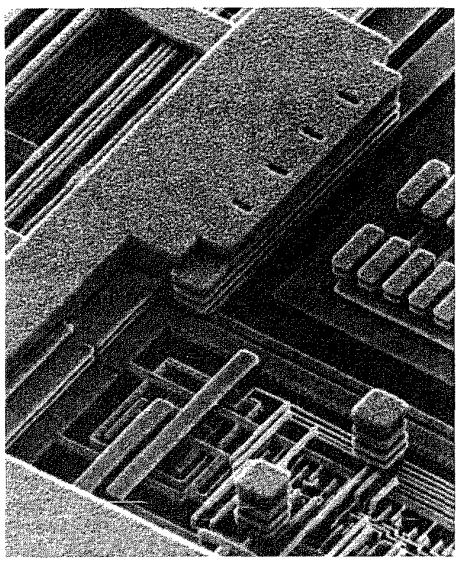


FIGURE 2.19 First copper interconnection in an IC. *(Courtesy of Prismark)*

2.4.6 Production of ICs

About 60 billion ICs were produced in the year 1999. The IC production can be divided into four major categories: 1) CMOS microprocessors, 2) CMOS logic, 3) linear/analog, and 4) memory ICs, as indicated in Figure 2.20.

While discrete semiconductors account for over 75% of the total semiconductor unit shipments, they represent less than 15% of the total dollar value for the semiconductor industry. Included in the discretes category are devices such as diodes, transistors, rectifiers and thyristors, which are still produced in billions of units. Table 2.3 lists the various devices and their unit consumptions in 1998.

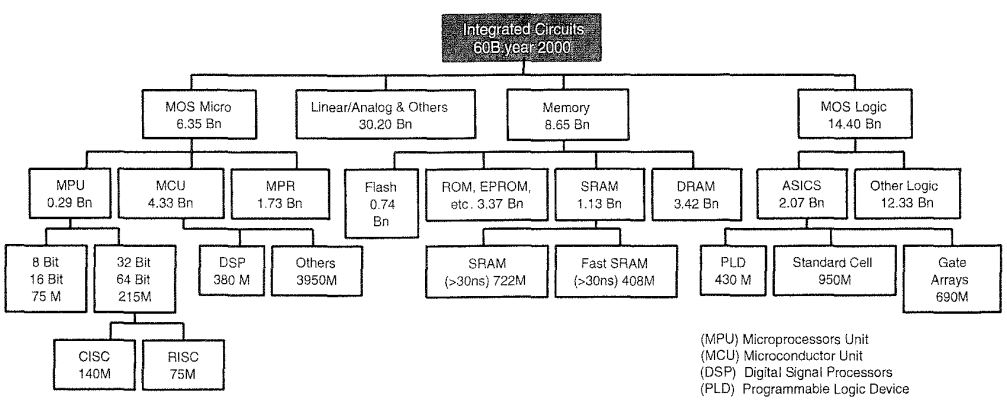


FIGURE 2.20 Annual IC production.

Samsung, EX1028, p. 64

TABLE 2.3 Production of discrete semiconductors (1998).

Discrete Semiconductors		Billions of Units
Diodes	Total	70
	Small signal junction diode	50
	Voltage reference/regulator diodes	18
	Transient protection device	2.0
Small Signal Transistors	Total	71
	Bipolar small signal transistors	67
	Field effect transistors	4.0
Power Transistors	Total	10.2
	RF and microwave power transistors	0.1
	Regular power transistors	5.0
	Field-effect general purpose transistors	5.0
	Insulated gate bipolar	0.1
Rectifiers	Total	22
Thyristors	Total	2.0
Other Discretes	Total	4.0
Optoelectronics	Total	18
Total Discretes and Optoelectronics		197.2

2.4.7 Wafer Processing

Integrated circuits are processed on a large piece of semiconductor substrate called a wafer. Wafer sizes can vary from three to 12 inches in diameter. Depending on the size of an individual IC, there may be hundreds to thousands of ICs on a wafer. After the IC's fabrication process on the wafer is finished, the wafer enters a wafer dicing process. In this process, the wafer is cut by a fully automatic dicing saw whose blade is tipped with diamonds; individual ICs or chips are the result at the end of the dicing process. This marks the completion of the IC fabrication process where a piece of semiconductor material is transformed into a functional microelectronic part. To be able to use these ICs, they have to be packaged, tested, and assembled on the system board.

2.5 IC PACKAGING

Figure 2.21 below describes the entire process of forming a system from the wafer. The single most important step in this process is the packaging of the IC.

The functions of an IC package are to protect, power, and cool the microelectronic device described above and to provide electrical and mechanical connection between the part and the outside world. Each chip has its unique packaging process, but for the purpose of discussion, a generic flow for Dual-in-Line Packages (DIP) is shown in Figure 2.22. DIPs were the first complex package solution developed shortly after the invention of ICs in the 1960s. Many package solutions have been developed since, but because of its low cost and high reliability, it continues to be used even four decades after its first introduction. The package is generally fabricated independent of the ICs. When the IC

Samsung, EX1028, p. 65

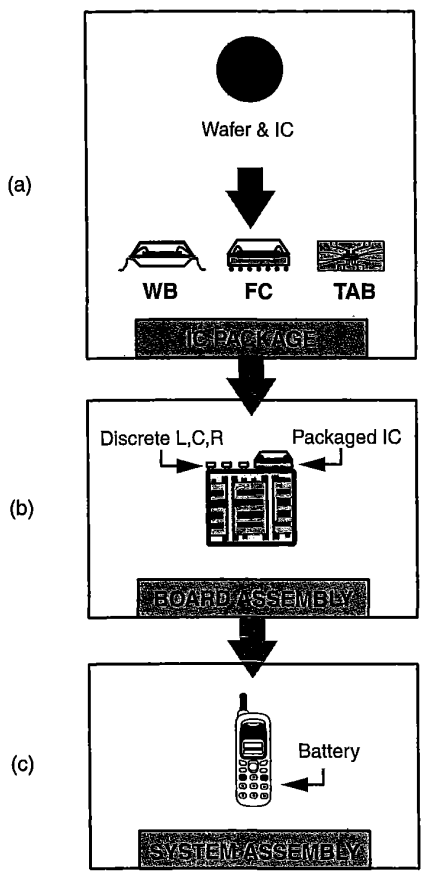


FIGURE 2.21 Wafer to system process: (a) IC package, (b) board assembly, (c) system assembly.

or the die is ready to be packaged, it is bonded to a package. The input/output (I/O) connection pads on the IC are connected to corresponding terminal pads on the package by the use of gold or aluminum wirebonds. The I/Os are connections that pass electronic signals in and out of the chip. Once the die is wirebonded to the package, the entire structure is sealed or molded with a plastic material such as epoxy resin. Only the package terminals or pins are seen, and everything else is completely sealed. Once the IC is sealed, the package is marked with the manufacturer's name, model number, and other identification information.

2.5.1 Different Types of IC Packages

Since there are so many types of ICs and IC requirements, as seen previously in Figure 2.20, it is impractical to have one package solution for all ICs. To resolve this problem, many types of IC package technologies have been developed that vary in their structures, materials, fabrication methodology, bonding technologies, size, thickness, number of

Samsung, EX1028, p. 66

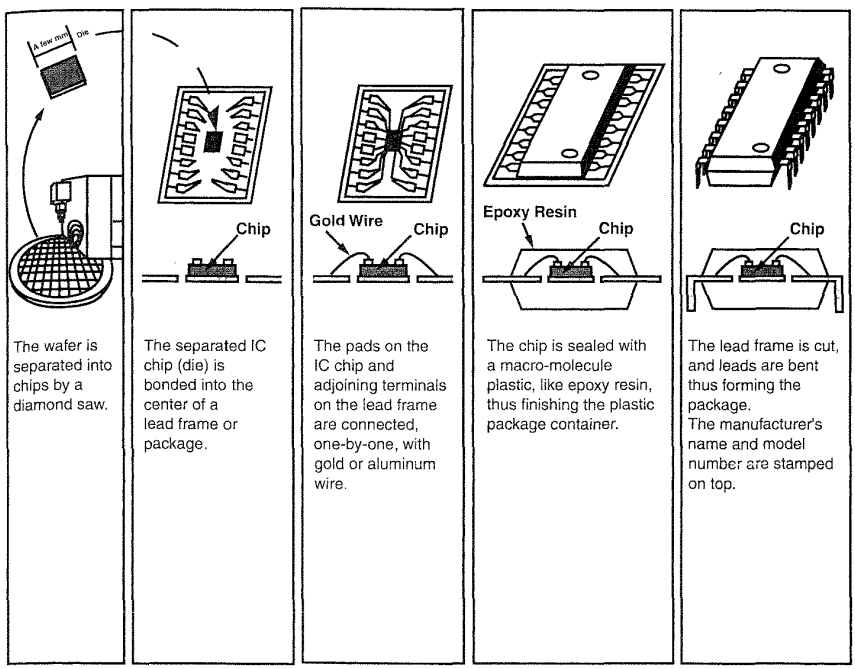


FIGURE 2.22 An example of dual in-line IC packaging with wirebonding.

I/O connections, heat removal capability, electrical performance, reliability, and cost. In general, IC packages can be classified into two categories: 1) through-hole, and 2) surface mount. These two categories refer to the methodology used in assembling the packages to the printed wiring board (PWB). If the packages have pins that can be inserted into holes in the PWB, they are called through-hole packages. If the packages are not inserted into the PWB, but are mounted on the surface of the PWB, they are called surface mount packages. The advantage of the surface mount package, as compared to through-hole, is that both sides of the PWB can be used, and therefore, higher packing density can be achieved on the board. These two categories of packages are shown in Figure 2.23. *Dual-in-line packages* (DIP) and *pin grid arrays* (PGA) are through-hole packages. In DIPs, the I/Os, or the pins, are distributed along the sides of the package. To achieve higher I/O connections, PGAs are used where the pins are distributed in an area array fashion underneath the package surface. Most of the advances and varieties are observed in the surface mount packages. The small outline (SO) package is the most widely used package in modern memory for low I/O applications because of its extremely low cost. The quad flat package (QFP) is an extension of the SOP with larger I/O connections. Both the SOP and QFP have leads that can be attached to the PWB. There are also leadless packages such as LCC and PLCC, but their usage is very limited. In the late 1980s, packages with solder balls were developed as an alternative to packages with leads. The solder balls can be placed underneath the surface of the package in an area array and significantly increase the I/O count of surface mount packages. Ball grid array (BGA)

Samsung, EX1028, p. 67

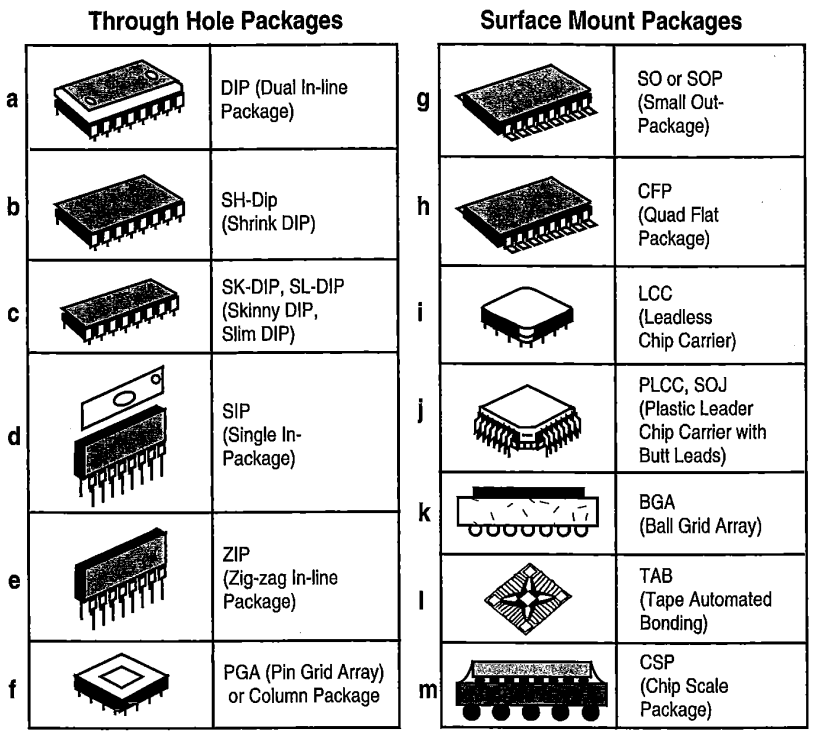


FIGURE 2.23 Through-hole packages and surface mount packages.

packages are an example of this technology. Smaller, thinner, and lighter packages are required in the modern age of portable and hand-held products. *Chip scale packages* (CSPs) have been developed to address these demands of modern electronics. For example, the DIP of the 1960s was roughly 100 times the size of the die. Since then, the evolution in packaging technology has reduced the ratio of package area to die area by 4 to 5. The CSP, by definition, is a package whose area is less than 1.2 times the area of the IC it packages.

2.6 SEMICONDUCTOR ROADMAP

To assure steady growth of the semiconductor industry in the future, a technology roadmap is put together and updated every two years by a group of semiconductor companies known as the Semiconductor Industry Association (SIA). This roadmap is called the International Technology Roadmap for Semiconductors (ITRS). The ITRS projects the technology requirements for six different product categories up to 10–15 years ahead. These product categories are low-cost, hand-held, cost/performance, high performance, memory, and harsh environment. Table 2.4 further elaborates these categories by listing typical products that may fall under each category.

Samsung, EX1028, p. 68

TABLE 2.4 ITRS product categories.

Product Category	Product Description
Low Cost	<$300: consumer products, microcontrollers, disk drives, displays
Hand-held	<$1000: battery powered products such as mobile and cellular products
Cost/Performance	<$3000: notebooks, desktop personal computers (PCs)
High Performance	>$3000: high-end workstations, servers, avionics, supercomputers
Harsh Environment	Under-the-hood and other hostile environment products
Memory	DRAMs, SRAMs

Recent projections for IC technology are shown in Table 2.5. Compared to today's IC, which has 20 million transistors, the ICs of the next decade are expected to have approximately a billion transistors and be almost 3 cm in size. For high-performance applications, this will require power of about 200 watts. Such an IC will have 5000 meters of wiring through eight to 10 levels of metal and will require as many as 8258 signal I/O connections. These are all phenomenal technology developments that are expected to take place in our lifetime.

2.6.1 Most Important IC Packaging Parameters

Referring to Table 2.5, which lists the IC roadmap parameters, the three most important parameters for packaging ICs are: (1) I/O which controls the pitch of the IC package as well as the wiring needs at the system level, (2) size of the IC which controls the reliability of the IC to package connection, and (3) power which controls heat dissipation properties of IC and system-level packaging.

The data of Table 2.5 are represented as semi-log Moore plots in Figure 2.24 (where all plots relate to the left-hand ordinate scale, except "gates/chip"). The power per die

TABLE 2.5 ITRS IC technology roadmap.

	1992	1995	1998	2001	2004	2007
Feature Size (μm)	0.5	0.35	0.25	0.18	0.12	0.07
Gates/chip	300K	800K	2M	5M	10M	100M
Transistors (cm^{-2})	0.01B	0.04B	0.1B	0.22B	0.6B	2.1B
Chip Size (mm^2)						
Logic/Uniprocessor	250	400	600	800	1000	1250
DRAM	132	200	320	500	700	1000
Maximum Power (W/Die)						
High Performance	10	15	30	40	40–120	40–200
Portable	3	4	4	4	4	4
Power Supply Voltage (V)						
Desktop	5	3.3	2.2	2.2	1.5	1.5
Portable	3.3	2.2	2.2	1.5	1.5	1.5
No. I/Os	500	750	1500	2000	3500	5000

Samsung, EX1028, p. 69

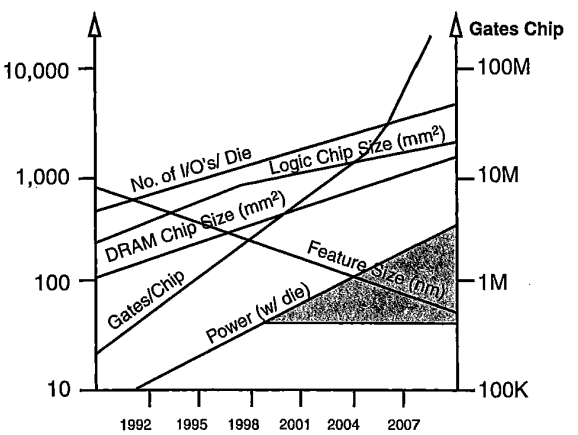


FIGURE 2.24 Semi-log Moore's plots of the Roadmap data.

split represents two extreme viewpoints, one that cooling technology has reached a limit, and the other that new techniques will emerge to maintain the past rate of progress as predicted by the Moore plot. The continued use of Moore's Law projections, which have served the semiconductor industry so well in the past, may falter in the next decade as current technologies approach their fundamental limits. Continued progress will require new technologies and new techniques.

Rent's Rule is an empirical law (i.e., one based on observation, with no theoretical foundation) with great significance for microelectronics packaging. The Roadmap I/O (input/output) data has been re-plotted in log-log Rent's Rule format in Figure 2.25. (The deviation from linearity corresponds to the Roadmap's jump in gates/chip in 2007, with no corresponding I/O increase predicted.) The empirical parameters α and β are different for each technology (e.g., memory, random logic, microprocessors).

The electronic packages have to be consistent with IC roadmap projections. One section of the ITRS specifically defines the packaging requirements of these ICs used in the six different product categories described above. The most demanding projections for some of the important parameters such as cost, power, I/O count, and operating frequency are given in Table 2.6 for low-cost and high-performance product categories. These two categories are chosen, because they provide the lower and upper limit for the parameters. The low-cost category has the lowest cost per pin requirement, whereas cost is not as demanding for high performance products. On the other hand, the high performance products provide the upper level requirements for power, I/O count, and electrical performance. The power usage by these products is projected to be as high as 129 watts by 2002 and 183 watts by 2014. The I/O counts also increase from 2248 to 8758, and the package is required to operate in the multi-gigahertz (GHz) range (1 GHz = 1000 MHz). These are the projections of the ITRS and are not an indication of available package technologies. In fact, while projecting these numbers, the Roadmap clearly states that no known solutions exist to address these requirements of future ICs. The Roadmap projections are to be used as guidelines by the package designers to develop novel technologies that can meet the requirements of future ICs.

Samsung, EX1028, p. 70

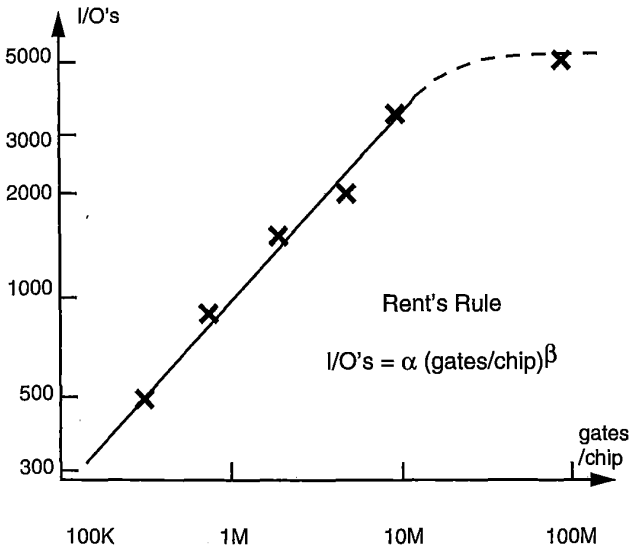


FIGURE 2.25 Rent's rule plot of roadmap data.

The status of current package technologies is explored in Figure 2.26, where the emphasis is placed on the history, with a modest peek into the future. There is also a backward extrapolation to 1980, so a sequence of four major technologies can be represented. These are represented schematically by the chip-to-package and package-to-board connections, except for the *direct chip attach* (DCA) example where the "package" is by-passed, and the chip is directly attached to the board. The chip-to-package connections are all assumed to be wire-bond. The DIP package is mounted to the PWB in the near-obsolete *pin-through-hole* (PTH) configuration, while both the QFP and chip scale package (CSP) are attached by *surface-mount-technology* (SMT). The SMT attachment of the CSP, however, is by ball grid array (BGA), rather than by any of the QFP peripheral leaded configurations.

TABLE 2.6 ITRS packaging requirements.

Year	2002	2005	2008	2011	2014
Low Cost					
Cost (Cents/pin)	0.34–0.77	0.29–0.66	0.25–0.57	0.22–0.49	0.19–0.42
Power (Watts)	2.0	2.4	2.5	2.6	2.7
I/O count	101–365	109–395	160–580	201–730	254–920
Performance (MHz)	100	100	125	125	150
High Performance					
Cost (Cents/pin)	2.66	2.28	1.95	1.68	1.44
Power (Watts)	129	160	170	174	183
I/O count	2248	3158	4437	6234	8758
Performance (MHz)	800	1000	1250	1500	1800

Samsung, EX1028, p. 71

Package	DIP	QFP	CSP	Flip Chip
Top View Showing Chip to Package Connections	DIP	QFP	CSP	Flip Chip
Plane View Showing Package to Board Connections	PTH	SMT	SMT/BGA	DCA
Chip Size (mm × mm)	5 × 5	16 × 16	25 × 25	36 × 36
Chip Perimeter (mm)	20	64	100	144
Number of I/Os	64	500	1600	3600
Chip Pad Pitch (μm)	312	128	625	600*
Package Size (in. × in.)	3.3 × 1.0	2.0 × 2.0	1.0 × 1.0	1.4 × 1.4
Package Lead Pitch (mils)	100	16	25	24
Chip Area (mm^2)	25	256	625	1296
Feature Size (μm)	2.0	0.5	0.25	0.125
Gates/Chip	30K	300K	2M	10M
Chip Area/(70× Feature Size)^2	1.25K	200K	2M	16.4M
Maximum Frequency (MHz)	5	80	320	1.28 GHz
Power Dissipation (W)	0.5	7.5	30	120
Chip Power Density (W/cm^2)	2.9	4.8	9.3	2.0
Package Power Density (W/cm^2)	0.024	0.3	4.8	9.8
Supply Voltage (V)	5	3.3	2.2	1.5
Supply Current (A)	0.1	2.3	13.6	80
*Chip Supply Lead Current Density (A/mm^2)	5.23	143	11.08	31.43
*Package Supply Lead Current Density (A/mm^2)	1.6	11.5	11.08	31.43

*Calculation assumes: (1) bond wire or solder pad diameter, or square lead width, equal to half the pitch, and (2) one supply and ground connection lead per 100 leads.

FIGURE 2.26 Package lead pitch, frequency, power dissipation estimates and comparisons.

MEMS231425

CONFIDENTIAL BUSINESS INFORMATION

Samsung, EX1028, p. 72

It is important to note that the package sizes, given in inches, are progressively approaching the size of the chip, given in millimeters. The package lead pitch is the center separation of two adjacent leads on the package surface. The package lead pitch has to shrink considerably to about 2 mils or 50 μm (1 mil = 25 μm), in order to accommodate a higher number of lead or I/O counts as projected by the ITRS. The I/Os also have to be distributed in an area array fashion instead of peripherally. Similarly, the maximum frequency and the power dissipation also have to follow the roadmap projections closely. The BGA and DCA type technologies may prove to be possible solutions.

2.7 IC PACKAGING CHALLENGES

The package is already the bottleneck to the system performance, as demonstrated for the maximum digital system clock frequencies in Figure 2.27. The clock frequency on the IC is higher than that possible for the package. Therefore, the package limits the IC technology. The simple fact is that the on-chip silicon system can out-perform the speed capability of the package. Furthermore, as volume production techniques continue to drive the cost of bare silicon chips down, the cost of the packaging constitutes a greater and greater proportion of the total system cost. For both of these reasons, the challenges for future electronic packages are becoming extremely complex and the demand for microelectronics packaging engineers is growing.

These challenges of high performance and low-cost packages must be resolved by microelectronics packaging engineers and systems integration engineers. In order to find effective solutions to these problems, the engineers must acquire a solid educational background in multidisciplinary areas. These areas include electrical, mechanical, and thermal design, fabrication of systems based on materials (metals, ceramics, polymers), material interface challenges, testing, assembly, and package reliability. The field is inherently multidisciplinary, and the chapters in this book describe each of these disciplinary areas in more detail.

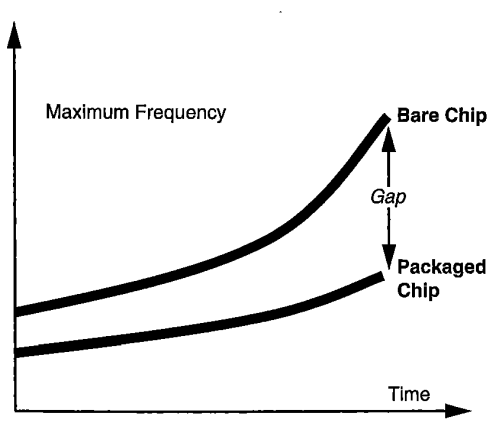


FIGURE 2.27 IC-Packaging performance gap.

Samsung, EX1028, p. 73

2.8 SUMMARY AND FUTURE TRENDS

The microelectronics industry is based on semiconductor materials, which can be modified into N- and P-type materials by impurity doping. Microelectronic devices such as PN diode, bipolar junction transistor (BJT), and metal oxide semiconductor field effect transistor (MOSFET) are constructed from semiconductor materials. The devices are interconnected to form circuits, and the many circuits constitute a fully functional part and result in integrated circuits (ICs). Many different types and categories of ICs are produced annually that vary in performance, size, and function. Each of these ICs, whether it be integration of a few transistors, millions of transistors, or a system-on-chip (SOC), have to be packaged to be usable in the practical world. The function of an electronic package is to protect, power, and cool the microelectronic chip and provide an electrical and mechanical connection between the chip and the rest of the system. Following the evolution of IC technology, package technology has also evolved from simple dual-in-line packages (DIP) to more complex structures such as chip scale packages (CSPs) and wafer-level packages. To date, developments in the semiconductor industry have been accurately predicted by the empirical Moore's Law. The success of this law over 30 years has enabled the industry to create "roadmaps" for future technological development, which are used to identify key "roadblocks" to be removed to ensure the industry's continuous success. The semiconductor roadmaps also project the package requirements of various types of IC technologies. To package these ICs and provide high performance and low-cost solutions is an exciting, challenging task for future generations of package and system integration engineers. In this chapter, the foundations of microelectronics and microelectronics packaging have been succinctly described. The technological progress of the ICs and the requirements placed upon packaging technology are also presented.

Packaging technology that is designed without keeping the IC in mind is useless. Likewise, the IC designed without keeping a potential package solution in mind is a futile effort. Traditionally, a generic package solution would apply to most ICs, and close collaboration between the IC and the package designer was not required. But as product specifications become more demanding, the system design without such collaborations cannot be accomplished.

In order to continue further into the book, the reader needs to understand the operation of the MOSFET device and CMOS circuits, and in particular, that:

- The MOSFET gate input is capacitive.
- The input capacitance slows circuit operation due to the RC time constant for short interconnections (e.g., on-chip).
- The input capacitance causes signal reflections on long interconnections (e.g., in the package or on the circuit board) which act as transmission lines. (This point is developed in the following chapter.)
- A quiescent CMOS circuit draws no current, since one series device is ON and the other OFF in either logic state.
- CMOS current is drawn only during switching, and the average current is therefore proportional to clock frequency.
- The system power dissipation therefore also increases with clock frequency.

Samsung, EX1028, p. 74

- CMOS clock frequencies also increase with decreases in device dimensions, which therefore lead to increased power dissipation.
- As device dimensions decrease with time, chip dimensions are generally increasing, both effects contributing to increased numbers of circuits/chip.
- Rent's Rule states that the number of I/O leads required by an IC increases with the number of chip circuits.
- So as time goes on, higher and higher I/O lead counts are required.
- Higher lead counts require closer lead spacings, which produces "fine-pitch" manufacturing problems, alleviated by the switch from peripheral leads to area array connections.
- Higher lead counts may also require thinner, weaker leads and decreased fatigue lifetimes.

2.8.1 What Is Next After Silicon?

Researchers have been talking about fundamental limits of silicon IC integration. Table 2.7 indicates one of the fundamental limits in IC integration. The 1 ps time constant listed for 1 μm technology is much less than the inherent 10 ps delay for signal transmission at one third of the velocity of light. The primary fundamental limit stems from the fact that as on-chip wiring becomes so long, the RC charging time constant increases on-chip delays by almost 100×, a "latency" effect further exacerbated by the higher resistance of thinner lines. RC and signal speed latencies are both reduced by the replacement of SiO_2 insulation with lower dielectric constant (low-k) materials.

The four technologies identified thus far are capable of correcting this problem. They are described below and illustrated in Figure 2.28.

The Optical Computer

This computer replaces electricity with laser light beams. Unlike wires, light beams can pass through one another, making possible three-dimensional microprocessors. An optical transistor has already been invented, but unfortunately the components are still rather large and clumsy, and the optical counterpart of a desktop computer would be the size of a car. Optical interconnect and data storage, however, are viable current technologies.

The DNA Computer

One of the most ingenious ideas being pursued is to compute using DNA, treating the double-stranded molecule as a kind of biological computer tape (except that instead of

TABLE 2.7 Giga-scale IC integration challenges.

Technology Generation	MOSFET Instrinsic Switching Delay	RC Response Time (L_{int} = 1 mm)
1.0 μm	~10 ps	~1 ps
.01 μm	~1 ps	~100 ps

Samsung, EX1028, p. 75

FIGURE 2.28 Four scenarios of what comes after silicon transistors.

MEMS23142

Samsung, EX1028, p. 76

encoding 0s and 1s in binary, it uses the four nucleic acids, represented by A, T, C, G). This approach holds much promise for crunching big numbers, and hence for the supercomputer market. But, a DNA computer is still an unwieldy contraption, consisting of a jungle of tubes of organic liquid, and miniaturization will be challenging.

Molecular and Dot Computers

Other exotic designs include the molecular computer and the quantum dot computer, which replace the silicon transistor with a single molecule and a single electron, respectively. But these approaches face formidable technical problems, such as mass-producing atomic wires and insulators, in order to take advantage of the potential size reductions. No viable prototypes of the molecular computer yet exist. The true quantum dot prototypes require cryogenic temperatures, but complex digital systems have already been built of modified single-electron transistors.

The Quantum Computer

The darkest horse to emerge in this race is the quantum computer, sometimes dubbed the ultimate computer. The idea is to direct a laser or radio beam on a carefully arranged collection of atomic nuclei, each of which is spinning like a top. As the beam bounces off the atoms, the spin directions can be changed. With spin up and down representing 1 and 0, complex computations can be performed by changing and analyzing spin orientations. The quantum computer is sensitive to the tiniest disturbance, since even a passing cosmic ray can change the spin orientations of its computational atoms, spoiling the calculation. At present, quantum computers can perform only trivial calculations on perhaps five atoms, and to do any useful work they would need to work on millions of atoms.

Clearly, none of these concepts are ready for mass production. Moore's Law is based on the continued scaling of device dimensions, which must clearly approach a discontinuity at the atomic level. If it were to somehow continue unabated, then by 2050 computers will be calculating well beyond 500 trillion bytes per second.

2.9 HOMEWORK PROBLEMS

1. (a) Find I_B, I_C, V_{CE} for the bipolar junction transistor circuit of Figure 2.P1,

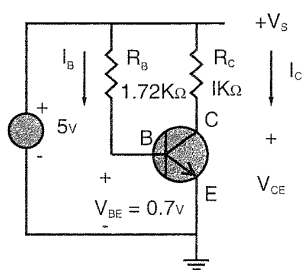


FIGURE 2.P1

assuming $I_C = 100\ I_B$.

(b) In Figure 2.8, label the axes, assuming the same BJT current gain in (a), and that the four curves correspond to $I_B = 10$, 20, 30, and 40 μA. Choose a 5 volt point towards the

Samsung, EX1028, p. 77

right hand end of the x-axis. Draw the line $I_C = (5v - V_{CE})/1\ K\Omega$ on the characteristics, and estimate where it crosses the I_C vs. V_{CE} curve for I_B in Figure 2.P1. Compare your result with part (a).

2. (a) Find V_{GS}, I_D, and V_{DS} for the N-channel MOSFET in Figure 2.P2,

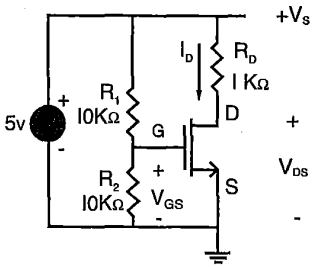


FIGURE 2.P2

with $V_{TN} = 1.5$ volts and $K_N = 1\ mA/V^2$.

(b) Repeat for $K_N = 5\ mA/V^2$. Interpret both results with reference to the MOSFET characteristics of Figure 2.10.

3. Refer to Figure 2.11b, and assume $+V = 5$ volts, $V_{TN} = V_{TP} = 1.5$ volts. Estimate the input noise margins for this circuit (i.e. the maximum noise level that will be attenuated by the circuit).
4. A chip using the CMOS system described in Problem 3, and with $K_N = K_P = 5\ mA/V^2$, dissipates 1 watt running at a clock frequency of 1 MHz. (a) How is the power dissipation affected if MOSFET device dimensions are reduced by 50%? (b) What if the power supply voltage is then reduced to 4 volts, and the clock frequency is raised to 5 MHz?
5. Calculate the gate RC time constants and line delays from 1992 to 2007 using the data of Table 2.5, assuming aluminum interconnect ($\rho = 2.8\ \mu$ohm-cm), SiO_2 ($\varepsilon_r = 3.9$) gate oxide, with all widths and thicknesses equal to the minimum feature size. Estimate line lengths for a square array of gates on the chip, and assume a signal velocity of 0.5× the speed of light.

$$(R = \rho \times \text{length} / [c/s\ \text{area}];\ C = \varepsilon_r \epsilon_o \times \text{area/thickness})$$

6. The Rent's Rule parameters vary with technology. According to Bakoglu (pages 416–420), the constants α and β are as listed in the table below. What technology is assumed for the I/O specifications in Table 2.5 and Figure 2.25?

Technology	α	β
Static memory (RAM)	6	0.12
Microprocessor	0.82	0.45
Gate array	1.9	0.50
Computer chip/module	1.4	0.63
Computer board/system	82	0.25

7. Locate the current SIA Roadmap on the Internet at www.semichips.org. Has it changed? What are the differences? What are the implications of these changes?
8. Find other packaging roadmaps on the Internet, for example, at http://www.lirmm.fr/~w3mic/Docs/roadmap.pdf. What are the differences between them (and the one in this chapter) and why do you think there are such differences?

Samsung, EX1028, p. 78

2.10 SUGGESTED READING

1. H. B. Bakoglu. *Circuits, Interconnections and Packaging for VLSI*, Chapter 2. Boston: Addison-Wesley, 1990.
2. J. Dally. *Packaging of Electronic Systems*, Chapter 2. New York: McGraw Hill, 1990.
3. R. Jaeger. *Microelectronic Circuits Design*. New York: McGraw-Hill, 1997.
4. W. Maly. *Atlas of IC Technologies: An Introduction to VLSI Processes*. Boston: Benjamin-Cummings, 1987.
5. J. D. Meindl. "Gigascale Integration: Is the Sky the Limit?" *IEEE Circuits and Devices*. Vol 12. No 6. 1996.
6. *1999 International Technology Roadmap for Semiconductors*. ITRS Update. <http://notes.sematech.org/NTRSRdmpmem.nsf/pages/Mhome>
7. A. Sedra, and K. Smith. *Microelectronic Circuits*. New York: Oxford University Press, 1998.
8. B. G. Streetman. *Solid State Electronic Devices*. 4th Edition. New Jersey: Prentice Hall, 1995.
9. N. Weste and K. Eshragia. *Principles of CMOS VLSI Design*. Boston: Addison-Wesley, 1985.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contribution of Chirag Patel of the Georgia Institute of Technology for critically reviewing this chapter.

Samsung, EX1028, p. 79

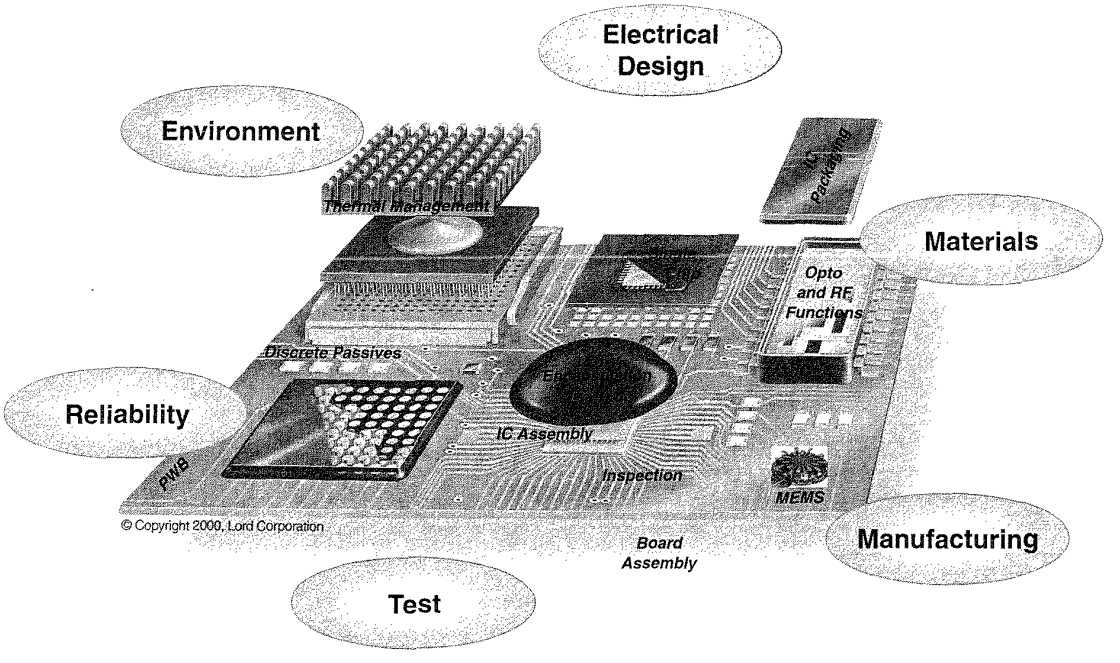
CHAPTER 4

FUNDAMENTALS OF ELECTRICAL PACKAGE DESIGN

Prof. Madhavan Swaminathan
Georgia Institute of Technology

Prof. Andrew F. Peterson
Georgia Institute of Technology

Prof. Joungho Kim
Korea Advanced Institute of Science and Technology, KAIST, Korea



Samsung, EX1028, p. 80

4.1 What Is Electrical Package Design?
4.2 Fundamentals of Electrical Package Design
4.3 Electrical Anatomy of Systems Packaging
4.4 Signal Distribution
4.5 Power Distribution
4.6 Electromagnetic Interference
4.7 Design Process
4.8 Summary and Future Trends
4.9 Homework Problems
4.10 Suggested Reading

CHAPTER OBJECTIVES

- Introduce the general philosophy, principles, and terminology of electrical package design
- Summarize the three major issues encountered by an electrical design engineer: ensuring signal integrity, ensuring an adequate power/grounding system, and designing for minimal electromagnetic emissions
- Develop the mathematical equations and tools needed for calculating critical design parameters

CHAPTER INTRODUCTION

In this chapter, the entire electrical design process is explored. The major issues facing the electrical design engineer from electrical circuit theory are presented with the basic equations needed to understand the cause and effect relations associated with these issues. The two major design tasks are those of 1) creating a signal distribution system and 2) creating a power supply and grounding system. The major challenge of electrical design is in overcoming the natural parasitic reactances present in every physical layout. Finally, the topic of electromagnetic interference, and the design principles necessary to ensure that the resulting system is not a source of interference, are presented.

Samsung, EX1028, p. 81

4.1 WHAT IS ELECTRICAL PACKAGE DESIGN?

Electrical package design is the process that defines the electrical signal and power paths through the package in a way that meets the overall system requirements. Ultimately, the end result of the design process is the geometrical layout of interconnects and the specification of materials and their geometries needed to meet the system requirements. In the future, it is likely that the process will also include the design of embedded passive components and embedded optical waveguides located within the package. Primarily, the electrical functions of a package are two-fold: provide signal distribution and power distribution.

4.1.1 Package Functions

Packages provide semiconductor ICs with signal and power distribution, physical support and chemical protection against the environment. The physical support helps during manufacturing, device testing, product storage, and attaching the chips to a board. The package also provides the capability to remove heat produced by the chip, to enhance the reliability of the packaging structures, and to secure the operation of the chip. Otherwise, the device performance of the chip degrades as the temperature of the chip rises. In addition, the package helps in distributing signals between chips and in supplying voltage and current to the circuits within a chip, as well as to other ICs in a given system, for their operation.

4.1.2 Electrical Functions

The electrical functions of a package are shown in Figure 4.1. In this figure, the package provides the signal path from the driver circuit in a chip to receiver circuits in other

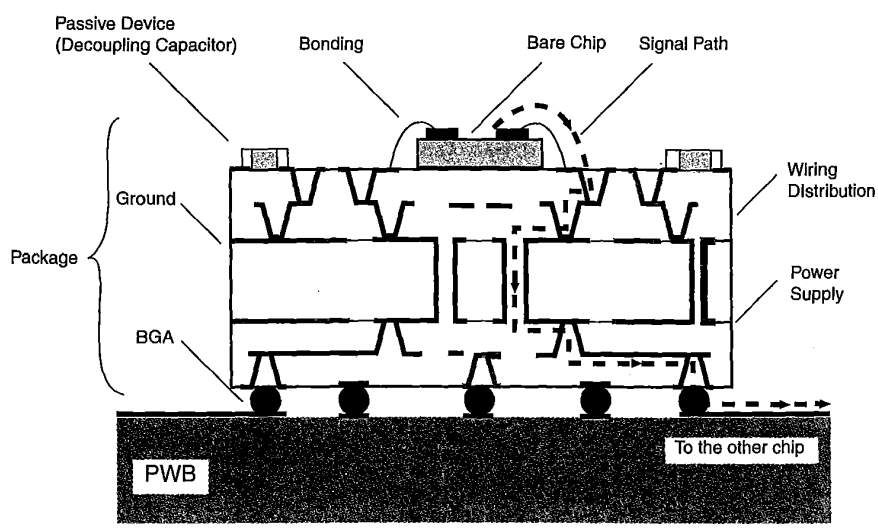


FIGURE 4.1 Electrical design issues.

Samsung, EX1028, p. 82

chips, provides the power supply and ground connection to the chips, and supports connections between passive devices. As shown by the dotted line in this figure, the signal path in the package is composed of bonding structures that act as an interface between the chip and package, transmission lines on the substrate, and vias that provide vertical connections. Through the signal path, chips can exchange data, address, clock, and control signals between each other.

For generating signals, chips require electrical power that is supplied through the package. Power enables transistor switching within the chip and data transfer between chips. The package supplying the power to the chips needs to have sufficient charge storage capability to supply the necessary amount of current with negligible power supply voltage fluctuations. The parasitic inductance of the power distribution and the bonding structures in the package, therefore, need to be minimized to ensure that they do not degrade the power supply. This is enabled by using power plane structures in the package that provide more capacitance and less inductance.

The first level package also serves as a space transformer from the chip interconnections to the board interconnections. The width and pitch of the interconnection lines within a chip have submicron dimensions. However, the interconnect dimensions on the board are much larger than a micron (around 50 μm). Therefore, the package acts as a space transformer between the fine line dimensions of the chip and the coarse line dimensions of the board. This translates into hundreds of I/Os for the chip, resulting in complex wiring inside the package for connecting the chip I/Os to the bonding pads on the board. To accommodate the wiring, several signal layers are often used in the package.

In the future, the package will contain lumped passive circuit elements such as resistors, inductors and capacitors. The capacitors in the package may be used to supply charge to the power line of the chip and to isolate the inductive parasitics of bonding structures. Lumped chip capacitors that are currently mounted on the surface or at the bottom of the package may be replaced with embedded capacitors buried inside the multilayer package. The embedded resistors may be used for the termination of signal lines to avoid signal reflections from the ends that cause unwanted high-frequency noise and signal propagation delay. Since the termination resistors of the signal lines determine the power consumption of the drivers, signal propagation delay and package complexity, their design can become very important.

Although the focus of this chapter is on the package, the overall design will not be successful unless the system, including the chips, package and *printed wiring board* (PWB) meet the intended electrical specifications. From one perspective, the electrical design of all three parts is similar, since one cannot tell from a schematic where a chip ends and the package or PWB begins. In practice, however, the range of parameters in each environment (chip, package, PWB) is so different that their designs are traditionally approached in relative isolation.

From Figure 4.1 it is clear that there are two aspects of electrical design that are of importance: providing suitable communication paths for signals and providing suitable channels for power distribution. The primary technical challenge for electrical design is driven by the frequency spectrum of the signals. At low frequencies, signal and power paths are easily realized since the physical geometry of the interconnects has little effect. At higher frequencies (1 GHz and up), the realization of appropriate interconnections is much more difficult. At high frequencies, interconnects are physically longer than the packets of energy routed along them, and their behavior depends on the properties of the

Samsung, EX1028, p. 83

materials and the electromagnetic fields that comprise the signal. Effects such as propagation delay, the characteristic impedance associated with interconnect configurations, and parasitic reactances determine the behavior of the signal. Hence, the degree of distortion of the signal and the time required for the signal to reach its destination are functions of the interconnect parameters. Since the frequency spectrum of the signal generally dictates the rate at which power is required by the chips, similar concerns apply to the power and ground paths as well.

In summary, electrical design of the package constitutes the following:

- Providing signal paths between the chips, including wiring that acts as a space transformer between the chip and board to match the dimensions of the interconnections and the design of embedded passive lumped components
- Providing suitable power distribution to enable the circuits to function

4.2 FUNDAMENTALS OF ELECTRICAL PACKAGE DESIGN

4.2.1 Ohm's Law

Electrical design involves electricity: the movement of electrons in a conductor. In packaging applications, electricity is used in both its fundamental forms: DC (direct current) and AC (alternating current). DC electricity is constant in time, while AC electricity varies in a sinusoidal manner over time. The power supplied to a chip is DC, while signals in and out of the chip vary with time. Most electrical signals are neither DC nor AC; instead, they vary with time in some nonconstant, nonsinusoidal manner. But for initial design purposes, it is usually convenient to work with DC and sinusoidal signals.

Electricity involves voltage and current. Voltage is measured in units of volts, where one volt is the work done moving one coulomb of charge one meter in the presence of one newton of force. An electron has a charge of -1.6×10^{-19} coulombs. Current is measured in amperes, where an ampere is one coulomb of charge per second of time. For DC signals, the ratio of voltage to current in a circuit is resistance, measured in ohms. In equation form, the voltage across a resistor is $V = IR$, where I is the current and R is the resistance. The relation $V = IR$ is also known as Ohm's Law. Resistance dissipates electrical power by converting it to heat; consequently, it is usually desirable to minimize the resistance of wires and conducting traces.

4.2.2 Skin Effect

At DC, current flows uniformly within a conductor. At higher frequencies, current tends to crowd along the surface of the conductors. This behavior is known as the *skin effect.* Because of the skin effect, the AC resistance of conductors is considerably higher than the DC resistance. For AC signals, the voltage and current sinusoids may also be out of phase with each other; this indicates the presence of capacitance or inductance within the circuit. Unlike resistance, which dissipates energy, capacitance and inductance act as energy storage mechanisms within a system. In essence, they capture energy over part of a cycle of the sinusoid and release it over a different part of a cycle. The formal definition of capacitance is the amount of charge a system of conductors can store per

Samsung, EX1028, p. 84

volt. The voltage and current at a capacitor are related by $I = C\,dV/dt$, where C is the capacitance in farads, and dV/dt is the time derivative of the voltage. The formal definition of inductance is less intuitive; inductance is the ratio of magnetic flux linked by a loop of current to the current. A more useful definition is embodied in the relation $V = L\,dI/dt$, where V is the voltage across an inductor, L is the inductance in units of Henrys, and dI/dt is the time derivative of the current.

4.2.3 Kirchhoff's Voltage Laws

Voltages and currents in an electrical circuit are related through Kirchhoff's Laws. *Kirchhoff's Voltage Law* (KVL) states that the voltage drops around any closed loop within the circuit must add to zero. Kirchhoff's Current Law states that the sum of the individual currents into a node of the circuit must add to zero. Kirchhoff's Laws combined with the fundamental equations of a resistor, capacitor, and inductor can be used to determine the voltages and currents throughout a circuit. The voltage and current relations in active devices (transistors) are generally more complex than these equations and usually nonlinear.

4.2.4 Noise

The presence of any undesired signal within a system can prevent that system from operating as intended. Such an undesired signal is known as noise. There are a variety of different sources of noise; some noise is present in all electrical systems due to inherent movement of atomic-scale particles within materials. However, the "noise" sources of most interest to an electrical designer are various non-ideal effects occurring within the circuitry of the system. Those effects can distort the ideal shape and amplitude of signals, and can also result in signals appearing in parts of the system where they don't belong. One such effect is parasitic capacitance or inductance. Parasitic capacitance is the inherent capacitance between any two pieces of conductors within a system. Similarly, parasitic inductance arises within any current-carrying structure. Any circuit layout involves some parasitic capacitance and inductance. These parasitics provide current paths that are not apparent from the physical connections within a circuit. Proper electrical design attempts to reduce parasitics to levels that will not disrupt system performance.

4.2.5 Time Delay

The combination of resistance with either capacitance or inductance introduces time delay into the system. Because of time delay, a signal traveling from one part of the circuit to another does not instantly arrive at the receiver. The time delay of a resistor and capacitor is given in terms of the time constant $\tau = RC$. RC delay is of considerable concern to chip and package designers, since most interconnections contain some resistance and appreciable capacitance. RC delay can be the limiting factor in the speed of a system. The presence of inductance and resistance also contributes a time delay in the form $\tau = L/R$. An L/R delay affects the ability of an on-chip power supply to instantly respond to a changing demand by the chip circuitry and contributes to *simultaneous switching noise* (SSN) within that circuitry.

Samsung, EX1028, p. 85

4.2.6 Simultaneous Switching Noise (SSN)

SSN is the fluctuation in signals that result from a brief reduction of the local DC supply voltage at some point within the system due to the power supply's inability to instantly respond. One remedy to the SSN problem is to place decoupling capacitors throughout the system to attempt to compensate for the system inductance.

4.2.7 Transmission Lines

Although the time delays associated with electrical signals can be modeled as RC and L/R delays, they are actually a consequence of the fact that electricity is carried by electromagnetic waves. Electromagnetic waves travel at the velocity of light in a particular material. In air, the velocity of light is about one foot per nanosecond. For most packaging applications, it is sufficient to consider EM waves that travel in one dimension. These waves can be modeled by voltages and currents on transmission lines. While interconnections within chips can usually be modeled as RC circuits, interconnections in printed wiring boards and in packages must generally be modeled as transmission lines. Any pair of conductors carrying a signal or power through part of a system can act as a transmission line.

Transmission lines account for the finite propagation velocity of electrical signals and also for the presence of reflected waves from environmental discontinuities. The ratio of the voltage to the current carried by a wave in a particular direction on a transmission line is the characteristic impedance of the line. The characteristic impedance of a transmission line is a function of the materials and geometry of that line. Discontinuities in the characteristic impedance (changes in materials, shape of conductors, etc.) cause partial reflections of the waves on the line. Reflections give rise to signals traveling the wrong way on an interconnect, which can be thought of as additional noise within the system. To prevent reflections, transmission lines must be properly terminated at their ends by "matching" the load impedance to the characteristic impedance of that line. Due to capacitive and inductive coupling, closely spaced interconnects act as coupled transmission lines and are vulnerable to crosstalk.

4.2.8 Crosstalk

Crosstalk noise is the result of a signal on one line inducing a signal on a nearby line, despite the absence of any physical connection. As explained above, crosstalk is caused by parasitic capacitance and inductance. Crosstalk is a complex phenomenon that will be discussed later in this chapter. Transmission line modeling can be used to predict and alleviate reflection and crosstalk noise.

4.2.9 Electromagnetic Interference (EMI)

Many of the effects described above are related to the presence of stray or parasitic capacitance and inductance within a system. Electrical design must include procedures for estimating these parameters and incorporating them into the system simulation. Often this involves electromagnetic simulation and modeling of subsystems and the extraction of SPICE-type circuit models to test the resulting designs with realistic time-varying

Samsung, EX1028, p. 86

signals. Undesired electrical effects that disrupt a system's performance or interfere with nearby systems are also known as *electromagnetic interference* (EMI). A number of design principles for reducing EMI to acceptable levels are described toward the end of this chapter.

4.2.10 SPICE Model

A number of general purpose circuit analysis tools exist for simulating electronic circuits. One of these, SPICE (*Simulation Program with Integrated Circuit Emphasis*), was developed in the early 1970s at the University of California at Berkeley and has spawned a number of successors, some in the public domain and some available through commercial software vendors. This type of simulation tool contains models for all common circuit elements and many active devices, and can perform time-domain and frequency-domain simulations. Although a variety of versions exist and are in widespread use, the generic name "SPICE" is commonly used to identify this type of tool.

The preceding discussion indicates that a substantial part of the electrical design process is devoted to ensuring the signal integrity of the system. Signal integrity is the aspect of the design process that ensures that noise levels, from all sources, are maintained below an acceptable threshold to ensure that the system functions as intended. Throughout the remainder of this chapter, the various concepts introduced above will be explained and integrated together to assist the reader in learning the fundamentals of the electrical design process.

Table 4.1 includes a summary of nomenclature typically used in electrical aspects of packaging.

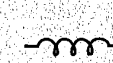
4.3 ELECTRICAL ANATOMY OF SYSTEMS PACKAGING

Electrical design involves the estimation of package performance. The electrical system specifications for the package include parameters such as delay, skew, loading, impedance, reflections, crosstalk and power/ground fluctuations. To estimate the electrical performance, circuit simulation methods are often used which require the circuit models of packaging structures. The circuit models are derived from the physical structure using electromagnetic simulations, analytical equations and/or high-frequency measurements. This section provides some details on the electrical equivalent of typical packaging structures.

The signal lines within the package are used for the transmission of signals between chips. Based on the characteristics of the signal lines, the driver and the receiver circuitry, they can cause delay, skew and reflections, which degrade the transmitted signal. To evaluate the signal line performance, they are represented by their characteristic impedance Z_0 and propagation velocity v_p, as shown in Figure 4.2. During signal transmission, the signal line can couple energy to adjacent signal lines, which results in crosstalk. Crosstalk can cause false switching of circuits and can increase delay. Crosstalk can be simulated by extracting the coupling between signal lines, such as mutual inductance L_m and mutual capacitance C_m, as shown in Figure 4.2. The chips are connected to the signal lines through bonding structures and vias. Similarly, signals exiting the package pass through pins and connectors. These structures add parasitics to the signal lines due to their resistance and reactance and can degrade the signal. The electrical representation

Samsung, EX1028, p. 87

TABLE 4.1 Electrical design nomenclature.*

Symbol	Nomenclature	Schematic	Units	Definition
R	Resistance		Ohms	The magnitude of the in phase opposition to the flow of current. $R = \rho L/A$
R	Resistor		Ohms	A component used to introduce resistance into a circuit.
ρ	Resistivity		Ohm * meter	Resistance of a conductor with unit length and unit cross-sectional area.
C	Capacitance		Farad	The concept of energy storage in an electric field. Restricted to physical dimensions and spacing of capacitor plates, as well as the property of the dielectric material.
C	Capacitor		Farad	Electrical device $C = \frac{\varepsilon A}{d}$ used to introduce capacitance into a circuit.
	Dielectric			Solid, liquid, or gaseous substance that affects the flow/propagation of electrical current.
ε	Dielectric Constant			The ratio of the capacitance of a capacitor with a given dielectric to that of the same capacitor having vacuum dielectric.
L	Inductance		Henry	The concept of energy storage in a magnetic field. Restricted to physical dimensions, spacing and number of coils, as well as the property of the magnetic material.
M	Mutual Inductance		Henry	Property that exists between 2 current carrying conductors when their magnetic lines of force link.
L	Inductor		Henry	Electrical device (usually conductor or metal) used to introduce inductance into a circuit.
	Transformer			A device that converts AC voltage and current to different levels at essentially constant power and frequency.
Q	Quality Factor			The ratio of reactance (combination of capacitance and inductance effect) to its equivalent series resistance (loss).
X	Reactance		Ohms	The phase quadrature opposition to current flow.
Z	Impedance		Ohms	The total opposition offered to the flow of an alternating current.
Y	Admittance		Siemens	Reciprocal of impedance.
H	Magnetic Field			State produced by electric charge in motion (electric current) and evidenced by a force exerted on a moving charge in the field.
Φ	Magnetic Flux		Weber	Term for lines of magnetism.
B	Magnetic Flux Density		Tesla	Measure of strength and direction of a magnetic field at a given point.
E	Electric Field			State of a region in which charged bodies are subject to forces by virtue of their charge, the force acting on a unit positive charge.
	Harmonic Frequency		Hertz	Integral multiples of fundamental frequency. For 50 Hz, the harmonic frequencies would be 100, 150, 200 . . .

Samsung, EX1028, p. 88

TABLE 4.1 Electrical design nomenclature* (*Continued*).

Symbol	Nomenclature	Schematic	Units	Definition
I	Current		Ampere	Time rate of change of charge.
DC	Direct Current		Ampere	A constant current.
AC	Alternating Current		Ampere	A sinusoidal current.
V	Voltage		Volt	The potential difference between two points.
P	Power		Watt	Time rate of change in the energy delivered to the network by the source. $P = V * I$.
	Semiconductor			A group of materials having conductivities between those of metals and insulators.
	pn-junction			Metallurgical interface of 2 regions in a semiconductor where one region produces positive charge carriers, and the other region produces negative charge carriers.
D	Diode			Two terminal solid-state semiconductor device that has a low impedance to current flow in one direction, but a high impedance to current flow in the other direction.
	Rectifier			A device that allows current to flow through it in one direction only (i.e., a diode).
BJT	Bipolar Junction Transistor			A multi-junction semiconductor device that, in conjunction with other circuit elements, is capable of current gain, voltage gain, and signal-power gain.
MOSFET	Metal Oxide Semiconductor Field Effect Transistor			Semiconductor device capable of voltage gain and signal power gain.
Tline	Transmission Line			A structure that allows for the wave/current/voltage propagation and transfers energy from one point to another.
	Microstrip Line			A planar transmission line consisting of a conductor strip printed or etched on a grounded dielectric substrate.
	Active Device			A device that requires a source of energy for its operation and has an output that is a function of present and past input signals.
	Passive Device			A device that does not require a source of energy for its operation.
	Amplifier			Active 2 port device with signal of higher amplitude than the input signal while retaining the essential signal characteristics of the input signal. $G = P/Pi$.
	Ideal Filter			A system that completely rejects sinusoidal inputs of the form $x(t) = A\cos(wt)$, where t is between negative infinity and infinity, and where w is in certain frequency ranges and does not attenuate sinusoidal inputs whose frequencies are outside these ranges.

Samsung, EX1028, p. 89

TABLE 4.1 Electrical design nomenclature* (*Continued*).

Symbol	Nomenclature	Schematic	Units	Definition
IC	Integrated Circuit			Small electric circuit containing large numbers of electronic devices such as transistors and packaged as a single unit with leads extending from it for input, output, and power supply connections.
	Antenna			A physical device for transmitting or receiving propagating waves.
	Oscillator			A source of sinusoidal electrical signals.
VFO	Variable Frequency Oscillator			An oscillator that can tune in frequency.
BFO	Beat Frequency Oscillator			A fixed oscillator.
	Mixer			A physical device that takes two or more input waveforms and combines them into one output waveform.
	Superheterodyne Receiver			Solves the problem of a fixed filter pass band by adding a variable oscillator, another mixer, and another image-reject filter.
	Speaker			A device that changes electrical signals into sounds loud enough to be heard at a distance.

*Information taken from the following: Dorf, Richard C. (1997). *The Electrical Engineering Handbook.* New York: IEEE Press.

of some of the parasitic structures is shown in Figure 4.2. Finally, the passive components that are mounted on the package or board have equivalent circuits that define them. One example is the decoupling capacitor, which is represented as a series *resistor-inductor-capacitor* (RLC) circuit in Figure 4.2.

The primary challenge associated with electrical design is dealing with the various non-ideal effects that impact electrical signal transmission. These non-ideal effects are typically those caused by the nonzero propagation delay of the actual electrical signal and the various parasitic reactances present in every circuit realization. For electrical design purposes, it is often convenient to model interconnects as transmission lines, similar in many respects to the cables used to route television, telephone, and digital communication signals throughout a community. The non-ideal effects associated with actual circuits can usually be modeled as noise that combines with the desired signals, resulting in distortion. Electrical designers must balance the trade-off between noise and layout complexity in order to produce a package that meets system specifications.

Samsung, EX1028, p. 90

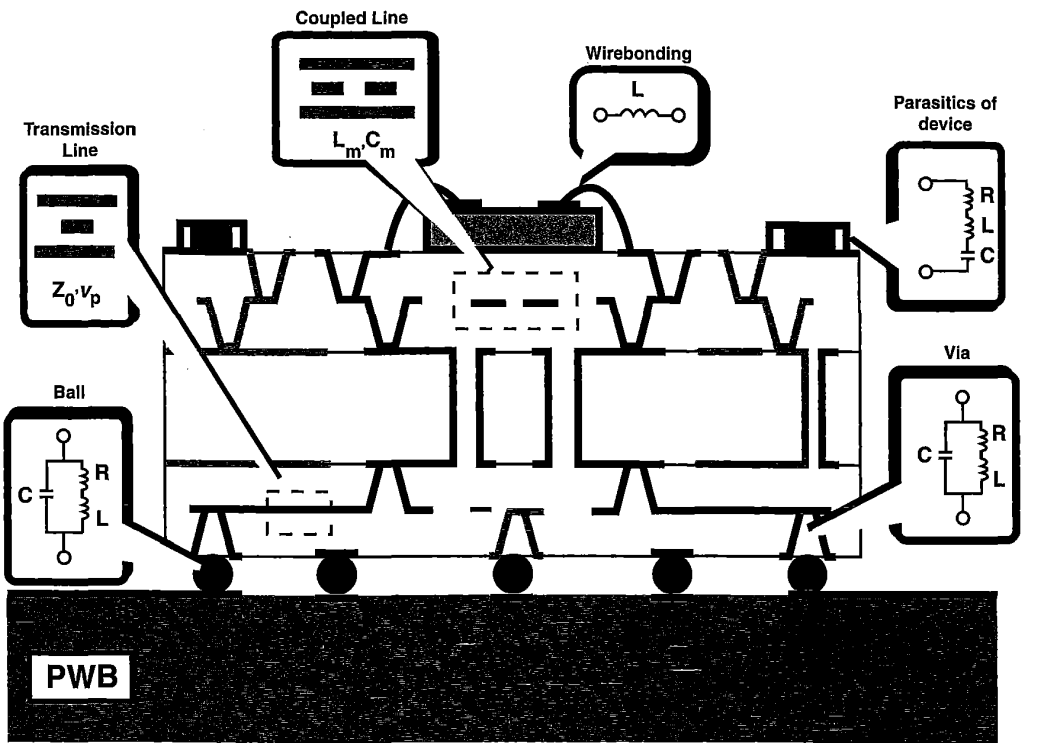


FIGURE 4.2 Package parasitics.

4.4 SIGNAL DISTRIBUTION

4.4.1 Devices and Interconnections

Signals are sent around a system to convey instructions or data from one point to another. The signals are produced by a driver circuit on one chip, then travel over an interconnect and terminate at a receiver circuit that may be located on the same chip or a different chip. The communication path between a driver and receiver often passes through an interconnect in the package. The process is illustrated in Figure 4.1. The interconnections can be either on the chip, on the package, or on the PWB. Both the circuits and the interconnections are required to complete a communication path. In this section, some basic concepts associated with circuits and interconnections will be discussed.

Circuits consist of transistors. The most popular transistor used today is the *metal-oxide-semiconductor field effect transistor* (MOSFET), due to its low power and high integration capabilities. This transistor has been used to develop the *complementary metal oxide semiconductor* (CMOS) technology, which combines P-channel and N-channel MOS transistors. CMOS technology is one of the more popular technologies in the world today and is being used to build microprocessors for a host of computer applications.

Samsung, EX1028, p. 91

Both the NMOS and PMOS transistors are three terminal devices consisting of a source, gate and drain, as shown in Figure 4.3. The transistors can be viewed as a switch, which either passes or stops a signal based on the input signal to the gate. For binary digital logic, signals consist of two states: a binary 1 state which corresponds to the high level and a binary 0 state which corresponds to the low level. Using these levels, the PMOS and NMOS transistors can be represented as a normally closed switch and a normally open switch, respectively. The operation of the transistors is shown in Figure

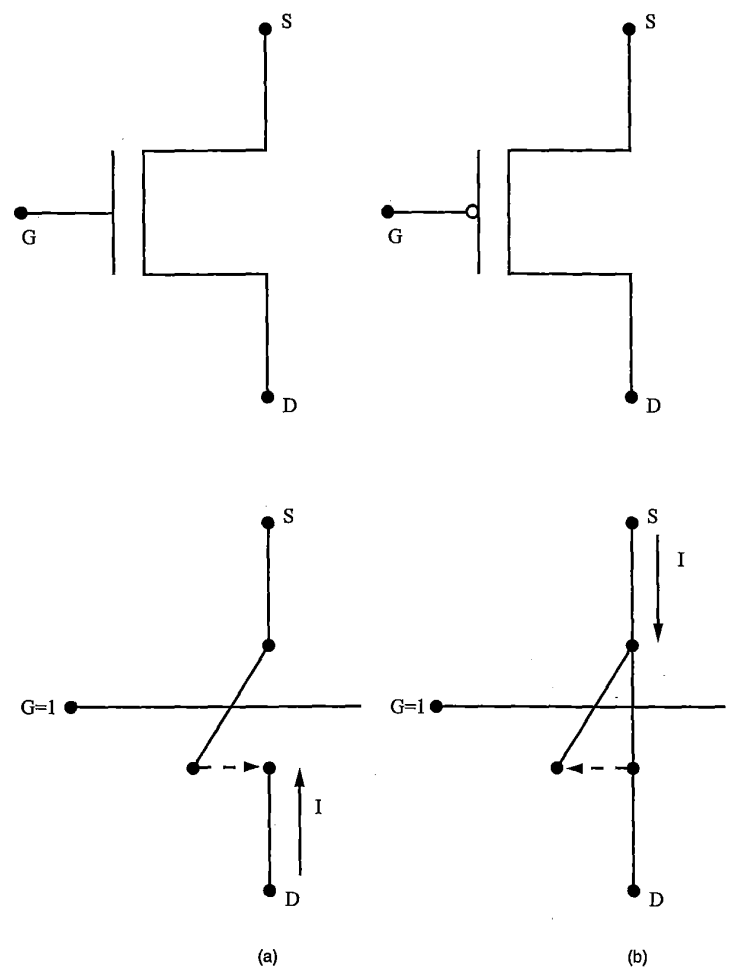


FIGURE 4.3 Transistor switch. (a) NMOS; (b) PMOS.

Samsung, EX1028, p. 92

4.3, where the binary state 1 at the input gate determines whether the switch opens or closes for the two transistors.

The basic building block in CMOS technology is the inverter, which combines a PMOS and NMOS transistor in series, as shown in Figure 4.4, which shows both a transistor and a switch-level implementation. A binary 1 at the input causes the NMOS transistor to close and the PMOS transistor to open while the reverse happens when the input is a binary 0. Hence, the output of the inverter is connected to either the GND or the V_{dd} connection of the circuit, depending on the input signal. Now envision what happens when two such circuits are connected to each other through an interconnection, as shown in Figure 4.5. With a binary 0 input, the PMOS transistor is closed while the NMOS transistor is open, causing current to flow into the interconnection from V_{dd}. With binary 1 input, the PMOS transistor is open and the NMOS transistor is closed, causing current to flow into GND from the interconnection. The direction of current determines whether the interconnection charges to V_{dd} or discharges to GND (this concept will be explained later), causing the input of the second transistor to change to either a binary 1 or 0, resulting in a transition of the second circuit. This completes the communication path between the two circuits. The circuit generating the signal is called a driver while the circuit receiving the signal is called a receiver.

The transistors are nonlinear devices, which produce a current I for a voltage V. Since the transistors are non-ideal devices, they have an on-resistance R_{on}. Though the signals in this section have been represented as a binary 1 and 0, in reality they are analog signals with a finite pulse width, rise and fall times. Since the interconnections transport the signals between the circuits, their characteristics affect the shape of the signal. This

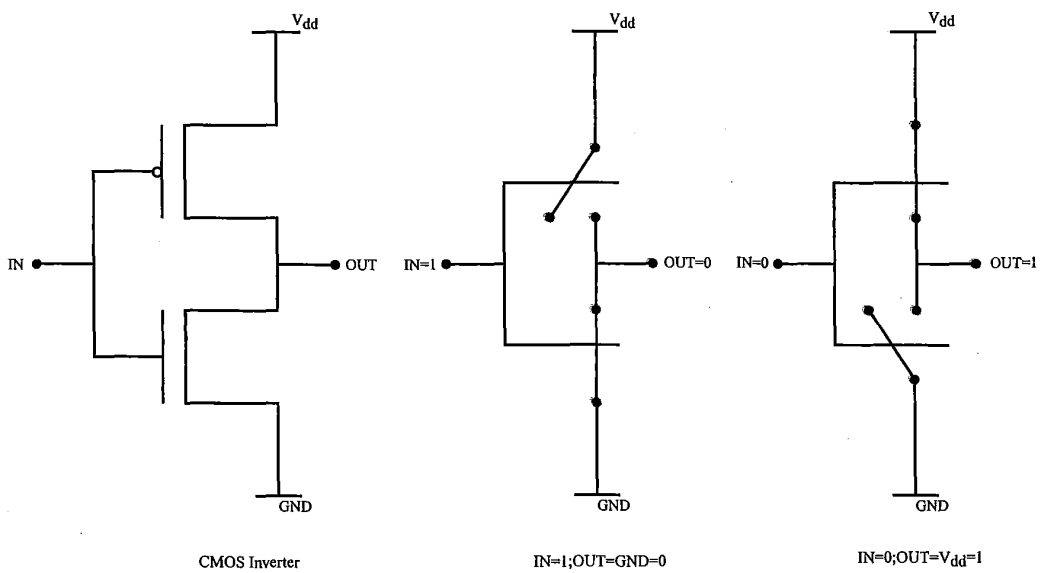


FIGURE 4.4 Inverter circuit.

Samsung, EX1028, p. 93

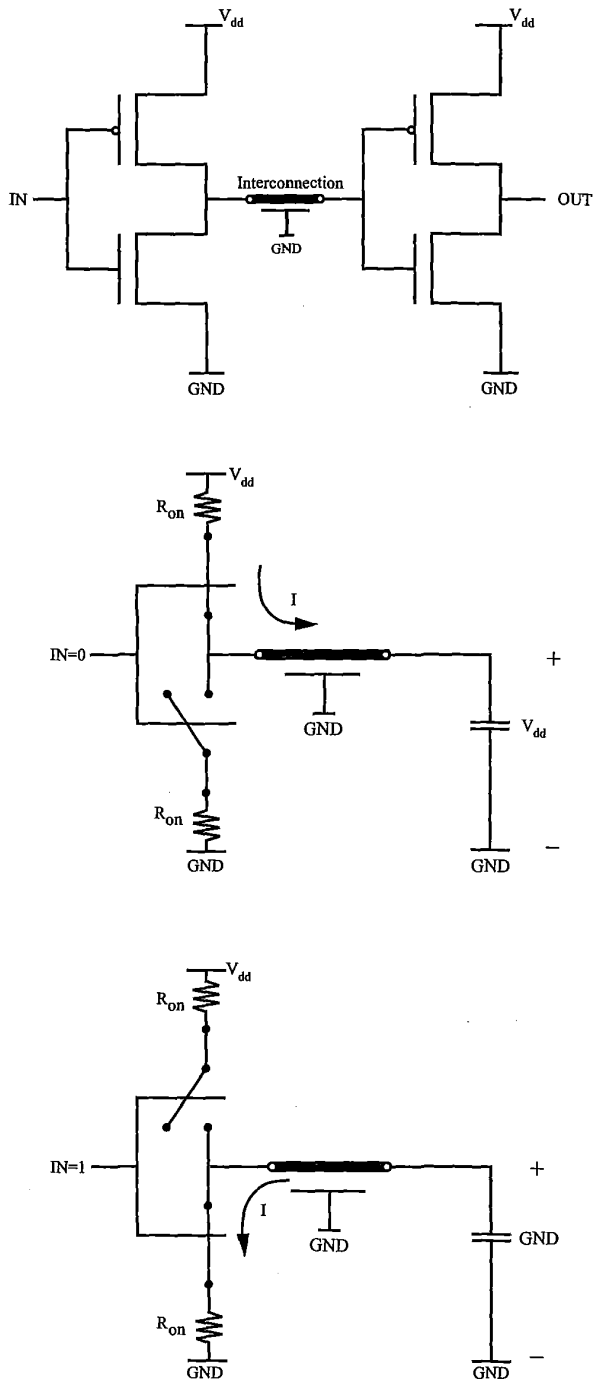


FIGURE 4.5 Charging and discharging interconnections.

Samsung, EX1028, p. 94

chapter focuses on the design of the package interconnects which enable the communication between chips, as shown in Figure 4.1.

4.4.2 Capacitive Delay of Interconnections

An interconnection can be treated as a capacitor which can either be charged or discharged using the inverter circuit described in the previous section. The capacitance arises in part from the metal conductors of the interconnect and ground circuitry. Stray capacitance is also introduced by the physical proximity of other interconnects, the wiring pads that are used to attach integrated circuits to packages, bends in the traces, etc. Figure 4.6 shows an equivalent circuit for the system of Figure 4.5, where the interconnect is represented as a capacitor. The switch represents the PMOS transistor that connects the voltage V_{dd} to the capacitor through the on-resistance R_{on}. The input of the receiver circuit is represented as a capacitor C_g. By combining Kirchhoff's Voltage Law with the equation

$$I = (C + C_g)\, dV/dt \tag{4.1}$$

describing the current through a capacitor, one can derive a differential equation in terms of the load voltage that has the form:

$$V_{dd} - V_{\text{load}} = R_{\text{on}}\, I = R_{\text{on}}(C + C_g)\, dV_{\text{load}}/dt \tag{4.2}$$

This equation can be expressed as:

$$dV_{\text{load}}/dt + V_{\text{load}}/\{R_{\text{on}}(C + C_g)\} = V_{dd}/\{R_{\text{on}}(C + C_g)\} \tag{4.3}$$

Since the switch is closed at time $t = 0$, Equation (4.3) has the general solution:

$$V_{\text{load}}(t) = A[1 - e^{-t/\tau}]\, u(t) \tag{4.4}$$

where A is a constant to be determined, and $u(t)$ is a unit step function that transitions from zero to one at $t = 0$. The parameter τ is known as the time constant, and is given by:

$$\tau = R_{\text{on}}(C + C_g) \tag{4.5}$$

The coefficient A will be determined by the condition that the capacitor charges to a steady-state value, or:

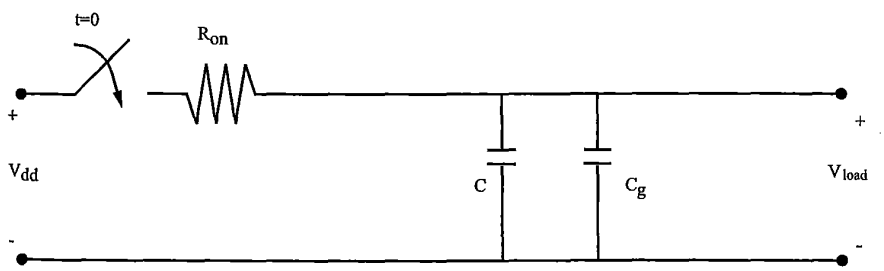


FIGURE 4.6 Capacitance of interconnections.

Samsung, EX1028, p. 95

$$dV_{load}/dt \rightarrow 0 \quad \text{as } t \rightarrow \infty \tag{4.6}$$

Therefore, $A = V_{dd}$, and

$$V_{load}(t) = V_{dd}[1 - e^{-t/\tau}]\, u(t) \tag{4.7}$$

EXAMPLE 4.1

Consider the situation with $V_{dd} = 5$ (V), $R_{on} = 50\ \Omega$, and $(C + C_g) = 10$ pF. A plot of Equation (4.7) is depicted in Figure 4.7. The time constant associated with the capacitor is $\tau = R_{on}(C + C_g) = 0.5$ (ns). The delay associated with the capacitor charging prevents the load voltage from reaching 50% of its steady-state value until $t = 0.35$ (ns), and keeps the load voltage from reaching 4.5 (V), 90% of its steady-state value until $t = 1.15$ (ns). Thus, the time needed for the capacitor to charge to 90% of its steady-state value in this case introduces a significant delay for the communication between circuits.

Digital circuit function is based on threshold logic, which means that the circuit changes state when the input reaches a certain level. In most circuits, 50% of the logic swing is chosen as the threshold level. Hence, based on Equation (4.7), the delay for the output to reach 50% of the input is given by:

$$T_{50\%} = 0.69R_{on}(C + C_g) \tag{4.8}$$

This equation is called the delay equation and is based on a lumped model where the interconnect is treated as a lumped capacitor. Delay equations are often used for generating timing information for the circuits. These equations are much more complex than Equation (4.8) since all the delay associated with parasitic capacitances have to be accounted for in this equation. For example, the delay equation for an inverter with on-resistance R_{on} driving a long interconnection with resistance R and capacitance C terminated in a receiver with capacitance C_g can be represented as:

$$T_{50\%} = 0.69R_{on}(C + C_g) + R(0.4C + 0.69C_g) \tag{4.9}$$

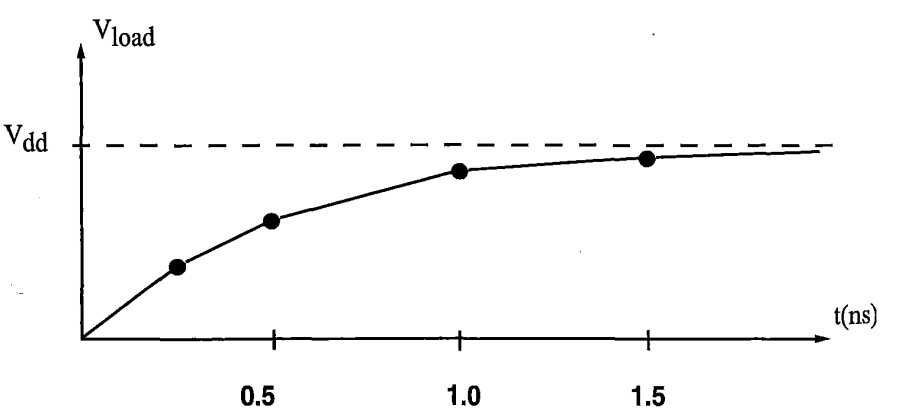


FIGURE 4.7 Capacitive delay.

Samsung, EX1028, p. 96

4.4.3 Kirchhoff's Laws and Transit Time Delay

For most of the past 200 years, electrical circuits have been modeled using Kirchhoff's Laws: (a) the sum of the individual voltage drops around a loop of the circuit must vanish and (b) the sum of the currents into a node of the circuit must vanish. Kirchhoff's Laws are exact for DC circuits and are good approximations for low frequencies, where the circuit dimensions are much smaller than the wavelength of the signal of interest. The RC model in the previous section is a good example where Kirchhoff's Laws were applied to derive a delay equation. However, at higher frequencies, these "laws" don't always work; they are actually just approximations of the true behavior of electrical signals.

The laws are limited by one important fact: electrical signals propagate at the speed of light, which is approximately 3×10^8 (m/s) in air, and somewhat slower than that on a typical printed wiring board, module, or semiconductor device. Kirchhoff's Laws neglect the finite velocity of an electrical signal, and therefore fail when the time delay or phase shift due to that finite velocity, become significant. This is seldom a concern at lower frequencies found in audio amplifiers, household appliances, and many other devices. However, at higher frequencies the effect can be pronounced.

From a time domain perspective, the critical parameter is the transit time (or *time of flight*) of a signal relative to the size of the circuit. The velocity of 3×10^8 (m/s) in air is roughly equivalent to a time delay of 1 (ns) per foot of travel. A time interval of 1 (ns) may seem small, but it isn't if the clock rate of the circuit is 1 GHz! Under these conditions, it may take an entire clock period for a signal to travel across a six inch printed wiring board, and part of the circuit may be a full clock period behind. Kirchhoff's Laws neglect this time delay. Early integrated circuits involved *transistor-transistor logic* (TTL) devices with internal delays of 15 (ns) or more, limiting the clock rates to lower levels and making the transit time of the signals negligible in comparison. Modern devices have improved to the point where transit delays can be the limiting factor in digital circuit design. To systematically incorporate this time delay into package design, one must employ transmission line theory, the topic of the following section.

When thinking in terms of the frequency domain, the key parameter is the wavelength of the signal relative to the size of the circuit. The wavelength is given by:

$$\text{Wavelength} = (\text{speed of light})/(\text{frequency}), \quad \text{or } \lambda = c/f \tag{4.10}$$

As the frequency increases, the wavelength decreases. At 1 GHz, the wavelength in air is 30 (cm). One wavelength corresponds to 360° of phase shift. Kirchhoff's Laws assume that all parts of a circuit have the same phase. If the signal accumulates more than a small fraction of 360° of phase shift in traveling across a circuit, Kirchhoff's Laws no longer accurately describe the situation; transmission line theory does.

Most situations of interest to packaging engineers will involve dielectric materials, which are described in terms of a parameter known as the relative permittivity ε_r, also called the dielectric constant. Magnetic materials, most often encountered in embedded inductor applications, are described by a relative permeability μ_r. The parameters ε_r and μ_r are unitless scale factors. The speed of light of any electrical signal, in a general material, is given by:

$$\text{speed of light} = (2.998 \times 10^8)/\sqrt{(\varepsilon_r \mu_r)} \quad \text{(m/s)} \tag{4.11}$$

Therefore, a signal propagating on an interconnect in a dielectric medium with relative

Samsung, EX1028, p. 97

permittivity $\varepsilon_r = 4.0$ will propagate with a velocity of 1.5×10^8 m/s. Hence, the signal travels a length of 1 m in a time of 6.67 ns which is often referred to as the time of flight or propagation delay. This velocity is always slower than that of light in air.

The wavelength of a single-frequency signal in a medium with parameters ε_r and μ_r is given by:

$$\lambda = (2.998 \times 10^8)/\{f\sqrt{(\varepsilon_r \mu_r)}\} \tag{4.12}$$

where f is the frequency in Hertz and λ is the wavelength in meters. The propagation delay and wavelength for common dielectric materials is shown in Table 4.2.

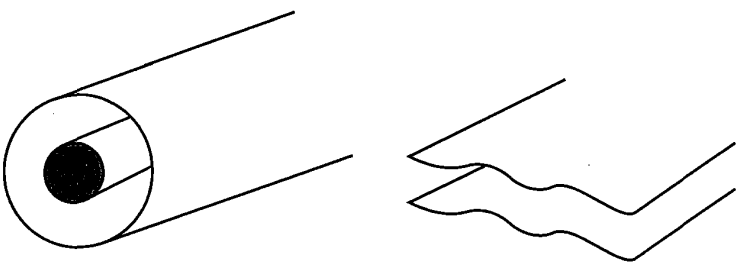
4.4.4 Transmission Line Behavior of Interconnections

To illustrate transmission line theory, consider the two cable cross sections shown in Figure 4.8. This figure shows a coaxial cable—a transmission line widely used in microwave measurements and CATV systems—and a parallel strip transmission line, which might be a simple structure for interconnects in a package or PWB. When electricity flows on either of these cables, there is a physical movement of charge carriers, or electrons, down one conductor and back on the other. Because of this movement, the charge has some momentum and, in essence, wants to keep moving once it has started. This effect is essentially equivalent to some series inductance. There is also some series resistance since the metal is not a "perfect" conductor of electricity, and therefore some of the electrical energy is converted to heat as the current flows. At the same time, equal

TABLE 4.2 Relative permittivities of assorted materials (ε_r at approximately 1 GHz) and the propagation delay in (ps/cm).

air	1.0	33.3
alumina	9.4	102.0
A-35 ceramic	5.6	78.9
glazed ceramic	7.2	89.4
Gallium arsenide	13.0	120.0
germanium	16.0	133.0
glass epoxy	4.0	66.7
FR4	4.9	73.8
lucite	2.6	53.7
mica	6.0	81.6
nylon	3.5	62.4
plexiglass	2.6	53.7
polyethylene	2.3	50.6
polyimide	3.5	62.4
polystyrene	2.6	53.7
quartz	3.5	62.4
Rexolite 1422	2.5	52.7
silicon	11.8	114.5
silicon dioxide	3.9	65.8
teflon	2.1	48.3

Samsung, EX1028, p. 98



Coaxial cable Parallel conducting strips

FIGURE 4.8 Common transmission lines.

and opposite charge is stored instantaneously on the two conductors, giving rise to some shunt capacitance. If the material separating the conductors is not a perfect insulator, there will also be some leakage current from one conductor to another, which can be modeled as shunt conductance.

Thus, an equivalent circuit for a short section of transmission line, say, of length Δz, where Δz is much smaller than a wavelength or the equivalent time of flight, is depicted in Figure 4.9. The inductance and capacitance present in this equivalent circuit provide the time delay and phase shift lacking in the treatment of these transmission lines using circuit theory, which would assume that they provide a direct connection from one end to the other. Consequently, this equivalent circuit can be treated using Kirchhoff's Voltage and Current Laws.

In Figure 4.9, the quantity $L\Delta z$ (H) is the total series inductance of the equivalent circuit, which depends on the inductance per unit length L (H/m). $C\Delta z$ (F) is the total capacitance, which depends on the shunt capacitance per unit length C (F/m). The total series resistance $R\Delta z$ and shunt conductance $G\Delta z$ depend on the resistance per unit length R (Ω/m) and the conductance per unit length G (S/m), respectively.

The application of Kirchhoff's Voltage Law to the equivalent circuit in Figure 4.9 yields the equation:

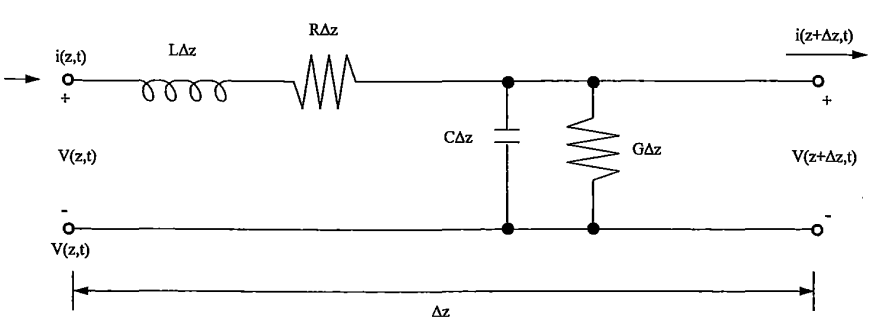


FIGURE 4.9 Transmission line equivalent circuit.

Samsung, EX1028, p. 99

$$V(z + \Delta z,t) + (L\Delta z)\partial i/\partial t + (R\Delta z)i(z,t) = V(z,t) \tag{4.13}$$

which can be rewritten in the form:

$$\{V(z + \Delta z,t) - V(z,t)\}/\Delta z = -Ri(z,t) - L\partial i/\partial t \tag{4.14}$$

In the limiting case as Δz tends to zero, this equation becomes:

$$\partial V/\partial z = -Ri - L\partial i/\partial t \tag{4.15}$$

An application of Kirchhoff's Current Law to the circuit in Figure 4.9 produces:

$$i(z + \Delta z,t) - i(z,t) = -(G\Delta z)V(z + \Delta z,t) - (C\Delta z)\partial V/\partial t \tag{4.16}$$

which, in the limiting case as Δz vanishes, yields:

$$\partial i/\partial z = -GV - C\partial V/\partial t \tag{4.17}$$

Equations (4.15) and (4.17) are known as the *Transmission Line Equations*. Formerly, these were known as the *Telegrapher's Equations,* since they were originally used to describe signals on telegraph lines.

Equations (4.15) and (4.17) are a coupled system of two partial differential equations in terms of the voltage $V(z,t)$ and current $i(z,t)$. These equations can be simplified by eliminating one variable. In addition, it suffices to consider the lossless case, where $R = 0$ and $G = 0$. For instance, by differentiating (4.15) with respect to z and differentiating (4.17) with respect to t, the equations can be combined to produce:

$$\frac{\partial^2 V}{\partial z^2} = LC\frac{\partial^2 V}{\partial t^2} \tag{4.18}$$

Equation (4.18) is well known to the math and science community, and has been given the name "one-dimensional wave equation." An alternate approach is to differentiate (4.15) with respect to t and differentiate (4.17) with respect to z, to eliminate the voltage and obtain:

$$\frac{\partial^2 i}{\partial z^2} = LC\frac{\partial^2 i}{\partial t^2} \tag{4.19}$$

Thus, the voltage and current satisfy the same second-order differential equation.

4.4.5 One-Dimensional Waves

The general solution to equation (4.18) has the form:

$$V(z,t) = V^+ f\left(t - \frac{z}{v_p}\right) + V^- g\left(t + \frac{z}{v_p}\right) \tag{4.20}$$

where V^+ and V^- are unknown coefficients (to be determined by initial or boundary conditions at the ends of the interconnection) and f and g are arbitrary functions of one variable, also to be determined by initial or boundary conditions. By substituting (4.20) into (4.15) and (4.17), the current is obtained as

Samsung, EX1028, p. 100

$$i(z,t) = \frac{V^+}{Z_0} f\left(t - \frac{z}{v_p}\right) - \frac{V^-}{Z_0} g\left(t + \frac{z}{v_p}\right) \tag{4.21}$$

The parameter v_p in these equations has units of velocity (m/s) and is given by:

$$v_p = \frac{1}{\sqrt{LC}} \tag{4.22}$$

while Z_0 has units of resistance (Ω) and is given by:

$$Z_0 = \sqrt{\frac{L}{C}} \tag{4.23}$$

The parameter Z_0 is known as the characteristic resistance, or more commonly, the characteristic impedance of the line.

To demonstrate that (4.20) and (4.21) are the solutions to (4.18) and (4.19), simply substitute them into the equations. This is left as an exercise for the reader.

The general solution in (4.20) and (4.21) is a one-dimensional wave. To understand the behavior of a wave, consider the function $f(u)$ depicted in Figure 4.10. The wave $f(t + z/v_p)$ can be constructed for specific values of space and time; Figure 4.11 shows three specific plots versus space (z) at fixed times for $v_p = 300$ (m/s). Observe that $f(t + z/v_p)$ represents a function that retains its shape as time marches forward but changes its position in z; specifically it travels in the $-z$ direction with a constant velocity v_p. A similar plot of the function $f(t - z/v_p)$, also shown in Figure 4.11, indicates that it travels in the $+z$ direction with constant velocity v_p while retaining its shape.

Therefore, a one-dimensional wave is a function that propagates with a constant velocity without changing its shape. Furthermore, the voltage and current waveforms that satisfy the lossless transmission line equations are one-dimensional waves. The minus sign in the argument of the wave function indicates the direction: a plus sign means the wave is traveling in the $-z$ direction, while a minus sign indicates the wave is traveling in the $+z$ direction.

EXAMPLE 4.2

The simplest type of transmission line might well be the ideal parallel strip transmission line illustrated in Figure 4.12. This line is "ideal" in the sense that the fringing nature of the fields at

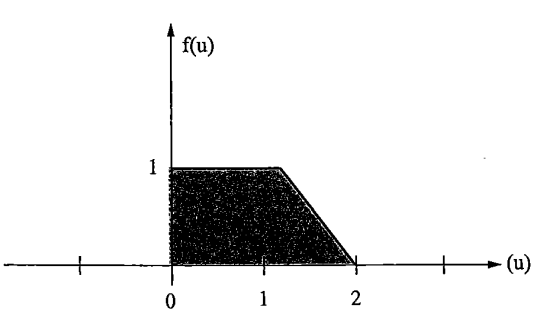


FIGURE 4.10 Wave function.

Samsung, EX1028, p. 101

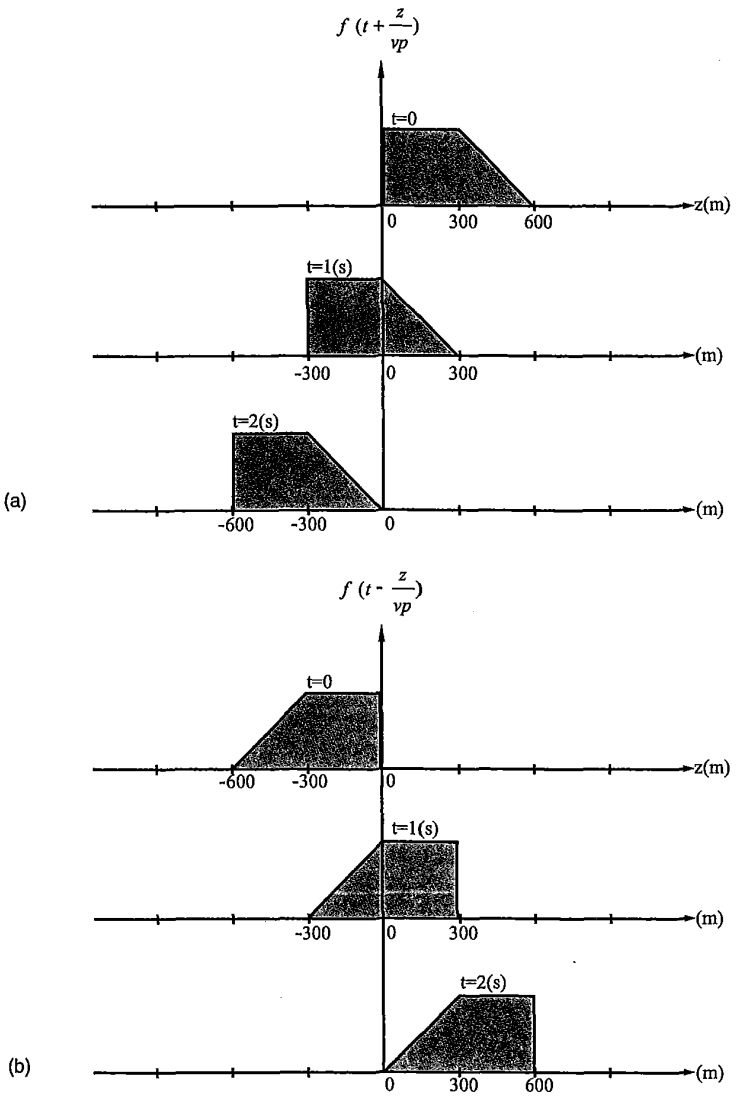


FIGURE 4.11 (a) Backward traveling wave and (b) forward traveling wave.

FIGURE 4.12 Parallel strip transmission line.



Samsung, EX1028, p. 102

the open ends is ignored. The fringing fields are usually too severe to ignore in the real world, however, and better models are available in the form of microstrip and stripline transmission lines, which are provided in a later section. Under lossless assumptions, the ideal stripline has an inductance per unit length of:

$$L = \mu d/w \tag{4.24}$$

and a capacitance per unit length of:

$$C = \varepsilon w/d \tag{4.25}$$

where μ and ε are the total permeability and permittivity parameters, given by:

$$\mu = (4\pi \times 10^{-7})\mu_r \quad \text{(H/m)} \tag{4.26}$$

$$\varepsilon = (8.854 \times 10^{-12})\varepsilon_r \quad \text{(F/m)} \tag{4.27}$$

Consequently, the propagation velocity of a signal on this line is given by:

$$v_p = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{\mu\varepsilon}} \tag{4.28}$$

while the characteristic impedance is given by:

$$Z_0 = \sqrt{\frac{L}{C}} = \frac{d}{w}\sqrt{\frac{\mu}{\varepsilon}} \tag{4.29}$$

Observe that the propagation velocity of an electrical signal on this line is independent of the height and width of the line geometry, and is the same as the velocity of light, or any other electromagnetic wave, in the same material. This is always true if the line is constructed from a uniform (homogeneous) material. Thus, the dimensions of the metal strips play no role in the velocity of an electrical signal on the line. The characteristic impedance does depend on the geometry of the line, or the cross-sectional dimensions, as well as the material. By adjusting the height and width of the strips, Z_0 can be set at any desired value.

4.4.6 Typical Transmission Line Structures Used as Package Interconnections

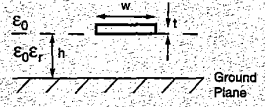
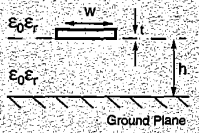
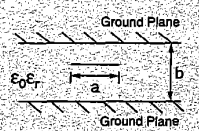
Though the parallel strip is an easy transmission line to design, other structures are typically used as package interconnections due to their high-frequency behavior. These can be classified into three main categories: the microstrip, the embedded microstrip and the stripline. The microstrip is an interconnection structure that is at the interface between air and a dielectric as shown in Table 4.3, and is commonly used on the top layer of a package. The embedded microstrip is a microstrip line that is embedded within a dielectric as shown in Table 4.3. These interconnections are commonly used in the top few layers of the package. Finally, the stripline consists of an interconnection sandwiched between two metal layers as shown in Table 4.3 and is commonly used as a buried interconnection in the package. The impedance Z_0 and propagation velocity v_p for these interconnection structures are shown in Table 4.3.

4.4.7 Waves on Transmission Lines between Digital Gates

Figure 4.13a illustrates a typical problem involving two generic digital logic gates connected by a trace that will be modeled as a transmission line. At time $t = 0$, the output

Samsung, EX1028, p. 103

TABLE 4.3 Formulae for typical package interconnects.

Microstrip Transmission Line		$\varepsilon_{eff} = \varepsilon_0\left[\frac{\varepsilon_r+1}{2} + \frac{\varepsilon_r-1}{2}\frac{1}{\sqrt{1+12b/a}}\right]$ $v_p = \frac{1}{\sqrt{\mu\varepsilon_{eff}}}$ $Z_0 = \frac{1}{2\pi}\sqrt{\frac{\mu}{\varepsilon_{eff}}}\ln\left(\frac{8b}{a}+\frac{a}{4b}\right) \quad a<b$ $Z_0 = \sqrt{\frac{\mu}{\varepsilon_{eff}}}\frac{1}{\frac{a}{b}+1.393+0.667\ln\left(\frac{a}{b}+1.444\right)} \quad a>b$
Embedded Microstrip Transmission Line		$v_p = \frac{1}{\sqrt{\mu\varepsilon}}$ $Z_0 = \frac{60}{\sqrt{\varepsilon_r+1.41}}\ln\left(\frac{5.98h}{0.8w+t}\right)$
Stripline Transmission Line		$v_p = \frac{1}{\sqrt{\mu\varepsilon}}$ $Z_0 = \frac{30\pi}{\sqrt{\varepsilon_r}}\frac{b}{a_{eff}+0.441b}$ $a_{eff} = \begin{cases} a & a>0.35b \\ a-\left(0.35-\frac{a}{b}\right)^2 b & a<0.35b \end{cases}$

$\varepsilon = \varepsilon_0\varepsilon_r$; ε_r: Relative dielectric constant; $\varepsilon_0 = 8.854 \times 10^{-14}$ F/cm
v_p: Propagation velocity in cm/s
Z_0: Characteristic impedance in Ω

MEMS23149

CONFIDENTIAL BUSINESS INFORMATION

Samsung, EX1028, p. 104

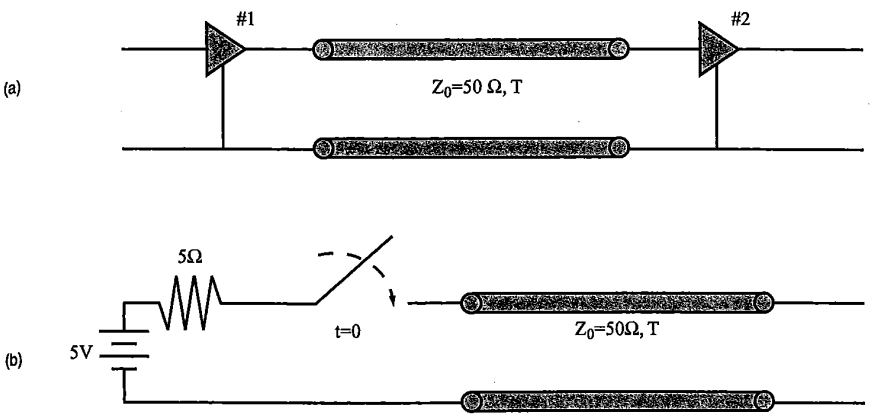


FIGURE 4.13 (a) Communication path and (b) equivalent circuit.

of gate 1 will transition from low state ($V = 0$) to high state ($V = 5$). Assume that the gates represent a CMOS driver and receiver pair, and the transmission line has characteristic impedance of 50 Ω. Therefore, the system is equivalent to the circuit shown in Figure 4.13b.

The switch in Figure 4.13b closes at $t = 0$, placing a voltage across the input end of the transmission line. The signal on the transmission line exhibits a finite velocity of propagation, given by

$$v_p = \frac{1}{\sqrt{LC}} \tag{4.30}$$

Thus, the signal will not immediately appear across the load end of the line. Instead, a voltage and current wave will propagate gradually down the line. Since this wave is traveling from left to right in Figure 4.13a, the ratio of V^+ to I^+ is Z_0 (from Equation 4.21), and the input end of the line appears to the external world to be identical to a resistor with resistance Z_0. Thus, at $t = 0^+$, the input end of the transmission line has the equivalent circuit shown in Figure 4.14a. Consequently, in this situation the voltage across the line is $V^+ = 50/(5 + 50) \times 5 = 4.545$ (V) and the current into the line is $I^+ = V^+/Z_0 = 90.9$ (mA).

Suppose that the line has a total one-way propagation delay of T seconds. At time $t = T$ the wave launched at $t = 0$ reaches the load end of the line. The load end is terminated by a CMOS receiver gate, which has a very high resistance, or capacitive load, and can be modeled as an open circuit ($R_{load} \rightarrow \infty$). The equivalent circuit for the load appears in Figure 4.14b. Since the ratio of V^+ to I^+ is Z_0, this voltage and current do not satisfy Ohm's Law at the load circuit, and a reflected wave is generated in order to simultaneously satisfy Equations (4.20) and (4.21), constrained by $V_{load}/R_{load} = 0$, or:

$$V^+ + V^- = V_{load} \tag{4.31}$$

$$I^+ + I^- = V_{load}/R_{load} = 0 \tag{4.32}$$

Samsung, EX1028, p. 105

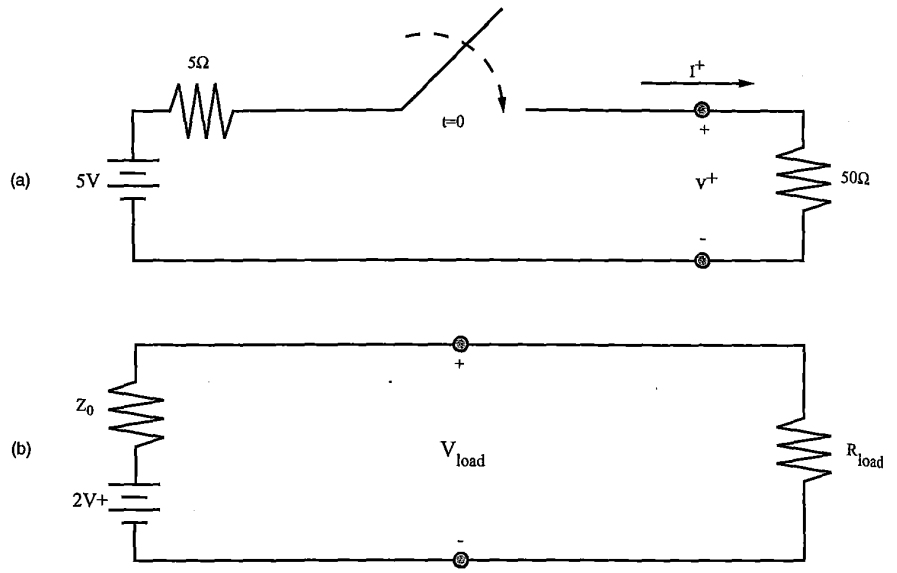


FIGURE 4.14 (a) Equivalent circuit at generator and (b) equivalent circuit at load.

From Equations (4.20 and 4.21), $V^-/I^- = -Z_0$, and therefore a solution of (4.31) and (4.32) yields:

$$V^- = V^+ \tag{4.33}$$

$$I^- = -I^+ \tag{4.34}$$

$$V_{\text{load}} = 2V^+ \tag{4.35}$$

Thus, there is a reflected voltage and current wave launched back toward the source end of the transmission line. This wave has $V^- = 4.545$ (V) and $I^- = -V^-/Z_0 = -90.9$ (mA).

It is convenient to introduce a unitless parameter known as the reflection coefficient at the load (Γ_L) to simplify the analysis to follow. Γ_L represents the ratio V^-/V^+, and is given by the general solution of (4.31) and (4.32) (without using $R_{\text{load}} \rightarrow \infty$) as:

$$\Gamma_L = V^-/V^+ = (R_{\text{load}} - Z_0)/(R_{\text{load}} + Z_0) \tag{4.36}$$

In the present example, $\Gamma_L = 1$, and at time $t = T$, the load voltage jumps from 0 to $2V^+ = 9.09$ (V). Simultaneously, a reflected wave is launched back down toward the source end of the line. At $t = 2T$, this reflected wave arrives back at the driver. The voltages and currents at the source end must satisfy the equations:

$$V^+ + V^- = V_{\text{source}} \tag{4.37}$$

$$I^+ + I^- = V_{\text{source}}/Z_0 \tag{4.38}$$

Since the existing values of V^+, V^-, I^+ and I^- do not satisfy (4.37) and (4.38), an

Samsung, EX1028, p. 106

additional reflection is generated toward the load end of the line with a voltage wave of the form:

$$V^{++} = \Gamma_G V^- \tag{4.39}$$

that is superimposed with the previous waves. In (4.39), by analogy with the situation at the load,

$$\Gamma_G = (R_G - Z_0)/(R_G + Z_0) = -0.818 \tag{4.40}$$

The process of waves reflecting at the load and source ends of the transmission line continues until the line voltage stabilizes as $t \rightarrow \infty$. Eventually, the entire transmission line charges up to a constant voltage, which in this case is $V = 5$ (V). Observe that this is the voltage that would be obtained from a conventional circuit analysis of the system in Figure 4.13a, without using transmission line theory. However, the use of circuit theory predicts that this steady-state voltage is reached instantly, while transmission line theory dictates that in actuality it takes a significant time for the system to converge to within 10% of this voltage. In a high-performance system, the proper prediction of this time delay is critical. Figure 4.15 shows a plot of the load voltage as a function of time.

For convenience, the calculations obtained in the previous example can be collected together on a single plot of voltage or current, as a function of position (horizontal axis) and time (vertical axis). Such a plot is often called a bounce diagram. Figure 4.16 shows the bounce diagram for the previous example. Plots of the voltage or current as a function of space or of time at any value of the other variable can be obtained directly from the diagram.

The bounce diagram tracks the position of the leading edge of the waveform as a function of position z and time t. The upper left corner of the diagram in Figure 4.16 represents $z = 0$, $t = 0$. The leading edge of the waveform follows a straight-line trajectory to $z = L$, $t = T$. At that point, the wave is reflected from the load end of the line, creating a new trajectory to the input end ($z = 0$) at time $t = 2T$. This process continues indefinitely, resulting in the zigzag pattern in Figure 4.16. Each triangular region of the plot exhibits a constant value for the voltage as shown. It is also convenient to track the leading-edge voltages on the diagram (also shown).

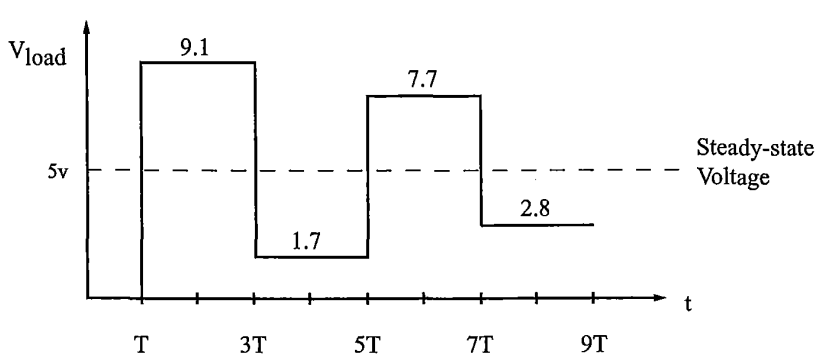


FIGURE 4.15 Waveform at the load.

Samsung, EX1028, p. 107

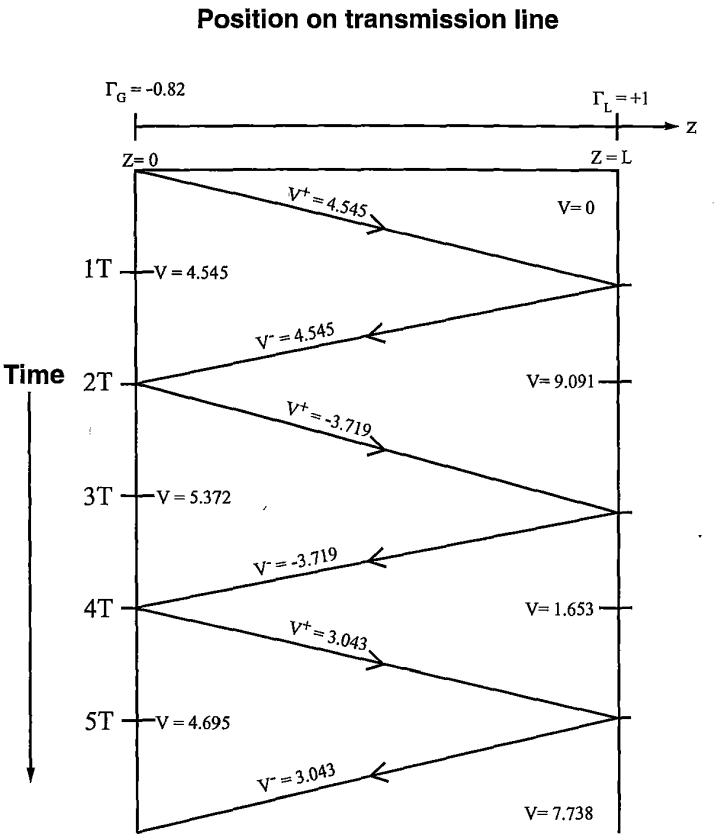


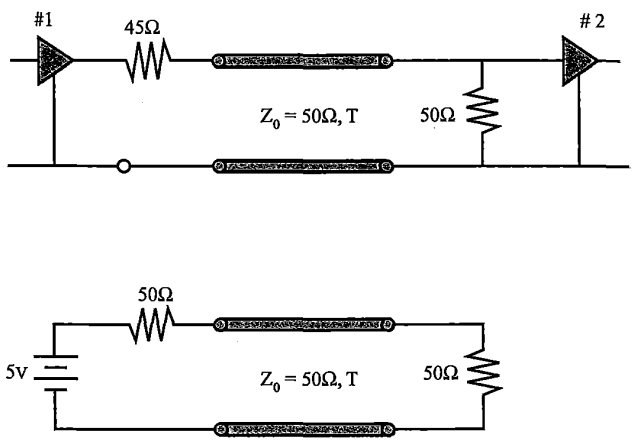
FIGURE 4.16 Bounce diagram.

4.4.8 Termination Schemes that "Match" the Transmission Line

Clearly, the severe reflections in the previous example can prevent a high performance digital circuit from operating properly. How can these reflections be eliminated or reduced so that they do not detrimentally effect the performance of a system? The process of eliminating these reflections is known as *matching,* or *impedance matching,* the line. There are a variety of ways in which this can be accomplished.

In traditional high-frequency transmission line applications, the common approach is to match the line at both ends, by adjusting the source and load resistances in order that they each equal Z_0. For the example of Figure 4.13, a series resistor of 45 Ω could be added at the generator end, while a shunt resistor of 50 Ω could be added at the load end. The effect of these additional resistors is to make both Γ_G and Γ_L vanish, eliminating reflections. Figure 4.17 shows the resulting system and the reflection diagram. Unfortunately, there is a problem evident in Figure 4.17: the voltage across the load is only 2.5 (V), half of the generator voltage. Thus, this approach won't always work for digital applications where the supply voltage might be limited to 5 (V) or less. In traditional

Samsung, EX1028, p. 108



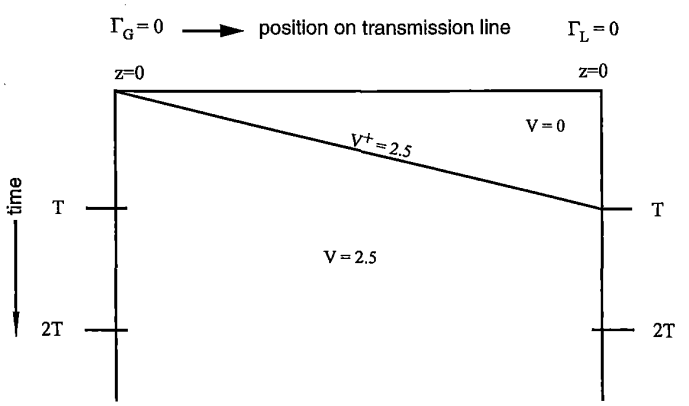
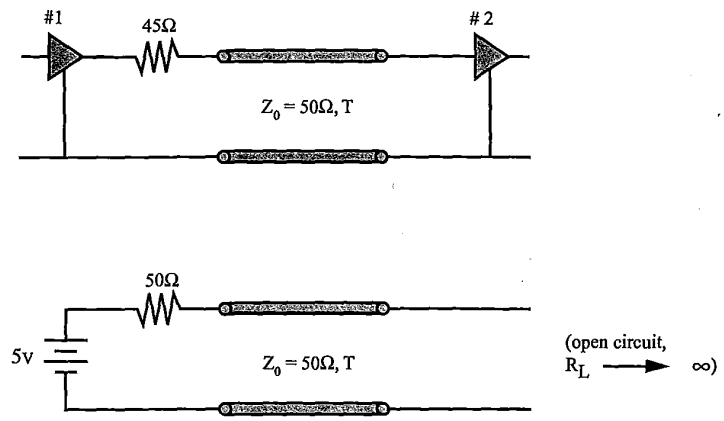


FIGURE 4.17 Series and parallel termination.

high-frequency applications, the generator voltages are easily adjusted to compensate for this factor of two.

Alternate matching schemes are possible, and might be more practical for the digital application. One approach is the *series termination,* illustrated in Figure 4.18. In the series termination scheme, a series matching resistor is inserted at the generator end of the line, while the load end is left unterminated. Consequently, the reflection coefficient at the generator is changed to:

Samsung, EX1028, p. 109



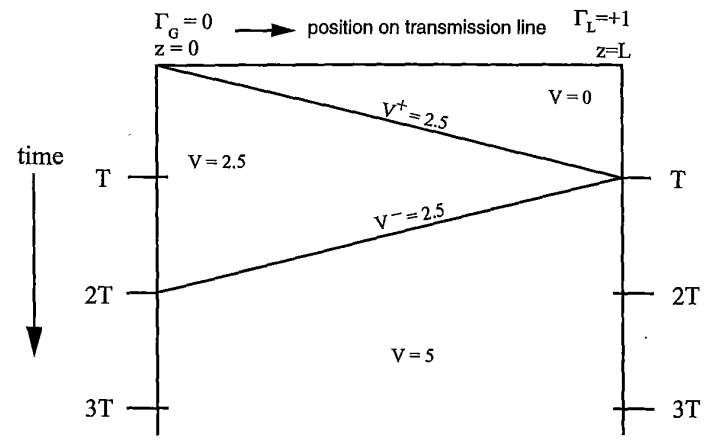


FIGURE 4.18 Series termination.

$$\Gamma_G = 0 \tag{4.41}$$

while the reflection at the load remains unchanged,

$$\Gamma_L = 1 \tag{4.42}$$

The series matching resistor changes the initial voltage to $V^+ = 2.5$ (V). The reflection diagram is shown in Figure 4.18, and illustrates the result. Despite the fact that there is

Samsung, EX1028, p. 110

a severe reflection at the load end, the voltage at the load jumps from zero to 5 (V) at $t = T$, and stays at that value. In fact, the reflection at the load is necessary in this case to provide voltage doubling, since the incident voltage is now only 2.5 (V).

The second approach to matching is the *parallel termination,* illustrated in Figure 4.19. In the parallel termination scheme, a shunt-matching resistor is inserted at the load end of the line, while the generator end is left mismatched. Consequently, for the previous example, the reflection coefficient at the generator remains:

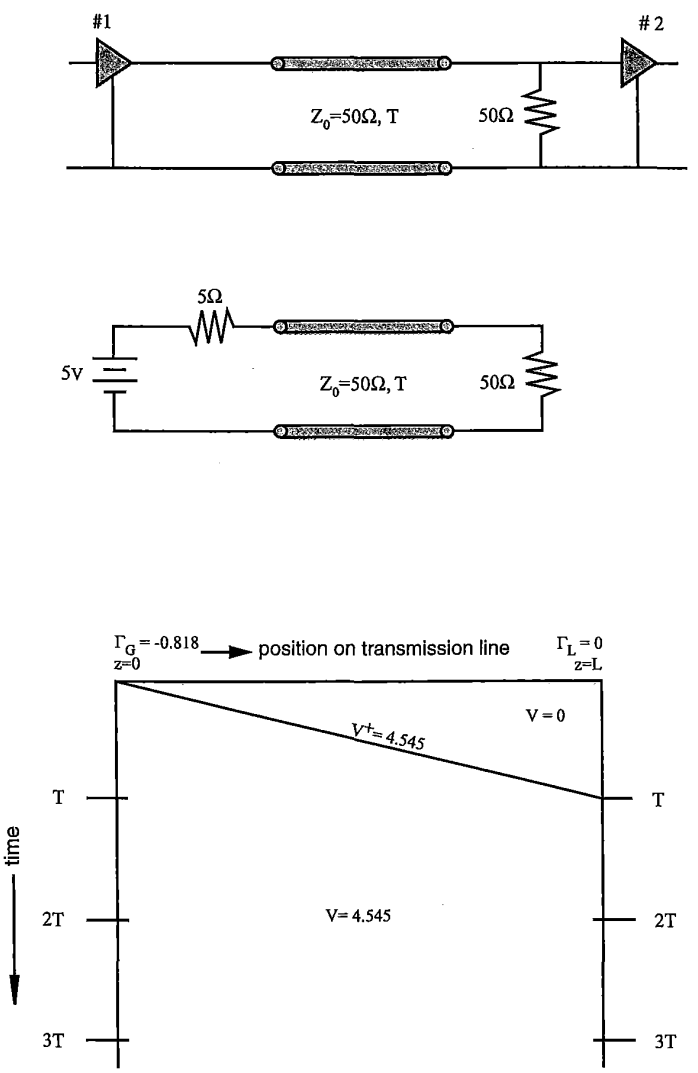


FIGURE 4.19 Parallel termination.

Samsung, EX1028, p. 111

$$\Gamma_G = -0.818 \tag{4.43}$$

while the reflection at the load is eliminated,

$$\Gamma_L = 0 \tag{4.44}$$

The initial line voltage remains $V^+ = 4.545$ (V). The reflection diagram is shown in Figure 4.19 and illustrates the result. The mismatch at the generator end never enters into the analysis, since there is no negative-going wave on the line.

An important aspect of the series versus a parallel matching approach is the power dissipated by the matching resistor. In the series matching case, once the line is charged, the current flow through the 45 Ω matching resistor is zero,

$$I_{in} = (5.0 - 5.0)/50 = 0.0 \text{ (A)} \tag{4.45}$$

meaning that the matching resistor dissipates no power. In the parallel matching case, the total load current is:

$$I_L = 4.545/50 = 90.9 \text{ (mA)} \tag{4.46}$$

meaning that the 50 Ω matching resistor dissipates 413 (mW) of power. The lower power demand of the series match is more suitable for CMOS logic, the most common digital circuitry in use today. It is also worth noting that the parallel matching approach might be the preferred approach when used with high-performance *emitter-coupled logic* (ECL), because it can play a dual role as a matching resistor for the transmission line and a pull-down resistor for the device circuitry.

Since many of the transmission lines arising in packaging applications are terminated with an active device, it might be convenient to modify the device circuitry to provide matching, instead of adding resistors; in other words, fabricate a matching network on the chip itself. An example of this is the *diode termination* circuit. A pair of diodes at the load end of the line hold the voltage within a set range, and limit the over-shoot and under-shoot possible with a mismatched line.

4.4.9 When Are Transmission Line Effects Important?

Not all interconnects in the package require analysis as transmission lines. Depending on the signal characteristics, the finite transit time of the signal may be so small that the transmission line effects can be ignored. In such cases, the transmission lines can be treated as lumped capacitors, inductors or resistors.

A digital signal waveform is shown in Figure 4.20. The signal takes time duration t_r to transition from 10% to 90% of the voltage swing. This time is called the *rise time* of the signal. This is similar to the *fall time* t_f which is the time duration for the signal to transition from the 90% to 10% of the voltage swing. For an interconnect to behave as a transmission line, the rise time of the signal has to be less than the round trip time of flight delay on the interconnect. For an interconnect in a dielectric medium ε_r, this can be represented as:

$$t_r \leq 33.33\sqrt{\varepsilon_r} \times 2 \times l \quad \text{ps} \tag{4.47}$$

where $2 \times l$ represents the round trip length of the interconnection in cm and 33.33 ps /cm is the time of flight in free-space.

Samsung, EX1028, p. 112

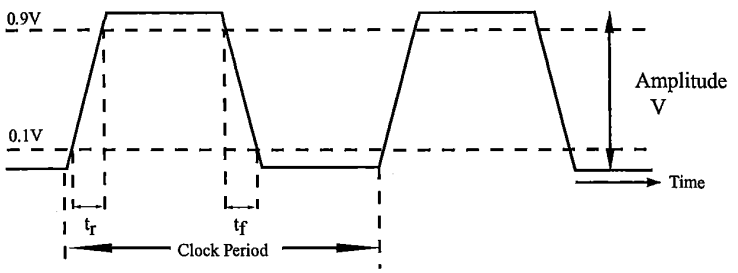


FIGURE 4.20 Clock waveform.

EXAMPLE 4.3

For a repetitive signal such as a clock, the rise time t_r can be estimated from the clock frequency. The reciprocal of the clock frequency provides the clock period, as shown in Figure 4.20. The rise time is typically around 10% of the clock period. Hence, for a 500 MHz clock, the clock period is 2 ns and the rise time is 200 ps. For a dielectric with $\varepsilon_r = 4.0$, interconnects with length greater than 1.5 cm need to be treated as transmission lines.

4.4.10 Crosstalk between Interconnections

Figure 4.21 shows the cross-section of two closely spaced traces in a uniform, homogeneous environment. Traces that are closely spaced give rise to the phenomenon of *crosstalk,* where some of the energy carried by a signal on one line is transferred to the second line.

Crosstalk occurs because of capacitive and inductive coupling between the lines. This coupling occurs continuously along the lines, and makes it possible for some of the energy carried by a signal in one line to become transferred to the other. In most practical situations, only a small percentage of the energy is transferred to the second line. The lines are said to be *weakly coupled* or *loosely coupled.*

The analysis of coupled transmission lines is rather complex. In the weakly coupled case, a signal traveling from left to right on the first line (the aggressor line) is not significantly perturbed by the second line (the victim line). However, when a signal is launched on the aggressor line, it creates two waves or signals on the victim line. One signal propagates towards the driver end and is called the backward or near-end crosstalk. The other signal propagates towards the receiver end and is called the forward or far-end crosstalk. An aggressor line coupling energy into the victim line is shown in Figure 4.22.

The forward crosstalk at the far end of the victim line is:

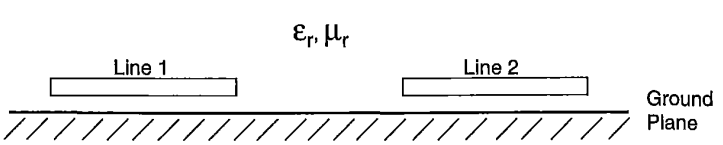


FIGURE 4.21 Coupled interconnections.

Samsung, EX1028, p. 113

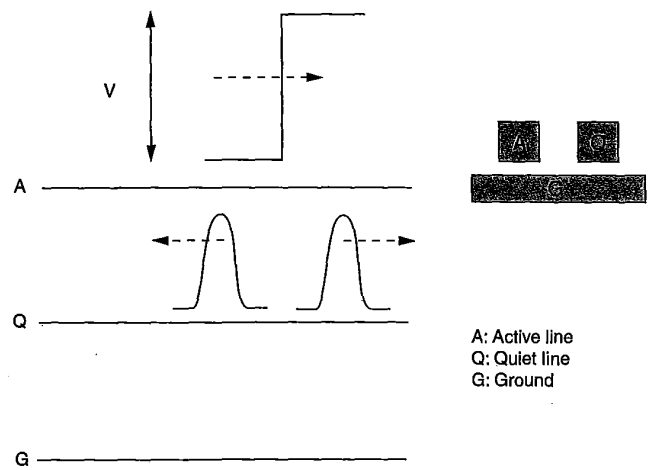


FIGURE 4.22 Crosstalk waveforms.

$$v_2(l,t) = K_F \times \left(\frac{\partial V^{\text{inc}}\left(t - \dfrac{l}{v_p}\right)}{\partial t} \right) \times 2 \times \tau \times l \tag{4.48}$$

where K_F and τ are the forward coupling coefficient and delay given by:

$$\begin{aligned} K_F &= \frac{1}{4}\left(\frac{E}{C} - \frac{M}{L}\right) \\ \tau &= \sqrt{LC} \end{aligned} \tag{4.49}$$

Parameters M and E are the mutual inductance and mutual capacitance per unit length. Similarly, the backward crosstalk at the near end of the victim line is:

$$v_2(0,t) = K_R \left\{ V^{\text{inc}}(t) - V^{\text{inc}}\left(t - \frac{2l}{v_p}\right) \right\} \tag{4.50}$$

where K_R is the backward coupling coefficient given by:

$$K_R = \frac{1}{4}\left(\frac{E}{C} + \frac{M}{L}\right) \tag{4.51}$$

In the above equations, V^{inc} is the voltage launched from left to right on the aggressor line in Figure 4.22 and v_p is the propagation velocity. The incident voltage is of the form $V^{\text{inc}}(t)$. The voltage $V^{\text{inc}}(t - l/v_p)$ is the same signal delayed in time by l/v_p seconds. Similarly, the voltage $V^{\text{inc}}(t - 2l/v_p)$ is the same signal delayed in time by $2l/v_p$ seconds. It is interesting to note that the far-end crosstalk varies linearly with length, while the near-end crosstalk saturates to a maximum value.

Samsung, EX1028, p. 114

Although these equations were obtained for homogeneous lines, the idealized situation of homogeneous lines is seldom realized; even if there is only one medium surrounding the trace, since nearby vias, traces crossing the path of the line in question, etc. perturb the environment enough to cause some coupling. Therefore, both K_F and K_R are assumed to be nonzero. An example will illustrate the coupling process.

EXAMPLE 4.4

Consider two coupled signal lines as shown in Figure 4.23. A voltage wave is launched on the aggressor line. To apply the preceding equations, the incident signal on line 1 and its derivative need to be determined. Figure 4.24 depicts the derivative of the signal at the far end. Using equation

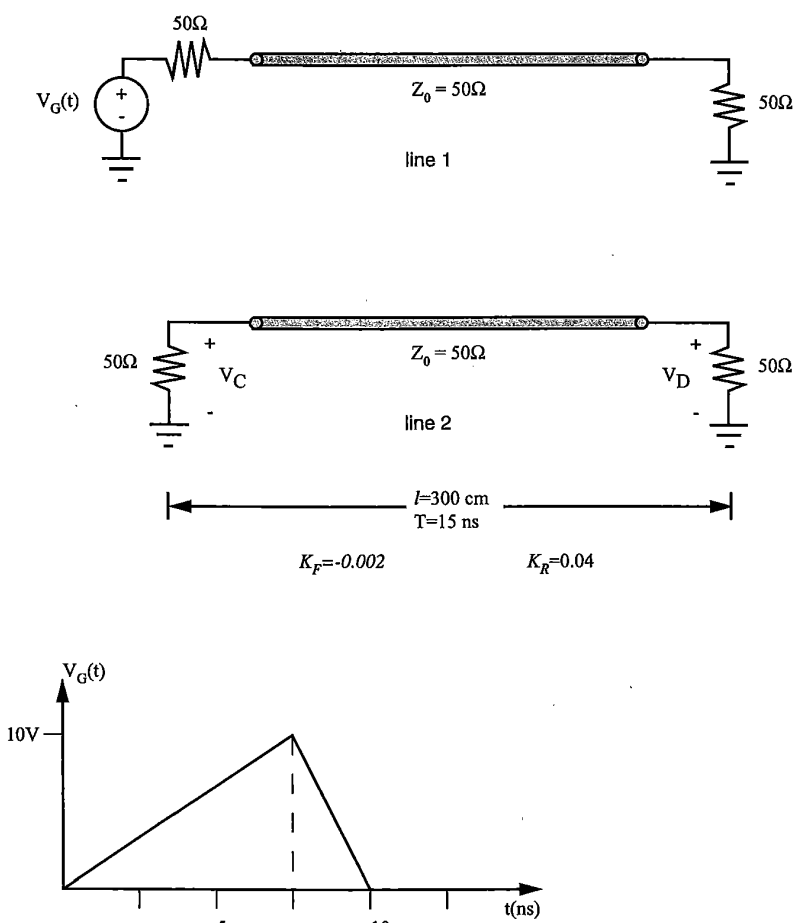


FIGURE 4.23 Coupled interconnection circuit.

Samsung, EX1028, p. 115

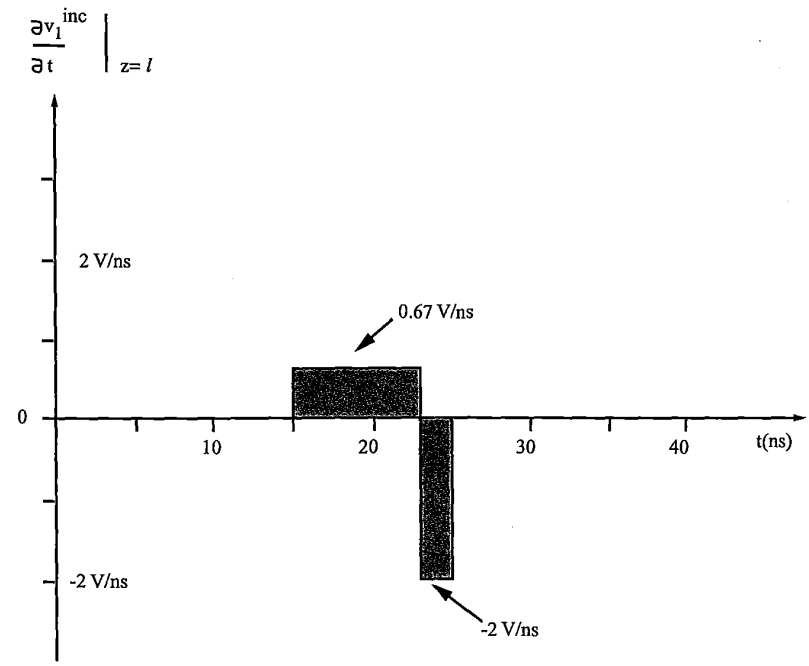


FIGURE 4.24 Derivative at far end.

(4.48), the forward crosstalk at the far end of the victim line is constructed as shown in Figure 4.25, while the backward crosstalk at the near end of the victim line is shown in Figure 4.26 using equation (4.50).

From the standpoint of a circuit designer, crosstalk is undesired and should be reduced or eliminated if possible. One way to minimize crosstalk is to design the transmission line cross-section in order to reduce the coupling coefficients as much as possible. This can be accomplished by increasing the spacing between traces, by increasing the coupling to ground planes, or by routing unused or grounded traces between the active lines. If the medium is homogeneous and the lines are symmetric, K_F will vanish or at least be very small. Another obvious strategy is to limit the length of lines running parallel to each other.

Other design parameters can be modified to reduce crosstalk. For example, a reduction in the rise time or fall time of a signal will reduce the forward crosstalk on nearby lines, which is proportional to the time derivative of the signal. Aspects of the circuit layout also directly impact the amount of crosstalk present, since critical parts of the circuit can be physically separated to minimize coupling. As a general rule, designers should attempt to avoid crowding signals together on any one part of a system.

4.5 POWER DISTRIBUTION

In the previous sections, the interconnects were viewed as lines carrying signals. Depending on the rise time of the signal and the length of the interconnect, they were

Samsung, EX1028, p. 116

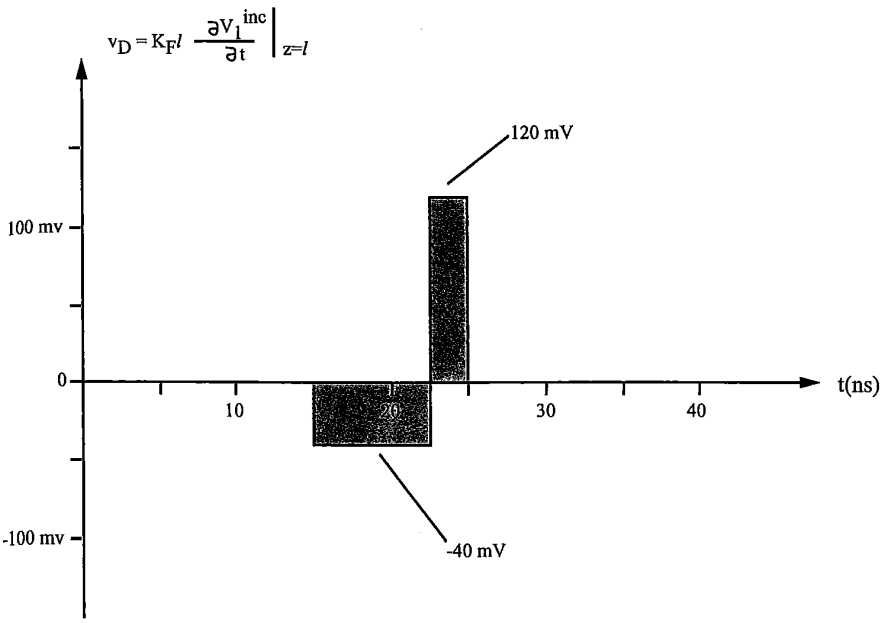


FIGURE 4.25 Far-end crosstalk.

treated as capacitive loads or as transmission lines. However, both the driver that generates the signal and the receiver that receives the signal require voltage and current to function, as shown in Figure 4.1. This aspect of electrical design is called *power distribution.*

Two important issues need to be addressed in power distribution: the *IR* voltage drops and the inductive effects. When DC current I is supplied to the circuits, the finite resistance R of the package metal layers causes a voltage drop given by Ohm's Law $V = IR$, where I is the current drawn by the circuit. Since the *IR* drop can vary across the chip, the supply voltage for all the circuits may not be the same. This variation of the DC supply voltage on the package metal lines can cause the false transitioning of the circuits for spurious input signals. Figure 4.27 shows a simple circuit for estimating the impact of the package *IR* drop on chip performance. In the circuit, R is the resistance of the package metallization, which includes vias, interconnects and planes.

A more dominant effect occurs when the circuits transition, since a time varying current now passes through the metal layers in the package. Since metal layers are inherently inductive, the time varying current causes the supply voltage at the chip to vary with time. Hence, the supply voltage oscillates around the DC level with time. This can once again cause the false transitioning of the circuits. Both these effects have to be minimized for the system to perform appropriately.

Inductive effects associated with power distribution represent a major bottleneck for future systems and will be discussed in the remaining sections.

Samsung, EX1028, p. 117

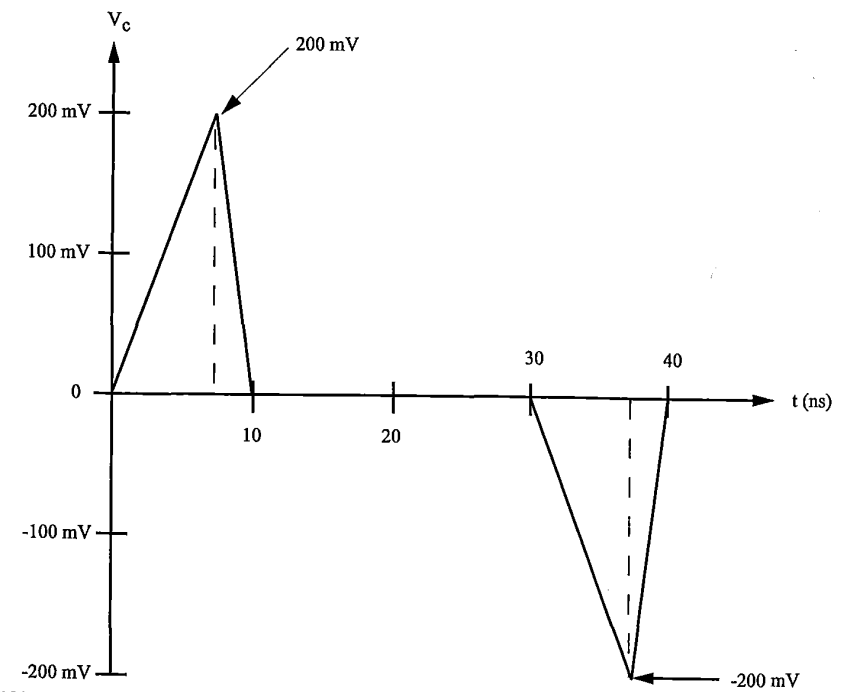


FIGURE 4.26 Near-end crosstalk.

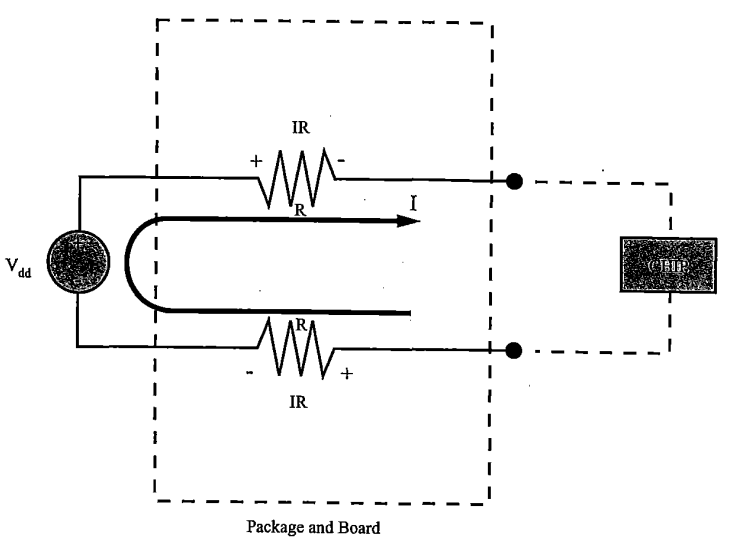


FIGURE 4.27 Resistive drop.

Samsung, EX1028, p. 118

4.5.1 Inductive Effects

Due to the current delivered to the different parts of the chip through the package by the power supply, all physical circuits contain some inductance in series with the power supply. The inductance associated with the power and grounding circuitry of a product can be appreciable, and generally cannot be neglected by a serious electrical designer.

Consider a PMOS transistor with a supply voltage V_{supply} driving a load. Figure 4.28 depicts a simple equivalent circuit where V_{supply} is the ideal power supply. Inductors L_1 and L_2 represent the inductance in the package for the power and ground leads, respectively. The switch represents the transistor with zero on-resistance and is driving a load of resistance R. Hence, the circuit depicts a chip being powered through the package inductance. Since the voltage and current for an inductor are related by:

$$V = L\frac{di}{dt} \tag{4.52}$$

the following equation can be obtained for the current loop in Figure 4.28 by applying Kirchhoff's Voltage Law (KVL):

$$V_{supply} = (L_1 + L_2)\frac{di}{dt} + Ri(t) \tag{4.53}$$

An analysis similar to that outlined in the capacitance section when the switch is closed at time $t = 0$ yields the solution:

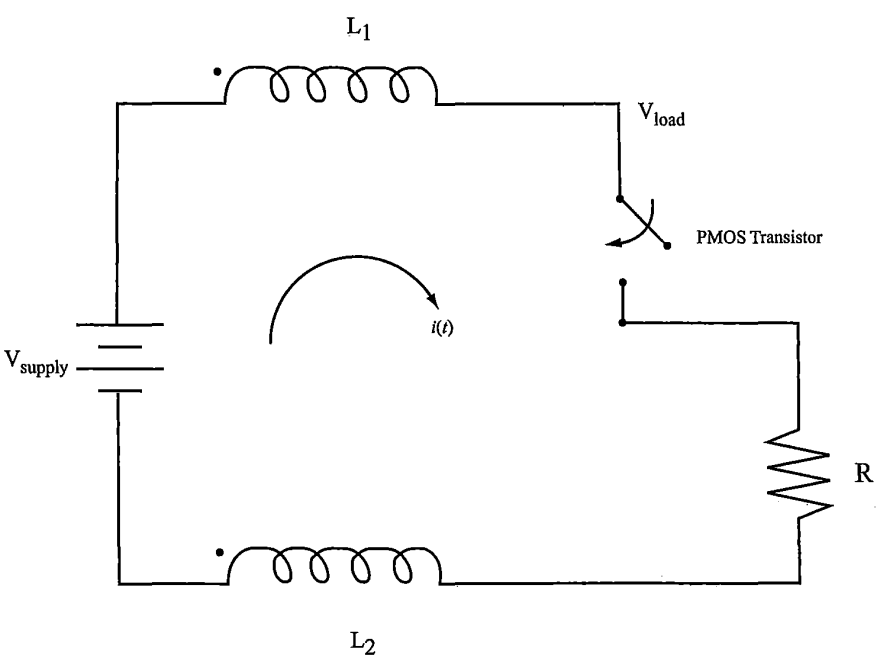


FIGURE 4.28 Power distribution inductance.

Samsung, EX1028, p. 119

$$i(t) = V_{\text{supply}} \frac{(1 - e^{-t/\tau})}{R} u(t) \tag{4.54}$$

where the time constant τ is determined to be:

$$\tau = \frac{L_1 + L_2}{R}$$

By combining this result with Equation (4.52), the load voltage is found as:

$$V_{\text{load}}(t) = V_{\text{supply}} (1 - e^{-t/\tau}) u(t) \tag{4.55}$$

The load voltage is the supply voltage to the PMOS transistor, which has a time constant, associated with it.

EXAMPLE 4.5

For illustration, consider the specific numerical values $V_{\text{supply}} = 5$ (V), $R = 50\ \Omega$, $L_1 = 1$ nH and $L_2 = 1$ nH. For these values, the time constant associated with the response of the system is:

$$\tau = L/R = 4 \times 10^{-11} = 40 \quad \text{(ps)} \tag{4.56}$$

Figure 4.29 shows a plot of the load voltage for these values. Under these conditions, the supply voltage to the transistor rises with a time constant of 40 ps. The transistor requires the supply voltage V_{supply} to function correctly. Any delay in the response of the power supply, measured as $V_{\text{load}}(t)$, leads to a delay in the output of the transistor. Hence, the inductance of the package power distribution network causes the circuits to slow down by introducing appreciable delays to the supply voltage of the circuits.

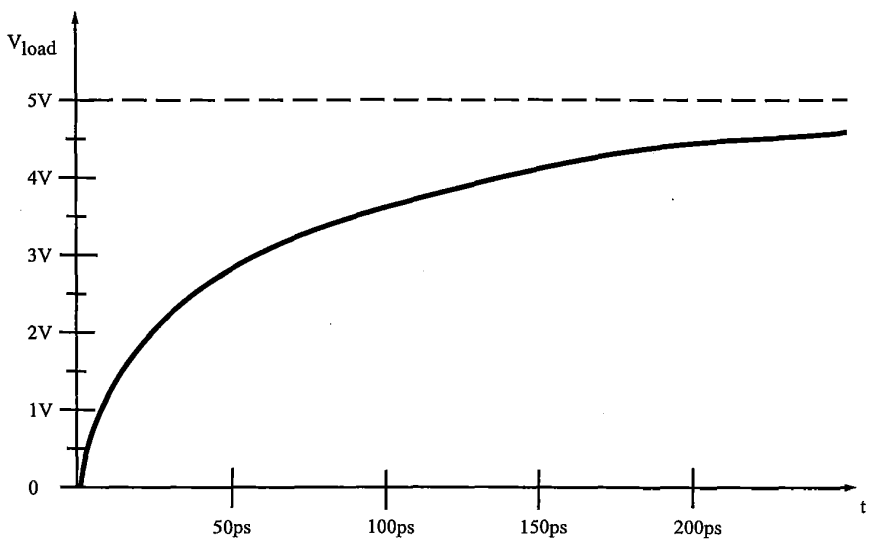


FIGURE 4.29 Inductive effect.

Samsung, EX1028, p. 120

4.5.2 Effective Inductance

In this section, a term called effective inductance L_{eff} will be introduced. The effective inductance is used by designers to evaluate the power distribution performance of a package.

Consider Figure 4.28. From Equation (4.53), the inductance looking into the power supply from the transistor is $L = L_1 + L_2$. This is for the current loop shown in the figure. This inductance represents the effective inductance L_{eff} for the circuit in Figure 4.28. It is very important to note that the effective inductance can only be defined for a loop of current and has no meaning in the absence of this loop—a mistake often made by package designers. In the figure, the loop is completed through the inductors, the switch, the load resistor and the power supply. The effective inductance for the previous example is $L_{eff} = 2$ nH.

As mentioned earlier, inductors L_1 and L_2 are the inductances of the voltage and ground leads of the package. The physical proximity of these leads often causes a magnetic flux linkage whereby current flowing through inductor L_1 causes a current to be induced in inductor L_2 and vice versa. The direction of the induced current is governed by Lenz's Law. This effect can be captured using a mutual inductance M between the two inductors in Figure 4.28.

Equation (4.53) can now be modified to include the mutual inductance whose contribution to the loop inductance can be either *additive* or *subtractive.* The contribution is additive if the current in the two inductors is in the same direction and is subtractive otherwise. In Figure 4.28, the current loop causes the current to flow in opposite directions in the two inductors, thus causing a reduction in the loop inductance. Assuming a mutual inductance M between the two inductors, Equation (4.53) becomes:

$$V_{supply} = (L_1 + L_2 - M - M)\frac{di}{dt} + Ri(t) \tag{4.57}$$

leading to a time constant:

$$\tau = \frac{L_1 + L_2 - 2M}{R}$$

and an effective inductance $L_{eff} = L_1 + L_2 - 2M$. The effective inductance is therefore reduced due to the mutual inductance between the voltage and ground leads of the package. The overall inductance can be minimized by routing the supply and ground paths as close together as possible.

EXAMPLE 4.6

For the previous example, let $M = 0.5$ nH. For this value of mutual inductance, the time constant is $\tau = 20$ ps and the effective inductance $L_{eff} = 1$ nH. Under these conditions, the supply voltage to the transistor rises faster and increases the speed of the circuit. For a designer, the goal is always to reduce the effective inductance of the package from the chip to the power supply on the board.

Samsung, EX1028, p. 121

4.5.3 Power Supply Noise

Noise is present in every system. In this section, noise sources associated with the power distribution network are identified and the impact of that noise on package design is considered. As mentioned earlier, circuits within a chip require voltage and current to switch binary states. Earlier in this chapter, CMOS transistors were discussed. In CMOS transistors, the input node of the transistor behaves like a capacitor, which acts as a load for the preceding stage. Hence, the current from a driving circuit is used to charge or discharge the load capacitance, thereby raising or lowering the voltage at the input node. This transition causes the succeeding stage to switch binary states. The current for charging the load is drawn from the V_{dd} supply. Similarly, the current generated by discharging the load is drained into the GND supply. Both these currents are time varying currents. The voltage and current to the chip are supplied from the power supply on the motherboard through the metal layers on the package and PWB. This is shown as an equivalent circuit in Figure 4.30 where the package and PWB metal layers add inductance to the power distribution network. Thus, the route that the power supply current follows can be relatively long and circuitous.

Figure 4.30 shows a packaged chip drawing current from an ideal power supply. An ideal power supply can provide infinite current; real power supplies cannot. The local supply on the chip acts as a non-ideal supply, primarily because the route taken by the power supply current through the board, package, and chip introduces a series inductance L into the system. In the previous section, the time delay associated with an LR circuit was considered. There is an additional, more serious side effect of series inductance in the power supply circuitry: whenever the power supply current changes in the presence of series inductance, the local V_{dd} level will drop since the current flowing through an inductor cannot change instantaneously. Consequently, the transient supply currents flow-

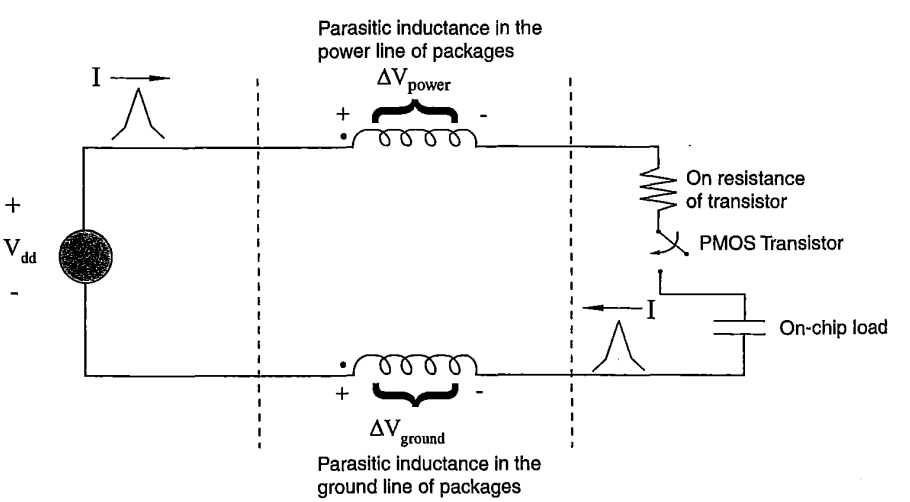


FIGURE 4.30 Power supply noise.

Samsung, EX1028, p. 122

ing through these interconnections cause voltage fluctuations on the power supply rails of the chip, as described by the equation:

$$\Delta V = L_{\text{eff}} \, (dI/dt) \tag{4.58}$$

In the above equation, L_{eff} is the effective inductance of the power distribution system, dI is the change in current in a time period dt and ΔV is the resulting voltage fluctuation or power supply noise. This noise is often called *simultaneous switching noise* (SSN), delta-I noise, ΔI noise or dI/dt noise. In equation (4.58), the part of the power distribution system within the package is usually the main contributor to the inductance.

As mentioned earlier, L_{eff} represents the inductance of a loop of current from the circuitry on the chip to the ideal power supply on the board. In a package containing thousands of vias and interconnections, it is very difficult to traverse the path taken by the current. Hence, software tools are used to extract the L_{eff} of the package.

With improved circuit speed, the rate at which currents are switched is increasing. As a result, within practical systems the chip current may change by large magnitudes in small time intervals (large dI/dt) causing power supply fluctuations. Since power supply fluctuations can be catastrophic, methods to reduce the inductance L_{eff} of the package need to be investigated. Current trends in chip design are that the voltage and signal levels are decreasing according to CMOS scaling rules, so the power supply noise budget must also decrease for future systems. The need for low inductance has motivated the migration towards vertical power distribution in packages such as MCMs due to their relatively small power supply inductance properties.

In order to mitigate the effects of power supply noise, an electrical designer must use the noise budget to determine the maximum permitted inductance L_{eff} associated with the power and ground distribution network. The designer must then design the supply routes to ensure an inductance below the maximum level. Below, techniques for minimizing inductance and estimating the power supply noise are considered.

4.5.4 Effect of Packaging Technology on Inductance

It is apparent from Equation (4.58) that minimizing inductance reduces power supply noise. The main contributor to the inductance is the package. In this section, let's take a careful look at the physical parameters of the package that contribute to the inductance and the role of technology in minimizing inductance.

In an earlier section, the effect of the mutual inductance between the voltage and ground leads of the package for reducing the effective inductance was discussed. The amount of inductance is dictated by the physical structure and is defined by the packaging technology.

In packages such as the *quad flat pack* (QFP), the leads on the periphery of the package are used to supply power to the chip. These leads are equivalent to wires with diameter a and spacing d, as shown in Figure 4.31a. Two adjacent leads supplying voltages V_{dd} and GND to the chip are similar to wires carrying equal and opposite currents. The loop formed between a pair of wires results in an effective inductance, which can be calculated from the physical parameters, as shown in Table 4.4. The L_{eff} for a pair of wires decreases with a decrease in the spacing d between the wires since this results in an increase in the mutual inductance between the wires. For example, a wire with di-

Samsung, EX1028, p. 123

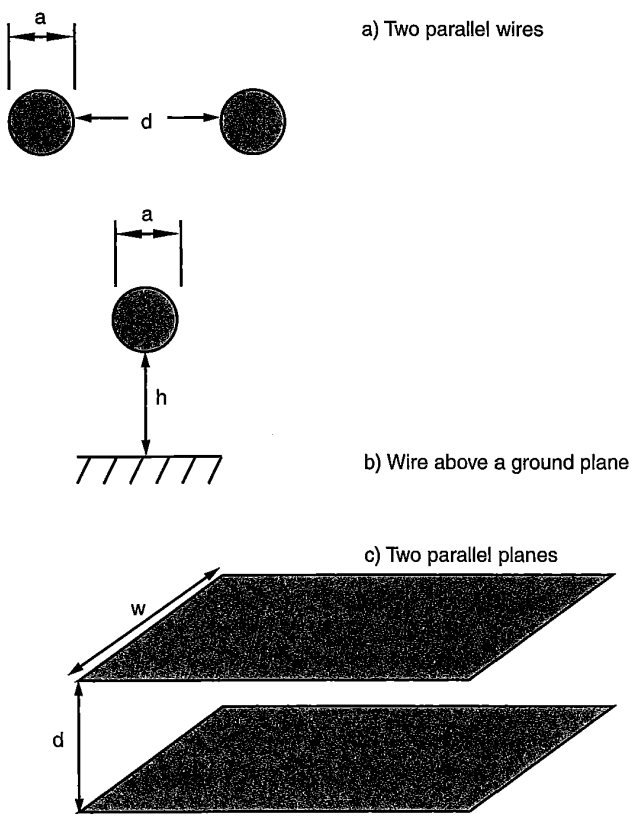


FIGURE 4.31 (a) Parallel wires (b) wire above a ground plane (c) parallel planes.

ameter 1 mil, length 1 cm and spacing 0.8 mm results in an inductance of 4.14 nH. For the same dimensions, the inductance reduces to 3.45 nH for a spacing of 0.4 mm. Since the total leads on a QFP are limited, very few voltage and ground leads are available to supply the power. This leads to an increase in the effective inductance of the package.

Packages such as the QFP can contain a plane of metallization. Planes are layers in a package where the entire area is covered with metal. The plane could be used to supply the GND voltage while the traces on an adjacent layer could be used to supply V_{dd}. This results in the structure shown in Figure 4.31b. The inductance for a wire above a ground

TABLE 4.4 Inductance comparison.

Structure	Inductance/ Length	Package	Inductance
Parallel wires	$\mu/4\pi \cosh^{-1}(d/a)$	DIP, QFP	3–50 nH
Wire above a ground plane	$\mu/4\pi \ln(4h/a)$	TAB, QFP w/ground plane	1–10 nH
Parallel planes	$\mu\, d/w$	PGA, BGA, DCA	0.25–1 nH

Samsung, EX1028, p. 124

plane can be calculated from its physical parameters, as shown in Table 4.4. The L_{eff} for a wire above a ground plane decreases with a decrease in the height h due to an increase in the mutual inductance between the wire and the ground plane. As an example, a wire with diameter 1 mil, height 0.8 mm and length 1 cm has an effective inductance of 4.84 nH. When the height is reduced to 0.4 mm, the effective inductance reduces to 4.14 nH. The usefulness of employing a ground plane is that it can be shared by many V_{dd} traces. This can lead to uniformity in the effective inductance across the package.

For multilayered BGA packages, plane layers are used to supply both V_{dd} and GND, as shown in Figure 4.31c. For a pair of planes, the effective inductance can be calculated from the physical dimensions, as shown in Table 4.4. For a pair of metal planes, the inductance decreases with a decrease in the separation d between the layers due to an increase in the mutual inductance. As an example, a plane of width 1 cm, length 1 cm and spacing 6 mils has an effective inductance of 0.19 nH. The inductance reduces to 0.1 nH when the spacing is reduced to 4mils. Clearly, the inductance for a pair of planes is much lower than what can be achieved with a pair of wires. For high performance packages, planes in the package are absolutely required for supplying power to the chip. The effective inductance due to the plane is sometimes referred to as the *spreading inductance.*

The three structures shown in Figures 4.31a, 4.31b and 4.31c are the basic structures used in a package to supply voltage and current to the circuits. Of the three structures, a pair of planes always has the smallest inductance. Vias provide the connection to the planes and between layers. Vias are like wires. Their effective inductance can be reduced by having adjacent voltage and ground vias to ensure opposite current flow. In addition, vias can be connected in parallel to further reduce the inductance. Since wires have larger inductance than spheres, use of solder balls instead of wirebonds can help reduce the inductance. Based on the application type, any one of these structures may be used to design the package. Table 4.4 summarizes the packages that use these structures and the approximate inductance that they support.

For multilayered packages containing planes, the calculation of the inductance depends on the distribution of the current on the planes. Since the current distribution in a package is a function of the frequency of operation, the spatial position of the signal lines and the spatial position of the vias, inductance calculations can be very complex. Table 4.4 captures the fundamental concepts related to inductance associated with the power distribution network in a package.

4.5.5 Inductance and Noise Relationships for On-Chip Circuitry

On-chip circuits are logic gates contained within a chip that communicate with each other. Noise resulting from these circuits is often called "core" noise. Consider a CMOS inverter circuit driving a capacitive load, as shown in Figure 4.30. As discussed earlier, the inverter can be represented by a pair of time dependent switches that connect the load capacitor to either V_{dd} or GND, thereby charging or discharging the capacitor. In Figure 4.30, the load capacitor is being charged to V_{dd} through the PMOS transistor and the current loop is being completed through the ground connection. Both the voltage and ground connections have inductance, as shown in the figure.

Let the load capacitor be initially uncharged. Based on the circuit in Figure 4.30, the electric charge required to energize the load capacitor to a voltage V_{dd} is:

Samsung, EX1028, p. 125

$$Q = CV_{dd} \tag{4.59}$$

This charge will be supplied by a current increase ΔI that will occur for a time Δt. Since $Q = \Delta I \times \Delta t$, using a first order approximation, charging the load capacitor will require a change in the circuit current given by:

$$\Delta I = CV_{dd}/\Delta t \tag{4.60}$$

The current in Equation (4.60) has to be supplied by the power supply V_{dd} and is for a single gate switching. When N gates switch simultaneously, the current drawn from the power supply is:

$$\Delta I_{\text{total}} = NCV_{dd}/\Delta t \tag{4.61}$$

The voltage drop across the inductor is equivalent to the voltage drop between the supply and chip terminals. From Equation (4.58), it follows that

$$\Delta I_{\text{total}} = \Delta V\, \Delta t/L_{\text{eff}} \tag{4.62}$$

In steady-state, $\Delta V = 0$. Hence the power supply and the chip terminal are at the same potential. However, when the circuits switch, current flows to charge the load capacitor. This causes the chip terminal voltage to diverge from the power supply due to the finite inductance of the voltage and ground leads. Remember that the inductance L_{eff} is for the loop defined in the circuit.

Since logic circuits have a specification on the maximum voltage fluctuation ΔV they can tolerate, Equations (4.61) and (4.62) can be combined to estimate the maximum effective inductance L_{eff} that can be tolerated on the power distribution:

$$L_{\text{eff}} \leq \frac{\Delta V\,(\Delta t)^2}{NCV_{dd}} \tag{4.63}$$

EXAMPLE 4.7

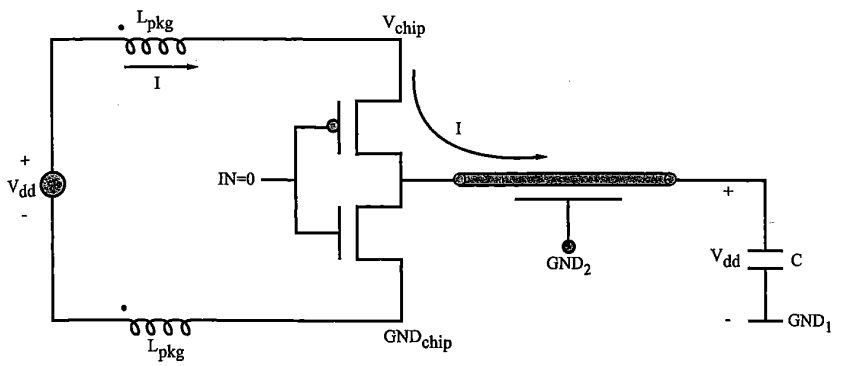
Consider an inverter circuit driving a 1 pF capacitive load. For the inverter to operate correctly, a power supply voltage variation not to exceed 10% of V_{dd} is desired. If the circuit switches binary states in time 10 ns, the effective inductance for the power supply is 10 μH, based on Equation (4.63). As the switching time is reduced to 100 ps, the effective inductance reduces to 1 nH. When 1000 of these circuits switch simultaneously, the effective inductance reduces to 1 pH. A modern microprocessor such as the Intel Pentium II may have hundreds of circuits switching simultaneously.

From Equation (4.63), for a given power supply noise and capacitive load, the maximum package inductance that can be tolerated is inversely proportional to the number of switching drivers and directly proportional to the square of the transition time. Equation (4.63) assumes that the time of flight delay is insignificant and the transition time Δt for the current and voltage are the same.

4.5.6 Inductance and Noise Relationship for Output Drivers

Many contemporary electrical circuits consist of IC chips that communicate with each other. Communication is made possible through output drivers that support signaling between chips. An output driver connected to a package interconnect behaving as a transmission line is shown in Figure 4.32. The driver functions as a switch as discussed

Samsung, EX1028, p. 126



Nodes GND_{chip}, GND_2, GND_1 indicate that the ground potential can vary across a system

FIGURE 4.32 I/O noise.

earlier. The driver circuit is powered by the V_{dd} and GND supply of the chip which contain parasitic inductance from the package, as shown in Figure 4.32. When the driver switches binary states, it launches a voltage and current wave on the interconnect which are related by:

$$I = V/Z_0 \tag{4.64}$$

where Z_0 is the characteristic impedance of the transmission line. Equation (4.64) describes the relationship between the voltage and current at any position on a transmission line for a forward traveling signal, as discussed earlier. Both the current and voltage are time dependent waveforms that propagate at the speed of light in the dielectric medium. As before, the current is drawn from V_{dd} to charge a load capacitor or drained into GND to discharge a load capacitor, where the load capacitor is located at the far end of the transmission line. Assuming that the transistor switches from GND to V_{dd}, the current drawn from the supply can be calculated from Equation (4.64) as:

$$\Delta I = \frac{V_{dd}}{Z_0} \tag{4.65}$$

Since the current flowing through the V_{dd} and GND inductance takes a finite time Δt, it causes a time varying voltage drop across the inductor, given by Equation (4.58). This voltage drop across the inductor causes the chip terminal voltage to diverge from the supply voltage. When N drivers switch simultaneously, the total current drawn from the power supply is:

$$\Delta I_{\text{total}} = N\frac{V_{dd}}{Z_0} \tag{4.66}$$

Combining Equations (4.58) and (4.66), the maximum effective inductance that can be tolerated on the chip terminals is:

Samsung, EX1028, p. 127

$$L_{\text{eff}} \leq \frac{\Delta V Z_0}{N \frac{V_{dd}}{\Delta t}} \tag{4.67}$$

The above equation assumes linearity in the operation of the drivers. However, drivers by their very nature are nonlinear. When a small number of drivers switch simultaneously, they produce noise on the power supply. This can be estimated based on a linear model by rewriting Equation (4.67) to calculate ΔV. As the number of switching drivers increases, the noise on the power supply increases. When this noise becomes excessive, it slows down the drivers by reducing the current output of the drivers during the time interval Δt. This in turn reduces the power supply noise voltage ΔV. Thus, the noise is initially linear and saturates as the number of switching drivers is increased.

EXAMPLE 4.8

Consider a 64 bit wide bus that is used to communicate between two chips. The bus transitions from binary state one to binary state zero, simultaneously. As a specification for the power supply noise, let's allow a maximum variation of 10%, which for 2.5 V logic is 250 mV. This limit will guarantee that the circuits function correctly. If the interconnections have a characteristic impedance of 50 Ω, a transition time of 1 ns requires a power supply inductance of 78 pH or less.

From Equation (4.67), for a given power supply noise and bus width, the required inductance is directly proportional to the impedance of the interconnections and inversely proportional to the voltage transition rate.

4.5.7 Power Supply Noise Characteristics

When circuits switch simultaneously, the noise generated on the power distribution network may cause the supply voltages to compress or expand. Power supply compression causes the voltage between V_{dd} and GND on the chip to be less than the supply voltage while power supply expansion causes this voltage to be larger than the supply voltage. Both these effects can affect the performance of the system. In addition, the inductance of the package can resonate with the capacitance within the chip to cause oscillations on the power supply.

There is a larger likelihood that the output drivers cause more power supply noise than the internal circuits, since they are nonrandom events. For example, a wide bus can transition simultaneously from binary 1 to a binary 0 state. In addition, based on the impedance of the interconnections, larger currents are necessary to charge the interconnections. Since the on-chip circuitry is very sensitive to noise and operates at higher speed, the power distribution network for the on-chip and off-chip circuitry are isolated for high speed systems. With increasing system speed, power supply noise needs to be suppressed, which can only be accomplished by reducing the inductance of the power distribution network in the package.

Equations (4.63) and (4.67) provide a simple relationship for computing the maximum effective inductance that can be tolerated on the power supply. This is based on a first order approximation. Neither equation accounts for the spatial position of the circuits nor the time associated with signal and noise to propagate between points that are spatially separated. In addition, linearity of noise is not preserved beyond a certain number of

Samsung, EX1028, p. 128

switching drivers since noise saturates, as explained earlier. For output drivers, the noise is also dependent on the path of the return current, a concept that is very important when working with transmission lines. For computer systems operating at high frequency, the switching circuits cause waves to propagate between planes on the power distribution network. These waves reflect from the edges of the package, causing resonance in the structure. This resonance can cause the noise to increase in packages, which no longer can be described using an effective inductance model. Hence, estimation of power supply noise for modern high speed systems requires a more distributed model, or transmission lines, as compared to the single inductor model used in Figures 4.30 and 4.32. Modeling and simulation of SSN is one of the most challenging problems today in the design of high performance packages.

In all systems, suppression of power supply noise is critical. If not controlled, it can have the following effects:

1. Circuits can produce less output current in the presence of power supply noise, which increases circuit delays causing excessive time delay in the system.
2. Noise glitches on the power supply may propagate to quiet receiving chips causing false switching.
3. Noise glitches may cause false switching of the circuits on the sending chip.

4.5.8 Decoupling Capacitors

An earlier section emphasized the fact that charge is required to energize a load capacitor. This charge is supplied by a current from the power supply. When this extra current flows through the package inductance, it produces a voltage drop across it. This was described by the circuit in Figure 4.30. The extra current required to charge the capacitor can also be supplied by capacitors. These capacitors are called decoupling capacitors.

Let's revisit the circuit in Figure 4.30 with an added *decoupling capacitor* C_d across the chip terminals, as shown in Figure 4.33a. Assume that the decoupling capacitor is initially charged to voltage V_{dd}. When the transistor switches, charge $Q = CV_{dd}$ is required to charge the load capacitor. This charge is supplied by a current ΔI that flows in the circuit. However, instead of this current being supplied by the power supply, it is supplied by the decoupling capacitor. Hence, the current loop changes. Since there is no inductance in the path of the loop current in Figure 4.33a, the chip terminal voltage is maintained at V_{dd}. This is true provided that the charge stored in the decoupling capacitor is larger than the charge required by the load capacitor and can be supplied within the time Δt. The time Δt represents the switching speed of the circuit.

Thus, capacitors act as a reservoir of charge by charging to the supply voltage when the circuits are idle and supplying current when there is a demand. This is depicted in Figure 4.33a where the capacitor on the chip supplies charge to the switching circuits before the power supply does. The role of decoupling capacitors is therefore to reduce the size of the current loop in the system. As an example, the size of the current loop in Figure 4.33a is reduced compared to Figure 4.30. Hence, the power supply noise has been reduced as well due to the absence of any inductance in the current loop.

In a computer system, the time interval Δt over which current ΔI is required is a function of the switching speed of the circuits. For example, when the circuits within the core of a processor are switching, large current may be required over a small time interval.

Samsung, EX1028, p. 129

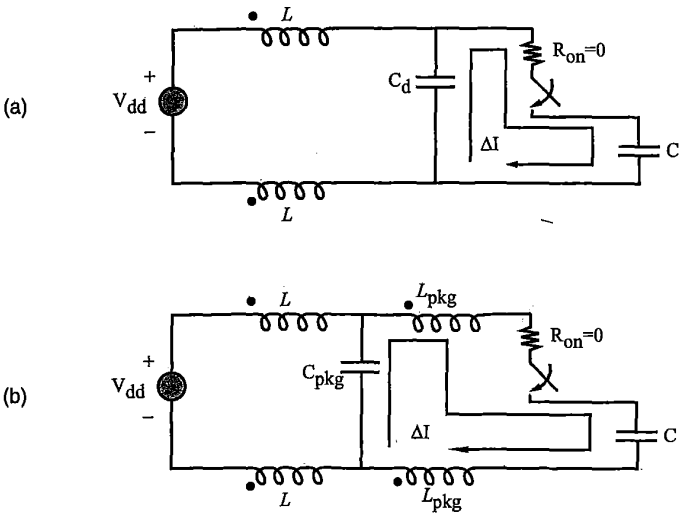


FIGURE 4.33 (a) On-chip capacitor and (b) package capacitor.

However, when the processor is transferring data to memory, the time interval may be much larger. Hence, decoupling capacitors are often distributed across the entire system to supply charge for the switching circuits. The goal once again is to reduce the size of the current loop, so that the capacitor can support the current and hence maintain the appropriate voltage across the chip terminals. Depending on the position of these capacitors, the inductance of the package layers carrying the current from the capacitor may change, thus changing the effective inductance in the circuit. As an example, consider the circuit shown in Figure 4.33b where the capacitor C_{pkg} is mounted on the package. This capacitor provides the current through the package inductance to charge the load capacitor. Since the effective inductance of $2L_{pkg}$ is smaller compared to the effective inductance in Figure 4.30, the power supply noise is reduced. However, the effective inductance $2L_{pkg}$ in Figure 4.33b limits the maximum frequency of operation of the capacitor. The capacitor C_{pkg} resonates with the inductance $2L_{pkg}$ at a frequency:

$$f_{max} = \frac{1}{2\pi\sqrt{2L_{pkg}C_{pkg}}} \tag{4.68}$$

Beyond the frequency f_{max}, the loop in Figure 4.33b becomes inductive and hence cannot suppress the noise voltage.

Based on the resonant frequency described in equation (4.68), capacitors can be categorized into high-frequency, mid-frequency, and low-frequency capacitors. This is based on where the capacitors are mounted. The frequency f_{max} is the frequency up to which the capacitor can maintain the power supply variation within specifications. Low-frequency capacitors are mounted on the printed wiring board and are effective up to f_{max} = 100 MHz. Mid-frequency capacitors are mounted on the package, especially on MCMs with f_{max} = 500 MHz. High-frequency capacitors are buried in the chip as trench capacitors and have $f_{max} > 1$ GHz.

Samsung, EX1028, p. 130

In Equation (4.68), the capacitors resonate with the effective inductance in the circuit. In addition, capacitors are never ideal and exhibit a self resonant frequency. This is due to the parasitic inductance and resistance which are produced by the leads and mounting pads of the capacitor. These parasitic parameters are known as the *equivalent series inductance* (ESL) and *equivalent series resistance* (ESR), as shown in Figure 4.34. The impedance of the capacitor is:

$$Z = j\omega L_{\text{ESL}} + \frac{1}{j\omega C} + R_{\text{ESR}} \tag{4.69}$$

where $\omega = 2\pi f$ is the angular frequency.

A plot of the impedance as a function of frequency is shown in Figure 4.34. The impedance is a minimum at the self resonant frequency:

$$f_{\text{SR}} = \frac{1}{2\pi\sqrt{L_{\text{ESL}}C}} \tag{4.70}$$

and has a magnitude of R_{ESR}. Based on the earlier discussion, the capacitor is therefore effective up to a frequency f_{SR}. A combination of Equations (4.68) and (4.70) dictates the maximum frequency up to which the power supply noise can be minimized by using the capacitor.

Earlier, it was mentioned that planes result in the minimum spreading inductance. The planes shown in Figure 4.31 also form a parallel plate capacitor with capacitance:

$$C = \varepsilon wl/d \tag{4.71}$$

where w is the width of the plane, l is the length of the plane, ε is the dielectric constant of the insulator material and d is the dielectric thickness. In equation (4.71), maximum capacitance is obtained for large plane area and small plane separation. Since the planes are in close proximity to the chip (a small separation along the vertical axis), they act as a very good high frequency capacitor. Hence, planes act as a good charge source and have a small inductance associated with them. These are desirable properties for high-

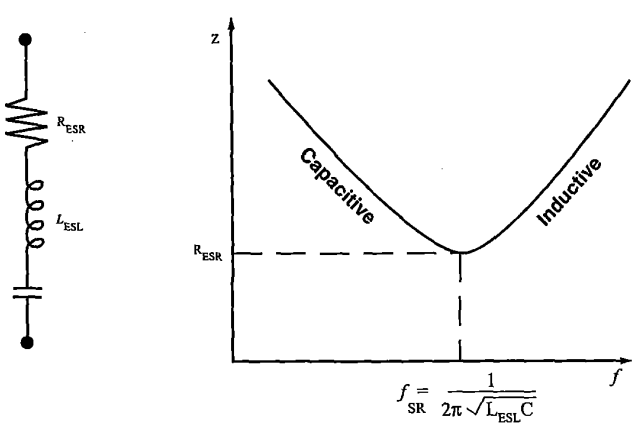


FIGURE 4.34 Parasitics of decoupling capacitors.

Samsung, EX1028, p. 131

speed packages. As new technologies are developed, the capacitance of the dielectric layers could increase, thus increasing the buried capacitance between the planes in the package. In addition, thin separation between the planes will reduce the effective inductance. As these technologies mature, the capacitance between the plane layers of the package could be used to maintain the low noise characteristics of the power supply for frequencies higher than 500 MHz.

EXAMPLE 4.9

Consider a package that has a power supply inductance of 10 pH and has to support the switching of 1000 on-chip circuits. These circuits draw 10 A of current in time 0.25 ns. From Equation (4.58), this results in a power supply noise voltage of 400 mV. Assume that the goal is to reduce the noise voltage to 200 mV. This can be made possible by using decoupling capacitors. The capacitance required can be computed as:

$$C = \Delta I \, \Delta t / \Delta V \tag{4.72}$$

For $\Delta I = 10$ A, $\Delta t = 0.25$ ns and $\Delta V = 200$ mV, this results in a capacitance of 12.5 nF. This capacitance limits the power supply noise to no more than 200 mV. The current flowing through the inductance along the supply path inhibits the maximum frequency f_{max} of the capacitor. For an effective inductance of 5 pH (a capacitor placed very close to the chip), $f_{max} = 637$ MHz. Hence, the capacitor cannot maintain the power supply noise within 200 mV for frequencies beyond 637 MHz. It is assumed here that the parasitic inductance and resistance of the capacitor are negligible.

The capacitor value can also be calculated in the frequency domain. The impedance needed to limit the power supply noise to 200 mV for a 10 A current change is:

$$Z = \frac{V}{I} = 20 \text{ m}\Omega \tag{4.73}$$

This is supplied by a capacitor:

$$C = \frac{1}{2\pi f \, 20 \times 10^{-3}} \tag{4.74}$$

at a frequency of 637 MHz, $C = 12.5$ nF. For a resonant frequency of 637 MHz, the effective inductance along the current path is 5 pH.

4.6 ELECTROMAGNETIC INTERFERENCE

The widespread use of radio-frequency and high-speed digital devices and systems has drawn attention to *electromagnetic interference* (EMI). As the integration of device, package, and board is improved, the likelihood of electromagnetic interference between the circuits, packages and systems increases. As the circuitry has become smaller and more sophisticated, more circuits are being crowded into less space, thus increasing the probability of interference. Moreover, as the clock frequency of the digital circuits and systems passes 1 GHz, the high-frequency noise of the fundamental clock signal and its harmonics are generated in the device, package and system, resulting in significant radiated emis-

Samsung, EX1028, p. 132

sion. An EMI-free design has become more difficult. The electromagnetic interference has the form of free-space electromagnetic radiation and of direct conductive noise through power/ground lines and signal lines.

To ensure the proper operation of nearby circuits and to meet EMI regulations, substantial attention must be paid to EMI from the early stage of the design process. Controlling the emission from digital systems cost effectively can be as complicated and difficult as designing the logic itself. Often the functional performance of a product is not the primary difficulty associated with meeting product schedules; passing the required EMI emission test is the challenge.

In the United States, the *Federal Communications Commission* (FCC) imposes allowable radiated emission and conductive emission for the use of radio and wire communications. Part of its responsibility concerns the control of interference. Part 15 of the FCC Rules and Regulation sets forth technical standards and operational requirements for radio-frequency devices. The FCC rule of general interest is part 15, subpart J, because it applies to almost all digital electronics. Digital devices covered by this definition are divided into two classes. Class A devices are those marketed for use in a commercial, industrial or business environment. Class B devices are those marketed for use in a residential environment. Satisfying these technical standards is the obligation of the manufacturer or importer of a product. Figure 4.35 represents the radiated emission limits for a product when measured at a distance of 3 m. As can be noticed from the figure, the allowed radiated emissions are greater at the higher frequencies, meaning that it is harder to control the higher frequency EMI emission.

Traditionally, EMI problems were mostly confined to the design of the PWB, cable and chassis of radio-frequency devices and high-speed digital systems. However, as frequencies increase, the device and package must also be regarded as a source of EMI.

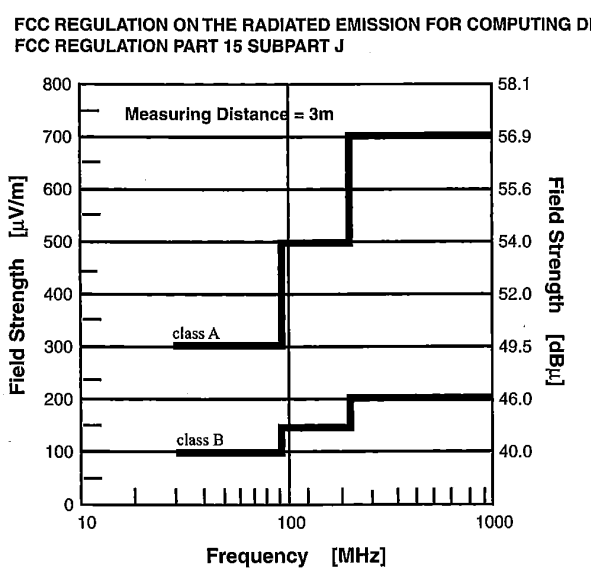


FIGURE 4.35 FCC regulations.

Samsung, EX1028, p. 133

High-frequency noise currents arise from a number of sources, including reflections along signal lines and transients on power/ground lines. These noise currents can flow through parasitic reactances within the package, which are no longer negligible at higher frequencies. The radiated emissions, or electromagnetic fields, are proportional to the frequency of the noise current, the area of the noise current path, and the amount of noise current. Therefore, the most significant radiating structures are large size conducting lines, such as board traces and cabling between boards. These structures are the most likely to behave as an antenna and radiate energy. The package itself is typically not a substantial radiating emitter, unless the package size is much larger than average and the noise frequency is sufficiently high. Usually the size of the package is much smaller than that of the board or cables. Consequently, the EMI issue of primary concern to a package designer is the frequency content and the amplitude of the noise currents.

The control of EMI is a complex problem involving many aspects of design, and strong collaboration between the chip, package, and system designers is needed to address it. Reflections, improper placement of decoupling capacitors, large current loops, and ground voltage fluctuation all contribute to undesired radiated emissions. Reflections on signal lines due to impedance mismatches must be avoided. This is especially true for signal traces exposed on the surface of a PWB, such as a microstrip configuration. Reflections that result in ringing along the signal trace on the PWB cause large radiated emissions at the ringing frequency. A capacitive discontinuity along a transmission line causes an under-shoot of the signal waveform, while an inductive discontinuity causes an over-shoot of the waveform. The amount of reflection generated by the discontinuity depends on the parasitics of the package and bonding structures and the rise time of the digital signal. Therefore, as the transition time of the digital signal becomes shorter, it is more important to tightly control the parasitics of the package and the termination resistors used to match the transmission lines.

The basic function of the decoupling capacitors inside a chip, package or board is to supply current necessary for the proper operation of the chip. Decoupling capacitors stabilize the voltage levels on the power lines, hence minimizing high-frequency noise. When a significant voltage fluctuation is present in the power line of the package, and is coupled to the board power line trace or a power cable, a large radiated emission can be generated. Therefore, it is very important to provide adequate decoupling capacitance inside the package. On the PWB, a decoupling capacitor should be placed at every intersection of the power and ground line. These decoupling capacitors minimize the current loops for the current path. By minimizing current loop size, the magnitude of current and the noise frequency, EMI can be reduced.

When package pins are assigned, a poor selection of ground pins can contribute to EMI difficulties. To minimize the radiated emissions, each signal line and ground return loop must be kept small. It is desirable to position the ground planes and pins very close to the signal lines. Figure 4.36 illustrates the current loop on the PWB. Since the amount of radiation is proportional to the loop area, the ground return areas on the PWB should be placed close to the signal lines to constitute a small loop current path.

As seen in previous sections, the inductance of the package ground network causes voltage fluctuations due to simultaneous switching noise. Differences in ground potential from one non-ideal ground point in a system to another can generate large high-frequency ground currents. If the ground current constitutes a loop, the loop works as a loop

Samsung, EX1028, p. 134

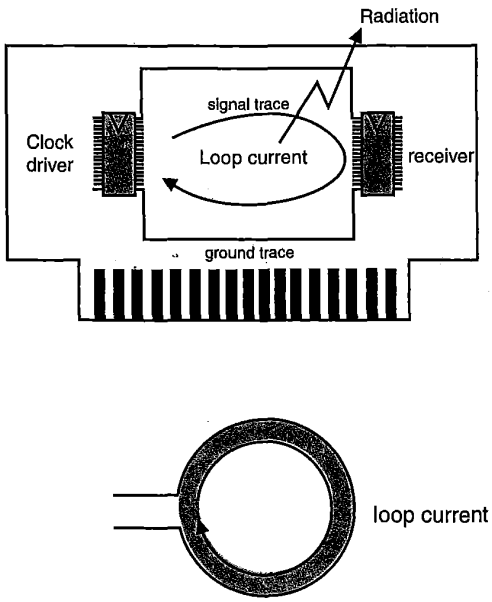


FIGURE 4.36 Loop antenna model.

antenna, generating EMI. If a cable shield is connected to the noisy ground point in a board, the cable shield becomes a dipole antenna, exacerbating the problem. This situation is illustrated in Figure 4.37. Hence, it is very important to control ground voltage fluctuations by minimizing the impedance of the grounding system. These issues become more crucial as the clock frequency is increased.

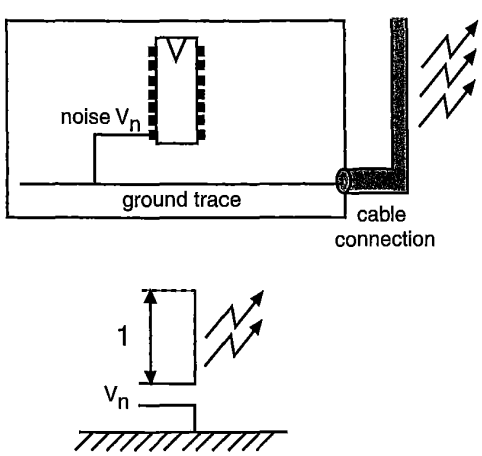


FIGURE 4.37 Dipole antenna model.

EXAMPLE 4.10

Figure 4.38 shows the time-domain voltage and current waveforms assuming capacitive load for the case f_{clock} = 100 MHz, T_{clock} = 10 ns, and t_r = 500 ps. When the signal rise time is longer than the propagation time along the transmission line, the assumption is valid. The figure also shows the frequency spectrums of the voltage and current waveforms. Due to the symmetry of the time-domain waveforms, the frequency spectrum exhibits odd number harmonics of the fundamental clock frequency.

Figure 4.39 shows the current spectrums of two different clock frequencies. One is for the case of f_{clock} = 100 MHz, T_{clock} = 10 ns, and t_r = 500 ps. The other case is for f_{clock} = 400 MHz, T_{clock} = 2.5 ns, and t_r = 125 ps. As can be seen from the figure, the current waveform for higher clock frequency and fast transition time contains more high-frequency elements. Since the radiation efficiency of the antenna increases at higher frequencies, the current waveform at higher clock frequencies and fast rise time has the potential to generate greater emissions. The figure also illustrates the radiated emission measured at 3 m distance from the radiator. The antenna structure is modeled as a dipole antenna structure, radiating common-mode radiation. When the cable shield is connected to the noise ground, the radiation pattern and behavior follows the dipole antenna model.

4.7 DESIGN PROCESS

The previous sections reviewed the fundamental issues associated with the electrical design of packages. The package design engineer often uses a methodology and a process for designing packages as shown in Figure 4.40, which is the subject of this section.

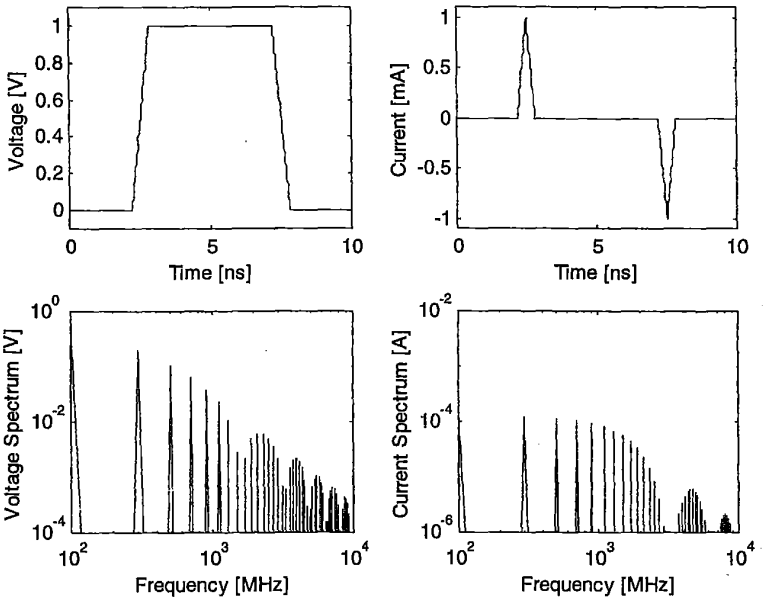


FIGURE 4.38 Time-domain waveform and frequency-domain spectrum.

Samsung, EX1028, p. 136

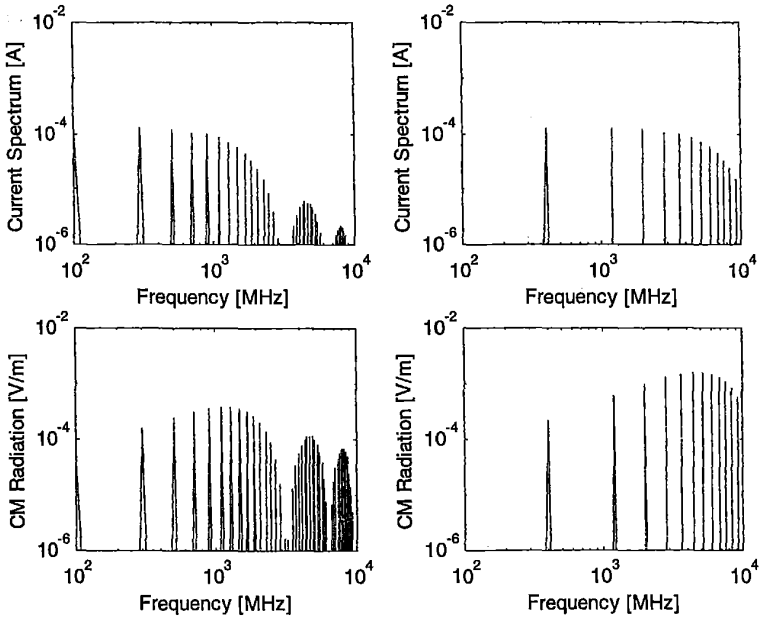


FIGURE 4.39 Current spectrum.

The electrical design procedure of the package starts with the determination of the electrical specifications for the package. Basically, these specifications are based on the system and chip performance, and are obtained from discussions with the system and chip designers. These specifications could include a bubble diagram showing the connectivity between chips, logic level, output impedance of the circuits, desired speed of communication, sensitivity of the circuits and the power consumption of the chips. Using these specifications, the electrical designer can extract information pertinent to the package such as single-cycle nets, multicycle nets, single-drop nets, multidrop nets, maximum delay, noise budget and power supply current which become the parameters that the package needs to support.

The package designer has access to the various technologies that can be used. Each technology is supported by a set of ground rules and material properties such as minimum line width (w), spacing (s), thickness (t), dielectric thickness (d), dielectric constant (ε_r) and conductivity (σ). Using these physical parameters, the electrical parameters for the interconnections such as resistance (R), inductance (L), capacitance (C) and conductance (G) are extracted using electromagnetic tools. Examples of such tools are Maxwell (Ansoft), EM (Sonnet) and HFSS (Agilent). These tools are also used to extract the parasitics in the package. These parameters are then used to construct circuit models of the interconnections for simulation in a circuit simulator such as SPICE.

SPICE is used for two types of simulation: signal distribution and power distribution. The simulation of the interconnections provides details on the reflections, crosstalk and signal attenuation. The DC drop and power supply noise are obtained from the power

Samsung, EX1028, p. 137

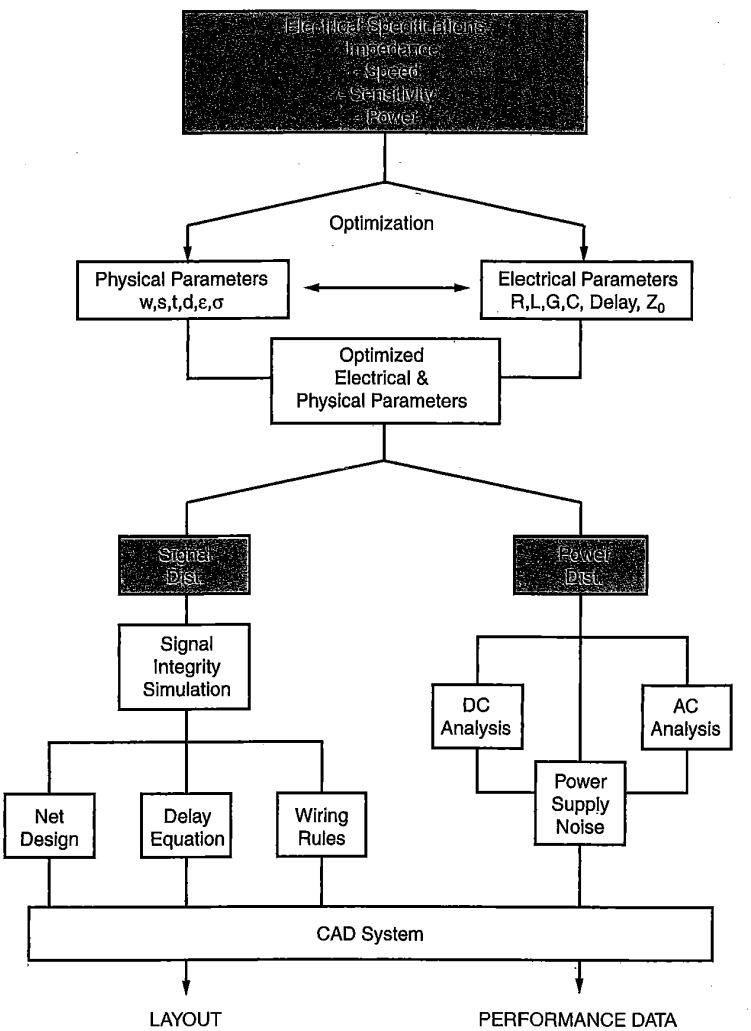


FIGURE 4.40 Design process.

distribution simulation. These simulations are initially done separately to define and optimize the package cross section in terms of the impedance and noise budget.

This is followed by an extensive simulation that includes the driver and receiver circuits powered by the package power distribution and driving package interconnections. These simulations include statistical variations in the process and operating temperature. For high performance systems, this often includes a simulation of every net in the package. These simulations result in the generation of wiring rules, delay equations and noise budgets. The wiring rules place constraints on the layer to be used for the net and the

Samsung, EX1028, p. 138

tolerance for the net length. This information is fed to a *computer-aided design* (CAD) system for the physical layout of the package.

The final output of the electrical package design process is a physical layout of each layer in the multilayer package; the dimensions of signal lines, stack-up of the power/ground planes, material parameters and dimensions of the bonding structures.

4.8 SUMMARY AND FUTURE TRENDS

Electrical design performs two critical functions: signal distribution and power distribution. Signal distribution is the process of providing electrical communication paths between circuits on chips. However, this would not be possible unless the circuits are powered. The process of supplying voltage and current to the chips is called *power distribution.*

The quality of signal communication is a function of the parasitics of the interconnections. In the early 1980s, the interconnections were primarily capacitive in nature due to the speed of the products that were designed at that time. The most important effect related to the quality of signal transmission was the RC time constant that slowed the risetime of the signal. Along the communication path, the resistance was dominated by the device, while the major contributor to the capacitance was the interconnection. The large capacitance was due to the wide line widths, long line lengths and high dielectric insulating materials that were used at that time. However, in the 1990s, the speed of the products increased dramatically. This was made possible due to the scaling of the transistor, which resulted in small device delays. With increasing integration levels, the interconnections connecting the circuits became shorter. The combination of transistor scaling and increased integration resulted in the RC time constant of the communication path becoming insignificant. Now, a new phenomenon took over: the finite velocity of the electrical signal propagating in the insulating medium. The signal started behaving more like an electromagnetic wave and the finite delay incurred by the signal due to the properties of the medium became a fraction of the clock cycle. The interconnections therefore had to be treated as transmission lines.

As the clock frequency increased, low dielectric insulating materials became very important since they reduced the delay of the signal. To maintain signal quality, reflections had to be suppressed through various termination schemes. As the density of lines increased to accommodate higher integration, the line spacing decreased, producing crosstalk on the lines. Controlling crosstalk became critical since it affected the quality and timing of the signal. However, though the interconnections behaved as transmission lines, they were low loss structures and hence the transit time delay of the signal was the only critical parameter. Interconnections could therefore be treated as lossless transmission lines. At the present time, the resistive nature of the interconnections are becoming very important. This is due to the demand for faster products with higher levels of integration leading to narrow lines and spaces and very low dielectric constant materials. For these lines, as the frequency increases, the resistance of the interconnections becomes frequency dependent. This is caused by the skin effect (the concentration of current on the surface of the conductor). In addition, the dielectric loss of the insulating material is becoming very important as clock rates increase beyond 1 GHz. Both these loss mechanisms cause the rounding and slowdown of a fast edge, which is a major problem for digital logic.

Samsung, EX1028, p. 139

Hence, in the future, all interconnections will behave as lossy transmission lines, and maintaining the quality and timing of the signal will become a major problem. New methods for signal distribution are therefore necessary in the future.

Supplying clean power to the chips will be the major bottleneck in the future. In the 1980s, the era of the DIP and QFP packages integrated on a board, chips dissipating a few watts of power needed to be supported. The clock frequency was small and hence, a large time window was available. Power supply inductance in the 10–50 nH range was sufficient to ensure the circuits did not generate excessive noise in the system. During the 1990s, it was clear that the power supply inductance had to be reduced. The current through the package increased dramatically due to chips dissipating larger amounts of power. In addition, the clock frequency increased, providing a smaller time window for transistor operation. The DIP and QFP packages that supplied current through leadframes and wirebonds could no longer support these products due to large inductance. This led to ball grid array (BGA) packages with area array connections for the chip and the package. BGA packages enabled a vertical path for the current, thus reducing the inductance. Planes in the package and board became necessary to support a low inductance path. In addition, the chip, package and board had to be populated with large amounts of decoupling capacitors to supply transient power to the chips.

At the beginning of the New Millenium, the power supply inductance has been minimized to ~10 pH for high-performance products using state-of-the-art packages and boards. Packages and boards have become thinner and contain a multitude of planes to support large current transients in the system. With such small inductance values, the power supply inductance will no longer be a major bottleneck for supplying clean power to the chips. However, the noise in the system will be dominated by bouncing planes, a phenomena occurring due to the propagation of electromagnetic waves between power and ground planes in the package and board. To avoid noise coupling caused by the bouncing planes, the power distribution for the core and I/O need to be separated. This requires new and efficient ways of supplying power to the chips. Decoupling at frequencies greater than 1 GHz will become a major problem which require sophisticated methods such as integral decoupling on the package and board. A large fraction of the real estate on the chip will be dedicated for decoupling, which could be a major problem.

At higher frequencies, digital logic is becoming more and more analog in nature. Computer engineers are beginning to appreciate the importance of electromagnetic theory and microwave theory in the design of their products. Electrical engineers are therefore beginning to play a critical role in the design of computers.

4.9 HOMEWORK PROBLEMS

The authors of this chapter encourage students to use SPICE for solving some of the problems. This can be used to confirm analytical solutions for simple networks or for obtaining waveforms where analytical solutions are difficult to derive. SPICE is a circuit simulator, some versions of which are available in the public domain. An educational version of Micro-Cap, a SPICE simulator with schematic entry capability, can be downloaded from http://www.spectrum-soft.com.

1. In a given interconnection, $R_{on} = 40\ \Omega$, $C_g = 3$ pF, and it is determined that the interconnect contributes an additional capacitance of 3 pF. Determine the delay $T_{50\%}$.

Samsung, EX1028, p. 140

2. Construct an RC circuit to represent an interconnection with $R = 100\ \Omega$. The capacitance of the capacitor varies from 1 pF to 5 pF in steps of 1 pF. Using a pulse as the input, calculate the 50% delay produced by the circuit. The response is equivalent to that of a CMOS inverter charging a capacitive load.
3. In the chapter, lumped and distributed RC networks were discussed for representing interconnections. Consider an interconnection with $R = 1000$ ohms and $C = 5$ pF. Simulate the interconnection using SPICE in the following ways:
   a) As a lumped RC network
   b) As a distributed network with 5 segments
   c) As a distributed network with 10 segments
   Is the 50% delay the same in all the cases? Briefly discuss your results. Use a pulse source with a logic swing of 5 V, rise and fall time of 100 ps and pulse width of 50 ns.
4. A trace with length 4 cm is routed through a polyimide substrate.
   a) Determine the transit time delay associated with this interconnect.
   b) If the system in question has parameters $R_{on} = 50\ \Omega$ and $C_g = 2$ pF, does the capacitive gate delay $T_{50\%}$ dominate the delay time, or does the transit time delay?
   c) If the trace has width 40 μm and is separated from a solid ground plane by 10 μm, determine its capacitance per unit length from Equation (4.25).
   d) If the total capacitance of the interconnect (obtained from the capacitance per unit length) was used in a traditional delay equation, such as Equation (4.8), is the answer essentially the same as found in part (b)?
5. Consider the system shown in Figure 4.13a and 4.13b, but suppose that the characteristic impedance of the interconnect is changed to 35 Ω. The driver is switched to a high state at $t = 0$, as modeled in Figure 4.13b.
   a) Determine the voltage across the input to the receive gate during the time interval $0 < t < 8T$.
   b) In an attempt to match the system, a parallel matching resistor equal to 50 Ω is placed across the receive end of the 35 Ω line. Determine the voltage across the input to the receiver in the presence of the matching resistor for $0 < t < 4T$.
6. Repeat problem (5), but use a characteristic impedance of 70 Ω.
7. In the system depicted in Figure 4.23, suppose that the coupling coefficients are changed to $K_F = +0.004$ and $K_R = 0.07$. Determine the forward and reverse crosstalk (V_D and V_C) for $0 < t < 40$ ns.
8. Table 4.2 lists the dielectric constant of insulator materials. Assuming a signal with risetime of 200 ps has to propagate on the package interconnection, estimate the minimum length for which the transmission line properties of the interconnection become important for each of these insulator materials. Tabulate the results (include the material name, dielectric constant and the minimum length).
9. Consider a driver driving a package interconnection with inductance $L = 3.75$ nH/cm, capacitance $C = 1.5$ pF and length $l = 100$ cm. The signal propagating on the interconnection has a risetime of 100 ps.
   a) Does the interconnection behave as a transmission line? Give reasons.
   b) Calculate the impedance and delay of the interconnection.
   c) Assuming the driver has a resistance of 1000 ohms, plot the waveform at the far end of the interconnection using SPICE. Use a pulse source with a logic swing of 5 V, rise and fall time of 100 ps and pulse width of 300 ns. The duration for the transient analysis is 300 ns. Use a transmission line model for the interconnection. Calculate the 50% delay and comment on the waveform. What is the problem with this circuit?
   d) Redo c) with a driver resistance of 10 ohms. Calculate the 50% delay and comment on the waveform. What is the problem with this circuit?

Samsung, EX1028, p. 141

e) Assuming the driver resistance is 10 ohms and the far end of the interconnection is terminated with a 50 ohm impedance, plot the waveform at the far end using SPICE. Calculate the power dissipated by the 50 ohm resistor. (You may want to plot the voltage and current for the 50 ohm resistor and use this to calculate power.)

f) Assuming the driver resistance is 10 ohms and a series resistance of 40 ohms is used at the near end of the interconnection, plot the waveform at the far end using SPICE. Calculate the power dissipated by the 40 ohm resistor. (You may want to plot the voltage and current for the 40 ohm resistor and use this to calculate power.)

g) Compare the two termination schemes used in e) and f).

10. Two chips are communicating with each other through a 32 bit wide bus in the package. The I/O drivers have a voltage transition of 2.5 V in 250 ps. The inductance in the package supplying power to the drivers has an inductance of 50 pH. As a designer, you're asked to decide on the characteristic impedance of the interconnections in the bus. The allowed range is Z_0 = 25 ohms–100 ohms, in steps of 5 ohms. Plot the noise voltage on the power supply as a function of the characteristic impedance Z_0 and provide reasons why you would choose a certain impedance.

11. A circuit with a supply voltage of 2.5 V is being powered through a package. The package has a resistance of 100 mΩ and a certain inductance L. The circuit requires at least 2.25 V to function correctly. If this voltage is required in a time of 1ns, calculate the effective inductance L for the package.

12. Table 4.4 lists the inductance expressions for three different package structures that are used to supply power to the circuits.

a) Consider a leadframe package. As an approximation, the leads can be assumed to be wires. The wires have a diameter of 25 μm, length of 1 cm and the separation between the wires is 100 μm. Calculate the effective inductance for a pair of wires: i) if adjacent wires are used to supply voltage and ground; ii) if the voltage and ground wires have a signal wire between them; and iii) if the voltage and ground wires have two signal wires between them. Assume no mutual inductance between the wires.

b) Consider next a leadframe package where a ground plane is used. The leads are still used to supply the voltage while the plane is used to supply ground. The wires have a diameter of 25 μm and length of 1 cm. Calculate the separation between the wires and the ground plane to obtain an effective inductance that is similar in value to the adjacent pair of wires in part a). For this separation, what is the advantage of using a ground plane as compared to part a)? Ignore the mutual inductance between the wires.

13. A decoupling capacitor has an ESR of 10 mΩ, ESL of 60 pH and capacitance of 32 nF.

a) Calculate the self resonant frequency of the capacitor.

b) Plot the frequency response of the capacitor (impedance versus frequency) and indicate the frequency range where it behaves like a capacitor.

c) Plot the frequency response for a pair of capacitors in parallel. How does it differ from the frequency response of a single capacitor?

14. A packaging technology has the following ground rules:
Line width = 100 μm, line thickness = 25 μm and dielectric thickness = 150 μm from the nearest ground plane. The dielectric constant of the insulator is 4.0. Calculate the impedance, propagation velocity and propagation delay if these ground rules are used to construct a microstrip line, an embedded microstrip line and a strip line.

4.10 SUGGESTED READING

1. H. B. Bakoglu. *Circuits, Interconnections and Packaging of VLSI.* Addison Wesley, 1985.
2. W. D. Brown, ed. *Advanced Electronic Packaging—With Emphasis on Multichip Modules,* Chapter 3 & 4. IEEE Press, 1999.

Samsung, EX1028, p. 142

3. H. W. Johnson and M. Graham. *High Speed Digital Design: A Handbook of Black Magic.* Prentice Hall, 1993.
4. David M. Pozar. *Microwave Engineering.* Addison-Wesley, 1990.
5. S. Rosenstark. *Transmission Lines in Computer Engineering.* McGraw-Hill, 1994.
6. R. Senthinathan and J. L. Prince. "Simultaneous Switching Ground Noise Calculation for Packaged CMOS Devices." *IEEE Journal on Solid-State Circuits.* Nov. 1991, 1724–1728.
7. R. Tummala, E. Rymaszewski, and A. Klopfenstein, eds. *Microelectronics Packaging Handbook,* Chapter 3. Chapman & Hall, 1997.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contribution of Dr. George Katopis of IBM for critically reviewing this chapter.

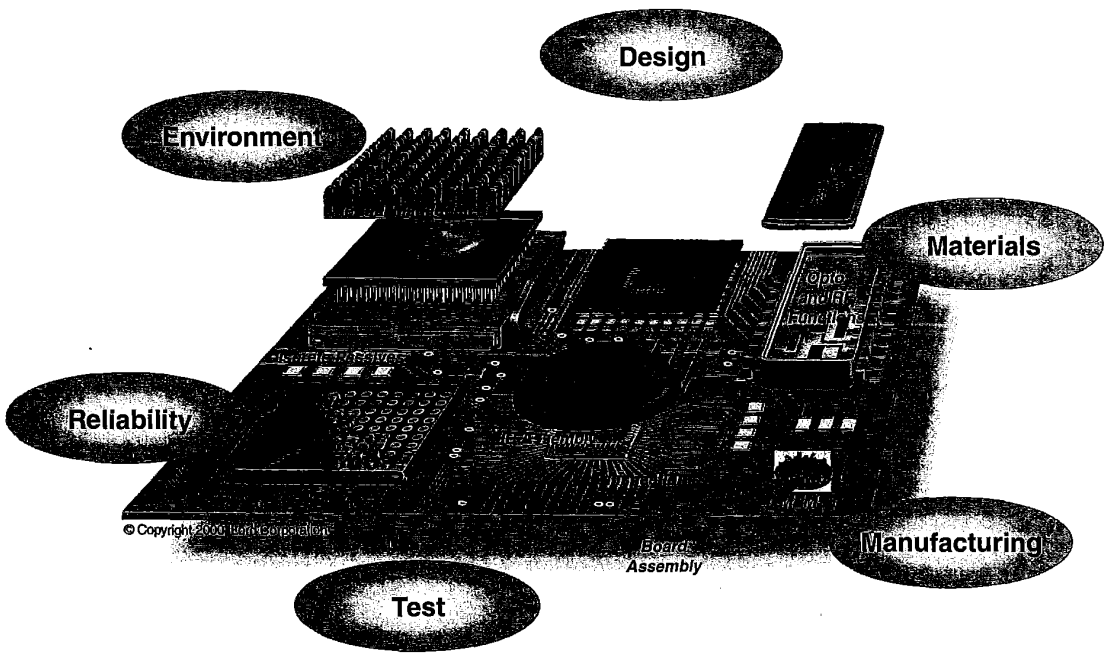
Samsung, EX1028, p. 143

CHAPTER 8

FUNDAMENTALS OF MULTICHIP PACKAGING

Prof. Harry K. Charles Jr.
Johns Hopkins University

Prof. Rao R. Tummala
Georgia Institute of Technology



8.1 What Are Multichip Modules?
8.2 Multichip Module Functionality
8.3 Multichip Module Advantages
8.4 Multichip Modules at the System Level
8.5 Types of Multichip Module Substrates
8.6 Multichip Module Design
8.7 Multichip Module Technology Comparisons
8.8 Alternatives to Multichip Modules
8.9 Summary and Future Trends
8.10 Homework Problems
8.11 Suggested Reading

CHAPTER OBJECTIVES

- Define multichip modules and basic application areas
- Describe multichip module types and construction
- Present elements of multichip module design
- Develop tradeoffs between multichip module types and alternative packaging methods

CHAPTER INTRODUCTION

Multichip modules (MCMs) have become a major thrust in electronic packaging technology. Because of the high performance and density of the integrated circuits packaged in multichip modules, stringent new demands are being placed on package structures, electrical interconnects, and sealing (encapsulation) materials. This chapter presents details of MCM technology, including different types, basic design rules, fundamental applications, and packaging alternatives to the MCM. Thin-film MCMs offer great promise for extremely high-performance (density) packaging at very low cost.

Samsung, EX1028, p. 145

8.1 WHAT ARE MULTICHIP MODULES?

A *multichip module* (MCM) or package is defined as a single unit ("package") containing two or more chips and an interconnection substrate which function together as a system building block. The MCM, or *multichip package* (MCP), provides signal interconnect and input/output (I/O) management for the chips contained within the module, thermal management (typically cooling), mechanical support, and environmental protection. MCMs or MCPs are an important technique for the packaging of integrated circuits (ICs) and associated subsystem and system elements. The MCM was developed in the early 1980s and is an evolution of the traditional hybrid package that appeared soon after the birth of the IC. Today's high-performance supercomputers and mainframes require many, many chips to form one central processing unit (CPU), in contrast to today's personal computers (PCs) that typically use one chip for each processor. MCM technology allows chips to be spaced more closely with less overall volume and weight than individually (single chip packaging) packaged parts. Since MCM packages are very dense with minimum volume and weight, they have widespread application in aerospace, medical, consumer, and all portable products in addition to supercomputers and mainframes. Figure 8.1 illustrates an example of an early 100-chip MCM used by IBM mainframe computers (circa 1980s), compared to a single chip PC processor from Intel.

Clearly, the MCM, by its very nature—multiple chips on a substrate or carrier—is different from its individually packaged counterpart. But what sets it apart from its precursor, the hybrid? It is the chip density. It is commonly held that for a package containing multiple chips to be classified as an MCM, 50% or greater of the substrate (carrier) or packaging area must be covered with active semiconductor devices. This can be represented by the simple formula:

$$A_C = \geq 0.5 A_S \tag{8.1}$$

where A_C equals the area of the semiconductor (or chip) and A_S equals the area of the substrate (package or carrier).

The substrate or carrier is the key element in any MCP and especially in the MCM. The substrate, as shown in Figure 8.2, provides the mechanical attachment for the chips,

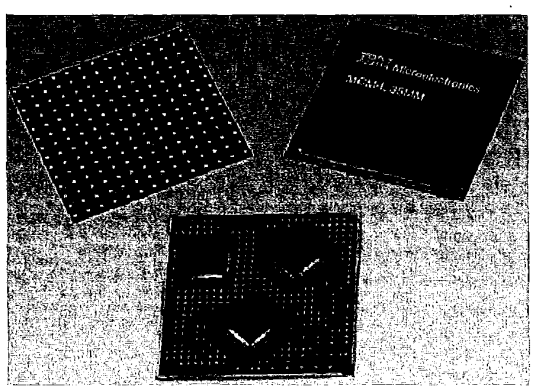


FIGURE 8.1 Examples of a multichip module central processing unit (left) and an individually packaged microprocessor (right) such as those used in a personal computer.

Samsung, EX1028, p. 146

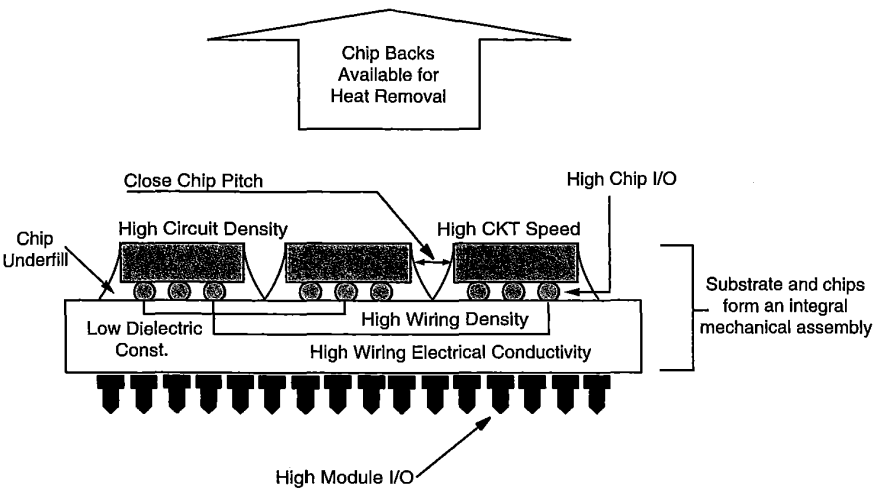


FIGURE 8.2 A schematic drawing of a multichip module, illustrating its major components and their role in supporting the four major principles of electronic packaging: signal I/O, thermal management, mechanical support, and environmental protection.

handles the interchip signals, provides power and ground for all chips, and interfaces the module with the next level system elements (see Packaging Hierarchy, Chapter 2). In addition, the substrate can play a major role in module thermal management, as well as help protect the chips from environmental impact. MCM substrates come in three main forms: 1) multilayer, thin-film structures on semiconductor or ceramic base layers, with deposited metal conductors and *dielectrics* (MCM-Ds); 2) either thick-film or cofired multilayer *ceramic* technology (MCM-Cs); or 3) organic *laminate* multilayer board structures (MCM-Ls), similar to standard *printed wiring boards* (PWBs), but with denser component placement and greater wiring densities. Details of these substrate types will be presented later in this chapter.

8.2 MULTICHIP MODULE FUNCTIONALITY

As mentioned above and schematically illustrated in Figure 8.2, the MCM provides the circuit designer and packaging engineer with a highly functional building block for the implementation of complex circuits, subsystems, and even small systems in certain applications. To be a highly functional MCM, the following criteria must be satisfied:

1. Chip-to-chip spacing in the MCM must be held to a minimum to minimize signal propagation delays. Close chip spacing requires the substrate to have high-density wiring to support the I/O requirements of each chip without an undue size or layer penalty. Because of their small size, conductors must be made of high conductivity metals to avoid excessive lead resistance. In addition, the traces must be properly designed to prevent signal distortion and minimize crosstalk and noise.

Samsung, EX1028, p. 147

2. The MCM must provide a means of thermal management to limit the junction temperature of the semiconductor chips to temperatures less than 85–100°C. Thermal management is typically accomplished through substrate conduction of the heat when chips are mounted to the substrate in the face-up configuration. If chips are flip-chipped—mounted in a face-down configuration—as shown in Figure 8.2, then thermal management can either be accomplished through the leads (solder balls) in low power applications or by the use of thermally conductive fingers for heat removal in high power applications.
3. The MCM must provide reliable I/O connections to the next level of assembly in the system. Typically, this takes the form of an interface with a PWB, although direct module interface with external system elements is possible. The number of module I/O is significantly reduced from the sum of the module's individual die or chip I/O numbers. If the MCM doesn't reduce overall I/O, then the partitioning of the circuit, or the overall requirement for the MCM, must be questioned. A method for calculating module I/O based on the constituent chips and their respective I/O is called *Rent's Rule* and will be discussed later in the chapter.
4. The MCM must provide protection from the environment. Some technique for sealing the packages, or at least overcoating/encapsulation of the chips and their interconnects, must be provided.

8.3 MULTICHIP MODULE ADVANTAGES

The need for the MCM arises because the circuit designer cannot obtain enough functionality in a single chip. As the march of integrated technology continues rapidly, it ultimately may be possible to have the entire system function on a large single chip or wafer. This *wafer scale integration* (WSI) holds the prospect of going back to single chip packaging techniques, but until it is a practical reality, multichip packaging will be required. In fact, MCP will always be used to address special circuit needs and the requirement for system redundancy. Thus, in today's electronic packaging world, there appear to be three major options for system-level packaging, as shown in Figure 8.3:

1. WSI with the system on a single chip (*system-on-chip* [SOC])
2. MCMs formed from multiple chips on a common substrate/package structure
3. Individually packaged chips on a PWB

It should also be noted that if all the functionality that could be packaged in an MCM was integrated into a single chip, the chip would become very large and the device feature size extremely small. This would result in very complex technical requirements and, hence, high development and manufacturing costs. Since the MCMs use less complex chips and, for the most part, can be repaired, they can be less expensive, and easier and faster to develop. Thus, until WSI (SOC) can be performed with high yield and relatively low development costs, the MCM will be the high performance (speed, volume, weight) mainstay for system packaging.

MCMs have four major advantages over individually packaged parts: 1) higher packaging efficiency (area of chip to area of the board), due to the fact that the bare chips can be placed much closer together than the chips in single chip packages (because of

Samsung, EX1028, p. 148

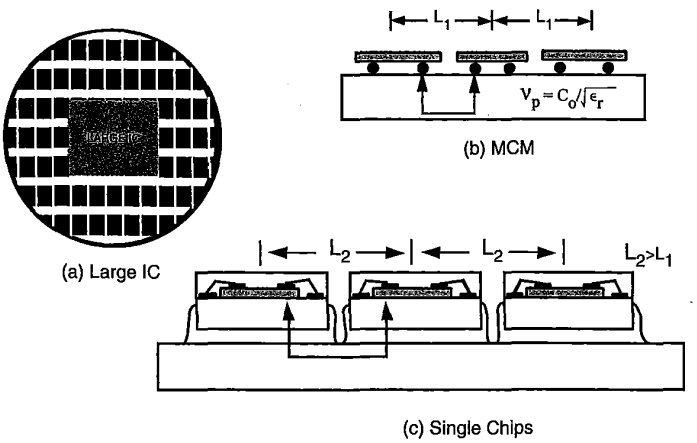


FIGURE 8.3 Schematic representation of system packaging methods: (a) system-on-chip or wafer scale integration; (b) multichip module; or (c) single chip packaging. v_p = velocity of propagation, C_o = speed of light and ϵ_r = relative dielectric constant.

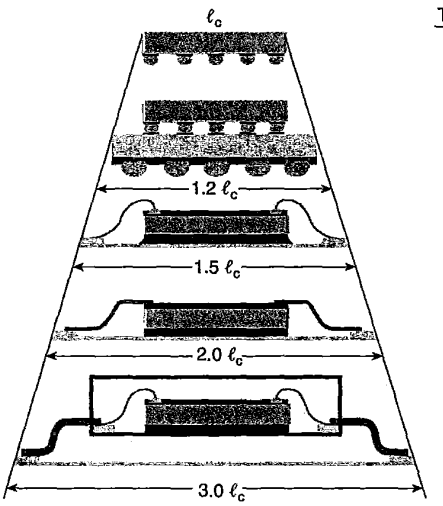
the wasted space of the packages themselves); 2) better electrical performance due to the shorter distances between chips and, hence, reduced substrate wiring length as well; 3) greater reliability due to reduced number of interconnects between the chip and the board (the MCM eliminates one level of interconnect, see Figure 8.3); and 4) the MCM (in high volume) offers the potential for lower cost because of the elimination of the individual IC packages and the reduced substrate size.

8.3.1 Packaging Efficiency

As mentioned above, one way of distinguishing an MCM from other single and MCP styles is to consider its *packaging efficiency.* Packaging efficiency is defined as the ratio of the area of all the base chips to the area of the MCM substrate or, for individually packages parts, the area of the system-level board. Single chip packages such as *dual-in-lines* (DIPs) or *quad flat packages* (QFPs) require a lot of board real estate due to the packaging size itself and the need to have peripheral leads. Figure 8.4 illustrates the packaging efficiency associated with each packaging style, while Figure 8.5 illustrates overall module packaging efficiency. The board-level efficiency with such packages (DIPs, QFPs) is quite low, typically around 10%. Even with standard ball grid arrays and chip scale packages, the packaging efficiency rarely approaches 50%.

MCMs with bare chips directly bonded to the substrate can be much more efficient, as illustrated in Figure 8.5, with efficiencies approaching 80%. The ultimate packaging efficiency occurs when the entire system (MCM) is integrated into a large, single chip (SOC or WSI) which can support all circuit functions—digital, analog, *radio frequency* (RF), and microwave. The next densest packaging occurs when the MCM integrates all chip types and required interconnect functionality into a single module. This integrated package structure has been called *system-on-package* (SOP) by some authors. In applications where the MCM incorporates more than digital functions, such as a cellular

Samsung, EX1028, p. 149



Technology	Area ratio: package/chip
Flip chip	1:1
Chip-scale package	1.5:1
Chip-on-board	2.25:1
Tape automated bonding	4.0:1
Quad flat package	9.0:1

FIGURE 8.4 Size comparisons of flip chip (direct chip attach technology) with alternate packaging styles (ℓ_c = length of chip side). For small chips, the package-to-chip area ratios may be even larger for chip-on-board, tape automated bonding, and quad flat packages.

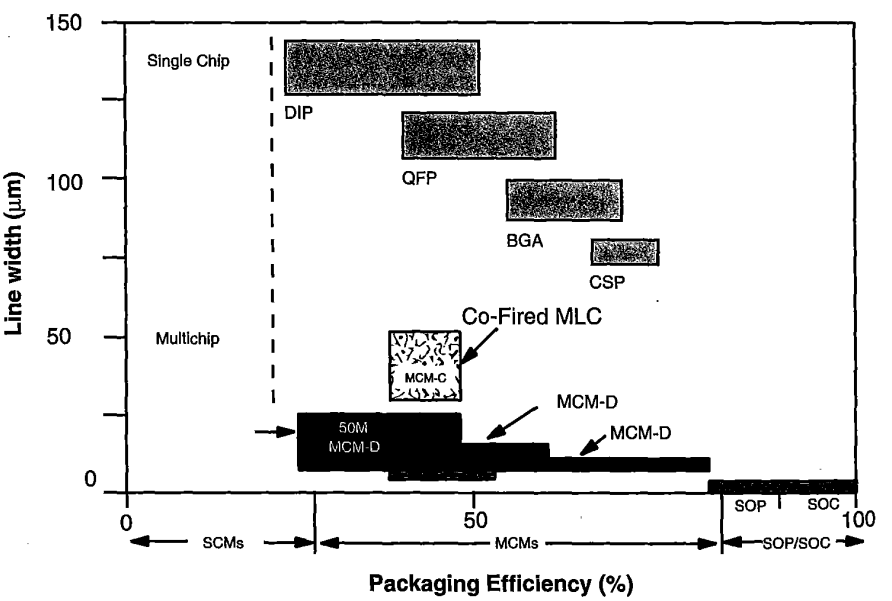


FIGURE 8.5 Line width versus packaging efficiency for various packaging styles. Single chip (DIP = dual-in-line, QFP = quad flat package, BGA = ball grid array, CSP = chip scale package, SCM = single chip module). Multichip (MCM-L = laminate, MCM-C = ceramic, MCM-D = thin-film deposited, SOP = system-on-package, SOC = system-on-chip, MCM = multichip module, MLC = multilayer ceramic).

Samsung, EX1028, p. 150

telephone, the MCM can contain large numbers of passive devices (R, L, C), microwave chips, switches, relays, and connectors, in addition to standard digital chips.

A cellular phone MCM can contain over 500 components that will occupy approximately 25 cm^2 of area at today's component density of 20 components/cm^2. If most of these non-standard IC components, such as the passive devices and the specialized RF and microwave devices, could be embedded inside the board as thin- or thick-film layers, thus reserving the top (and bottom) surfaces for the large, regularly sized ICs, then the packaging efficiency could approach that of SOC or WSI. As mentioned above, such an integrated packaging structure would be called SOP.

As shown in Figure 8.5, the packaging efficiency of MCMs varies widely, with laminate and ceramic MCMs providing the lowest and thin-film multilayers providing the highest. This packaging efficiency is directly proportional to component I/O and the wiring density supported by the substrate. Thin-film MCMs possess the highest wiring density, measured as length of minimum width track possible per unit area, and typically provide the highest packaging efficiency because they can interconnect the ICs in the smallest area. Small-area, high-density packages such as MCM-Ds offer significant benefits in addition to size, including better electrical performance due to shorter wiring length, and improved reliability due to elimination of one level of interconnect (see Figure 8.4).

8.3.2 Electrical Performance

Fundamental MCM performance can be measured by *functional throughput rate* (FTR). FTR is defined as the product of the equivalent number of gates per module times the maximum clock rate of such gates (f_{max}). The maximum clock rate is usually defined to be equal to 0.25 t_D, where t_D is the delay associated with a typical gate. Today's high-density MCMs can have FTRs in the range of 10^8 to 10^{10} Gate-MHz. The processing rate of any complex system (MCM) is heavily influenced by architectural decisions and techniques, like parallel processing. Another measure of performance is a throughput parameter related to the cycle time, the gate delay, and the number of cycles required to execute a given instruction. This throughput parameter is most often measured in the number of *millions of instructions per second* (MIPS) and is given by:

$$\text{Number of MIPS} = 10^3/[(\text{cycle time}) \times (\text{cycles per instructions})] \qquad (8.2)$$

The larger the number of MIPS, the shorter the cycle time and the fewer the number of cycles per instruction must be. Figure 8.6 illustrates the influences of both system architecture and MCM packaging on the number of MIPS. For computers that perform scientific and engineering computations, MIPS is not the most useful measure of performance. This is because scientific and engineering computers are vector machines. A *vector* is a collection of N variables, and a single instruction in a vector machine can initiate thousands of floating-point arithmetic operations. A more appropriate measure is then *millions of floating-point logic operations per second* (MFLOPS) or *megaflops*. Faster machines perform at the *billions of floating-point logic operations per second* (GFLOPS) or *gigaflops*. State-of-the-art machines now approach the *trillions of floating-point logic operations per second* or the *teraflops* regime.

Other measures of electrical performance include clock speed, operation frequency (RF and microwave modules), power dissipation (power time delay products), etc.

Samsung, EX1028, p. 151

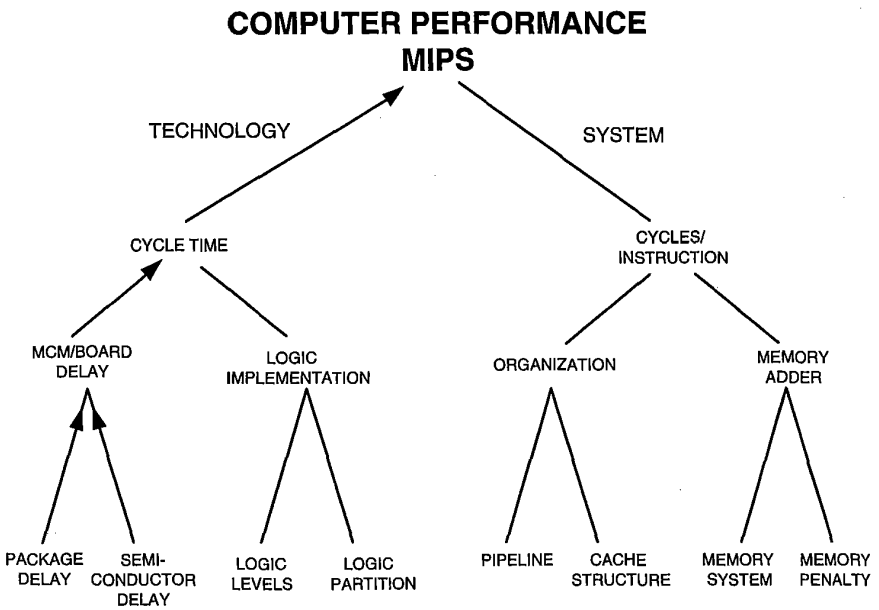


FIGURE 8.6 Packaging and computer architectural influences on the performance of a computer system in MIPS.

8.3.3 Reliability

Reliability is an essential element to all MCM design, fabrication, and testing processes. Figure 8.7 illustrates the yield of an MCM as a function of the number and quality of the ICs used in the module. Even with high-quality chips, or those with a high probability of *known good die* (KGD) greater than 95%, module yield can be extremely low for modules containing a large number of chips. The curves in Figure 8.7 assume that there are no defects or errors in the substrate and the assembly processes. The inclusion of such factors would further detract from the yield. Thus, repair appears to be critical in large modules containing complex, expensive chips. Figure 8.8 illustrates the improvement in yield as a function of the KGD probability and the amount of repair for a 25-chip MCM.

Inherently, the MCM packaging operation must be more reliable than the individual ICs that it uses. Such packaging reliability must be designed and built into the product. To ensure this reliability, three different areas must be addressed:

1. Design for reliability with a minimum number of connections.
2. Construct the module using six-sigma manufacturing processes as described in Chapter 22.
3. Perform accelerated and other screening tests on the MCM to remove defect-induced failures before the product is shipped to the customer.

The reliability indicators for IBM's *thermal conduction module* (TCM), the industry's first MCM, are illustrated in Figure 8.9. In this diagram, a comparison between single

Samsung, EX1028, p. 152

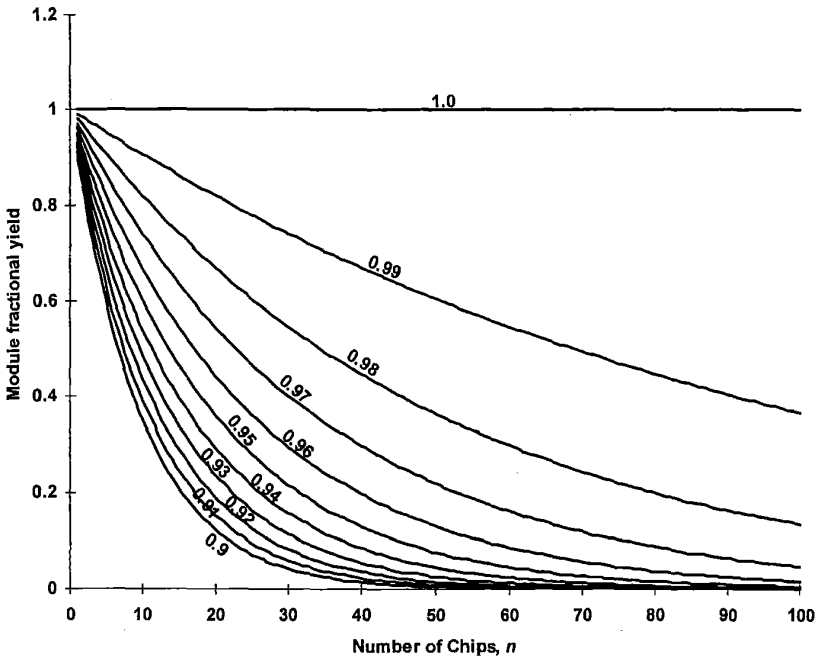


FIGURE 8.7 MCM fractional yield as a function of the number of chips in the module for various chip known good die (KGD) probabilities, p_c (p_c ranges from 0.9 to 0.99).

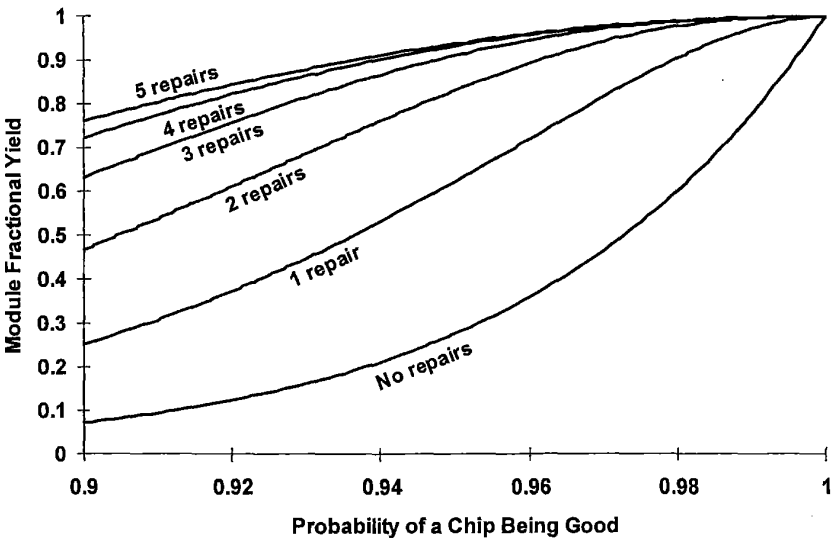


FIGURE 8.8 Multichip module fractional yield as a function of the KGD probability for a 25-chip MCM with and without repairs (the number of repairs ranges from 0 to 5).

Samsung, EX1028, p. 153

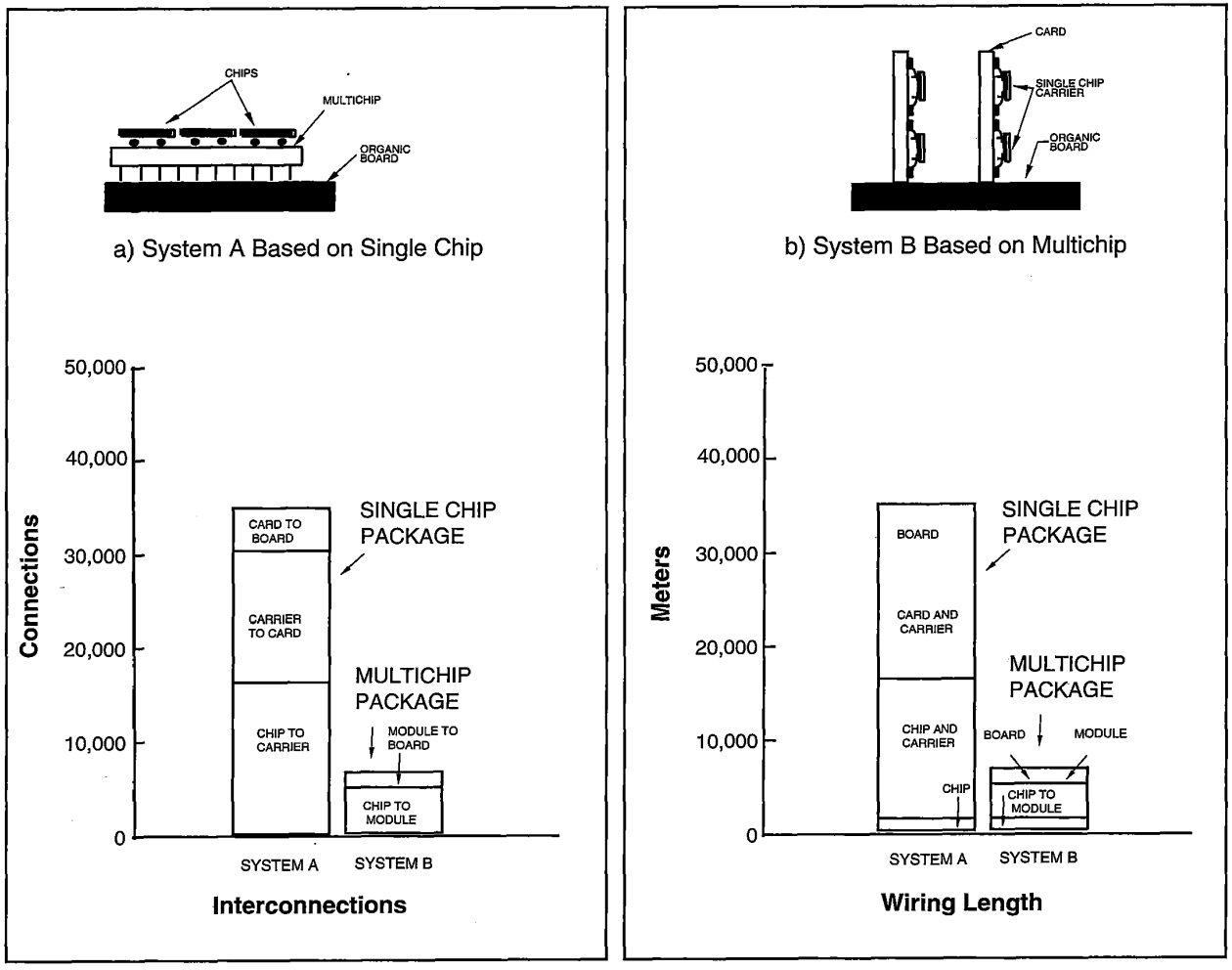


FIGURE 8.9 The improvement in wire length and number of connections for IBM's thermal conduction module (an MCM) versus the same system implemented in single chip packages.

MEMS23165

Samsung, EX1028, p. 154

chip packaging and the multichip TCM approach is presented. System A, the individual chip packaging approach, required 50,000 total connections between the ICs and the packages and between the packages and the boards. System B, the multichip approach, reduced the number of connections by a factor of almost 10, and the total wiring length was also reduced by a similar factor, thus increasing the system reliability by about a factor of 10.

8.3.4 Cost

The cost of MCMs at the system level, in theory at least, is expected to be lower than the alternative single chip package implementation. This is due to reducing or minimizing the number of interconnects and minimizing the actual substrate area and system volume. MCM substrates are typically fabricated using the finest feature (line widths, vias, and spacings) processing technology, which can significantly reduce board area and, because chips are packed in a very dense manner without packages, system volume is minimized. In quantity, the cost of MCM substrates should be comparable to other substrates supporting single chip packaging because, even though the cost per unit area is higher, the overall board size (area) required is much, much smaller. One of the reasons that MCM per unit area costs are high is that the high-density MCM substrates are typically batch processed at or near final size or small area processing. Large area processing can help lower MCM substrate costs.

8.4 MULTICHIP MODULES AT THE SYSTEM LEVEL

Since the typical MCM is a subsystem or system in most applications, it requires the packaging engineer to be familiar with the following:

1. Multilevels of the Packaging Hierarchy
2. Electrical Design (Chapter 4)
3. Chip-level Interconnect (Chapter 9)
4. Module or Package-level Interconnect
5. Heat Removal (Chapter 6)
6. MCM Substrate Technologies (this chapter)
7. Electrical Testing (Chapter 19)
8. MCM Wiring and Connectivity (this chapter)
9. Repair, Rework, and Change Integration (this chapter)
10. Module Sealing and Encapsulation (Chapter 15)

Figure 8.10 illustrates the technologies or concepts critical to MCM design. At the IC level, one of the most important concerns is heat removal. This includes heat transfer issues such as the thermal conductivity and surface area of heat sinks and boards as well as the thermal diffusion properties of encapsulants. At the chip-MCM substrate interface, properties such as fatigue, selection of the right materials (which obey a temperature hierarchy where solders used at the chip level do not melt when the MCM is attached to the board) and the type of electrical interconnect are important. The MCM substrate must have an appropriate dielectric constant and conductors that are of sufficient electrical

Samsung, EX1028, p. 155

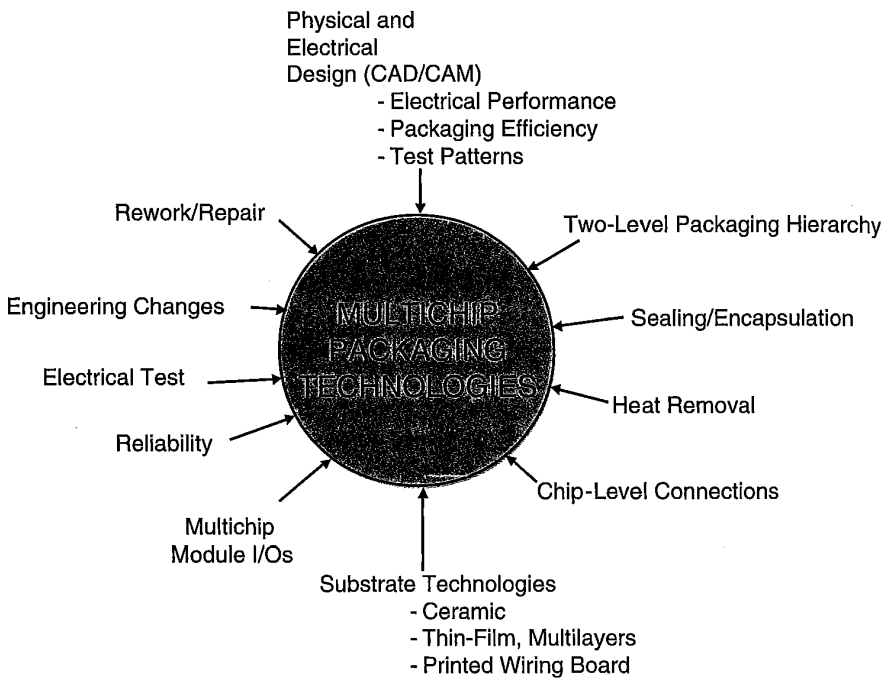


FIGURE 8.10 Technologies, important parameters, and concepts critical to MCM.

conductivity to ensure rapid signal propagation. Its wiring density must be able to accommodate the number of ICs and their I/Os with a minimum number of board layers and in the smallest practical area. In connecting the MCM to the board or system, one is particularly concerned about electrical inductance and resistance as well as mechanical properties of the interconnect, such as fatigue resistance. Board or system electrical and mechanical properties may play a role as well.

8.4.1 Electrical Design

The major electrical design requirement for MCMs is to control basic circuit parameters such as resistance, capacitance, and inductance. In particular, signal paths must be short with controlled impedances and low loss. Deviations from the package's electrical design specifications can result in spurious signals, increased delays, and signal waveforms that are distorted. The electrical design for MCMs has been presented in more detail in Chapter 4. In this electrical design process, the packaging engineer must address the following: dielectric constant; signal line geometries, thus controlling resistance, capacitance, and inductance; interline spacings to control crosstalk and spurious signals; and the distribution and location of power and ground. As the frequency of MCM operation goes up, the need for controlled impedance signal paths becomes increasingly important.

Samsung, EX1028, p. 156

8.4.2 Sealing and Encapsulation

Hermetic sealing or encapsulation of the MCM is extremely important and can be a major contributor to overall module reliability. MCPs historically have been hermetically sealed. Today, MCMs are either hermetically sealed in ceramic or metal packages or they are encapsulated. The tremendous improvements in polymeric encapsulants has allowed effective sealing of MCMs without the need for hermetic packages. Polymer sealing is effective, provided the polymer coat remains in intimate contact with substrate and chip surfaces, thus preventing liquid water accumulation and subsequent corrosion. In high-performance, high-value MCMs, where the module must be repaired rather than discarded, the encapsulants must be reworkable. In other words, the encapsulants must be relatively easy to remove, so that the malfunctioning chip(s) can be replaced and then allow recoating or re-encapsulation. This subject is presented in more detail in Chapter 14.

8.4.3 Heat Removal

Heat removal from the MCM is very important for efficient and reliable operation. Module power dissipations have risen over the years from a few watts per module to 30–180

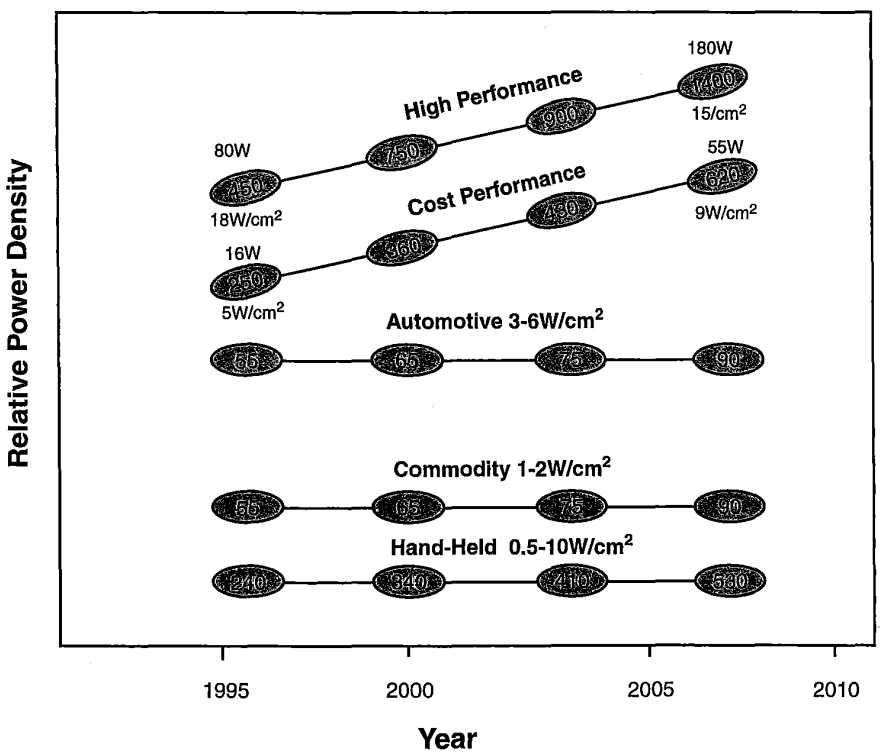


FIGURE 8.11 Integrated circuit chip power dissipation density as a function of time where the numbers in the ovals equal the die size in mm².

Samsung, EX1028, p. 157

W. The trend in increasing power density for MCMs is illustrated in Figure 8.11. Despite such power densities, the chips (ICs) must be maintained at 100°C or below. The substrate is a critical element in the heat removal process, but the actual thermal transfer depends on how the chips are interconnected to the substrate. The three most common ways for chip-to-substrate interconnection are wirebonding, flip chip, and *tape automated bonding* (TAB), as illustrated in Figure 8.12. In wirebonding and conventional or standard TAB, the chip is mounted face up and the chip's rear surface is directly bonded to the substrate. In this configuration, the heat removal path is directly through the chip's back surface and into the substrate. The heat transfer process is facilitated by a high thermal conductivity die attach medium. In flip chip and flip TAB, the chip is not back bonded to the substrate. The solder balls or flip TAB leads provide the only substrate contact. These methods of interconnect do not depend on high thermal conduction to and through the substrate. In the facedown mode, heat can be easily removed from the upward facing back surface of the chip using a metal heat sink or cold fingers placed in contact with the back surface. Inexpensive heat sinks can be made of copper, aluminum, aluminum nitride, or other high thermal conductivity materials. Cold fingers are typically liquid cooled. Using facedown interconnect techniques, the topside removal of heat can be very efficient, handling thermal power densities of 10–100 W/cm^2. Almost all supercomputers

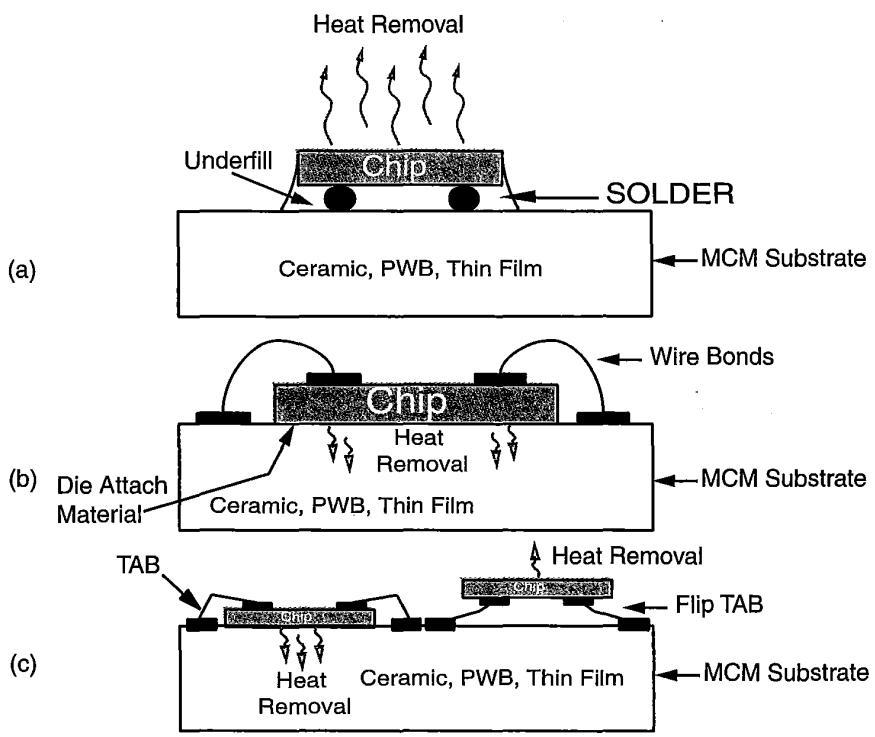


FIGURE 8.12 Methods of chip interconnection to the MCM substrate: (a) flip chip, (b) wirebonding, (c) tape automated bonding (or TAB).

Samsung, EX1028, p. 158

use some form of this technology. As packaging densities increase, the heat removal challenges become more formidable.

In wirebonding and standard TAB, the thermal conductivity of the substrate or carrier is very important. PWB substrates with special thermally-conductive layers have thermal conductivities equal to about 1 W/m·K. Alumina ceramic (thin-films and high temperature cofired) can raise this number to approximately 25 W/m·K. *Low temperature cofired ceramic* (LTCC), because of the large amount of glass in the dielectric layers, has thermal conductivities in the range of 1–10 W/m·K. Thin-film circuits can be built on high thermal conductivity substrates such as silicon (145 W/m·K) and aluminum nitride (200 W/m·K).

8.4.4 Electrical Interconnections

It is highly possible that some MCMs will include all three first-level chip interconnection types on a single substrate. In addition, other components such as chip resistors, capacitors, quad flat packages, and even a *pin-through-hole* (PTH) component may also have to be mounted. Regardless of the interconnection type, the basic requirements for the interconnect are:

- Fatigue and creep resistance
- Corrosion resistance
- Electromigration resistance
- High conductivity
- Processing temperature consistent with downstream heating environments

Each of the basic interconnection methods has its advantages and disadvantages. Wirebonding is the dominant form of first-level interconnect because of its flexibility, low interconnect cost, relatively low capitalization cost, and ease of use. Wirebonding is extremely flexible. It can bond various chip types, metallurgies, pad sizes and configurations, etc., by simply changing machine parameters, software programs, and, perhaps, bonding tips (capillaries and wedges), wire size, and material. Such changes are relatively low cost and can be performed very quickly. Similarly, the cost of a fully automatic wirebonding machine and a suitable wirebond testing machine, along with an organic die attach system, can be acquired for less than $200K. Wirebond connections in volume production have been in the range of 0.1 to 0.2 cents per interconnect. The other major interconnect types, flip chip and TAB, require major tooling and capital investment to produce the bumps and the prepatterned tape, respectively. Similarly, any minor changes in chip pad geometry will require costly phototooling changes, in addition to a procurement cycle for new tape or mating substances. This complexity, coupled with the cost of the increased processing involved, despite the advantage of gang bonding or mass reflow, produce a per-interconnect cost of between 5 and 10 cents. Given that the cost of interconnection is much greater (by a factor of 50 or more) than wirebonding, and the number of I/Os possible with wirebonding is rapidly increasing, why are interconnection types other than wirebonding considered important? Table 8.1 compares the three first- or chip-level interconnection technologies.

Flip chip has four major advantages: 1) it can effect the highest number of interconnects per unit area (and the interconnect density is constant); 2) all the interconnects are contained within the chip area, i.e., there is no second bond or outer lead bond location

Samsung, EX1028, p. 159

TABLE 8.1 Properties of current first level chip interconnection technologies.

	Flip Chip	TAB	Wirebond
Connections	3000–6000	200–600	600–1000
Interconnect Material (form)	Solder (ball)	Copper on Polymer (planar lead)	Gold, Aluminum, Al-Si (wire lead)
Process	Reflow (soldering)	Thermocompression (welding)	Thermocompression Thermosonic Ultrasonic (welding)
Pad Metallurgy (Chip)	Cr/Cu/Au	Ti/W/Au	Al-Si, Ti/W/Au
Pad Metallurgy (Substrate)	Cu/Ni/Au Ni/Au	Cu/Ni/Au Ni/Au Cu/Pb-Sn	Ni/Au Cr/Cu/Ni/Au
I/O Density	1600/cm^2	400/cm^2	200–400/cm^2
Inductance	5–10 pH	1–6 nH	1–6 nH
Mutual Inductance	1 pH	5 pH	100 pH
Rework	Good	Poor	Poor
Failure Rate (%/1000 hours)	10^{-8}	10^{-5}	10^{-6}

beyond the chip perimeter; 3) it has extremely low inductance and capacitance per joint, thus allowing operation at very high frequency; and 4) it has the most robust replacement process for preserving the underlying board or substrate.

TAB has one interesting advantage over the other two major interconnection forms. This advantage is that in the TAB process, the IC is attached to its final lead frame, or inner lead bonding, prior to placement in the package or on the substrate. Such attachment allows testing both at speed and temperature in a configuration that is closely representative of its final state, thus, in principle, solving the KGD problem. Similar techniques for wirebonded interconnects have been developed, but they lack TAB's flexibility and always involve some form of wirebond lifting, or removal, and replacement, which carries with it inherently greater risk of good die loss.

Given the entire discussion above, it is clear that wirebonding will be the dominant interconnect for some time to come in all applications that can afford the size perimeter extension required for the second bond, have a pitch large enough to allow wirebonding to effect the required number of I/Os with perimeter bonding pads, and a frequency of operation low enough (<10 GHz), so that wire loss is a manageable issue.

If requirements exceed those listed above for wirebonding, then flip chip appears to be the interconnect choice. Flip chip usage is increasing rapidly worldwide and is clearly the second most important first-level interconnection where its special advantages of small size and high electrical performance are required. On occasion, special testing requirements may necessitate the use of TAB.

8.4.5 MCM Connections to System-Level Boards

The I/O connections that supply power and allow the MCM to communicate with the rest of the system and peripheral equipment are extremely important. If the I/O number

Samsung, EX1028, p. 160

is large enough, this can only be accomplished with an area array type MCM interconnect. Large MCMs (>5 cm × 5 cm), however, have been produced with peripheral leads exceeding 400 in number. Area array packages of all types have been produced in the past, including pin grid arrays, pad grid arrays, and, more recently, ball grid arrays (BGAs). If the MCM is encapsulated rather than packaged in a conventional ceramic or metal package, then the substrate must provide this board-level interconnect. Probably multilayer ceramic and multilayer PWB type substrates provide the most effective means of providing these module contacts. In both of these technologies, the ability to create area array back contacts is relatively easy. With silicon as a carrier, such as in thin-film MCM-Ds, wafer via technology is not well developed. Although it can be done, and its use will increase in the future, silicon-based MCMs are typically limited to peripheral I/Os. The trend in I/O growth for both single and multichip packages is shown in Figure 8.13.

8.4.6 Repair, Rework, and Engineering Change

As shown in Figure 8.7, the yield of a large MCM is relatively low, even if the individual chips have a high probability of being good, or they have a high KGD probability. Thus, either MCMs must be built so inexpensively that they can be discarded, as is usually done with single chip packages, or they must be able to be repaired or reworked. In

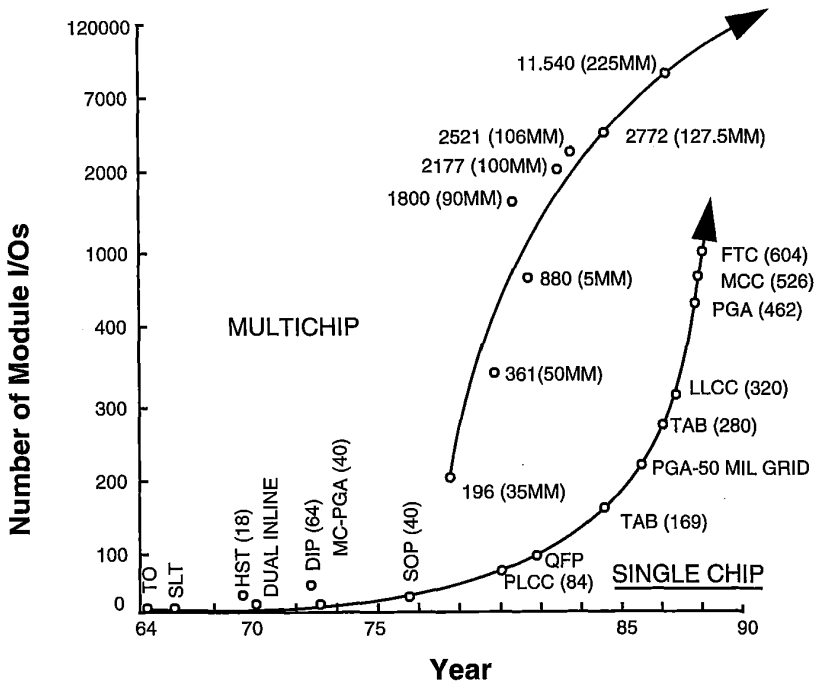


FIGURE 8.13 Trends in the growth of package inputs and outputs for both single chip and multichip packages.

Samsung, EX1028, p. 161

addition, the modules are typically very complex and, until they are actually built, the module designers, even after extensive simulation, may still need to make engineering changes. Repair and *engineering changes* are most often directed at the substrate and its structure. Typically, circuit lines can be shorted, open, or connected to the wrong circuit node. The removal of some line shorts and wrong connections can be done by a laser or similar technology, which ablates or removes the conductive path. Line opens or reconnections, after cutting a misplaced trace, can be accomplished by the deposition of metal conductors on an insulating surface, either directly or through a mask, using physical vapor deposition, chemical vapor deposition, printing polymer thick-films, electroplating, etc. The vias that connect wiring on one layer with that of another can also be open. This is generally caused by insulating materials and/or debris that are left at the via bottom prior to metal deposition. If the via accesses one of the MCM substrate surfaces, then it can sometimes be repaired by laser ablation of the insulating material at the bottom of the via and subsequent metal redeposition.

Rework generally refers to the need to remove bad chips and replace them with good ones. This is done by identifying the bad chips and selectively removing them by a combination of heat and mechanical means. Flip chips can be repaired by localized heating—ducted hot air, thermode, etc.—of the defective die in question, redressing the removal site (solder removal, leveling, and cleaning), and rebonding of the repair chip (again by directed hot air or some type of furnace heating).

8.4.7 Electrical Testing

Electrical testing is done at several different levels for MCMs. First, it must be ensured that the substrates are defect-free prior to the start of the assembly process (adding chips, interconnections, and some form of encapsulation). Defect-free substrate validation is usually accomplished by a series of design rule checks to ensure compatibility with the MCM fabrication process, followed by electrical testing using a bed of nails, or probe card, or flying probe testers. The electrical tests verify that all networks on the substrate are connected in the appropriate manner, consistent with the circuit's schematic diagram, and that there are no opens, shorts, or nets with resistance/impedances outside the prescribed range. These substrate electrical tests are often supplemented with visual inspections and *destructive physical analysis* (DPA) on a lot sample basis. The inspectors look for uneven metallizations, fine cracks that may have escaped electrical probe tests, and via misalignments. The DPA looks at the quality of internal layers by cross sectioning the sample.

High module yields require testing of the chips prior to assembly. These KGD are produced by additional electrical testing and associated costs prior to assembly. Much time and effort over the last few years has been expended on the mounting of bare die for full functional testing, at speed and temperature, and then demounting them for actual attachment to the MCM.

Following assembly, the resultant MCM must be electrically tested to ensure a working module, thus validating the defect-free nature of the assembly process and certifying the results of chip (KGD) and substrate testing. Even if the module is working after assembly, it still must be encapsulated or packaged and then environmentally stressed (temperature, humidity, bias voltage) to ensure its long-term reliability due to latent propagation of defects, corrosion, etc.

Samsung, EX1028, p. 162

In an MCM, if one die fails, the whole module fails or its performance is so reduced that the MCM cannot be used for its intended purpose or sold at a price consistent with the full cost recovery. Since chip and substrate pretesting alone cannot ensure that all dies will perform as planned after assembly, MCMs must either be built so inexpensively that they are disposable throwaways that use a low-cost, high-yield, high-volume manufacturing process, or they must be able to be repaired by replacing defective chips.

Testing to find the defective chips is a key element in the repair process. Without appropriate procedures to locate defective die, repair can be extremely costly or impossible. MCMs must be designed from the beginning for testability and repair, including the complete test protocol necessary to locate defective die. The need to test and repair influences substrate design, including extended bond pads, room for die attach and removal tools, robust board metallizations, and, perhaps, extra test points, traces, and/or chips to support the testing process. The chips themselves may require extra circuit elements and contacts to facilitate the testing process. The design for testability of both substrates and integrated circuits is beyond the scope of this work, although more details on module testing are presented in Chapter 19, which is focused exclusively on electrical tests.

8.5 TYPES OF MULTICHIP MODULE SUBSTRATES

Recent advances in integrated circuit density and speed have allowed the IC to outstrip the capacity, or the number of interconnects per unit area, trace density, etc., of conventional PWBs. Although PWBs have adapted well over the years to changing IC technology by adding layers to provide increased wiring density, modifying dielectric characteristics to accommodate higher speed circuitry, and reducing feature size to allow improved packaging density, the new requirements far exceed the capabilities offered by today's standard PWB technology.

MCMs were invented to replace the standard PWBs in circuits that required interconnecting many high-density ICs and other complex components. In addition, since modern circuits have to be compact lightweight and operate at high speed and/or high frequency, MCMs also were developed to reduce delay times and crosstalk between components and signal traces. As a rule of thumb, a packaged electronic circuit is considered to be an MCM if more than half of its area is covered with active devices [see Equation (8.1)], although some circuits adapt more readily to this definition than others. As mentioned previously, MCMs come in three basic styles: MCM-Ls, based on organic laminate technology; MCM-Cs, thick-film or cofired ceramic technology; and MCM-Ds, deposited thin-film multilayers.

Each type of MCM has its place in the technological hierarchy. MCM-L is derived from conventional PWB technology and makes use of enhanced materials and processes, high definition photolithographic techniques, and strict process control to allow the fabrication of small feature sizes and the more accurate placement of components than are customarily found with PWBs. MCM-L component assembly procedures differ dramatically from PWB assembly practices; most noticeable is the replacement of standard packaged-part soldering processes associated with surface mount technology, with bare die interconnected with wirebonds or flip chip processes. In this respect, MCM-L is similar to what once was called hybrid microelectronics with two notable exceptions: 1) the substrate is a laminated organic structure rather than ceramic or silicon, and 2) the

Samsung, EX1028, p. 163

module package need not be hermetically sealed. MCM-L provides the lowest density of the three major MCM technologies and is usually the least expensive to implement. MCM-L is also known as *chip-on-board* (COB). While some authors make a distinction between COB and MCM-Ls, the differences are minor.

MCM-C is similar to the older ceramic, thick-film hybrid circuits, with some notable improvements. In conventional ceramic hybrids, the multilayer circuitry is screen printed onto a rigid ceramic substrate, usually Al_2O_3 or alumina. Printed layers alternate between metal traces and dielectrics, with metal-filled "vias" to connect one layer to the next. The ceramic dielectric layers typically have a dielectric constant greater than 9. A true MCM has only bare ICs and, perhaps, some passive chip components on its surface. A few specialized packaged parts, or reference crystals, are occasionally used. The packaging density [see Equation (8.1)] is proportional to the fraction of the substrate area occupied by circuit comments, but is limited by via size and alignment difficulties when screen-printing several layers, which become less planar as the number of layers increases.

A particularly interesting MCM-C substrate approach is offered by *low temperature cofired ceramic* (LTCC). This fused ceramic technology can be processed at conventional thick-film temperatures (800–1000°C) rather than the normal cofired temperatures of up to 1400°C. To fire at the low temperatures, glass is incorporated with the ceramic. The glass not only reduces the fusing point but also lowers the relative dielectric constant, thereby improving circuit performance at high frequencies.

In the LTCC process, each layer is screen-printed on a flexible, unfired ceramic tape. Layers typically have laser-drilled via holes that are filled with conductor material by screen-printing. The layers are subsequently aligned and subjected to high-pressure consolidation being fired together all at once. The small laser-drilled vias and the planarity of each screened layer serve to increase the wiring density over the conventional "print, dry, and fire" operations associated with the traditional multilayer thick-film process. The LTCC process adapts well to the incorporation of blind vias—those that do not extend completely through the substrate—and also permits the fabrication of precisely dimensioned cavities that have found applications in microwave circuitry. Recent modifications in this technology will allow somewhat higher packaging density to be achieved through the use of ultraviolet, light-sensitive materials that are photolithographically defined, rather than printed. These photo-defined conductors can have smaller line widths and better edge quality, thus making them suitable for high-frequency circuitry. Screen-printed thick-films have also been etched to produce better edge quality.

The various MCM-D technologies yield the highest circuit density while being the most process-intensive and expensive of the three MCM technologies. Although MCM-Ds tend to be expensive per unit substrate area, they typically use much less area for a given circuit application and, thus, can be cost competitive, especially in large volumes. In MCM-Ds, both the metal and dielectric layers are sequentially deposited in the form of thin-films, and patterned by a photolithographic pattern transfer process. The substrate base layer or carrier must be very smooth and flat. Polished metal plates, ceramics, glass, and semiconductor wafers have been used. Typically, silicon wafers are used as the MCM-D substrate base layer, because they are readily available, very smooth and flat, and relatively low cost, and, more importantly, they match the coefficient of thermal expansion of the dielectric layers, or low-stress polyimide, and, of course, the silicon die. In today's MCM-Ds, the silicon base is merely a platform for supporting the circuit

Samsung, EX1028, p. 164

TABLE 8.2 Comparison of MCM features.

Design Parameters	MCM-L	MCM-C (LTCC)	MCM-D
Feature size (line/space) (mm)	125/125	100/125	20/20[a]
Via size (μm)	250	200	20
Critical dimension uniformity $\pm$ (μm)	12	25[b]	5
Number of levels	10[c]	30[c]	5+
Dielectric constant	3.5 to 4.5	5.2 to 7.8	2.9
Dielectric thickness (μm)	112	100	1 to 10
Cutouts/cavities	Yes	Yes	No
Integrated resistors/capacitors	No[d]	Yes	Yes

[a]Smaller lines and gaps can be defined, subject to circuit impedance and delay requirements.
[b]Even though line widths are smaller for MCM-C (LTCC) than MCM-L, critical dimension uniformity is greater because of firing shrinkage.
[c]Typical values. The technology lends itself to the incorporation of many additional levels.
[d]Integrated capacitors and resistors are being developed in several laboratories.

constructed upon it; the semiconducting properties of silicon play no role, as opposed to their role in ICs. In the future, active devices, as well as passive components, will be integrated into this base or supporting layer. Feature (conductor widths and gaps, vias, etc.) sizes can be less than 25 μm. Table 8.2 presents a brief comparison of salient properties of the three MCM types.

8.5.1 MCM-L

MCM-L technology is derived from organic PWB fabrication and uses copper metal traces separated by fiberglass-reinforced organic laminates (the most common), with plated through-holes or vias interconnecting the layers. A typical fabrication sequence is shown in Figure 8.14. Compared with conventional PWBs, advanced processing techniques result in smaller feature size and allow the incorporation of vias that extend only part way through the board (blind vias) as well as those that connect between layers entirely within the board (buried vias). The use of blind and buried vias greatly increases the wiring density of components that can be connected on the board.

The label MCM-L refers specifically to the technology used to manufacture the wiring platform. The platform can be used to interconnect ICs, as in a pure stand-alone MCM, or to build up a complex circuit board containing many different components, active and passive. Once populated, the MCM-L uses organic coatings to protect the chips and their bonds (*glob top*), as well as the components and the entire board. Some complex boards may even have other MCMs attached to them.

There are three types of laminated (MCM-L) substrates: rigid, flex, and rigid-flex. Rigid MCM-L substrates are typically made of epoxy-based polymer resin dielectrics, reinforced with fiberglass. Other reinforcing materials include quartz, Kevlar®, and Aramid®. Different resin systems, as discussed in Chapter 17, include polyimide, Teflon® (polytetrafluoroethylene), and PEEK (polyether ether ketone). Polyimide provides excellent stability and a low dielectric constant, thus assuring excellent signal propagation and high frequency performance. There are two types of dielectric layers in typical MCM-L construction: cores and prepregs. *Core* material is fully cured when delivered from the

Samsung, EX1028, p. 165

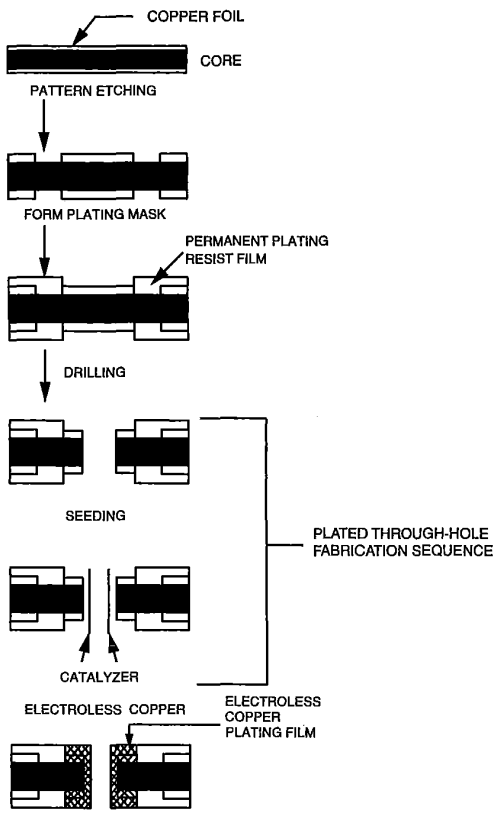


FIGURE 8.14 Schematic representation of a standard printed wiring board (MCM-L) fabrication sequence.

manufacturer and is usually clad on both sides with copper. The as-received *prepreg* material is partially cured and is used between core layers as a bonding agent to hold, or glue, them together. Both the cores and prepregs are typically fiber reinforced. If the dielectric layers have no reinforcing fibers, the resulting laminate in thin layers is quite flexible, leading to the name flexible or flex circuits. Combining both types of laminates into a single board structure yields an MCM-L substrate called rigid-flex.

MCM-L substrate process. There are five basic process steps in the manufacture of MCM-L substrates:

1. Selecting the appropriate core and prepreg layers. The layer thicknesses are chosen to meet the circuit electrical and mechanical performance criteria.
2. Photolithographic patterning and etching of the copper conductors on the core layers.
3. Drilling of vias (blind, buried, or full through-holes).
4. Lamination of the cores to each other using the prepreg layers. Multiple laminations are usually required to form blind and buried vias. If the board only

Samsung, EX1028, p. 166

has plated through-holes, then a single lamination step is usually possible with post-lamination via drilling.

5. Plating of drilled holes in single layers (buried vias), partially through several laminated layers (blind vias), and holes going all the way through the board (plated through-holes as used in PWBs).

Copper is the typical conductor material used in MCM-Ls and PWBs. The copper comes laminated to the core material. The thickness of the copper is specified by an arcane measure of its weight, in ounces, rather than its direct thickness. A 1 oz. copper foil is 1.4 mils or about 35 μm in thickness, based on the density of copper and the unit board area. The ounce system refers to the weight of the copper layer per square foot.

Inner layer processing. The copper surfaces are cleaned in preparation for pattern processing. Photoresist is applied to the copper surfaces by laminating (heat and pressure) of a dry film resist material or by dipping in a liquid photoresist bath and drawing the copper material out at a uniform rate. Other techniques for resist application have been used including spray coating, screen printing, and/or electrophoretic deposition. Liquid resists typically allow finer line definition than the dry film used in PWB fabrication. Following setting of the resist, the pattern is exposed through a mask by ultraviolet light, developed to remove the unwanted resist areas, and hardened by baking against the etch. The copper foil is etched in an ammonia-based alkaline system or by using such materials as ferric chloride. The photoresist is then chemically removed, or stripped, and the copper surface is treated to form an oxide layer that promotes the adhesion of the overlaying dielectric layer during lamination.

Via formation. As mentioned above, PWBs utilize holes drilled through the entire board thickness to electrically connect the desired metal planes, and the patterned conductor layers that intersect the hole. After drilling, the holes are copperplated, a combination of electroless plating followed by electroplating, to make the required electrical connections, called plated through-holes. Blind vias extend from one surface to the desired internal plane while buried vias only interconnect internal planes. Both blind and buried vias are then copperplated as above to effect the electrical interconnection.

The use of blind and buried vias is considered expensive by many experts, due to the difficulties encountered in handling the thin core and prepreg layers with the multiple drilling, plating, and lamination operations required. However, the use of blind and buried vias in place of plated through-holes greatly increases board (substrate) wiring density.

Surface layer finishing. The treatment of copper surface layers defines a major difference between PWBs and MCM-Ls. Conventional PWBs are made for surface mounting or the through-hole soldering of leaded packaged components. The surface pads are typically solder plated with tin-lead solder, which is suitable for mass surface mount reflow or the wave soldering of through-hole devices. MCM-L substrates potentially require the selective plating of several metals, including gold for wirebonding, tin-lead solder for chip components and TAB leads, and hard wear-resistant gold for edge connectors. When gold is plated, it must be segregated from the copper by use of a nickel layer. Since these surface layers are usually required to be plated after patterning, electroless (autocatalytic) deposition processes must be used. Gold for wirebonding must be of medium hardness ($\leq$80 kg/m^2 on the Knoop Hardness Scale) and at least 0.5 μm in thickness for good bonding results.

Samsung, EX1028, p. 167

Advanced MCM-L substrates. The major limitation of a standard PWB for MCM-type applications is the lack of wiring density required to meet the routing and I/O requirements. This is due in large part to the need for the plated through-hole vias produced by mechanical drilling. In a typical PWB, the drilled holes are relatively large (nominally 12–15 mils, 300–475 μm) with large cover pads. As the size of the via is reduced, the cost of hole drilling goes up tremendously, as shown in Figure 8.15. Other factors which may limit a PWB's performance as a true MCM-L include the thickness of the basic laminate layers, the resolution of fine lines and gaps over relatively large areas on the "rough" laminate due to the weave of the reinforcing fibers, the thickness of the metal to be patterned, relatively high coefficient of thermal expansion compared to the CTEs of silicon and GaAs, and polyimide boards required to meet the demands of high-temperature processing, such as with lead-free solders, thus increasing cost.

All advanced PWBs or MCL-Ls are based on thin-film technologies which can be processed over large board areas, thus reducing cost per unit area, much the way large area semiconductor processing reduces individual chip cost. The important technologies include dielectric deposition, via formation, and conductor deposition and patterning. Figure 8.16 illustrates various options for each of these key areas.

Today, a typical MCM-L, or chip-on-board substrate, will have 75-μm lines and spaces, 200-μm vias, and 400-μm lands or plated through-hole cover pad. On such a board, one cover pad can block out 3 to 4 line routing channels. In advanced, high-density applications such as microprocessors, telecommunications, and workstations, the

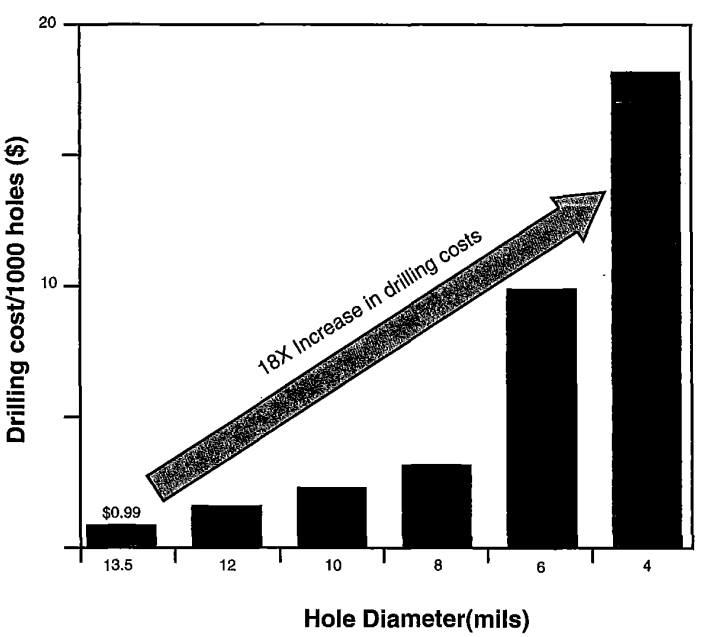


FIGURE 8.15 Drilling cost of holes in organic laminate material as a function of hole diameter. As the size (diameter) of the hole decreases, the drilling cost rises rapidly.

Samsung, EX1028, p. 168

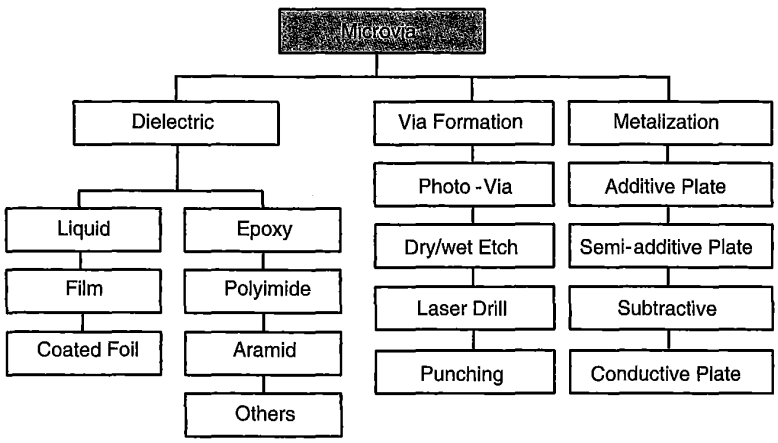


FIGURE 8.16 Options for fabrication of microvias in advanced printed wiring board or MCM-L applications.

loss of this much wiring density cannot be tolerated. A new and emerging technology, which can provide high wiring density at low cost, is shown in Figure 8.17. The illustration depicts a ball grid array (BGA) with a hybrid or built-up substrate. The substrate consists of a conventional PWB center or core, using standard through-hole vias with two multilayer thin-film circuits on either side of the core. The multilayer thin-film circuits contain blind and buried vias.

Built-up technology began in Japan (IBM Japan Yasu Technology Application Laboratory) with an innovative technology called *surface laminar circuits* (SLC). The SLC is very simple in concept, the basis of which is to use low-cost materials such as epoxy, or polyimide, and copper metallization, which are already used in both conventional PWB and MCM-D construction, but apply them together over large board areas. This is a

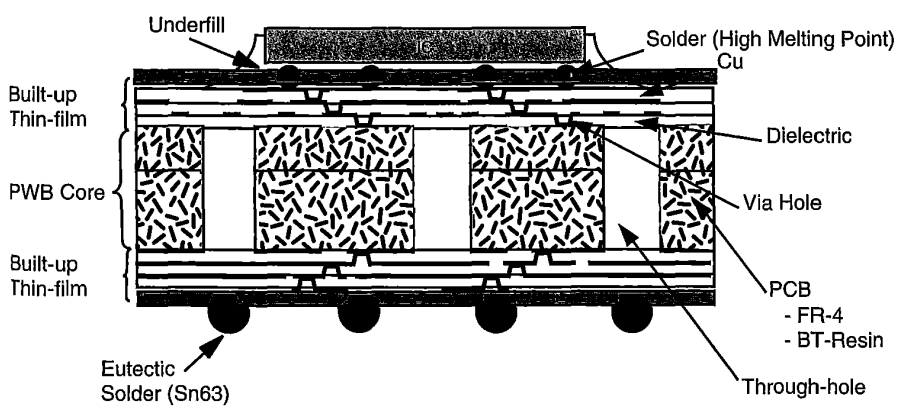


FIGURE 8.17 Schematic representation of a built-up MCM-L. (Thin-film multilayer signal layers are placed on either side of a PWB core.)

Samsung, EX1028, p. 169

natural evolution of the PWB, which requires higher wiring densities, and MCM-D, which requires cost reduction. Built-up technology can be considered a merger of these two technologies.

The built-up SLC is a two-part structure (Figure 8.17), comprising a copper-clad, epoxy-glass (FR-4) carrier substrate that has been built-up on both sides with alternating layers of photo-imageable dielectric coating, either epoxy or polyimide, and copper metallization, either by deposition or plating. Signal interconnection between wiring planes in the built-up layers is accomplished by via holes produced in the photo-imageable dielectric layer by UV exposure and subsequent developing. Contact between the front and back thin-film multilayers is accomplished by plated through-holes in the PWB-like core. These holes may need to be plugged with a conductive material to maintain planarity of the thin-film layers.

Key advantages of built-up technology include:

1. Wiring density greater than a conventional PWB's due to elimination of drilled through-holes (except in the core)
2. Capability to support very high device I/O densities
3. More flexible design options for EMI shielding than with MCM-D technology alone
4. Utilization of conventional interconnect technologies: wirebonding, TAB, and flip chip
5. Room to grow (more thin-film layers and/or core could be multilayer) as technology needs increase

8.5.2 MCM-C

The middle rung on the three-step MCM classification ladder uses ceramic-based substrates. These MCM-C substrates have evolved from traditional thick-film fabrication techniques in response to the need for increased packaging density and enhanced performance. Density was increased by shrinking the size of features such as the vias used for interconnecting substrate layers, the conductor traces used for signal routing, and the gaps between different traces or vias. In the traditional thick-film process, these features are created by screen-printing, in which the screen mesh geometry limits the achievable line widths and spaces. Several parameters determine the resolution of printed elements in addition to the screen itself: the viscosity of the thick-film paste, the ability of the emulsion to form a good gasket with the substrate, and the squeegee speed and pressure. Combining and optimizing all these elements results in a process than can routinely produce 200-μm vias and 125-μm lines or spaces. With care, 75–100-μm lines and gaps can be produced in selective areas. An alternate and more advanced ceramic processing method is called *cofired* technology.

The main difference between conventional thick-film and cofired ceramic processes is the nature of the dielectric. In the thick-film process, the dielectric layers are screen-printed, as are the conductor layers, by screening the appropriate paste onto the ceramic carrier substrate, usually with the via holes delineated and aligned with the metal pattern, and then fired. Subsequent layers are built on top of the previous layers like a stack of bricks set on a foundation.

Samsung, EX1028, p. 170

With the cofired ceramic process, the dielectric layers are sheets of unfired ceramic, green state ceramic, or green tape, if the layers are tape cast. Each sheet is separately patterned with its metal conductor traces and then collated and stacked together like pages in a book. The cofired ceramic process can produce 200-μm or smaller vias because they are either mechanically punched or laser machined into each individual unfired ceramic layer instead of being screen-printed. The vias are filled by extruding the conducting paste into the via holes through a stencil made from either thin brass or a polyimide sheet. The reduction in line and space sizes with cofired technology results chiefly from being able to print each conductor layer on its own slice of unfired ceramic. The surface is significantly smoother than thick-film ceramic and essentially flat, which allows the screen emulsion to seal against the surface during the printing process, thus ensuring high pattern integrity. Finally, since all of the circuit features during cofired ceramic production are created on unfired ceramic, they undergo identical shrinkage of about 15–20% when fired—that is, all the layers are fired at once, hence the use of the word "cofired" in describing this technology. Cofired fired ceramic technology falls into two major areas based on their firing temperature: *high temperature cofired ceramic* (HTCC) and *low temperature cofired ceramic* (LTCC). The HTCC technology utilizes primarily alumina, with small amounts of glass and various organic components, green sheet ceramic dielectric layers, which fire at temperatures between 1400 and 1600°C. Because the HTCC must be fired at these extremely high temperatures, it uses refractory metal conductor materials such as tungsten and molybdenum, mixed with alumina and small amounts of glass (4–11%).

The LTCC technology utilizes ceramic dielectric layers containing large amounts of glass, thus producing firing temperatures in the range of 800–1000°C. These heavy glass-bearing ceramic layers have a dielectric constant about half that of alumina, nominally 5, and because firing is done at relatively low temperatures, the refractory metal electrodes can be replaced with higher conductivity materials, nominally 3 or more times better than tungsten or molybdenum, such as copper and gold. The large glass content of the dielectric also gives the composite module a coefficient of thermal expansion very close to that of silicon and GaAs. The glass, however, reduces the LTCC's thermal conductivity significantly when compared to an HTCC module's. Cross-sectional diagrams of the various LTCC process stages are shown in Figure 8.18.

Computer-aided design tools can be used to create a cofired ceramic substrate layout from a circuit schematic. In addition to a set of design rules (feature sizes, etc.) customized for each process, one must consider certain issues that could affect manufacturability, such as via and cutout size and location, and metal balance. The individual unfired ceramic layers are analogous to pieces of paper; they are flexible and can be cut or torn, but unlike paper, they are somewhat brittle in that they can fracture if the stresses are too great. A well-designed layout places vertically adjacent vias in a staggered pattern rather than in a straight line, and the design should avoid the formation of "ceramic islands," or areas separated from the remainder of the substrate by a perforation of vias. During the processes of via filling and conductor printing, the pieces of ceramic are held in position by a vacuum that passes through a porous stone fixture to the back side of the single-layer substrate. Large-area cutouts and a large number of vias reduce the amount of surface area in contact with the vacuum fixture, making it difficult to hold the layer securely during printing. Care must be taken to limit the openings in each ceramic layer to no more than 40% of the total surface area.

Samsung, EX1028, p. 171

FIGURE 8.18 Schematic representation of low temperature, cofired ceramic (LTCC) substrate (MCM-C) fabrication sequence.

MEMS23167

Samsung, EX1028, p. 172

The other issue, metal balance, is important for two reasons. First, the metal shrinks at a slightly different rate than the ceramic during firing; and second, the mechanical strength of the fired metal is lower than that of the fired ceramic. To ensure uniform shrinkage and mechanical integrity, the design should strive to maintain the same cross-sectional metal content across the substrate. Individual layers can be "unbalanced" as long as the sum of all the layers provides a reasonably overall even distribution of metal. Sound designs lead to robust cofired ceramic substrates capable of yielding high-performance MCM packages. Taking appropriate care, ceramic cofired technology substrates can be made with as many as 100 layers, each with 70- to 90-μm vias.

Figure 8.1 shows an example of a high-density MCM-C circuit using cofired ceramic technology. Figure 8.19 is a direct digital synthesizer circuit fabricated with LTCC technology. While relatively low in packaging density, the circuit had some unique challenges, including mixing digital and radio frequency circuitry together, as well as dissipating high power. Figure 8.20 is a 14-GHz transmitter/receiver module which contains five GaAs microwave ICs fabricated by a foundry using the MESFET (metal semiconductor field effect transistor) process. These dies, along with passive components and thin-film 50-Ω transmission lines, must be packaged within a waveguide structure designed to be beyond its cutoff frequency at 14 GHz. This circuit was fabricated using LTCC technology. LTCC's unique ability to include complex cavity structures with accurate dimensions was a perfect match for this application. The substrate contains 23 dielectric layers, with the microwave cavity formed by the top 14 layers. The components are mounted on the cavity floor (layer 9), and all of the layers below contain thermal vias to facilitate heat transfer from the back of the die. Signal routing occurs on all 23 layers; the package input/output is brought to the top layer. In this configuration, the transmit/receive func-

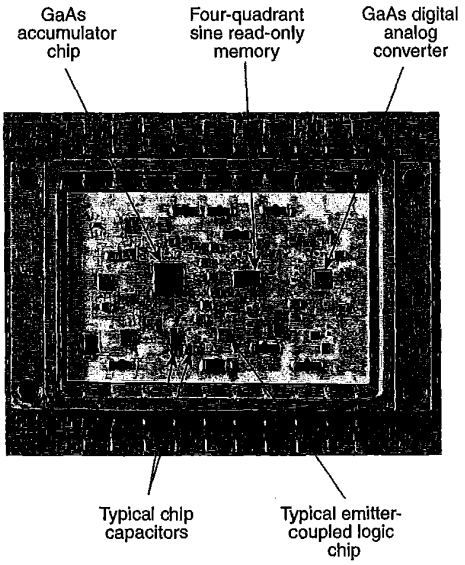


FIGURE 8.19 Direct digital synthesizer operating on the MSX spacecraft. (The circuit, built with LTCC technology, converts the digital clock to a sinusoidal reference signal. It operates at frequencies up to 1 GHz and dissipates over 30 W of power.)

Samsung, EX1028, p. 173

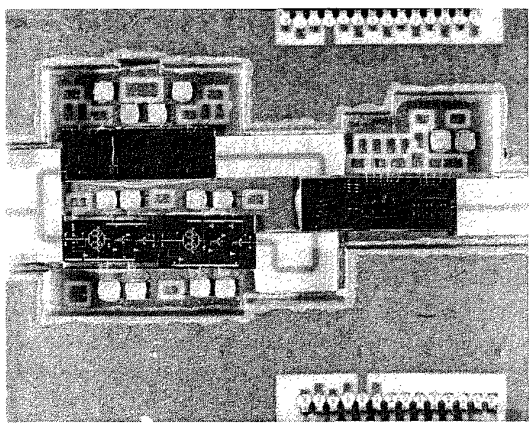


FIGURE 8.20 A 14-GHz transmit/receive module packaged in a waveguide beyond cut-off cavity structure fabricated using LTCC MCM-C substrate technology.

tion is achievable with only a 0.25-dB reduction in gain of the low noise and power amplifiers.

The thick-film industry's introduction of photo-printable thick-film pastes promises to further reduce feature size in MCM-C substrates by eliminating screen patterning altogether. This technology has been shown to yield a process capable of producing vias, lines, and spaces down to 75 μm. In this process, the conductor and dielectric pastes are photosensitive. They are applied to the entire substrate by printing with a blank screen, a screen with all of the emulsion removed. Then, after drying the layer, it is exposed through a photomask to ultraviolet light. The unexposed regions are subsequently removed in development, leaving behind the desired circuit pattern, which is then fired using a standard thick-film furnace profile. Since this system optically delineates substrate features, it has significantly greater resolution than screen-printing. In addition, blind and buried vias are easily fabricated; no critical processing is required, as is the case in MCM-L circuits. Other vendors are introducing etchable thick-films that hold similar potential, especially for high-frequency circuitry.

8.5.3 MCM-D

The combination of superior materials and the dimensional resolving power of thin-film technology enables MCM-D to dominate the MCM arena as the clear leader in packaging density and circuit speed. Feature sizes can be smaller than 10 μm, which is an order of magnitude smaller than either MCM-L or non-photodelineated MCM-C. The fabrication processes are similar to those used in the manufacture of ICs in that all of the features are photolithographically defined.

The classification MCM-D applies to several different dielectric/metallization technologies that share a common technological philosophy of photolithographically patterning sequential thin-film layers. Spin-deposited polymers such as polyimide and benzocyclobutene are the mainstay dielectrics; chemical-vapor-deposited silicon oxides,

Samsung, EX1028, p. 174

nitrides, and oxynitrides are sometimes used as well. The conductors are usually sputter-deposited copper, aluminum, or gold.

A popular MCM-D technology uses a single-crystal silicon wafer base layer with spun-on polyimide dielectrics containing buried copper conductors and gold as the top-level metallization. The first layer deposited on the silicon carrier can be either a conductor or a dielectric, depending on the circuit design. When forming a dielectric layer, spin coating of the polyimide is done in several steps. Initially, an aminosilane solution is spun onto the surface and baked to provide a monolayer that increases the adhesion of the polyamic acid precursor. Then, to ensure a defect-free dielectric layer and to increase the degree of planarization, a second coat of the polyamic acid precursor is applied and soft-baked. The complete curing process entails further heating up to a peak temperature of 350–400°C, which provides enough thermal energy to crosslink the polymer and form a true polyimide. The vias are formed in the polyimide by reactive-ion etching in an oxygen plasma using a photo-patterned metal mask.

The next step is to remove the metal mask and lightly plasma-etch the top surface of the polyimide to increase bonding sites for the subsequent metal deposition. The metal layer is actually a tri-layer of two metals, with a thin adhesion metal underneath and on top of a thicker high-conductivity metal. After photo-patterning, the metal stack is etched in various acid solutions, the photoresist is removed, and the process is repeated with the next dielectric layer. The cross-sectional drawings in Figure 8.21 depict the process steps.

As with other MCM technologies, computer-aided design tools are valuable for laying out MCM-D substrates, and some processing knowledge beyond design guidelines can facilitate building a circuit design that enhances yield and function. The weakest link in the MCM-D substrate fabrication process is forming the dielectric. It is difficult to create large planes of defect-free polymer films, and thus a major concern when generating a substrate design is the amount of *crossover area.* This area is created whenever the metal that sits on top of a given dielectric layer crosses over any metal on top of the layer immediately below.

Any open defect in the dielectric layer in the crossover area will cause a short between the two metal layers, effectively compromising the substrate. A design that minimizes the crossover area is therefore desirable to maximize yield. Most importantly, full area ground and/or power planes must be avoided, since adjacent layer ground and power planes will almost certainly cause failure. Whenever possible, distributed power and ground lines should be used, and if full planes are required, they must not be next to each other in the layer stack. The use of gridded ground planes can also help in this regard. Also, when routing signal traces on more than one layer, it is important to keep most of the traces within each layer running parallel to each other, and to alternate the direction of adjacent signal layers, so that the traces on those layers are orthogonal to each other.

Progress in MCM-D technology will not likely focus on further shrinking of feature sizes to increase packaging density, since the module surface, in some cases, may already be nearly completely covered with devices. The relatively high cost of MCM-Ds is still an important consideration. However, increasing substrate yield can reduce costs. This can be accomplished by developing more efficient processing equipment, like polymer precursor dispensing systems, and by reducing the process cycle time through the use of

Samsung, EX1028, p. 175

FIGURE 8.21 Schematic representation of multilayer thin-film polyimide substrate (MCM-D) fabrication sequence.

MEMS231679

Samsung, EX1028, p. 176

photosensitive polymer dielectric, which simplifies the process while enhancing via reliability.

8.6 MULTICHIP MODULE DESIGN

The design of a multichip module is a complex process involving many elements, including the required electrical function (schematic), the size and weight of the overall MCM, environmental performance, and reliability. The size of the overall MCM substrate is determined by the physical chip size, the interconnect perimeter, the wiring density, and, of course, the system physical constraints, like a hand-held cellular phone, a portable computer, etc. In order to determine the basic size possible, a wireability analysis must be performed. The objective of a wireability analysis is to determine, before detailed design and layout, the ability to route a substrate. Often the board size is fixed for a particular application and the board cost may be constrained as well. Wiring analysis can predict the number of layers required to wire the components for a particular circuit design. Since the routing effort and the layer count are directly proportional to board cost, wireability analysis is a useful tool for projecting substrate cost.

The basic concepts behind the wireability analysis include estimation of wiring demand, wiring capacity, average wire length, and connectivity. Wiring demand, D, is the amount of wiring required to interconnect a given circuit. Wiring capability, C, is the amount of wiring available for interconnection. Wiring demand is related to wiring capacity through the equation:

$$D = \varepsilon C \tag{8.3}$$

where ε is the wiring efficiency. Typically, the wiring efficiency is in the 30–70% range, depending on circuit type.

Wiring capacity, which is a function of the minimum signal line pitch, P_S, that can be fabricated on a given MCM substrate technology is usually normalized to a given square area such as cm^2 or inches^2. Sometimes it is normalized to the size of the pitch, P_P, of the packaged part or bare chip attached to the MCM. In this case, neglecting via holes and through-holes, for N_T signal layers, the total wiring capacity, C_T, is given by:

$$C_T = P_P N_T / P_S \tag{8.4}$$

To obtain the wiring demand, the average length, $\bar{L}$, per interconnection must be estimated. This length between any two points (1 and 2) is not a direct route, since alternate layers of signal line are usually routed in x and y orthogonal channels or streets. The average wiring path length is sometimes referred to as the Manhattan length, which is given by $|x_1 - x_2| + |y_1 - y_2|$. The wiring demand is simply the number of pin-to-pin connections, N_{pp}, times the average length $\bar{L}$. If considering that, for an average number of pins per net, N_{AP}, there are $(N_{AP} - 1)/N_{AP}$ number of wires, then for a given number of inputs and outputs, N_{IO}, the wiring demand is:

$$D = \left(\frac{N_{AP} - 1}{N_{AP}}\right) N_{IO} \bar{L} \tag{8.5}$$

To estimate the average wire length, it is assumed that the chips on an MCM substrate

Samsung, EX1028, p. 177

interact with either their nearest neighbors or their next nearest neighbors. This yields an average length given by:

$$\bar{L} = (P_P + 2P_P)/2 = 1.5P_P \tag{8.6}$$

Rickets estimates this length to be approximately:

$$\bar{L} = 0.77P_P N_C^{0.245} \tag{8.7}$$

where N_C is the number of chips to be interconnected. Other methods for estimation of $\bar{L}$ are based on Monte Carlo simulations and interconnect probabilistic weightings.

Estimating the wiring demand in the absence of a preliminary layout requires the use of Rent's Rule, which is given by the relation:

$$N_{IO} = ag^b \tag{8.8}$$

where N_{IO} is the number of inputs and outputs for the chip, package, or substrate; g is the number of fundamental circuits (gates) contained within the chip, package, or substrate; a is the Rent's coefficient or the average number of connections per circuit or I/O; and b is the Rent's exponent. Approximate Rent's Rule exponents and coefficients for various devices and systems are given in Table 8.3.

Rent's Rule is especially useful in determining required circuit density and optimal device placement for a particular circuit design. In the following example, the model described by Schmidt for a PWB with an I/O connector on one edge, is extended to estimate the maximum wiring required by a four-sided MCM with inputs and outputs on all sides. The total number of gates, or circuits, G, on an MCM containing an arbitrary number of circuits (chips), can be found from the relation:

$$G = \sum_{i=1}^{m} N_i g_i \tag{8.9}$$

where N_i is the number of type i components, g_i is the number of gates in the ith component type, and m is the number of different devices. This expression assumes that all of the individual circuits or components have the same Rent's relationship. The total package I/O count, N_{IO}, assuming no interaction between the circuits of different types (a worst case wiring scenario), is given by:

TABLE 8.3 Rent's Rule coefficients and exponents for various device and system technologies.

Device/System	Rent's Coefficient, a*	Rent's Exponent, b*
Dynamic Random Access Memory (RAM)	6.20	0.085
Static RAM	6.00	0.120
Microprocessors	0.82	0.450
Random Logic (gate arrays)	1.90	0.500
Computer Systems	2.50	0.600
Chip/Model Level	1.40	0.630
Board/System Level	82.00	0.250

*a and b refer to the coefficients and exponents, respectively, as shown in Equation (8.10), for example.

Samsung, EX1028, p. 178

$$N_{IO} = aG^b \tag{8.10}$$

Consider the circuit shown in Figure 8.22a, which contains two gate array devices with 5000 gates each and eight 64-MB static random access memory (SRAM) chips. Using the values for a and b given in Table 8.3, the number of gate array chip I/Os is estimated to be 190 each, and the SRAM I/O to be 52 leads each. The Rent's exponent for the gate array is assumed to be 0.5, the theoretical value for pure random logic. Since the SRAM circuits have a different Rent's exponent, it is necessary to develop an equivalent number of circuits for them, so that they have the same Rent's exponent and coefficient as the gate arrays. This can be done by solving the equation:

$$1.9(g_{SRAM})^{0.5} = 52 \tag{8.11}$$

Although each of the SRAMs has 64 million circuits, the equivalent number of random logic circuits [solving Equation (8.11) for g_{SRAM}] is approximately 749.

Now applying Equation (8.9) with the total number of equivalent logic gates, G, equal to 15,992, the estimated total number of package leads can be estimated from Equation (8.10):

$$N_{IO} = 1.9(15{,}992)^{0.5} \cong 240 \text{ leads} \tag{8.12}$$

Thus, an MCM package with approximately 60 leads on a side can be used.

In this example, the maximum number of wires crossing a horizontal line can be estimated. It is assumed that the maximum number of vertically traveling leads will occur at the coordinate point where there are $G/2$ gates above the line and $G/2$ gates below the line. The horizontal and vertical coordinates of this point will be given by $y(G/2)$ and $x(G/2)$, respectively (refer to Figure 8.22b and Figure 8.22c). To estimate the number of leads whose wires will cross this horizontal line, an imaginary box is placed around the lower portion of the circuit, as shown in Figure 8.22c. Using Rent's Rule, the number of leads, N, whose wires will exit this imaginary box is given by:

$$N = a(1/2G)^b = (1/2)^b N_{IO} \tag{8.13}$$

Also, it is assumed that the number of leads going to the package I/O pins located below this line will not affect the wire congestion in the vertical direction. Similarly, the number of package leads must be reduced by a factor of $1 - (N'_{IO}/N_{IO})$ where N'_{IO} is the number of package leads available to the $G/2$ circuits without crossing over the imaginary line $y(G/2)$. Thus, the maximum number of vertically going leads is given by:

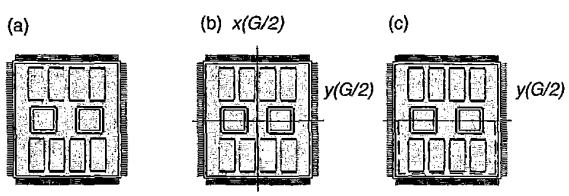


FIGURE 8.22 Circuit example for illustrating Rent's Rule as a method for estimating the number of multichip module package leads: (a) basic MCM; (b) imaginary dividing lines such that half of the logic gates fall on either side of the two lines; (c) imaginary box incorporating half of the total number of logic gates.

Samsung, EX1028, p. 179

$$N_h = (1/2)^b N_{IO} \left(1 - \frac{N'_{IO}}{N_{IO}}\right) \tag{8.14}$$

The number of wires crossing that horizontal line is W_h given by:

$$W_h = aN_h \tag{8.15}$$

In this example, the number of package leads, N'_{IO}, available without crossing the $y(G/2)$ line is 120. Therefore, assuming an average of 1.9 leads per the maximum number of vertically traveling wires that cross the horizontal line $y(G/2)$, the number of wires is estimated to be 161. By symmetry, in this particular example, the maximum number of horizontal wires crossing the vertical line $x(G/2)$ would also be 161. If the MCM substrate is approximately 1.5 in. × 1.5 in. and the line-to-line pitch is 10 mils, then the wiring capacity is 150 wires per layer, neglecting vias and keep out areas. However, assuming a 50% routing efficiency, the actual wiring capacity is 75 lines per layers. Thus, for a total of 322 wires, at least five signal layers would be required to route the design.

For signal wiring even more layers need to be added than those suggested by the above analysis. Conductors to distribute power and ground planes to control impedance must be included. A rule of thumb might be to add a power layer and a ground layer for every 4–5 signal layers, with all MCMs having at least one power layer and one ground layer. Other requirements (electromagnetic interference shielding, impedance control, etc.) may dictate the use of even more ground planes. Increasing levels of integration place greater demand on the number of layers needed in all MCM technologies. Most of the MCM technologies have been able to cope with such demands by increasing the number of layers. Increasing the number of layers has inherent disadvantages, including long vias and multiple vias in series, which have high inductance as they pass through many layers. These high inductance vias produce voltage transients upon current switching, which contributes unwanted noise (ΔI noise) to the circuit. As a rule of thumb, the higher the inductance in switching circuits, the greater the noise as described in Chapter 4.

There is a certain requirement on the spacing of the wires on the substrate in relation to the spacing of the I/Os on the chip or package. If the spacing of the wires is larger than the I/O spacing on the chip, then it is difficult or impossible to fit wires in the pin pattern, thus causing a major increase in the number of board layers to get 100% connectivity. The inevitable progress of increasing chip functionality and I/O density will force the MCM substrate developer towards thin-film techniques (MCM-Ds, built-up MCM-Ls, etc.). Thin-film circuit traces are made with high conductivity metals patterned by photolithography in a manner similar to IC fabrication. Thin-film processes promise to permit the use of much narrower lines on MCM substrates. Table 8.2 illustrates the relative differences between all three MCM technologies. Clearly, thin-films present the opportunity to use much narrower lines on MCM substrate, which translates directly into smaller packaging per chip or fewer layers of wire.

As the levels of integration increase, larger pitch technologies, like MCM-L and MCM-C, will be forced to add more and more layers to accommodate the required routing and interconnection. Because of the vias and other multilayer wiring difficulties as more layers are added, each layer becomes less efficient at providing the required wiring, until adding a layer makes no increase in wiring capability. Thin-film can accommodate most current circuit requirements in, at most, a few layers. Advanced thin-films

Samsung, EX1028, p. 180

can support conductor widths and gap widths down to 5 μm, thus providing increased flexibility for reducing circuit layers, even as the complexity of ICs rapidly increases.

8.7 MULTICHIP MODULE TECHNOLOGY COMPARISONS

Some of the basic properties of MCM substrates, regardless of the process by which they are made and their resulting structure, are dielectric constant: wiring pitch, wiring density, trace or wire conductivity, coefficient of thermal expansion, thermal conductivity, dimensional uniformity (see Table 8.2), mechanical strength, and cost. Some of these properties are included in Table 8.4 for MCM-Ls, MCM-Cs, and MCM-Ds.

The wiring density or trace density requirements of MCMs are due to the number of available circuits on each chip, the total number of chips, and the interconnectability of a given chip I/O (fan-out or fan-in) as calculated using Rent's Rule. Figure 8.23 illustrates wiring or trace density for all three MCM technologies. Clearly, most common circuits, or trace densities in the 40–400 cm/cm^2 range, can be accommodated with all three MCM technologies at the expense of using more layers with lower pitch technologies.

The *coefficient of thermal expansion* (CTE) is important in applications involving high-powered chips or modules. If the chip is bonded face-up, the major thermal sinking path is through the back of the chip into the substrate. With flip chip bonding, heat can be taken from the chip back using a heat sink or cold finger. Figure 8.24 indicates the fatigue life as a function of CTE for common MCM substrate materials. CTEs range from low CTEs perfectly matched to silicon (glass-ceramic) to large, mismatched CTEs

TABLE 8.4 Material properties of multichip modules.

	MCM-C			MCM-D		
Parameters	**Thick-film**	**HTCC**	**LTCC (Low K)**	**Inorganic Dielectrics on Si**	**Organic Dielectrics on Si**	**MCM-L Laminated Boards**
Dielectric Constant	9–10	9	5	3.6–4.0	2.4–4.0	3.5–5.0
Dissipation Factor at 1 MHz	0.0004	0.001–0.01	0.001–0.01	0.03–0.3	0.002–0.006	FR-4 0.02–0.03 Polyimide 0.01–0.015
Metallizations	Au, Cu, Pt-Au, Pt-Ag, Pd	Mo, W	Au, Cu	Al	Cr/Cu/Cr Cr/Cu/Au	Cu/Ni/Au
Coefficient of Thermal Expansion (10^{-6}/K)	7.1	7	3–5	2–3	2–3	FR-4 14–20 (x,y) Polyimide 50–70 (z) 12–16 (x,y) 40 (z)
Thermal Conductivity (W/m·K)	26	15–20	1–5	145 (Si) 0.1–1.0	145 (Si) 0.01–0.1 polyimide	1–2

Samsung, EX1028, p. 181

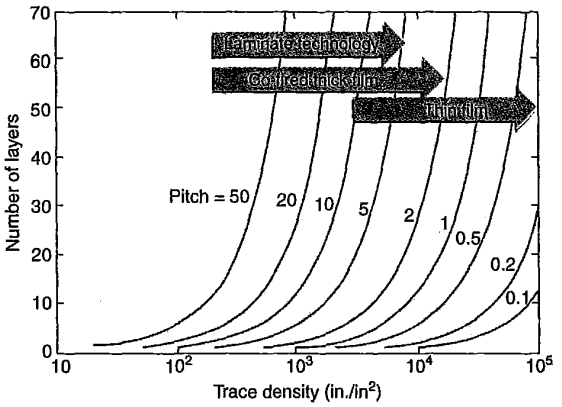


FIGURE 8.23 Number of layers versus effective trace density for various multichip module substrate technologies and trace pitches (line plus gap in mils). [The penalty factor is assumed to be less than 10%.]

such as that of epoxy-glass (FR-4) circuit board. As discussed in Chapter 5, the use of organic underfill technology between silicon chips and organic substrates has allowed an acceptable fatigue life.

MCMs have demonstrated high wiring density and superior electrical performance, but the high cost of fabrication and lack of infrastructure for handling bare die have

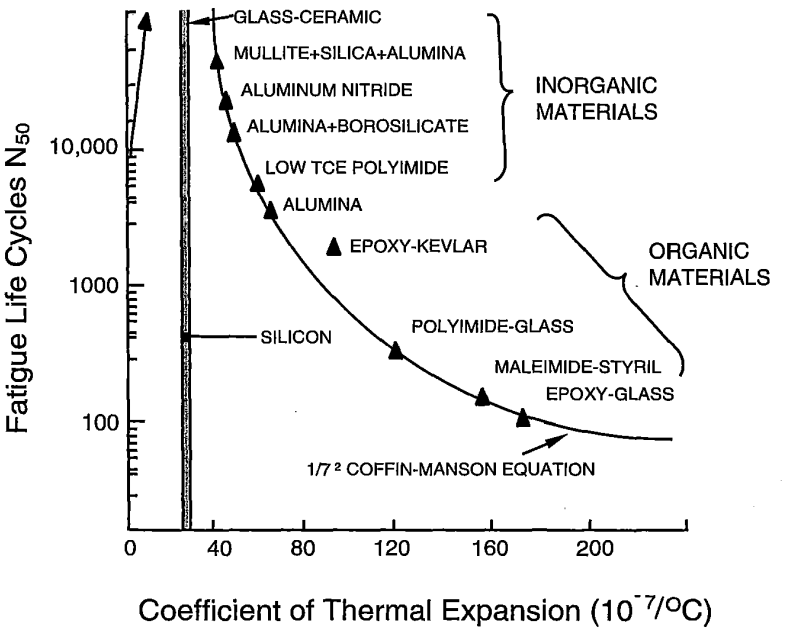


FIGURE 8.24 Fatigue life for MCM solder joints on various substrate materials versus the material's coefficient of thermal expansion.

Samsung, EX1028, p. 182

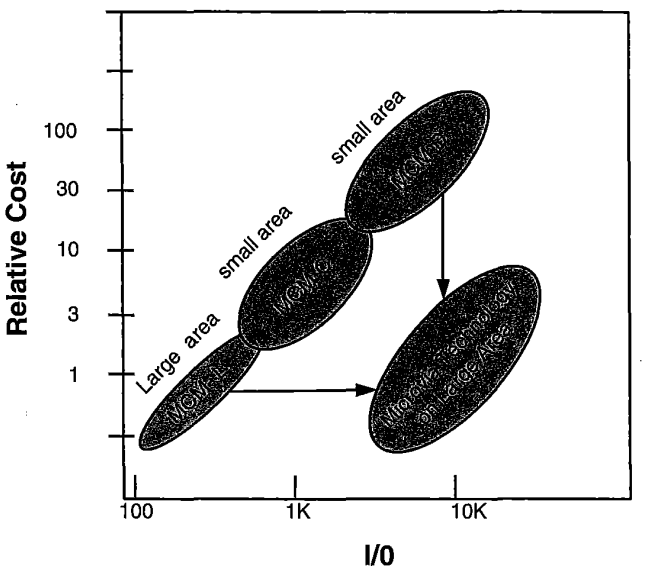


FIGURE 8.25 MCM relative cost versus module I/O for the various MCM technologies.

limited their applications. The high cost of MCMs is illustrated in Figure 8.25. The cost is almost linearly proportional to wiring density, with MCM-D being the most expensive, and MCM-L the least. The high cost of MCM-Ds is due primarily to the fact that MCM-D substrates are manufactured on small area wafers (a few up), and the demand is low. This cost, however, can be shown to dramatically drop as a function of both the size of the MCM panel, containing multiple patterns, fabricated as well as the manufacturing demand, as shown in Figure 8.25. The fundamental reason for expected low cost from large area processing is analogous to the wafer-IC strategy in that all individual MCM substrates on the panel do not have to be 100% functional. The merger of MCM-D processes with large area processing associated with MCM-Ls (PWB-like) is emerging as a very cost-effective technology. This technology is also called *microvia technology* when currently applied to BGA and chip scale packaging (CSP) applications. Microvia technology is already a $2B business.

8.8 ALTERNATIVES TO MULTICHIP MODULES

MCMs were defined as single substrates capable of providing integral signal interconnection, power, thermal management, and protection for two or more chips that are directly mounted on their surfaces. These substrates provide very high packaging efficiencies, as previously shown in Figures 8.4 and 8.5, and very high electrical performance as indicated in Table 8.1. The bare chips can be interconnected by wirebonding, TAB, or flip chip techniques, with flip chip being the preferred connection method for high performance circuits, due to its low resistance, capacitance, and inductance.

CSPs, as presented in Chapter 7, are defined as integrated circuit packages that are no more than 50% larger in area than the bare chip itself, as given by:

Samsung, EX1028, p. 183

$$A_{\text{package}} \leq 1.5A_{\text{chip}} \tag{8.16}$$

If the chip and package are square, then the length of a package side compared to the length of the chip side is about 20% greater. If CSP were to be used on an MCM substrate structure, they can meet the same requirements as an MCM at the system packaging level, without the need to handle bare chips. From Figure 8.4, the calculated packaging efficiency with CSPs is very close to those of other MCMs using bare die.

In addition, there is an emerging trend of wafer scale packaging (WSP), or wafer scale CSP, in which packaging is done directly on the integrated circuit wafer prior to die, or chip, separation. Some examples of WSP have included mini and micro BGAs as well as microspring CSPs. All of these are fabricated at the wafer level, similar to the original flip chip (controlled collapse chip connection [C4]). Thus, there are alternatives to MCMs using miniaturized packaged parts that can be mounted by solder reflow, thus extending surface mount technology lines into the future. Other advantages may include: compliant leads to mitigate the need for underfill; the ability to completely functional test the individual packaged part and, hence, the die; and the development of standard package styles and pitches which will facilitate board design and assembly.

8.9 SUMMARY AND FUTURE TRENDS

Modern electronic packaging is a rapidly changing field with many nuances and a growing myriad of technologies. The historic drivers for electronic packaging have changed as the electronics business becomes almost entirely dominated by commercial products. The portable/wireless revolution is driving toward small, lightweight, high-performance, reliable systems, all at costs far below those associated with historic electronic products and markets. Such trends have produced major paradigm shifts in the electronic world. As seen in this chapter, phrases such a flip chip, *direct chip attach* (DCA), ball grid array, chip scale package, and multichip modules now are just as familiar in electronic packaging vocabulary as the dual-in-line package and the chip and wire hybrid of yesteryear.

Figure 8.26 summarizes many of the packaging concepts described in this chapter and places them in a family tree-like form with some of their more familiar ancestors. It also shows how the minimal package and packageless parts have evolved from more traditional electronic packaging products. In fact, it is this contest between CSP and DCA on MCM substrates that is posing the greatest challenge for MCMs in particular, and the packaging world in general. The advent of the CSP allows SMT line utility to be extended well into this decade, without major capital reinvestment, extensive training, and facility upgrades. Many people believe that this is the correct approach. Table 8.5 highlights many of the advantages of CSP. The advantages of DCA on MCM substrates also are listed. It is clear that the MCM with DCA offers significant advantages when one considers size, performance, reliability, and cost in high volume.

The trend towards full integration of the package at the wafer level (WSP) needs to be carefully assessed as the technology and infrastructure for manufacturing grows. WSP offers the potential to further reduce the cost of packaged parts and, with some of the inherent handling and testing advantages of packaged parts, it may pose a future challenge to the MCM.

Samsung, EX1028, p. 184

FIGURE 8.26 Packaging taxonomy comparing the evolution of surface mount technology with the "packageless" or direct chip attach approach to the fabrication of microcircuits.

MEMS231688

CONFIDENTIAL BUSINESS INFORMATION

Samsung, EX1028, p. 185

TABLE 8.5 Comparison of the advantages of chip-scale packages with those of an MCM substrate using direct chip attach (DCA).

Chip-Scale Packages	MCM Substrate with DCA
Test at speed (known good die)	Smaller footprint (and height)
Burn-in at part level	Low cost in volume
Die protection	Lower weight
Packaged part procurement infrastructure	Smaller circuits and systems
Base of handling and rework	Improved electrical performance
Solder reflow assembly, extends surface-mount technology	Improved thermal performance
Package standardization	More reliable (fewer interconnects)
Can accommodate die shrinks or expansion	

8.10 HOMEWORK PROBLEMS

1. Given the relative sizes of various packaging schemes as shown in Figure 8.4, calculate the minimum size of a supporting substrate for a 25-chip square array. Assume that the chips or packaged parts can be spaced as close as 0.2 times their linear length (chip length plus interconnect length). Also assume that a perimeter of one half the linear length exists around the periphery of the substrate.
2. Referring to the problem above, which of the packaging schemes produces a true multichip module?
3. List and describe the major advantages and disadvantages of the three major MCM types.
4. Using Rent's Rule, calculate the number if I/Os required for a 10,000-gate array. Develop a curve for estimated I/O requirements for memory chips with gate counts ranging from 256K to 1,024M.
5. Find the wiring density required for a specialized signal processing MCM that contains 4 custom, 10,000-gate arrays and 8 static random access memories with an equivalent circuit count of 128 MB. Assume the chips are placed symmetrically with respect to both the vertical and horizontal center lines.
6. Calculate the thickness of ½ oz. copper printed wiring board. Why is making the metal thinner important for the fabrication of advanced MCMs?

8.11 SUGGESTED READING

1. R. J. Baron and L Higbie. *Computer Architecture*. Reading: Addison-Wesley Publishing, 1994, 26–30.
2. C. E. Bauer, K. W. Passe, and K. T. Wilson. "Good Enough Die? A Cost Model for KGD Test Decisions." *Proceedings of the 8th International Microelectronics Conference*. Omiya, 1994, 260–270.
3. O. J. Bodendorf, K. T. Olsen, J. F. Trinko, and J. R. Winnad. "Active Silicon Chip Carrier." IBM Technology Disclosure Bulletin, 1972, 256.
4. C. J. Brierley and J. P. McCarth. "The Reliability of Surface Mounted Solder Joints Under PWB Cyclic Mechanical Stresses." *Proceedings of the International Electronic Packaging Society Conference,* 1985, 312–319.
5. H. K. Charles, Jr. "Assembly Constraints in High Speed Packaging Applications." *Proceedings of the Technical Program of NEPCON*. Anaheim, 1994, 960–971.

Samsung, EX1028, p. 186

6. H. K. Charles, Jr. "Cost Versus Technology Trade-offs for Multichip Modules." *Proceedings of the 28th International Microelectronics Symposium.* Los Angeles, 1995, 553–558.
7. H. K. Charles, Jr. "Design Rules for Advanced Packaging." *Proceedings of the 26th International Microelectronics Symposium.* Dallas, 1993, 301–306.
8. H. K. Charles, Jr. and G. V. Clatterbaugh. "Thin Film Hybrids." in *ASM Handbook, Electronic Materials Handbook.* Materials Park: ASM International, 1989, 313–331.
9. H. K. Charles, Jr. and M. E. Uy. "Design and Aid for Multichip Modules." *Proceedings of the International Microelectronics Symposium.* Boston, 1994, 42–48.
10. G. V. Clatterbaugh and H. K. Charles, Jr."The Applications of Photosensitive Polyimide Dielectrics in Thin Film Multilayer Hybrid Structures." *Proceedings of the 21st International Microelectronics Symposium.* Seattle, 1988, 320–334.
11. A. S. Francomacaro, G. V. Clatterbaugh, and N. A. Blum. "High Speed Digital Circuits on Low Stress, Thin Film, Polyimide Multichip Modules." *Proceedings of the 6th SAMPE Electronics Conference,* 1992, 358–370.
12. R. A. Gardner and R. W. Nuter. "Properties of Multilayer-Ceramic Green Sheets." *Solid State Technology,* Vol. 17, 1974, 38–43.
13. P. J. Holmes and R. G. Loasby. *Handbook of Thick Film Technology.* Scotland: Electrochemical Publications, 1976.
14. J. S. Kilby. "Invention of the Integrated Circuit." *IEEE Transactions Electron Devices,* 23rd Edition, 1976, 648–654.
15. B. S. Landman and R. L. Russo. "On a Pin Versus Block Relationship for Partition of Logic Graphs." *IEEE Transactions on Computers,* Vol. C-20, 1971, 1469–149.
16. R. E. Meister, D. J. Shanefield, and R. B. Runk. "Tape Casting of Ceramics." in *Ceramic Processing.* New York: John Wiley & Sons, 1978, 411–448.
17. G. Messner. "Price/Density Trade-offs of Multichip Modules." *Proceedings of the 1988 International Microelectronics Symposium.* Seattle, 1988, 28–36.
18. L. F. Miller. "Controlled Collapse Chip Connection." *IBM Journal of Research and Development,* 1969, 239–250.
19. M. Nachnani, L. Nguyen, J. Bryan, and H. Takiar. "A Low-Cost Multichip (MCM-L) Packaging Solution." *Proceedings of the International Electronics Manufacturing Symposium,* 1993, 464–468.
20. M. G. Pecht. *Integrated Circuit, Hybrid, and Multichip Module Package Design Guidelines: A Focus on Reliability.* New York: John Wiley & Sons, 1995.
21. J. M. Petek and H. K. Charles, Jr. "Known Good Die, Die Replacement (Rework), and their Influences on Multichip Module Costs." *Proceedings of the 48th IEEE Electronic Components and Technology Conference.* Seattle, 1998, 909–915.
22 K. J. Pullitz. "Flip Chip Replacement Within the Constraints Imposed by Multilayer Ceramic (MLC) Modules." *Journal of Electronic Materials,* 1984, 29–46.
23. R. G. Rickets. "Design Tradeoffs." in *Electronic Materials Handbook,* Vol. 1, M. L. Minges, ed. Materials Park: ASM International, 1989, 18–44.
24. D. C. Schmidt. "Circuit Pack Parameter Estimation Using Rent's Rule." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. CAD-1.* 1982, 186–192.
25. B. Schwartz. "Ceramic Packaging of Integrated Circuits." in *Electronic Ceramics.* New York: Marcel Dekker, 1988, 1–44.
26. F. N. Sinnadurai, ed. *Handbook of Microelectronics Packaging and Interconnection Technologies.* Scotland: Electrochemical Publications, 1985, 7–10.
27. T. Takahashi, E. N. Rutter, Jr., E. S. Moyer, R. F. Harris, D. C. Fry, et al. "A Photodefinable Benzocyclobutene Resin for Thin-Film Microelectronic Applications." *Proceedings of the International Microelectronics Conference.* Yokohama, Japan, 1992, 64–70.
28. R. R. Tummala. "Multichip Module—A Tutorial," *Proceedings of the IEEE,* 1992.

Samsung, EX1028, p. 187

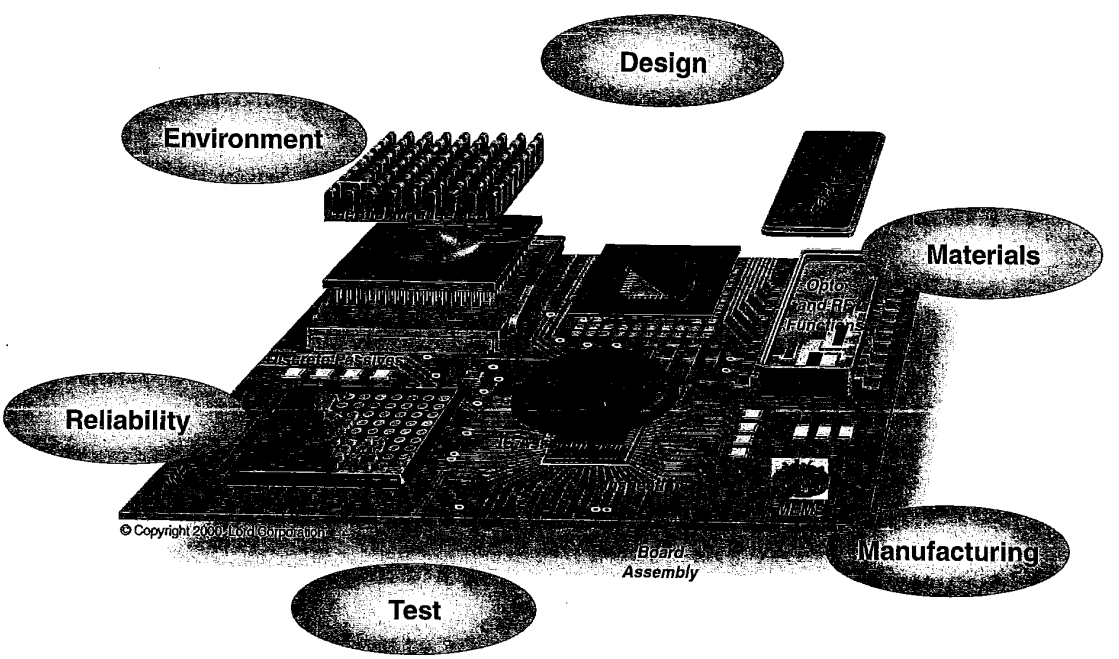
29. R. R. Tummala. "Multichip Packaging in IBM, Past, Present and Future." *Proceedings of International Conference on Multichip Modules,* 1995, 1–11.
30. R. R. Tummala, E. J. Ryamaszewski, and A. G. Klopfenstein, eds. *Microelectronics Packaging Handbook.* New York: Chapman & Hall, 1999, I-38–I-40.
31. R. R. Tummala, E. J. Ryamaszewski, and A. G. Klopfenstein, eds. *Microelectronics Packaging Handbook.* Part I. New York: Chapman & Hall, 1997, 130–195.
32. R. R. Tummala, E. J. Ryamaszewski, and A. G. Klopfenstein, eds. *Microelectronics Packaging Handbook.* Part I. New York: Chapman & Hall, 1997, 130–195.

Samsung, EX1028, p. 188

CHAPTER 9

FUNDAMENTALS OF IC ASSEMBLY

Prof. Daniel F. Baldwin
Georgia Institute of Technology



9.1 What Is IC Assembly?
9.2 Purpose of IC Assembly
9.3 Requirements for IC Assembly
9.4 IC Assembly Technologies
9.5 Wirebonding
9.6 Tape Automated Bonding
9.7 Flip Chip
9.8 Summary and Future Trends
9.9 Homework Problems
9.10 Suggested Reading

CHAPTER OBJECTIVES

- Define and describe the purpose of IC assembly
- Introduce and describe each of the three primary IC assembly technologies: wirebond, TAB, and flip chip
- Project future trends

CHAPTER INTRODUCTION

IC assembly is the most important first step in the use of an IC which will go through a number of process steps before it can be used in an electronic system. This chapter introduces and describes the three most important IC assembly technologies.

Samsung, EX1028, p. 190

9.1 WHAT IS IC ASSEMBLY?

IC assembly is the first processing step after wafer fabrication and singulation that enables ICs to be packaged for systems use. IC assembly is defined as the process of electrically connecting I/O bond pads on the IC to the corresponding bond pads on the package. The package in this case can be a single chip package, a multichip package, or a system level board. Such an assembly process involves three interfaces: (1) metallurgical bond pad interface on the IC; (2) metallurgical bond pad interface on the package; and (3) electrical interconnection between these two interfaces. Figure 9.1 illustrates the three primary interfaces.

9.1.1 Metallurgical Bond Pad Interfaces

While the metallurgical bond pad interface on the package side varies with the package type, the primary metallurgy of the IC bond pad is aluminum. Copper metallization is being introduced because of its higher electrical conductivity that translates into higher clock speeds in PCs, workstations and supercomputers.

The primary concerns at the metallurgical bond pad interfaces have to do with interdiffusion of metals into each other, the effect of the diffusion on thermomechanical reliability and electrical stability, the formation of brittle intermetallic compounds, volumetric change occurring due the formation of some intermetallic compounds (such as $AuAl_2$), and corrosion of the bond pad metallurgy. Diffusion at the macroscopic scale can be easily visualized. If a drop of ink is placed into a glass of water, the ink will spread out due to diffusion. To relate the macroscopic picture of diffusion to a microscopic description, some basic atomic diffusion mechanisms must be noted. In short, there is atomic motion and rearrangement in a crystalline solid, where atoms are arranged in an ordered pattern with translational symmetry. Two major concepts must be under-

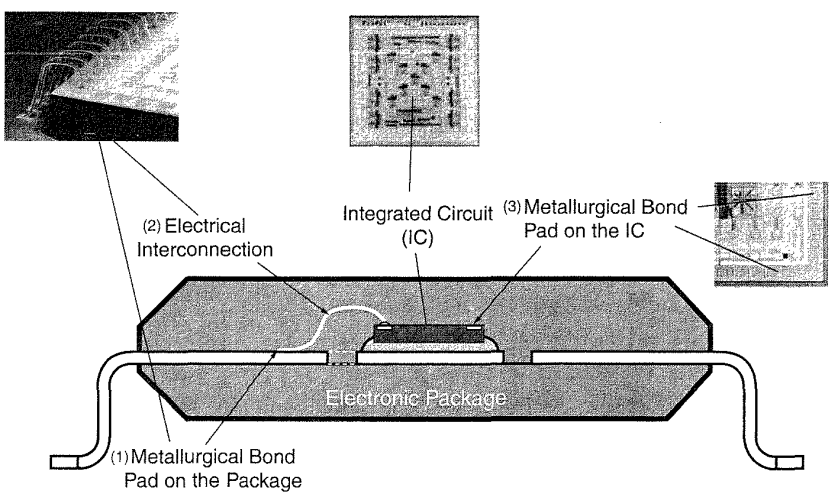


FIGURE 9.1 Three major interfaces in IC assembly.

Samsung, EX1028, p. 191

stood to clarify atomic diffusion. First, solids always contain defects due to entropy. Second, diffusion occurs in a solid when an atom changes places with a neighboring defect. Possible diffusion mechanisms include: (a) atomic diffusion where atoms jump into neighboring vacant lattice sites; (b) interstitial diffusion which occurs only for small atoms in the solid; (c) interstitial pushing of an atom from its lattice site to an interstitial site; (d) atomic diffusion, where two neighboring atoms swap positions directly; and (e) ring rotation of four atoms.

Interdiffusion of the bond pad metallurgical species (chip and package) and the interconnection metallization can compromise the integrity of the interface and package reliability. Hence, it is critical that an energetically stable interface is formed by judicious selection of the metallurgy selected for the interconnection and package bond pad interface. Layers of metallization often are used, each performing separate functions such as adhesion layers, diffusion barriers, and sacrificial layers to promote wetting.

The metallurgical bond pad interface on the package side is typically finished with gold metallization for nobility and bondability, but barrier and adhesion layers are also used. Details for each of the major IC interconnection techniques is reviewed below.

9.2 PURPOSE OF IC ASSEMBLY

The purpose of IC assembly is to enable an IC to be electrically interconnected to the package, so as to allow that IC to be handled, tested and "burnt in" to guarantee the quality of the IC. Such a "qualified" IC is now ready for use, together with other components, in an electronic product such as a cell phone, camcorder or a personal computer. The end product, however, requires a number of active ICs and passive components with the desired functions to be assembled on a system level board to form that end product.

In order for integrated circuits to provide useful functions, they must be interconnected to other ICs by means of organic or ceramic wiring boards to form systems. The other components include passives, other ICs, flat panel displays, keyboards, sensors, wiring harnesses, connectors, antennas, switches, etc. providing a physical interface with the surrounding environment. The primary purpose of IC assembly is to enable ICs to be interconnected with the rest of the system, which could be a cellular phone, a computer, an industrial controller, a hearing aid, an engine controller, etc.

There are three primary functions of IC assembly: (a) to provide the signal and power distribution of the packaged IC to the system; (b) provide mechanical support and robustness to the fragile IC; and (c) provide for environmental protection of the IC. Each of these functions is accomplished using a series of material systems and manufacturing processes.

9.3 REQUIREMENTS FOR IC ASSEMBLY

There are five primary requirements for IC assembly:

1. To provide *acceptable electrical properties,* including capacitance, resistance, and inductance. Each IC assembly has unique electrical characteristics, along with mechanical characteristics and manufacturability. For example, wirebonds have long lengths and parallel proximity, resulting in high impedance and longer signal delay times.

Samsung, EX1028, p. 192

2. IC assembly technologies should provide a *low cost* solution for the electrical interface between the chip and package. Tape automated bonding is not typically a low cost interconnection technique. It requires specialty tooling, equipment, and circuit tapes for production significantly increasing cost.
3. The third requirement is *high throughput manufacturing*. For finer pitch packaging applications, tape automated bonding provides a low cycle time, high throughput IC assembly approach.
4. The fourth requirement is *high reliability*. Flip chip on ceramic technology has been a highly reliable interconnection technique. For instance, IBM has used this IC assembly technology for over 40 years with no reported field failures due to thermal cycle fatigue.
5. The final requirement is *repairability or replaceability* where the interconnection between the IC and package should provide for removal of a failed IC, and replacement with a new high quality IC or a new part number.

9.4 IC ASSEMBLY TECHNOLOGIES

Chip to package interconnection can be accomplished using three primary interconnection technologies as illustrated in Figure 9.2. These include wirebonding, *tape automated bonding* (TAB), and flip chip. Wafer-level technology is covered in Chapter 10.

9.5 WIREBONDING

9.5.1 Concept

Wirebonding is a chip-to-package interconnection technique where a fine metal wire is attached between each of the I/O pads on the chip and its associated package pin, one

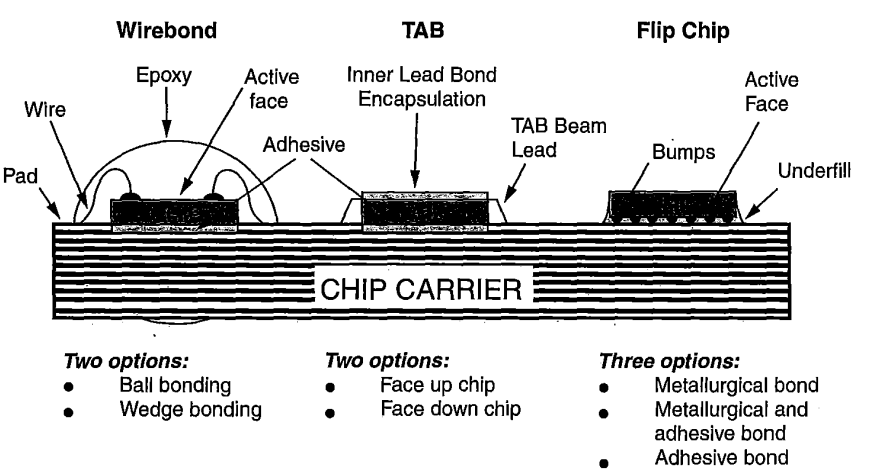


FIGURE 9.2 Chip to package or substrate interconnection techniques.

Samsung, EX1028, p. 193

at a time. This technology originated with AT&T's beam lead bonding in 1950s. In this technology, a fine wire—typically gold wire 25 μm in thickness—is bonded using ultrasonic bonding between the IC bond pad and the matching package or substrate bond pad. The flexibility derived from this point-to-point process is one of the major advantages of wirebonding. This is referred to as compliancy. Wiring changes are simply modifications to a data-driven machine program requiring no hard tooling or package design changes, such as new tapes for TAB packages, saving time and money. Wirebonding accounted for well over 90% of all the chip-to-package interconnections formed in 1999. Figure 9.2 shows a typical wirebond interconnection structure for plastic packaging, which is the dominant packaging configuration. There are two major types of wirebonds used in high volume production: gold *ball bonding,* which comprises over 95% of all wirebonds formed; and gold or aluminum *wedge bonding.* The remaining sections will focus on ball bonding, since it dominates this technology. Brief references will be made to wedge bonding and other wirebonding techniques.

The advantages of wirebonding are:

- Highly flexible chip-to-package interconnection process
- Low defect rates or high yield interconnection processing (40–1000 ppm)
- Easily programmed or taught bonding cycles
- High reliability interconnection structure
- Very large industry infrastructure supporting the technology
- Rapid advances in equipment, tools, and materials technology

The disadvantages of the wirebonding interconnection include:

- Slower interconnection rates due to point-to-point processing of each wirebond
- Long chip-to-package interconnection lengths, degrading electrical performance
- Larger footprint required for chip to package interconnection
- Potential for wire sweep during encapsulation overmolding

9.5.2 IC Bond Pad Interface

The predominant IC bond pad interface for wirebond interconnection is aluminum metallization with which the IC is fabricated. No other metallurgical interface is necessary. In order to increase the performance of ICs, copper metallization is being introduced as a wiring material for IC metallization. This requires further refinements to the wirebonding process for robust interconnection yields.

9.5.3 Package Bond Pad Interface

The package bond pad interface for wirebond interconnection can consist of a metallized lead frame, a metallized chip carrier—organic laminate, polymer film, or ceramic—or a metallized printed circuit board. In the case of lead frames, the most common lead frame materials are copper alloys (e.g., Cu-Fe, Cu-Cr, Cu-Ni-Si, and Cu-Sn), and in some instances Fe-Ni or Fe-Ni-Co alloys are also used. These precision rolled and stamped/etched films are electroplated in a reel-to-reel process for the final package bond pad

Samsung, EX1028, p. 194

interface metallization. The most common metallizations used are nickel, gold, or silver. In the case of chip carriers and PWBs, the most common conductor is copper, which is plated with nickel, and soft or immersion gold for good wirebondability.

9.5.4 Overall Processes

To understand wirebonding, it is best to put it into perspective in terms of the most common packaging configuration in which it is used: plastic packaging. Plastic packaging comprises a major portion of electronics packaging used worldwide. The basic process flow is shown in Figure 9.3. The lead frame which is the primary package metallurgical bond pad interface, or interconnection structure, for plastic packages is formed by a stamping and plating operation. Next, the chip is mounted onto the lead frame or chip carrier using a die attach process. The die attach process is highly automated and uses high precision, high speed chip bonders. Die attach is achieved using a variety of materials including conductive epoxies and solders. Next, fine wires are bonded between each chip I/O and the package lead frame pads or chip carrier traces. Thermosonic and ultrasonic welding methods are the most commonly used for wirebonding. The first wirebonding technique, thermocompression bonding, developed by Bell Laboratories, is now rarely used. Next, the package is overmolded with a polymer encapsulant to provide mechanical support and environmental protection, while presenting a cooling surface for heat transfer. Molding is accomplished using a transfer molding or injection molding process. Next, the packages are singulated from the lead frame carrier, and the leads are bent to form the package leads. In the case of area array packaging, the chip carriers are solder balled. J lead or gull wing leads are common for lead frame packages. The lead frame carrier or chip carrier pallet, configured in a multiple up format, is next subject to functional test and burn-in testing.

Ball Bonding: The most common wirebonding technique, capturing over 95% of all wirebonds formed, is ball bonding using a thermosonic welding technique. A major advantage of ball bonding is that the capillary bonding tools are round, such that the wire can be paid-out or pulled through the capillary at any angle radiating from the ball, enabling simple high speed x-y motion of the bonding head. The basic steps of thermosonic ball bonding are shown in Figure 9.4. The process sequence consists of forming a ball bond with the IC's I/O pad. This is done under a controlled capillary bonding force, typically, less then 100g, moderate temperatures of 150–200°C, and ultrasonic excitation of the capillary and wire between 60–120 kHz. Next, the wire is paid-out of the capillary to form the wire loop. The crescent bond is now formed between the wire and the package lead frame or chip carrier trace. The wire is then paid-out, the tail

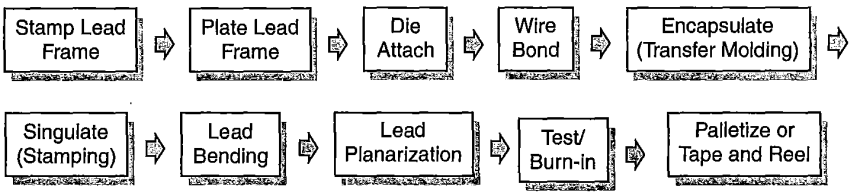


FIGURE 9.3 Processing steps for construction of a wirebonded plastic package.

Samsung, EX1028, p. 195

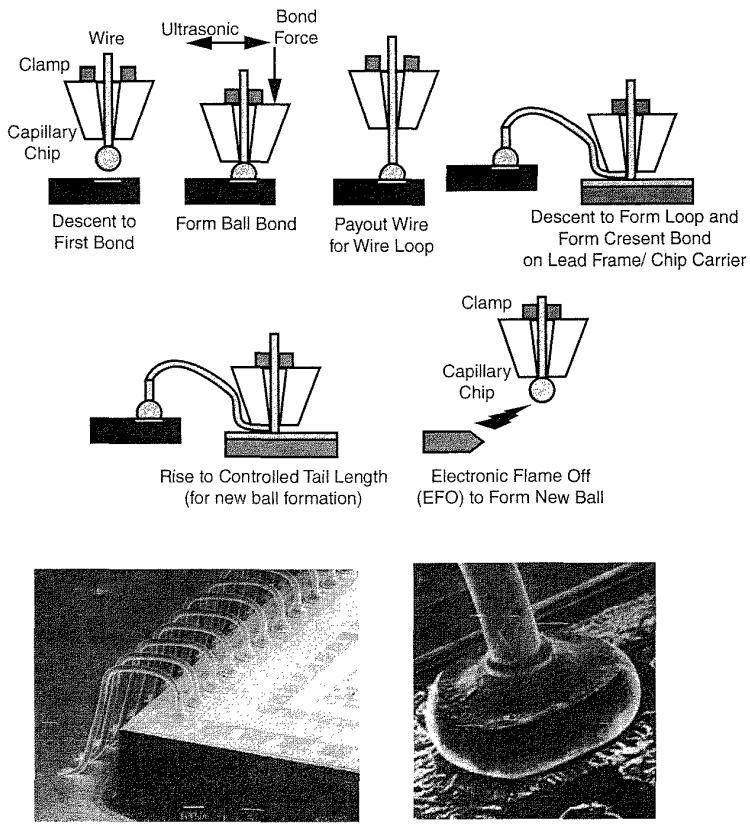


FIGURE 9.4 Thermosonic ball bonding process and illustrative pictures of ball bonds.

fractured, and the *electronic flame off* (EFO) fires to form the ball for the next bond. Process cycle times for a full bond cycle can be less than 20 ms for thermosonic bonding.

Prior to initiating the bond sequence, the die and lead frame pattern are located in space, using a high precision vision system. This enables the bonder to calculate the bonding cycle tool paths to a high degree of accuracy. In some advanced packaging applications such as BGAs and CSPs, the wirebonds must have multiple bends rather than a single loop. This requires the capillary and bond head to move in a controlled fashion in x, y, and z axis.

Wedge Bonding: Wedge bonding has consistently demonstrated the finest pitch bonding capabilities, largely due to the fact that the bonds can be formed by deforming the wire only 25–30% beyond the original diameter, compared with 60–80% for ball bonds. Moreover, wedge bond yields are often higher than ball bonding yields. The major disadvantage to wedge bonding is that the bonding head must be able to provide a precision rotation, such that the bonding tool (i.e., wedge) is aligned along the axis between the chip pad and lead frame or chip carrier pad. It is necessary to provide this radial fan out of the bonded wires to prevent buckling and wire fracture during the bond cycle.

Samsung, EX1028, p. 196

The process sequence consists of moving the clamped wire in the bonding tool into contact with the IC I/O pad. The chip side wedge bond is then formed. The ultrasonic weld is formed under a controlled wedge bonding force, moderate wire temperatures—for gold wire 125–150°C is used—and ultrasonic excitation of the capillary and wire between 60–120 kHz. Next, the wire is paid-out of the capillary to form the wire loop by raising the wedge. The wedge is now positioned over the package side bond site and a wedge-off bond is formed between the wire and the package lead frame or chip carrier trace. Again, ultrasonic energy and a controlled bond load is applied to form the wedge off bond. With the wire clamped, the wedge is pivoted forward to fracture the wire at the heel of the wedge-off bond and the bonding cycle continues with a new site. Process cycle times for a full bond cycle can be less than 80 ms for ultrasonic bonding.

Fundamentals of Ultrasonic Bonding: While the fundamental mechanisms for ultrasonic and thermosonic bonding are not completely understood, a number of theories have been proposed and are finding acceptance. In one such theory, the ultrasonic energy applied through the bonding tool, called a capillary or wedge is thought to increase the dislocation density of the wire and bond site metallization. As the density of dislocations in the wire and bond site microstructure increases, the modulus of elasticity and the plastic flow stress of the wire decrease, and the intermolecular diffusion between the wire and bond pad microstructures increases. The ultrasonic energy allows the materials to plastically deform at much lower stress compared to pure thermal or mechanical energy.

As the material plastically deforms, or flows, at the interface between the wire and bond site metallization, microscopic slip planes form and shear across one another. As the slip planes slide over one another, new surfaces of metal are formed at the interface that are clean, free of oxides, and highly energetic. These metallurgically clean surfaces, in contact at the bond site and wire interface, promote the solid state diffusion of the molecular species across the interface boundary. A diffusion weld forms between the wire and bond metallization. With continued exposure to ultrasonic energy, the bond areas increase. The addition of thermal energy enhances the diffusion and solid state bonding. Higher ultrasonic frequencies increase the strain rate, and enable the material to more efficiently transfer energy from the capillary through the wire to the bond interface.

Materials Used in Wirebonding: Bonding Wires: The most common bonding wire materials are gold for ball bonding and aluminum for wedge bonding. A typical wire diameter is 25 μm. It is important to specify the properties of the wires used in bonding. Typical properties specified include *break strength* (BS) and *elongation* (EL), based on standard stress–strain curves. Gold wire is typically alloyed with small amounts of beryllium (Be) and Ca (5–10 ppm by weight) to improve the drawability of the soft gold (i.e., 99.99% pure). Gold wires are heat treated post drawing to increase ductility and soften the wire. Aluminum wire is typically alloyed with silicon (1% Si) or magnesium (0.5–1.0% Mg) to promote drawing of the wire to fine diameters (<50 microns). A disadvantage of the Al-Si wire is that at room temperature the solubility of Si in Al is lower than that of the alloy (0.02% at 20°C), resulting in the precipitation of Si out of the wire at the bond, forming a solid phase of silicon. The presence of the silicon nodules can adversely affect reliability.

Die Attach Materials: Solders, conductive adhesives, and glass adhesives are the material systems most often used for die attach in wirebonded packages. Solder die attach

Samsung, EX1028, p. 197

is advantageous because of its high temperature properties, good fatigue, and good creep resistance. Disadvantages of solder chip attach are the higher stresses on the chip resulting from thermal expansion mismatch and the lack of plastic flow. Solders used include single alloy, binary alloy, or ternary alloy compositions.

Polymer adhesives (e.g., epoxies or polyimides) filled with noble metals have found wide use as die attach materials in low-cost packaging applications or plastic packaging. They are bonded at lower processing temperatures, reducing stress and providing a thermally and electrically conductive path from the chip to the leadframe. Significant disadvantages of polymer adhesive die attach materials are that they are not thermally stable and absorb moisture, hence cannot be used in hermetically-sealed packages. While conventional materials can take hours to cure, recent developments in adhesive chemistry have resulted in fast-cure, or snap-cure, epoxies capable of curing in a few minutes or less.

Silver-filled specialty glass materials have also been used for die attach. These materials are thermally stable and can provide a void-free bond improving reliability and thermal management. They require high bonding temperatures (~400°C), oxidizing environments for good adhesion, and solvents leading to several processing concerns.

9.5.5 Electrical Performance

Of all of the chip-to-package interconnection types, the electrical performance of wirebonds is the lowest. This follows largely from the relatively long lengths of the wires interconnecting the chip and package lead frame. This tends to increase impedance, promote inductive coupling, and slow package operating speeds. The lower electrical performance has limited the use of wirebond interconnection in high speed applications over the last several years. Applications such as microprocessors, high speed ASICs, high speed memory, high speed RF and analog have tended to implement flip chip and TAB interconnection solutions, due in part to their higher performance.

9.5.6 Reliability

Wirebond interconnection structures tend to be very reliable. Wirebond interconnections are used in a very wide range of applications including very high reliability products—medical applications, space applications, automotive applications and aerospace applications—and low cost commodity electronics such as toys, smart cards, RF tags, and AM-FM radios. A common failure in wirebonded plastic packages tends to occur due to delamination of the encapsulant molding compound or die attach, followed by highly localized stress concentrations in the wirebonds, causing fatigue failures in the wire or bond. A number of other failures can occur in wirebonded packages, several of which are discussed below.

Chip Fracture: Failure can occur due to chip fracture. Surface defects on the semiconductor, predominantly single crystal silicon, significantly lower the fracture strength of these brittle materials. These initial microcracks can be caused by dicing along the edge of the chips, back side grinding/polishing defects, and die pick marks produced by the die bonding process. The residual stresses resulting from die attach and the coefficient of thermal expansion (CTE) mismatch in the packaging materials, can lead to a cata-

Samsung, EX1028, p. 198

strophic device failure. In this case, the residual stresses, in combination with the thermomechanically induced stresses, can exceed the effective fracture strength of the silicon, causing die cracking along the crystallographic plane of maximum principle stress.

Chip Passivation Cracking: Cracking of the chip passivation can occur due to the shear stresses imposed by the mold compounds. Passivation layers are usually made of brittle glass films such as silica (SiO_2), phosphosilicate (PSG), or silicon nitride (Si_3N_4). A larger distribution of passivation cracks tend to be found near the edges of the device, as expected from the residual stress patterns on the molded chip. Often, passivation cracking is associated with ball bond lift-off and shearing. This follows due to the close proximity of the I/O bond pads to the passivation openings and high stress chip edge. Often times, polymer passivation materials, such as polyimide or BCB, are used as stress buffers over the glassy passivations to help minimize cracking.

Chip Metallization Corrosion: Corrosion of the chip metallization can result in various electrical failures, including current leakage, opens, and shorts. Several factors contribute to the corrosion of the chip metallization. Since aluminum is the most common metallization, the remaining discussion will focus on this system. Typical factors affecting corrosion include system acid level, metal composition, encapsulant material, passivation glass, ionic contamination, temperature, absorbed moisture, and applied voltage. A major driving force for corrosion is the presence of ionic contaminants, largely resulting from the encapsulant molding compound, and absorbed moisture. In this case, the moisture diffuses through the molding compound, or along failed interfaces, reacting with chlorine and other ionic contaminants, resulting in corrosion of the metallization. Galvanic corrosion occurs under the presence of applied voltage. Minimizing the exposed chip I/O bond pad reduces the likelihood of corrosion. Unfortunately, wirebonding typically leaves exposed bond pad metallization. Advances in low ionic content electronics grade polymers has significantly reduced corrosion failures ion non-hermetic packages.

Wire Sweep: Wire sweep is the deformation of wirebonds resulting from the flow of viscous polymer molding compounds during encapsulation, typically a transfer molding process. Wire sweep tends to occur on wires oriented perpendicular to the polymer flow direction in the mold. Wire sweep is detected using x-ray analysis and is evidenced by kinks at the attachment point, typically near the gate where the polymer enters the mold cavity. Process factors affecting wire sweep include polymer viscosity, local high flow rates, due to shear thinning of the polymer, void entrapment, and filler particle collision with bond wires.

Cratering of Wirebond Pad: Cratering can occur during the wirebonding process and is a fracture of the silicon under the chip I/O bond pad. It results from the improper setup of the bonding parameters, such as temperature, ultrasonic power, bonding force, and bond time. The ultrasonic energy applied can result in dislocation formation, and stacking faults in the single crystal silicon in the direction perpendicular to the bonding tool motion, lowering the fracture strength of silicon. Thin metallization layers tend to promote cratering. Micron-size silicon nodules in the aluminum bond pads can also act as stress risers, leading to the rapid formation of cracks during bonding. Approaches to minimize cratering are using higher temperatures (~250°C), lower ultrasonic power, ramping the applied ultrasonic amplitude, reduced molding stress, and removing any brittle phosphorus glass layer under the bond pad.

Bond Fracture and Lift-Off: The weakest point in wirebonds tends to be at the heel of the crescent bond and at the neck of the ball bond. Cracks in the heel of the bonds

Samsung, EX1028, p. 199

can arise as a result of flexure of the wire loop during formation and temperature cycling. Fracture of the wire at the neck occurs as a result of wire embrittlement due to interdiffusion of metallic compounds and formation of inter-metallics. Thallium contamination is particularly known to cause neck embrittlement. Finally, bond lift-off and package failure occur as a result of either tensile or shear forces induced by thermal stresses, or the flow of encapsulants during molding.

Interfacial Delamination: Interfacial delamination in plastic packages tends to result from poor adhesion between the molding compound and the chip, the chip passivation, the bond wire, and the lead frame. Residual stresses and thermomechanical stresses promote the propagation of the interfacial microcracks and delamination. Moisture tends to aggravate delamination. Under conditions of good adhesion, moisture tends to diffuse through the bulk encapsulant. If poor adhesion exists, moisture tends to diffuse along the interfacial boundary. This can cause the polymer at the interface to hydrolyze, causing a further reduction in interfacial adhesion of the polymer to the lead frame, etc. Delamination tends to result in highly localized stress concentrations and potential wirebond fatigue fracture.

Package Cracking: Package cracking can occur due to fatigue fracture of the encapsulant, and due to a phenomenon called "popcorning." Fatigue fracture of the polymer encapsulant results from the mismatch in thermal expansion caused by thermal gradients and rapid temperature changes. Popcorning results from the internal stresses produced at high temperatures and rapid thermal excursions during reflow or high temperature operation. The thermal mismatch between the encapsulant, lead frame, and chip can produce interfacial stresses large enough to promote interfacial delamination. Hydrostatic stresses, due to polymer swell from absorbed moisture, produce further stresses. At high temperatures, the encapsulants are typically above their glass transition temperature, further weakening interfacial adhesion and making the system more susceptible to delamination. Absorbed moisture further exaggerates the situation to volatilize water-vapor, resulting in nucleation of voids at the interface. The resulting voids increase the interfacial stresses due to pressure in the voids, further promoting delamination. Common locations for popcorning are at the encapsulant-lead frame interface and the encapsulant-chip surface. The most common location for these types of cracks is at the chip paddle to the bottom of the package.

9.5.7 Production Examples

Wirebond interconnection is used in all forms of plastic packaging, multichip modules, hermetic packaging, and ceramic packaging. Examples of plastic packages that are wirebonded to lead frame structures or chip carriers include thin small outline packages (TSOPs), plastic quad flat packs (PQFPs), ball grid arrays (BGAs), chip scale packages (CSPs), plastic leaded chip carriers (PLCCs), small outline ICs (SOICs), and small outline transistors (SOTs).

9.6 TAPE AUTOMATED BONDING

9.6.1 Concept

Tape automated bonding (TAB) is an IC assembly technique based on mounting and interconnecting ICs on metallized flexible polymer tapes. It is based on the fully auto-

Samsung, EX1028, p. 200

mated bonding of one end of an etched copper beam lead to an IC, and the other end of the lead to a conventional package or PWB. TAB was invented by Frances Hugle in Silicon Valley and commercialized in 1966 by General Electric Research Laboratories (New York, NY), who used it with *small-scale integration* (SSI) devices. The goal of this new technology introduction was a lower cost replacement of wirebonding technology by providing a highly automated reel-to-reel "gang bonding" technique for packaging high-volume, and low I/O devices. Through the 1970s, TAB received strong consideration and attention but experienced relatively little industry acceptance, except in Japan. Japan had the largest number of companies to commercialize TAB. TAB saw its most widespread adoption in the 1980s as *surface mount technology* (SMT) flourished. With its ability to handle high-density *input/output* (I/O) and high speed circuitry of *very-large-scale-integration* (VLSI), TAB has been applied to a variety of consumer, medical, security, computer, peripheral, telecommunication, automotive and aerospace products. Its principal use is as driver connections on LCD panels.

Some of the advantages of TAB include:

- Ability to handle small bond pads and finer pitches on the IC
- Elimination of large wire loops
- Low profile interconnection structures for thin packages
- Improved conduction heat transfer for thermal management
- Improved electrical performance
- Ability to handle high I/O counts
- Ability to burn-in on tape before device commitment
- Reduced weight

Some of the disadvantages of TAB include:

- Basically a peripheral interconnection technique with no active circuitry under the chip bond pads
- Package size tends to increase with larger I/O counts
- Process inflexibility due to hard tooling requirements of the flex circuit, bond heads, etc.
- Relatively little production infrastructure
- Additional wafer processing steps required for bumping
- Gang bonding no longer useful for the larger chips; replaced by single point thermosonic bonding
- Large capital equipment investment required
- Difficulty in assembly rework
- Additional engineering requirements
- Specialty materials and equipment requirements
- Coplanarity of the beam leads with the chip and substrate
- Long parallel interconnections with poor electrical performance
- System testability

An example of the TAB package with inner lead bonds and an outer lead bonding frame is shown in Figure 9.5.

Samsung, EX1028, p. 201

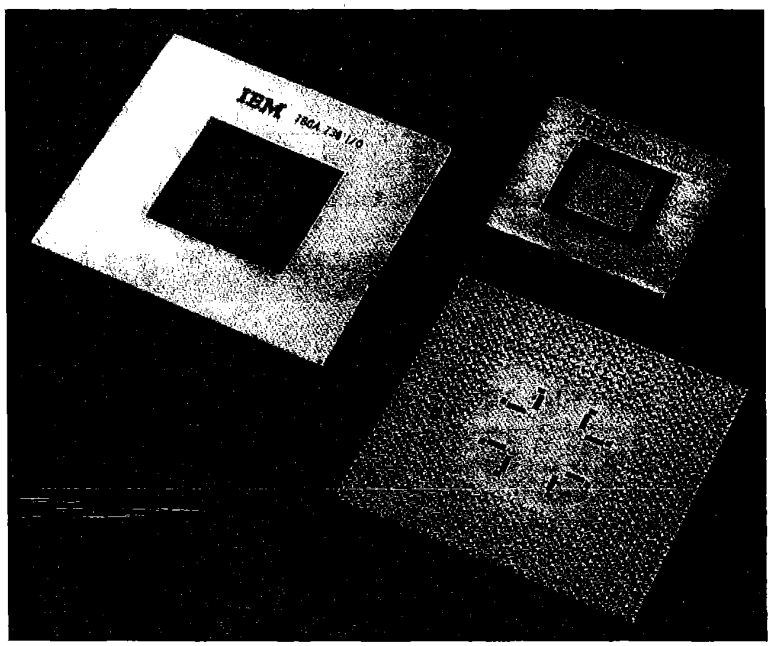
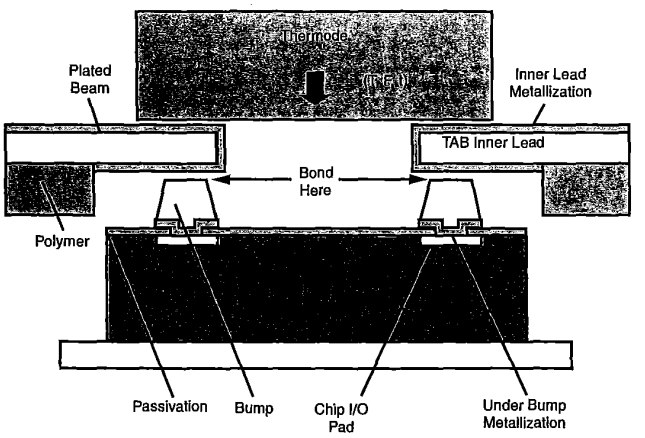


FIGURE 9.5 Example of a TAB package and a polymer tape. (*Source: IBM*)

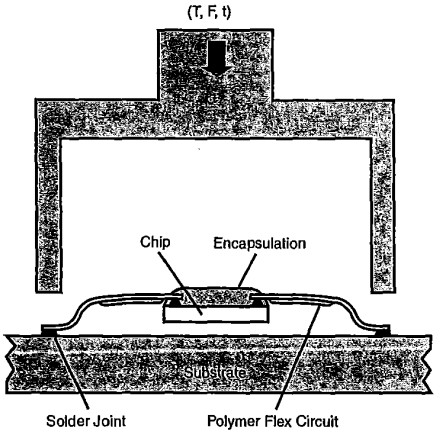
9.6.2 IC Bond Pad Interface

To form a TAB interconnection, an IC is first bumped to provide the necessary bonding metallurgy for inner lead bonding shown in Figure 9.6a. This bump acts as a physical stand-off between the chip and circuit tape to prevent lead-chip shorting, protect the chip's thin-film Al I/O bond pad metallization from corrosion and contamination, and provide a deformable, ductile buffer for the bonding process. The bumping consists of deposition of an *under bump metallurgy* (UBM) on the metallized I/O pads of the chip, shown in Figure 9.6a. The UBM includes an adhesion metal layer pattern deposited on the final IC metallization, typically aluminum or copper. This layer insures a high strength interface and low contact resistance between the chip metallization, and the bump. Next, barrier metal layers are deposited over the adhesion layer. The purpose of barrier layers is to prevent diffusion of metal atoms from subsequent metal layer deposits into the chip I/O pad metallization, resulting in brittle intermetallic compound formation, and to prevent diffusion of ionic contaminants from encapsulant compounds, resulting in corrosion of the interconnection system. Both of these phenomena decrease the reliability of the interconnection system. Common adhesion and barrier layer metals include thin layers (about 200 nm) of titanium, tungsten, titanium-tungsten, copper, molybdenum, nickel, platinum, palladium, and chromium, deposited using evaporation or sputtering processes. Next, the bump metallurgy is pattern deposited onto the barrier layers. The most common bump materials are gold, copper, and solder (Sn/Pb), deposited using an electroplating

Samsung, EX1028, p. 202



(a) Inner Lead Bonding



(b) Outer Lead Bonding

FIGURE 9.6 Tape automated bonding elements: (a) TAB inner lead bonding structure and materials; (b) TAB outer lead bonding structure and materials.

process. Post bumping, the wafer is commonly annealed at elevated temperatures to reduce the as-plated hardness of the bumps to a level suitable for inner lead bonding.

9.6.3 Package Bond Pad Interface

As shown in Figure 9.6a, the package bond pad interface is typically comprised of a metallized inner lead from the TAB tape. To promote a high yield interconnection via ultrasonic bonding for the inner leads, the final metallization on the inner leads is "soft"

Samsung, EX1028, p. 203

gold, which is deposited, using an electroless or chemical plating process, also called immersion gold. The gold metallization is deposited over a nickel metallization, or directly on the copper conductor of the TAB tape. Deposition of gold using an electrolytic plating process results in a "hard" gold deposition, which is not considered to be ultrasonically bondable. The immersion gold layer must be free of pin holes to prevent surface oxidation formation due to the underlying metal layers. A second package to bond pad interface metallization used is tin, which is electroplated over nickel, which is electroplated over the copper conductor of the TAB tape. During thermocompression bonding of the gold bump on tin leads, studies have shown that the liquid tin and solid gold form a gold rich Au-Sn eutectic intermetallic compound, hence forming a metallurgical bond between the chip and tape conductors. While other metallurgical systems are also used for the package to bond pad interface in TAB, gold and tin comprise the majority of the applications.

9.6.4 Structure and Processes

The basic structure and fabrication process for TAB interconnection systems is shown in Figure 9.7, illustrating a typical bumped chip and planer polymer circuit tape configuration. The principal system for TAB has been Sn-plated Cu beams on Au bumps. The heated thermode, applied with pressure, creates tiny nuggets of AuSn eutectic intermetallic at the interface for the metallurgical bond. Alternately, Au over Ni on Cu beam is thermocompression bonded. More recently, there has been thermosonic welding of the

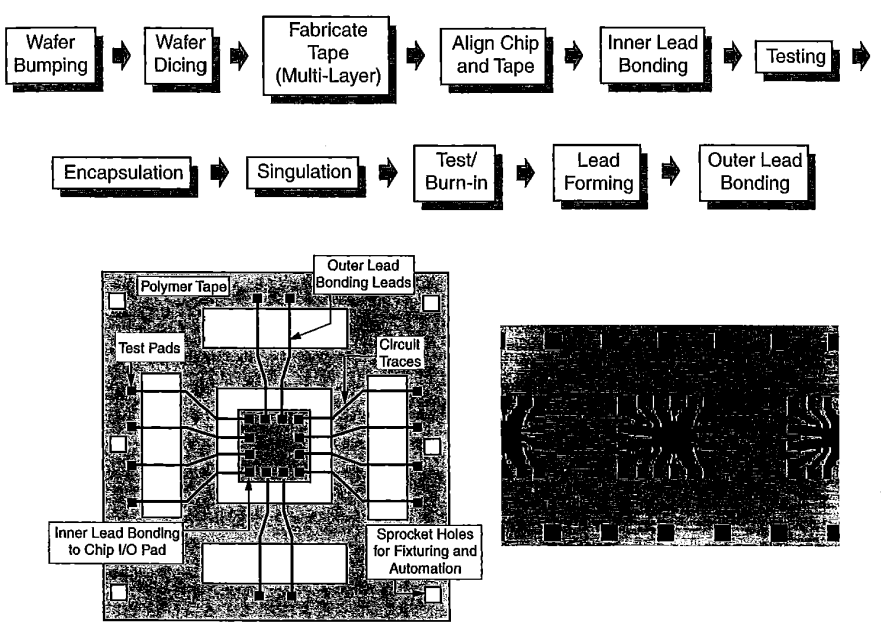


FIGURE 9.7 Typical TAB fabrication process and schematic.

Samsung, EX1028, p. 204

TAB tips with a single point thermosonic bonding machine in a serial fashion. This avoids over- or under-bonding when the thermode is not parallel, or when some bumps are out of planarity.

Dicing: After wafer bumping to form the IC bond pad metallurgical interface, the wafer is mounted on an elastic tacky dicing tape and diced. The tacky mounting tape is secured to a mounting ring, which is used to affix the wafer into the dicing saw. The dicing process results in the singulation of the bumped chips and is accomplished using a high speed saw blade embedded with diamond particles. Typical dicing blades are 7 cm in diameter, rotate at a speed of 30,000 rpm, and have a feed rate of 7 mm/s on cutting passes at a blade height of 0.04 mm.

9.6.5 Inner Lead Bonding

The TAB interconnection process proceeds with an *inner lead bonding* (ILB) operation that interconnections with bumped chips to the polymer circuit tape, as shown in Figure 9.6a. More specifically, the ILB process bonds the plated inner leads, also called *beams,* to the bumped chips, forming a mechanically robust interconnection, and electrically functional metallurgical bond between the chip and tape. Several ILB processes are used in production. The first is called *thermocompression gang bonding,* where all of the interconnections between the chip and substrate are formed simultaneously at elevated temperature and pressure, generated by a thermode. This process can be performed in a reel-to-reel format, where the TAB tape is fed via reels into an automated bonding machine which feeds the bumped chips to a bonding head; The machine then aligns the bumped IC I/O to the corresponding tape beam leads by adjusting x, y, and θ coordinates. The interconnections are formed by thermocompression bonding at elevated temperature, and pressure, and for fixed dwell times using a *thermode.* After bonding at one chip site, the tape increments to the next chip site and a new bumped chip is fed into position for bonding this site. Additional ILB processes include *single point bonding, eutectic solder hot gas bonding, laser bonding* and *laser sonic bonding.* United States companies predominantly use single point bonding, while Japanese companies typically use thermocompression gang bonding. Several factors impact the ILB process, including bump metallurgy, bump hardness, bump flatness, tape metallurgy, tape beam hardness, tape flatness, bump design, UBM, thermode flatness (over temperature range), thermode design, thermode mass, thermode heat transfer characteristics, thermode wear resistance, tape and die alignment, bump-beam interface temperature, stage temperature, thermode temperature, and bonding pressure or thermode force.

9.6.6 Polymer Circuit Tape

The *polymer circuit tape,* also called the *tape carrier* or *film carrier,* can be configured in several ways, including one layer tape, two layer tape, and three layer tape. The polymer circuit tape has a form factor similar to the common 35 mm photographic film, including tractor feed fixturing holes along the edge of the film. One layer tapes are typically 35–70 μm thick and have copper circuit traces that are etched or stamped. Two and three layer tapes are typically 50–70 μm and 75–125 μm thick, respectively. While copper is the most common conductor, aluminum, steel, alloy 42, and thick-film conductors are also used for polymer trace conductors. The polymer films used for TAB

Samsung, EX1028, p. 205

tapes comprising the primary dielectric of the package include polyimide, epoxy-glass, polyester, and BT resins, with polyimide being the most common.

9.6.7 Testing and Encapsulation

At this point, the circuit can be functionally tested using standard probe type circuit testers, followed by single or double sided polymer encapsulation. The encapsulation process is similar to that used for wirebond package encapsulation discussed above. A common encapsulation process would involve dispensing a liquid polymer, typically a filled epoxy compound, onto the chip, and curing the polymer producing material properties compatible with the interconnection system. The primary purpose of encapsulation is to protect the chip, bumps and leads from environmental and mechanical damage. Of particular importance is protection against mechanical loads (bending, flexure, etc.), shock, vibration, and edge shorts. This must be done without causing excessive stresses at the chip interface. Common encapsulant materials for TAB structures are filled epoxies, fused silica and crushed quartz are the most common fillers—and silicone elastomers.

9.6.8 Singulation and Burn In

Next, the TAB circuits are singulated from the tape and tested or *burnt-in.* The *singulation process* is accomplished using a shearing operation. One of the advantages of TAB is that after ILB, the component can be tested and burned-in prior to assembly. The purpose of testing is to identify assembly problems and to ensure electrical functionality. Burn-in is the screening of functional devices under operating conditions to produce highly reliable components and identify components that fail prematurely, so called infant mortality failures, under operating conditions and elevated temperatures. This ensures that known good devices proceed to the next level of assembly for outer lead bonding.

For single chip packaging cases, the leads of the package are now formed (also called *downsetting*) using a bending operation. The most common lead shape is a gull-wing lead. The purpose of the lead forming process is to bring the leads for outer lead bonding from the plane of the chip to the plane of the substrate top side, to provide a large flat area for thermode or bonding tool, and to increase compliance.

9.6.9 Outer Lead Bonding

The final process step for TAB interconnections is *outer lead bonding* (OLB), shown schematically in Figure 9.6b. This operation transfers the chip, interconnected with the tape leads, to the next level package or PWB. Similar to ILB, OLB can be accomplished using gang bonding or single point bonding. There are many processes available for OLB, including thermocompression, thermosonic, hot bar, laser, ultrasonic, lasersonic, infrared, hot gas, vapor phase, and conductive adhesive interconnection processes. These processes are reviewed further in Chapter 17.

9.6.10 Electrical Performance

Relative to wirebond interconnection structures, TAB interconnections tend to have improved electrical performance. This is due to short circuit lead lengths between the chip,

Samsung, EX1028, p. 206

and substrate reducing impedance and signal delays. In contrast, wirebond interconnections tend to have long wire loops between the chip and package lead frame, increasing line impedance and signal delays. It should be noted that improvements are continuous in wirebonding to make shorter, lower profile bonds.

9.6.11 Production Examples

Tape Ball Grid Array: The tape ball grid array (TBGA) is a ball grid array TAB package developed by IBM. The TBGA uses area-array first level interconnections and a standard ground plane, as shown in Figure 9.5, achieving lower lead inductance, lower power-supply inductance, and lower signal delay. A stiffener is used to maintain planarity of the polymer tape during package construction and second level assembly. The stiffener, which is Ni plated Cu, also helps to reduce the effective thermal expansion coefficient between the chip, package tape, and substrate. Area array solder ball interconnections are used to interconnection the TBGA to the second level of packaging, typically a PWB. The ILB process can use conventional peripheral beam lead bonding, gang or single, or area array flip chip interconnection. After bonding, the assembly is encapsulated with a curable liquid polymer epoxy. The TBGAs are typically 1.4 mm in height and weigh less then 5 g.

TapePak®

The TapePak® was developed by National Semiconductor in the mid-1980s. It is a fully testable, plastic molded, *quad flat pack* (QFP), having leads on all four sides of the surface mountable package. The TapePak uses a single layer bumped copper metallized tape. Thermocompression bonding is used for the inner lead bonds and the devices are molded with epoxy compounds. For outer lead bonding, the lead trim and form operation are done at the *pick-and-place operation* at board assembly. The compact size of the TapePak offers improved electrical performance.

Pentium® TCP

For mobile applications, Intel introduced the TCP version of the Pentium for notebook, laptop, palm top computers, and related portable products. The TAB tape used consists of a JEDEC-style UO-018, 48 mm three layer 35 μm, one ounce copper tape. ILB is achieved using a single point thermosonic bonding process. The inner lead bonds are encapsulated with a high temperature epoxy material, leaving the back side of the chip exposed for thermal management purposes. Intel recommends three outer lead bonding processes, including hot bar soldering, hot gas soldering, or laser soldering. For these high power packages, thermal management is a particular challenge, where thermal vias and thermally-conductive heat sink-attach materials are used to promote heat dissipation.

ETA Supercomputer

ETA Systems, Inc. implemented one of the first TAB applications in the 1970s for supercomputers. The packages were 248 pin tape quad flat packs. The chips were bumped with high lead solder having a diameter of 100 μm. ILB was accomplished using a *solder reflow* process which provides a higher performance interconnection than com-

Samsung, EX1028, p. 207

parable gold bump thermocompression bonding techniques. The bumps accommodate variable bump heights or non-parallel thermode for better gang bonding.

9.7 FLIP CHIP

One of the significant developments to improve cost, reliability and productivity in the electronic packaging industry has been the advancement of flip chip technology. The flip chip process was first introduced for ceramic substrates as the *Solid Logic Technology* by IBM in 1962. It was later, in 1970, that IBM converted this to the *Controlled Collapse Chip Connection,* or C4, for ICs. Flip chip technology is an advanced form of surface mount technology, in which bare semiconductor chips are turned upside down, and hence called *flip chip* (i.e., active face down), and bonded directly to a printed circuit board or chip carrier substrate. Development of solder-bump flip chip interconnections was initiated in an attempt to eliminate the expense, unreliability and low productivity of the manual wirebonding of that era. In contrast to wirebonding, which is a peripheral and time consuming bond technique, in which bonds are formed sequentially, the flip chip allows all I/Os to be connected simultaneously.

Like its predecessors, flip chip technology was initially applied to peripheral contacts, but quickly progressed to area arrays which allow for high input/output (I/O) counts at larger pitches and reduced die size because solder bumps can be put over active device areas on ICs, unlike wirebonds. The C4 flip chip technology uses high lead solder, usually 93%Pb/7%Sn, bumps deposited on solder wettable metal terminals on the active surface of the semiconductor chip that connect to matching wettable pads on the ceramic substrate.

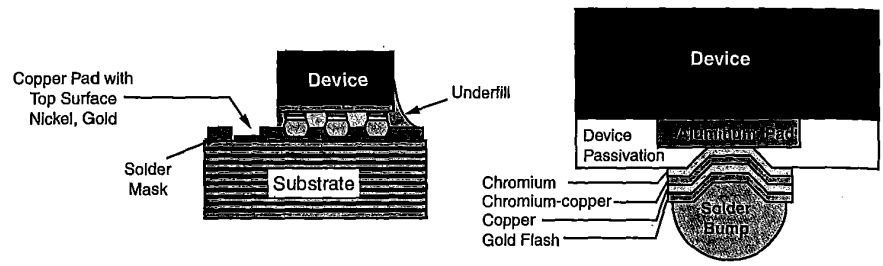
Next, the solder bumps are aligned to the corresponding substrate metal pads and then reflowed at high temperature to simultaneously form electrical and mechanical connections. During reflow, the wetting action of the solder driven by surface tension forces will align the chip's bump pattern to the corresponding substrate pad.

9.7.1 Concept

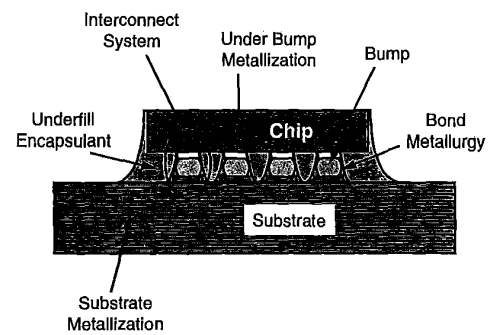
Flip chip interconnection is the connection of an integrated circuit chip to a carrier or substrate with the active face of the chip facing toward the substrate. Interconnection between the chip I/O and substrate is achieved using a bump structure on the chip and a bonding material, typically on the substrate, forming an electrical interconnection between the chip and the substrate. Flip chip bonding typically involves solder interconnections that make the electrical and mechanical connection between the chip and the carrier, although alternate material systems such as conductive adhesives can also be used. A schematic representation of a flip chip interconnection configuration is shown in Figure 9.8, where a bumped chip is interconnected to a substrate with the active face of the IC towards the substrate surface.

In order to better understand the flip chip interconnection system, it is helpful to systematically break down the structural elements of the assembly. One such breakdown is shown in Figure 9.8. The basic structure of a flip chip consists of an IC or chip, an interconnection system, and a substrate. The ICs can be made of silicon (the most common), gallium-arsenide (GaAs), indium-phosphide (IP), silicon-germanium (Si-Ge), etc.

Samsung, EX1028, p. 208



(a) Solder Interconnection



(b) Solder and Conductive Adhesive Interconnection

FIGURE 9.8 Schematic of the flip chip interconnect system.

The substrate materials could be ceramic (used in IBM's original controlled collapse chip connection—C4 process), epoxy-glass laminate, polymer thin-film build-up, *resin-coated copper* (RCC) build-up, glass, silicon, dielectric-coated metal, liquid crystal polymer, dielectric metal matrix composite, *low temperature co-fired ceramic* (LTCC), ceramic thick-film, multilayer high temperature co-fired ceramic, etc.

9.7.2 IC Bond Pad Interface

The interconnection system can be subdivided into four functional areas: *under bump metallization* (UBM); *chip bumps,* bond materials between the bump and substrate metallization; *encapsulant;* and *substrate metallization.* A representative structure is shown in Figure 9.9. The IC bond pad interface metallization for flip chip applications consists of the UBM, chip bump, and bond materials.

Under Bump Metallization

The under bump metallization (UBM) serves as a compatible layer between the bump metallization and final chip metallization. The most common final chip metallization is aluminum (Al), but gold is also used (mainly for GaAs applications), and copper (Cu) is gaining in popularity due to its improved electrical performance. The UBM also helps

Samsung, EX1028, p. 209

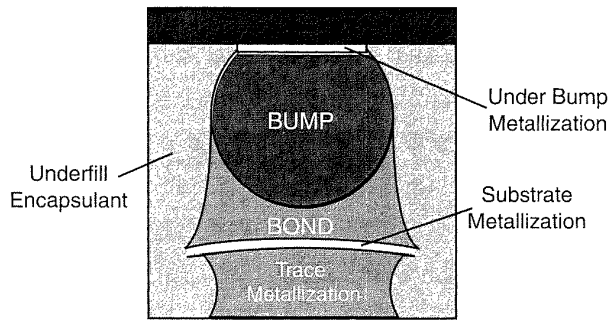


FIGURE 9.9 Schematic of the flip chip interconnect system.

prevent the corrosion of the chip metallization due to diffusion of contaminant ionics from the encapsulants and environment. The structure of the UBM consists of an *adhesion layer* covering the chip metallization, a *barrier layer,* a *wetting layer,* and an *anti-oxidation barrier.* The adhesion layer promotes a strong interface between the bump, the chip metallization, chip passivation, and any dielectric passivation. Careful material selection is necessary to provide good adhesion between these diverse surfaces. Chip passivation provides protection from the environment and insulation of the devices on the silicon surface, and may act as a buffer layer for stress relief (e.g., polyimide and BCB). Typical *adhesion layer* materials are chromium (Cr), titanium (Ti), nickel (Ni), tungsten (W), titanium-tungsten (TiW), and zincate (Zn derivative). The purpose of the *barrier layer* is to prevent diffusion of metal species and ionic contaminants into the chip metallization and adhesion layer. Such diffusion can result in the formation of brittle intermetallics and corrosion of the chip/adhesion metallization, significantly lowering the reliability of the interconnection system.

Based on Fick's second law for diffusion along one direction, Equation (9.1), the approximate time required for elemental diffusion can be derived, and is given by Equation (9.2).

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left(D\,\frac{\partial c}{\partial x}\right) \tag{9.1}$$

$$t \approx \frac{h^2}{D} \tag{9.2}$$

where c = instantaneous or local concentration of the diffusing species [kg(diffusing species)/kg(base material)]
D = diffusivity of a molecule (species) in a base matrix [m^2/s]
h = characteristic length over which diffusing species diffuses [m]
t = characteristic time over which diffusion takes place [s]

Using Equation (9.2), the characteristic diffusion time can be estimated for intermetallic formation.

Common metallurgies used for barrier layers include chromium (Cr), tungsten (W), titanium (Ti), titanium-tungsten (TiW), nickel (Ni), or chromium-copper (Cr-Cu).

After the diffusion layer, there is a *wetting layer.* This metallization provides a consumable layer for the subsequent bump metallization to wet and react, forming inter-

Samsung, EX1028, p. 210

metallics. Molecular diffusion of reactive bump species into the wetting layer is common and often results in the formation of intermetallic compounds. For example, for an a eutectic lead-tin bump, the tin will wet and diffuse into a copper metallized wetting layer and form copper-tin intermetallics such as Cu_6Sn_5 and Cu_3Sn. Typical wetting layers consist of copper (Cu), nickel (Ni), palladium (Pd), or platinum (Pt).

The final UBM layer is an *antioxidation barrier* which is optional, and is typically a very thin layer of gold (Au). Thin layers of gold are used in order not to embrittle the UBM-bump interface due to the formation of intermetallics. For example, in the case of lead-tin bumps, gold will rapidly dissolve in liquid tin, forming brittle tin-gold intermetallics compromising the interconnection reliability. Minimizing the gold thickness, while providing a pin hole-free oxidation barrier, ensures robust UBM performance for subsequent bumping operations.

UBMs used commercially today for solder bumping operations include [Cr, Cr-Cu (deposited phased), Cu, Au], [Ti, Ni, Au], [Ti, Pt, Au], [Ti, W, Ni, Au], [Al, Ni-V, Cu], and [Zn, Ni-P, Au]. Each UBM has been qualified for a range of applications and solder compositions. In general, UBM selection must be based on the bump metallization desired, the operating conditions of the chip, current carrying requirements, and the process history required (i.e., multiple reflow cycles, etc.).

The under bump metallization (Figure 9.10) is necessary for several reasons. A chip passivation is necessary to protect the Al metallized chip from moisture and corrosion. A polyimide layer is then applied for stress relaxation. The first metal layer is the *adhesion layer.* This metal must adhere to the aluminum oxide, the chip passivation, and the polyimide. It must also have low electrical contact resistance. Often, argon sputter etching of the Al surface is used to remove Al_2O_3 and reduces interfacial resistance. In the original flip chip design, IBM recommended a 0.15-μm chromium layer. The second layer necessary is the *barrier layer.* This layer prevents the solder, primarily the tin, from migrating too quickly and corroding the aluminum device and/or dewetting the Cr layer during multiple reflows. Traditionally, a chrome/copper alloy has been used as barrier layer (0.15 μm). The third layer deposited is the *wetting layer.* This layer, usually a thick layer of copper (1.0 μm), is necessary because the solder will not wet the metal deposited for the barrier layer nor the bare aluminum. Finally, a flash of gold is placed on the top of the wetting layer to inhibit copper oxidation in a dual evaporation process.

There are several processes for deposition of the UBM layers. Evaporation has been used the longest and was the original deposition technique used by IBM for its flip chip

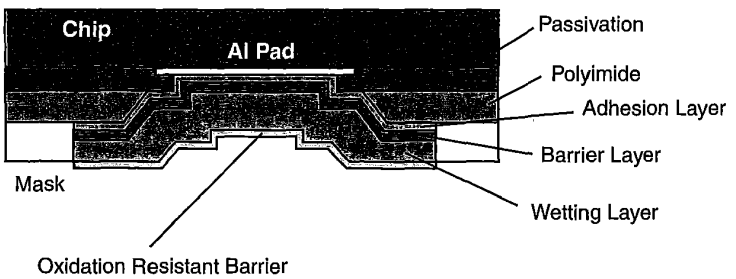


FIGURE 9.10 Structure of the under bump metallization.

Samsung, EX1028, p. 211

technology. Evaporation is a process based on the vaporization of a metal compound in a vacuum chamber, resulting in a uniform coating over the entire chamber surface, including the wafer. It requires patterning based on physical masks or lift-off techniques. In some cases, blanket metal coatings are deposited on the entire wafer and later etched back. UBMs can also be sputtered. Sputtering uses a directed metal ion plasma between an anode and cathode in a vacuum chamber, based on targets made from the materials to be deposited. The directed plasma reduces waste; however, to minimize oxidation problems, sputtering systems require multiple targets, one for each metal layer, and large chamber volumes, resulting in very expensive tooling. Sputtering is perhaps the most popular UBM deposition process.

Additional processes for deposition of UBMs are based on plating processes: *electrolytic plating* and *electroless plating.* Electroless nickel plating, based on a nickel-phosphorous plating chemistry, has gained in popularity. Alternate chemistries can also be used. Electroless nickel plating requires a pretreatment of the wafer using a zincate surface preparation technique. This surface treatment acts to seed the aluminum metallization with zinc to allow for good nickel adhesion. Electroless nickel plating is an autocatalytic process utilizing a hypophosphite reducing agent. Deposition rates are typically 0.25 microns per minute. There are several steps in the electroless plating process. The process steps would include micro-etching, rinsing, acid stripping, zincate, rinse, and conditioning to strip the zinc from the wafer passivation, rinse, electroless nickel plating, and a final rinse. Post plating, an amorphous deposit of nickel and phosphorous, is left on the chip I/O pads. In general, the deposited nickel imparts a large stress on the I/O pad and underlying silicon. This stress is a strong function of the plating thickness and stability of the nickel microstructure, so careful control of the plating step is important. The amorphous microstructure of the nickel-phosphorous layer can also lead to unstable microstructures for the nickel for some applications. This follows since the amorphous microstructure is energetically metastable. From a free energy perspective, the nickel will always drive towards a minimum energy state, which is its crystalline form. Alternately, it may react to form stable crystalline intermetallic compounds, or morph into larger nodule structures.

Overall, the UBM process is very costly and time consuming. In general, these metallization steps are completed in a clean, large area vacuum environment. In addition to the metal layers necessary for flip chip, an underfill, usually silica-filled epoxy, is typically applied between the joined chip and substrate. This underfill is completed after the solder is reflowed. Any rework and repair must be done before the application of underfill. The underfill is needed to improve reliability. Because there is a mismatch in the coefficient of thermal expansion among the chip, the solder interconnections, and the substrate, stresses will be introduced. The underfill distributes the thermomechanical stresses over a larger area, significantly increasing reliability.

IC Bumps: The bumps in a flip chip interconnection provide four functions: (1) electrical connection between the chip and the substrate; (2) a heat dissipation path from the chip; (3) environmental protection; and (4) a structural link between the chip and the substrate. The materials and processes involved in the manufacture of the flip chip interconnection system determine its performance.

The most common materials used for bumping flip chip devices include solder, metallic stud bumps, and compliant polymer bumps. Schematics of the various types are shown in Figure 9.11. Solders can be divided into three primary material systems:

Samsung, EX1028, p. 212

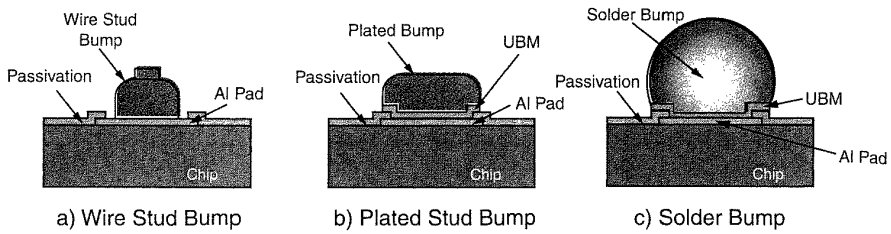


FIGURE 9.11 Flip chip bump: (a) wire stud bump; (b) plated stud bump; and (c) solder bump.

1. High temperature with melting points in excess of 250°C (examples include 95% Pb–5% Sn and 97% Pb–3% Sn)
2. Moderate temperature with melting points between 200 and 250°C (examples include 95.5% Sn–3.5% Ag–1.0% Cu, CASTIN® Cu-Ag-Sb-Sn, and 85.9% Sn–3.1% Ag–10% In–1.0% Cu, and 96.5% Sn–3.5% Ag)
3. Low temperature with melting points less then 200°C (examples include 37% Pb–63% Sn eutectic, 88% In–12% Pb, 100% In, and 48% Sn–52% In).

An example of a solder-bumped flip chip is shown in Figure 9.12. Notice the uniformity of the solder spheres, which is typical for a high quality bumping process.

Solder bumps are deposited onto the UBM using several processes. The earliest process used was *evaporation,* developed by IBM for its flip chip technology. It is still commonly used today and requires a physical mask, typically molybdenum, that is aligned to the wafer I/O pads and released after deposition. The mask must have a draft angle though the apertures in order for the solder to release from the mask. A second process for deposition of solder is *electroplating.* An example of a commercial solder bumping process based on electroplating is shown in Figure 9.13. Notice that the solder is electroplated through a polymer photoresist mask to control volume and registration. The bumps are reflowed after plating to homogenize the solder microstructure, since the electroplated solder is phase separated on deposition.

A third process for deposition of solder is based on *solder paste screening* or *printing.* Squeegee printing of specially designed solder paste can be accomplished using a precision stencil and automated stencil or screen printers. At finer pitches, this process becomes more difficult (less than 250-μm pitch). Bump printing can also be achieved using a polymer photoresist mask similar to that shown in Figure 9.13; however, the solder is doctor-bladed or squeegeed into the holes in the resist. A third printing process commercialized is based on jet printing technology similar to ink jet printing used for

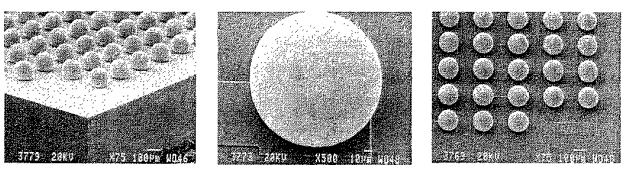


FIGURE 9.12 Typical solder-bumped area array flip chip die. (*Source: Flip Chip Technologies*)

Samsung, EX1028, p. 213

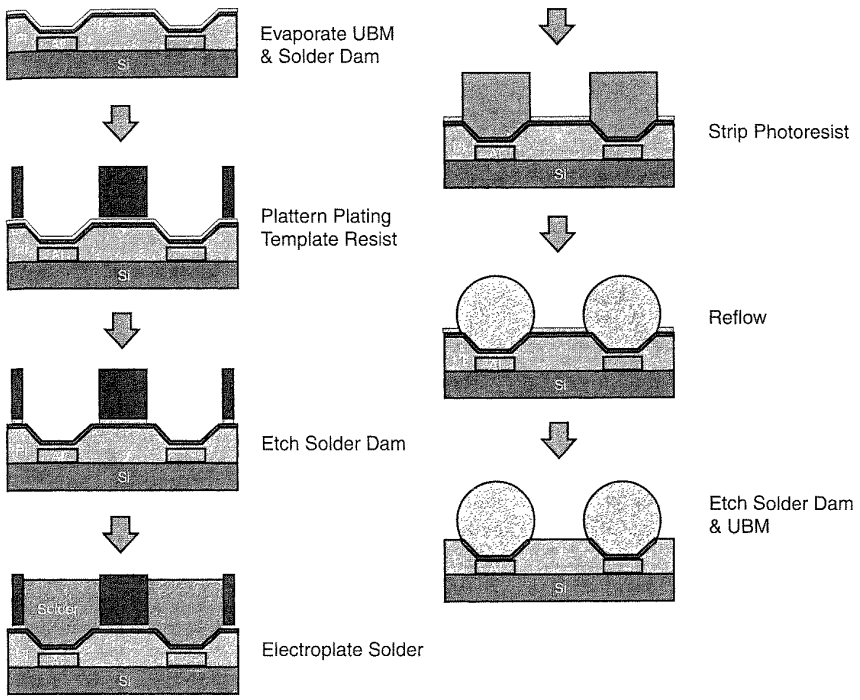


FIGURE 9.13 Example of an UBM and solder bumping deposition process based on evaporation and plating. (*Source: Unitive*)

document printing. In this case, molten solder drops are jetted onto the bond pads, forming the bumps. Solder jet printers can be configured for drop-on-demand deposition or the higher speed continuous drop deposition.

Stud bumps can be plated or deposited using a wire stud bumping technique. Plated stud bumps are largely derived from the systems used for TAB bumping, including gold, nickel, copper, gold-tin, nickel-gold, and nickel-copper. Both electrolytic and electroless plating, techniques are used where electroplating requires an electrical path for plating and achieves nearly an order of magnitude faster deposition rates. Wire stud bumps are formed using a ball bonding technique where the wire is fractured after ball bond formation. The wire stud bumps can be coined to form a uniform structure for bonding. The most common wire stud bumps are made with 25-μm diameter gold wire. Also used are Pb-Sn (98/2), Sn-Ag, Cu, Pt, and Pd with varying degrees of success.

Compliant polymer bumps have several configurations, the most common consisting of a polymer elastomer, often filled with conductive metal particles, that is over-coated with gold. Conductive polymer bumps are most often screen or stencil printed onto the wafer I/O pads metallized on a UBM. The gold over coat, if used, is based on electroless or immersion gold plating.

Samsung, EX1028, p. 214

9.7.3 Flip Chip Processing

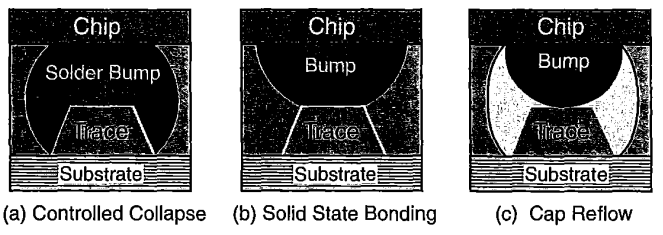
Interconnection of the bump to the substrate metallization is achieved using a bond material providing both an electrical and mechanical interface between the bump and substrate trace. Often the bump and bond materials are the same, or the bond and encapsulant materials are the same. That is to say, the function of the bump and the bond, or the bond and the encapsulant, is accomplished using a single material solution such as eutectic lead-tin solder or anisotropic conductive adhesives, respectively.

Solder Interconnection Processing: In the case of solder interconnection systems, the bond can consist of three basic types.

1. The first is a fully *wetted controlled collapse solder interconnection* where the entire solder bump is raised above its liquidus temperature, wets, and reacts with the substrate metallization as shown in Figure 9.14a. IBM's flip chip process is based on such an interconnection system. In this case, the standoff gap between the chip and substrate is controlled by the surface tension of the solder, the solder volume, chip weight, and the pad sizes on the chip and substrate.
2. Although not used commercially, the second solder interconnection system is a *solid state bond* between the bump and substrate metallization, as shown in Figure 9.14b. Solid state bonds are commonly formed using thermocompression, or thermosonic bonding techniques, much like wirebonds or TAB bonds. In this case, the interconnections are gang-bonded, forming all interconnections simultaneously.
3. The third solder interconnection structure is shown in Figure 9.14c, consisting of a *cap reflow configuration.* In this case, a high melting point bump is bonded to the substrate trace with a lower melting point alloy. The most common configuration is a high lead solder bump (95% Pb–5% Sn or 97% Pb–3% Sn with liquidus temperatures of 312°C) bonded with a eutectic Pb-Sn solder cap (liquidus temperature of 183°C) that is plated on the substrate trace.

With the use of an appropriate flux material and reflow profile, the eutectic solder will liquify, wet and react with the substrate metallization and high lead bump, forming a robust interconnection with a known standoff gap height controlled by the bump height. This interconnection configuration is commonly used in flip chip package applications for microprocessors and ASICs.

Conductive Adhesive Interconnection Systems: Adhesive interconnection systems are also used for bonding bumped flip chip components to the substrate metallization. The



(a) Controlled Collapse (b) Solid State Bonding (c) Cap Reflow

FIGURE 9.14 Solder interconnection structures consisting of a bump, bond and encapsulant system.

Samsung, EX1028, p. 215

basic structures of the two major adhesive interconnection systems, isotropic conductive adhesive and anisotropic conductive adhesive bond systems, are shown in Figure 9.15.

Isotropic Conductive Adhesives: Isotropic conductive adhesives are typically thermosetting polymers filled with conductive particles. The most common material system is epoxy filled with silver flake particles. They become conductive in all directions upon cure, yielding an electrically functional interconnection. Isotropic adhesives can be printed onto the substrate bond pads, printed on the chip bumps, or dip-transferred to the bumps. In the latter case, the bumped chips are dipped in a precision thin-film of the conductive paste, transferring a controlled amount of the adhesive onto the bumps. Upon assembly, the conductive adhesive forms a bond between the bump and substrate metallization. The most common bumps used for these interconnections are gold wire and gold plated stud bumps. Gold surface metallization on the substrate traces are used to minimize corrosion of the interconnection systems. These material systems form a mechanical contact, versus a metallurgical bond, as is the case with solder interconnection systems. The electrical performance of the mechanical contacts is adequate for many applications but can be problematic under conditions where corrosion failure modes can occur. In general, unless pin hole-free noble metal finishes are used over the bumps and substrate metallization, the electrical contact resistance between the bump—to conductive adhesive or conductive adhesive—to substrate metallization can increase in value over time. Isotropic conductive adhesives do not provide self-alignment during assembly like solders requiring high precision assembly processes and slower assembly tools. They are commonly used for flip chip on glass assembly for flat panel display production. If the substrate has a tendency to warp during assembly, particularly during cure, conductive adhesive interconnections require the chip to be held under force to tack cure the adhesive. The underfill can then be applied to the flip chip and simultaneously cured with the conductive adhesive. Without the tack cure of the conductive adhesive, any warpage of the substrate, or chip, can cause the interconnection to fail, since it is quite brittle prior to completion of cure.

Anisotropic Conductive Adhesives: Anisotropic conductive adhesives can be film-type or paste-type polymer materials filled with conductive spheres. The conductive spheres can be metal (nickel, gold-coated nickel) or metal-coated polymer (gold-coated or nickel–

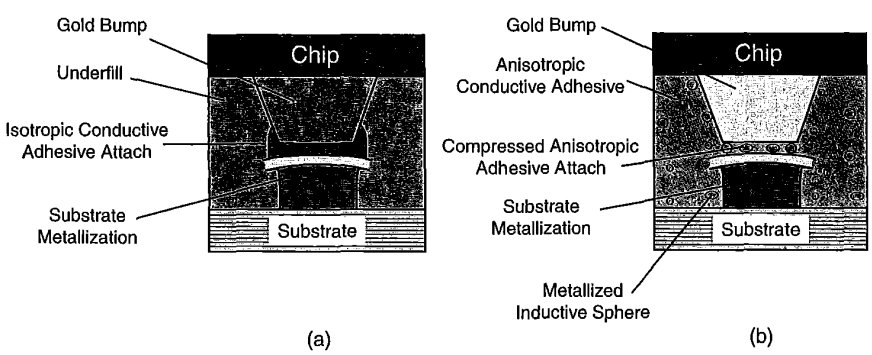


FIGURE 9.15 Adhesive interconnection structures consisting of a bump, bond and encapsulant system: (a) isotropic conductive adhesives; (b) anisotropic conductive adhesives.

Samsung, EX1028, p. 216

gold-coated). In the latter case, both solder and hollow polymer spheres are used. Anisotropic conductive adhesives do not conduct electric current until compressed to the point where the conductive spheres are compressed and captured between the bump and substrate trace, as shown in Figure 9.15b. Similar to isotropic conductive adhesives, bumps and substrate metallization are typically noble metals, to minimize corrosion failures. Assembly of anisotropic conductive adhesives requires high-accuracy placement since the materials do not provide self-alignment capabilities like solders. In addition, anisotropic conductive adhesives require relatively high bonding forces in order to make good electrical contact. The assembly process involves a high-accuracy placement using a flip chip bonder, followed by a high force bonding tool having a high degree of planarity between the bonding head and the substrate holder. Typical bonding forces are 20–100 g per bump, which limits I/O counts of the chips that can be assembled.

9.7.4 Flip Chip to Organic Substrates

The original flip chip processes implemented by IBM and others was for ceramic substrates. For example, IBM used this technology successfully from 1960 to 1990, making hundreds of billions of flip chip connections without a failure related to flip chip during this time. Over this time, IBM learned many benefits of this technology:

1. Lower cost than wirebonding since the bumping was done at wafer level and since all connections are made simultaneously
2. Higher reliability than wirebonding and beam lead bonding
3. Better electrical performance than wirebonding due to its lower resistance, capacitance and inductance
4. And most importantly, reparability: If an IC was defective, either during assembly or during usage, it can be removed and a new IC is flip chip bonded on the same ceramic substrate.
5. Improved chip designs may be substituted on multichip modules.

In the early 1990s, *organic carriers* emerged as alternatives to ceramic for better and cheaper packages. However, organic carriers have one major problem. Their coefficient of thermal expansion (CTE) of about 17–22 ppm/°C is far from alumina ceramics having a CTE of about ~7 ppm/°C and silicon of ~3 ppm/°C. Any attempts of flip chip bonding to organic carriers, therefore, resulted in poor joint reliability, often fatigue-failing in less than 100 thermal cycles.

This is not surprising if one looks into the fundamental fatigue Equation (9.3) governing this phenomena:

$$\text{Strain on Outer Most C4 Ball} = \frac{\text{DNP}((\text{CTE}_{\text{Carrier}} \cdot \Delta T_{\text{Carrier}}) - (\text{CTE}_{\text{Si}} \cdot \Delta T_{\text{Si}}))}{H} \tag{9.3}$$

where ΔT = relative thermal excursion
H = C4 interconnection height
DNP = distance to neutral point
CTE = coefficient of thermal expansion

Samsung, EX1028, p. 217

Decades of both fundamental and experimental reliability testing has generally validated the fatigue life plot shown in Figure 9.16. The mean time to failure, according to this plot, is related to the thermal expansion coefficient of the substrate, everything else being equal. This led to the belief that ceramic materials such as alumina, because of their low CTE, are inherently better and that as the size of the IC is increased beyond 300 I/Os at 0.25 mm pitch, further lowering of CTE was necessary. This was the basis of glass-ceramic/copper technology with its CTE designed to exactly match silicon, which shows almost infinite fatigue life in this figure.

9.7.5 Major Discovery

IBM and Hitachi were credited with starting the journey of flip chip to organic boards. Tsukada of IBM Japan found, based on some observations with sealing using AIP-10 polymers in the 1970s and 1980s, that underfilling the gap between a flip chip IC and the organic substrate or board, would result in tremendously improved reliability. This work started in 1985. By 1988, Hitachi reported the use of polymer underfills for improved flip chip reliability. IBM-Endicott reported the use of flip chip underfill in 1990. Tsukada at IBM Japan discovered, as illustrated in Figure 9.17, that the strain on the solder with an organic underfill is almost an order of magnitude lower than without underfill. It is also noted that IBM data indicates that eutectic solder has higher reliability than high lead solder with underfill. Interestingly, the bulk of IBM's flip chips mounted on high density SLC organic substrates are comprised of 97/3 Pb/Sn solder bumps from the IC side with eutectic 37/63 Pb/Sn solder tips (from the board side) connecting the high Pb-solder bumps to the substrate pads. This way, IBM didn't have to change its wafer-level bumping from its needs with ceramic substrates.

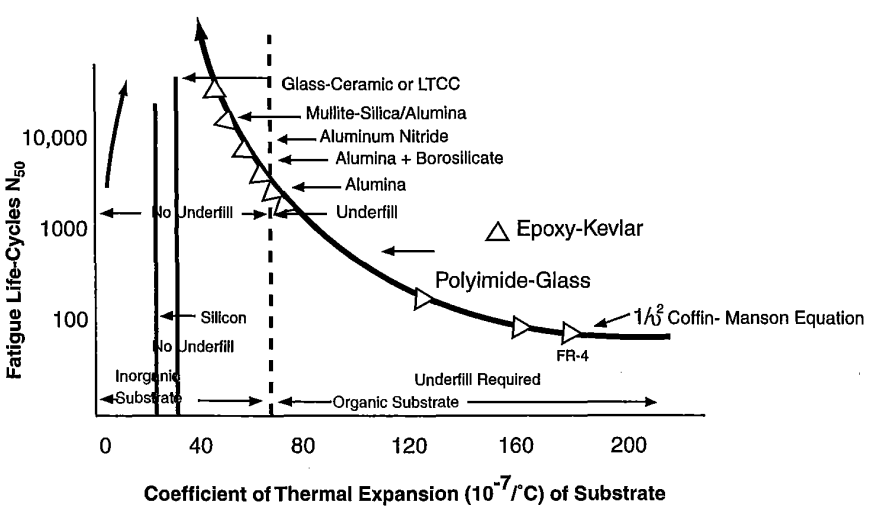


FIGURE 9.16 Mean time to fatigue failure as a function of thermal coefficient of expansion of substrate.

Samsung, EX1028, p. 218

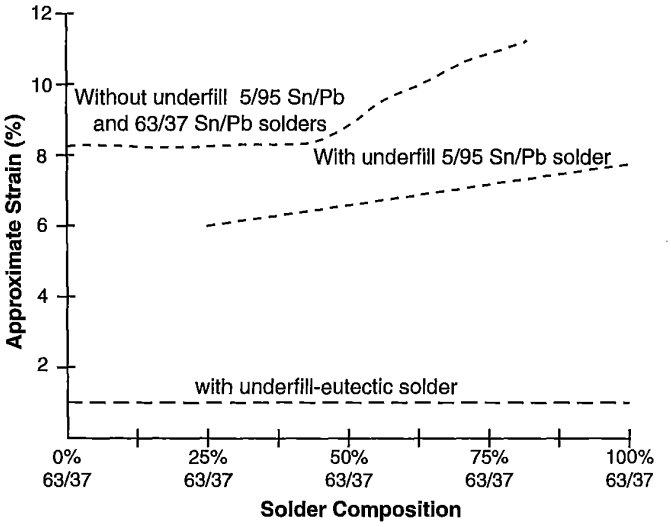


FIGURE 9.17 The strain reduction in solder with underfill.

9.7.6 Underfill Encapsulants and Processing

Underfill encapsulants are applied at the interface between the chip and substrate, and serve to compensate for the coefficient of thermal expansion (CTE) difference between chip and substrate. For flip chip on organic board applications, the silicon device has a CTE of ~3 ppm/°C and the FR4 organic substrate has a CTE of 17–22 ppm/°C. For flip chip on alumina ceramic applications, the CTE of the alumina ceramic is ~7 ppm/°C, resulting in significantly lower thermomechanical stresses on the interconnections. Even today, for relatively small flip chip components mounted on ceramic substrates, underfill materials are not used. In all of the early applications of IBM's flip chip technology, flip chips were mounted on ceramic substrates without underfill. IBM used AIP-10 polymers on MC substrates and silicone gel on SLT to avoid solder corrosion between fine features. It was not until the early 1990s that the concept of underfills was introduced into production.

One of the primary purposes of underfill is to couple the chip and substrate over the entire area of the chip, or at least around its perimeter, lowering the effective thermomechanical stress on the flip chip interconnections. By coupling the chip and substrate, the effective composite system CTE falls between that of the chip and substrate, hence increasing reliability. The effectiveness of underfills depends on the relative difference in CTE between the chip and substrate. Flip chips on ceramic packages with underfill show very high reliability relative to flip chips on organic substrates with underfill. In both cases, package level qualification reliability can, and has been, achieved commercially.

Underfills also serve the function of protecting the interconnections from environmental effects (i.e., moisture) and absorb harmful alpha particle emission from the lead in solders which can cause soft errors in logic circuits. For flip chip applications, encapsulants consist of filled polymer underfills or uncompressed anisotropic conductive adhesive, which provide encapsulation properties.

Samsung, EX1028, p. 219

The predominant method of encapsulant underfill processing in flip chip assembly is based on capillary flow. Figure 9.18 shows a schematic of the basic process. The underfill is applied after interconnection of the device. Underfill comprised of a filled liquid polymer—silica filled epoxy compounds being the most common—is dispensed along a single edge or dual adjacent edges of an assembled flip chip. The reservoir formed supplies the capillary flow of the underfill between the standoff gap of the device. Surface tension forces draw the material under the chip. The material passes through the interconnection geometry and subsequent dispense steps replenish the reservoir, assuring that the material completely fills the volume underneath the chip. Subsequent dispensing of the material is usually required to completely fill the area beneath the chip and to form the fillets. Fillets are a key component in the compensation of stresses. Following the dispense process, the material undergoes a thermal cure process, producing a rigid polymer composite.

After a final fillet dispense pass, the underfill is cured in an oven. Typical cure temperatures for underfill are below that of the melting point of the solder interconnections, in the range of 130–175°C. In addition to capillary flow, underfills can also be preapplied on the substrate prior to chip assembly, and then compressed during chip assembly. Yet another underfill process involves the injection of the viscous polymers under the chip, using a transfer molding process or discrete injection process.

Underfill encapsulants in use today are typically silica-filled epoxy based materials. Appropriate materials should exhibit good flow properties and short cure times. Short fill and cure times are essential for more cost effective and high volume manufacturing. Materials should provide mechanical shock and bending protection, as well as a barrier for environmental protection. Epoxy based materials cured with anhydride exhibit a high glass transition temperature (T_g), good chemical resistance, low moisture absorption, good flow properties and adhesion strength to other packaging materials.

The most critical properties of underfill encapsulants are CTE, T_g, pot-life, viscosity, filler size, alpha particle emission, and extractable ionics. Typical viscosities for underfill materials at room temperature range anywhere from 10–30 Kcps at low shear rates. Flow rate during processing depends on such factors as viscosity, temperature, surface tension, gel time, gap height, and surface wettability.

Process parameters associated with capillary flow underfill processing are the dispense pattern, amount of material, temperature, dispense needle speed, and the distance between the material and the chip. Dispense patterns are performed in one of two ways. Material can be dispensed along one edge of the chip. Once the material has flowed underneath the chip, a second dispense on the remaining sides is used to guarantee complete coverage and fillet formation. The other method dispensed material along two adjacent chip edges.

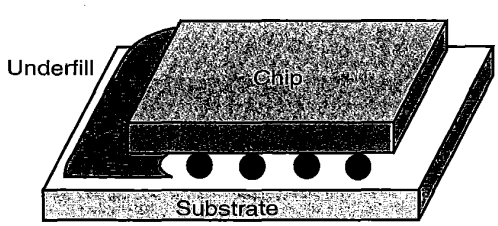


FIGURE 9.18 Schematic of underfill processing via reservoir dispense and capillary flow within the standoff gap height.

Samsung, EX1028, p. 220

This is typically referred to as *L-shape dispensing*. Increased board and chip temperature are used to aid the flow of the underfill (Suryanarayana, et al., 1991).

Potential problems that can occur with underfills are voiding, delamination, and moisture penetration. Voids can be caused by air pockets between the chip and the substrate that form during the dispense process. Delamination can occur due to poor wettability of the underfill material. Voiding and delamination typically lead to solder cracking and moisture penetration.

Capillary Flow Processing: Based on a first-order Newtonian fluid flow analysis of capillary flow in a thin gap (see Figure 9.19), the underfill flow time under the flip chip component can be estimated by Equation (9.4). Notice from Equation (9.4) that the capillary flow time varies with the square of the chip size and inversely with the standoff height, which are continuously increasing and decreasing, respectively, for next-generation electronic assemblies. The trend towards longer underfill flow time puts significant pressure on conventional flip chip processing cycle time, throughput, and cost competitiveness.

$$\tau_{\text{fill}} = \frac{3\mu L^2}{\sigma h \cos(\alpha)} \tag{9.4}$$

where μ = the underfill apparent viscosity
L = the chip size (chip is assumed square of dimension L)
σ = the underfill surface tension
h = the standoff gap height between chip and substrate
α = the wetting angle of underfill

Injection Flow Processing: A second technique for underfill processing is based on injection flow. In this process, underfill is applied using a pressurized mold that mates and seals with the edge of the chip, schematically shown in Figure 9.20. An alternative technique injects the underfill through a gate in the bottom of the substrate, such that the polymer flow expands outward from the center of the chip. A third technique underfills and overmolds the mold compound based underfills in a specially designed transfer molding process and mold. Underfill is injected into the standoff gap between the chip and substrate under pressure. This eliminates slow capillary flow, dramatically reduces underfill flow times and allows for high speed underfill processing.

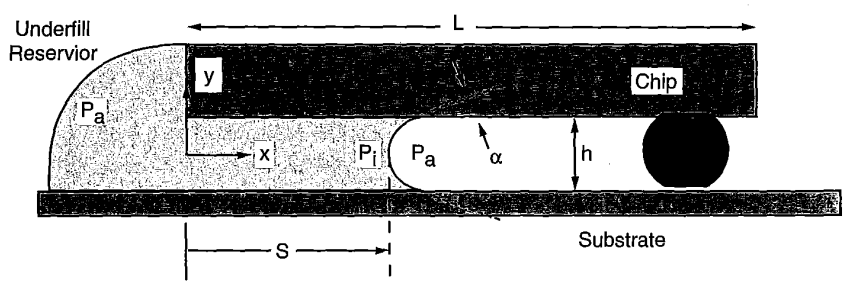


FIGURE 9.19 Capillary flow underfill processing configuration and analysis schematic.

Samsung, EX1028, p. 221

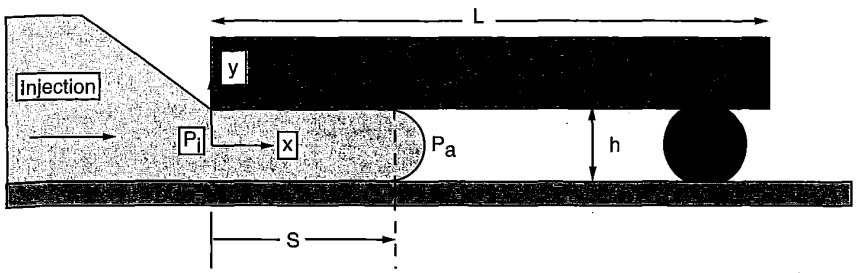


FIGURE 9.20 Schematic of injection flow underfill processing.

The injection flow can be analyzed as a viscous fluid flowing between parallel plates, as shown in Figure 9.20. As a first order approximation, the influence of the interconnections and gravity are neglected. Equation (9.5) gives an approximate fill time for a power law underfill flow.

$$\tau_{\text{fill}} = \frac{2L}{h}\frac{2n+1}{n+1}\left(\frac{2mL}{(P_i - P_a)h}\right)^{1/n} \tag{9.5}$$

The underfill time for a Newtonian fluid is given by Equation (9.5) setting $n = 1$ and $m = \mu$.

$$\tau_{\text{fill}} = \frac{6\mu L^2}{(P_i - P_a)h^2} \tag{9.6}$$

where L = the chip size
n = the underfill power law coefficient
m = the underfill power law constant
h = the standoff gap height
P_i = the underfill injection pressure
P_a = the atmospheric pressure
μ = the underfill apparent viscosity

Compression Flow Underfill Processing: A third technique for underfill processing is based on an innovative compression flow technique, as shown in Figure 9.21. In the compression flow underfill process, the underfill is applied to the substrate prior to assembly. During chip assembly, the underfill is compressed or squeezed between the chip and substrate.

In order to analyze the compression flow underfill process, the chip and underfill geometry are approximated as circular equivalents of a square chip of size L, such that the circular and square geometries have equivalent wetted areas. Based on this criterion, the circular equivalent chip radius is $R = L/\sqrt{\pi}$. Estimating the underfill flow time for the compression flow process is relatively straightforward. Many commercial flip chip placement systems available place components at a constant speed using *placement velocity feedback control.* Typically, force limit sensors are used to insure that the programmed placement force is not exceeded during chip placement. Therefore, the underfill

Samsung, EX1028, p. 222

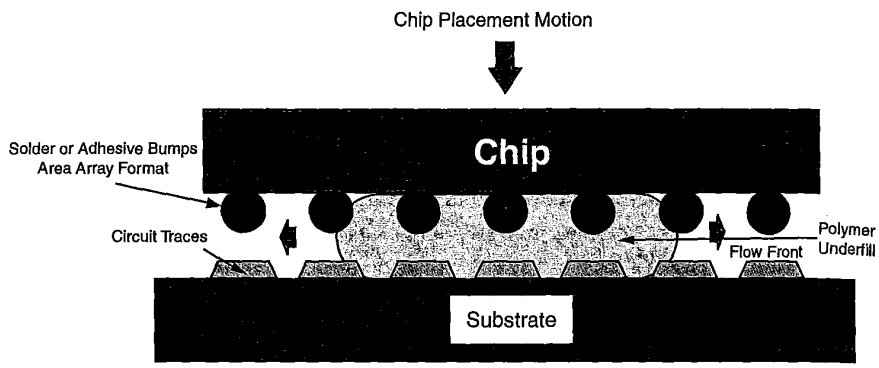


FIGURE 9.21 Compression flow underfill processing.

flow time for the compression flow process is simply estimated by the chip placement time given by equation 9.7:

$$\tau_{\text{placement}} = \frac{(H_0 - f_f)}{v_{\text{placement}}} = \frac{(H_0 - h_f)}{\dot{h}} \tag{9.7}$$

where $\dot{h}$ is the placement velocity of the chip relative to the substrate and h_0 is the initial height of the chip above the substrate.

Anisotropic Conductive Adhesive Encapsulation: The encapsulation process, based on anisotropic conductive adhesives, follows naturally from the basic structure of the materials. Since the materials are not conductive until highly compressed, the anisotropic conductive adhesive, film or paste-type, encapsulates the gap between the chip and substrate upon assembly of the chip. Once the adhesive is cured under bonding forces, the adhesive mechanically couples the chip and substrate in the same manner as traditional underfill. The primary difference is that the anisotropic conductive adhesives have higher CTEs than filled epoxy underfills, due to lower filler loading amounts.

Substrate Metallization: Substrate metallization is best described in terms of the bond materials used in the flip chip interconnection system. For solder interconnection/bond systems, the most common metallizations are copper coated with an *organic surface protectant* (OSP), and copper plated with electroless nickel (100–150 microinches) and immersion, or electroless gold (4–8 microinches). In some instances, electroplated nickel and electroplated gold are used. It is important to control the gold thickness in order to minimize the formation of gold-tin intermetallics and embrittlement of the solder bump and UBM. For this reason, immersion gold is the most popular configuration, since it is a self-limiting plating process that tends to limit the gold thickness. Additional metallizations used for solder interconnections include thick-film conductors (silver, platinum, palladium, copper, etc.), electroplated palladium, and electroplated solder, so called tinned pads.

For adhesive interconnection/bond systems, immersion gold, or soft gold, is the most popular substrate metallization. In some instances, electroplated gold, or hard gold, is used in addition to platinum and other noble metals. The noble metal surfaces are necessary to minimize corrosion in the assembled flip chip components.

Samsung, EX1028, p. 223

9.7.7 Flip Chip Assembly Processes

This section will discuss a conventional *eutectic solder flip chip-on-board assembly* process and is intended to present an example of one of the many possible flip chip assembly processes. A typical *flip chip-on-board* (FCOB) assembly process uses the same chip bonding principles as the classic C4 flip chip assembly process. The primary exception being that the high lead solder reflow is replaced by a lower temperature eutectic solder reflow. A typical process flow for this type of interconnection technology is shown graphically in Figure 9.22. This assembly technology involves the connection of a chip bumped with eutectic lead-tin solder to a Cu-Ni-Au surface metallization on an organic high density microvia substrate using a "no clean" process.

The process begins with a *known good substrate* (KGS) and a *known good die* (KGD), having appropriate chip passivation, under bump metallization, solder bumps, substrate dielectric, substrate solder mask, and encapsulant underfill. All of these factors have to be carefully considered for flip chip applications in order to achieve target reliability levels. This follows since failure tends to occur at the interfaces between these materials. Optimum adhesion between these interfaces is critical for robust reliability performance.

Stencil printing solder paste on the board for the surface mount packages and discretes is the first process step. This often requires a high-precision vision alignment operation for finer pitch surface mount components. Next, the surface mount packages and discretes are picked, vision aligned, and placed onto the substrate. The flip chips are now picked, fluxed, aligned, and placed onto the substrate. The fluxing operation consists of dipping the bumps of the flip chip into a precision film of a low solids, no-clean tacky flux transferring the flux onto the solder bumps. The flux compounds can also be applied on

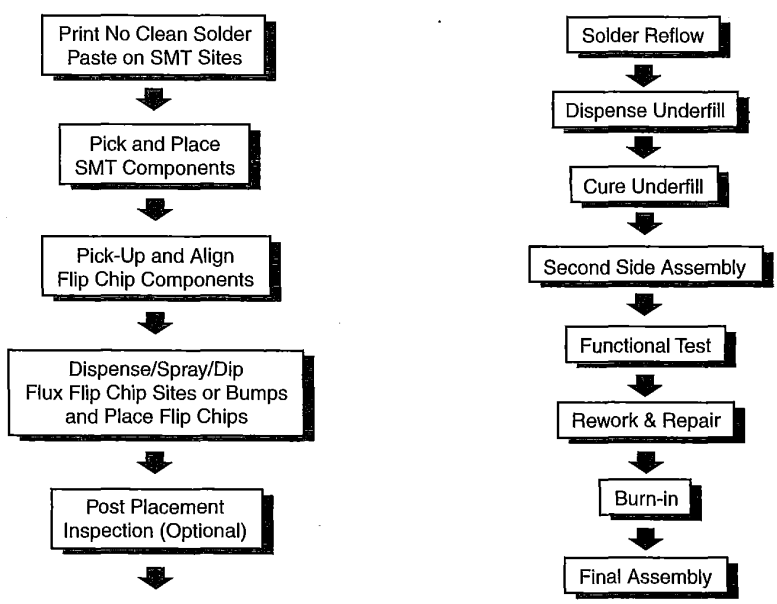


FIGURE 9.22 Conventional flip chip-on-board (FCOB) assembly process integrated with surface mount assembly.

Samsung, EX1028, p. 224

the chip site using a dispensing, spray coating, or jetting operation. After placement, the flux acts to temporarily hold the chip in place prior to reflow. Flux also aids in the removal of the metal oxides during the reflow process, and protection of the metallization from reoxidation. On the downside, flux residues after reflow can have an adverse affect on assembly reliability. Flux residues that are left behind after reflow must be cleaned, or in the case of no-clean processing, the flux materials must be carefully selected to be compatible with the chip passivation, solder mask, and underfill. After placement, an optional process step would be to use *automated optical inspection* (AOI) of the components and chips to insure that all components are placed in the correct orientation and within a high enough degree of accuracy.

During the reflow process, the low melting point solder melts, and wets the substrate metallization. Surface tension forces drive the self alignment mechanism, and a metallurgical bond is made between the chip and substrate. During wetting, the solder reacts with the substrate metallization to form stable intermetallic compounds. In the case of a eutectic solder bump on a copper trace, Cu_6Sn_5 and Cu_3Sn intermetallics form due to diffusion of the Sn into the copper and of Cu into the tin. The standoff height between the chip and substrate is controlled by the chip weight, solder bump volume, chip pad size, and substrate pad size. A typical interconnection is shown in Figure 9.23.

After the reflow process, the interconnections have been formed. For a capillary flow underfill process, underfill is dispensed on single or dual adjacent sides of the chip and allowed to completely flow under the chip filling the standoff gap. Depending on the underfill material's self filleting characteristics, fillets may need to be dispensed on the remaining sides of the chip. The underfill is now cured at a temperature ranging from 100–175°C. An epoxy-based resin filled with fused silica particles serves as the underfill encapsulation material. This material serves to compensate for the CTE mismatch between the board and the chip, as well as provide environmental protection. The capillary flow of underfill can be a time consuming process, especially when considering larger chip sizes. In addition, longer fill times require the use of a material with an extended

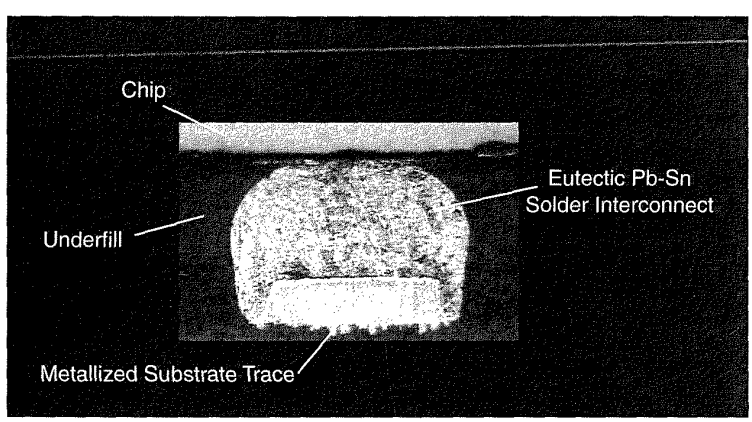


FIGURE 9.23 Typical controlled collapse chip interconnect based on a eutectic solder bump and a Cu-Ni-Au metallized trace.

Samsung, EX1028, p. 225

cure time. Once the chip has been underfilled, the assembly is placed into a cure oven, where the underfill cure process takes place.

The second side of the substrate is now processed using a similar process sequence. After second side assembly, a functional test or in-circuit test of the assembly takes place. If there are yield defects, the surface mount components can be reworked. Underfilled flip chips are not generally reworkable, but with the introduction of new thermally reworkable underfill systems, this will be a reality. During rework, the components and solder interconnections are heated, and the component is removed from the substrate. Once the component is removed, the site is redressed with the appropriate solder volumes consistent with assembly requirements.

Next, the electronic systems go through burn-in, where the system is operated at temperature, sometimes elevated temperatures, to identify infant mortality failures. After burn-in, final assembly takes place, where the electronic substrate is integrated with remaining system to form a product. Final assembly may include connection of peripherals, assembly of key pads, assembly of displays, and mounting of product covers. After final assembly, a completed product—cellular phone, computer workstation, pager, engine control module, etc.—is ready for shipment.

9.7.8 Electrical Performance

Flip chip interconnection provides the shortest possible chip-to-package interconnection distance. As such, its electrical performance tends to be quite good in terms of minimum impedance, minimum resistance, minimum capacitance, and minimum inductance. However, electrical performance of flip chip assemblies is highly dependant on the interconnection design layout and material systems selected, including the interconnection materials, underfill, substrate metallization, substrate solder mask, chip passivation, etc. Hence, care should be taken when designing the system and selecting the materials solution.

9.7.9 Reliability

In general, flip chip interconnection systems can range in reliability from high reliability to adequate reliability to meet the demands of a specific application. IBM's flip chip assemblies to ceramic have demonstrated high reliability. The chip metal is effectively sealed by the flip chip process and makes the chip resistant to common environmental exposures. Solder fatigue due to thermal cycling is the primary reliability issue for flip chip assemblies. Underfilling of chips greatly enhances their fatigue life, however, it drives an alternate failure mechanism of interfacial delamination of the underfill, followed by rapid solder fatigue. Other reliability concerns for flip chip include alpha particle emission and increased *electrostatic discharge* (ESD) sensitivity.

The encapsulation of flip chip structures is required primarily for the compensation of the CTE mismatch between silicon and high CTE substrates. Thermal mismatch is the cause of significant shear strain on the interconnections during thermal cycling. These lead to plastic stress damage, fatigue cracks, and eventual electrical failure. Underfill materials form a rigid layer between the chip and substrate, and enhance the thermal shock reliability of a flip chip assembly. In other words, these materials mechanically

Samsung, EX1028, p. 226

couple the chip and substrate, and locally constrain the CTE mismatch that exists between the two.

Flip Chip Defects and Failure Modes: A number of defects and failure modes are commonly observed for flip chip interconnection systems and assemblies; most apply both to flip chip on organic laminate and flip chip on ceramic assemblies. However, due to the large CTE difference for flip chip on laminate, they are typically observed earlier in the product life for organic laminate structures over their ceramic counterparts.

A number of common failure modes have been observed in flip chip devices. The most common failure mode is *delamination* at the underfill to die passivation interface that leads to rapid fatigue crack growth in the solder joints. A second failure mode is *die cracking,* and this generally occurs in one of two places. The most common form of die cracking occurs along the edge of the device. A center die crack is less common. Other failure modes include *underfill fillet cracking* and *solder migration.* Failure modes are generally common to particular materials, although there are exceptions.

Delamination/Void Growth: Delamination is a very common failure mode on interfaces in flip chip assemblies. It typically occurs between the chip passivation to underfill interface (more common) or the underfill to solder mask material interface. Delamination can also occur at the underfill/solder interconnection interface or solder mask/substrate interface. Delamination is caused by low adhesion strength between the surfaces. This low interfacial adhesion can be caused by cleanliness issues, flux/underfill incompatibility, underfill to passivation incompatibility, underfill to solder mask incompatibility, etc. Dust particles, inorganic contaminants, and organic contaminants can cause underfills not to adhere effectively. Material incompatibility can also lead to adhesion problems, increasing the likelihood of delamination occurring during either operational life or during reliability testing. Figures 9.24 and 9.25 show a delamination failure at the chip passivation/underfill interface via *C-mode scanning acoustic microscopy* (C-SAM). The cross-sectioned bump corresponds to the bump indicated on the C-SAM image. Notice that the bump in the C-SAM image is nearly indistinguishable, which is the result of the

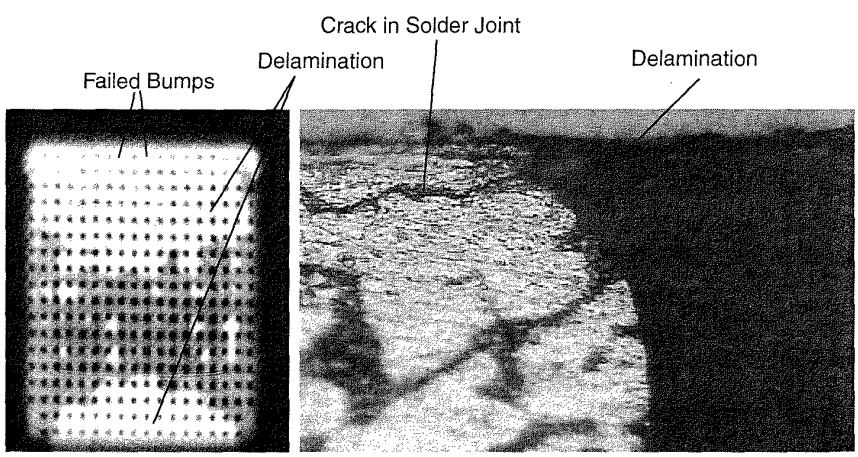


FIGURE 9.24 Delamination of the die to underfill interface (left) causing solder joint crack (right).

Samsung, EX1028, p. 227

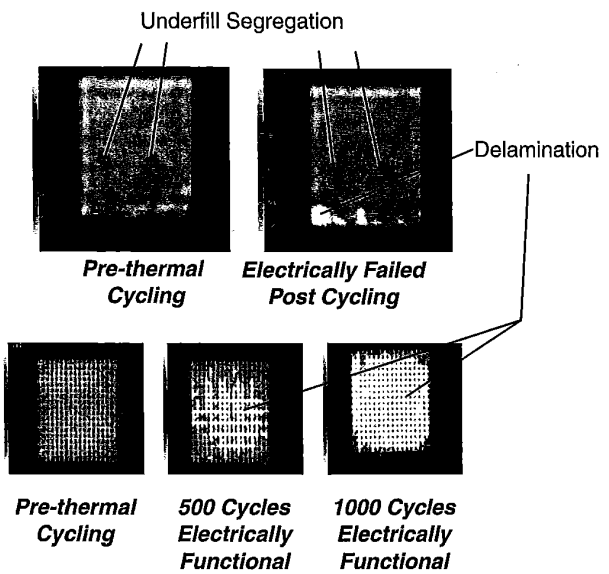


FIGURE 9.25 Propagation of a typical delamination occurring at the chip to underfill interface.

large fatigue crack shown in the micrograph on the right. The delamination is a very thin region between the silicon die passivation and the top of the underfill. As the delamination increases, more stress is applied to individual solder bumps leading to rapid solder joint fatigue failure. In the majority of cases, solder joint fatigue cracks occur on the chip side.

Solder Migration: Solder migration may occur during flip chip processing, reliability testing or operational life. Solder migration during processing is generally caused by voids, or delamination next to a joint, and is more common in processes where the cure step is done at high temperatures or after thermal excursions. These cure steps are common to some underfill materials that require long high-temperature cures or materials that cure during a reflow cycle, such as snap-cure underfills and no flow underfills. Solder migration also occurs in parts exposed to multiple reflows during processing or preconditioning. Migration can also occur in reliability testing when the solder joints are given time to creep, especially in air-to-air thermal cycling. In all cases, solder migration is due to the expansion of the solder at high temperatures, combined with a void adjacent to the solder. As the solder expands, or conversely the underfill shrinks, the solder is forced into the void. The solder relieves internal pressure by extruding into the adjacent void. The solder can be in a melt state, or viscoplastic solid state, for solder migration to occur. As material creeps or flows into the void, the effective stress concentrations on the solder joint rise during subsequent thermal cycling. Solder migration generally shortens the device life, but can be avoided by minimizing or eliminating voids or delamination next to solder joints by changing underfill pattern, substrate design, cure temperatures, underfill adhesion, etc. In Figure 9.26, the voiding occurred during a high ramp rate cure of a snap cure underfill in a conventional reflow oven. Due to the voids next to the solder joints, the solder expands into these areas during the spike of the profile. In the upper right hand corner, an entire solder joint migrates into the void.

Samsung, EX1028, p. 228

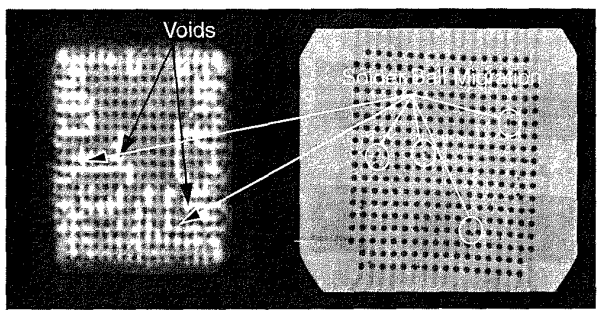


FIGURE 9.26 Solder ball migration due to voids caused by the cure. C-SAM image on left and x-ray on the right.

In the case of a solder bridge, a bridge of material joins two adjacent bumps together. In this case, a capture void between the two bumps allowed the solder to bridge. In the case of Figure 9.27, the bridge occurred during J-Standard 020 Moisture Sensitivity testing, which requires three reflow cycles after extended moisture exposure.

Die Cracking: Die cracks occur in flip chip assemblies due to excessive stresses on the brittle semiconductor, typically silicon, in combination with surface defects on the semiconductor chips that reduce the fracture strength of the single crystal semiconductor materials. The die cracks occur when the maximum principle tensile stress in the silicon reaches the effective fracture strength of the single crystal silicon. Die cracks are a catastrophic failure for active die. There are generally two types of die cracks, center cracks and edge cracks, although other active surface die cracks can occur in the presence of high surface stresses due to metallization layers, active circuitry, and chip passivation. The more common die crack occurs along the edge of the flip chip device; both types of die cracks are explained in more detail below.

Edge Cracking: An edge die crack can be caused by local defects resulting from wafer dicing. These localized defects include blade scoring, chipping, die chip out that cause surface cracks, subsurface damage, and other abnormalities along the edge of the flip

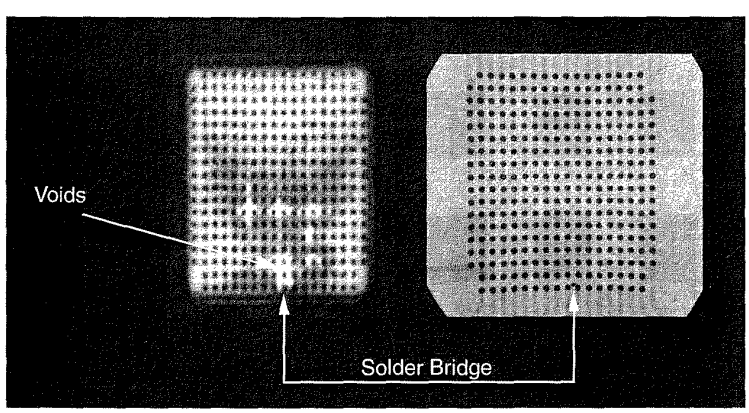


FIGURE 9.27 Solder bridge in a capture void that occurred during underfill dispensing. C-SAM (left) and x-ray (right).

Samsung, EX1028, p. 229

chip device. As the devices are cycled, these defects have stress concentrations due to their irregular features, and can cause the silicon to crack at lower stresses than defect-free silicon. In Figure 9.28, all of the rows of bumps are visible in the acoustic image taken prior to edge crack formation (left), while the outer most row at the bottom of the die has almost disappeared after crack formation (right).

Center Die Cracks: An example of a center die crack is shown in Figure 9.29. A center die crack can take one of two forms. In the first case, the crack starts at the surface and works its way down to the underfill interface, as shown in Figure 9.29. In the second case, the crack again starts at the surface but only penetrates to a neutral plane, at which point the crack moves horizontally to the die edge. Similar crack propagation can occur with edge cracks. A center die crack is a catastrophic failure, caused by bending stresses on the die due to the coupling of the die to the substrate and the relative CTE mismatch. These bending stresses can be compounded by imperfections in the die, lowering the effective fracture strength of the silicon. In Figure 9.29, the die crack propagated from a pick mark on the top of the die, caused by the die pick system used to remove the die from the wafer tacky tape. Die cracking can also be caused by large fillets on the chips that cause high stresses to be concentrated on the top of the die rather than through the bulk of the die. Another cause of die cracking is the underfill material, as some underfills can exert very high stress concentrations upon the die.

Fillet Cracking: Underfill cracking is a common failure mode for no flow underfills, although it also occurs in filled capillary underfills. Fillet cracking occurs mainly in

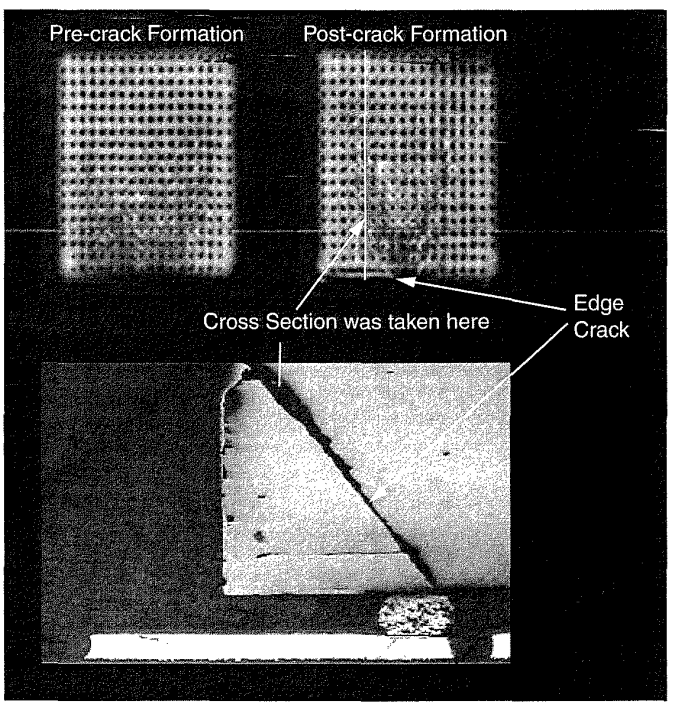


FIGURE 9.28 C-SAM image before (left) and after (right) edge crack formation and cross section of edge crack (bottom).

Samsung, EX1028, p. 230

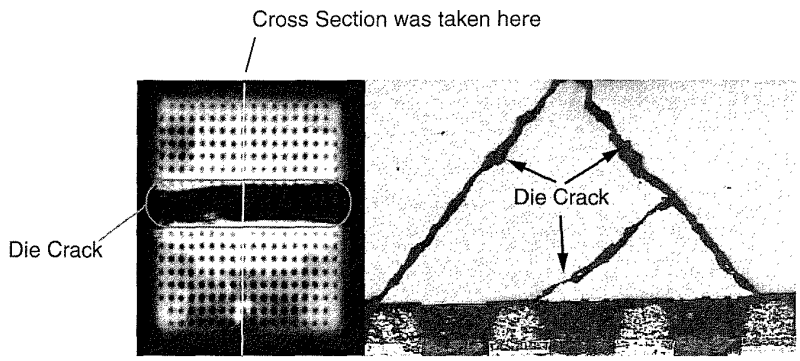


FIGURE 9.29 C-SAM image (left) and cross section (right) of a device with a die crack through the center.

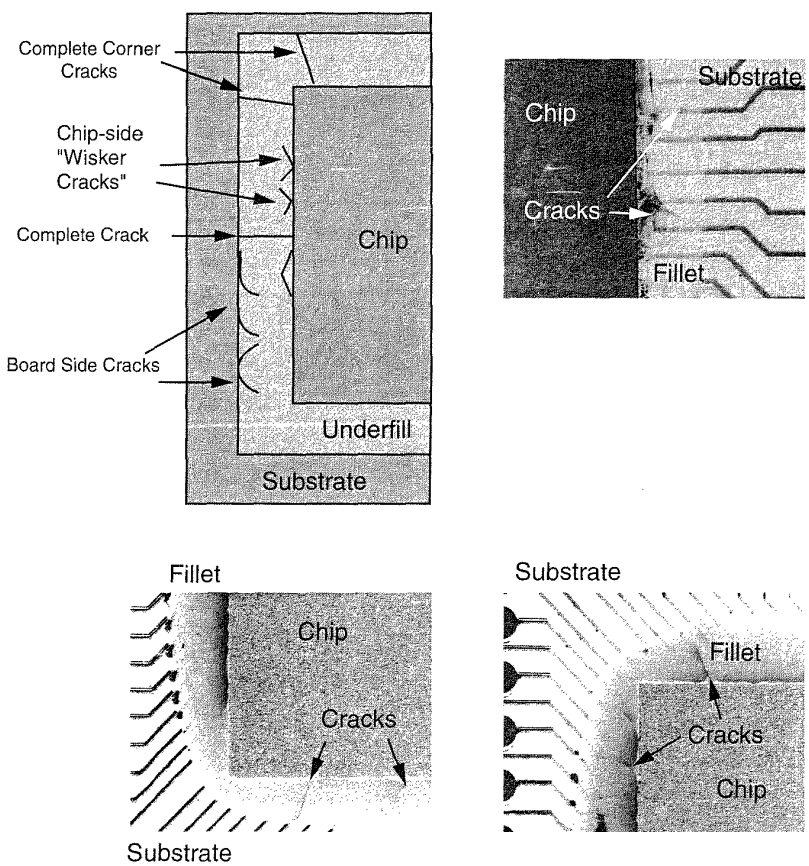


FIGURE 9.30 Schematic and examples of underfill fillet cracks.

Samsung, EX1028, p. 231

assemblies subjected to thermal cycling. Fillet cracks may be classified as one of three types: chip side cracks, board side cracks, and complete cracks. Figure 9.30 shows examples of the three crack types.

Chip side cracks are small cracks that originate at the chip/underfill interface and usually form a V-shaped pattern that extends partially down the fillet "whisker cracks." Board side cracks are short cracks that extend up the fillet or run parallel to the fillet edge. Chip side cracks can lead to small areas of delamination between the underfill and the side of the chip, while board side cracks can lead to small areas of delamination between the underfill and substrate. A complete fillet crack extends from the top of the fillet to the substrate and propagates completely through the fillet. In cases of extreme cracking, the underfill fillet separates from the chip, eliminating the stress relief properties of the underfill, thereby promoting rapid delamination of the underfill at the chip to underfill interface.

Solder Fatigue Cracking: Solder fatigue cracking is also a common failure for flip chip assemblies. Fatigue is the initiation and propagation of solder cracks caused by repeated cycling from a high positive stress state to a high negative stress state. The end result of solder fatigue in flip chip devices is a crack through the interconnection, as shown in Figure 9.31. The solder fatigue failure can occur as the primary failure mode, or as a secondary failure mode resulting from some other primary failure mode. The most common case is solder fatigue as a direct result of the increasing delamination of the underfill to die passivation interface. This delamination of the underfill causes stress concentrations at the solder joints in the delaminated region. The fatigue causes a crack which propagates through the interconnection, eventually causing an electrical open. Solder fatigue generally occurs on the die side as opposed to the substrate side. Solder fatigue typically occurs more rapidly in air-to-air testing than in liquid-to-liquid testing.

Bulk Underfill Cracking: In addition to underfill fillet cracking, cracks may be seen in the bulk underfill underneath the chip occurring during thermal excursions or thermal cycle testing. Typical cracks in underfills are shown in Figure 9.32. It is important to notice that the cracks typically occur adjacent to the solder joints propagating between

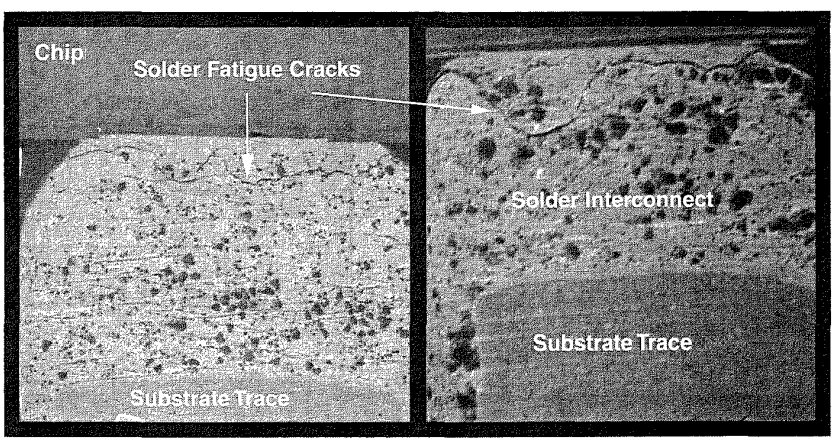


FIGURE 9.31 Typical solder fatigue cracks in eutectic solder joints.

Samsung, EX1028, p. 232

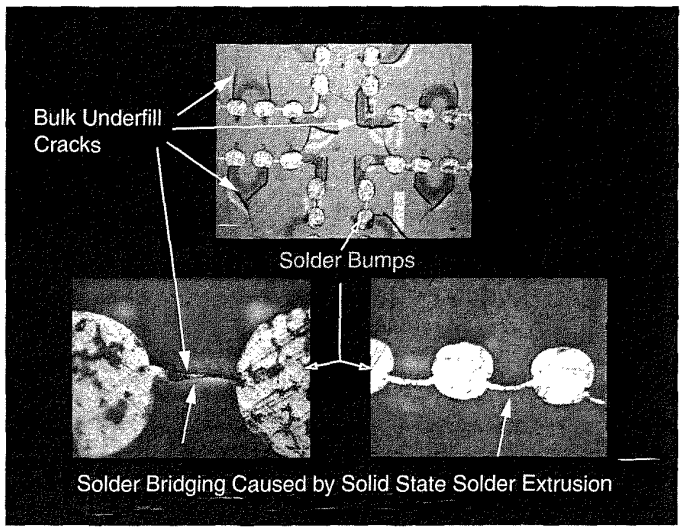
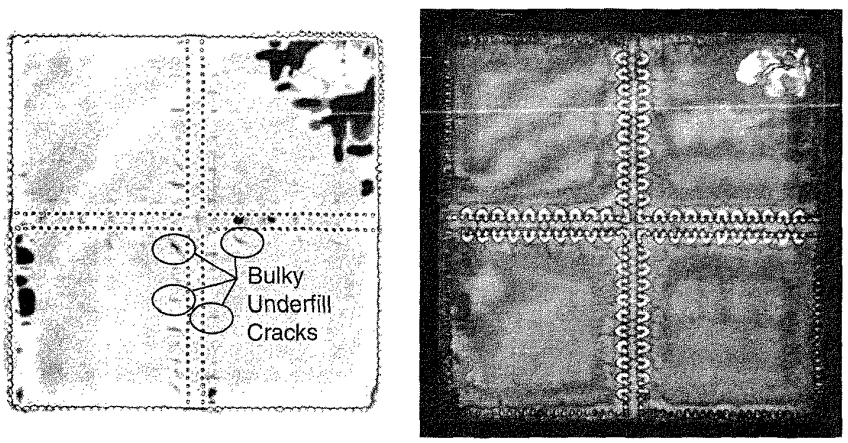


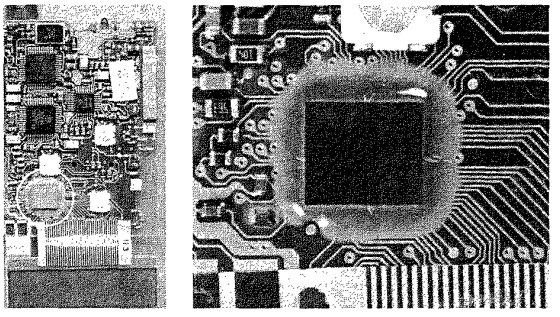
FIGURE 9.32 Examples of bulk underfill cracking, solder extrusion into underfill cracks between adjacent bumps.



Chip to Underfill Interface Underfill to Board Interface

FIGURE 9.33 C-SAM images of the chip/underfill interface (left) and underfill/board interface (right) showing underfill delamination and bulk underfill cracking.

Samsung, EX1028, p. 233



(a) (b)

FIGURE 9.34 Flip chip on organic circuit board application. (*Source: Prismark Partners*)

the joints. In addition to cracks propagating between solder joints, cracks also initiate and propagate along the metallization pattern on the substrate as seen in the second image in Figure 9.32b. Moreover, it has been observed that solder from the joints tends to extrude into underfill cracks during cycling, which has the potential to cause electrical shorting as shown in the figure.

Figure 9.33 shows a series of C-SAM images at the chip to underfill interface, where fine lines near the solder joints appear, and expand as cycling continues. Cross sections show that these fine lines correspond to cracks in the bulk underfill, indicating that C-SAM analysis can be used to detect bulk underfill cracks.

9.7.10 Product Examples

Flip Chip on Board and Surface Mount: Figure 9.34 shows a flip chip on organic circuit board application. In this case, the solder interconnection is mounted directly on a PWB, underfilled and overmolded with a glob top encapsulant.

Figure 9.35 shows another flip chip on board application consisting of a Casio portable radio. In this application, five flip chip die are interconnected to the PWB using conductive adhesive.

Low Cost Underfill Process: Since underfills are necessary for the reliability of flip chip to organic board, key factors to be considered in production are processability (e.g.,

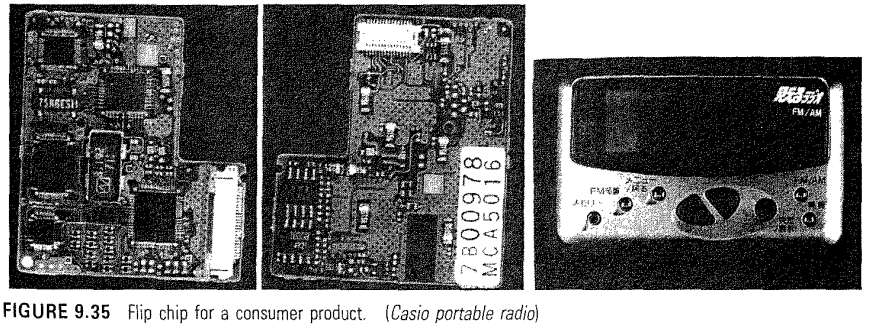


FIGURE 9.35 Flip chip for a consumer product. (*Casio portable radio*)

Samsung, EX1028, p. 234

flow time, cure time, shelf life, flow characteristics) and reliability (e.g., moisture sensitivity, adhesion, thermal, mechanical, corrosion, electromigration). From a processability standpoint, some of the main reasons for incompatibility between FCOB and SMT are the underfill flow time and the underfill cure time. Conventional underfills can take anywhere from 1–5 min to flow under a 5-mm^2 die, while requiring a cure time of 30–90 min at temperatures ranging from 125–165°C. New commercial fast-flow underfills take between 2–30 s to flow under a 5-mm die. Cure times have also been cut with the introduction of snap-cure underfills. The cure schedules have been reduced to 5–10 min at temperatures between 150–165°C. A target processing metric for high throughput flip chip processing based on fast-flow, snap-cure underfills is a maximum flow time of 5 s on a 5-mm chip, and a maximum cure time of 5 min.

Fast Flow Snap-Cure Underfills: The advent of snap-cure underfills has also made possible a new flip chip assembly process leveraging double-sided board assembly. The new low-cost flip chip assembly process is shown in Figure 9.36. The process flow is to print paste for the top side of surface mount technology (SMT) components, place the SMT components, flux the flip chip bumps or bond site, place the flip chips, reflow the top side forming interconnections, and perform an electrical test. The in-circuit test allows rework of any failed devices prior to dispensing underfill, but can be eliminated based on throughput requirements. If there are no defects, the underfill is dispensed on the flip chip sites, underfill fillets are dispensed if necessary, the board is flipped, the second side SMT components are processed, and the second side reflow is used to simultaneously cure the underfill and form the second side interconnections. The second side reflow profile is slightly modified to accommodate underfill snap-cure in the soak stage of the profile. Typically, this involves increasing the soak time and possibly the soak temperature of a conventional eutectic solder reflow profile. A key element of the process is verifying that the underfills are fully converted during the soak temperature exposure. This is done using *differential scanning calorimetry* (DSC) analysis to verify underfill conversion. The new flip chip assembly process is much more conducive to high throughput, double-sided SMT than conventional flip chip processes.

No Flow Underfills: In many cases, present state-of-the-art flip chip-on-board assembly technology is not capable of achieving the high throughputs required for integrated high volume SMT processing and low-cost electronics packaging. Elimination of processing steps will enable high throughput, reduce process complexity, reduce capital equipment requirements, reduce equipment maintenance, and increase process robustness. Potential processing steps for elimination are flux application, underfill flow, flux residue cleaning, underfill fillet processing, and secondary thermal curing of underfill. Highly populated flip chip assemblies and large chips compound current flip chip process limitations. Based on semiconductor and packaging roadmaps, it is clear that trends for larger devices and higher packaging densities will continue to increase, thereby inhibiting cost effective flip chip assembly based on conventional process techniques. Underfill flow and cure processes dramatically reduce assembly line throughput, and therefore exert significant pressure on profit margins. This has been verified by cost modeling and analysis comparing a typical industry flip chip process, the proposed low-cost high-throughput process, and surface mount assembly.

To enable high throughput that is compatible with high-speed in-line automated assembly, a new process shown in Figure 9.37 has been developed.

Samsung, EX1028, p. 235

FIGURE 9.36 Fast flow snap-cure underfill process.

MEMS231736

CONFIDENTIAL BUSINESS INFORMATION

Samsung, EX1028, p. 236

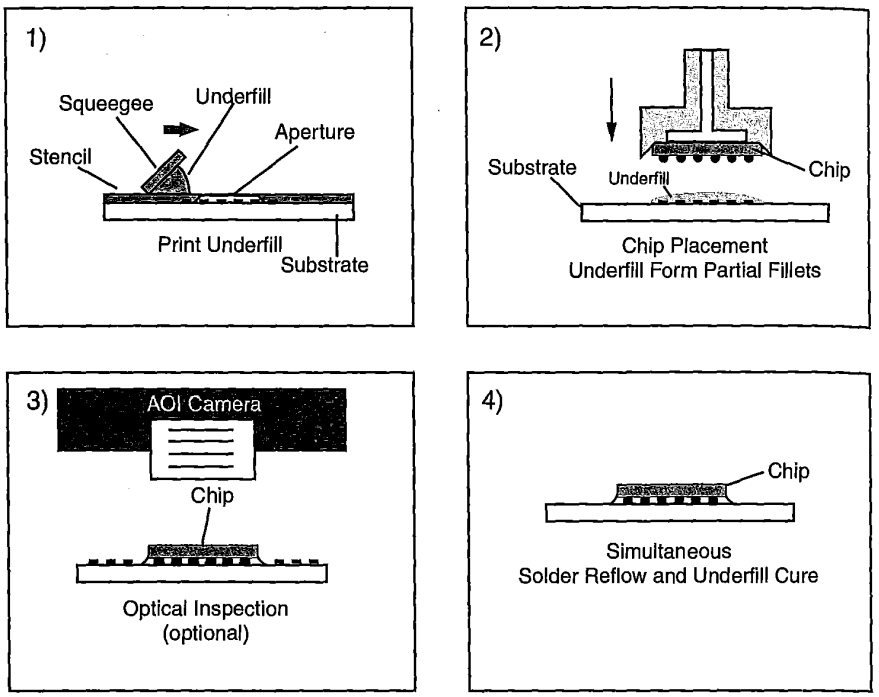


FIGURE 9.37 No-flow underfill process.

The overall process flow begins with a known good substrate and die. A controlled volume of underfill material is stencil-printed over the bond pads on the substrate, an area-based process compared to conventional point-to-point underfill dispensing. Next, solder paste is printed onto the board using a blind stencil for SMT component assembly. Alternately, the no-flow underfills can be dispensed on the chip site post solder paste printing. Next, the SMT components are placed. The bare chips are then aligned using a vision system to orient the chips relative to bond site fiducials on the substrate. The chips are placed on the substrate, compressing the liquid underfill. Note, the placement forces required are well within those of commercial high speed flip chip placement systems, and the process requires no special high force capability, unlike anisotropic conductive adhesives or compression bond adhesives processing. A significant advantage here is the elimination of lengthy capillary flow times for underfill processing, particularly with large devices. Finally, the solder interconnections are reflowed simultaneously while curing the polymer encapsulant underfill. Care is taken here to prevent premature gelation of the underfill prior to solder reflow, using innovative underfill materials called no-flow underfills, providing latent gelation such that the cure reaction is inhibited until a critical temperature is reached above the solder liquidus temperature. Moreover, it is critical that proper fluxing action be achieved during reflow. The major advantages of the new process are that it increases throughput of flip chip processing, transforms a chip-based process into an area-based process, reduces the required number of processing steps, and reduces the ratio of wafer-to-assembly cost by two to five times over the competing processes.

Samsung, EX1028, p. 237

TABLE 9.1 Chip-level connection parameters.

	Wirebonding		TAB	Flip Chip
Connection Metallurgy	**Al**	**Au**	**Cu**	**Pb/Sn**
Resistance (ohms)	0.035	0.03	0.02	0.002
Inductance (nH)	0.65	0.65	2.10	0.200
Capacitance (pF)	0.006	0.006	0.04	0.001
I/O Density	400	400	400	1600
Rework	Poor	Poor	Poor	Good
Failure Rate (%/1000h)	1×10^{-5}	1×10^{-5}	N/A	$<1 \times 10^{-8}$

9.8 SUMMARY AND FUTURE TRENDS

IC assembly is the most important first step in the use of ICs. Wirebonding is the most used technology today, and will remain so in the foreseeable future, at least up to 700 I/Os on a single IC. For higher I/Os, flip chip or TAB are the dominant technologies. For very high I/Os beyond 2000, very high electrical performance, and very high reliability, flip chip becomes virtually the only viable IC assembly. While all these developments are taking place, an entirely new paradigm is taking shape that may affect these directions. This is referred to as wafer-level packaging, as presented in the next chapter.

Tables 9.1 and 9.2 summarize the key parameters of wirebond, TAB and flip chip as practiced in the year 2000. Table 9.1 includes the electrical parameters such as resistance, capacitance and inductance, as well as failure rate. Flip chip clearly dominates in all of these. The problem with its limited use has been due to lack of adequate infrastructure.

TABLE 9.2 Assembly technology comparison (2000).

Technology	Wirebond	TAB	FC-Solder	FC-Adhesive
Max. I/O (2000)	<700	800–1000	>1000+	>50+
Max. I/O (2010)	1000	2000	10,000	10,000
Min. Pitch Area Array (2000) (μm)	N/A	N/A	200	200
Min. Pitch Area Array (2010) (μm)	N/A	N/A	150	150
Min. Pitch Perimeter Array (2000) (μm)	75–100	50–75	165	165
Min. Pitch Perimeter Array (2010) (μm)	50–75	50	50	50
Footprint	Medium	Med.–Lg.	Small	Small
Pretest and Burn-in	Difficult	Good	Yes	Yes
Electrical Performance	Fair	Good	Excellent	Excellent
Thermomechanical Performance	Fair	Excellent	Good	Good
Repairability	Difficult	Yes	Yes	Yes
Infrastructure	Good	Growing	Minimum	Minimum
Approx. Cost	1.0	1.2	0.8	0.4

Samsung, EX1028, p. 238

Table 9.2 indicates the I/O range, pitch, pretest and burn-in and relative cost in volume manufacturing of the three IC assembly technologies. The projections for 2010 are included for I/Os and pitch.

9.9 HOMEWORK PROBLEMS

1. Discuss the advantages and disadvantages of IC assembly-based wirebond interconnections versus flip chip interconnections. Focus your discussion on cost, equipment requirements, infrastructure, process cycle time, chip interconnection counts, and electrical performance.
2. Estimate the cycle time on a per-package basis for construction of two different plastic packages housing the same IC chip. The IC has 1500 I/O and is 12 by 14 mm in size. The first package is a wirebonded lead frame plastic quad flat pack (PQFP). The cycle time estimate should include the die attach through lead planarization steps in Figure 9.3. Assume the transfer molding operations has a cycle time of 60 seconds. Since both package production methods require a test and burn-in step, neglect it in the cycle time comparison. Use reasonable cycle time approximations for the stamping and lead processing steps. Note that prior to singulation, the packages are processed in multiple up format. After singulation, each package is processed separately. The second package is a flip chip plastic ball grid array (PBGA) package shown in Figure 9.P2. In estimating the process cycle time, assume that the high density organic substrates (PWBs) are supplied to the assembly contractor.
3. Estimate the standoff gap height, h, for the FBT250, FBT500, PB8-2, PB8-4, FA10-2 and FB10-4 flip chip components listed in Table 9.P3 based on a controlled collapse connection of eutectic solder. Base the analysis on a one-to-one chip pad to substrate pad design rule (i.e., equal bond pad areas on the chip and substrate). Also, estimate the flow gap for capillary flow underfill processing, assuming that the solder mask thickness is 0 μm and 25 μm.

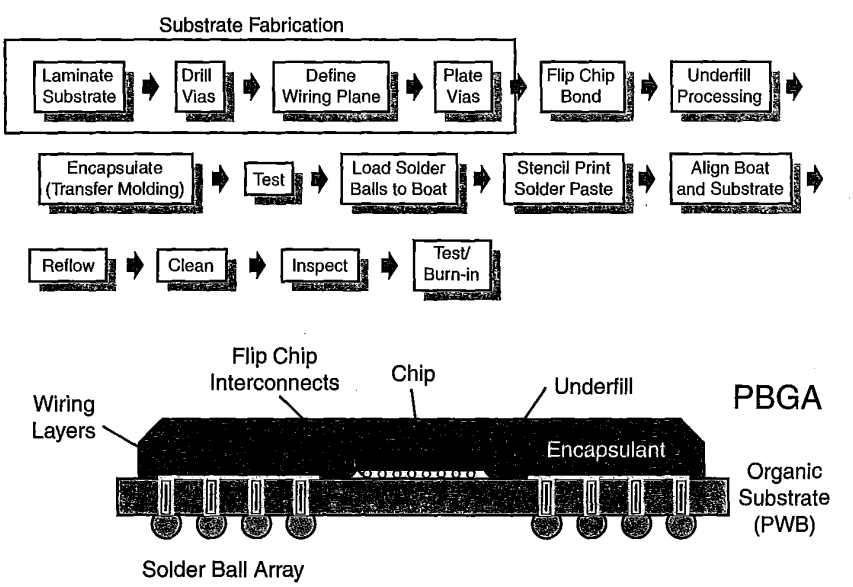


FIGURE 9.P2 Example of a plastic ball grid array package fabrication process and schematic.

Samsung, EX1028, p. 239

TABLE 9.P3 Flip chip parameters.

Test Die (I/O Format)	Die Size (mm)	Bump Pitch (μm)	Passivation Via (μm)	UBM Diameter (μm)	Bump Height (μm)	Bump Diameter (μm)	Number of Bumps	Max. Bump Diagonal (mm)	Die Metal Pad Size (μm)	Die Metal Type
FBT500 (Perimeter Array)	12.6	457	102	178	140	190	96	—	193 × 193	98/1/1 Al/Cu/Si
FBT250 (Perimeter Array)	6.3	457	102	152	140	190	48	7.7	193 × 193	98/1/1 Al/Cu/Si
FA10-2 (Area Array)	4.98	254	80	102	120	135	317	7.9	127 × 127	98/1/1 Al/Cu/Si
FA10-4 (Area Array)	10.1	254	80	102	120	135	1268	7.9	127 × 127	98/1/1 Al/Cu/Si
PB8-2	4.98	203	80	102	95	120	88	—	115 × 115	98/1/1 Al/Cu/Si
PB8-4	10.1	203	80	102	95	120	352	—	115 × 115	98/1/1 Al/Cu/Si

Samsung, EX1028, p. 240

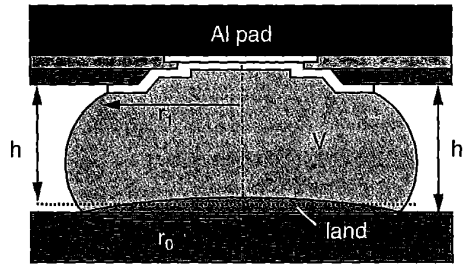


FIGURE 9.P3 Schematic of a controlled collapse solder connection.

Estimation of the standoff gap height is simplified by approximating the solder interconnection shape as a sphere, with the upper and lower sections removed, defined by the chip and substrate pads, as shown in Figure 9.P3. Assuming the chip pad size has a radius of r_1, the substrate pad size has a radius of r_0, and the solder joint volume is V, geometric analysis yields the standoff gap height given by:

$$V = \frac{\pi h}{6}\left[h^2 + 3(r_0^2 + r_1^2 v_1)\right] \tag{9.8}$$

where

$$V = \frac{4}{3}\pi R^3 - \frac{\pi}{3}v^2(3R - v) \tag{9.9}$$

$$r_1^2 = v(2R - v)v_1 \tag{9.10}$$

R = radius of the chip solder bump prior to flip chip assembly and reflow

What alternate analysis could be used to estimate the standoff gap height based on the chip equilibrium position?

4. Estimate the underfill volume *and* mass required for the FBT500, FBT250, PB8-2, PB8-4, FA10-2 and FA10-4 flip chip test chips based on the calculations from the problem 3 predictions, assembly schematic in Figure 9.P4, and Table 9.P3. Assume the density of the underfill $\rho = 1.05$ g/cm^3. Compare the volume of the underfill in the fillets with that under the chip. Discuss these values in relative terms, based on controlling underfill dispense volume and controlling the fillet width during the dispensing process.

The underfill volume can be estimated by accounting for the underfill volume filling the standoff gap under the chip, less the volume of the solder interconnections, plus the volume of the underfill in the fillets around the chip edges. The underfill volume can be approximated by:

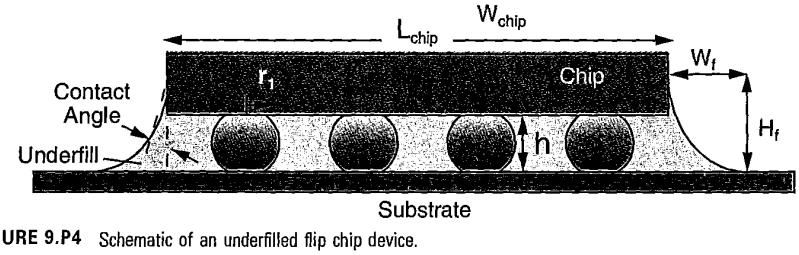


FIGURE 9.P4 Schematic of an underfilled flip chip device.

Samsung, EX1028, p. 241

$$V_{\text{underfill}} = V_{\text{standoff gap}} - V_{\text{bumps}} + V_{\text{fillet}} \tag{9.11}$$

$$V_{\text{standoff gap}} = L_{\text{chip}} W_{\text{chip}} h \tag{9.12}$$

$$V_{\text{bumps}} = N_{\text{bumps}} V \quad \text{(see problem 3)} \tag{9.13}$$

$$V_{\text{fillet}} = \lfloor 2(L_{\text{chip}} + W_{\text{chip}})(H_f W_f) + (W_f^2 H_f)/3 \rfloor \text{SF} \tag{9.14}$$

SF = shape factor

Contact Angle (°)	Shape Factor
1	0.1109
5	0.3202
10	0.4739
15	0.5849
20	0.6742
25	0.7505
30	0.8187
35	0.8817
40	0.9416

5. Estimate the underfill flow time, t_{fill} based on capillary flow underfill processing for the FBT500, FBT250, PB8-2, PB8-4, FA10-2 and FA10-4 flip chips using predictions calculated in problems 3 and 4. Assume the wetting angle α is equal to the contact angle used for problem 4. Assume the underfill viscosity $\mu = 0.8$ Pa s, and the underfill surface tension $\sigma = 33$ mN/m. Discuss the differences between the values predicted for each chip in terms of the chip layout and size. What issues do you see affecting the predictive accuracy of the underfill flow analysis relative to the chip layout?
6. Estimate the underfill flow time, t_{fill} based on injection flow underfill processing for the FBT500, FBT250, PB8-2, PB8-4, FA10-2 and FA10-4 flip chips using predictions calculated in problems 3 and 4. Assume the wetting angle α is equal to the contact angle used for problem 4. Assume the underfill shear thins under higher shear rates and follows the power law approximation with the flow parameters of $m = 0.8$ Pa $s^{0.74}$ and $n = 0.74$. Assume, the underfill surface tension is $\sigma = 33$ mN/m. Discuss the differences between the values predicted for each chip in terms of the chip layout and size. What issues do you see affecting the predictive accuracy of the underfill flow analysis relative to the chip layout?
7. This problem focuses on the design of the flip chip placement process for the High Throughput Flip Chip Assembly Process Based on No Flow Underfills using a compression flow underfill process. Based on compression flow underfill process, the underfill film under the chip produces a hydrostatic pressure under the chip resisting the placement head motion. The force exerted by the underfill can be estimated, based on squeeze flow of a Newtonian fluid using quasi-steady-state approximations.

   The force the underfill exerts on the chip, F_{fluid}, can be approximated with the Stefan equation below, assuming the chip and substrate surfaces are parallel. In the Stefan equation, the underfill viscosity is μ, the chip is modeled as having an effective radius of R where the chip and substrate are separated by a distance h_1.

$$F_{\text{fluid}} = \frac{3\pi R^4 \mu(-\dot{h})}{h_0^3} \tag{9.15}$$

   The effective chip radius, R, can be approximated assuming a circular equivalent area of the square chip of size L.

Samsung, EX1028, p. 242

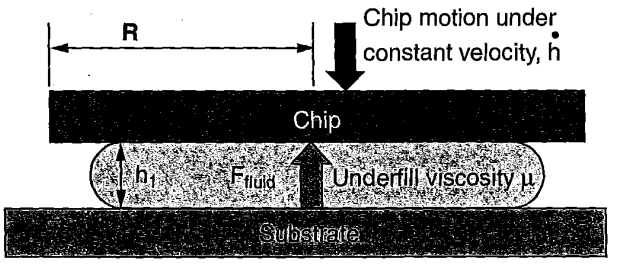


FIGURE 9.P7a Flip chip motion under constant applied chip velocity.

$$R = \sqrt{\frac{L^2}{\pi}} \tag{9.16}$$

The Stefan equation can be used to estimate the force a placement system must exert to place a chip based on the compression flow process. Typically, placement machines move the gantry at high speed to the chip position determined by the vision system, based on substrate fiducials. The chip is then moved at high speed towards the substrate to a programmed search height. Now the chip is moved towards the substrate at a constant programmed velocity, $\dot{h}$, until the placement head force sensor reaches a preprogrammed value, $F_{\text{placement}}$. A schematic of the process is shown in Figure 9.P7a. The placement machine then maintains the set force for a specified length of time, the dwell time or bond time, τ. Using the Stefan equation, F_{fluid} can be calculated at any height, h_1, under conditions where the chip is fully wet by the underfill. Similarly, the Stefan equation can be used to determine the height h_1 where the underfill hydrodynamic force equals that of the preprogrammed placement force limit, i.e., $F_{\text{fluid}} = F_{\text{placement}}$.

Once $F_{\text{fluid}} = F_{\text{placement}}$, the chip moves with a constant force, $F_{\text{placement}}$, rather than a constant speed. The motion of the chip under constant applied force is described by the equation below and illustrated in Figure 9.P7b. Here τ estimates the time required for the chip to move towards the substrate from h_1 to h_f where h_f is equal to the flip chip bump height, i.e., $h_f = h_{\text{bumps}}$. This expression is derived from the Stefan equation by solving the differential equation in h with an initial condition of $h(t = 0) = h_1$. The parameter τ is the time $F_{\text{placement}}$ is applied, h_1 is the initial separation of the chip and substrate at the point where $F_{\text{fluid}} = F_{\text{placement}}$, and h_f is the final separation distance between the chip and substrate equal to the bump height.

$$\tau = \frac{3\pi R^4 \mu}{2F_{\text{placement}}}\left[\frac{1}{h_f^2} - \frac{1}{h_0^2}\right] \tag{9.17}$$

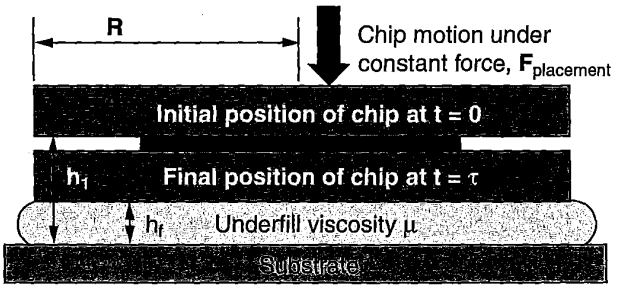


FIGURE 9.P7b Flip chip motion under constant applied placement force.

Samsung, EX1028, p. 243

Based on these expressions, design the compression flow chip placement process for two cases. In both cases, the viscosity of the underfill is 1.5 Pa s. Base the analysis on the PB8-2 and PB8-4 flip chip devices in Table 9.P3. For the first case, assume the placement velocity is 5 mm/s and the placement force is 400g, (i.e., $F_{placement} = F_{fluid} = 3.92$ N). In the second case, assume the placement velocity and placement force are 0.1 mm/s and 1600g (i.e., $F_{placement} = F_{fluid} = 15.7$ N), respectively.

For each case, estimate the distance above the substrate at which point the placement force limit is reached. Based on these results, estimate the minimum dwell time required to ensure the chip bumps make contact with the substrate bond pads. Which placement process variable set would you select? Discuss the pros and cons of each case.

9.10 SUGGESTED READING

1. W. D. Brown, ed. *Advanced Electronic Packaging,* New York: IEEE Press, 1999.
2. Doane and Franzon, ed. *Multichip Module Technologies and Alternatives,* Van Nostrand Reinhold, 1993.
3. Electronic Materials Handbook, *Volume 1—Packaging,* ASM International, 1989.
4. G. Harman. *Wirebonding in Microelectronics: Materials Processes, Reliability, and Yield,* 2nd ed., New York: McGraw-Hill, 1997.
5. C. A. Harper, ed. *Electronic Packaging and Interconnection Handbook,* McGraw-Hill, 1991.
6. Johnson, Teng, and Balde, ed., *Multichip Modules,* IEEE Press, 1990.
7. J. Lau, ed. *Ball Grid Array Technology,* McGraw-Hill, 1995.
8. J. Lau, ed. *Chip on Board Technologies for Multichip Modules,* Van Nostrand Reinhold, 1994.
9. J. Lau, ed. *Flip Chip Technologies,* McGraw-Hill, 1996.
10. J. Lau, ed. *Handbook of Fine Pitch Surface Mount Technology,* Van Nostrand Reinhold, 1994.
11. J. Lau, ed. *Handbook of Tape Automated Bonding (TAB),* Van Nostrand Reinhold, 1992.
12. J. Lau, ed. *Solder Joint Reliability: Theory and Applications,* Van Nostrand Reinhold, 1991.
13. J. Lau and Pao, ed. *Solder Joint Reliability of BGA, CSP, Flip Chip, and Fine Pitch SMT Assemblies,* McGraw-Hill, 1997.
14. L. Levine. "Wirebonding," *Advanced Packaging,* Vol. 9, No. 4, 47–52, April 2000.
15. P. Messner, I. Turlik, G. Balde, and P. Garrou, ed. *Thin-film Multichip Modules,* ISHM, 1992.
16. R. P. Prasad. *Surface Mount Technology: Principles and Practice,* Van Nostrand Reinhold, 1989.
17. Seraphim, Lasky, and Li, ed., *Principles of Electronic Packaging,* McGraw-Hill, 1989.
18. R. R. Tummala, E. J. Rymaszewski, and A. G. Klopfenstein. *Microelectronics Packaging Handbook Part II: Semiconductor Packaging,* 2nd ed. New York: Chapman & Hall, 1997.
19. R. R. Tummala, E. J. Rymaszewski, and A. G. Klopfenstein. ed. *Microelectronics Packaging Handbook: Part 3 Subsystem Packaging,* 2nd ed. New York: Chapman & Hall, 1997.
20. C. H. Van Vlack. *Elements of Materials Science and Engineering,* 5th ed. Reading, MA: Addison-Wesley, 1989.
21. J. Vardaman, ed. *Surface Mount Technology-Recent Japanese Developments,* IEEE Press, 1993.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contribution of Mr. Paul Totta of IBM for critically reviewing this chapter.

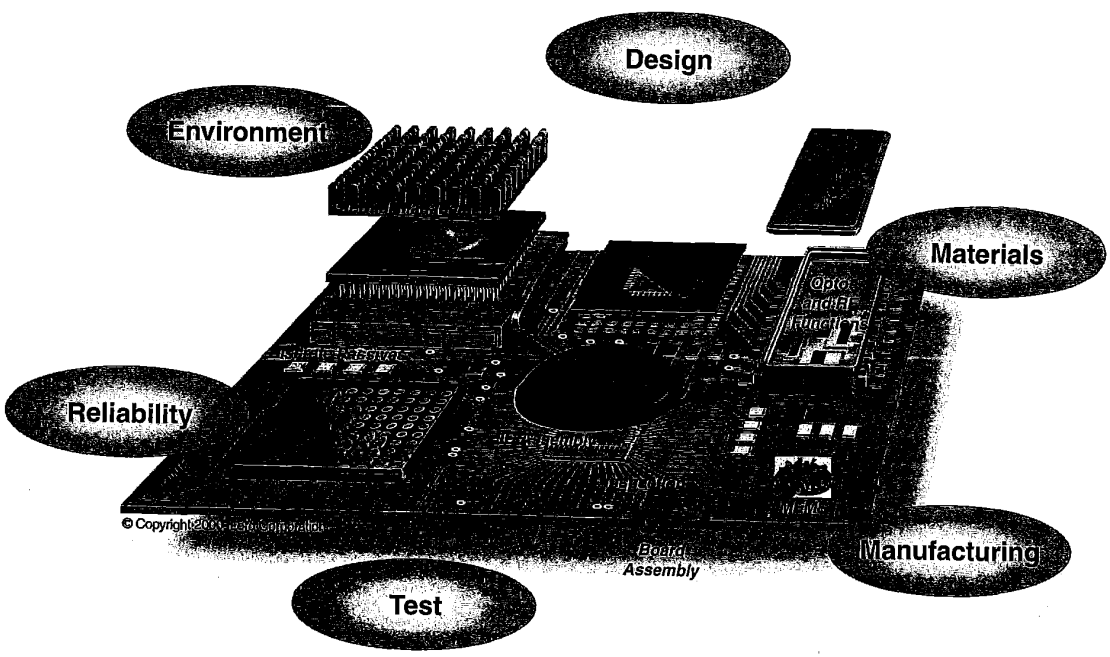
Samsung, EX1028, p. 244

CHAPTER 15

FUNDAMENTALS OF SEALING AND ENCAPSULATION

Prof. C. P. Wong
Georgia Institute of Technology

Dr. Treliant Fang
Motorola



15.1 What Is Encapsulation? What Is Sealing?
15.2 Why Is Encapsulation Necessary?
15.3 Fundamentals of Encapsulation and Sealing
15.4 Encapsulation Requirements
15.5 Encapsulant Materials
15.6 Encapsulation Processes
15.7 Hermetic Sealing
15.8 Summary and Future Trends
15.9 Homework Problems
15.10 Suggested Reading

CHAPTER OBJECTIVES

- Explain the need for encapsulation and sealing.
- Describe the basic encapsulation and sealing of electronic packages.
- Discuss various encapsulation materials and processes.
- Discuss hermetic and nonhermetic packaging and sealing processes.
- Describe various methods for hermetic sealing.
- Discuss how reliability is achieved without hermeticity with plastic packaging.

CHAPTER INTRODUCTION

Encapsulation and sealing are two of the major protecting functions of IC packaging. They are used to protect IC devices from adverse environmental and mechanical effects. This chapter provides some fundamental understanding of encapsulation and sealing processes used in IC packaging. The chemistry of some of the common encapsulants such as molding compounds, liquid and glob-top encapsulants and underfills are presented. The important material properties of these materials are reviewed. Finally, the difference in hermetic vs. non-hermetic packaging, and how to achieve reliability without hermeticity, is indicated.

Samsung, EX1028, p. 246

15.1 WHAT IS ENCAPSULATION? WHAT IS SEALING?

An electronic device cannot perform its designed functions until it is packaged such that it is interconnected with the rest of the system and protected. One such protection technique that has been commonly used is encapsulation, typically done by means of low temperature polymers. Encapsulation provides an economical way to protect device packages by isolating the active devices from environmental pollutants, and at the same time offering mechanical protection by structural coupling of the device to the constituent packaging materials into a robust package. Figure 15.1 shows one of the most common device packages, QFP (quad flat package), protected this way by a molding process. All these types of packages require encapsulation or sealing to protect the die within. The protection can be an organic overcoat, in which case it is called encapsulation as illustrated in Figure 15.1(a), or inorganic as illustrated in Figure 15.1(b), in which case it is called sealing. Encapsulation materials are typically molded onto the IC or dispensed under the die, such as with flip chip ceramic BGA packages. This type of organic coating is a very inexpensive way of protecting devices, but their protection is not permanent, typically controlled by permeation properties of the polymeric resins used. The inorganic sealing, however, is permanent, by being hermetic, but the cost of this process is high.

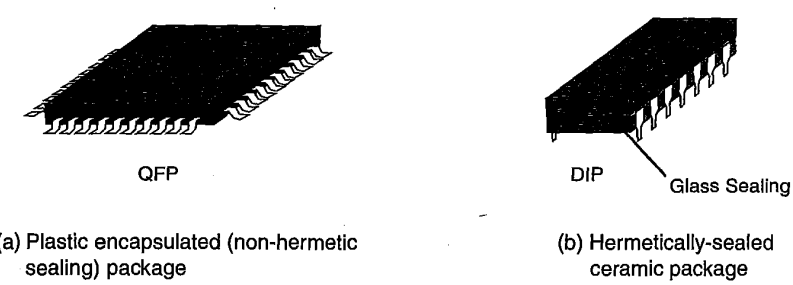
Performance of encapsulation is determined by its dimensional stability, its resistance to thermal excursions, its permeation providing isolation of environmental pollutants and its thermal dissipation providing dissipation of the heat generated by the packaged device. The progress in encapsulation technology during the last two decades has allowed plastic-encapsulated microcircuits to be used in some of the most demanding applications, including space-borne electronic systems, as a cost-effective device protection.

15.2 WHY IS ENCAPSULATION NECESSARY?

Encapsulation provides both chemical and mechanical protection of the IC, such that a reasonable life expectancy can be achieved under field conditions in automotive, telecommunications, computer, consumer, medical and other industries.

15.2.1 Chemical Protection

Protecting electronic devices from environmental pollutants by encapsulation constitutes a vital part of a robust packaging practice. Moisture, as well as other atmospheric pollutants, are key factors that seriously degrade electronic products' service life. Corrosive



(a) Plastic encapsulated (non-hermetic sealing) package

(b) Hermetically-sealed ceramic package

FIGURE 15.1 (a) Plastic encapsulated package; (b) hermetically-sealed package.

Samsung, EX1028, p. 247

ingredients, such as salts and other biological secretions, are among many harmful components present in the field. For example, mobile ions, such as sodium, can diffuse rapidly to the device junction, pick up an electron and ruin the device; chloride ions under bias will accelerate the corrosion of the aluminum metallization relatively easily. These contaminants must be prevented from diffusing to the IC.

Protection from Moisture

Moisture is a major contributor to packaging failures in electronic products, degrading their performance and reliability. Rapid water desorption from polymeric packaging during board level assembly is a major cause of delamination (popcorning effect in surface mount components) between device surface and the plastic molded cases. The vapor pressure build-up within the packages sometimes cracks the plastic cases. Delamination and cracking of the package may seriously reduce the device reliability after assembly on system boards. These concerns have forced the use of dry packing methods by IC manufacturers and moisture bake-out by system manufacturers just prior to board assembly. It was found that if the uncured encapsulants were aged in the presence of moisture at room temperature, and then cured, the resulting properties, such as glass transition temperature (T_g) and tensile modulus (E), were significantly lower than the ultimate properties which developed when the resins were stored in the absence of water and then cured. Hydrolysis of fully cured encapsulants under elevated temperatures and high humidity, that were previously stored in a dry state, have also been documented. These are clear indications that moisture attacks both the cured and uncured epoxy materials, degrading the intrinsic mechanical properties that are designed to protect the electronic packages. The adhesion strength between solder balls and these epoxy underfill materials was also found to deteriorate when aged under humid conditions.

The swelling of the encapsulants caused by moisture pickup, and the resulting hygrostress, is a major driving force of failures at the interconnection level. Water absorption by epoxy resins follows an initial linear relationship with the square root of exposure time. This linearity suggests that the absorption is diffusion controlled.

Fick's Law of Diffusion states that the amount of water pickup, W_t, at time, t, before reaching equilibrium, is proportional to the square root of time, and inversely proportional to the film thickness L:

$$\frac{W_t}{W_e} = \frac{4}{L}\sqrt{\left(\frac{Dt}{\pi}\right)}$$

The equilibrium water content, W_e, of an encapsulant is an exponential function of humidity:

$$W_e = KH^{\alpha}$$

Where K and α are temperature-dependent constants. The typical equilibrium water content for epoxies is 1.5–2% for the neat resin, and diffusion constant D for water in epoxy is 1.5×10^{-13} m^2/s. Moisture can be relatively easily absorbed into plastic molded components. However, when these components go through the solder reflow process, the absorbed moisture vaporizes, and it generates such an enormous pressure as to crack the component. This is called the "popcorn effect" in plastic surface mount assembly, as illustrated in Figure 15.2.

Samsung, EX1028, p. 248

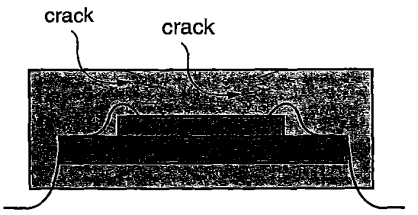


FIGURE 15.2 Popcorn effect in plastic packages.

Protection from Salts

Ionic contaminants, such as sodium, potassium and chloride, affect the reliability, and in some instances electrical performance, of encapsulated *integrated circuit* (IC) devices. Sources of salts can be biogenic (e.g., sweat) or abiogenic (e.g., from sea water or dicing operation). In the presence of these salts, corrosion of the IC metallization is accelerated. Microelectronic materials are frequently chosen for their electrical performance without regard for their corrosion inhibition properties. This, along with the applied operating voltages, is sufficient to cause electrolytic corrosion on most devices. Since most IC conductors are submicron in line widths and micrometer or less in pitch, very small amounts of localized corrosion are all that is necessary to produce an open circuit, or change the electrical characteristics of the device sufficiently, to cause noise or errors in the signal. The most damaging corrosion effect comes from submicron-sized ionic particulates such as airborne ammonium sulfate particles. These salt particulates can cause serious atmospheric corrosion of copper, if proper encapsulation is not rendered. The coating thickness of the encapsulant, the solution pH of which the device is exposed to, and the applied device voltage, are all critical factors in determining the accelerated corrosion.

Protection from Biological Organisms

Insects can be attracted by the electric field generated by a functioning electronic device. Fireants, for example, have been known to cause damage to some telephone switches in the field due to their corrosive secretion and droppings left within the switching boxes. Frequency dependence studies have revealed that insects are more responsive to electric fields generated at higher frequencies.

Protection from Atmospheric Contaminants

Other corrosive gases in the atmosphere that can be even more harmful to electronic devices include nitrogen oxides (NO_x) and sulfur dioxide (SO_2); both are present in air as by-products from fossil fuel burning. Nitrogen oxides and sulfur dioxide form corrosive acids when reacting with moisture in the air. Acid rain can reach a pH as low as 5 in the presence of these oxides, and will become extremely corrosive to all exposed devices, significantly reducing their effective service life.

15.2.2 Mechanical Protection

A second important protective function that an encapsulant can offer to an electronic package is mechanical protection. Both the wirebond and flip chip devices have very fine

Samsung, EX1028, p. 249

input-output (I/O) interconnects in the form of wire (for wirebond devices) or bumps (for flip chip devices). The wire can be as fine as 25 μm in diameter, and the bumps can be as small as 50 μm in diameter. Therefore, the structural integrity provided by these I/Os to the next level of packaging is very minimum.

Mechanical protection is achieved in two ways:

- Prevention of mechanical damage by the use of encapsulation over the IC
- Minimization of strain in the solder joined by underfill between IC and package substrate

The encapsulant, whether in the form of molding compound or liquid underfill, can serve to embed these fine structures into a more robust format.

15.3 FUNDAMENTALS OF ENCAPSULATION AND SEALING

15.3.1 Hermetic versus Non-Hermetic Sealing

A particular electronic product may be appropriate for many uses. For example, a mobile phone may see several different environments in its lifetime, such as its use in an air conditioned office or in a hot automobile on a summer day. When all factors related to the device are counted in, such as the cost of manufacture, production volume, expected use environments or markets and expected volume of product going to the various markets, then packaging issues and decisions become very complex. These complexities require the packaging engineer to make critical decisions on how to cost-effectively package the device, while obtaining acceptable reliability in the field. Hermetic versus non-hermetic packaging selections can be seen as a compromise between cost and performance, where performance includes reliability, as well as other design considerations.

Inorganics are hermetic; organics are not. A hermetic package is theoretically defined as one that prevents the diffusion of helium below a leak rate of 10^{-8} cm^3/s. The fundamental basis has to do with the atomic structure of solids, liquids and gas. Figure 15.3 illustrates the engineering property that describes this behavior as being the permeation rate as illustrated in this figure. Hermeticity starts with glasses and continues on to metals. Polymers and gases, as illustrated on the left part of the figure are, of course, non-hermetic.

In the early days of single transistors, hermetically-sealed metal cases were used to protect sensitive chips from environmental stresses. Moisture related problems prompted the military and aerospace industries to require hermetic packages in order to achieve their long-term reliability. Hermetic packages utilize a sealed environment that is impervious to gases and moisture to protect the devices. Final sealing is accomplished with caps or lids using glass or metal seals. Moisture content within the package is limited to a volumetric maximum of 5000 parts per million (ppm), which was chosen to eliminate any possibility of condensation occurring inside the package enclosure. For extremely sensitive devices, however, moisture limits are lowered even further, increasing both the time and cost to manufacture these hermetically-sealed packages.

Hermetic packaging proved to be such a robust method of achieving long field life that little has changed in hermetic packaging technology in recent years. After a slow development and initial reliability problems, hermetically-packaged devices gave way to lower-cost non-hermetic plastic-molded packaging, especially for electronics destined for

Samsung, EX1028, p. 250

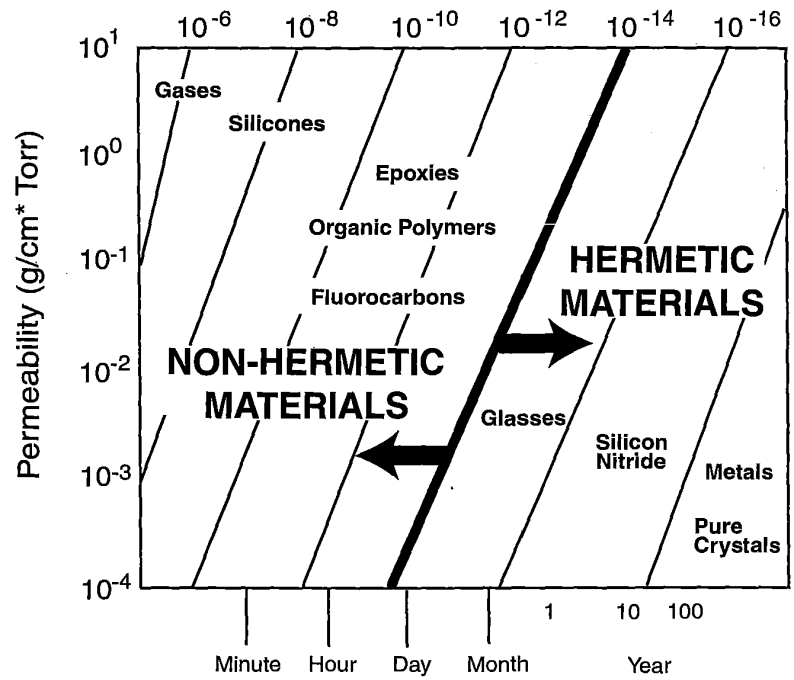


FIGURE 15.3 Permeability of water through organic (non-hermetic) and inorganic (hermetic) materials.

commercial and industrial markets. Today, however, non-hermetic does not imply non-reliability. Plastic packages are not hermetic, yet demonstrate acceptable reliability and now account for approximately 90–95% of all device packages.

15.3.2 Moisture Absorption of Encapsulants

Moisture Effects on Plastic Packages

Moisture can also have deleterious effects on the long-term adhesion between organic materials and the substrate being protected. Moisture acts as a debonding agent through a combination of the following mechanisms: (1) the moisture-reacted metal surface can form a weak, hydrated oxide surface, (2) moisture-assisted chemical bond breakdown, and (3) moisture-related degradation or depolymerization. The natural protection of metals and surface oxides can actually be detrimental to the durability of the polymer's adhesion in wet environments. Both metal and oxides are relatively polar. Water (H_2O) preferentially absorbs onto the oxide surface and creates a weak boundary layer at the metal–polymer interface that can lead to adhesion problems. In wet environments long-term corrosion prevention properties of a polymer can be adversely affected by this weakening interface region. Besides, the moisture diffusion rate depends on the material, as well as its thickness and the diffusion time, as illustrated previously in Figure 15.3.

Samsung, EX1028, p. 251

Many obstacles had to be overcome with respect to understanding materials and processes associated with plastic packaging before their use was widespread. Organic materials are not hermetic and allow moisture to penetrate and be absorbed. Different organic material families have greatly varying material properties relative to their moisture permeation rates and moisture-absorption characteristics. In plastic packaging's infancy, corrosion was found to be the primary cause of failure, due to poor adhesion and high levels of mobile ionic contaminants in the materials. The presence of moisture, along with mobile ions, were identified as the major contributors to corrosion, among other failure mechanisms. Improvements in plastic packaging materials and processes have all but removed corrosion as a source of failure. These improvements in materials and processes have lead to reliability that approaches that of hermetic packages in preventing the ingress and egress of moisture at the package perimeter during its operating life, thus achieving excellent long-term reliability. The word hermetic is defined as completely sealed by fusion, solder, and so on, so as to keep air, moisture or gas from getting in or out; in other words, airtight. In hermetic packaging practices, such seals are nonexistent.

Small gas molecules typically enter the package over time through diffusion and permeation. Eventually, these gases will reach equilibrium within the cavity of the package. In spite of this permeation, long life can still be obtained in the field because of the extremely slow nature of this activity. Accelerated tests and leak testing for screening and qualification are specified in the military standard MIL-STD-883 and represent the foundation and industry-accepted approach to testing for reliability in hermetic packages.

15.3.3 Organics Came a Long Way

The widespread availability of non-hermetic packaging followed that of hermetic packaging by many years due to problems associated with the initial high purity polymer materials' availability for encapsulation. Both the traditional plastic-molded package and polymer-encapsulated circuits fell into this category. The inability of these early polymers to sufficiently retard the deleterious effects of moisture, once it reached the delicate surfaces of the IC and its assembly, led to poor performance, both in accelerated testing and in the field. Inadequate adhesion, contaminants within the material itself, incompatible thermal expansion, and resultant stress-related problems and a relatively immature knowledge of filler technology—all combined prevented the plastic packaging's immediate acceptance. With significant efforts in the areas of resins, fillers, material formulations, and process development work, polymer packaging finally began to make its presence felt in the early 1970s. During this time, significant progress was also made in improving the quality of the glass passivation layer over the active areas of the device, as a first line of defense against moisture-related problems. The combination of these technological advances acted as the fundamental base that was needed for polymer packaging to be accepted, ultimately leading to its widespread use.

During plastic packaging's infancy period, many failure modes were identified and material concerns were addressed which resulted in different polymers and polymer formulations that more closely matched the application requirements. This has led to recent acceptance of plastic packages as approaching the reliability of hermetic packages in many applications and environments. Clearly, plastic-molded packages are the most dominant packaging method in use today for commercial- and industrial-grade electronics. It

Samsung, EX1028, p. 252

is estimated that over 90% of all integrated circuits are marketed in this form. Non-hermetic packaging encompasses not only the plastic-molded packages but also plastic and ceramic cavity packages that are sealed with a polymer rather than expensive inorganics. More recently, chip-on-board (Figure 15.3) and multichip module have also migrated to non-hermetic packaging. Most of the high-performance MCMs being produced, such as MCM-C (ceramic) and MCM-D (deposited), are mainly packaged in hermetic packages, but the lower-cost MCM-L (laminated), which is either wirebonded or TAB or flip chip bonded to a printed wiring board, tends to be encapsulated with organic polymers in an overcoat fashion. This is typically referred to as chip-on-board (COB).

Much remains to be understood in the use of polymers for protection of semiconductor devices. Generalities of the required material properties and how to avoid the failures associated with them are known, but specific degradation processes and reactions at the material–surface interface remain the object of current research. As mentioned earlier, corrosion was the primary cause of failure in early polymer packaging. High purity resins with low mobile ions such as Na^+, K^+, Cl^- have improved these materials' performance. New formulations with better filler technology resulted in materials that do not impart stress-related failures of ICs and their interconnect. In the newer technologies of COB and MCMs, corrosion is once again a primary cause of failure with the existing encapsulation materials, especially with wirebonded devices, as in most plastic-molded packages, where the thin aluminum bond pad on the device is exposed to the coating interface.

Adhesion Is Very Critical

Good interfacial adhesion between polymers and packages is important. Adhesion between metallic–organic interfaces is facilitated by a combination of mechanical interlocking and chemical and physical bonding. If the density and strength of the molecular bonds between polymer and substrate is not high enough to prevent an aqueous phase from forming, corrosion is likely to occur. Different families of materials have different ways of attaining corrosion protection. For example, although typical epoxies do not have a much higher strength of adhesion than silicones, they do have a much higher modulus, which can lead to higher stress and degradation of the interface. Corrosion protection and adhesion properties are closely linked, and long-term reliability requires long-term adhesion. This polymer matrix density perspective carries over into the bulk materials as well. The use of silane coupling agents has greatly improved the polymer adhesion properties. Although it is certainly desirable to have a polymer material that is completely void of ionic species to begin with, it is also highly desirable for the material to be able to limit the migratory capabilities of ionic species from the ambient environment.

Accelerated Testing Helps to Select the Right Material

Accelerated testing is usually the means by which non-hermetic packaging is assessed during screening and qualification during the manufacturing process. Temperature cycling is the most common thermomechanical environmental test. Temperature cycling does not test the corrosion-resistant properties of the package and the polymer system, but it tests the ability of the assembly to endure the stresses imparted by the various materials that make up the device, interconnect and polymer encapsulation. An accurate correlation of accelerated testing to field life is underway. As more and more of this correlation is understood in relation to the fundamental principles surrounding the use of polymers in

Samsung, EX1028, p. 253

packaging, its use is expected to increase, and more difficult markets, such as the military, will become important uses. In fact, the military has chosen to accept plastic components that meet the new MIL-STD-883, which consists of:

1. Thermal shock: test method: −65°C to 150°C; cycle time: 10 s; dwell time: 5 min at each extreme 1000 cycles
2. Salt-spraying time: 24, 48, 96, 240 hours; salt concentration: 0.5–3% (NaCl) pH = 6.0–7.5, 95°F; deposition rate: 10,000–50,000 mg/m^2 for 24 hours at 35°C
3. Autoclave: 121°C, 100% relative humidity (RH), 30 psi (2 atm), with or without bias (pressure pot).

High performance plastic packages that are based on careful material selections and processes are expected to pass all of the above new standards and achieve the reliability without hermeticity (RWOH).

15.4 ENCAPSULATION REQUIREMENTS

Encapsulants must have the required mechanical, thermal and chemical properties. Flow and adhesion are the two primary physical properties that any encapsulants should be optimized to perform. The *coefficient of thermal expansion* (CTE) in the desirable range, with glass transition temperature (T_g) outside of reliability testing window (−65°–150°C), are essential for robust encapsulated packages. It is desirable for the encapsulant to have a CTE close to that of the solder joints, a glass transition temperature that ensures dimensional stability during reliability testing, an elastic modulus that will not contribute to large stresses during thermal cycling, an elongation at break that is greater than one percent, and a low moisture absorption. However, these properties are not present in a single epoxy. Therefore, an epoxy formulation comprising a mixture of epoxies and filler provides a practical approach to fulfill most encapsulation needs. Table 15.1 lists some important flip chip underfill material requirements, which are also applicable to most other encapsulating materials.

15.4.1 Mechanical

Mechanical Properties

One of the most important mechanical properties of a good encapsulant should be its good stress–strain behavior. Figure 15.1 shows stress–stain curves of two encapsulant materials. An ideal encapsulant should exhibit >1% elongation at break, a tensile modulus of 5–8 GPa, and a minimum shift in properties at temperatures close to T_g. Material A in Figure 15.4 fails at low strain and has low toughness, and is not an acceptable material. Material B, on the other hand, is ideal.

Thermomechanical Considerations

Table 15.2 lists coefficients of thermal expansion (CTE) of some materials of interest. Most unfilled polymeric systems show 4× CTE change above their T_gs. The CTE changes after T_g in filled encapsulants, such as underfills and molding compounds, are in the range of 2–5×. Ideally the CTE of a molding compound should be as close to Si (2.6

Samsung, EX1028, p. 254

TABLE 15.1 Flip chip underfill material requirements.

Properties	Desirable Values	Comments
Flow	>0.5 mm/s	Fast flow with no air bubbles entrapment
Adhesion	>50 MPa shear force	Key to device protection
CTE	18–30 ppm/°C	Matches CTE of solder (26 ppm/°C)
Elongation	>1%	Resists CTE mismatch stress
Modulus	5–8 GPa	Provides mechanical coupling
T_g	>130°C	Maintains dimensional stability
Stress after cure	<10 MPa	Minimizes internal stress caused by shrinkage of polymer
Water pickup	<1%	Reduces moisture-induced failures
Ionic impurities (Na^+, K^+, Cl^-, Br^-)	<10 ppm	Prevents corrosion and metal electromigration
Thermal stability, 1% weight loss	>260°C	Prevents underfill decomposition during solder reflow
Curing time at 160°C	<0.5 hr	Maintains good product output
Volatility during cure	<1% weight loss	Maintains correct stoichiometry
Pot life at RT, 20% increase in viscosity	>8 hr	Provides long usable underfill life

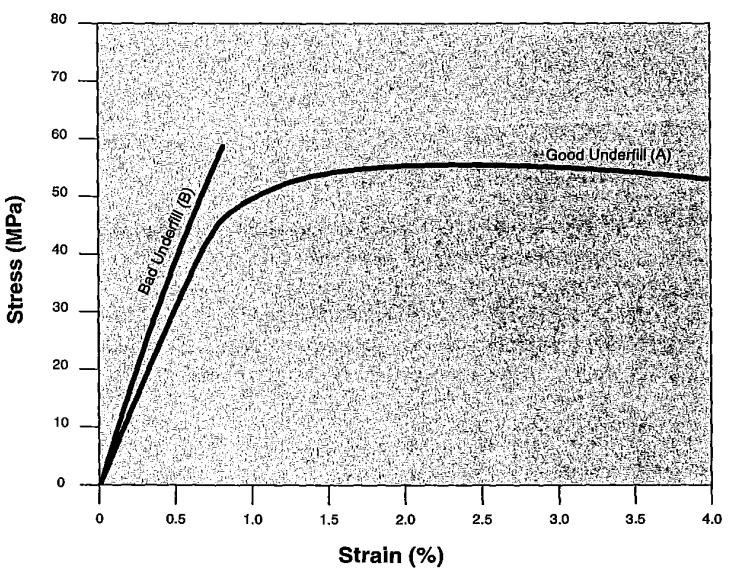


FIGURE 15.4 Stress–strain curves of good (material A) and bad (material B) encapsulation materials.

Samsung, EX1028, p. 255

TABLE 15.2 Physical properties of some materials of interest.

Material	CTE (ppm/°C)	Modulus (GPa)	Density (g/cc)
Silicon	2.6	107	2.33
Silicon dioxide	0.5	119	2.60
Alumina	6.6	345	3.90
Solder (63Sn/37Pb)	25	50	8.40
Aluminum	23	79	2.90
Molding compound	15	14.2	2.30
FR-4	16	20	1.85

ppm/°C) as possible, and the CTE of an underfill should be as close to the solder bump (25 ppm/°C) as possible. These properties ensure low stress between chip and underfill, as well as between solder and underfill.

Residual Stress

Most resin systems used in encapsulation develop various degrees of residual stress after cure. The two major sources of residual stress come from shrinkage of resin and thermomechanical loading due to mismatch of CTEs of constituent materials between cure temperature and storage (room) temperature. Volume shrinkage of polymers upon curing is very common; epoxy resins, for example, shrink between 3–6% by volume after cure, or 1–2% of the linear dimensions. This shrinkage alone produces a residual stress of ~20 MPa after cure. The degree of volume shrinkage is dependent on the type of curing reactions. Epoxy-amine systems tend to give more shrinkage than the epoxy-anhydride systems. Epoxy modifiers that expand upon cure can be added to the resin systems, to reduce or eliminate this shrinkage stress. The CTE-mismatch-induced stress is in the same order of magnitude as the shrinkage stress. The stress generated by the CTE mismatch can be estimated by the following equation:

$$\sigma = k \int_{25}^{T_g} E(\alpha_e - \alpha_s)\, dT$$

where σ is the epoxy film stress, E is the elastic modulus of the epoxy, k is a constant, $(\alpha_e - \alpha_s)$ is the difference in CTEs between the epoxy and the substrate, and dT is the change in temperature between T_g of the epoxy and room temperature.

15.4.2 Thermal Properties

Coefficient of Thermal Expansion (CTE)

The requirements for CTE vary significantly with the type of encapsulants in need. Most resin systems exhibit CTEs of 50–80 ppm/°C and therefore require addition of ceramic filler to lower the CTEs. The lower the CTE the better, but CTE is normally dictated by how much of the silica filler (CTE ~0.5 ppm/°C) can be formulated into the resin (CTE ~60 ppm/°C) without compromising flow and adhesion during molding. In reality, molding compounds with CTE of 10–15 ppm/°C and underfill with CTE of 20–35 ppm/°C are commonly accepted.

Samsung, EX1028, p. 256

Glass Transition Temperature (T_g)

Glass transition temperature is defined as the temperature at which the transition from solid to liquid takes place. Above this temperature, the modulus is low and nearly constant, and below which it is nearly three orders of magnitude as much as indicated in Figure 15.5. The T_g measures phase-transition of the polymers used in encapsulants with respect to the temperature change. Above T_g, most polymers exhibit tripling (or more) of their CTEs. In a confined space, such as within the gap of a flip chip assembly, such abrupt changes in CTE can cause catastrophic failures. For this particular reason, all robust underfill systems should have T_gs above the intended reliability testing temperatures (125° or 150°C). For the surface coating encapsulants, such as conformal coating materials, the T_g requirement may not apply. In fact, most conformal coatings have T_gs much lower than 125°C.

Flow During Encapsulation

The initial form of molding compound is in solid preform. The flow characteristics of the molten compound within the mold during the molding operation is very critical in a high yield manufacturing process. All other liquid form encapsulants, such as underfill, cavity fill, glob-top, and conformal coating, require good flow during operation. Good wetting to the surfaces to be encapsulated that produce void-free filling are considered good flow characteristics. The following Washburn equation depicts a simple model,

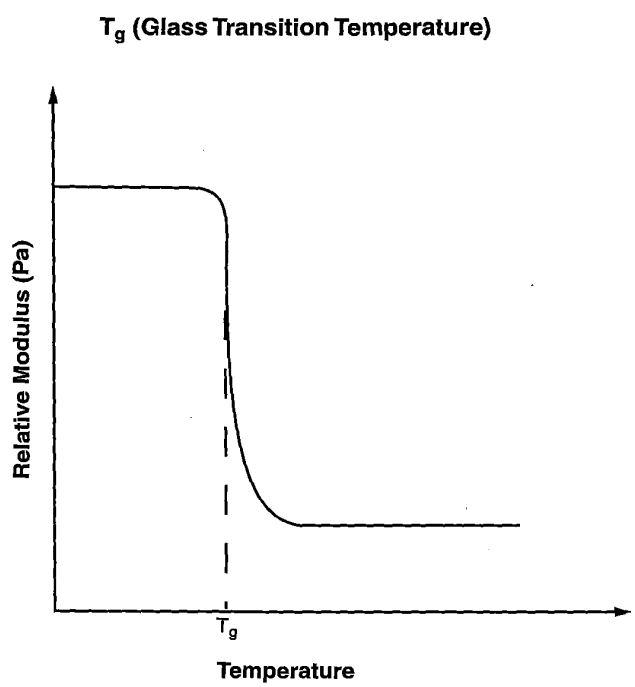


FIGURE 15.5 Glass transition temperature of polymers.

Samsung, EX1028, p. 257

derived from parallel plates flow, that describe the time required for an underfill to flow between the die–substrate gap:

$$\text{Underfill time} \cong \frac{3\eta L^2}{h\gamma \cos \theta}$$

where η = Viscosity
L = Distance of flow
h = Die–substrate gap
γ = Surface tension
θ = Wetting angle

It is clear that from an underfill flow time point of view, an underfill with the lowest viscosity, the highest surface tension and the smallest wetting angle, is the material of choice for a given die–substrate gap and die size.

15.4.3 Physical Properties

Adhesion

Adhesion is defined as the measure of the strength between two interfaces. A robust encapsulation system provides strong adhesion to the device encapsulant interfaces, such that the mechanical integrity of the package can be preserved under thermal stress. Both chemical (covalent) and mechanical (van der Waals) bonding can be incorporated into the system with desirable results. Examples of adhesion enhancement include the addition of adhesion promoters in the encapsulant formulations to improve chemical bonding and surface roughening of device–substrate interfaces by plasma etching. Adhesion failure in electronic packages proceeds with the following stage: damage initiation, microcrack formation, debonding growth, and finally, interfacial delamination. Any intervention prior to the arrival of delamination can potentially postpone or eliminate the structural damages caused by adhesion failure. Adhesion promoters provide good bonding between interfaces, such that damage initiated by thermomechanical load can be minimized. Additives such as tougheners incorporated in the encapsulant formulations also help to limit growth of microcracks and debonding progression.

Interfaces

Interface is defined as any physical or chemical layer, often in atomic scale between two materials. Robust interfacial adhesion between device hard passivation (silicon dioxide, silicon nitride, silicon oxynitride), or soft buffer coat (polyimide or bezocyclobutene on top of hard passivation), and the encapsulant, is the first line of defense against adhesion failure. Adhesion between encapsulant and bondwires (for wirebond device) or solder bumps (for flip chip device) is a secondary concern. Interfaces between solder mask, exposed conductor leads, and substrates (FR-4, BT, ceramics) and encapsulating materials all require good adhesion to prevent package failure.

15.5 ENCAPSULANT MATERIALS

The most common encapsulants fall into four generic categories: epoxy, cyanate ester, silicone, and urethane, as indicated in Table 15.3. Proper materials selection for encap-

Samsung, EX1028, p. 258

TABLE 15.3 Example of a flip-chip underfill composition.

Ingredient	Weight %	Functionality
Bisphenol A diepoxy	5.8	Resin
Cycloaliphatic epoxy ERL4221	12.5	Diluent, cross-linker
HMPA anhydride	13.8	Cross-linker, hardener
2-ethyl-4-methyl imidazole	0.3	Curing accelerator
Pigment (C-black)	0.1	Color coding
Spherical silica filler	67.5	CTE reducer

sulation applications can be realized by a prototype-modeling-reformulation sequence as shown in Figure 15.6. A prototype encapsulant can be prepared from readily available raw materials, its performance studied through empirical measurements, and any improvements can be further optimized by modeling.

All encapsulants involve some form of polymerization and cross-linking reactions that further enhance the mechanical properties of the packaging system. Addition reaction, which does not produce any volatile by-products, is the common path. The reaction (curing) usually occurs at elevated temperatures (80–180°C) for a period of 0.25–3 hours.

15.5.1 Epoxy and Related Materials

This is the most commonly used system in encapsulation technology. Epoxy polymerization is fast and clean, produces no volatiles and has been very well studied. Three common curing agents (hardeners) are used in the epoxy encapsulants: anhydrides, amines, and phenolics.

Anhydride-Epoxy

This system gives polyester linkage that can be susceptible to hydrolysis. The relatively high vapor pressure of the anhydrides at curing temperatures (~150–175°C) may pose

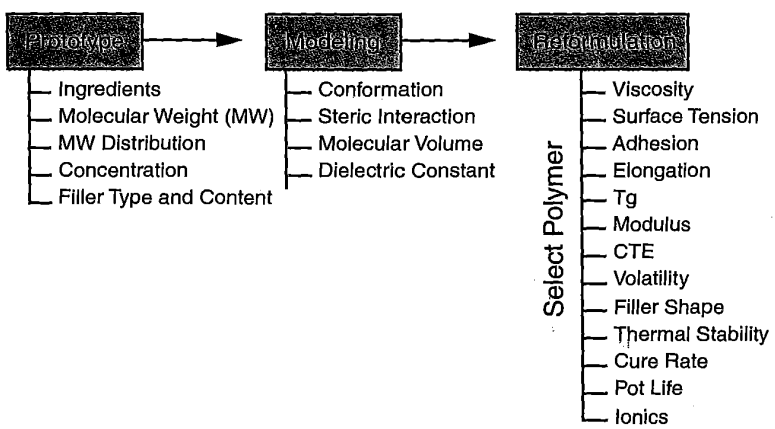


FIGURE 15.6 Encapsulation formulation optimization methodology.

Samsung, EX1028, p. 259

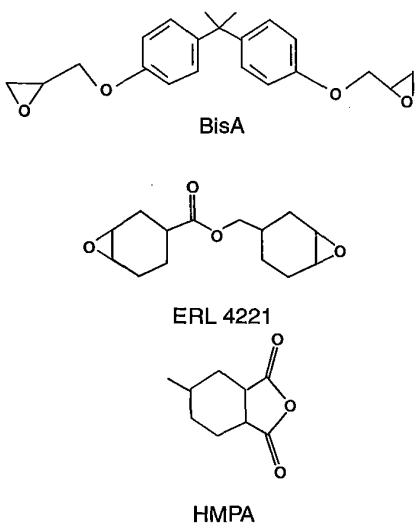


FIGURE 15.7 Structures of common ingredients used in underfill.

a stoicheometric problem due to the loss of anhydride by vaporization. To avoid such a problem, the anhydride containing encapsulants (underfill, cavity fill, and glob top) have to go to an intermediate gelling stage at a lower temperature (~125°C) before they can be cured at a higher final temperature. Regardless of the inconvenience, epoxy-anhydride systems are still the work horse for liquid encapsulation operation. Figure 15.7 shows a commonly used epoxy-anhydride encapsulant system that consists of a mixture of two epoxies and an anhydride.

Amine-Epoxy

This system is a less popular system due to the fact that most effective amines are in the solid or viscous liquid forms that often require a solvent to dissolve them in the formulation. The presence of solvents in the encapsulants can cause voids during cure, and may also plasticize epoxy structure, thereby weakening its mechanical strength. Nevertheless, teraryamide or amine-adducts are being used for some underfill formulations.

Phenolic-Epoxy

This system is used mostly in molding compounds, due to its solid nature. The phenolic curing reaction is the slowest among the three systems, often requiring prolonged post-mold curing for hours at 180°C. The epoxy polymers cured by phenolics also exhibit the most desirable mechanical properties, by lacking other disruptive structures such as esters and amines in epoxy backbone. These disruptive structures also cause higher moisture uptake of the polymer.

Commonly used aromatic epoxy resins include: bisphenol-A diepoxide, bisphenyl-F diepoxide, cresol novalac epoxide, dicyclopentadiene epoxide, tetramethylbiphenyl diepoxide, and 1,6-naphthalenediol diepoxide (see structures in Figure 15.8). Both tetramethylbiphenyl diepoxide, and 1,6-naphthalenediol diepoxide are solid at room

Samsung, EX1028, p. 260

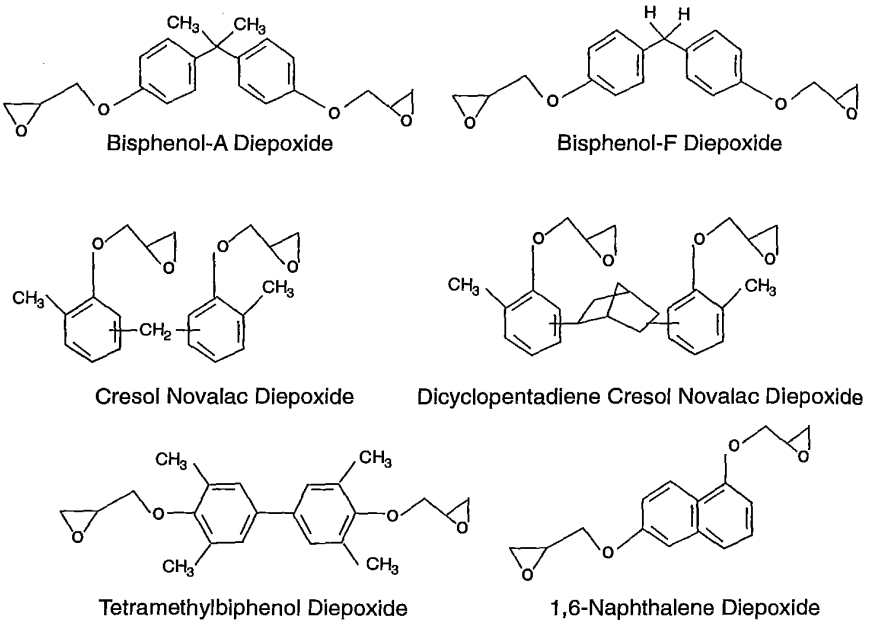


FIGURE 15.8 Structures of common aromatic epoxy resins.

temperature but melt to thin liquids at elevated temperatures, making them good starting materials for molding compound formulations. Other than aromatic epoxies, liquid aliphatic epoxies are also being used as reactive diluents in formulations requiring dissolution of solid curing agents or for viscosity adjustment. Butanediol diglycidyl ether and cyclic diepoxides such as 3,4-epoxycyclohexylmethyl 3,4-epoxycyclohexanecarboxylate (ERL4221 from Union Carbide) can be found in some underfill formulations. These aliphatics do not exhibit good mechanical properties after polymerization, and they tend to absorb more moisture than their aromatic counterparts. As compared to epoxies, the selection of anhydride curing agents is very limited. There are only three liquid anhydrides in common use: 4-methyl hexahydrophthalic anhydride (HMPA), tetrahydro-4-methylphthalic anhydride, and 5-methylnorbornene-2,3-dicarboxylic anhydride. 4-Methyl hexahydrophthalic anhydride is mostly used in underfill applications. Liquid imidazole curing agents such as *N*-cyanoethyl-2-ethyl-4-methyl imidazole, mixed with a liquid aromatic epoxy resin, is an example of the amine-epoxy system. Phenolics that are commonly used in the molding compound curing include cresol novolak and similar alkylated phenols.

15.5.2 Cyanate Ester

Cyanate esters are precursors of the high-performance triazine polymers. Chemically, these polymers exhibit higher T_gs (192–289°C) and lower water absorption than their epoxy counterparts. Due to higher raw material costs, cyanate ester resins are sometimes

Samsung, EX1028, p. 261

blended with epoxies to reduce their costs. Polymerization of cyanate ester to form triazine polymer is a unique process that requires participation of three cyanate groups to form a dendritic polymer network, and the reaction can never go to completion without end-groups "bite-back" to finish the cyclization reaction (see Figure 15.9). Another advantage of cyanate ester polymers is their self-dessicating property: the cyanate end groups of the polymer react with water to maintain a very low humidity in hermetically-sealed packages. Underfills formulated from cyanate esters have also proven to be more resistant to hydrolysis under steam bomb tests.

15.5.3 Urethanes

Urethane constitutes a major ingredient in most conformal coatings, which are used to protect circuit boards from thermal shock, engine fluid splashes, humidity, corrosive atmospheres and other adverse environments. Urethane coatings show good adhesion to the circuit boards and also provide stress relief to the electronic packages, and they sometimes offer equivalent protection to devices, as compared to the hermetically-sealed packages. These coatings are produced by the reaction of diisocyanates with hydroxy-containing acrylates. The hydroxy groups provide the acidic protons for the urethane backbone formation while the acrylate groups provide a photo–cross-linking capability, so that the coating can be cured using UV radiation (as shown in Figure 15.10). The urethane coatings can also be filled with some inorganic fillers to reduce the CTE for

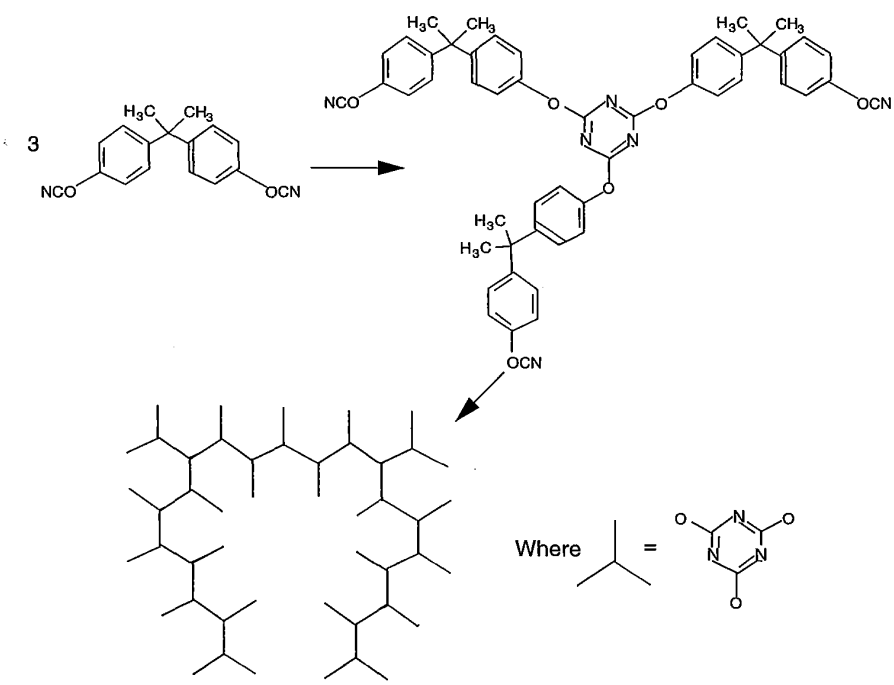


FIGURE 15.9 Structures of cyanate ester polymerization products.

Samsung, EX1028, p. 262

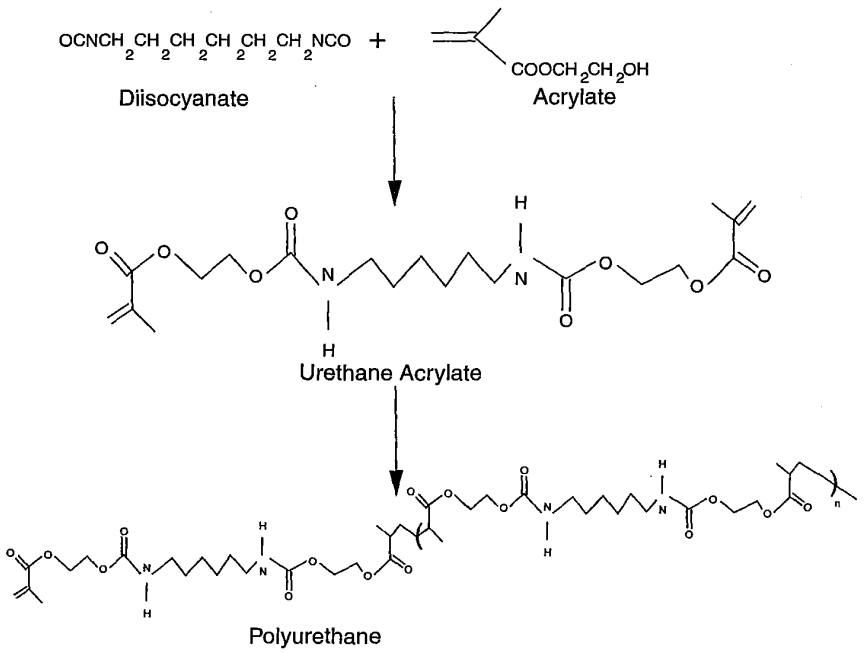


FIGURE 15.10 Structure of a urethane conformal-coating silicones.

better performance. As a matter of fact, polyurethane is one of the toughest tear-strength in most polymers.

15.5.4 Silicones

Silicone (organic polysiloxanes) was one of the earliest organic polymers used for electronics encapsulation purposes. Thermal stability and conformity are among the attractive properties that make silicone the material of choice. Its low T_g (~−125°C), extremely high CTE (300–800 ppm/°C), and low modulus, limit its applications in the conformal coating, sealing and cavity filling areas. The presence of even a trace of volatile low molecular weight cyclics/oligomers (i.e., D3, D4, D5, D6, etc.) in the raw silicone resin materials may present a surface contamination problem if bare conductor leads are exposed to those low molecular weight cyclics during encapsulation cure. Figure 15.11 depicts a cure mechanism of the heat-curable silicone, which can be tailored to prepare an elastomer or ultra-low stress gel. This class of silicones is used in potting ignition modules and voltage regulators in most automobile engine components, as well as in flip chip devices on silicon substrate and conformal coating of hybrid printed circuit boards. When silicone materials are properly chosen and components are properly cleaned and cured, they can provide the reliability without hermeticity encapsulation of most electronic systems. However, poor solvent resistance and low mechanical strength are major drawbacks of the material.

Samsung, EX1028, p. 263

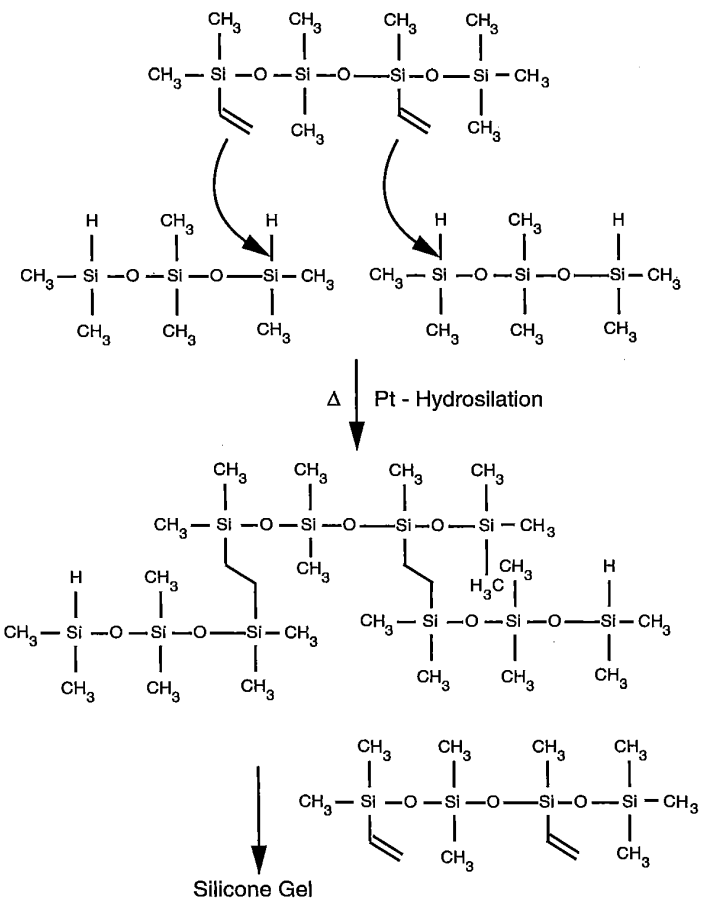


FIGURE 15.11 Silicone cross-linking reaction.

15.6 ENCAPSULATION PROCESSES

Encapsulation processes can be classified into two major categories: molding and liquid encapsulation.

15.6.1 Molding

The majority of encapsulating processes in the IC packaging market are transfer molding. By applying pressure, the heated molten molding compound is transferred from a pot (plunger), through the runners, and into the mold cavities. This is a simple mass production and low-cost method for most device encapsulation. However, the transfer molding method is hard to apply to some new applications such as flip chip and cavity fill type PGA (pin-grid array, Figure 15.12). Molding operation can be done individually on

Samsung, EX1028, p. 264

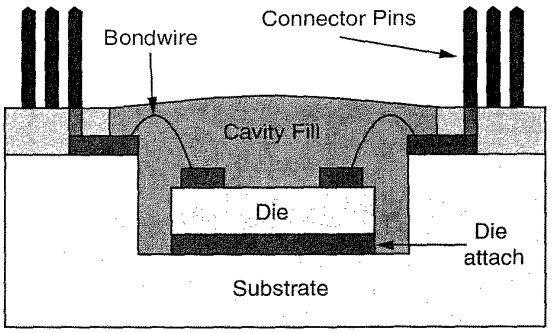


FIGURE 15.12 Cavity filling of PGA.

a single device or in *molded array packages* (MAP). The latter requires careful design of the array mold so that defects, such as wire sweep and void entrapment, are minimized to a single device molding level. Spherical silica fillers used in the molding compounds for the MAP operation also help to minimize defects.

Transfer Molding Process

A transfer mold contains a series of in-line cavities, one for each eventual semiconductor package. Each cavity has very small exhaust channels, called vents, which allow air to exit the cavity as the molding compound enters. A runner lays parallel to the strip cavity set, and it feeds the molding compound to each cavity through small entrance ports called gates. Usually, the runners are cut into only one half of the mold. Large molds also contain primary runners, which connect these strip runners to the main transfer pot. The transfer pot is a cylinder cut through one of the mold halves, and it receives the molding compound initially. It is often near the center of the mold. Knockout pins lay flush with the runner and cavity surfaces during the molding cycle and then extend into those regions to push the molded devices and runners out when the mold opens. The mold may contain heater elements within its body, or it may depend on heat transfer from the press structure.

The operational process is simple and usually automated: The mold opens and new strips of several in-line, pre-assembled devices, load into the in-line cavities. The actual active silicon devices are somewhere near the middle of the cavities. The mold closes and the molding compound, almost always a solid tablet, is placed into the transfer pot. This tablet may be pre-warmed, or it may rely on heat transfer from the mold. Then the transfer ram compresses the tablet into the runner system. By the time material starts to enter the cavities, it is fully molten and usually at its viscosity minimum. This transfer continues at rather near-atmospheric pressures until all cavities fill. The pressure then increases to about 6.4 MPa to "pack-out" the molding compound. This increase in pressure effectively reduces the diameters of any voids to a fourth of their original values, and it ensures that the sharp cavity corners fill out completely. This pressure holds until the molding compound cures, and becomes rigid enough to support the devices and the strips. The mold opens, the knockout pins automatically extend, and all molded packages and runners eject into collection systems. Brushes then clean away any residue or "flash" molding compound, and the cycle begins again.

Samsung, EX1028, p. 265

"Conventional" molds dominated in the 1970s through the 1980s. These were huge, complex molds with a single, large pot often holding up to a half-kilogram EMC tablet. There were several dozen to a thousand individual package cavities. The strip loading, tablet loading, parts removal and cleaning operations were usually manual because of the complexity of the molds. While these molds were extremely economical, they were also "unbalanced," meaning that some cavities were filled sooner than others. This design allowed variations in gold interconnect wire movement (called "wire sweep"), and variations in void levels and sizes. Cycle times ran from 90 to 240 seconds.

Newer molds now use several smaller pots and rams and the cavities are symmetrically located. These are balanced and all cavities experience the same flow profiles and pressure profiles during fill. These are called gang-pot molds because of the chain-gang alignment of the plungers and pots.

The presses, which hold and operate the mold, have three main functions: they clamp the molds closed, they move the transfer ram and they control the mold temperature. Automated presses, typical for gang-pot molds, manage all other functions like device loading, device/runner ejection and cleaning robotically with almost no operator input. Hydraulic clamping and transfer control is trending toward stepping motor control. Stepping motors do not vent airborne oil, generally require less maintenance and are readily controlled with digital electronic processing. Cycle times under a minute are typical and some operate as low as 15 seconds. A state-of-the-art auto, gang-pot press costs at least several hundred thousand dollars. Figure 15.13 illustrates the typical molding process.

15.6.2 Liquid Encapsulation

An effective way to ensure reliability of fine-pitch, low-gap devices, is to first dispense an encapsulant in a liquid form, and then cure to form a solid encapsulated package. By

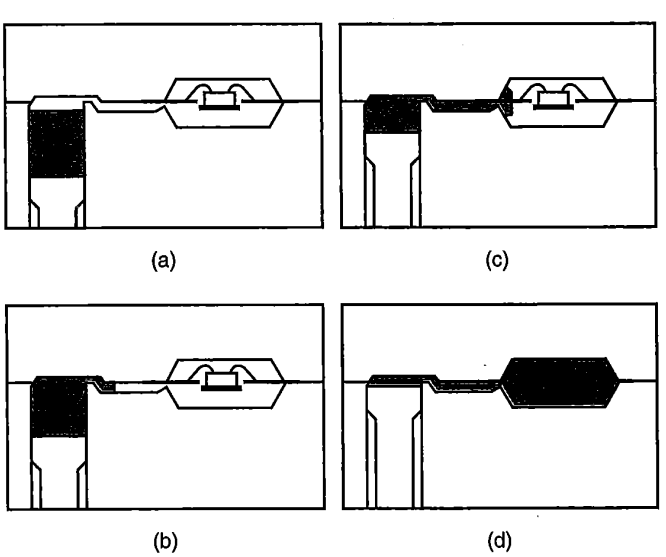


FIGURE 15.13 A typical transfer molding process.

Samsung, EX1028, p. 266

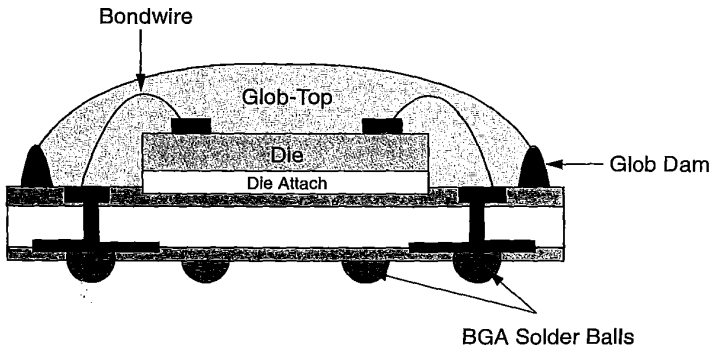


FIGURE 15.14 Glob-topping a ball grid array (BGA) device.

doing so, encapsulation materials utilization can be maximized, as compared to the rather inefficient use of the molding compounds in the transfer molding operation. Liquid encapsulants can be designed into different viscosity grades to fit into different flow requirements. The three most used liquid encapsulation processes are cavity filling, glob-topping, and underfilling.

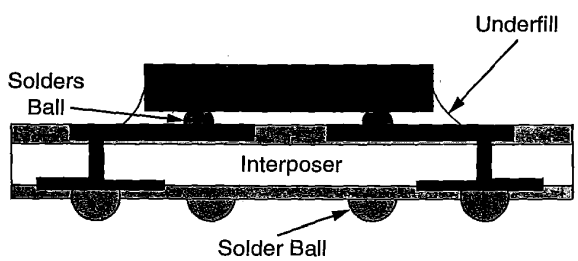
Cavity Filling

Cavity filling is mainly used in ceramic chip carriers, but can be adapted to many different substrates. A cavity is machined into the ceramic substrate that has interconnect leads and pads prefabricated to accommodate the die. After die attach and wirebonding, the cavity is flooded with a liquid encapsulant to protect the device from environmental pollutants, as well as mechanical protection. The process is very simple, but the liquid filling nature requires the device to be placed in a cavity of predefined size and shape, which may not be as flexible as the glob-topping approach. The package can be hermetically sealed with a metal lid if higher reliability is desirable.

Glob Topping

Similar to cavity filling, glob-topping, as illustrated in Figure 15.14, uses a damming material to form an enclosure so that a liquid encapsulant can be dispensed on top of a wirebonded device. The glob dam may be optional if the profile of the glob is not critical, or the device to be protected is so small that bondwire exposure is not likely. A standard

FIGURE 15.15 A typical flip chip ball grid array.



Samsung, EX1028, p. 267

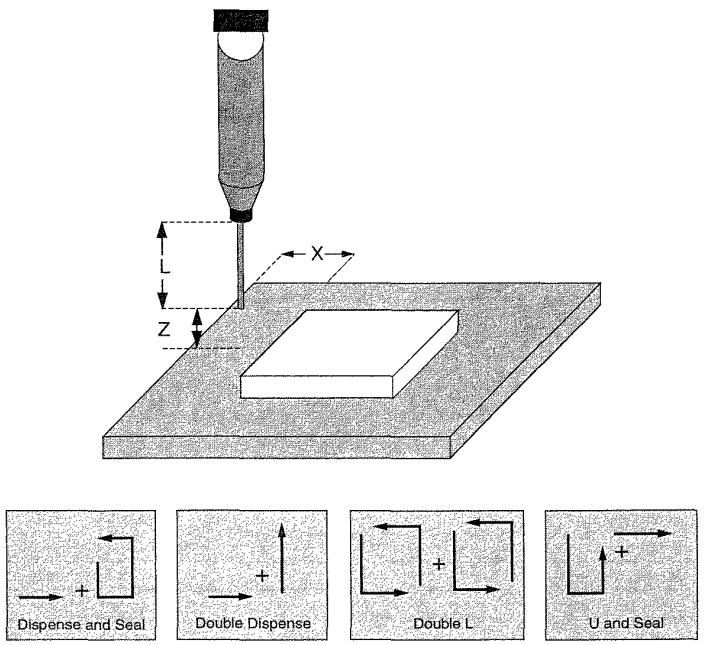


FIGURE 15.16 Underfill dispensing parameters and optimization.

glob-topping process consists of die attach, wirebonding, dam dispensing and cure, glob-top dispensing and cure. Glob-topping provides a convenient alternative for liquid encapsulants to be used in applications other than the cavity-filling process. There is no special requirement for the substrates.

Underfilling

Underfilling is the most critical operation in flip chip assembly. The flip chip bonded device is placed in a temperature-controlled stage with a temperature set at 70–100°C to aid underfill flow, and a syringe of underfill is dispensed from the barrel through a needle to the edge of the die. Both screw-driven or linear displacement pumps are used in underfill dispensing, with the latter producing better controlled volume. Different dispense patterns can be applied, e.g., single straight stroke, single L stroke, double L diagonal, and full-circle seal, etc., depending on required fillet sizes and profiles. Other underfill parameters, such as head speed, needle size and length (L), needle to substrate (Z) and to die edge (X) distances are all critical in specifying a successful underfill operation. Figure 15.15 and Figure 15.16 illustrate these underfilling process technologies.

15.7 HERMETIC SEALING

15.7.1 Sealing Processes

Sealing is a process that contains the electronic packages within an inert environment. This usually refers to hermetic packaging. The process consists of the following types:

Samsung, EX1028, p. 268

1. Fused Metal Sealing
2. Soldering
3. Brazing
4. Welding
5. Glass Sealing

These processes are discussed below.

Fused Metal Seals

Metal hermetic packages with enclosed volumes of a tenth of a milliliter or more are commonly welded, soldered, or brazed. Ceramic packages can also be sealed by glass sealing. To facilitate the sealing of the ceramic substrates by soldering or welding, a metal seal band should be provided on the substrate surface. In hard glass packages, a seal frame made of Kovar (17% Co–29% Ni–53% Fe) or ally 42 (42% Ni–58%Fe alloy) is first attached to the substrate by using a borosilicate glass. In ceramic packages, the seal band is formed by thick film, co-sintered molybdenum or tungsten metallurgy. The seal frame is then suitably plated, and a metal lid is attached by soldering or welding. The large throughputs, high yields, and reliability associated with the welding technique are spurring a change from glass sealing to welding for ceramic packages (see Figure 15.17a, b, c). The major considerations in selecting a sealing method are the availability of equipment and the cost of the hybrid circuit. Welding is more economical because of its high speed process yields, and reproducibility. Solder or braze sealing is commonly employed when it is required to allow for the ability to remove the cap and reseal the lid. Of these, the most popular method of hermetic sealing is welding.

Soldering

The solders used for hermetic sealing are selected, based on the required temperature hierarchy for the processes that precede and follow the sealing operation, the desired minimum seal strength, and cost. The lid seal has to remain intact, for example, during the soldering of chip carriers to printed circuit boards. In this case, the solder for the seal should have a higher melting temperature than the solder for direct mounting. Where cap rework is required for pin-grid array packages, the sealing solder should have a considerably lower melting point than the melting point of the solder or braze used to attach the pins to the substrate. Although straight tin-lead solders are widely used for hermetic sealing, alloying additions such as indium and silver, are sometimes added to improve the strength or fatigue resistance. Use of bismuth-tin alloys (lower melting point than eutectic tin-lead solder) for sealing has been suggested. It has been found that it is the property of these alloys to slightly expand on solidification, which helps to minimize the shrinkage voids in such seals.

Brazing

Brazing with a eutectic (80:20) Au-Sn alloy is used in place of soldering when the need exists for a stronger, more corrosion-resistant seal, and where the use of flux has to be avoided. The braze is usually used as a thin, narrow perform tack-welded to a gold-plated Kovar lid. The metal seal band on the substrate is also gold plated for good wettability and corrosion protection. In furnace sealing, the typical reflow time is 2–4 minutes above

Samsung, EX1028, p. 269

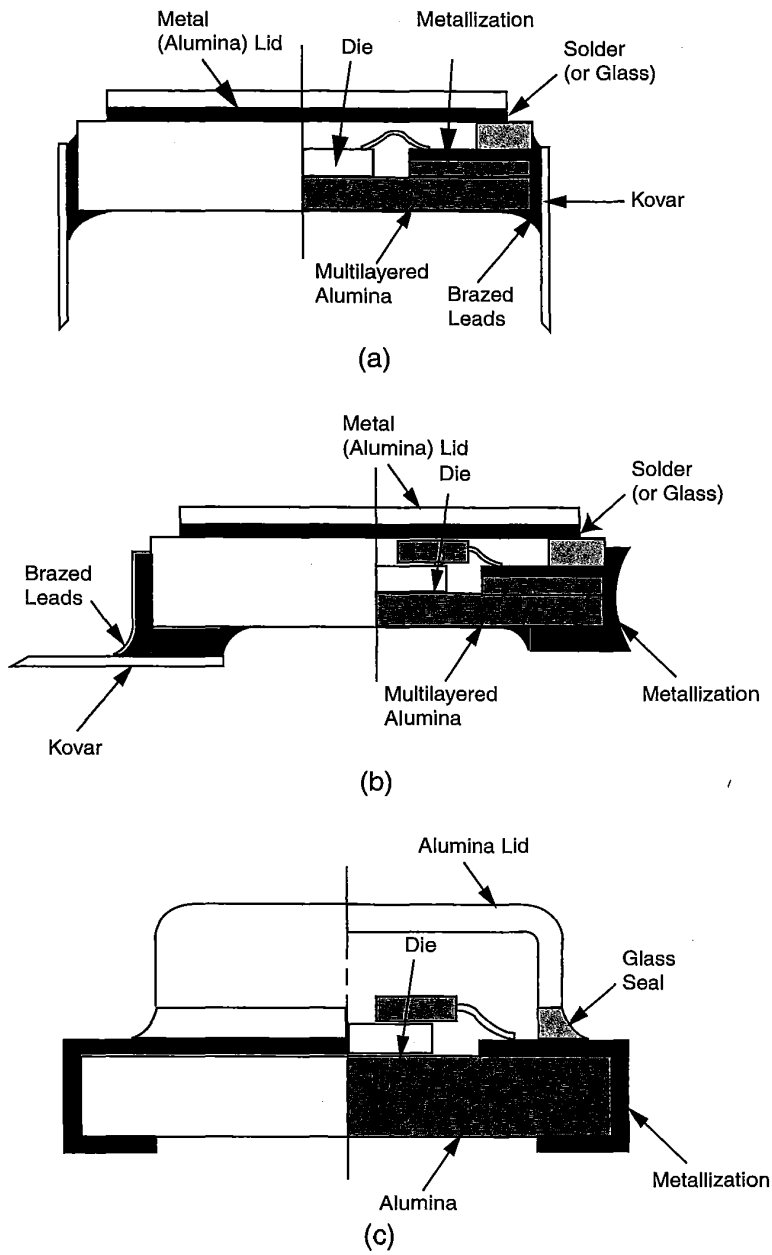


FIGURE 15.17 Sealing of various ceramic package types: (a) side braze ceramic packages; (b) chip carrier; (c) chip carrier (SLAM).

Samsung, EX1028, p. 270

the eutectic temperature of 280°C, with a peak temperature of about 350°C. Other methods of sealing can also be used for soldering with advanced braze sealing.

Welding

The most popular method for sealing high-reliability packages, such as those used in military applications, is welding. One survey indicated that about 80% of military packages are welded. Despite the higher initial cost for equipment, welding is popular because of the high yields and a good history of reliability. In welding, high-current pulses produce local heating between 1000°C and 1500°C, fusing the lid or the plating to the package. The local heating prevents damage to internal components.

In parallel seam welding (Figure 15.18), also called series welding, the package and the lid are passed under a pair of small tapered copper electrode wheels. The transformer produces a series of energy pulses, which are conducted from one electrode, across the package lid, to the other electrode.

In the opposed electrode welding (Figure 15.19), the package to be sealed is moved under a pair of small tapered copper electrode wheels. The power ply, usually a capacitor discharge, produces a series of welding pulses, which pass from the electrode wheel across the lid–package sidewall interface and return via the workbench, that generates heat at the electrode–lid interface and the lid–package sidewall interface. Welding can accommodate greater deviations from flatness for the package and the lid than soldering or brazing. The higher temperatures at the welding interface can volatilize most contaminants, so that cleanliness is not as critical a factor as it is in soldering and brazing.

Other less common methods for package welding are electron beam (e-beam) welding and laser welding. In general, they all provide reliability hermetic sealing of ceramic packages.

Glass Sealing

Glasses have been used in semiconductor package sealing since the beginning of the transistor age in the 1950s and 1960s. The first use of glasses was in device passivation,

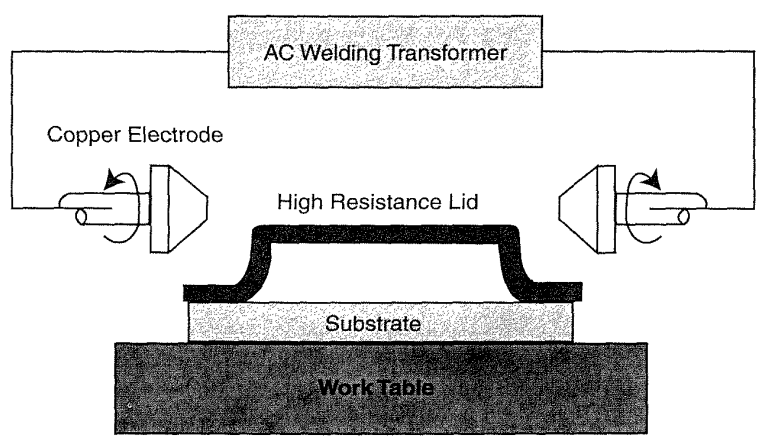


FIGURE 15.18 Parallel seam or welding.

Samsung, EX1028, p. 271

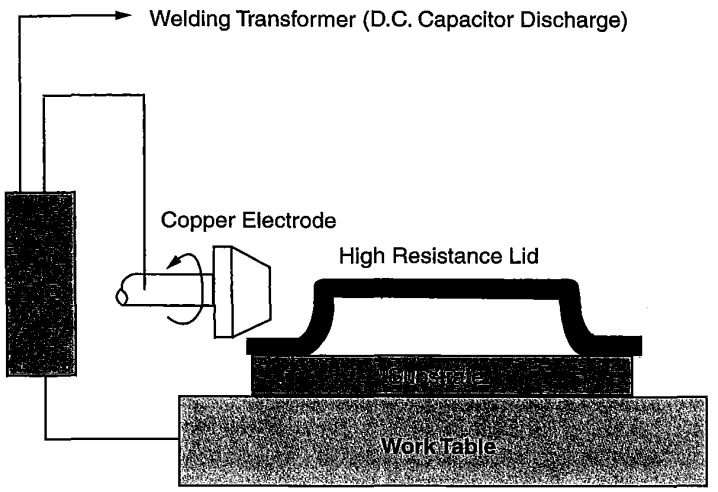


FIGURE 15.19 Opposed electrode welding.

and, to this day, glasses continue to be used for this important purpose, which serves as the last line of defense for a semiconductor device against external moisture and other contaminants.

Glass sealing is encountered in a wide range of package types, from the earliest T-O headers used to package the first transistors, to the latest ceramic packages. In the former, glass is used to form hermetic glass-to-metal seals for the package leads fed through holes in a metal plate or header. In the latter, the glass is used to form a sandwich seal between a ceramic cap or lid, and the ceramic substrate on which the devices are mounted. In Cerdip packages, the same glass seal performs both these functions. Low melting point lead-containing glasses are also used for low temperature sealing of some hermetic packages.

15.8 SUMMARY AND FUTURE TRENDS

The delayed availability of non-hermetic (plastic) packaging came after the hermetic packaging by many years, partly due to problems associated with the high impurity content of the initial polymeric materials used for encapsulation. Both the traditional plastic-molded package and polymer-encapsulated circuits fell into this category. The inability of these early polymers to block the deleterious effects of moisture once it reached the delicate surfaces of the circuitry, interconnect, and substrate, led to poor performance, both in testing and in the field application. Inadequate adhesion, contaminants within the raw material itself, incompatible thermal expansion, and stress-related problems, in addition to a relatively immature knowledge of filler technology to control the TCE mismatch, all combined to account for the delayed acceptance of non-hermetic plastic packages. With significant efforts in the areas of resins, fillers, material formulations, and process development work, polymer packaging finally began to make its presence felt in the early 1980s, although it had been used in some form as early as the

Samsung, EX1028, p. 272

late 1950s. During this time, significant progress was made in improving the quality of the glass passivation layer that is deposited on the active areas of the device, as a first line of defense against moisture-related problems. The combination of these technological advances provided the essential increase in reliability that was needed for polymeric packaging to begin in today's widespread use. We believe that the polymeric packages will continue to grow their shares in the device encapsulation area, and one day hermetic packages may become a custom-made item, only for those interested in applications in unusual environments such as in space and deep sea. However, collaborative efforts between materials scientists, chemists, physicists and chemical, mechanical and electrical engineers, are still needed to refine the next generation of high-performance, low-cost plastic packaging of electronics.

15.9 HOMEWORK PROBLEMS

1. A square die of 1.13 cm in length is flip chip bonded on a substrate with a gap of 50 μm. An underfill with a viscosity of 42.6 cps, a surface tension of 23.0 dynes/cm^2, and an average wetting angle between the die and substrate of 15° is used to underfill the package. What is the estimated underfill time if the underfill the follows the Washburn equation?
2. What is your prediction on underfill time of a device if you (a) raise the temperature, (b) increase the filler content, (c) shrink the die–substrate gap, (d) add a surfactant in the underfill?
3. A square die of 1.5 cm in length is populated with 2500 bumps. Assume the bumps are perfect spheres of 130 μm in diameter, and bonding does not change the volume of these bumps, and a gap of 60 μm is maintained between the die and substrate after bonding. How many grams of the underfill of density 1.82 g/cc is needed for packaging 5000 devices?
4. A cavity-fill package is encapsulated with an epoxy encapsulant having a diffusion constant of 2.5×10^{-11} m^2/s and the die of 625 μm in thickness is buried into a 3 mm cavity that is filled with the above material. Estimate the time for the die surface to reach equilibrium water concentration when exposed to a controlled humidity environment.
5. A good estimate for fatigue life enhancement by encapsulation is by 10× that of the unencapsulated devices. Assume the Coffin-Manson equation is correct, estimate the strain of the encapsulated bump that has an unencapsulated strain of 1.8% under identical thermal stress.
6. Explain why chemical bonding, not van der Waals bonding, is a preferred method for adhesion improvement on encapsulant–die interface.
7. Estimate the CTE-mismatch induced stress from room temperature to T_g on a die (CTE = 3 ppm/°C) that is encapsulated with an underfill having a T_g of 160°C, a CTE of 30 ppm/°C, and an elastic modulus of 11 GPa.
8. Explain why polymers of cyanate ester cannot assume a planar structure beyond high degree of polymerization (>>75%).
9. For most sealing applications, the UV-curable coatings are more desirable than their thermal curable counterparts. Explain why.
10. An epoxy encapsulant has an equilibrium water content of 1.8% at 85% relative humidity at 85°C. Assume the epoxy has an α = 1.2 at 85°C, then what would be the weight gain, in milligrams, of a glob-top package with a total dry weight of 2.5 g that contains 8% of the same epoxy at 30% relative humidity at 85°C in equilibrium?
11. What are the differences between hermetic (ceramic), versus non-hermetic (plastic), packages?

Samsung, EX1028, p. 273

12. What are the advantages and disadvantages of hermetic (ceramic), versus non-hermetic (plastic), packages?
13. What are the difficulties in sealing hermetic packages, in particular if the size of the hermetic package is larger then 6 × 6 square inches?
14. What are the three test methods for military qualified standard testing of electronic packages?

15.10 SUGGESTED READING

1. M. Bixenman and T. Fang. "Wafer Solder Bumping." *Advanced Packaging,* Vol. 48, 1999.
2. S. L. Buchwalter and L. L. Kosbar. "Cleavable Epoxy Resins: Design for Disassembly of a Thermoset." *Journal Polymer Science, Part A: Polymer Chemistry,* Vol. 34, 1996, 249.
3. L. Crane, A. Torres-Filho, C. K. Ober, S. Yang, J. S. Chen, and R. W. Johnson. "Development of Reworkable Underfills, Materials, Reliability and Processing." *IEEE Transactions on Components and Packaging Technology,* Vol. 22, 1999, 163.
4. T. Fang. *Environmentally Sound Assembly Processes in Wiley Encyclopedia of Electrical and Electronics Engineering.* New York: John Wiley & Sons, 1999, p. 126.
5. T. Fang and D. Shimp. "Polycyanate Esters: Science and Applications." *Progress in Polymer Science,* Vol. 20, 1995, 61.
6. P. N. Houston, M. Deladisma, and D. Baldwin. "Low Cost Flip Chip Processing and Reliability of Fast-Flow, Snap-Cure Underfills." *Proceedings of the 49th Electronic Components & Technology Conference,* 1999, 61.
7. J. H. Lau. *Chip on Board,* Chapter 12. New York: Van Nostrand Reinhold, 1994, p. 504.
8. J. Lau, C. P. Wong, and J. Prince, W. Nakayama. *Electronic Packaging Design, Materials, Process, and Reliability.* New York: McGraw-Hill, 1998.
9. S. Luo, M. Vidal, and C. P. Wong. "Study of Surface Tension and Adhesion in Electronic Packaging." *Proceedings of the 50th Electronic Components and Technology Conference,* 2000.
10. L. Nguyen, L. Hoang, P. Fine, et al. "High Performance Underfills Development-Materials, Processes, and Reliability." *Proceedings of the 1st IEEE International Symposium on Polymeric Electronic Packaging,* 1995, 300.
11. H. Quinones, A. Babiarz, and R. Ciardella. "Why Encapsulate Chip Scale Package-to-Printed Circuit Board Interconnects." *Proceedings of the Symposium on Advances in Packaging,* 1999, 83.
12. V. Sarihan and T. Fang. "Optimal Design Methodology for Filled Polymer System." *ASME,* Vol. 12, 1995.
13. D. Suryanarayana, J. A. Varcoe, and J. V. Ellerson. "Repairability of Underfill Encapsulated Flip-Chip Packages." *Proceedings of the 45th Electronic Components and Technology Conference,* 1995, 524.
14. R. Tummala, E. Rymaszewski, and A. Klopfenstein, eds. *Microelectronics Packaging Handbook.* New York: Chapman & Hall, 1997.
15. L. Wang and C. P. Wong. "Epoxy-Additive Interaction Studies of Thermally Reworkable Underfills for Flip-Chip Applications. " *Proceedings of the 49th Electronic Components & Technology Conference,* 1999, 34.
16. L. Wang and C. P. Wong. "Novel Thermally Reworkable Underfill Encapsulants for Flip-Chip Applications." *IEEE Transactions on Components, Packaging, and Manufacturing Technology,* Vol. 2, 1999, 46.
17. L. Wang and C. P. Wong. "Novel Thermally Reworkable Underfill Encapsulants for Flip-Chip Applications." *IEEE Transactions on Advanced Packaging,* Vol. 22, 1999, 46.

Samsung, EX1028, p. 274

18. L. Wang and C. P. Wong. "Syntheses and Characterizations of Thermally Reworkable Epoxy Resins, Part I." *Journal of Polymer Science, Polymer Chemistry,* Vol. 37, 1999, 2991.
19. C. P. Wong. *Polymers for Electronic and Photonic Applications*. California: Academic Press, 1993.
20. C. P. Wong, S. H. Shi, and G. Jefferson. "High Performance No-Flow Underfills for Low-Cost Flip-Chip Applications: Material Characterization." *IEEE Transactions on Components, Packaging, and Manufacturing Technology,* 1998, Vol. 21, 450.
21. C. P. Wong, L. Wang, and S. Shi. "Novel High Performance No Flow and Reworkable Underfills for Flip-Chip Applications." *Materials Research Innovations,* Vol. 2, 1999, 232.
22. S. Yang, J. Chen, H. Korner, T. Breiner, C. K. Ober, and M. Poliks. "Reworkable Epoxies: Thermosets with Thermally Cleavable Groups for Controlled Network Breakdown." *Chemical Materials,* Vol. 10, 1998, 1475.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contribution of Professor Ray Pearson of Lehigh University for critically reviewing this chapter.

Samsung, EX1028, p. 275

CHAPTER 16

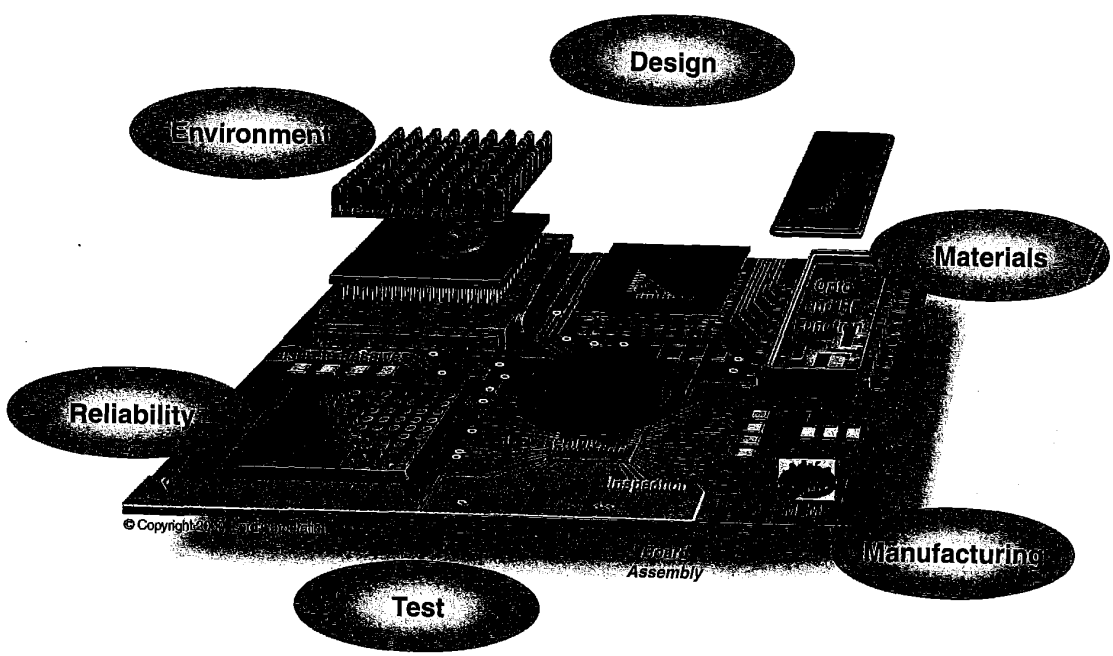
FUNDAMENTALS OF SYSTEM-LEVEL PWB TECHNOLOGIES

Mahesh Varadarajan
Indian Institute of Science, India

Prof. Zsolt Illyefalvi-Vitéz
UTE, Budapest

Dr. Joachim Zimmermann
Bosch, Germany

Prof. Rao R. Tummala
Georgia Institute of Technology



16.1 What Is a System-Level Printed Wiring Board?
16.2 Types of Printed Wiring Boards
16.3 Anatomy of a Printed Wiring Board
16.4 Fundamentals of Printed Wiring Boards
16.5 CAD Tools for Printed Wiring Board Design
16.6 Printed Wiring Board Materials
16.7 Standard Printed Wiring Board Fabrication
16.8 Limitations in Standard Printed Wiring Board Process
16.9 Microvia Boards
16.10 Printed Wiring Board Market
16.11 Summary and Future Trends
16.12 Homework Problems
16.13 Suggested Reading

CHAPTER OBJECTIVES

- Define printed wiring board (PWB) and describe its functions.
- Given the anatomy of a PWB, introduce and define all the critical elements of a PWB.
- Introduce and describe the CAD tools for PWB design.
- Describe standard materials and processes typically used in fabricating PWBs.
- Describe the limitations of PWBs and describe how microvia solves the problem.
- Describe various microvia technologies recently developed.

CHAPTER INTRODUCTION

Printed wiring boards allow both active and passive components to be interconnected, powered, cooled and protected to form electronic systems. This chapter describes the materials, processes, and designs, both for conventional laminated printed wiring boards and for advanced microvia-based boards.

Samsung, EX1028, p. 277

16.1 WHAT IS A SYSTEM-LEVEL PRINTED WIRING BOARD?

A *printed wiring board* (PWB), or *printed circuit board* (PCB), is a composite of organic and inorganic materials with external and internal wiring, allowing electronic components to be electrically interconnected and mechanically supported. In addition, a PWB must provide power to the components and conduct away heat when necessary. The boards are also called motherboards or system-level boards, because they carry all of the components required for that system or subsystem. They are also called back planes when some of the boards are interconnected to a larger motherboard.

Every electronic system contains at least one PWB, some of which are electrically interconnected components on both sides. The terms "printed circuit board" and "printed wiring board" are synonymous but the latter is becoming widely used in the industry. This chapter and the book, for the most part, use the terms interchangeably, with a preference for the newer printed wiring board terminology. The earliest "printed circuits" were made by printing a pattern of rubber-like resist on a copper plane, followed by etching. A laminate material made of phenolic and paper supported the copper. Holes punched in the laminate held the component leads that were soldered to the copper-printed pattern. Thus, the printed circuits provided mechanical support and electrical interconnection for the components.

The earliest developments of this nature were credited to Paul Eisler in the early 1940s when he was able to image copper circuitry using gelatin and ferric chloride etchant. Figure 16.1 illustrates a cell phone, and the printed wiring board used in that product.

The purpose of the printed wiring board therefore is:

1. To electrically interconnect all the components
2. To support and affix mechanically the electronic and electromechanical components
3. To provide power and dissipate heat generated by components

In accordance with everyday usage, a printed wiring board (PWB) is a substrate, which is made up of an insulating structure with copper foil tracks on its outer surfaces. Sometimes, some of the tracks are only on one of the outer surfaces, but more often in the

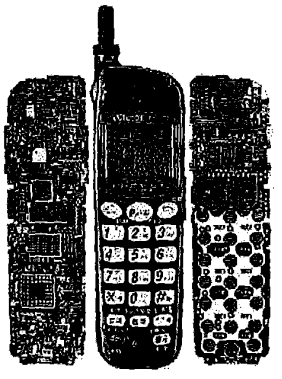


FIGURE 16.1 An example of the use of PWB for cell phone.

Samsung, EX1028, p. 278

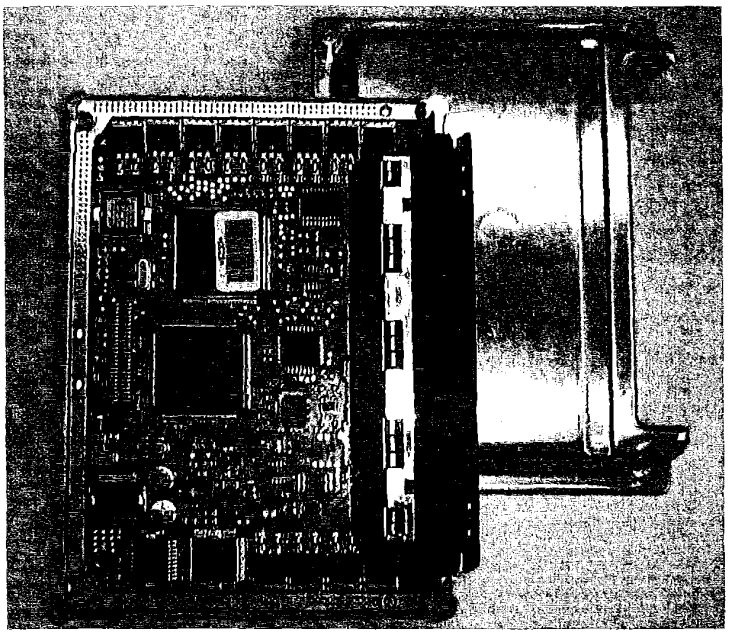


FIGURE 16.2 A finished PWB.

form of inner layers as well. The components are affixed to the PWB and connected to the conductive copper tracks by soldering. Figure 16.2 is a photograph of a finished PWB, ready for component assembly.

16.2 TYPES OF PRINTED WIRING BOARDS

16.2.1 Printed Circuit Boards

Printed circuit boards (PCBs) are largely made of the epoxy-glass-based laminates, the older ones of which were made of phenolic and paper laminates. They are typically less than 0.1 mm to several millimeters in thickness. The thinner laminates are used for pagers and calculators, whereas those in the mid-range of 0.5 mm, are found in notebook computers, camcorders, and radios. Thicker PCBs are used in printers and TV monitors, and are prevalent in personal computers and workstations. The thickest multilayers are found in office computers, mainframes, and telecommunications switches. Their thickness typically relates to the number of laminated layers, which is determined by both the number of components and their I/Os to be interconnected. For example, a mainframe of past era used 47 layers of PWB, because it had to interconnect over 300 VLSI packages, each with more than 500 I/Os.

16.2.2 Chip Carriers

Chip carriers are single-chip packages that include wirebond, *tape automated bonding* (TAB), and flip-chip bonded IC packages. They are of both polymer-based and ceramic-

Samsung, EX1028, p. 279

based materials. The polymer-based ones are largely epoxy-glass based because of their low costs.

16.2.3 Flexible-Circuit Carriers

Flexible-circuit carriers are printed circuits deposited on a flexible dry film. These films are predominantly polyimide and polyester-based materials. Their adaptability to roll handling manufacturing, and high-density conductor capability by photolithography, makes them popular for 3-D consumer applications such as cameras, disk drives and older telephones.

16.2.4 Metal Core Boards

Metal core boards are the boards with metal as the base material, on which glass or polymer is deposited as the insulator and conductive paste, or thin-film metal is deposited to serve as the conductor. The metal boards are used in special applications where either high temperature, as in automotives, or high heat dissipation, as in power supplies, is required.

16.3 ANATOMY OF A PRINTED WIRING BOARD

Figure 16.3 illustrates the anatomy of a printed wiring board. It includes one or more layers of conductors, interconnected by either through, blind or buried vias, separated by an insulatory epoxy-glass, or some other material.

The main considerations for selecting and categorizing PWBs are as follows:

1. Rigidity of the insulating board: rigid, flexible or combined (rigid-flex)
2. Number of conductive layers: single-sided, double-sided or multilayer

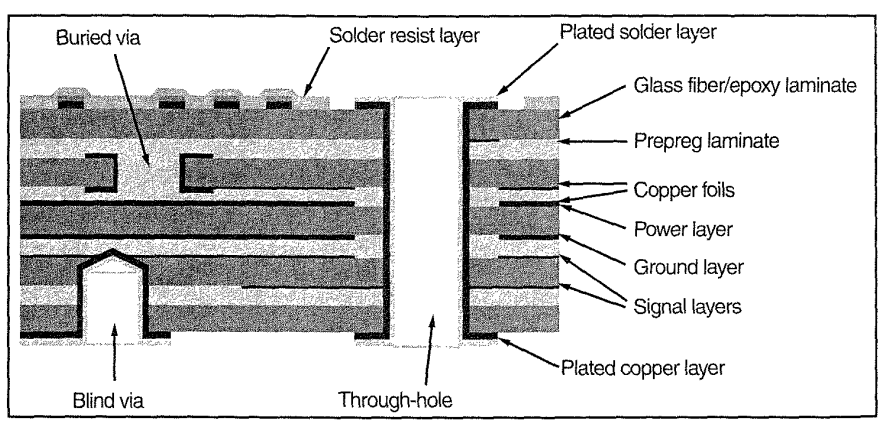


FIGURE 16.3 A typical printed wiring board configuration.

Samsung, EX1028, p. 280

3. Role and type of holes: non-metallized holes, *plated through-holes* (PTHs) or vias (including through-board), blind and buried vias
4. Quality of the conductive pattern: coarse, fine or very fine
5. Type of connector: edge connector, other standards

For most electronic systems, rigid printed wiring boards are used. Flexible printed wiring is very popular for 3-dimensional interconnection such as in cameras, while rigid-flex combinations are used in folded applications. Figure 16.4 shows rigid and flex PWB configurations.

16.3.1 Single- and Double-sided Printed Wiring Boards, Surface Mount Technology, and Multilayer PWBs

1. Single-sided boards have only one patterned conductive layer, and are used for very simple applications.
2. Double-sided PWBs are the most widespread, where there are two conductive layers in the laminate on the two outer sides of the board. The tracks formed from the two conductive layers are interconnected where necessary by plated through-holes. The components are usually assembled onto the board by the insertion of their leads through the same holes. This process is called through-hole assembly.
3. *Surface mount technology* (SMT) boards are used where components are assembled onto the surface of the board without the insertion of the leads into the holes; the holes only serve as interconnections, and they are usually produced with smaller diameters and are called vias.
4. Multilayer printed wiring boards contain more than two, typically four to eight, conductive layers for high-density applications. For prototype and special applications like military and aerospace, the layer count could be as high as 24 to 32.

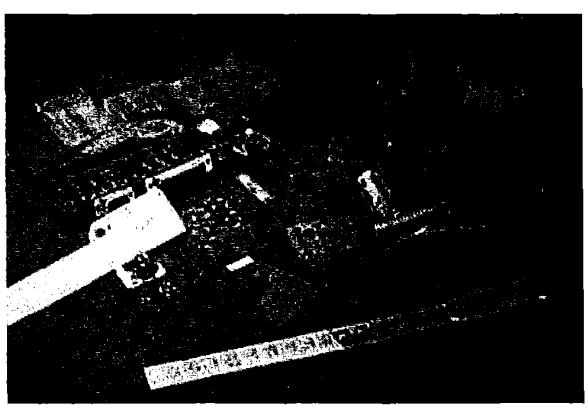


FIGURE 16.4 Rigid and flexible PWBs.

Samsung, EX1028, p. 281

16.4 FUNDAMENTALS OF PRINTED WIRING BOARD

A printed wiring board is typically a system-level board providing wiring to interconnect all the components of that system. This wiring acts as the power distribution and signal distribution. The following fundamental parameters are important in understanding this technology.

Panel is the name given to the square or rectangular board containing one or more conductive circuit patterns. This conductive pattern serves as the electrical circuit through which current flows. The pattern includes conductors, lands, and through connections.

Conductive pattern is the configuration or design of the conductive material on the base laminate through which electrical energy passes. The pattern includes conductors, lands, and through connections.

Land is the portion of a conductive pattern usually, but not exclusively, used for the connection or attachment of components. Land may also be referred to as a pad. A pad can contain a hole or a *via.*

Vias are not intended to insert component leads into them; they are only interconnections electrically or thermally joining different conductive layers of a multilayer structure. Figure 16.5 shows a multilayer structure with different vias.

The functions of different types of vias are as follows:

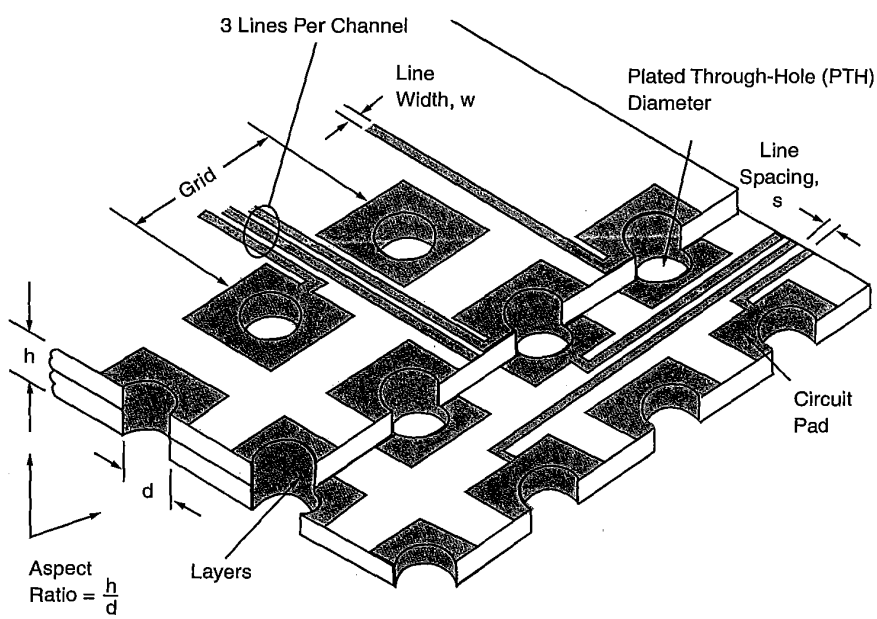


FIGURE 16.5 Fundamentals of PWB.

Samsung, EX1028, p. 282

1. A *through-board via* connects both outer layers and inner layers.
2. A *blind via* connects an outer layer and at least one inner layer, but does not connect both outer layers.
3. A *buried via* makes interconnection between inner layers of the board only.
4. The aim of a *thermal via* is to conduct heat from a component heat source to a heat sink layer of the structure.
5. *Microvias* are also used to make connection between two adjacent layers, but their diameter is typically less than 150 μm. Laser ablation, plasma etching, or photoprocessing are typically used to form microvias. (These are explained in later sections of this chapter.)

Pitch is the nominal distance between the grid lines, or the centers of adjacent features, on any layer of a printed board. Pitch is also known as center-to-center spacing.

Non-metallized holes are holes that are round, rectangular or any other shape, and are used for affixing components onto the board, either by inserting the component leads through them or by using other assembling methods.

Plated through-hole (PTH) is a hole in a circuit board that has been plated with metal, usually copper, on its sides to provide electrical connections between conductive pattern layers of a PCB. These metallized holes are used not only for producing electrical interconnections, but also for through-hole assembling of the components.

Aspect ratio is the ratio of height to diameter. This term is used for determining the size of the hole that can be drilled in a given thickness. It is also used in determining the lithographic via dimensions that can be formed. This is shown in Figure 16.5.

Wiring density is defined as the total wiring length contained within a unit square. It is measured in inches per square inch or centimeters per square centimeter. Greater wiring density means increasing levels of integration on the board. The conductor width, its pitch, and the number of layers, determine wiring density on the board.

Wireability is a measure of package capability to provide interconnections between components. It is measured in linear centimeters of wiring per square centimeter of package. The primary factors affecting the wireability of a package are the geometry and spacing of signal lines, the size and spacing of via holes, and the number of single layers in the package. To appreciate the effect line geometry has on wireability, as well as to gauge the progress in PWB technology, one may consider Figure 16.6. The total wireability of a package can be calculated as:

$$\text{Wireability} = (\text{Number of lines per channel}) \times (\text{Number of channels per cm}) \times (\text{Number of signal planes})$$

The units of wireability, as stated above, are cm/cm^2.

The wireability required for an application of N chips with an average distance between chips of P is determined as:

$$L = \{0.75 \times [1 + 0.1 \times \ln(N)] \times N\} \times P$$

where L is the average wire length, this can be approximated as:

Samsung, EX1028, p. 283

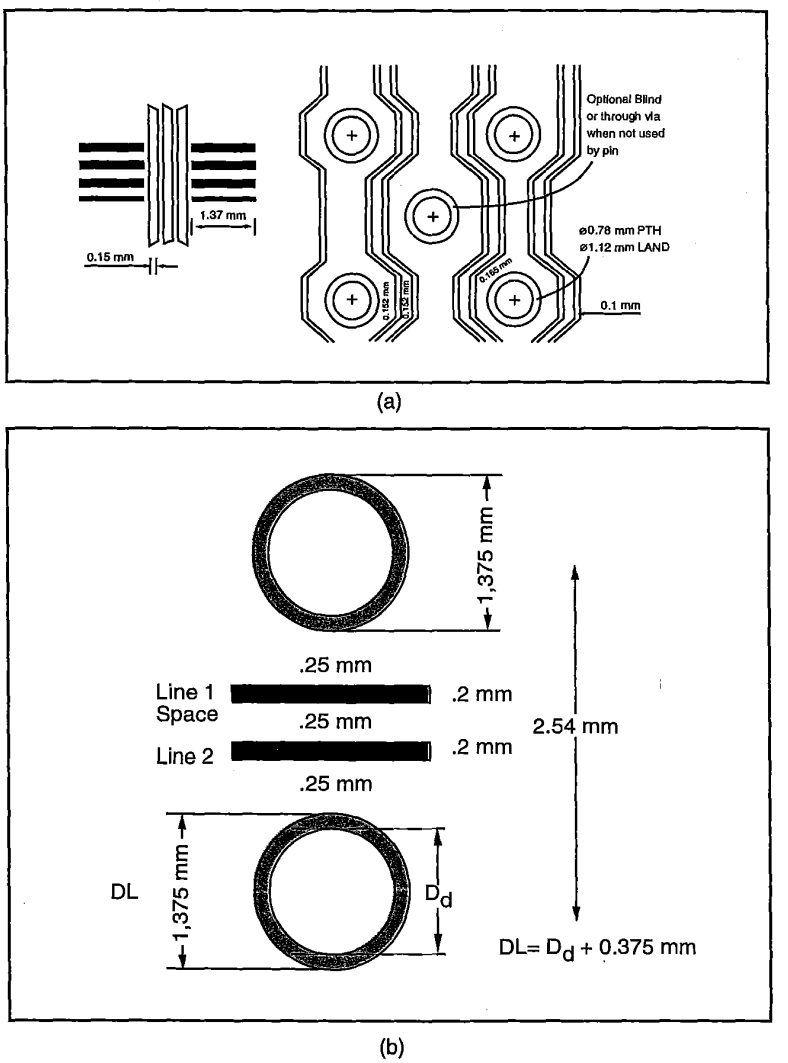


FIGURE 16.6 Definitions of (a) wiring density and (b) wireability lines per channel.

$$L = \sim 0.77 \times N^{0.245} \times P$$

Given that the application is wiring n pins (forming $n - 1$ nets), the total signal line length used for wiring is given by:

$$\text{Wiring length} = (\text{Number of I/Os per chip}) \times N \times (n - 1/n) \times 0.77 \times N^{0.245} \times P$$

The area available for wiring is:

Samsung, EX1028, p. 284

$$A = N \times P^2$$

Empirically, it has been found that $n = 3$ for a mid-range computer system and that, for most applications, only about 40% of the available wireability are used for I/O interconnections. Therefore, the area required to interconnect N chips is given by:

$$\text{Area} = 1.65 \times N^{1.49} \times \frac{(\text{Number of signal I/Os per chip})^2}{\text{Wireability}}$$

The wireability needed to provide the required level of interconnect is approximated by:

$$\text{Wireability} \sim 1.25 \times N^{0.25} \times \frac{\text{Average number I/Os per chip}}{\text{Pitch between chips, } P}$$

Lines per channel: In a conventional configuration, the distance between the drilled holes, also called the pitch, is the usual 2.54 mm (or 2.54 mm = 1/10 in)—see Figure 16.6b. A simple DIP package is a good example of this pitch. For coarse patterns, the typical value of the line width and spacing is 0.5 mm. It is also the same for the width of the ring pad around the hole. So, if the diameter of the hole is 1.1 mm, then there is no space for a track line between the holes. For fine patterns, the typical line width is 0.3 mm and the hole diameter is 0.8 mm, so a line can be fabricated between two holes. For very fine patterns, the resolution is less than 0.2 mm, and the hole diameter can be 0.3 mm or less, if necessary, so at least two conductive lines can be placed between the holes. It should be obvious that the number of lines that can fit into a channel of a given size is determined by the size of a given line, and the space between lines. Changing both of these parameters involves trade-off with other aspects of the package. In general, increasing the number of lines per channel exposes the package to cost, electrical, and reliability issues. These are primarily the result of defects due to processing problems. To increase the number of lines per channel, the size of the individual lines must be decreased. This has the effect of magnifying any small discontinuities in the line, resulting in an increase in defects known as opens and near opens.

FR-4 board is a composite material, constructed on multiple plies of epoxy-resin impregnated onto a woven glass cloth. It is the most widely used material in the PCB industry because its electrical, chemical and mechanical properties satisfy most system-level board applications. In addition, it can be processed in very large sizes up to, and including, 750 mm in linear size, thus making this technology one of the lowest cost technologies. Other resins widely used are based on polyester, phenolic, polyimide and Teflon.

Subtractive, additive and semi-additive process: There are two main types of fabrication processes of PWBs: subtractive and additive processes. The subtractive process uses copper clad laminates, and subtracts the unnecessary pattern of copper layer from the surface by chemical etching, leaving copper traces. In the course of the additive process, the wiring pattern is deposited physically, chemically or mechanically onto the original insulating surface of the board. The most popular board fabrication sequences are based on the subtractive process; however, they contain additive steps as well. The main advantage of the subtractive process is the excellent adhesion of the copper foil to the insulating board. The additive process is typically used in thick-film formation of ceramic substrates. The so-called semi-additive process combines the two process types.

Samsung, EX1028, p. 285

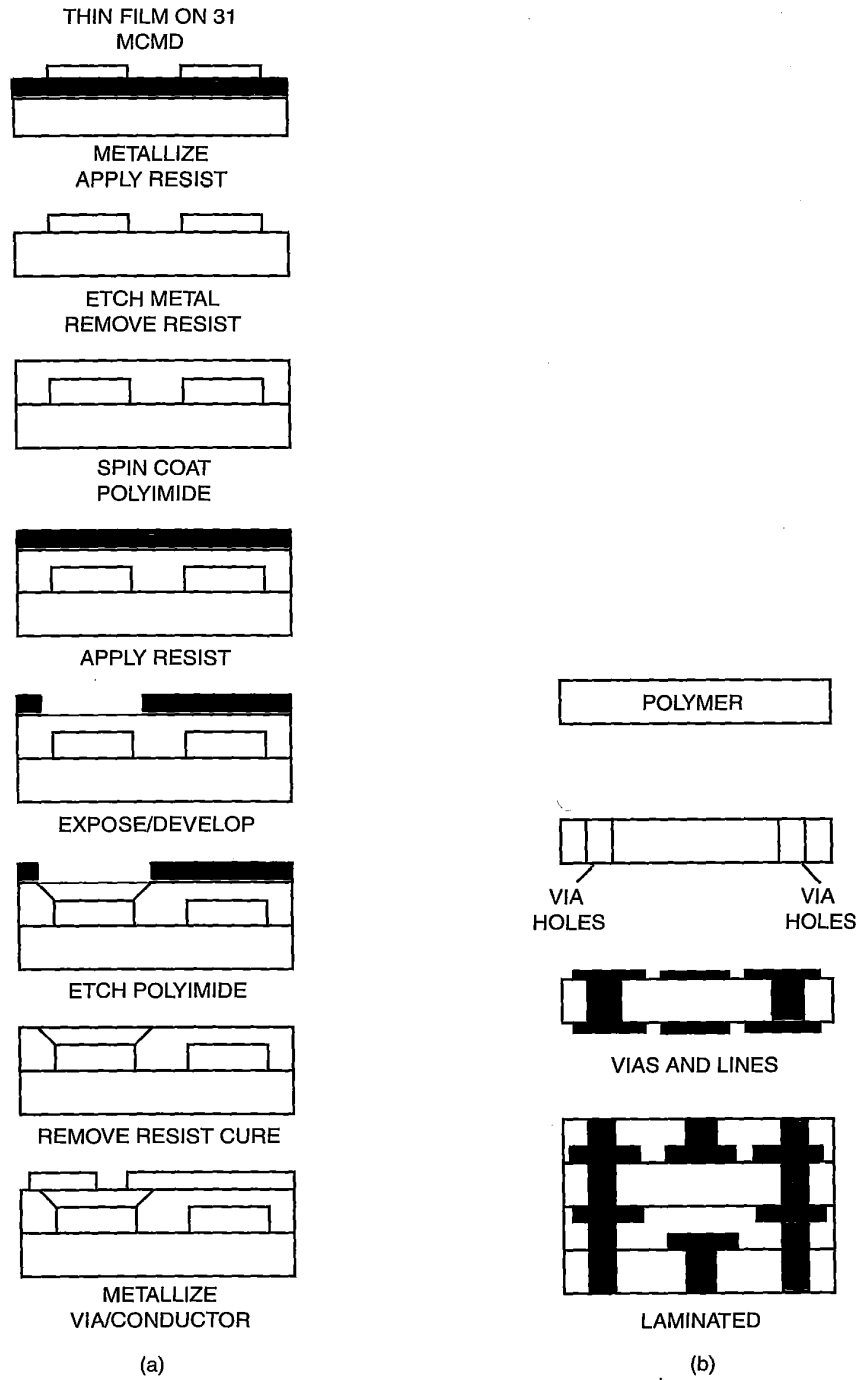


FIGURE 16.7 Sequential and parallel build-up processes for PWBs: (a) Sequential process; (b) parallel process

Samsung, EX1028, p. 286

It starts with laminates, covered by thin copper foils, then uses additive process steps and differential etching, resulting in very fine pattern.

Solder mask: With the requirement of finer and finer board assembly interconnection geometries, a solder resist mask, or solder mask, is required to be deposited over all parts of the board, except where solder joints are to be made. A solder resist is a heat-resisting coating material applied to selected areas of a PWB to prevent the deposition of solder upon those areas during subsequent soldering, and in particular, to prevent solder bridging between conductive tracks. The functions of the solder resist mask also include the control of outer layer impedance, minimizing handling damage during assembly, and increasing corrosion and flammability resistance. It is also applied to improve overall appearance of the board.

Build-up: There are two ways to the build-up process. One is sequential build-up and the other is parallel build-up. This is best described in Figure 16.7.

1. The sequential build-up process is a process of depositing dielectrics and insulators sequentially to build higher density PWB substrates, in contrast to the lamination and drilling process, which is typically applied for conventional PWB. This technology is very similar to the so-called MCM-D (D for deposited) technology, except for the fact that the thin-film is deposited on a very large PWB, often on top of the drilled PWB.
2. The parallel build-up process involves achieving high-density thin-film wiring, as with the sequential build-up process described above, by laminating a number of individually formed thin-film layers. The individual layers are typically copper-patterned polyimide or Kapton films similar to TAB tapes. The parallel nature of the process allows individual layers to be fabricated and tested before they are joined or laminated.

16.5 CAD TOOLS FOR PRINTED WIRING BOARD DESIGN

16.5.1 Design for Manufacture

Design for manufacture (DFM) means that all the necessary rules for the manufacture of printed wiring boards and the electronic devices have already been taken into consideration in the design. This permits the functionality and subsequent quality of the PWB and electronic device to be verified at an early juncture without the need for time-consuming, cost-intensive revision processes, avoiding unsuitability of the design for the necessary production processes. To simplify, the term DFM will be used here.

Figure 16.8 illustrates the major factors exerting an influence on the design as a result of PWB and electronic device production. A considerable interactive effect naturally exists between the different production engineering-related outline conditions. The type of components to be processed, for example, can determine the production class—minimum conductor structure, leading to the highest output, coupled with the lowest number of rejects in the defined production process—and the via technology. This mutual correlation is evident when considering the example of the use of *ball grid arrays* (BGAs). The type of connection technology used (soldering, adhesive or bonding) also exercises an impact on the type of wiring board surface which is finally used.

Samsung, EX1028, p. 287

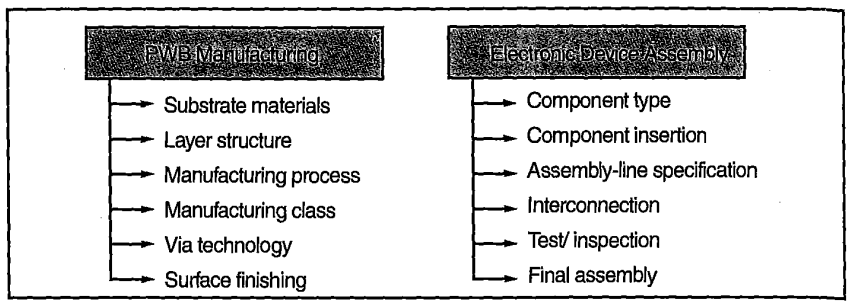


FIGURE 16.8 Influencing variables in DFM.

16.5.2 CAD Tools

Why CAD at All?

Complex PWBs, with high component and interconnection density, can no longer be designed without use of *computer aided design* (CAD) systems. CAD systems which are used in an electronic design environment are called e-CAD systems. They are the direct link between designer and computer.

CAD Process Flow Chart

According to Figure 16.9, the design process may be split into four major design steps: circuit design, physical design (or layout), design verification and manufacturing preparation. An e-CAD Program consists of several routines to support the design process, so-called e-CAD tools. They have to be of extraordinarily high performance to allow automated component layout at a high degree within an acceptable time frame.

Circuit design consists of the following elements:

- Schematic entry and simulation of the circuit
- Physical design grouped in placement of components
- Routing process of the traces

The physical design is followed by design verification and manufacturing preparation.

16.5.3 CAD Process

Schematic Entry

A PWB layout is possible by means of a wiring diagram or electrical schematic diagram, which contains all of the electrical functions and connections of the circuit. During schematic entry, which is illustrated in Figure 16.10, symbols are selected from a library, placed on the circuit and connected by traces, which represent the electrical connection between pins. Additional information for circuit analysis and layout, such as test points, current and voltage-related requirements, shielding of traces, noise suppression on the circuit, restrictions of heat transmission or grouping of components, must be defined and previously entered.

Samsung, EX1028, p. 288

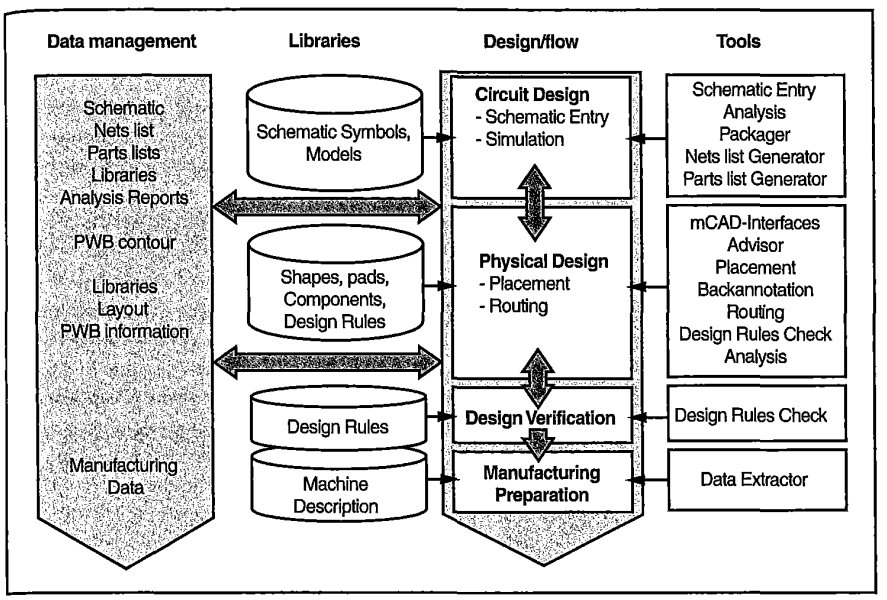


FIGURE 16.9 Main segments of a CAD system applied for PWBs.

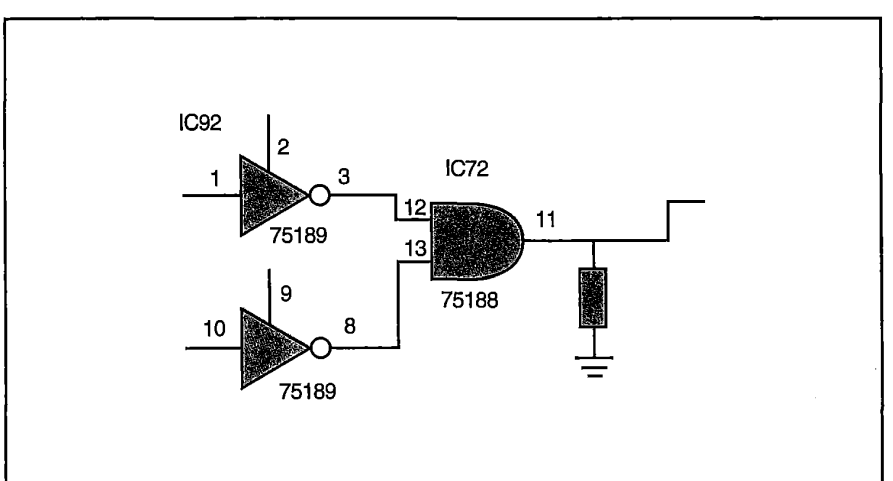


FIGURE 16.10 Typical schematic diagram of a simple circuit.

Samsung, EX1028, p. 289

Analysis Tools: Electrical, Thermal and EMI

A variety of analysis tools support the engineer in reducing development time. Based on the model descriptions mentioned above, signal integrity, including wave propagation delay, crosstalk, noise or waveform distortion, can be checked even in mixed-signal designs. Other tools are able to calculate accurate resistance, inductance, capacitance, and dielectric conductance parasitic for every trace and via, thus enabling the analysis of an impedance-controlled PWB, for example.

Component Placement

The objective of automatic placement systems is to place all components, an example of which is shown in Figure 16.11, to minimize the connection length and to consider the mentioned constraints. The calculation of the connection length has the highest priority. The step from through-hole to surface mount technology has increased the degree of difficulty for automatic placement dramatically, because all parameters of two-sided component placement have to be considered simultaneously. In general, modern placement tools apply different algorithms and procedures simultaneously. Automatic placement is normally done in more than one phase, to approach an optimal result interactively. These phases might be stopped to perform interactive changes for a better placement solution. The layout tool is, for example, allowed to move, unplace, fix/unfix, rotate, align, swap or change the placement side of components or groups of components.

Routing

Routing might be done interactively, or with assistance of automatic programs, which differ from the number and complexity of layout rules, which they are able to consider.

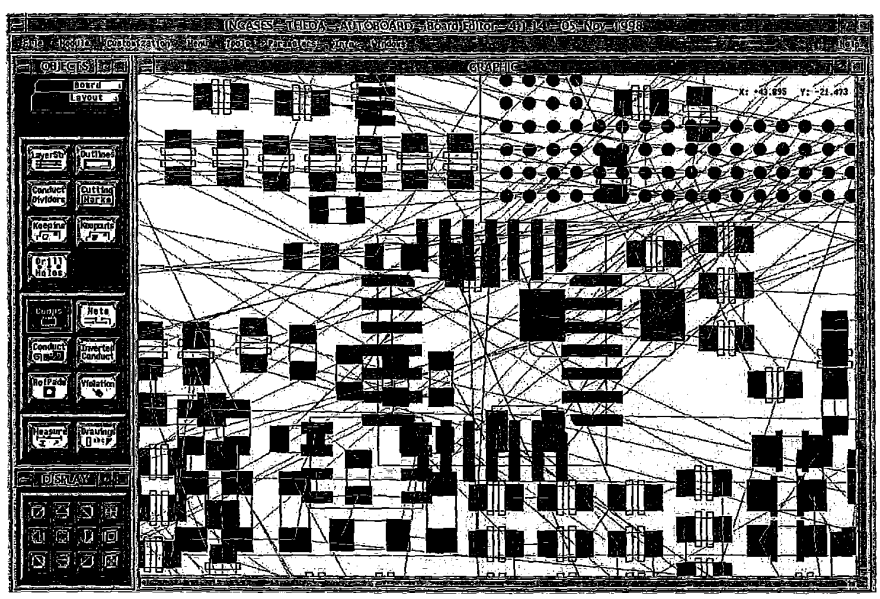


FIGURE 16.11 Component placement. *(Courtesy of Robert Bosch GmbH)*

Samsung, EX1028, p. 290

The automatic routing of PWBs and ICs has been under analysis for many years. While automatic programs have been able to route only digital boards with a multitude of integrated circuits sufficiently in the past, the efficiency of modern routing tools has increased in the last years. Autorouters might even be applied to analog or power circuits with a majority of discrete components.

Before starting the automatic procedure, several routing parameters must be defined. These might be fixed or pre-routed traces or areas, routing grid, cost values for vias, maximum connection lengths, shielding of a signal or definition of the preferred routing layer of a signal. Automatic routing is like placement, normally done in more than one phase to approach an optimal result interactively. The routing process might as well be stopped to perform interactive changes such as moving components, editing or routing of traces interactively or cleaning-up of the routing. An example of a routed design is shown in Figure 16.12.

Automatic routing programs also possess functions for subsequent optimization of conductor track routing. Examples are reducing the number of vias, beveling off traces under a 45° angle, reducing trace lengths, widening traces and mediating the position of bundles of traces between pins, as well as automatic or interactive generation of tear drops and test points.

Design Verification

Design verification is a process which can be performed during the layout process (on-line), and also following layout completion (batch mode). This verification process is known as the *Design Rules Check* (DRC), and in either case is based on a check of the

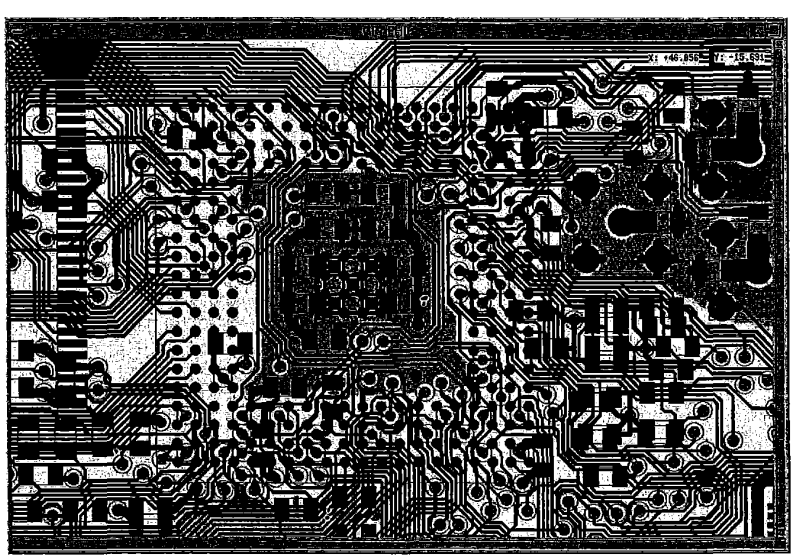


FIGURE 16.12 Layout of the conductive layers. (Only two of the six layers of the layout are represented here.) (*Robert Bosch GmbH*)

Samsung, EX1028, p. 291

TABLE 16.1 Documents and control programs for fabrication.

PWB Manufacturing	Assembly and Test
Plotting film masks for all layers	Film masks for solder paste print
Direct imaging on substrates	Lists of materials
Drilling	Pick and place
Milling contours	Network lists
Bare Board Test	In-Circuit Test
Manufacturing drawings	Assembly drawings

layout, to ensure adherence to the design rules necessary for manufacture of the PWB and of the electronic device.

Design Files for Manufacturing, Assembly and Test

Table 16.1 lists typical technology files that need to be generated from the CAD data for manufacturing the board. Each of them is generated in standard universal formats.

16.6 PRINTED WIRING BOARD MATERIALS

16.6.1 Key Components of a PWB Material

The three key components of a standard PWB are:

1. Organic resin
2. Inorganic filler
3. Copper conductor

The laminate material, which is used in the printed wiring board manufacturing, is a sandwich structure of conductors and dielectrics. The wiring pattern is produced in the conducting plane by etching-in high electrical conductivity tracks, whereas the dielectric supporting the tracks should be a good insulator.

16.6.2 Requirements of a Conductor

The conductor should be fabricated of materials of extremely high electrical conductivity and should be solderable, etchable and manufacturable in small lines and spaces. The most common material is copper.

16.6.3 Requirements of an Insulator

The insulator selected should:

- Have high insulation resistance and breakdown voltage
- Have good mechanical strength
- Withstand corrosive chemicals used in processing
- Not absorb water

Samsung, EX1028, p. 292

TABLE 16.2 Key properties of PWB base materials.

Resin/Filler	T_g (°C)	Lateral CTE (ppm)	Permittivity, ϵ_r at 1 MHz	Max. Water Absorption (%)	Peel Strength of Foil (N/mm)
Phenolic/Paper	125	14 . . 18	4.5	0.75	>2.0
Epoxy/Glass (FR-4)	130	14 . . 18	4.9	0.15	>2.0
Polyimide/Glass	250	12 . . 16	4.5	0.35	>1.4
Polyimide/Quartz	280	6 . . 8	4.0	0.35	>1.2
Epoxy/Aramid	180	7 . . 9	3.9	0.44	>1.7
BT-Epoxy	185	13 . . 14	4.3	0.19	>2.0
Cyanate Ester	240	—	3.7	0.40	>6.0
Polyimide/Aramid	230	7 . . 9	3.6	0.81	>1.6

- Not degrade at process temperatures
- Be able to drill through
- Not expand too much in the Z-direction
- Be able to dissipate heat in product use

The base material, or the insulating board of a rigid PWB, is a sheet of laminate reinforced resin. A large majority of the laminates are produced using epoxy, or less frequently phenolic, and polyimide resins are also in use for advanced applications. The reinforcing materials, or the fillers, are typically glass cloth, paper, asbestos, aramid, nylon, and so on. Table 16.2 provides information about the main characteristics of base materials of standard resin/filler combinations. The main parameters include the glass transition temperature, T_g, which is defined as the temperature at which an amorphous polymer changes from a hard and relatively brittle condition to a viscous or rubbery condition. When this transition occurs, many physical properties such as hardness, brittleness, elastic modulus, coefficient of thermal expansion, and specific heat undergo significant changes. Another important parameter is the *coefficient of thermal expansion* (CTE), which value numerically describes the dimensional stability.

Copper conductor tracks are fabricated as layers on the insulating board to electrically interconnect the components that are mounted onto the board and soldered to the tracks later. Since copper has excellent electrical and thermal conductivity, good mechanical properties and sufficient adhesion to the epoxy-glass layer, electroless and electroplated copper layers are the most popular conductors. For flexible boards, such as foil keyboards, cameras and other special applications, polymer conductor layers made of silver-filled *polymer thick-films* (PTF) are also in use.

16.6.4 Standard PWB Materials

1. The FR-4 epoxy-fiberglass laminate is the standard printed wiring board material for most board applications because of its acceptable dimensional stability, heat resistance, good adhesion and large area processability, resulting in low cost. The key components in a FR-4 material are the epoxy-resin, glass fiber and copper conductor, which are depicted in Figure 16.13. The resin is brominated, so that

Samsung, EX1028, p. 293

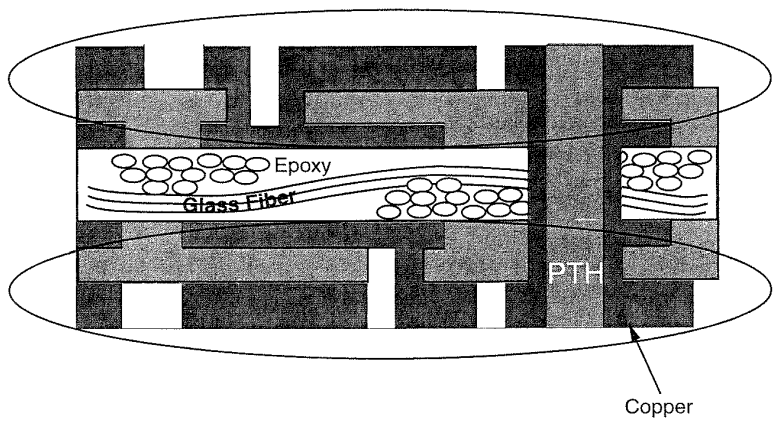


FIGURE 16.13 Elements of an FR-4 substrate (cross section).

the material has the property of "flame retardancy." FR in FR-4 stands for flame retardancy. As discussed in Chapter 22, this brominated flame retardancy gives rise to toxic gases when burnt, such as in accidental fires.

2. Polyimide-epoxy-resins are also used with fiberglass reinforcement for some special rigid PWBs requiring higher temperature for assembly, as they retain their flexural strength up to 350°C or higher. This value is much higher than the standard soldering temperature of 220°C, typically encountered in SMT assembly. On the other hand, polyimide laminates are considerably more expensive than their epoxy equivalents. For flexible PWBs, it is also the polyimide insulating material, which is used without any reinforcement, or with a low percentage of filler like quartz powder. In some cases, photosensitive polyimides like dry films are used in order to make via formation easier and more economical.

16.6.5 Production of Laminates

The boards are generally produced with one or both sides covered by copper foil, and called copper clad laminates. Electrolytic plating onto a stainless steel drum, slowly rotating in the liquid electrolyte, produces the foil. The side of the foil in contact with the drum is smooth and shiny, whereas the other side is matte and granular. The thickness of the copper foil most commonly used is 35 μm, but for fine line circuits, in order to obtain better resolution, foils as thin as 5 μm are also in use. The adhesion of the foil to the organic reinforced prepreg (pre-impregnated laminate) is achieved at the lamination stage, by pressing the granular side of the foil to the resin of the laminate and curing at increased temperature. The formation of a laminate is shown as a simple figure in Figure 16.14.

Let us take FR-4 as an example. The major raw materials required for producing a FR-4 grade laminate are:

Samsung, EX1028, p. 294

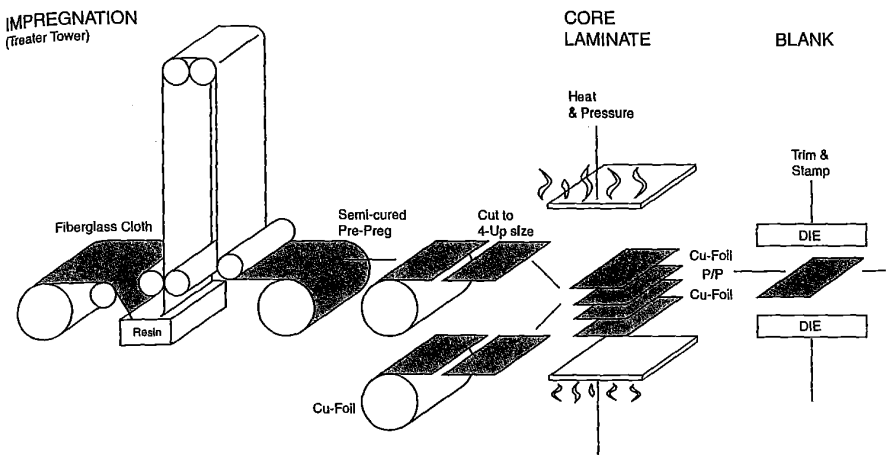


FIGURE 16.14 Production of laminates.

1. Epoxy-resin suitably blended
2. Filler material in the form of glass cloth
3. Copper foil of required thickness

Epoxy-Resin

The term *epoxy* is the generic name given to a series of resins, which are derived from the reaction between chemicals known by the names bisphenol and epichlorohydrin, in the presence of a hardener.

The stages of laminate production are as follows:

Stage A This is the first operation, known as *compounding*. It is the activation of the epoxy-resin by mixing together precise amounts of the resin components in a batch tank.

Stage B In order to impart mechanical rigidity to the final laminate, the above resin blend is reinforced with glass fabric as filler material. As the filler material constitutes about 40–50% of the laminate bulkiness, it contributes significantly to mechanical, electrical and chemical properties of the laminate.

Stage C The C-staging operation is the final curing process after laminating the copper foil to the prepreg material. It is done in batch pressings, or more recently, in continuous operation. A press-load of laminate is normally considered a production lot.

16.6.6 Advanced PWB Base Materials

The standard FR-4 described above, while it is the most widely used PWB material because of its lowest cost, does not meet the high-performance requirements. Its dielectric constant is too high; its water absorption is very high as well. Its thermal expansion

Samsung, EX1028, p. 295

coefficient of 17 ppm/°C in relation to 3 ppm/°C of silicon presents thermomechanical stress challenges. Its glass transition temperature (T_g) of 130°C, in relation to the 220°C surface mount assembly temperature requirement, makes it difficult for processing.

Advanced materials play a vital role in the manufacture of thinner, smaller, and faster interconnect systems. Recent developments in advanced materials provide new degrees of freedom in properties such as thermal resistance, ultra-thin dielectrics, and small-hole formation. These materials include high temperature resins such as BT epoxy, cyanate ester and polyimide. Other newer materials also include so-called linear laminates that provide additional benefits.

Recent developments in PWB materials have been directed toward improving their dimensional stability and surface smoothness. This allows patterning of smaller features, reducing their dielectric constant and dissipation factor to meet the requirements for very high frequency RF applications, and replacing glass reinforcement with laser processable materials to make laser drilling easier.

Linear laminates still use glass filaments for reinforcement, but instead of being woven, the very thin filaments are placed parallel to form a layer; such layers are oriented alternately perpendicular to one another to make a smooth reinforcing fabric, thus reducing the surface roughness of the final resin-impregnated insulating board.

Aramid paper reinforced laminates, using paper-like nonwoven aramid fabric with epoxy-resin impregnation, exhibit very good dimensional stability with near-to-silicon CTE, have a smooth surface, and can be easily processed by laser.

16.6.7 Advanced Filler Systems

The filler material provides the mechanical rigidity of laminates. There are three types of glass fibers: E-glass, S-glass and D-glass. Traditionally, this structure has been achieved with woven E-glass fabrics of various weave styles and various filament diameters. Glass filaments are the basis of woven and nonwoven glass fabrics.

E-glass comprises the vast majority of all applications because of its general properties and its low cost. This glass is similar in composition to the pink insulation that is typically used to insulate homes. The low cost comes from the volume of manufacturing as well as from low temperature processing. S-glass, on the other hand, is based on silica glass, as shown in Table 16.3, which is a much higher temperature glass, and hence the extra cost. D-glass is also a high temperature glass. Both S- and D-glasses are typically used

TABLE 16.3 Glass composition.

	E-glass	S-glass	D-glass	Quartz
SiO_2	52–56	64–66	73–75	99.97
Al_2O_3	12–16	22–24	0–1	—
CaO	15–25	<0.01	0–2	—
MgO	0–6	10–12	0–2	—
B_2O_3	8–13	<0.01	18–21	—
Fe_2O_3	—	0.1	—	—
Zr_2O_3	—	<0.1	—	—

Samsung, EX1028, p. 296

TABLE 16.4 Typical woven reinforcement properties.

Reinforcement Type	CTE, 10^{-6}/K	Dielectric Constant at 1 MHz	Styles Widely Available	Thickness/Ply	
				mm	mils
E-glass	5.04	5.80	106–7642	0.05–0.23	2.0–9.0
S-glass	2.80	4.52	1080,116,7628	0.06–0.18	2.5–7.0
D-glass	2.00	3.95	1080,2116,7628	0.06–0.18	2.5–7.0
Quartz	0.54	3.78	503,525	0.08–0.13	3.0–5.0
Kevlar 49	−5.20	4.00	120	0.1	4.0
Technora HM-50	−7.50	4.00	120	0.1	4.0

either to lower the thermal expansion or the dielectric constant of the board. Examples are given in Table 16.4.

Figures 16.15, 16.16, and 16.17 illustrate the electrical, and coefficient of thermal expansion (CTE) in both x-y and z-directions for various advanced materials. It should be noted that while the dielectric constant of these advanced materials is superior to ceramics, they lack significantly in dielectric loss except for PTFE. This material and its modifications, however, are expensive and hard to process.

16.6.8 Advanced Resin Systems

Resin systems provide the bond between the copper and the filler, and have a strong influence on the final electrical, mechanical and physical properties of the multilayer structure. Resin systems are moving toward higher T_g and better electrical properties than conventional epoxy. The T_g of 180°C is proving to be a very cost-effective balance for high-technology boards, which demand good thermal resistance for wirebonding and assembly. Also, the reliability of a 180°C-T_g epoxy board is improved over a 140°C-T_g

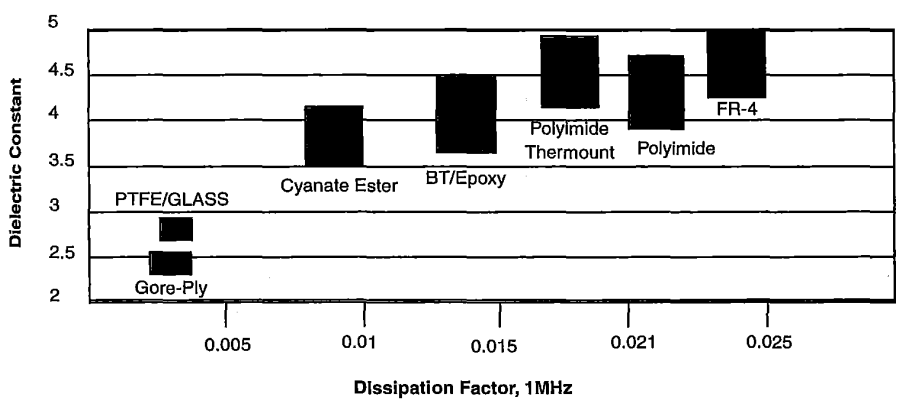


FIGURE 16.15 Electrical properties of various advanced PWB materials.

Samsung, EX1028, p. 297

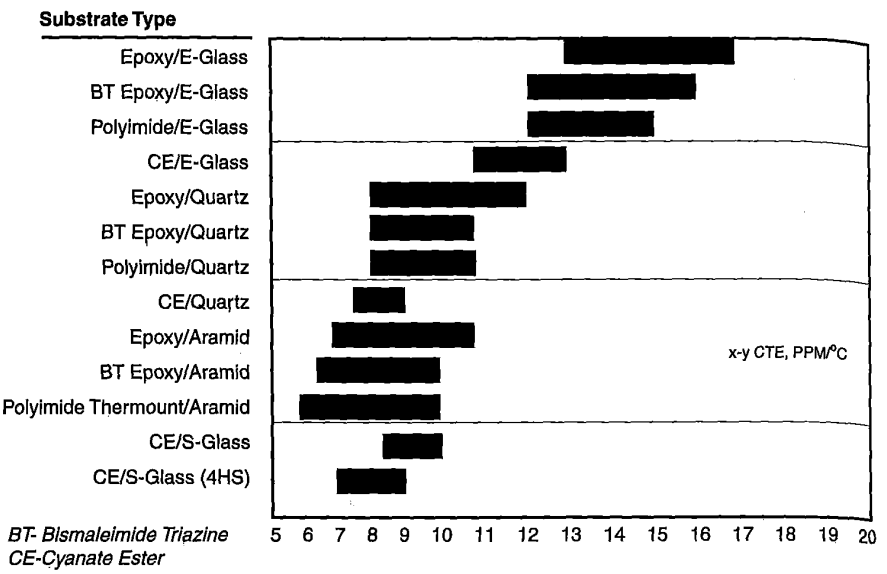


FIGURE 16.16 CTE of materials in the *x-y* direction.

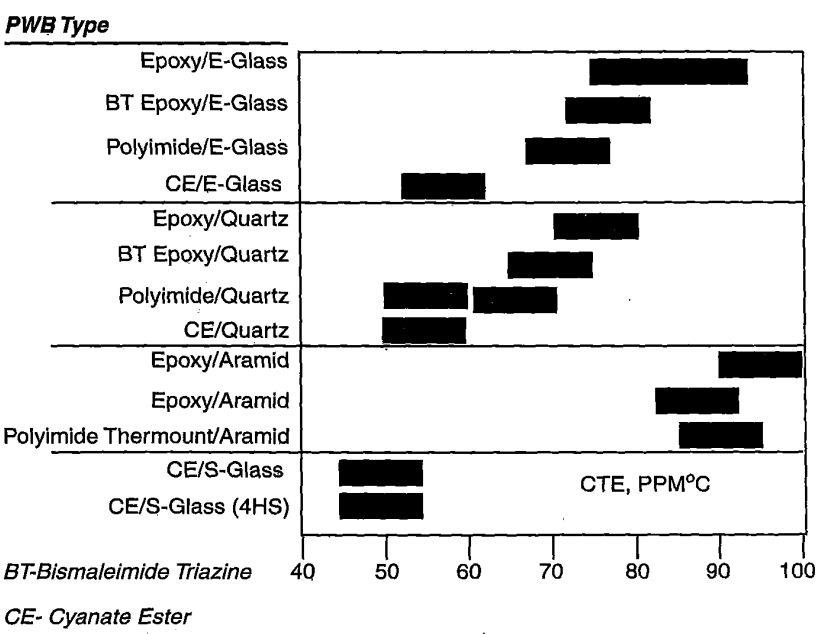


FIGURE 16.17 CTE of materials in the *z* direction.

Samsung, EX1028, p. 298

TABLE 16.5 Typical resin properties.

Resin System	Dielectric Constant at 1 MHz (a)	Glass transition Temperature (a) °C	°F	Relative Cost
FR-4 epoxy	3.50–3.60	125–135	255–275	1
Polyfunctional FR-4	3.50–3.60	140–150	285–300	1–2
High-temperature, one-component epoxy system	3.90–4.00	170–180	340–355	3–6
Bismaleimide triazine epoxy	3.20–3.30	180–190	355–375	3–6
Polyimide epoxy	3.50–3.60	250–260	480–500	10–20
Cyanate esters	2.80–3.00	240–250	465–480	20–30
Polyimide	3.30–3.40	>260	>500	10–20
PTFE (melting point)	2.03–2.09	327	620	10–15

(a) Numbers for cast resin samples

board during reliability thermal cycling. In addition, during thermal shocks, through-hole survives with less impact to the hole from the solder. Pad lifting and various other defects, such as voiding, are greatly reduced. Some of the advanced resin systems are BT/epoxy-resin, polyimide resin and cyanate ester resin. Their properties are listed in Table 16.5.

BT/Epoxy-Resin

Bismaleimide triazine (BT) is a high-performance resin system with good electrical and thermal properties. The BT boards are primarily used in IC packaging applications as BGAs, and, system-level packaging applications as in MCM-Ls. The glass transition temperature of BT-epoxy is 180°C.

Polyimide Resin

Polyimide resins have extremely high heat resistance, but they are very expensive. They are utilized where heat resistance is critical, such as burn-in boards, oil exploration electronics and military avionics. They have a T_g of 260°C.

Cyanate Ester Resin

This is a premier resin for advanced applications, because it combines the best electrical and thermal properties that are crucial to miniaturization and high performance. The dielectric constant is ~3.5 and T_g is 250°C. It has a narrow manufacturing process window and thus require specialized circuit processing.

16.7 STANDARD PRINTED WIRING BOARD FABRICATION

A standard PWB process sequence can best be described by the flow chart in Figure 16.18.

Samsung, EX1028, p. 299

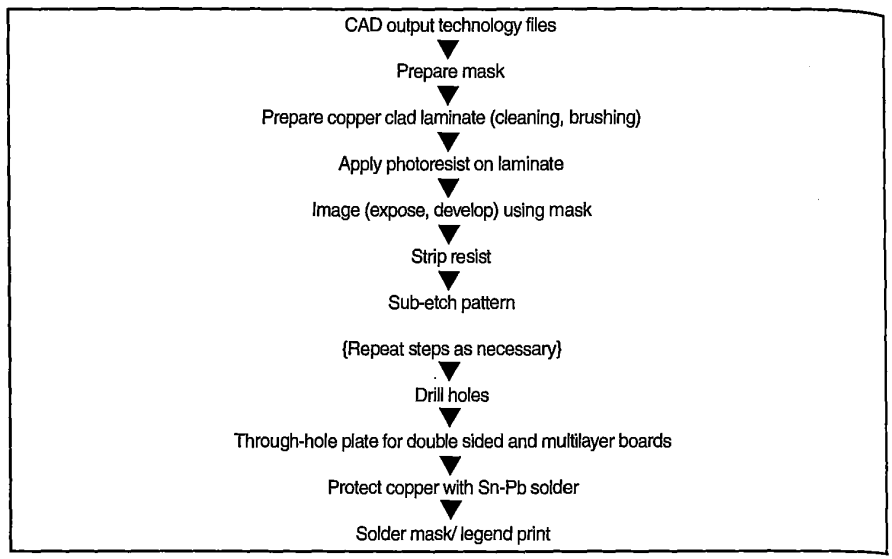


FIGURE 16.18 PWB process flow chart.

16.7.1 Fabrication Process Sequence

Some of the individual processes for PWB fabrication are described below.

Imaging

The imaging process is comprised of several sequential steps, which together allow for the metal interconnect pattern to be formed on bare substrate. The steps are interactive, and careful control is required in order to reproduce a master pattern onto the substrate with high fidelity. The process involves the coating of a polymeric material onto the substrate, and then patterning the material, either by depositing the initial layer in a patterned fashion (screen-printing), or by the use of a master to the polymer coating. The choice really depends on the feature sizes in the design. Feature sizes smaller than 200 microns (8 mil) can be formed by the photolithographic process.

The process sequence is as follows:

1. Clean surface
2. Apply photoresist
3. Expose photoresist
4. Develop photoresist image
5. Pattern transfer image (plating or etching)
6. Strip photoresist

The photoresist materials are either liquids or dry films of an initial liquid system. Both types of materials are widely used. The key factor to the functioning of these

Samsung, EX1028, p. 300

materials is their solubility in the developing solution after exposure to light. Dry film photoresists are commonly used for pattern formation, by both plating and etching of underlying materials. On the other hand, liquid resists are used for precision work [typically 50 microns (2 mil) and less]. They are coated by roller, spray, curtain, meniscus and electrostatic methods. The cleanliness of the substrate and the coating and drying operations are critical for a high-yield process. The chemistry of these materials is mostly similar, but there is a choice of developing agents namely, aqueous, semiaqueous or solvent.

Drilling

The purpose of drilling is to produce an opening through the board which will permit a subsequent process to form an electrical connection between top, bottom, and, sometimes intermediate conductor pathways, and to permit through-hole component mounting. The quality of a drilled hole is measured by its interface with plating, soldering and forming a high-reliability electrical and mechanical connection.

The elements of the laminate drilling process are drill entry and backup foils, drill bits and equipment-related parameters. When all of these elements are in order, high-quality drilled holes are the result. Even with the best equipment, drill holes need to be examined and treated properly before plating copper.

Drill bits are usually made of tungsten carbide. The entry and back-up foils are aluminum-clad materials. This helps in removing excess heat generated during the drill process. The most common defects during drilling are delamination, smear, burr, and debris.

For high manufacturing throughput, drilling machines with over 100 spindles operate simultaneously at high speed and feed. Representative operating ranges in the industry are spindle speeds of 50,000–100,00 rpm, a feed of from 50–400 cm/min, and high retraction rate. The selection of drill design, materials, and construction, as well as the drilling parameters, are based on a proper balance of many economic and technical factors, for example, drill breakage, drill wear, resin smear, resin-glass fiber fracture, hole exit and entry burr, nail heading (burr on an inner plane), and drill wander. Drill breakage is primarily caused by torsional stress and bending (buckling) stresses initiating a fracture around a weak region of the drill body. Increasing the cobalt content in the tungsten carbide drill provides even more resistance and thus leads to duller drills for the same number of holes drilled. Microsurface finish has been reported to be effective in reducing drill friction and lengthening wear life.

Drill smear is probably the most important factor determining hole quality. Drill smear is caused by grinding the heated resin chip debris into the hole wall. In the repetitive drilling steps, the drill itself is gradually heated to a steady state when the heat dissipation through the hole wall to the surrounding air and transport of the swarf are balanced by energy from drilling and friction. For a set of board and drill parameters, the temperature reached during drilling of a copper internal plane can be as high as 150°C in a functional land and 230°C in a nonfunctional land. Worn drills can cause temperatures to climb as high as 205°C in a functional land and 650°C in a nonfunctional land. With a glass transition temperature of 130°C for epoxy-resin, it is not surprising to observe epoxy-resin smear over the copper inner plane and land.

Samsung, EX1028, p. 301

Plating

A major part of manufacturing involves the wet process. The plating aspects of wet chemistry include deposition of metals by electroless and electrolytic processes. The typical final requirements of plated copper are:

- High electrical conductivity
- Good mechanical strength
- High ductility and elongation
- Excellent solderability
- Good tarnish and corrosion resistance
- Good etchant resistance

The steps involved in electroless copper plating include cleaning, copper microetching, hole and surface catalyzation with palladium, and electroless copper plating by using a reducing agent like formaldehyde. In electrolytic plating, where external DC current is applied, cleaning and treatment of the surface are important prior to plating. Copper can be plated in two routes, panel and pattern plating. Panel plating is the process wherein the entire surface area and the drilled holes are copper plated. Pattern plating, on the other hand, is a selective process, wherein only the desired circuit pattern and holes receive copper build-up and etch-resist metal plate. The preferred copper-plating bath is acid copper sulfate.

Electroplating is also used to protect copper by plating Sn or Sn-Pb, and by deposition of nickel as a solder barrier as an undercoat, prior to application of gold in select areas of the board.

Etching

Etching here refers to the removal of copper to achieve the desired circuit patterns. Etching is also used for surface preparation with minimal metal removal (microetching) during electroless or electrolytic plating. The most common chemical etching systems are based on alkaline ammonia, hydrogen peroxide–sulfuric acid, and cupric chloride. Other systems include persulfates, ferric chloride and chromic–sulfuric acids. The process steps include:

- Resist stripping
- Precleaning
- Etching
- Neutralization
- Water rinsing
- Drying

Etching results are dependent on proper image transfer on boards and qualities of plated-metal etch resists or photoresist, as the case may be. Other controls include agitation, temperature, pH, bath concentration and regeneration of bath constituents.

16.7.2 Single-sided Board Fabrication

The variations of manufacturing routes to progress from a sheet of copper clad laminate to a PWB are numerous. The process proceeds through different types of masking meth-

Samsung, EX1028, p. 302

ods, etching and deposition techniques. Here, three characteristic sequences are discussed. For single-sided boards, the simple masking and etching technique is used which is illustrated in Figure 16.19.

Starting with a copper clad laminate, the pattern of the wiring is screen-printed onto the surface as a positive mask. Alternatively, photoresist technology can also be used to make the same mask, but the much cheaper screen-printing is sufficiently good for the rough resolution of common single-sided boards. From the uncovered places, the copper foil is removed by etching, then the mask layer is stripped, and in the last step the holes are drilled.

16.7.3 Double-sided PWB Fabrication

The basic process flow for fabricating a double-sided through-hole connected PWB is illustrated in Figure 16.20.

This sequence includes main groups of processing steps as follows (pattern plating):

1. Starting with a double-sided copper clad laminate, holes are fabricated by drilling. Then they are deburred and cleaned.
2. The panels are sensitized and an electroless copper flash is deposited to make the holes conductive for electroplating.
3. A plating resist, usually a negative mask, is produced by the following steps: dry film resist lamination, transferring the image through photo-mask UV light exposure and development.
4. A pattern of copper layer is electroplated onto the uncovered areas. The minimum thickness of the copper layer required in the holes is usually 20 μm.
5. A second metal, typically a tin-lead solder, is electroplated over the copper to act as an etch resist mask, to improve solderability during subsequent board assembly process and to provide protection against the corrosion of the copper layer.
6. Photoprinted mask is stripped, and copper layer is etched to get the required pattern.

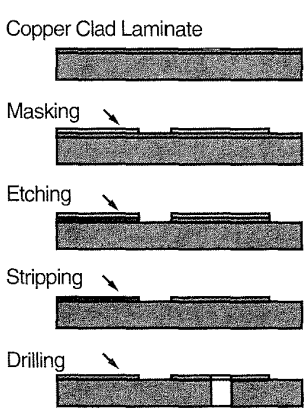


FIGURE 16.19 Simple masking and etching sequence for single-sided PWBs.

Samsung, EX1028, p. 303

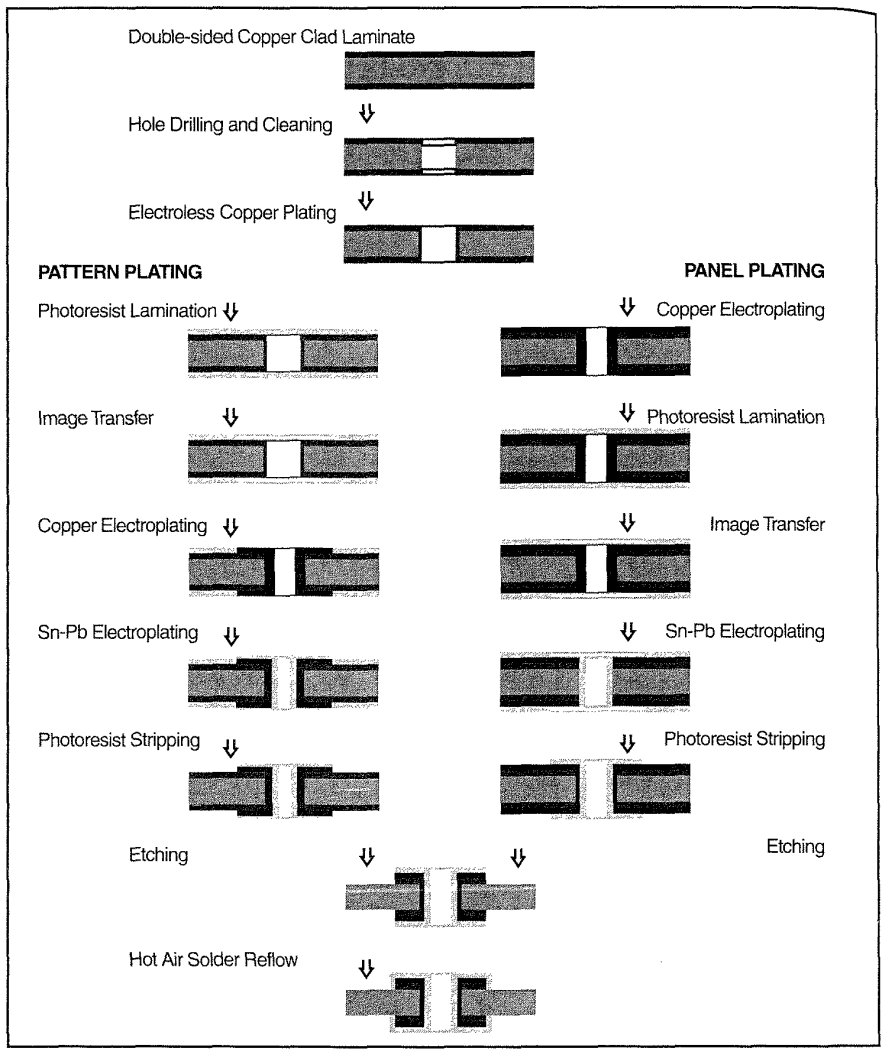


FIGURE 16.20 Fabrication sequences of double-sided PWBs.

7. Finally, the tin-lead coating is reflowed to cover and protect the edges of the copper layer, and to improve the overall appearance of the board. This solder reflow is usually carried out by a process called *hot-air leveling* by blowing hot air to the surface.

The second fabrication route used for double-sided printed wiring boards is known as panel plating which is also depicted in Figure 16.20. The main steps are as follows: drilling holes and cleaning; activation and deposition of electroless copper; electroplating

Samsung, EX1028, p. 304

copper across the whole panel; laminating with photoresist; image transfer by exposure and development; electroplating tin-lead solder; removal of photoresist; etching unprotected copper; and reflow of the solder coating.

A modern electrochemical process called direct plating is also in use. It substitutes electroless plating and electroplating, thus simplifying the processing sequence. The modified sensitizing step prior to the direct plating process makes the surface slightly conductive, which launches the electroplating process directly.

16.7.4 Multilayer PWB Fabrication

Fabrication of multilayer printed wiring boards contains three separate process sequences, which can be understood from Figure 16.21.

1. First, the panels with the inner conductive layers are produced by simple double-sided etching from copper clad prepreg laminates;
2. Then, these etched panels, adhesive prepregs and unetched outer panels are laminated, pressed and cured (heat treated) to get a rigid board;
3. Finally, this board is processed by a sequence of process steps corresponding to pattern or panel plating of double-sided boards.

The special and most important steps in the manufacturing process of multilayer PWBs involve the lamination of etched and unetched panels. Any combination of single-sided and double-sided prepreg panels can be used to form the required multilayer structure, with the only condition that all inner layers must be patterned before lamination (Figure 16.22). The resin used in these prepregs must have a property of great flow, in order to equalize the thickness variation of the etched inner panels. Other important steps in the manufacture of multilayer boards is that the holes are drilled after lamination, and the interconnections between the inner and the outer copper patterns are produced by plating through the via holes using direct copper plating (as in Figure 16.21), or a sequence containing electroless and electroplating processes.

16.7.5 Solder Mask

There are several types of solder masks. Each type of solder resist mask has advantages and limitations. Selection of a solder resist mask for a PWB is influenced by the board fabrication, the assembly process and the final application. There are three primary types of solder resist masks in use today. They are:

1. Screen-printed
2. Dry film
3. Liquid photoimageable (LPI)

A screen-printed mask only requires the thick-film application step before it is ready for the final cure. Both dry film and liquid photoimageable (LPI) are photosensitive materials. Therefore, they require an imaging and developing process after the film application. However, dry film and LPI have the ability to produce finer features than the screen-printed solder masks, making them more suitable for designs with finer pitch. There are two types of developing processes. Those based on solvents and those based

Samsung, EX1028, p. 305

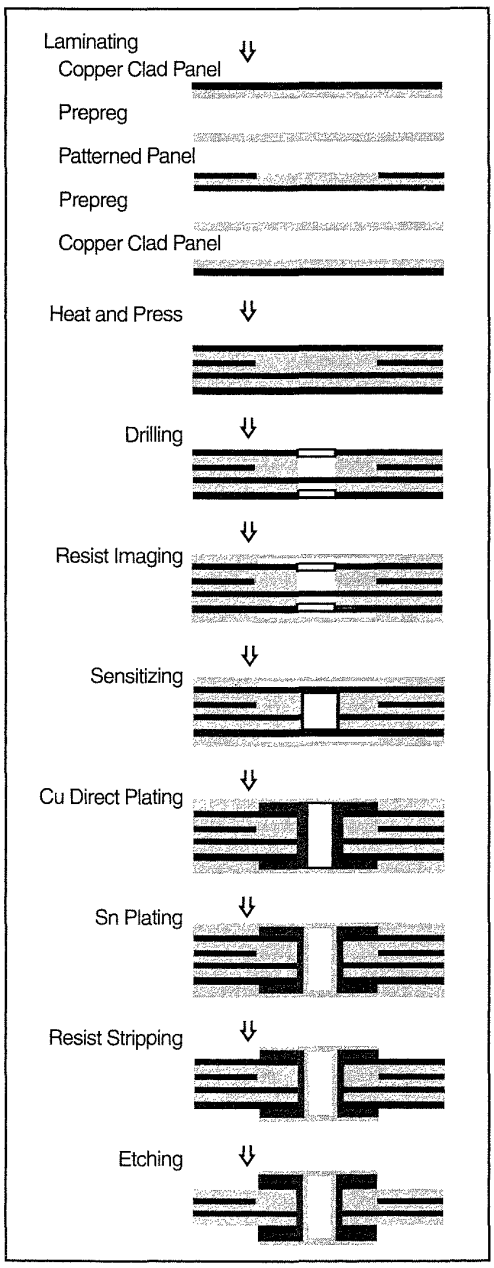


FIGURE 16.21 The major steps in the fabrication of multilayer PWBs.

Samsung, EX1028, p. 306

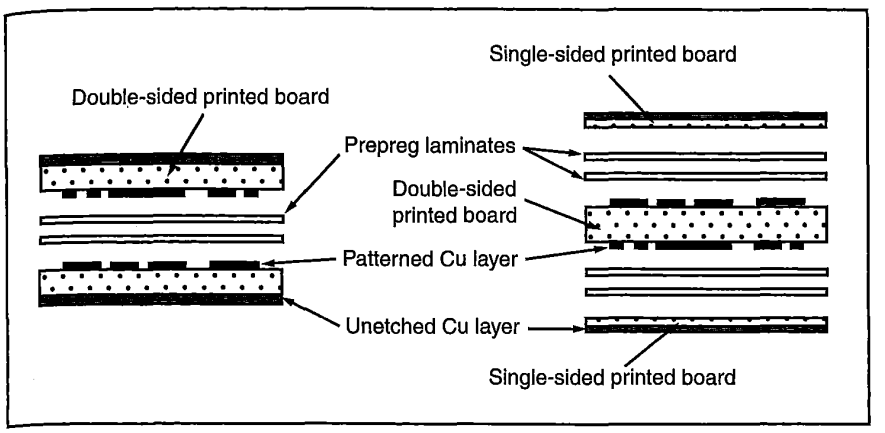


FIGURE 16.22 Lamination combinations of four-layer PWBs.

on water. Generally, aqueous developing solder resist masks are not as robust as the solvent developing types, and they are less environmentally friendly. With the increasing dominance of fine pitch surface mount components, the liquid photoimageable solder resist is becoming the most prevalent.

All exposed copper surfaces not covered by solder mask need to be protected by one of the finishes described below.

1. Electroplated nickel + matte tin. Typical thickness: 7.5-μm tin over 5-μm nickel. Features: solderable surface and good shelf life.
2. Electroplated nickel + hard gold. Typical thickness: 0.75–1.25-μm gold (99.7%) over 5-μm nickel. Features: excellent corrosion resistance, shelf life, hardness and wear resistance, best for surface rotary switches, on-off contacts, and edge connectors.
3. Electroplated nickel + soft gold. Typical thickness: 0.75–1.25-μm gold (99.9%) over 5-μm nickel. Features: excellent corrosion resistance and shelf life, fair wear resistance, good for pressure contacts and aluminum or gold wirebonding.
4. Electroless plated nickel + immersion gold (99.9% gold). Typical thickness: 0.02–0.1-μm gold over 4.5-μm nickel. Features: excellent corrosion resistance, solderability and shelf life, good for aluminum wirebonding and for fine-pitch technology.
5. *Hot-air solder leveling* (HASL) (eutectic 63% tin–37% lead). Typical coating thickness: 1.5–5 μm, depending on application and design. Features: excellent solderability, good shelf life, only 0.625 mm pitch capability.
6. *Organic solderability preservative* (OSP). Typical coating thickness: 0.2–0.5 μm. Features: excellent solderability, surface coplanarity and hole size uniformity, good shelf life, excellent for use in fine-pitch technology. Board is not subjected to thermal shock (as with HASL).

Samsung, EX1028, p. 307

When reflow soldering is to be used for preparing the solder joints, additional solder material is required at soldering pads to supply the necessary amount of material for solder fillet and for filling up the holes. The solder fillet, which is the concave surface of the solder joint at the intersection of the metal surfaces, characterizes the quality of the solder joint. Its proper formation highly depends on whether the amount of solder at a joint is in correlation with the dimensions of the pad, metal lead and hole of the joint. Solder-paste is generally printed through a screen or metal stencil mask to provide the appropriate amount of solder for each joint. This step is not necessary when wave soldering is used, since the solder bath provides the required material.

16.8 LIMITATIONS IN STANDARD PRINTED WIRING BOARD PROCESS

16.8.1 New Products Require Higher Pad Densities

The pitch and area density of component connections ultimately drives PWB design and its processing technologies. IC package evolution to area array-based packages such as BGA and CSPs has driven dramatic escalation in I/O lead density, initially in high-performance computers and subsequently in hand-held electronics such as cellular phones and camcorders.

In general, increased lead or pad density of printed wiring boards is achievable through a trade off in substrate layer count, via size, and circuit line width and space. However, above a threshold of approximately 20–30 pads/cm^2, most designers are forced to deploy small-diameter blind via structures, which enable more efficient circuit routing. Leading portable equipment now has pad densities in the range of 35–90 pads/cm^2. This need obsoletes the conventional PWB technology. Figure 16.23 illustrates the gap in pad density between today's conventional laminated PWB that is available and the needed PWB. The higher pad density is achieved by so-called microvia technologies.

16.8.2 Drilled-Hole Technology Becomes Too Expensive

Drilled holes and vias are among the basic structures of PWBs; however, they have two main problems:

1. They take up too much space on the board and do not provide high pad densities. This is because the vias and pads block routing channels.
2. They become expensive as their size is lowered. Standard drill technology can produce a via pad size of 0.025-in with a 0.014- to 0.015-in drilled hole. But as the hole size decreases, they get very expensive, as illustrated in Figure 16.24, wherein the cost of drilling has gone up by a factor of 18 times. As a result, small-hole drilling alone can contribute as much as 30–40% of the total cost of the PWB. This can be realized from Figure 16.25. Also, as the holes get smaller for the same board thickness, the aspect ratio increases. This tends to decrease reliability of the assembled boards as high-aspect holes have problems in getting sufficient plating solution and solder into the hole.

Samsung, EX1028, p. 308

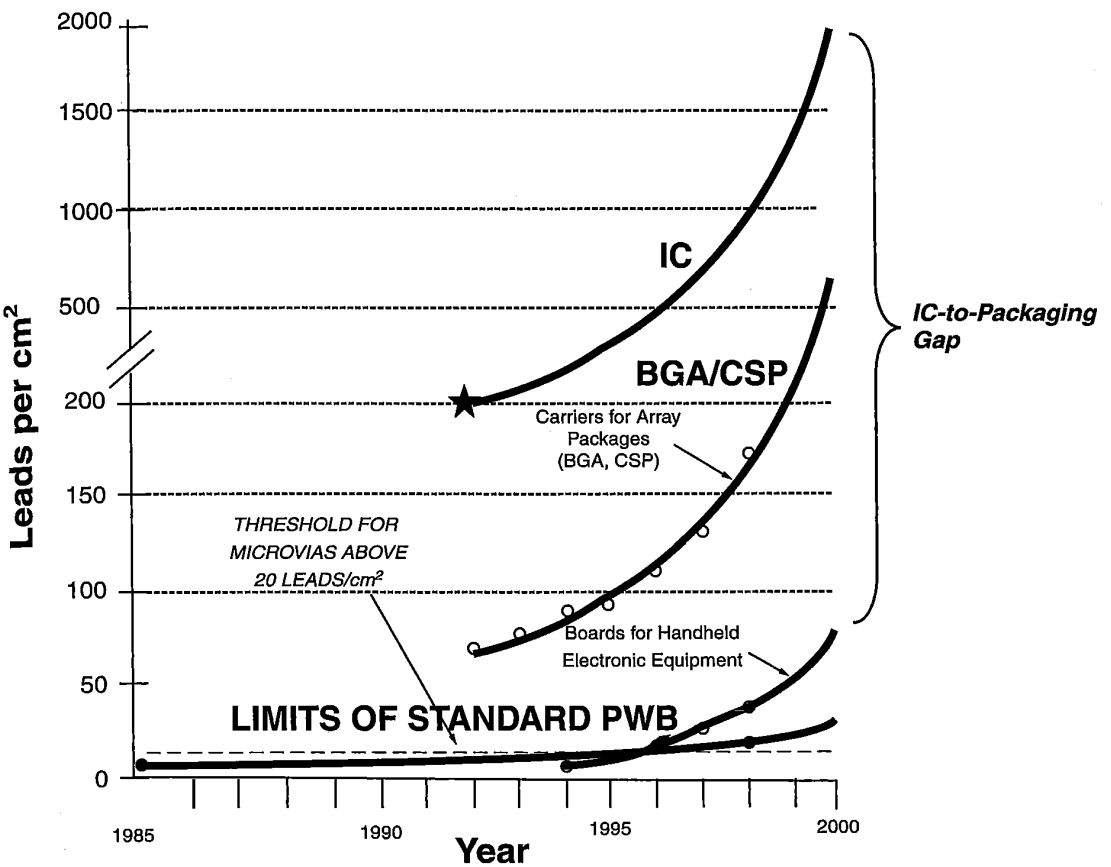


FIGURE 16.23 Limitations and technology gap in today's PWBs.

16.8.3 Microvia Solves these Two Problems

One of the responses to this problem is the use of so-called microvia, or blind via, in the form of non-drilled blind and buried vias. These are fabricated by three primary processes:

1. Laser drilling
2. Plasma or reactive ion etching
3. Photolithography

A PWB with microvias processed with one of the above processes seems to provide a good solution to the gap problem in pad density described above. This technology retains the low-cost materials and processes of PWBs by the use of lamination, chemical and electrochemical processes for the core of the newer PWB, combined with photo-,

Samsung, EX1028, p. 309

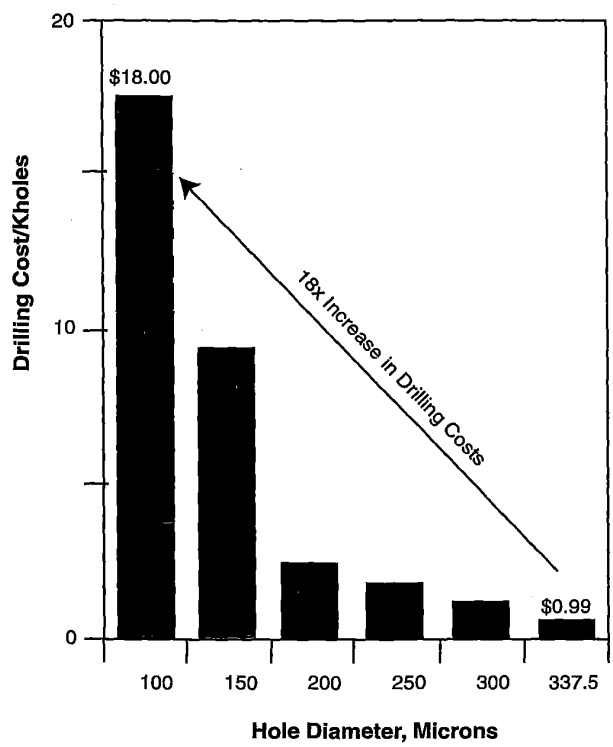


FIGURE 16.24 Drill-hole size vs. cost of drilling.

laser-, or plasmavia openings, as well as thin-film deposition by sputtering, spinning, meniscus coating techniques for the surface layers of high pad density.

PWBs with microvias offer the following advantages for high-volume production:

Increased circuit density. Microvias can be incorporated within the pad structure. The greater via density and the small via size enable a greater density of routing tracks.

Advanced packages enabled. Area array CSP and BGAs can easily be incorporated into the assemblies.

Better electrical performance. The size reduction of the vias and tracks reduces the noise and parasitic effects.

Improved reliability in comparison with drilled through-holes.

Improved thermal conductance through the thin dielectric films.

Lower PWB cost. Even at lower densities, a cost reduction can be achieved if the substrate size and/or the number of layers can be reduced by applying microvias.

16.8.4 Methods of Microvia Generation

The three most common via processes are illustrated in Figure 16.26.

Samsung, EX1028, p. 310

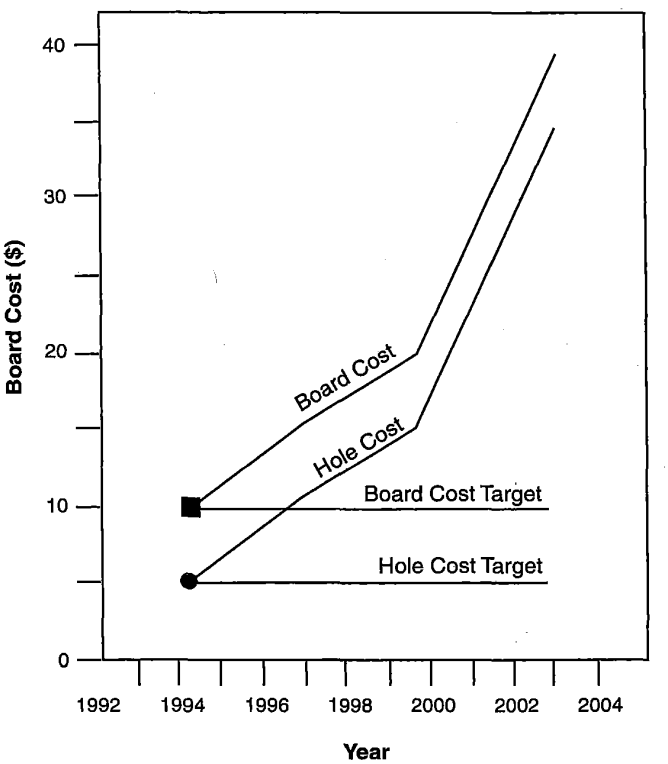


FIGURE 16.25 Drilling cost in comparison to the total cost of PWB.

Photovia Process

The low-cost photovia processing requires layers of photoimageable or photosensitive permanent dielectric materials. These photosensitized polymers, which are available in many chemistries such as polyimide or epoxy, work like photoresists. They are exposed through a mask to form the vias and windows of any geometry defined by the mask and subsequently developed. The exposed and developed films are then cured to obtain the final cross-link properties of that polymer.

Plasmavia Process

Plasma-etched via (PEV) technology applies vacuum processing to remove the polyimide dielectric layer. All vias for one layer are generated simultaneously. The previously patterned copper layer serves for masking—that is the polyimide is etched through the openings of the copper layer—and etching automatically stops at the inner copper layer. Blind vias can be generated from any side of the board. Typical via holes are 60–90 μm in diameter and require 200–300-μm capture pads. Regarding the processed geometry, PEV is a very flexible process: in addition to through and blind vias, it can create slots, windows, stepped windows (using controlled etch time), slanted vias and unique structures as well.

Samsung, EX1028, p. 311

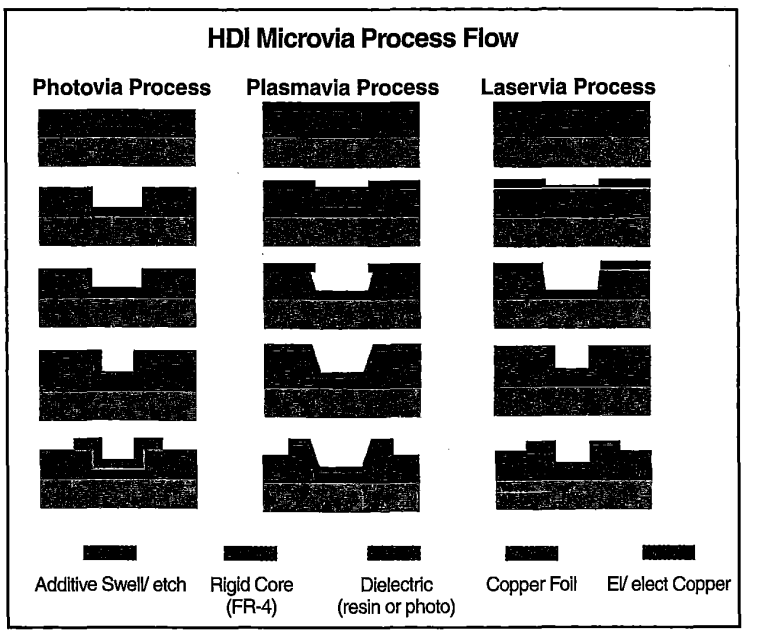


FIGURE 16.26 Common via preparation methods.

Laservia Process

The most important and the most successful microvia technology to date has been laservia generation technology. Similar to PEV, laservias can be generated through polymer films, applying the patterned copper layer for masking by exposing the entire surface with CO_2, UV excimer lasers, or by punching one hole after the other, using a frequency multiplied UV Nd:YAG defocused laser beam. PEV and excimer laser via generation technologies need expensive equipment. However, they can be economical for mass production. Nd: YAG laser technology has the advantage in that copper layers can also be processed with focused beam along the hole perimeter. Blind vias can also be generated by the combination of focused beam spiraling and defocused beam punching. In general, it is less effective, but more flexible; it is preferred for prototyping or smaller batch manufacturing.

Comparison of Via Processes

The debate about what is the best overall via process has been going on for two decades or more. Direct CO_2 laser drilling seems to have taken a clear lead in throughput and in the quality of holes generated, as shown in Figure 16.27. Recently, boards for cell phones have been made at speeds of 20,000 holes per minute per head. As a result, cell phone costs have come down by 30% over the last few years.

Samsung, EX1028, p. 312

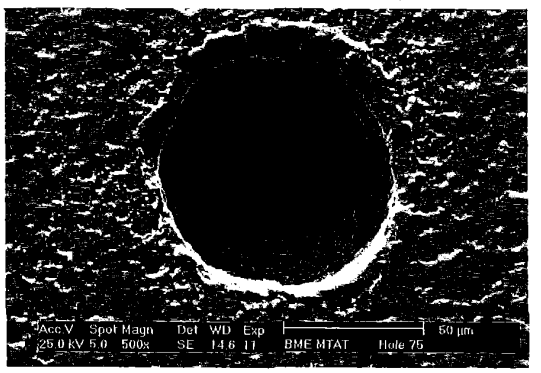


FIGURE 16.27 A photograph of a via generated by laser drilling process.

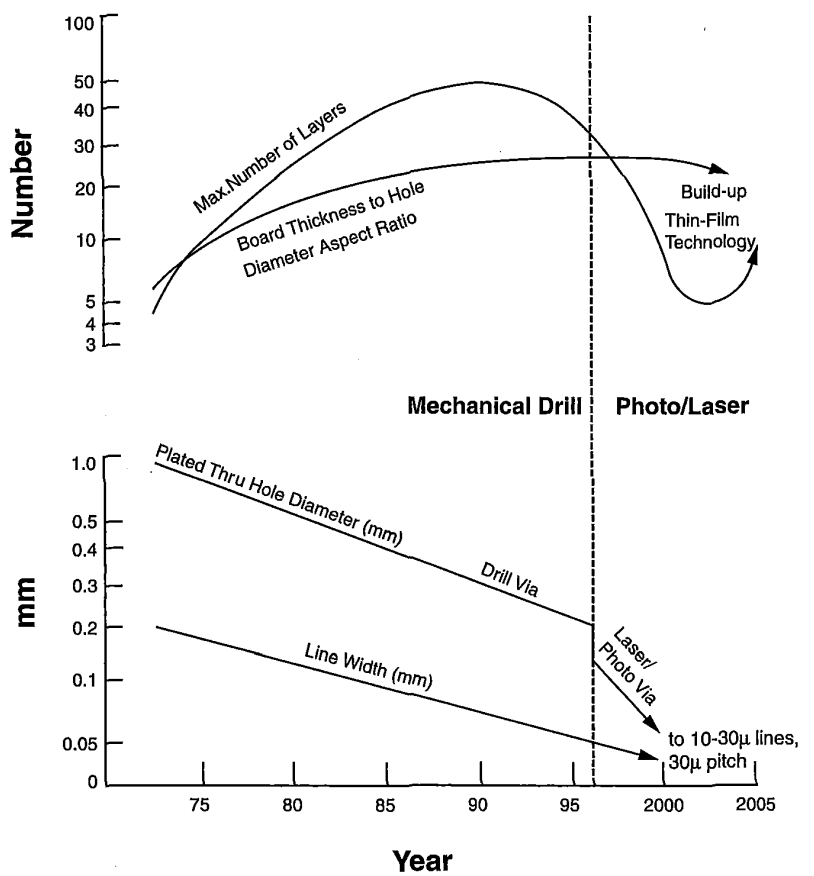


FIGURE 16.28 Comparison of various via generation processes.

Samsung, EX1028, p. 313

Paste-Via Process

Any of these three processes can be used for via hole generation in the paste-via process. After the holes are opened into a single dielectric layer, they are filled with conductive paste. The copper foils are laminated and photoimaged later on the dielectric layer. The complete board is finally prepared by co-lamination and curing.

All four via preparation and wall metallization methods have particular advantages:

- Photovia technology is very productive, resolution is high, but provides irregular hole wall quality, and needs special photodielectric materials
- Plasmavia process provides even hole uniformity and cleanliness, is productive, but needs expensive equipment
- Laservia drilling provides very clean surface and suitable wall shape, is very flexible, but less economic than photovia processing
- Paste-vias are very cheap, but less reliable than wall metallized vias

Figure 16.28 is a trend chart comparing mechanical drill, laser and photovia process features in terms of hole size and aspect ratio over the last two decades.

16.9 MICROVIA BOARDS

Microvia boards are fabricated by three major technologies:

1. *Surface Laminar Circuitry* (SLC) or build-up technology
2. *All Layer Internal Via Hole* (ALIVH) technology
3. *Buried Bump Interconnection Technology* (B^2IT)

The original microvia process invented by IBM Japan has been called surface laminar circuitry. There is a large number of variations of this in materials, via and dielectric processes that generally fall into this category. The two other processes that have gained acceptance are ALIVH and B^2IT. The original surface laminar circuit was for depositing thin-film wiring on the surface of conventional PWB with through-hole. This technology is now generally called *build-up* (BU) technology, sometimes *sequential build-up* (SBU), implying that the dielectric and metallization layers are deposited sequentially.

16.9.1 Build-up Technology

An example of build-up technology is shown in Figure 16.29. It involves the thin-film-on-PWB processing sequence, starting with a conventionally patterned, double-sided PWB, and then structured dielectric and interconnection thin-film layers are sequentially built up onto the board. The key to this process is the "via" formation, which is achieved either with photosensitive polymer technology or with laser technology. In photopolymer technology, photosensitive epoxy, polyimide, benzocyclobutene or other photopolymer dielectric materials, whose typical thickness after coating and curing is 20–40 μm, are used. Their dielectric constants range from 2.6–4.5. Vias and noncircular openings are developed after through mask exposure to broadband wavelength exposure, with high resolution resulting in via diameters of 50–100 μm. In laser process technology, the dielectric used is non-photosensitive.

Samsung, EX1028, p. 314

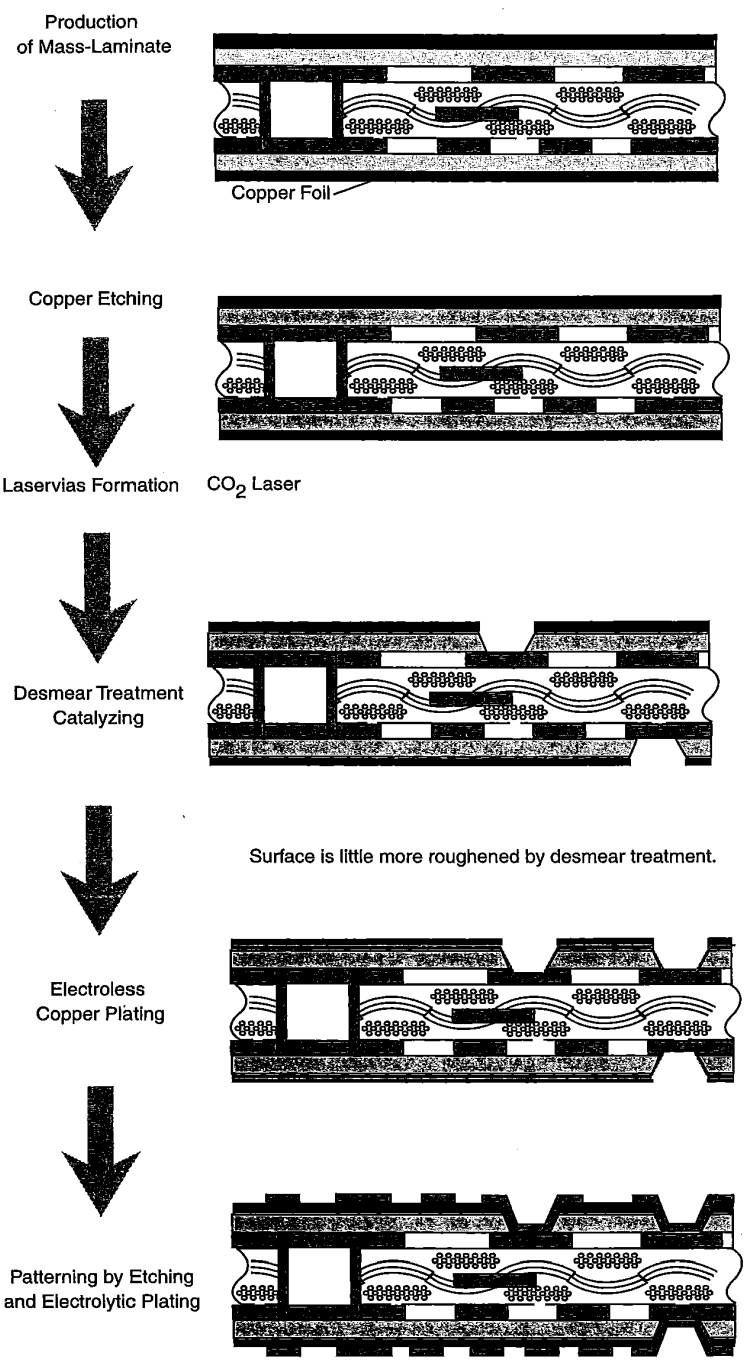


FIGURE 16.29 Sequential build-up (SBU) technology using photovias.

Samsung, EX1028, p. 315

Copper lines and spaces are patterned as small as 20 μm wide, using a fully additive electroless copper plating process. The conductive layer can also be fabricated by direct panel plating by a combination of electroless and electrolytic copper deposition onto the whole surface, followed by conventional photoimaging and wet chemical etching for patterning.

The dielectric is deposited by curtain or slot coating, and more recently by meniscus coating. Another alternative is spin coating of the dielectric and vacuum deposition of the conductive layer, whose small thickness provides the possibility of very small feature dimensions. Finally, electroless gold plating is used for wire bondable surface finish. Two to four thin conductive layers on both sides are typical of build-up technology.

16.9.2 All Layer Internal Via Hole (ALIVH) Process

The ALIVH process, invented by Matsushita in Japan, and illustrated in Figure 16.30, is also a high-volume manufacturing process, currently used in Japan for making boards for the cell phone industry. This process is very similar to the ceramic substrate process, except for the fact that the greensheet in the ceramic process is replaced with a permanent organic dry film.

The different layers, each containing insulating and conducting layers, are fabricated separately, and then they are unified during a single lamination process. The dielectric uses B-stage epoxy-aramid prepregs with laser-drilled vias that are filled with conductive paste. Then copper foils are laminated onto the substrates and patterned by photolithography and etching. A final co-lamination unifies the multilayer substrate.

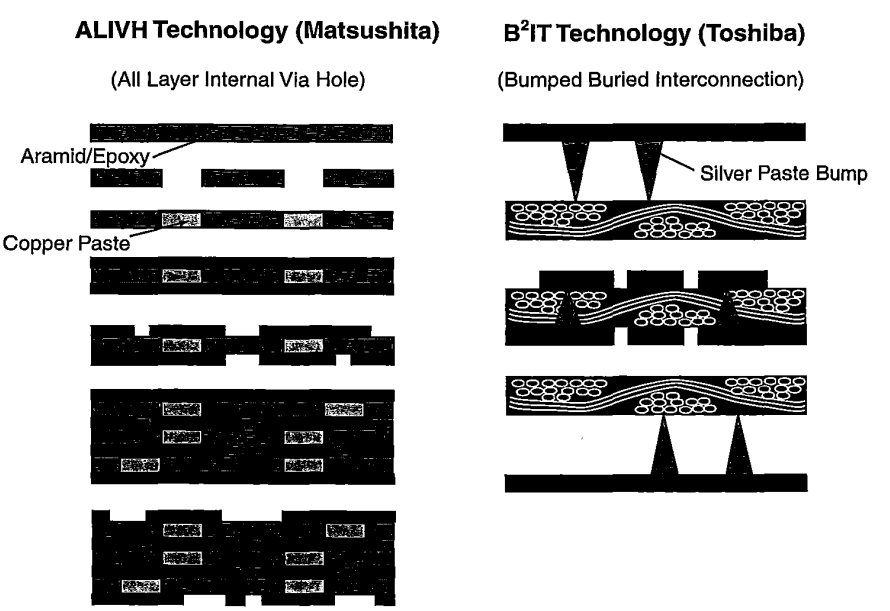


FIGURE 16.30 ALIVH and B²IT process sequence.

Samsung, EX1028, p. 316

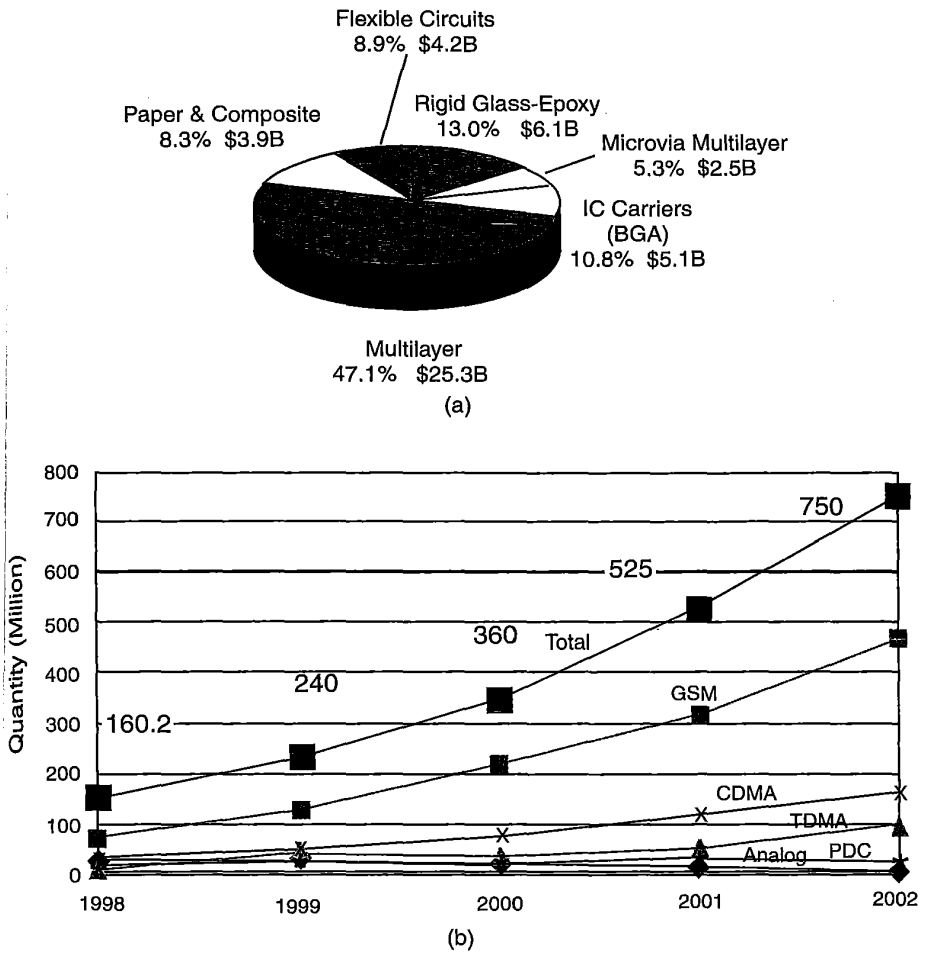


FIGURE 16.31 (a) PWB market; (b) microvia market.

16.9.3 Bumped Buried Interconnection Technology (B^2IT) Process

The B^2IT process, invented by Toshiba, is the concept illustrated in Figure 16.30. This process involves the use of silver bump paste metallization to pierce holes through the dielectric or prepreg, thus providing microvia interconnections. The wiring technology is similar to others described above.

16.10 PRINTED WIRING BOARD MARKET

Printed wiring board is the primary platform for interconnection of electronic components to form system-level boards in all electronic systems. The total PWB market in 1999 surpassed $35B worldwide, as shown in Figure 16.31, and is expected to become $47B

Samsung, EX1028, p. 317

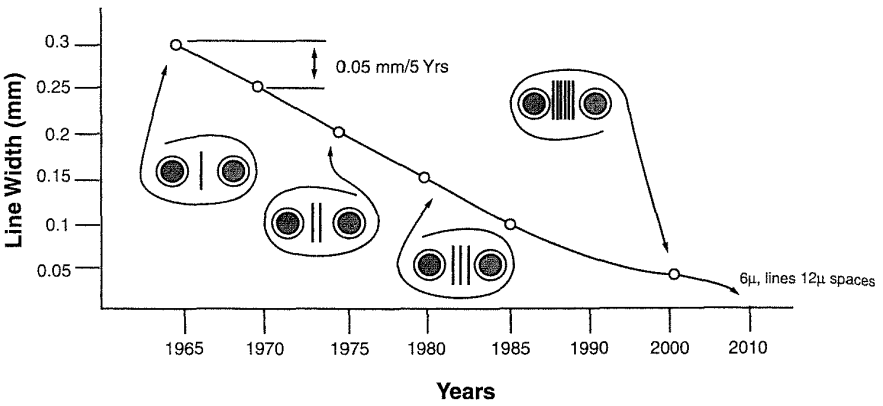


FIGURE 16.32 PWB roadmap.

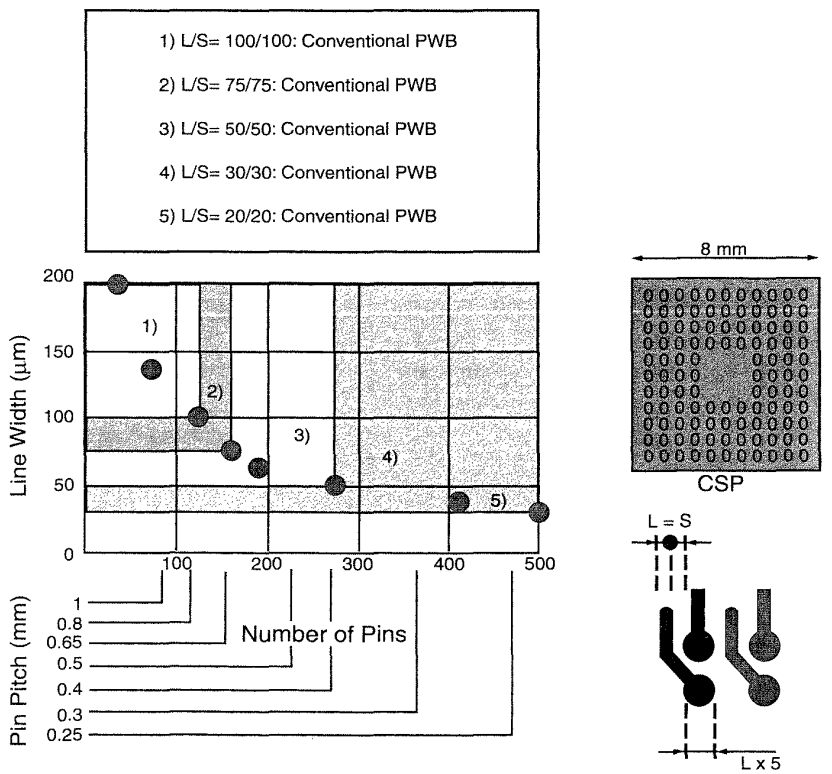


FIGURE 16.33 The relationship between line width, pin pitch, and number of pins.

Samsung, EX1028, p. 318

TABLE 16.6 PWB technology needs for the future.

	1998	2000	2005	2010
Notebook PC (μ)	50–100	50–70	30–70	15–70
Hand-held PC (μ)	100	75–80	30–50	15–25
PDA (μ)	70–100	50–100	50–75	30–50
Cellular Phones (μ)	75	75	50	10
Audiovisual Equipment (μ)	75–100	60–100	50–75	25–50

by 2005. The main categories of this PWB market include multilayer paper and composite, flexible-circuit, rigid glass and microvia. Until recently, the PWB technology was viewed as a system-level board for mounting system-level components. That has changed dramatically in recent years.

Accelerating adoption of chip size and ball grid array packages in portable electronics, combined with the transition to flip-chip IC connection in packages for high-performance microprocessors and ASICs, has spawned enormous investment in capacity to manufacture high-density printed circuit substrates based on new fabrication processes, materials, and process equipment. What was only a novelty a few years ago, now accounts for more than 6% of the value of PWBs produced worldwide, and by 2002 it is expected that IC packages and conventional microvia multilayers will account for almost 17% of the total market as shown in Figure 16.31.

16.11 SUMMARY AND FUTURE TRENDS

The progress over the past few decades has been developing the sophistication of features for various parameters that are typical of the computer industry. However, there is a big gap between IC pad density and PWB pad density, and until this gap is closed, intermediate packaging called IC packaging is necessary.

The microvia technology has come a long way, and paved the way to close the gap, to some extent, but much remains to be done. Figure 16.32 indicates the PWB roadmap that needs to be followed to close the gap. Until then, wafer-level packaging will not be a true wafer-level packaging; it will only be wafer-level packaging that is described in Chapter 10, as redistribution type which redistributes the IC I/Os that are on 100–150 micron pitch to those manufacturable PWB pitches.

Figure 16.33 illustrates the line widths and spaces required to form a chip scale package made of build-up technology. This figure assumes the I/O pitch to range from 1.0–0.25 mm.

Table 16.6 indicates the PWB board technology needs for various systems. Some of these require board lines and spaces of the order of 5–10 micron lines spaced on 20–30 micron centers.

16.12 HOMEWORK PROBLEMS

1. Describe the anatomy of a copper clad laminate.
2. Describe how woven glass cloth is made.

Samsung, EX1028, p. 319

3. How does the resin bond to the inorganic glass substrate?
4. What are fillers and how do they improve the laminate property such as thermal dissipation?
5. What is the limitation of FR-4 grade laminate?
6. What electrical, thermal and mechanical properties are important in a material suited for electrical design?
7. Enunciate additive process technology and subtractive process technology for PCBs.
8. What are the functions of a solder mask?
9. Why is copper protected with tin or tin/lead? What are the alternatives?
10. If line widths are 6 microns and the spaces are 12 microns, how many lines per channel would that PWB be? Assume the channel to be 2.54 mm.
11. Write a process flow chart for the manufacture of a single-sided PCB.
12. If the chip I/O pitch is 100 microns, what is the wiring density required to support that chip for direct chip attachment to the PWB?
13. If the build-up PWB has 10 micron lines and spaces, and has four such layers, what is its wiring density in cm/cm^2?
14. What does FR-4 laminate denote? Name the filler and the resin in FR-4.
15. Draw a flowchart for the manufacture of a multilayer rigid PWB.
16. Differentiate between electroless and electroplating, in a simple manner.
17. What are the factors that need control during etching process to obtain a high-quality etched board?
18. As a designer, what are the important characteristics you will look for in the selection of a laminate?

16.13 SUGGESTED READING

1. W. C. Bosshart. *Printed Circuit Boards—Design and Technology*. New York: McGraw-Hill, 1983.
2. W. D. Brown, ed. "Computer-Aided Engineering and Design." *Advanced Electronic Packaging*. New Jersey: IEEE Press, 1999.
3. R. H. Clark. *Handbook of Printed Circuit Manufacturing*. New York: Van Nostrand Reinhold, 1985.
4. R. H. Clark. *Printed Circuit Engineering—Optimizing for Manufacturability*. New York: Van Nostrand Reinhold, 1989.
5. C. Coombs, Jr. *Printed Circuits Handbook*, Fourth Edition. New York: McGraw-Hill, 1995.
6. G. Geragosian. *Printed Circuit Fundamentals*. Virginia: Reston Publishing Company, 1985.
7. G. L. Ginsberg. *Printed Circuits Design*. New York: McGraw Hill, 1991.
8. D. M. Goulet, Ed. *Bare Board Drilling*. San Francisco: Miller Freeman Inc., 1992.
9. C. A. Harper. *High Performance Printed Circuit Boards*. New York: McGraw-Hill, 2000.
10. H. Holden. "Introduction to High Density Interconnection Substrates." *The Board Authority, A Supplement to Circuit Tree*, Vol. 1, No. 2, 1999, 6–10.
11. T. L. Landers, W. D. Brown, E. W. Fant, E. M. Malstrom, and N. M. Schmitt. *Electronics Manufacturing Processes*. New Jersey: Prentice Hall, 1994.
12. "Laser Drilling Techniques." *Application Notes*. Portland: Electro Scientific Industries, Inc., 1996.
13. J. A. Scarlett. *Printed Circuit Boards for Microelectronics*, Second Edition. Great Britain: Electrochemical Publications, 1980.
14. C. J. Simon. *Computer Aided Design of Printed Circuits*. North Canton: Abbot, Foster and Hauserman Co., 1987.

Samsung, EX1028, p. 320

15. R. R. Tummala, et al. *Microelectronics Packaging Handbook*, Part III. New York: Chapman and Hall, 1997.

ACKNOWLEDGMENT

The editor gratefully acknowledges the contributions of Mr. Rolf Funer for critically reviewing this chapter.

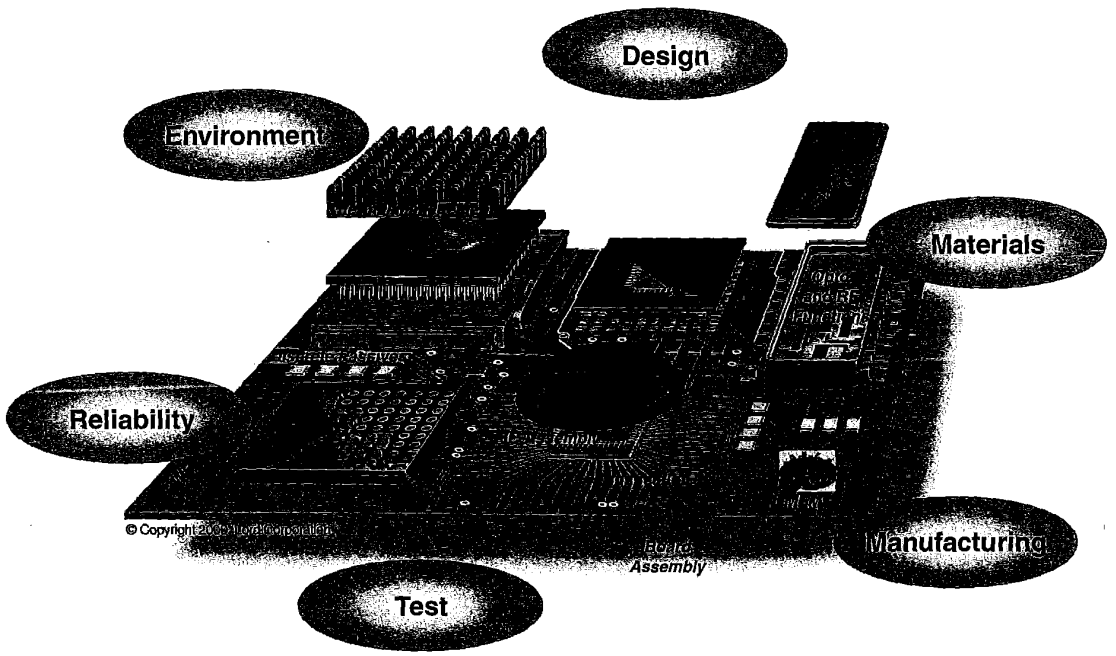
Samsung, EX1028, p. 321

CHAPTER 17

FUNDAMENTALS OF BOARD ASSEMBLY

Fredrik Mattsson
Ericsson, Sweden

Stefan Kandelid
Ericsson, Sweden



17.1 What Is a Printed Wiring Board Assembly?
17.2 Surface Mount Technology
17.3 Through-Hole Assembly
17.4 Generic Assembly Issues
17.5 Process Control
17.6 Design Challenges
17.7 Summary and Future Trends
17.8 Homework Problems
17.9 Suggested Reading

CHAPTER OBJECTIVES

- Understanding the assembly process, which typically involves mounting components on a wiring board and soldering their leads to that board
- Comparison of surface mount assembly and through-hole assembly
- Describe generic assembly issues
- Appreciate the influence of assembly technique on the size, manufacturing cost, performance and reliability of the product

CHAPTER INTRODUCTION

Printed wiring board assembly is the process of building functional electronic systems from individual electrical components. The interconnections created in this process are sometimes referred to as the second level of interconnection, where the chip to package connection would be the first, and the board to system would be the third. The assembly process typically involves mounting components on a wiring board and soldering their leads to that board.

Samsung, EX1028, p. 323

17.1 WHAT IS A PRINTED WIRING BOARD ASSEMBLY?

Printed wiring board assembly is the process of building functional electronic systems from individual electrical components. The assembly process typically involves mounting components on a wiring board, as illustrated in Figure 17.1, and soldering their leads to that board. The finished electronic system could be anything from a mainframe computer to a greeting card playing "Happy Birthday."

The PWB serves several purposes. Apart from providing electrical connections between the different components, the board also allows the heat generated by the components to be transferred to the board, where it can be dissipated. The use of a printed wiring board as the electrical and mechanical platform for assembly has been a successful concept since the introduction of semiconductors in the 1950s.

The most significant step in the evolution of the PWB assembly process is the change from through-hole assembly to *surface mount assembly* (SMA). Both methods are still widely used, but the use of SMA has grown rapidly in the past decade, which is shown in Figure 17.2. By the late 1990s, SMA accounted for more than 80% of electronic manufacturing.

The main reason for this change is size reduction. Both the size reduction of the component, and the real estate used on and inside the board, are greatly reduced. A hole used for through-hole assembly goes all the way through the PWB, using valuable board area in every layer of the board. All electrical traces have to be routed around the holes at a safe distance. Another important factor is that automated assembly has become simplified with SMA.

Through-hole assembly is achieved by inserting the leads of the components into plated holes in the board. While using surface mount assembly, the components are placed on the surface of the board. This fundamental difference, depicted in Figure 17.3, has an impact on virtually every step of the assembly process.

17.2 SURFACE MOUNT TECHNOLOGY

The abbreviations SMA, SMD and SMT are often used somewhat carelessly. There is, however, some difference in what they mean. SMA is the surface mount *assembly,* SMD is the *device* (or component), and SMT is the entire *technology.* PWB is the common abbreviation for *printed wiring board.* PBA is used for *printed wiring board assembly,* meaning the printed wiring board populated with components.

Surface mount assembly is done by stencil-printing solder paste to a board, placing the component on the board and then heating the entire assembly, so that the solder melts and forms solder joints. The flux is already mixed in with the solder paste, and apart from reducing metal oxides, it provides necessary stickiness to hold the components in their correct positions until the solder joints are created.

It is also possible to glue the surface mount components to the board and solder them by wave soldering, but this technique is mostly used when there are both surface mount and through-hole components on the boards. This method is further described in Section 17.3.5 under Mixed Component Assembly.

The main process steps of surface mount assembly are solder-paste printing, assembly of components and reflow soldering. Figure 17.4 shows a typical surface mount assembly line. The component assembly is divided into assembly of small passive devices and

Samsung, EX1028, p. 324

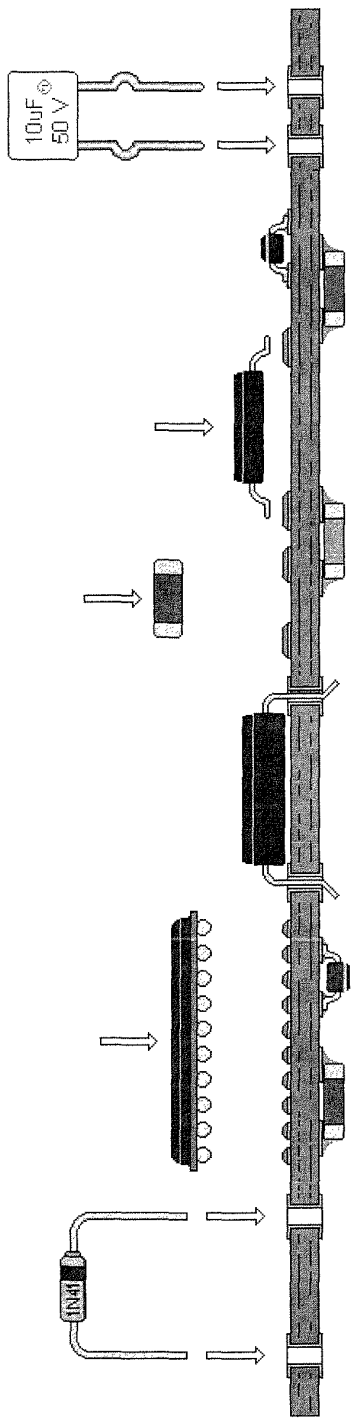


FIGURE 17.1 Printed wiring board assembly.

Samsung, EX1028, p. 325

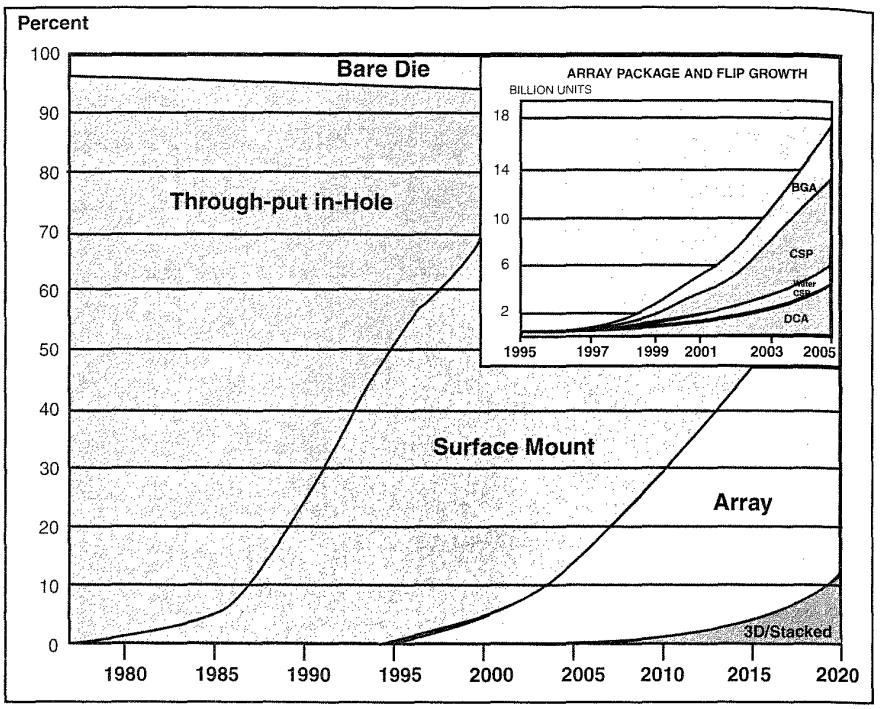
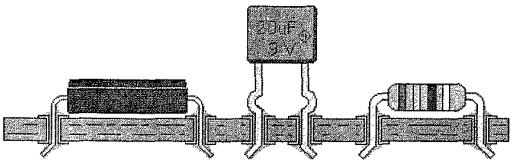
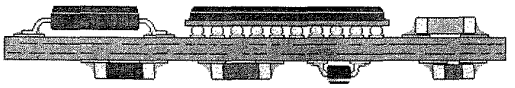


FIGURE 17.2 Growth of SMT and array assembly technologies. (*Courtesy of Prismark*)

FIGURE 17.3 Through-hole assembly versus SMA.



Through-hole Assembly



Surface Mount Assembly

Samsung, EX1028, p. 326

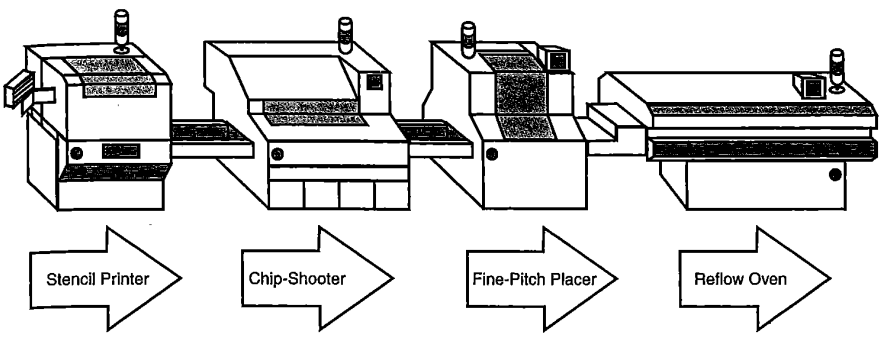


FIGURE 17.4 Example of surface mount assembly line.

larger components that require better placement accuracy, or are difficult to handle. Automatic optical inspection is sometimes performed as a part of the process.

Balancing the assembly line to maximize throughput is an important issue. This means that each machine should spend equal time processing the board, thus eliminating bottlenecks. The most common way to balance an SMD-line is to move the components to be assembled from one machine to another. Usually, mid-range components that can be assembled in different machines are moved.

17.2.1 Solder-Paste Printing

Screen-printing is an old technique, which has been used since the 6th century, mostly for printing paint on paper or cloth. For screen-printing, a screen or mesh is used. The screen has openings through which the paint is pressed and transferred.

Such a screen cannot be used for solder-paste printing of fine patterns in the electronic industry. Instead, stencils are more commonly used. A stencil is a thin sheet of metal; for example, brass, which has a pattern of holes through which the solder paste is pushed by a squeegee onto the desired places on the board. A stencil allows smaller openings than a screen, and it is far more accurate. Solder-paste printing is the most critical step to control in the entire surface mount process. It is also the process step where the most errors occur.

Solder-paste printing begins with the board being fixed by vacuum or by mechanical means to a flat plate, then the stencil is positioned on top of it. Solder paste is added to the stencil, and if the printing is done in an automatic machine, the stencil is then aligned using fiducials on the topside of the board and the bottom-side of the stencil.

After the alignment is done, the squeegee moves across the stencil pushing the solder paste in front of it, creating a downward and forward pressure. The paste starts to roll and fills the apertures (openings) in the stencil as they are passed. The solder paste is supposed to be completely wiped off, leaving the top surface of the stencil clean behind the squeegee. After printing, the board is lowered from the stencil, with the solder paste deposited on the soldering surfaces and then sent to next step in the process. The remaining solder paste is left on the stencil to be used on the next board.

Squeegee pressure is important and must be correctly set. If it is set too high, the stencil and squeegee blade will be worn down quickly and paste will be scooped up from

Samsung, EX1028, p. 327

the apertures. If it is set too low, the paste will not roll, the apertures will not be properly filled, and the stencil will not be wiped clean. The speed of the squeegee motion must also be correctly set in order to get good results. The use of proper support underneath the board during the print operation is vital; without that, it is impossible to get good printing quality.

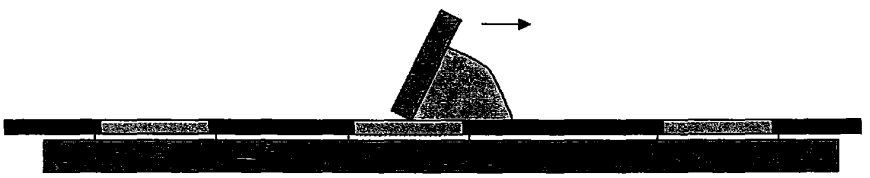
Stencil printing is typically done one of two ways: contact printing or snap-off printing. Contact printing means that the entire stencil is in contact with the boards during the print. In snap-off printing, a small distance is kept, allowing only the part of the stencil that is under the squeegee to touch the board. The difference is illustrated in Figure 17.5.

Some printers are equipped with automatic inspection systems to inspect the quality of the print. The data collected can be used to adjust parameters for improved performance. There are systems for both 2D and 3D inspection. The 3D alternative gives additional information about the height of the deposited solder paste. The disadvantage of 3D inspection is increased cycle time. The automatic inspection systems usually do not allow closed loop control of the print parameters.

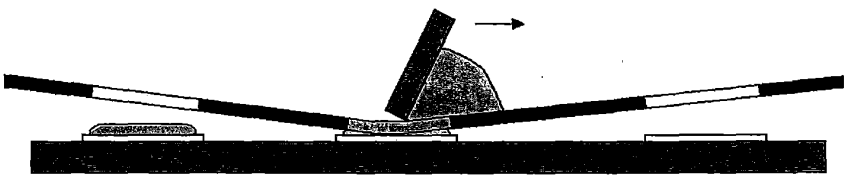
Stencils

Stencils are usually made of nickel, nickel-plated brass, stainless steel or plastic. Even copper or plain brass has been used. The thickness ranges from 4–8 mil (0.1–0.2 mm). Sometimes a stencil is made to have different thicknesses in different areas. The reason is that some components need more (or less) paste than others. The technique is called step down, step-etching, or simply stepped stencils.

In the past, chemical etching was the common method of producing stencils. Recently, however, demands for better accuracy, paste release and wear resistance have made laser cutting, or even nickel additive build-up techniques, more common. Nickel additive, build-up or electroformed stencils are different names for stencils that are manufactured



Contact printing



Snap-off printing

FIGURE 17.5 Contact printing versus snap-off printing.

Samsung, EX1028, p. 328

by electroforming. First, a mask of photoresist is created. After developing, a negative image remains which corresponds to the openings in the stencil. The stencil is then grown on the mask by nickel electroforming. The current and time determine the thickness of the stencil. The last step is to remove the mask from the apertures.

Chemically-etched stencils can have an "hourglass" shape to the aperture opening; this inhibits good paste release and decreases the solder-paste transfer rate. Both the use of laser cutting and additive technique give the desired tapered opening, which improves the paste release from the stencil. The additive, or electroformed stencils, also have a very smooth aperture wall, which further improves paste release. However, the slightly rougher laser cut surface can be electro-polished to improve its performance. Figure 17.6 shows aperture openings in stencils from different manufacturing methods.

Squeegees come with either metal or rubber blades. It may be easier to wipe the stencil clean with a rubber squeegee, but the rubber wears down faster and can scoop up paste from large apertures. Rubber squeegees come in several different blade geometries and hardnesses. A metal squeegee wears down the stencil faster, and it must be handled carefully, in order not to damage the squeegee edge. Metal squeegees also come in different materials and qualities. Regardless of the blade material, the length of the squeegee should be adapted to the printed width of the board.

In the last few years, new inventions have surfaced in the area of solder-paste printing. One novel idea is to replace the entire squeegee with a paste-filled chamber. This chamber would be pressurized and would move across the stencil, filling the apertures. The main benefit is that the pressure would be focused in Z-direction, which should help fill the apertures more effectively. Most modern automatic printers are equipped with wiping systems to keep the bottom-side of the stencil clean. The problem created, when the stencil's bottom-side contaminates the board with paste, is called smearing. Smearing can cause process defects such as solder balls and short circuits.

Another typical printing defect is called slumping and might lead to short circuiting. Slumping occurs when the paste printed on the pads is unable to hold its shape and

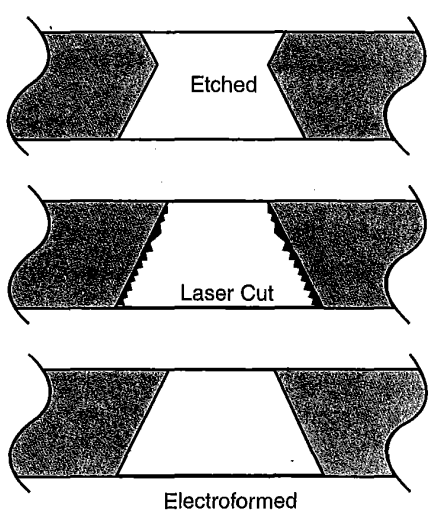


FIGURE 17.6 Stencils from different manufacturing methods.

Samsung, EX1028, p. 329

instead collapses and flows out onto surrounding surfaces. The effect becomes more significant as the pitch decreases. Slumping occurs more easily at elevated temperatures. The temperature and humidity where the printing is performed have a large impact on the printing result, and consequently on the assembly-yield.

Solder Paste

Solder paste consists of small solder spheres, flux and solvent. It is a thick gray paste with a viscosity close to that of toothpaste. The solder spheres are usually of the eutectic tin/lead alloy (63% tin and 37% lead). The pastes are separated into types 1, 2, 3, . . . etc., as shown in Table 17.1. The higher numbers have smaller solder spheres and are more suitable for fine-pitch devices. Solder paste for stencil printing has a metal content of around 90 weight-percent (wt%). Note that the flux has a much lower density than the solder spheres, so the volume percentage of solder in the paste is only about 50%. The solder volume after reflow is therefore much less than the paste volume printed. Paste for dispensing has a lower metal content to make it flow better through a nozzle.

A minimum amount of paste is required on the stencil to keep the printing results consistent. As the solder paste is consumed, a dispenser inside the machine can add more automatically. This way, the machine can run continuously without operator interventions. However, if the paste is left on the stencil during longer stops, the viscosity changes due to the evaporation of the solvent, eventually leaving the paste dry and useless. The viscosity of the paste will increase when the printer is idle and decrease when in use. Viscosity is also temperature dependent; it decreases with rising temperature. The paste on printed boards that are not assembled soon loses tackiness, and components may fall off before reflow soldering. Solder paste contains unhealthy materials, and the use of proper ventilation and protective gloves are absolutely necessary. Used solder paste must be handled appropriately and should be sent for recycling.

17.2.2 Pick and Place

Component assembly is almost always automated. The equipment used are commonly called pick-and-place machines and comes in several different varieties. Component assembly starts with the board being transported into the machine on a conveyor, and fixed in its position. Next, a camera is moved in over the board to find the marks used for optical alignment. These marks, as shown in Figure 17.7, are commonly called fiducials.

TABLE 17.1 Solder-paste types.

Type	Solder Sphere Diameter	
	μm	mil
1	75–150	3–6
2	45–75	1.8–3
3	20–45	0.8–1.8
4	20–36	0.8–1.4
5	15–25	0.6–1
6	5–15	0.2–0.6

Samsung, EX1028, p. 330

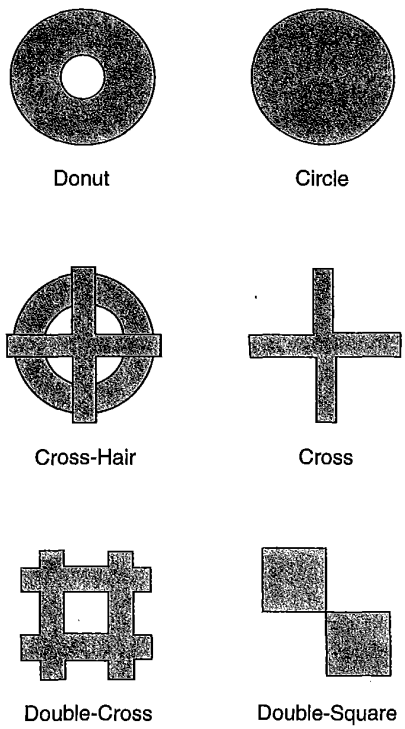


FIGURE 17.7 Board fiducials.

The accuracy of the mechanical alignment of the board is not good enough for assembly of today's small or fine-pitch components, and that is why optical alignment is used. The placement machine is programmed to know approximately where to look for the fiducials, and when it reaches that position, it begins to search for an expected image. Once all fiducials are found, the machine can calculate X, Y, and theta (θ) offsets. Even stretching or shrinkage of the board can be compensated for. A board should have two or three fiducials widely separated in X and Y directions, to allow optimal alignment. After this step is completed, a vacuum nozzle is used to lift the component up from its feeder and transport it to the correct position on the PWB. Since the pick-up position is not very accurate compared to what is required in placement accuracy, there is a need for centering after pick-up. Centering means aligning the center of the component with the center of the nozzle. Note that the accuracy of the placement machine is the ability to position the nozzle at the correct coordinates, and getting the component in the right position includes identifying the correct center of the component; this can be done mechanically or optically. Mechanical centering or alignment means that the component is actually pushed into the correct position on the nozzle with the help of centering fingers.

Optical alignment means that the component is presented to a camera, and from the image retrieved, the center of the component can be calculated. The component is not moved on the nozzle. The distance between the correct component center and the actual nozzle position on the component is then added to the placement coordinates. All of this

Samsung, EX1028, p. 331

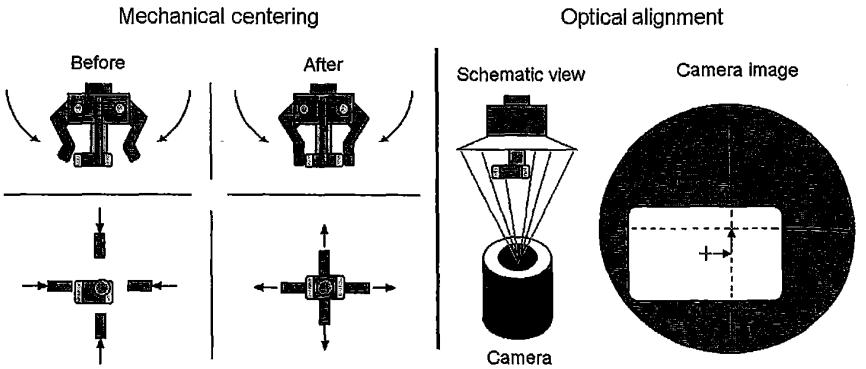


FIGURE 17.8 Mechanical centering versus optical alignment.

is done automatically while moving the component between pick-up and placement positions. These two methods are compared in Figure 17.8.

Many machines also use this information to "learn" where the correct pick-up position is. They use a mean value of the last few pick-up corrections to calculate what the right coordinates are, thus minimizing the centering correction. A mechanical gripper can also be used instead of a vacuum nozzle, but that technique is only used when the necessary vacuum level cannot be achieved due to the lack of a flat enough area from which to pick up.

Machines used for smaller components, like resistors, capacitors, or inductors, are called *chip-shooters.* They are very fast and have moderate placement accuracy. Assembly machines for fine-pitch and larger-sized components, have better accuracy and more flexibility, but are on the other hand slower. Because of the lower speed, it is more expensive to assemble a component in a fine-pitch placer than in a chip-shooter. Chip-shooters can be divided into a few separate types, depending on their basic component-handling concept. The types described here are called *turret head chip-shooter, revolver head chip-shooter* and *multimodule chip-shooter.*

The turret head chip-shooter (Figure 17.9) has a fixed component placement position. Instead, the board moves on an XY-table. The turret head is a horizontally rotating drum, which has a number of nozzles around the periphery. As the turret rotates, the nozzles

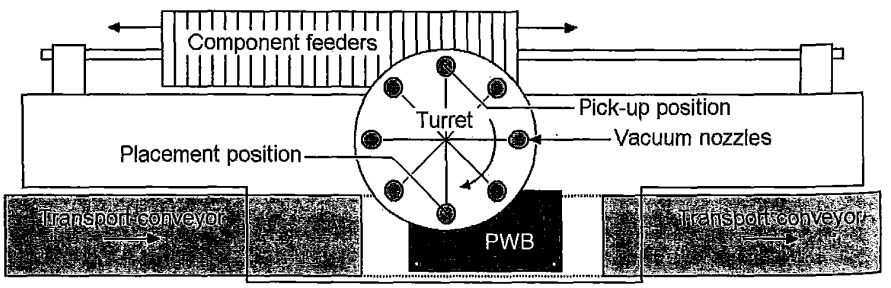


FIGURE 17.9 The turret head chip-shooter.

Samsung, EX1028, p. 332

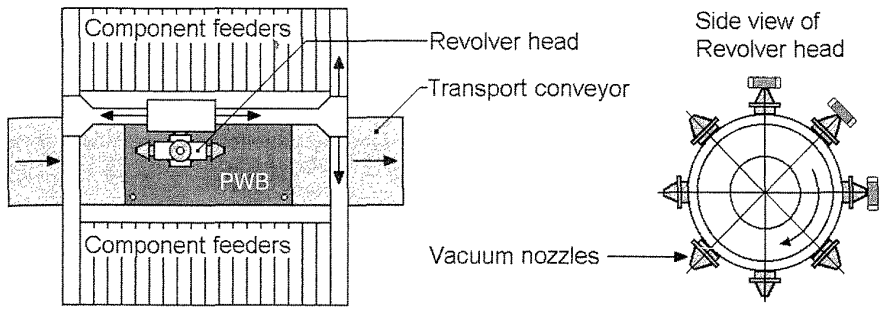


FIGURE 17.10 The revolver head chip-shooter.

move between different stations. The cycle starts at a pick-up station, moves to centering, component rotation, placement, and so on.

The component feeders are mounted on a moving carriage, which transports the next component to be picked up to the nozzle at the pick-up station. Fast moving XY-tables present a problem with heavy components due to G-forces, since the components are only held in place by the tackiness of the solder paste.

The revolver head chip-shooter (Figure 17.10) has a number of nozzles pointing outwards, mounted on a vertical wheel. The wheel is carried by a Cartesian robot between the component feeders, where it picks up a number of components, and the board where it mounts them. Centering, component rotation, etc., take place on the revolver head while moving. Neither the printed wiring board nor the feeders move during the assembly.

The multimodule chip-shooter (Figure 17.11) has several small, pick-and-place modules, densely packed together to form one machine. While most assembly machines work more or less sequentially, this type utilizes a number of robots in parallel. Each module may not work very fast, but since they are all moving simultaneously, the number of components placed per hour can be very high. The feeders do not move during the assembly. Since each pick-and-place module has a limited work area and only can reach a few feeders and a small part of the PWB at a time, the PWB is moved forward in steps, so that all parts of the board can be reached.

Fine-pitch placement machines are usually variations of Cartesian robots. The board is fixed in its position while the placement heads are moving, thus enabling heavy com-

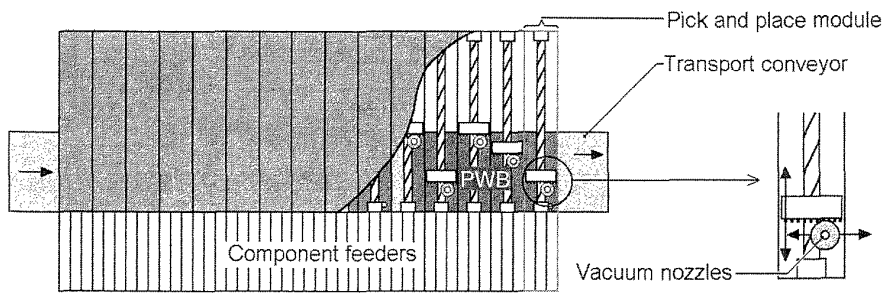


FIGURE 17.11 The multimodule chip-shooter.

Samsung, EX1028, p. 333

ponents to be mounted. To increase throughput without jeopardizing the placement accuracy, multiple nozzles, dual assembly heads and dual lanes can be utilized. Dual lanes are when two parallel conveyors are used for board transport. Optical alignment can be time consuming if the robot has to move the component to a separate alignment camera. Therefore, many machines use "vision on flight," which means that centering takes place during movement from the pick-up to the placement position, thus eliminating the transport to a separate alignment station. A separate camera offers other benefits, however, like higher accuracy or larger field of view. The lighting of the components during optical alignment is critical in order to acquire a good image for the computer to analyze. It should therefore be possible to vary the intensity and the angle of the light in many different ways. A fine-pitch placer has many different nozzles to choose from, enabling it to assemble many different types of components. The machine can change nozzles by itself, by moving to a nozzle station to pick up and return the different tools. The flexibility and accuracy of a fine-pitch placement machine is often very high, while the placement speed is less than that of a chip-shooter.

Coplanarity, illustrated in Figure 17.12, can be defined as the variation in height of component leads. If some leads are bent too much in any direction, they will not become soldered to the board. The coplanarity must be less than the solder-paste height for all leads to be soldered. Most fine-pitch placers have an optional coplanarity measuring station to remove components with poor coplanarity before placement.

Feeders for pick and place machines can be sorted into a few main groups.

Tape-on-reel feeders (Figure 17.13) are by far the most common and popular type. They feed components delivered on paper or plastic blister tape. In both cases, the component is placed in a pocket in the tape and covered with a plastic film, a top foil. The top foil is automatically peeled off right before the component is picked up. The tape width ranges from 8 mm up to 56 mm and sometimes even wider. Chip-shooters usually handle tapes up to about 16 mm, but there is no fixed limit. Tape-on-reel feeders are driven electrically, pneumatically or mechanically. In the latter case, a pusher in the placement machine provides the motion.

Tray feeders (Figure 17.14) usually have a separate unit on the fine-pitch placer to handle them. Usually, the tray feeder looks like a cabinet with several different shelves. Each shelf holds one or more trays with components. The tray-feeding unit transfers one component at a time and presents it to the placement machine. Tray feeders are only used for fine-pitch placement machines. The types of components fed in trays might be PLCC, CLCC, QFP, TSOP or any odd-shape component that is difficult or expensive to supply in tape.

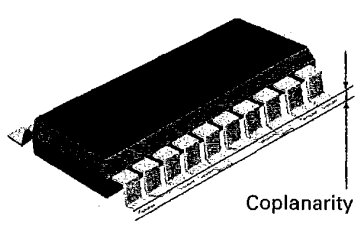


FIGURE 17.12 Definition of coplanarity.

Samsung, EX1028, p. 334

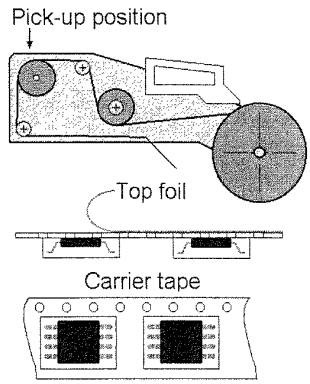


FIGURE 17.13 Tape-on-reel feeder.

Gel-pack is a sticky kind of tray, sometimes used for flip chip. The components are held in position, without being damaged, by a tacky gel.

Bulk-feeders are sometimes used for small passive components that can be delivered simply poured in a box. This delivery form is potentially the cheapest. The feeder orients the component automatically and presents it to the machine. Bulk-feeders can be used in chip-shooters.

Sticks are plastic tubes with a cross section that is designed to fit a certain component type. The sticks are usually tilted slightly forward when set up in the machine, and the components are moved forward by vibration and gravity. Sticks are only used in fine-pitch placers and mostly for low volume production. Sticks can sometimes be useful for programmable ICs, since it is easier to repackage components in them after programming. Sometimes, several sticks can be set up in the same stick feeder to make replacement less frequent. A schematic stick feeder is shown in Figure 17.15.

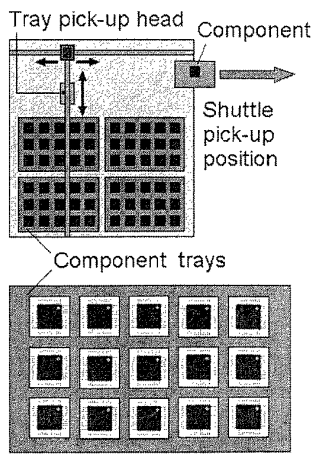


FIGURE 17.14 Tray-feeding unit.

Samsung, EX1028, p. 335

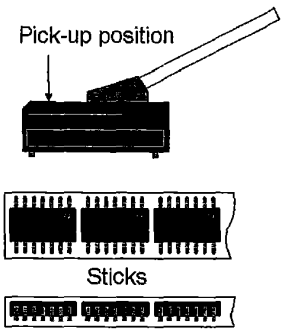


FIGURE 17.15 Stick feeder.

17.2.3 Soldering

The purpose of soldering is to get an electrical, mechanical and thermal connection between the components and the board. The basic materials needed to create a good solder joint are solder and flux.

Solder Joint Quality

All solder connections must indicate good wetting and adherence to the pads on the board. The amount of solder attaching the component to the board has great impact on the reliability of the solder joint. The shape of the joint is important since much of the strength comes from the fillet. Figure 17.16 shows the minimum and maximum amount of solder required for getting proper strength. If not enough heat is applied when soldering, the joint does not get the right intermetallic build-up. A good solder joint normally looks shiny and smooth; a poor joint looks dull and grainy.

17.2.4 Flux

The purpose of the flux is to reduce the oxides on the surfaces to be soldered. The flux activates first at higher temperatures; 150–160°C is a normal temperature range for ac-

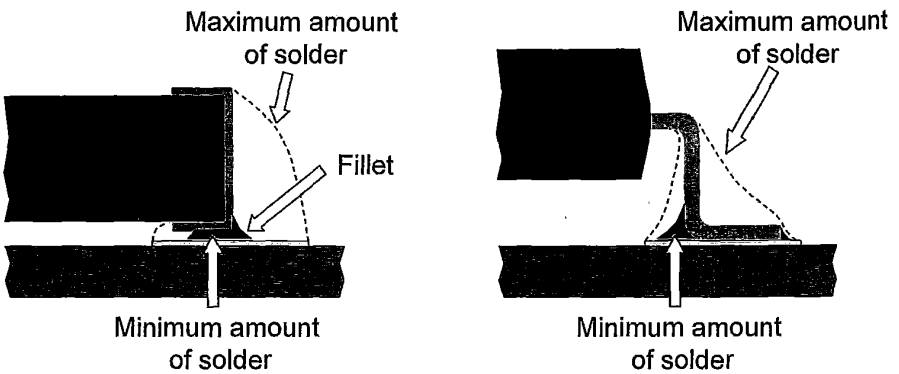


FIGURE 17.16 The recommended minimum and maximum amount of solder required for good reliability.

Samsung, EX1028, p. 336

tivation. When choosing which flux to use, one must consider the melting temperature of the solder, the surface finish on the component lead and on the solder pad, together with the activity of the flux. Traditionally, fluxes have been classified according to their activity level:

1. R (*rosin*): on-activated fluxes (no activators).
2. RA (*rosin activated*): The most active flux and also the most corrosive.
3. RMA (*rosin mildly activated*): mildly activated flux requires good solderability on solder pads and components.

Today there is also a more detailed classification done by International Printed Circuit (IPC) Association, J-STD-004.

Resin is, in the context of fluxes, the synthetic version of rosin with similar performance. The resin can be formulated to exact desired features, while rosin can have some lot-to-lot variations. When the ozone-depleting *chlorofluorocarbons* (CFCs) and freons, used in SMT cleaning, were prohibited due to environmental reasons, the RMA fluxes were improved to make what are called no-clean fluxes. The conversion to no-clean fluxes meant significant cost and timesavings in high volume production, since the cleaning step could be eliminated. The predominant flux systems today are: No-clean flux, which does not normally require any washing after soldering; and water-soluble flux, which does require washing after soldering.

17.2.5 Solder Alloys

Metal alloys used for solder paste include combinations of tin (Sn), lead (Pb), silver (Ag) and indium (In). The most commonly used alloy is tin-lead (63%Sn/37%Pb) or tin-lead-silver (62%Sn/36%Pb/2%Ag). These alloys are used for 90% of all surface mount applications; the reasons are their low cost, low melting temperature, proven reliability and ability to join common assembly metals. The 63/37 composition can be understood using the Sn/Pb phase diagram, as illustrated in Figure 17.17. The 63Sn/37Pb is a eutectic alloy, meaning that the transition between the liquid state and the solid state occurs over a very narrow temperature range.

The eutectic solder (63%Sn/37%Pb) has no plastic (pasty) phase between its solid and liquid states. When choosing flux and solder alloy, one must consider the finish on the surfaces to be soldered. Tin-lead solder is the most common surface finish on solder pads. When producing the board, it is dipped in a solder bath to apply the solder to the bare conductors; this process is called *hot-air solder leveling* (HASL). The process results in some height variations and hemispherical surfaces on the solder pads. For boards with decreased pitch, like in modern portable electronics, it is necessary to improve the flatness of the pads. Two methods of reaching that goal are *organic solderability preservative* (OSP) and nickel-gold (Ni-Au) application. These are both applied chemically and result in flat surfaces with little height distribution.

17.2.6 Reflow Soldering

Reflow soldering is a process where the solder is already present on the solder pads before the soldering starts. The solder paste used in this process is a homogeneous mixture of metal particles (solder powder) in a flux medium. Solvent, an activator and a

Samsung, EX1028, p. 337

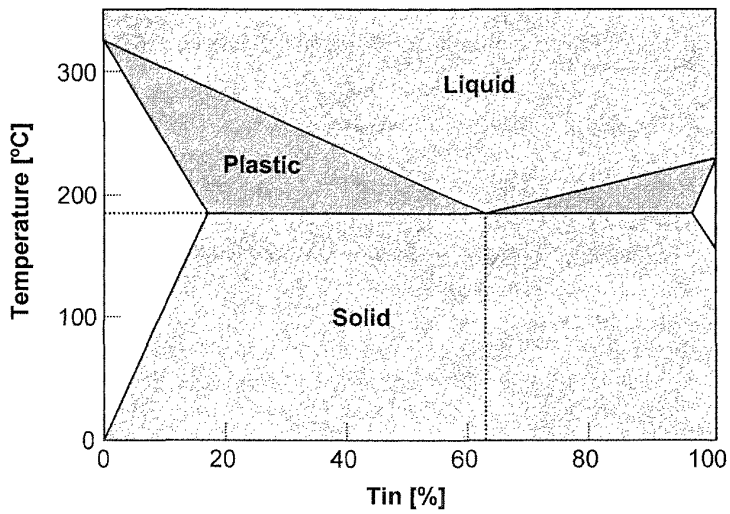


FIGURE 17.17 Phase diagram for the tin-lead alloy.

viscosity-regulating agent are usually included in the solder paste. The solder paste is screen- or stencil-printed onto the solder pads, followed by the placement of the components and reflow soldering. The increasing temperature in the oven activates the flux. It must then keep the surfaces deoxidized until the wetting of the solder.

Normally the SMA soldering process is achieved using a conveyorized reflow oven, as depicted in Figure 17.18. There are three different types of ovens that are commonly used: infrared (IR), forced-convection and vapor-phase soldering ovens. The first two processes have some similarities, but the forced convection offers far better heat distribution in the board assembly, since the heat transfer is independent of component color and surface structure. In the IR oven, the assembly is directly heated by infrared radiation,

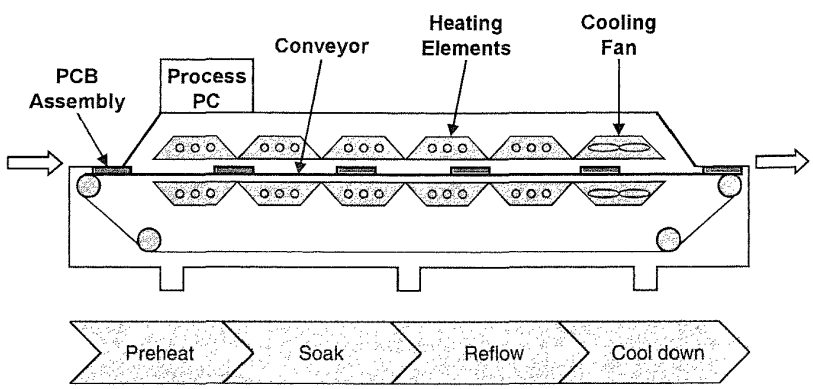


FIGURE 17.18 Reflow oven with five heating zones and one cooling zone.

Samsung, EX1028, p. 338

but in the forced-convection oven, the heat transfer is accomplished by circulating hot gas. Some ovens combine the two methods.

The reflow profile can normally be characterized with four different phases: preheat, thermal soak, reflow, and cool down (Figure 17.19). The reflow profile has direct impact on manufacturing yield, solder joint integrity and the reliability of the assembly. A typical reflow oven has five to nine heating zones and one to two cooling zones, so that the temperature can be easily adjusted. The objective of the preheat phase is to increase the temperature at a controlled rate to minimize any thermal damage to the components and the board. The ramp rate is typically 2–4°C/s. If the temperature is increased too rapidly, the solvent in the solder paste is not allowed to evaporate slowly enough. The solvent in the paste might start boiling, causing solder balls and short circuits between leads.

The main purpose of the soak phase is to equalize the temperature of all surfaces being soldered. The flux covers the surface of the solder powder and begins to wet the component leads and the solder pads on the board in the 150–180°C range. The flux reduces metal oxides, exposing the clean solderable metal.

In the reflow phase, the eutectic Sn-Pb solder reaches its liquidus temperature at 183°C. Rapid heating is used to ensure that the entire assembly quickly reaches a temperature above the melting point of the solder. The solder begins to wet the solder pads and component leads and intermetallic layers start to form. Thirty to ninety seconds is normally recommended to form a good electrical and mechanical connection. The peak temperature is normally 215–225°C in the eutectic Sn-Pb solder. Being at too high a temperature for too long, creates excessive intermetallic layers that can have negative influence with regard to solder joint reliability.

The cool down phase allows the solder to solidify before exiting the reflow oven. It is very important to keep this part of the process under strict control. If the board is cooled too slowly, larger grains will form in the solder, which may result in fragile connections. Cooling too fast, on the other hand, might cause other problems.

The principle behind vapor phase soldering is that a certain fluid has a specific boiling temperature. The fluid is formulated so that its boiling temperature is the same as the temperature at which the reflow is preferably taking place. The bath of fluid is heated

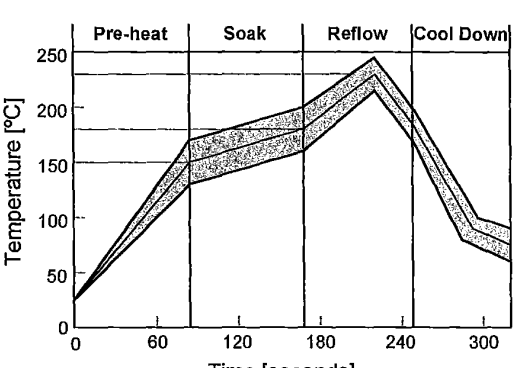


FIGURE 17.19 Typical reflow profile for eutectic solder.

Samsung, EX1028, p. 339

up until it starts to vaporize and then the board, with the components placed on it, is moved through the vapor. The vapor is heavier than air and is also held in place by cooling coils at the top of the machine. The whole assembly is heated up, the solder is reflowed and the solder joints are formed. As the heat is transferred from the vapor to the assembly, the vapor turns back into fluid and drops back into the fluid bath again.

These three processes can all be run in an inert environment. The advantage of running the reflow in inert gas, compared to air, is that no reoxidation occurs inside the oven. Another advantage is a stronger wetting force in an inert environment. The use of inert gas gives a wider process window; this is helpful if there are variations in component or board solderability.

Vapor phase soldering is always performed in an inert environment since no air is present. Infrared and convection soldering can be performed both in air and in an inert environment. Nitrogen, N_2, is the inert gas that is used most often to achieve the desired environment. The benefits with inert soldering (explained earlier) must be weighed against the disadvantage of the added cost for nitrogen.

When components are placed on the bottom-side of the board and are soldered a second time, the solder is melted on both the top and the bottom-side. The components on the bottom-side are then hanging upside down, kept in place only by the surface tension of the molten solder. Heavy components, or components with leads to only one side, might then fall off in the reflow oven. Vibrations in the conveyor make it even more difficult for the components to stay on the board. When there is risk for such components to fall off in the second reflow, the components can be kept in place using adhesive. After solder-paste printing, but before component placement, the adhesive is dispensed on the board. The heat in the reflow oven cures the adhesive and the component stays in place even if the component is reflowed upside down a second time.

Due to thermal mass differences between components and the design of the board, it is usually inevitable to get an uneven temperature distribution over the board when using IR or forced-convection ovens. Vapor-phase soldering is better in this regard, and forced convection is much better than pure IR soldering. It is therefore necessary to perform temperature measurements at different locations on the board to make sure that all solder joints follow the specified thermal profile. The result might otherwise be cold joints or overheated joints and components. These may result in lowered manufacturing yields or influence the reliability in a negative way. The measurement is carried out using thermocouples that are glued in place or soldered with high-temperature solder onto an assembled board. Special equipment is used to collect the data when the board is sent through the oven. The set temperature for the oven is altered and the procedure is repeated until the measured temperature is within the specified process window. The set temperature for the oven is normally much higher than the temperature in the solder joints.

Some of the most common defects in the reflow process are:

- Unsoldered joints, due to component coplanarity (lifted leads) or board warpage
- Cold joints, due to mass differences between components or bad thermal design of the board
- Solder balls, due to too rapid out-gassing of the solvent in the solder paste or bad design of the solder pads
- Bad wetting, due to bad solderability of boards or components

Samsung, EX1028, p. 340

TABLE 17.2 IPC/JEDEC J-STD-020A, component floor life (time before unsoldered components must be dried-out using baking).

Level	Time	Conditions
1	Unlimited	≦ 30°C/85% RH
2	1 Year	≦ 30°C/60% RH
3	168 Hours	≦ 30°C/60% RH
4	72 Hours	≦ 30°C/60% RH
5	48 Hours	≦ 30°C/60% RH

IPC/JEDEC J-STD-020A, Table 2, in Moisture/Reflow Sensitivity Classification for Non-Hermetic Solid State Surface Mount Devices, Copyright IPC, Northbrook, IL and Electronics Industries Alliance, Arlington, VA. Used by permission.

- Short circuit between leads, due to bad solder printing, slumping of the solder paste or too rapid out-gassing of the solvent in the solder paste

Many IC packages are more or less sensitive to humidity; these components absorb moisture from the air. If there is moisture trapped inside the component, it might result in what is called *popcorning* when the board is sent through the reflow oven. The moisture starts to vaporize and finds no fast way out of the package. These humidity-sensitive components are delivered in sealed bags, called *dry-pack*. The time, after opening the sealed bag, within which the component must be assembled and sent through the reflow oven, is specified in standard IPC/JEDEC J-STD-020A. This time is also called *floor life* (see Table 17.2).

If the dry-pack is opened and the components are not used within the specified time, the moisture must be dried out before reflow. In this process, called *baking*, the components are placed in an oven at a certain time and temperature specified by the manufacturer. The baking normally takes between 8 to 24 hours, depending on temperature in the oven.

Other Soldering Techniques

There are some other soldering techniques available that are mainly used for rework, prototyping and other special applications.

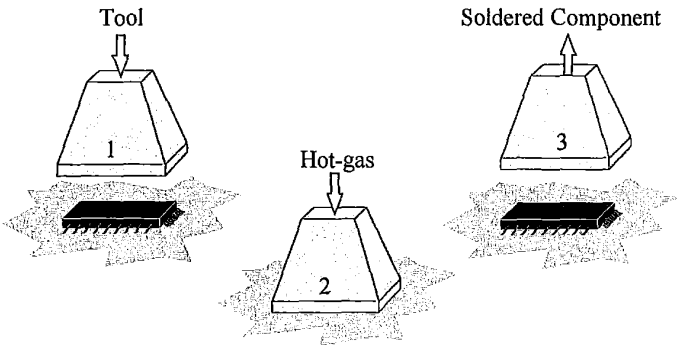


FIGURE 17.20 Hot-gas soldering of a flatpack component.

Samsung, EX1028, p. 341

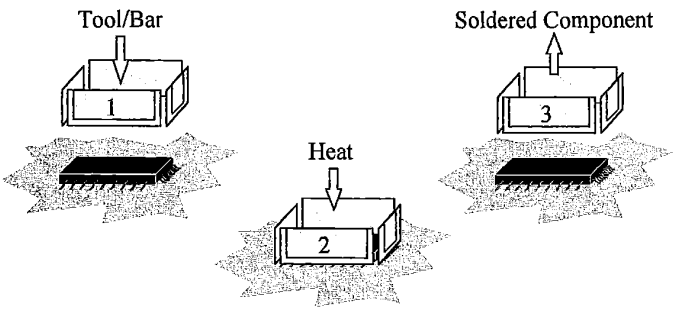


FIGURE 17.21 Hot-bar soldering of a flatpack component.

Hot-gas soldering (Figure 17.20) is a process where a nozzle, or housing, is placed over the component, and hot air or nitrogen is blown through the nozzle onto the component. For more complex boards, the use of backside heaters is essential to prevent warpage.

Hot-bar soldering (Figure 17.21) is similar to hot-gas soldering in that it also uses a special designed bar. The bar is heated and placed in contact with the leads on the component. It is then pressed down, and the heat is transferred through the leads down to the pads on the board.

Soft-beam soldering (Figure 17.22) and *laser soldering* (Figure 17.23) use different kinds of light to accomplish spot soldering (i.e., one lead at a time is soldered). These techniques are not as fast as wave or reflow soldering, where the whole assembly is soldered in one pass, but they can be very useful in mixed assembly. Both techniques use the beam to heat up the solder wire, or pre-deposited solder paste, together with the component and board.

17.2.7 Surface Mount Components

In the beginning of the surface mount assembly era, the pitch of the surface mount components was smaller than the pitch of through-hole components. Another advantage of SMT is that the components do not use as much area on the topside and do not use any area inside, or on the bottom-side, of the board. The size of chip components started with 1206 and 0805 components; the 0805 component, e.g., is 0.08 inch long and 0.05 inch wide (2.03 × 1.27 mm).

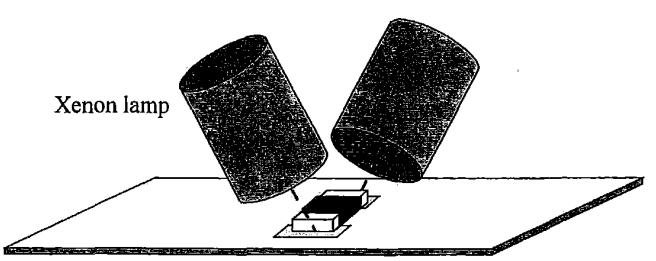


FIGURE 17.22 Soft-beam soldering of a chip component.

Samsung, EX1028, p. 342

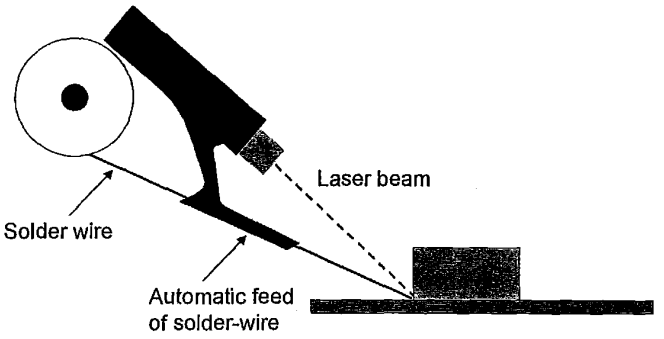
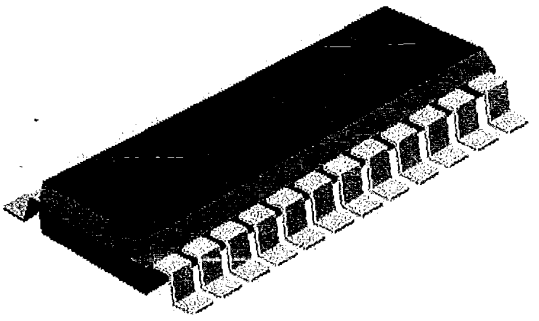


FIGURE 17.23 Laser soldering of a large component.

The IC packages started with a lead pitch of about 50 mil. Examples of IC packages are the small outline integrated circuit (SOIC), shown in Figure 17.24, with leads on two sides of the component, and the plastic leaded chip carrier (PLCC), with leads on all four sides. When the pitch of the components went down towards 20 and even 15 mil, there were a lot of challenges that made the industry hesitate to follow the path towards finer pitches. There was a need to come up with a new solution for the future.



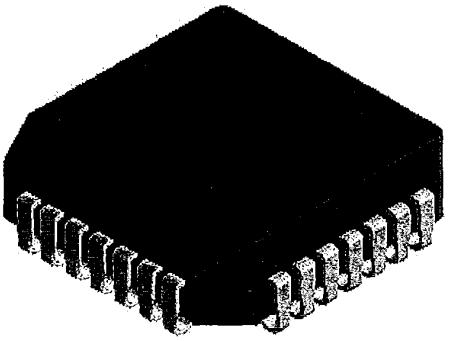


FIGURE 17.24 (a) SOIC and (b) PLCC components.

Samsung, EX1028, p. 343

Some of the more difficult hurdles to overcome with fine pitch surface mount assembly are:

1. Smaller apertures in solder-paste printing, requiring smaller powder sizes in the solder-paste and thinner screen stencils to avoid clogging in the apertures
2. Tighter process control in screen-printing and more careful choice of process materials and process parameters
3. More difficult pick up and handling of the small components, and the need for better placement accuracy for the pick-and-place machines

Examples of SMT components listed below are illustrated in Figure 17.25.

- Passive or discrete components (0402, 0603, 0805, 1206, etc.)
- Small outline transistor (SOT) (SOT 23, 89, 323, etc.)
- Small outline integrated circuit (SOIC) (SO14, SO16, etc.)
- Small outline J-leaded (SOIJ)
- Plastic leaded chip carrier (PLCC)
- Thin small outline package (TSOP)
- Quad flat pack (QFP)

17.2.8 Ball Grid Array

New advances in PWB manufacturing surfaced in the middle of the 1990s. Thinner traces and new build-up technologies made it possible to route out the traces from a new type of component called *area array package*. Instead of only using the perimeter of the

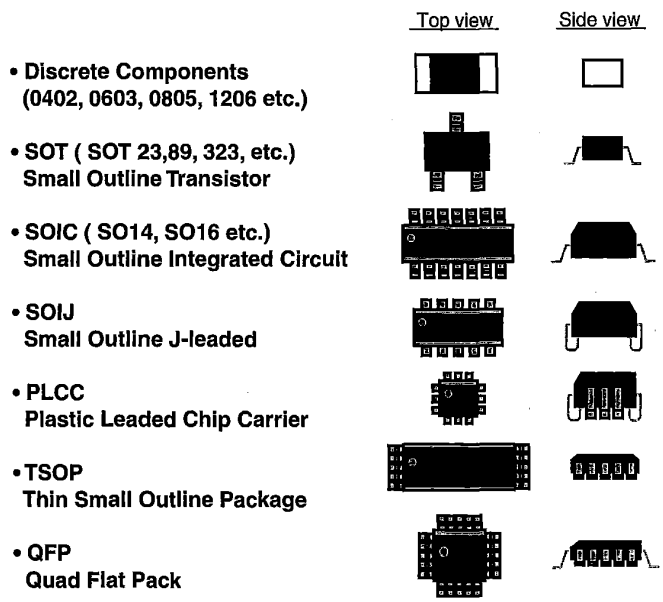


FIGURE 17.25 SMT components.

Samsung, EX1028, p. 344

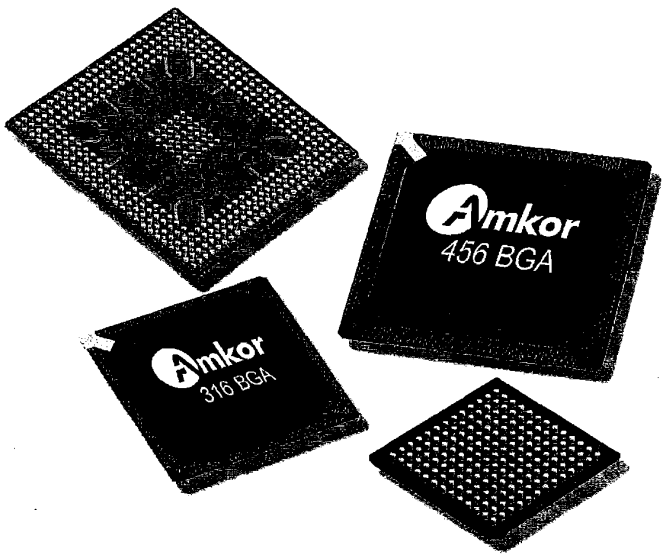


FIGURE 17.26 BGA components. (*Picture courtesy of Amkor*)

component, as done with traditional surface mount components, the area array components can utilize the whole bottom-side for interconnections. Even if the total I/O density for the area array component is higher, compared to the same chip in a QFP package, the area available for interconnections has increased. This makes it possible to actually increase the lead pitch, resulting in a more robust screen-printing process and less risk of short circuit between leads.

The most common area array component is the *ball grid array* (BGA), depicted in Figure 17.26, which utilizes small solder balls for the connection between the component and the board. The solder balls provide a shorter electrical path than the standard SMD leads. This creates less inductance and causes less degradation of the signal, improving the overall performance of the circuit. The active chip in a BGA package can be interconnected to the package by wirebonding, as illustrated in Figure 17.27, or flip-chip technology.

One big drawback with the BGA component is the difficulty to inspect the formed solder joints after assembly. In most cases, electronic manufacturers will use x-ray ma-

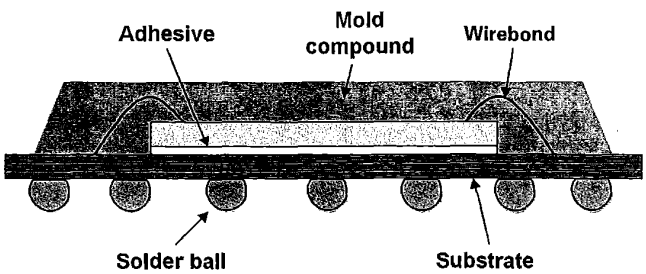


FIGURE 17.27 Typical BGA build-up using wirebonding technology.

Samsung, EX1028, p. 345

chines to inspect BGA components for errors in the assembly process. This method is mainly used for off-line inspection, and for process set-up, process control and failure analysis. The problem with x-ray is that it is very difficult to judge the appearance of the solder joint, to see if it is a cold joint and if there is good wetting towards the board. There is inspection equipment available that can be used to actually see the solder joints. They use a fiberoptic light source and a small optical instrument to get underneath the component body.

There are, however, design limitations for the layout on the board if such equipment is used (e.g., quite a large keep-out area around the component). The inspection is done manually and is rather time consuming compared to using a microscope for standard SMD components. The BGA assembly process is, however, so reliable that, if done correctly, no inspection is necessary.

17.2.9 Chip Scale Package

A *chip scale package* (CSP) is defined as having a side that is a maximum of 1.2 times the size of the chip or 1.5 times the area of the chip. A CSP component can very well be manufactured in the BGA format. Another example of a CSP package is the *land grid array* (LGA). The LGA has bare solder pads on the bottom side of the component. Only the stencil-printed solder-paste is used for interconnection. Flip-chip and wirebonding techniques, as illustrated in Figure 17.28, are both commonly used when building a CSP component. A rigid board, or a flex film, is normally used as a carrier for the chip. The chip in the upper component is attached to the substrate using wirebonding; the chip in the lower component is attached to the substrate using flip-chip technology.

17.2.10 CSP versus Flip Chip

The advantage of the CSP component, compared to direct chip attach of flip chip to a board, is that the CSP is designed to take up the mismatch in coefficient of thermal

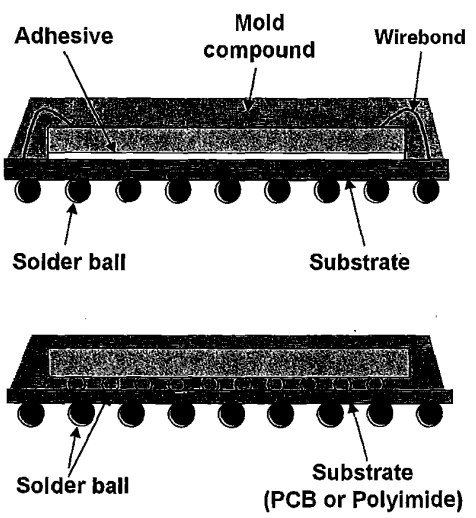


FIGURE 17.28 Build-up of two typical CSPs.

Samsung, EX1028, p. 346

expansion (CTE) between the chip and the board. Normally, it is therefore not necessary to underfill the CSP component after it is soldered to the board, which is necessary for the flip chip component. This is one of the reasons why the use of CSP packages took off in high volume in camcorders and cell phones in the late 1990s. The CSP component can normally be implemented into a standard surface mount process without problems.

17.3 THROUGH-HOLE ASSEMBLY

The number of pure through-hole assemblies is constantly decreasing. This trend is mainly due to the amount of board real estate wasted by the need for holes through every layer of the board. The through-hole share of the product volume has dropped significantly in the last decade.

17.3.1 Advantages of Through-Hole Assembly

Even though through-hole technology has some disadvantages, it will continue to be used for large and heavy components, due to the stronger mechanical fastening to the board. It is also beneficial for components that are exposed to large dynamic forces, such as connectors that are frequently plugged in and out. Through-hole assembly can also be necessary to use when large currents are conducted. A typical example would be power supply units, which have heavy transformers and coils, large connectors, and large currents. Most through-hole components can be separated into these groups: *axial, radial,* and *dual-in-line* (DIL or DIP), as depicted in Figure 17.29.

17.3.2 Component Insertion

Through-hole assembly is done by inserting the component leads into plated holes in the PWB and soldering them. The soldering operation can be performed using either a wave-soldering machine, a selective soldering system or a common solder iron. The components can be fixed in their positions before soldering by clinching the leads on the backside of the board. This is to keep the components in their proper positions during transport, or to keep lighter components from floating up when they pass the solder-wave in a wave-soldering machine.

Component insertion can be done either manually or automatically. In the case of manual insertion, the components are preformed to fit the distances between the holes, and they are sometimes also given a standoff due to thermal considerations, or for other reasons.

When doing manual assembly, reversible fixtures are sometimes used in order to access both sides of a board. First, the components are mounted, then the lid of the fixture is closed, and the board can be turned over for cutting of the leads or for soldering, without the components falling out.

Automated insertion can be separated into axial, radial and DIL insertion. Axial components can be set up in a sequencing machine, where the different components are put together in a mix to suit a certain board. The machine arranges a sequence of different components, as illustrated in Figure 17.30, to be mounted in the next step. The sequence is repeated many times and the resulting ribbon of components is wound up on a reel.

Samsung, EX1028, p. 347

FIGURE 17.29 Different through-hole component types: (a) axial; (b) radial; and (c) dual-in-line (DIL).

Samsung, EX1028, p. 348

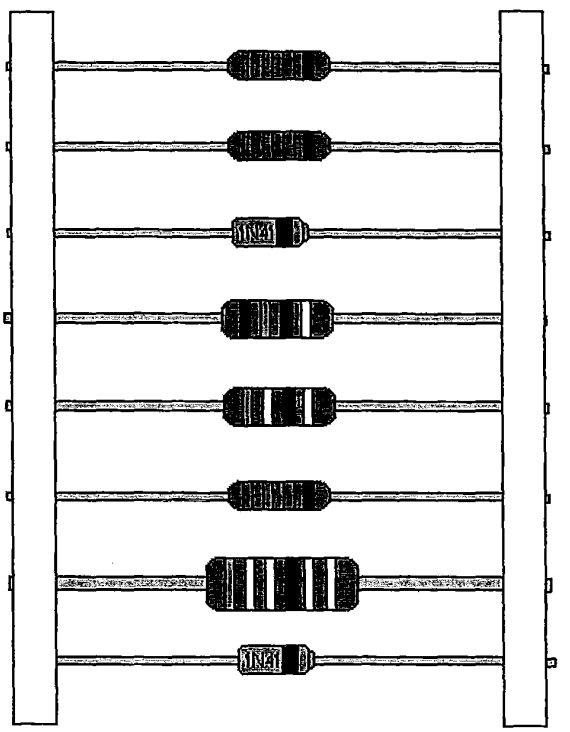


FIGURE 17.30 Component sequence.

This reel is set up in the axial insertion machine, which forms the leads, inserts them into the holes and cuts and clinches them on the backside, as shown in Figure 17.31.

A radial insertion machine directly assembles components from the original reels, without the sequencing step. It is also possible, however, to take a reel of axial components and have them cut and formed into radial components and inserted in a radial insertion machine. It depends on the product mix and the machines available as to which solution is preferred. When it comes to insertion machines for DIP, or dual-in-line packages, the components are set up directly in the insertion machine, without any sequencing. Dual-in-line packages are usually transported and delivered in sticks.

17.3.3 Wave Soldering

Wave soldering is the most common soldering method for through-hole assembly. Unlike reflow soldering, wave soldering both applies the solder, and provides the necessary heat for the solder joints to form. The assembled boards are usually placed in a fixture, and that fixture carries the boards through the wave-soldering machine. The fixture also holds information on how this particular board should be processed, which transport speed to use, which preheat settings, and so on. The fixture with the board travels through a fluxing station, preheat zone, solder wave and possibly cleaning, depending on which kind of flux is used. Passing the solder wave, the bottom-side of the boards just touching the

Samsung, EX1028, p. 349

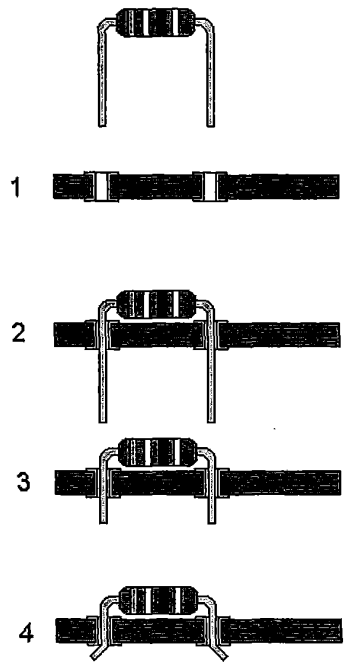


FIGURE 17.31 Axial insertion.

solder, and all component leads get soldered in one pass. Sometimes an inert atmosphere of nitrogen gas is used in order to improve the wetting properties. The wave is formed by molten solder that is continuously pumped up and falls over an edge, back into the reservoir. The solder makes contact with the entire bottom-side of the board, but it only wets the metal surfaces not covered by solder mask. A typical wave-soldering machine is depicted in Figure 17.32. Wave soldering is the most common soldering method for through-hole assemblies. Unlike reflow soldering, wave soldering both applies the solder and provides the necessary heat for the solder joints to form.

17.3.4 Press-Fit Assembly

Apart from surface mount and through-hole technology, there is also a technique called press-fit, which is mostly used for mounting connectors on boards. A press-fit connection between a printed wiring board and a connector pin is made by pressing the pin into a plated through-hole of the PWB. The fundamental idea is that the pin must have a larger diameter than the hole. This results in an overlapping of material at the pin and the through-hole, which must be taken up by deformation of either the pin or the hole. Usually this is accomplished by using a compliant pin that does not damage the wall of the hole.

17.3.5 Mixed Component Assembly

Mixed assembly (Figure 17.33) is the use of both SMD and through-hole components on the same PWB. Such an assembly is a very common technique. In 1999, 70% of all

Samsung, EX1028, p. 350

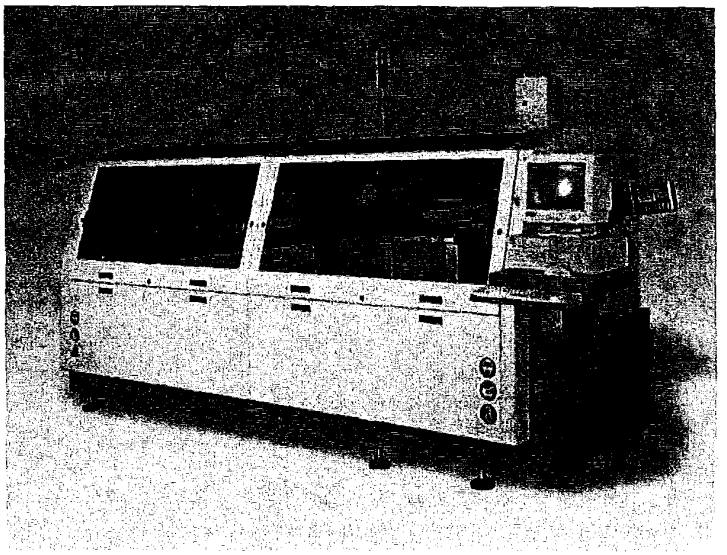


FIGURE 17.32 Wave-soldering machine. (*Picture courtesy of Sincotron, Vitronics Soltec DeltaMax*)

system-level boards produced had both SMD and through-hole assembly. The wave soldering process has been adapted to incorporate surface mount assembly as well. Using this process, the SMT components must first be attached to the secondary side of the board, using glue between the component body and the board. The glue used is an ordinary heat- or UV-cured adhesive. When the through-hole components have been assembled, the board is sent through the wave-soldering machine. Using SMT components in this process requires special orientation of the components on the board to achieve good solder joints and avoiding solder bridging. Mixed assembly is the use of both SMD and through-hole components on the same PWB.

Sometimes, board assemblies are divided into three different types, depending on which techniques have been used.

- Type 1: Only surface mount devices (no through-hole devices), both sides of the board are used and may include all sizes of actives and passives. Reflow soldering is used.
- Type 2: Surface mount devices and through-hole devices, the topside may have SMD and through-hole devices. The bottom-side is reserved for SMD only (passive SMD and small actives). Reflow and wave soldering are both used.

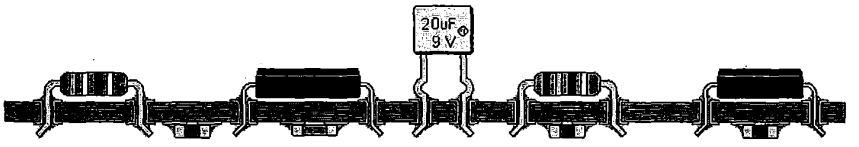


FIGURE 17.33 Mixed assembly.

Samsung, EX1028, p. 351

- Type 3: Surface mount devices and through-hole devices. The topside is only for through-hole, the bottom-side is reserved for SMD only (passive SMD and small actives). Only wave soldering is required.

When a board has mostly SMD components and only a few through-holes, it can become advantageous to skip the wave-soldering process and use a pure SMD process. After SMD assembly, the remaining few components can be separately soldered, either manually or using a soldering robot. Another way to get rid of the wave-soldering process is to use the "pin in paste" technology. This means depositing solder paste over the hole in the board, inserting the component (preferably using a surface mount machine) and reflowing the solder at the same time that the rest of the assembly is reflowed.

17.4 GENERIC ASSEMBLY ISSUES

The following areas are generic for surface mount assembly, through-hole assembly and mixed assembly.

17.4.1 Additive Materials

Apart from solder, additive materials can include flux, conductive adhesive, non-conductive adhesive, thermally conductive grease, conductive film, or other tape. Solder can have different compositions (alloys) and can also come in several different forms. It can be used as solder paste, solder wire, or solder bars to be molten into a wave-soldering bath, or as pre-forms. Pre-forms are pellets, or pieces of solder in different shapes, which can be placed in the desired positions and reflowed.

Conductive adhesives are usually thermally curable epoxies, filled with conductive particles of, for example, silver. It is sometimes used to replace solder in special applications. It has the advantages of being lead-free and does not have as high a processing temperature as solder. The disadvantages are that there is no self-alignment of the components. It is also difficult to match the thermal, mechanical and electrical performance of solder.

17.4.2 Cleaning

Cleaning is done to remove contamination on board assemblies that might cause corrosion or hinder the soldering process, causing assembly yields to go down. For high-performance applications, like military and medical equipment, cleaning is used to improve the reliability, and prolong the lifetime of the product. Reasons that might make it necessary to clean the assembly can be the use of high frequency, high speed, high voltage, or extreme reliability demands.

In high-volume production of consumer electronics, the use of no-clean fluxes means big cost saving and time saving, since the cleaning step can be excluded. When the ozone-harmful CFCs, used earlier in SMT cleaning, became prohibited, the fluxes were improved into what is called no-clean fluxes. The no-clean fluxes are soluble in organic liquids and do not normally require any washing after soldering.

Aqueous Cleaning: The water-soluble flux used in this process leaves residues that require washing immediately after soldering to prevent corrosion of the solder joints. The

Samsung, EX1028, p. 352

cleaning occurs after a chemical reaction of a saponifier, added to the water, converting the flux residues to a flux "soap," which is then washed off the board with de-ionized water.

Semi-aqueous Cleaning: The big difference from pure aqueous cleaning is that the assembly is first cleaned with an organic solvent before the water rinse takes place. Ultrasonic sound waves make the contamination come loose easier and can therefore be used to improve the result of the cleaning. In some cases, the ultrasonic treatment might have a negative impact on the reliability of high-performance devices.

17.4.3 Rework and Repair

The theoretical goal of automated PWB assembly is, of course, 100% yield. Yet failures have always been a part of processing, particularly as line speeds increase and component sizes decrease. Also, the move toward higher complexity and higher functionality products—including BGAs, flip chips and other advanced technologies—has increased the value of each assembly.

The choice of tool to use for rework is very much dependent on which type of component is to be replaced. Most components, from small chip resistors to rather big flat-pack components, can be replaced with a standard soldering iron or thermal tweezers, shown in Figure 17.34. The tool must be chosen so that no thermal or mechanical damage or degradation is caused to any traces, solder pads or other components on the board.

To achieve a reliable solder joint, it is recommended to remove the old solder before placing the new component; this can be done using a desoldering braid and a soldering iron. One reason to do this is to minimize the possibility of bridging and shorting. The operation can also be done with a desoldering vacuum tool. For both techniques, there

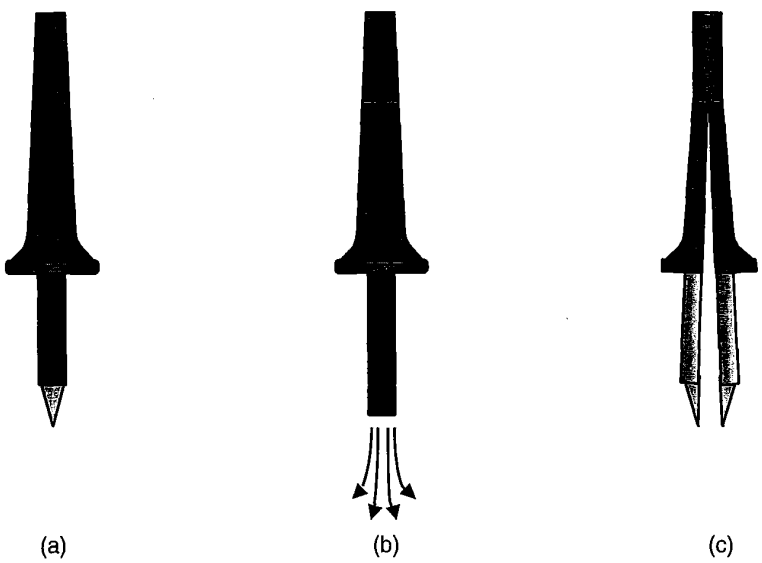


FIGURE 17.34 (a) Standard soldering iron, (b) a hot-gas tool, and (c) thermal tweezers.

Samsung, EX1028, p. 353

is always a risk of lifting pads or removing the solder mask on the board. Before placing the new component, the site needs to be cleaned from residues using alcohol solvent or equivalent.

More advanced SMDs, including BGAs and flip chips, require hot gas or IR rework systems for removal and replacement. These pieces of equipment are also very useful when replacing PLCCs and flatpacks. Most of the semiautomatic rework stations include preheaters and backside heaters. For more complex boards, the use of backside heaters is essential to prevent warpage. In hot-gas equipment, the gas is forced through a nozzle at a temperature determined by the preprogrammed thermal profile to reflow the solder without damaging the substrate or surrounding components. The thermal profile is set to follow the standard reflow profile as much as possible.

The type of nozzle varies with equipment and with technician preference. Some blow air above and below the component, while others move the air horizontally, and some to the top of the device only. Sometimes it is better to use a shrouded nozzle, trying to protect nearby components by focusing the hot air on the reworked component only.

The high-end rework systems also include an optical split vision system. This allows the operator to see a magnified image of the entire bottom-side of the component, superimposed over the corresponding land pattern on the board.

Last, but not least, it is not to be forgotten that a good rework result requires a trained and highly skilled operator. Trying to save money by using cheap or improper equipment, or letting a poorly trained operator perform the rework, can be very costly.

17.4.4 Electrostatic Discharge

Most people have probably experienced an electrostatic discharge (ESD), such as getting an electric shock when getting out of a car in the winter (Figure 17.35). The same phenomena can be experienced when petting a cat. When wearing certain types of fabrics, you can easily get charged up to 20,000 volts just by walking around. The charge is achieved when two materials are rubbed against each other or when two materials are separated. Even if the discharge is not very harmful to humans, the high voltage and power can easily damage a component or assembly; although the component does not necessarily fail at once. It might only be partially damaged at that moment, which can lead to a failure after some time.

The effect of ESD is something that is not easy to prove. In the worst case, ESD damage cannot be detected in production, but it will show up much later out in the field



FIGURE 17.35 Electrostatic discharge (ESD) sensitive material or product.

Samsung, EX1028, p. 354

when the customer uses the product. The extent of ESD damages to the components and assemblies in the production, leading to failure at a later time, can be very difficult to pinpoint.

17.5 PROCESS CONTROL

When producing high volumes of electronics, assembly yield becomes very important. In many other circumstances, a 99% correct result would be considered successful. However, imagine you have 333 components on a PWB. If the components have three terminations each as an average, that would add up to about 1000 soldered interconnections per board. If 99% are correct, that means there will be an average of 10 defects on every board and the production yield is 0%. Now, let us assume that a 99% production yield is desired instead. That means one defect on 100 boards, or 100,000 interconnections. It would be necessary to make 99.999% correctly formed solder joints. It is, however, more common to discuss parts per million (ppm) in the industry. In this example there would be a defect rate of 10 ppm. Maintaining a process like that is an interesting challenge, considering how many different materials and processes are involved in forming each joint. This is why it is very important to create margins in the processes; this means that the result of a process should come out well, even if there are slight variations in the materials or parameters. Usually, this is referred to as a process window, a range of settings within which the process is yielding a good result.

17.6 DESIGN CHALLENGES

At the early stages of design, many decisions are made that will have a substantial impact on the assembly result. Choice of component packages, spacing between components and design of soldering surfaces are especially critical. When choosing a component, it is important to consider not only the function but also the processability of that component. Considering the producability of a new product in an early stage of development will have a tremendous effect when it comes to reducing manufacturing cost and improving first pass yield. This means involving experienced production engineers in the design process and insuring good communications between manufacturing and research & development departments. This kind of work sometimes follows ready-made models and is often called *design for manufacturing* (DFM). DFM software tools are readily available on the market. The important thing, however, is to have guidelines for design based on the capabilities and limitations of the manufacturing processes. Such a document can become an excellent starting point for improved communications between manufacturing and design. Some simple examples are having a flat surface on top of the component, sufficient for a vacuum nozzle in a placement machine. The package should also be able to survive the soldering process and it should be possible to perform rework. If a cleaning process is going to be used, the component must be able to withstand it. Choosing the right size production panels for PWB, considering price, manufacturability, etc., is also important. The design of the soldering surfaces is also vital for the assembly yield. All in all, it is a matter of designing for manufacturing robustness, which could mean anything as simple as making guiding pins on connectors, tapered instead of straight, or anything else that simplifies assembly and minimizes the risk of defects. A

Samsung, EX1028, p. 355

mistake in the manufacturing will affect one product or a batch of products, but a mistake in design will affect every product.

17.7 SUMMARY AND FUTURE TRENDS

High expectations on increased performance, decreased size, weight and cost, in combination with tough competition, are powerful drivers for the future development of electronic assembly. The chart in Figure 17.36 shows a component density of 20 components/cm^2 in 1990 and 50 components/cm^2 in 1998.

Flip Chip: One promising techniques is the use of flip chips. Flip chips are unpackaged silicon dies that have been supplied with solder balls directly on the active side of the die. The reason they are called flip chips is that they are flipped upside-down, compared to a conventional wirebonded chip. Since the die is not packaged at the time of assembly, some kind of mechanical and chemical protection must be added. This is done by adding an epoxy-compound called underfill and curing it. The main benefits of flip chips are minimal size, low cost for IC package and potentially good electrical performance due to the short signal paths. One of the challenges to overcome when using flip chips on glass-fiber reinforced epoxy boards, is the CTE mismatch between the silicon of the die and the board material. This mismatch can cause reliability problems if not correctly resolved. Other challenges are high placement accuracy demands due to the fine pitch of the solder-bumps, and added cost due to additional process steps.

Integral Passive Components: The use of integral passive devices instead of discrete components can have a large impact on future electronic assembly if they become com-

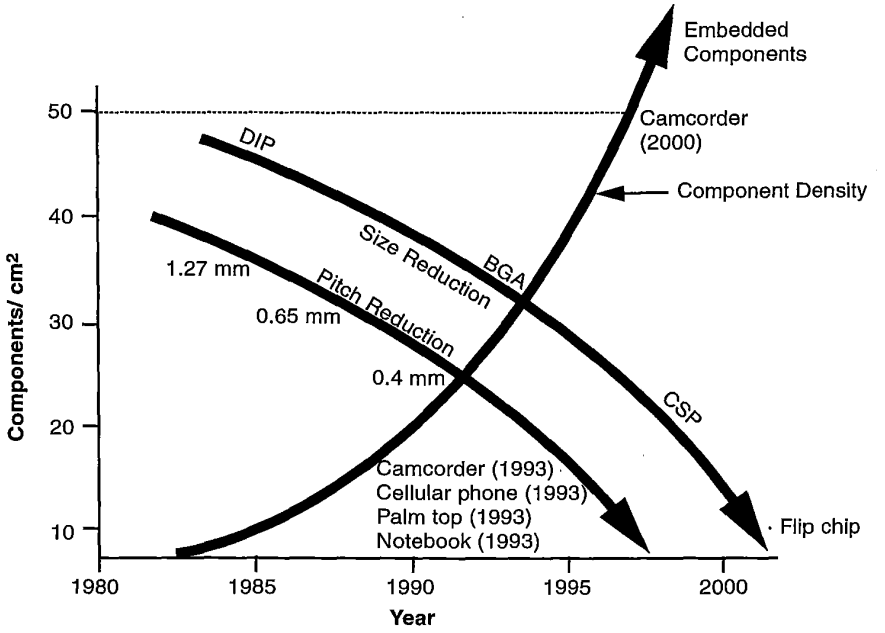


FIGURE 17.36 Japanese consumer product component density trend. (*Courtesy of Sony*)

Samsung, EX1028, p. 356

mon. Integral passive devices are resistors, capacitors and inductances that are embedded in the PWB, instead of mounted on top of it thus moving the passive devices into the PWB production processes.

Environment: Another important issue for the future of electronic manufacturing is the protection of the environment. The tremendous growth in this business segment, in combination with the many different materials and chemicals used, can seriously impact the environment. Eliminating toxic lead by using lead-free solders is one way of protecting the environment. In Japan, lead-free electronic products are already (year 2000) on the market, and Europe is considering legislation. The halogen-based flame-retardants used in PWBs and component packages are also being replaced more and more by halogen-free flame-retardant systems. Recycling of electronic products is also starting to catch on. More responsibility is expected from the manufacturers in the future. Increased miniaturization should also have a positive impact, since less material will be used for each product.

17.8 HOMEWORK PROBLEMS

1. What are the main benefits of surface mount assembly compared to through-hole technology?
2. What are the main benefits of through-hole technology compared to surface mount assembly?
3. Describe the assembly and soldering process for a mixed assembly.
4. Draw and describe a typical reflow profile, the critical times, temperatures and the different purposes of the zones.
5. What are the advantages and disadvantages of using nitrogen?
6. Why is the particular alloy of 63% tin and 37% lead advantageous to use?
7. Why should we consider environmental issues?
8. What are the benefits of area array packages?
9. What is solder paste? Describe the content and purpose of each ingredient.
10. Why is optical alignment of PWBs used in placement machines?
11. How are stencils for solder-paste printing manufactured?

17.9 SUGGESTED READING

1. *The Electronics Industry Report.* New York: Prismark Partners LLC, 1997–1998.
2. *The Electronics Industry Report.* New York: Prismark Partners LLC, 1998–1999.
3. R. Prasad. *Surface Mount Technology: Principles and Practice,* Second Edition. New York: Chapman & Hall, 1997.
4. *Printed Wiring Board and Contract Assembly to 2004.* New York: Allied Business Intelligence, Inc.
5. R. Tummula, E. Rymaszewski, and A. Klopfenstein, eds. *Microelectronics Packaging Handbook,* Second Edition. New York: Chapman & Hall, 1997.

ACKNOWLEDGMENTS

The editor acknowledges the initial contributions by Dr. Ron Lasky, M. Lee of Cooleson which led to this chapter. The editor gratefully acknowledges the contribution of Mr. James Wooldridge of Rockwell-Collins for critically reviewing this chapter.

Samsung, EX1028, p. 357

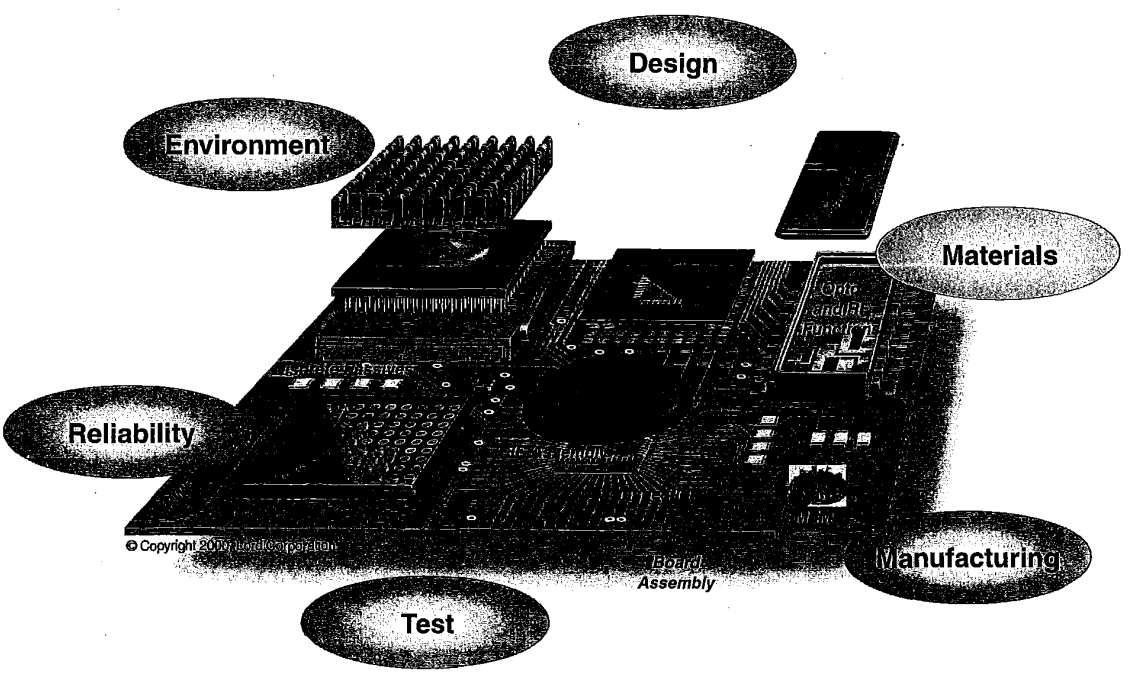
CHAPTER 18

FUNDAMENTALS OF PACKAGING MATERIALS AND PROCESSES

Dr. Pulugurtha Markondeya Raj
Georgia Institute of Technology

Prof. Johan Liu
Chalmers University of Technology, Sweden

Per Öhlckers
Olso University, Norway



18.1 The Role of Materials in Microsystems Packaging
18.2 Packaging Materials and Properties
18.3 Materials Processes
18.4 Summary and Future Trends
18.5 Homework Problems
18.6 Suggested Reading

CHAPTER OBJECTIVES

- Describe the diverse and crucial role of materials in microsystems packaging
- Describe the most important material properties relevant to packaging
- Discuss packaging processes for the IC and system-level packaging
- Look at future materials, relevant issues and bottlenecks

CHAPTER INTRODUCTION

Materials constitute the heart and soul of all microelectronic systems, because their properties dictate how efficiently their system-level performance, size, cost and reliability are met. This chapter reviews the role of materials in providing the system-level functions, and it translates these functions into specific electrical, mechanical, chemical and thermal properties. Finally, it describes both the thin- and thick-film processes by which these materials are processed into IC and system-level packages.

Samsung, EX1028, p. 359

18.1 THE ROLE OF MATERIALS IN MICROSYSTEMS PACKAGING

Materials provide several functions in microelectronic packaging. They transmit signals from IC to IC, supply power to ICs, provide interconnections to form the system-level hierarchy, mechanically and environmentally protect ICs, and dissipate heat. These functions are schematically depicted in Figure 18.1. The functions of the package dictate the required properties, based on which appropriate materials are chosen within the cost and processing constraints. A comprehensive list of packaging technologies, functions, properties, materials and processes are described in Figure 18.2.

18.1.1 Integrated Circuit Packaging

IC Packages

Packaging of an *integrated circuit* (IC), usually referred to as first-level packaging, should provide mechanical and environmental protection of the IC, remove the heat that is generated by the IC, and provide electrical connections to the rest of the components by means of a systems-level board. Ceramics such as alumina and glass ceramics offer good protection from the atmospheric moisture because of their hermiticity, have thermal con-

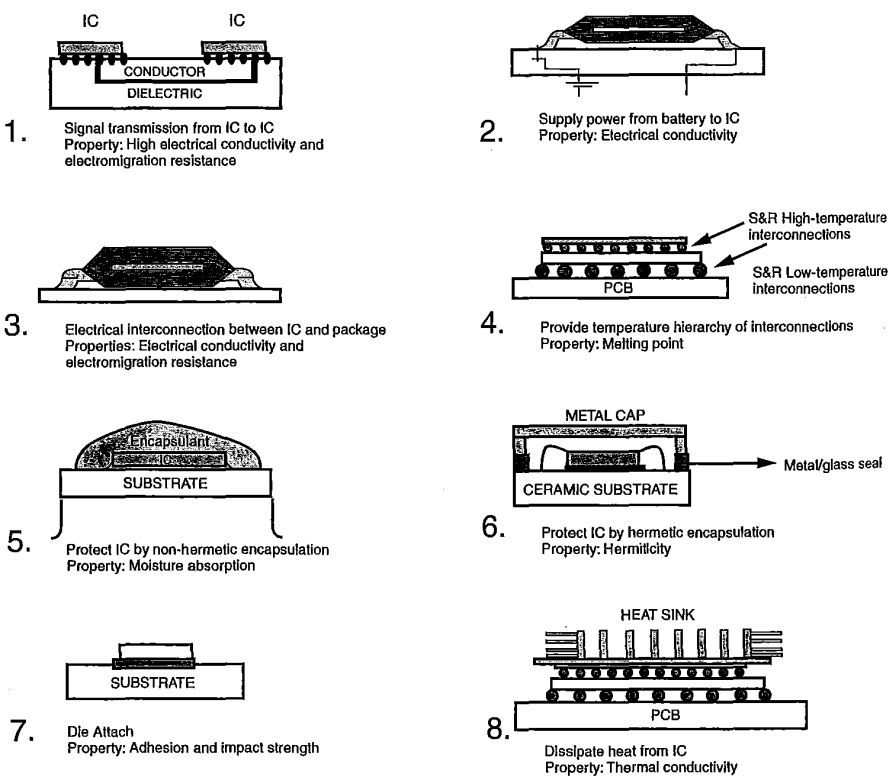


FIGURE 18.1 Role of materials in microsystems packaging.

Samsung, EX1028, p. 360

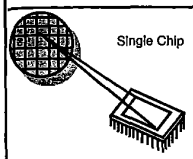
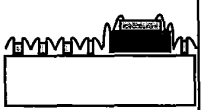
PACKAGING TECHNOLOGIES	FUNCTIONS	MATERIALS	PROCESSES	PROPERTIES
 Single Chip	• Protection • Interconnection • Heat dissipation	• Plastics • Epoxy + fillers • Ceramics	• Compress molding • Tape casting • Dry pressing	• Permitivity • Thermal Expan. Coeff • Thermal conductivity • Moisture absorption
IC Assembly 	• Electrical connection	• Cu • Al • Au	• Thermocomp. bonding • Ultrasonic bonding	• Electrical conductivity • Oxidation resistance • Fatigue • Creep
 TAB	• High I/O density	• Cu on polyimide	• Electroplated copper • Polymer deposition	• Heat resistance
 Flip Chip	• Reliability • Environmentally friendly	• Pb -Sn solder • Non-lead solders • Cr-Cu/Au • Cr/Cu/Ni/Au • Conductive adhesives: Ag-filled epoxy • Underfill: Silica filled epoxy	• Reflow • Evaporation plating • Evaporation dep. • Electroplating, etching • Screen printing • Curing • Dispensing	• Wetting • Surface properties • Melting point • Electromigration resistance • Interdiffusion • Electrical conductivity • Modulus • CTE • Nontoxicity
Passives 	• Band-pass filters • Decoupling • Tuning • EMI shielding • Power converters • Termination (resistors)	• $BaTiO_3$ • PMN-PT • Ta_2O_5 • Titania • Ferrites • Ru_2O_3 doped glass • NiCr, TaN • Conductive composites	• Screen printing • Thin film processing • Sputtering • Evaporation • CVD, MOCVD	• TCC • Leakage current • Break-down voltage • Inductance • Q-factors • Temp. Coeff. Resistance • Humidity Coeff. Resis. • Voltage Coeff. Resis.
System-level Board and Assembly 	• Wiring • Signal Speed • Electrical Connection • Reliability	• Glass-epoxy (FR4) • Polyimide • Bismaleimide triazine • Epoxy • Electroplated copper • Pb-SnSolders • Lead-free solders • Conductive adhesives	• Lamination • Electroplating • Photolithography • Wave soldering • Dispensing • Reflow • Curing	• Permittivity • Dielectric Loss • Elastic Modulus • CTE • Glass-transition temp. • Moisture absorption res. • Melting point • Conductivity • Modulus • CTE

FIGURE 18.2 Key packaging materials, processes and properties.

Samsung, EX1028, p. 361

ductivity that is high enough to help with heat dissipation and have thermal coefficient of expansion closely matching that of silicon. Therefore, ceramics provides thermomechanical reliability. Though ceramics are widely pursued by industry, they were almost completely eliminated by organic IC packages. Strongly driven by their low-cost and ease of processing, organic materials are widely used, in spite of their deficiencies when compared to ceramics. A cheap and low-temperature compression molding process is used to accomplish this. Polymers perform better electrically than ceramics because of their low dielectric constant, except for applications where ultra-low loss is required. The materials, properties and processes for IC assembly are summarized in the first row of Figure 18.2.

18.1.2 IC Assembly

The electrical interconnections between the chip and package are provided by metal wirebonding techniques, as indicated in the second row of Figure 18.2. The conducting wire should have a high electrical conductivity, oxidation resistance, and good wetting to the bonding pads and mechanical properties to withstand creep and fatigue. Aluminum (Al) and gold (Au) are the most favored materials for wiring because of their high electrical conductivity and corrosion resistance. Wirebonding needs any two of the three conditions that assist joining: heat, compression or ultrasonic vibration. Wirebonding has been supplemented by *tape-automated bonding* (TAB) in instances where finer pitches are required. TAB is a process where chips are joined to a patterned metal on polymer tape, using automated thermocompression bonding. A typical system is copper wiring on a polyimide tape. With ever-increasing input/output connections from the chip, peripheral or side-connections are being replaced by area-array interconnection packages such as a *ball grid array* (BGA), or a flip chip assembly. These are classified as leadless processes where long wires are not needed for the interconnections, hence reducing the delay time. Electrical and mechanical interconnections are achieved by providing an appropriate number of solder bumps on the wafer, which are then face-down bonded to the matching metal pads on a substrate. There is a strong drive to replace solders by conductive adhesives (polymer-based interconnects) to lower the processing temperature and thermomechanical stresses and eliminate hazardous materials such as lead. Passive components and related processes are described in Chapter 11.

18.1.3 System-Level Packaging

Boards

System-level packaging provides wiring that forms an electrical interconnection for all components within the system. The organic substrate that provides these functions is called a *printed wiring board* (PWB). MCM packages on PWB type laminates, referred to as MCM-L, limit the interconnection temperature to less than 200°C. An epoxy, reinforced with a glass cloth to provide enough stiffness, is widely used for making PWBs. The board should also provide a mechanical base for various devices and components that form the final system. With increasing integration on semiconductors leading to more input/output density, and a strong drive towards miniaturization and faster processing speed, it becomes necessary to use multilayered electrical wiring in a thin, low-

Samsung, EX1028, p. 362

permittivity medium as part of the PWB. Thin-film packaging techniques such as electroplating, etching and photolithography (called build-up processes, similar to the ones used to make an IC), are employed for this. For high-performance applications, the dielectric material directly affects the performance, since its permittivity limits the signal propagation speed, and its thickness determines the characteristic impedance of the signal lines. A low-dielectric constant material is desirable to minimize capacitative coupling between the signal layers. Ceramic thick films provide high wiring density because a greater number of layers can be used, while polymers with low-dielectric constant provide dense wiring within a layer. Hence, many companies have successfully employed a hybrid of thick- and thin-film packaging.

Multilayered ceramic (MCM-C) packaging is dominant in applications that must meet the stringent reliability requirements for defense and aerospace applications, and high wiring density for high-speed microprocessors in mainframe and supercomputer applications. Though alumina has been used for this, its high-temperature cofiring treatment precludes the use of high conductivity copper wiring. Alumina also has a high dielectric constant, thereby limiting the signal speed and high thermal coefficient of expansion, causing reliability concerns when joined to silicon chips with relatively low expansion coefficient. Various glass ceramics have been developed to overcome these deficiencies and are being successfully used.

Board Assembly

Surface mount technology (SMT) interconnections are achieved by soldering, with the most common soldering compound being an eutectic Pb-Sn alloy with a melting point of 183°C. A huge *coefficient of thermal expansion* (CTE) mismatch between the PWB and IC induces significant stresses that cause failure at the solder joints. This technical challenge increases with the chip size, power density and the drive to use low-cost board materials. The thermomechanical stresses that cause failure can be alleviated by using an underfill material such as epoxy, that mechanically couples the IC and substrate, hence reducing the strain on the solder. An underfill material with a CTE matching that of solder further helps this process, and hence various fillers are being used to lower the underfill CTE. The stresses at the joints can also be reduced by using a compliant organic conductive adhesive instead of the relatively high moduli solders. Thus, there is an increasing trend to pursue adhesives based on organics in place of solders. Other driving forces are cost reduction, lower process temperatures, fatigue resistance and the elimination of hazardous Pb-based materials. The adhesives are typically polymers filled with high electrical conductivity silver flakes. These are still mostly in the development state.

The diversity of the materials and processes used in manufacturing an electronic package is evident from the above discussion. This chapter provides more details on various materials, processes and properties that are used in packaging. The chapter describes the role of materials in achieving the system-level packaging functions. The physical origin of the related properties such as electrical conductivity, CTE, modulus, etc. are then discussed and the range of properties that the available materials can provide are presented. Basic processes that build the package are then described by broadly classifying them as thick- or thin-film technology. Thick-film processes employ powder or pastes to build layers unlike in thin-film technology, where layers are built from individual atoms or molecules. Thick-film processes are typically followed by a costly high-temperature

Samsung, EX1028, p. 363

heat-treatment process, while vapor-based thin-film processes need costly high-vacuum equipment. The selection of materials and processes dictates the interplay between system-level performance, cost, reliability and size.

18.2 PACKAGING MATERIALS AND PROPERTIES

Materials selection for any application is guided by its properties. Hence, it is important to understand the origin of these properties and the materials available for the required range of properties. The properties relevant to packaging are electrical and thermal conductivity, coefficient of thermal expansion, electrical permittivity, polymer glass transition temperature and Young's modulus. The spectrum of materials and the range of properties are broadly described in Figure 18.3. Other properties such as thermal stability, adhesion, wetting, etc. also dictate the utility of a material. These properties are determined by the lattice or molecular structure, the atoms that constitute the lattice and their interactions, and the extrinsic effects such as impurities. No single material has the required combination of properties. Design of materials by their suitable combination (composites) hence emerged as an important discipline in materials science and engineering. Composites are widely employed in packaging to modify and improve ceramic, polymer and metal properties. Various fillers and reinforcements are introduced into these matrices to increase their stiffness and thermal conductivity, alter electrical permittivity, lower thermal expansion coefficient and more.

18.2.1 Electrical Properties

Conductivity

When an electric field is applied onto a conductor, the electrons drift towards the positive potential, resulting in a current. Electrical conductivity is the ratio of current density and the applied electric field

$$J = \sigma E$$

where J is the current density [A/m^2], σ is the electrical conductivity [Ohm^{-1}m^{-1}] and E is the electric field [V/m]. σ is a measure of the number of available electrons for conduction and their mobility. The inverse of electrical conductivity is resistivity. Electrical conductivity depends on the material and temperature. It is generally, but not always, independent of applied voltage (Ohm's Law). Current results from electronic and ionic conduction. We confine our discussion to electronic conductivity in this section.

Microelectronic packaging requires the materials that are the best insulators and the best conductors. Fortunately, the electrical conductivity of solid materials varies over 25 order of magnitudes from a good insulator (quartz) to a good conductor like copper and silver. The spectrum of materials which cover these resistivities are shown in Figure 18.3. Most covalent and ionic solids are insulators, whereas metals are good conductors. Semiconductors form an intermediate group between these two.

The huge difference in conductivity can be explained with the band theory of solids. The innermost electron energy levels are typically completely filled and tightly bound to the nucleus. However, the outer shells are influenced by the fields of other nuclei. From quantum mechanical considerations, the energy levels that are influenced by a very large

Samsung, EX1028, p. 364

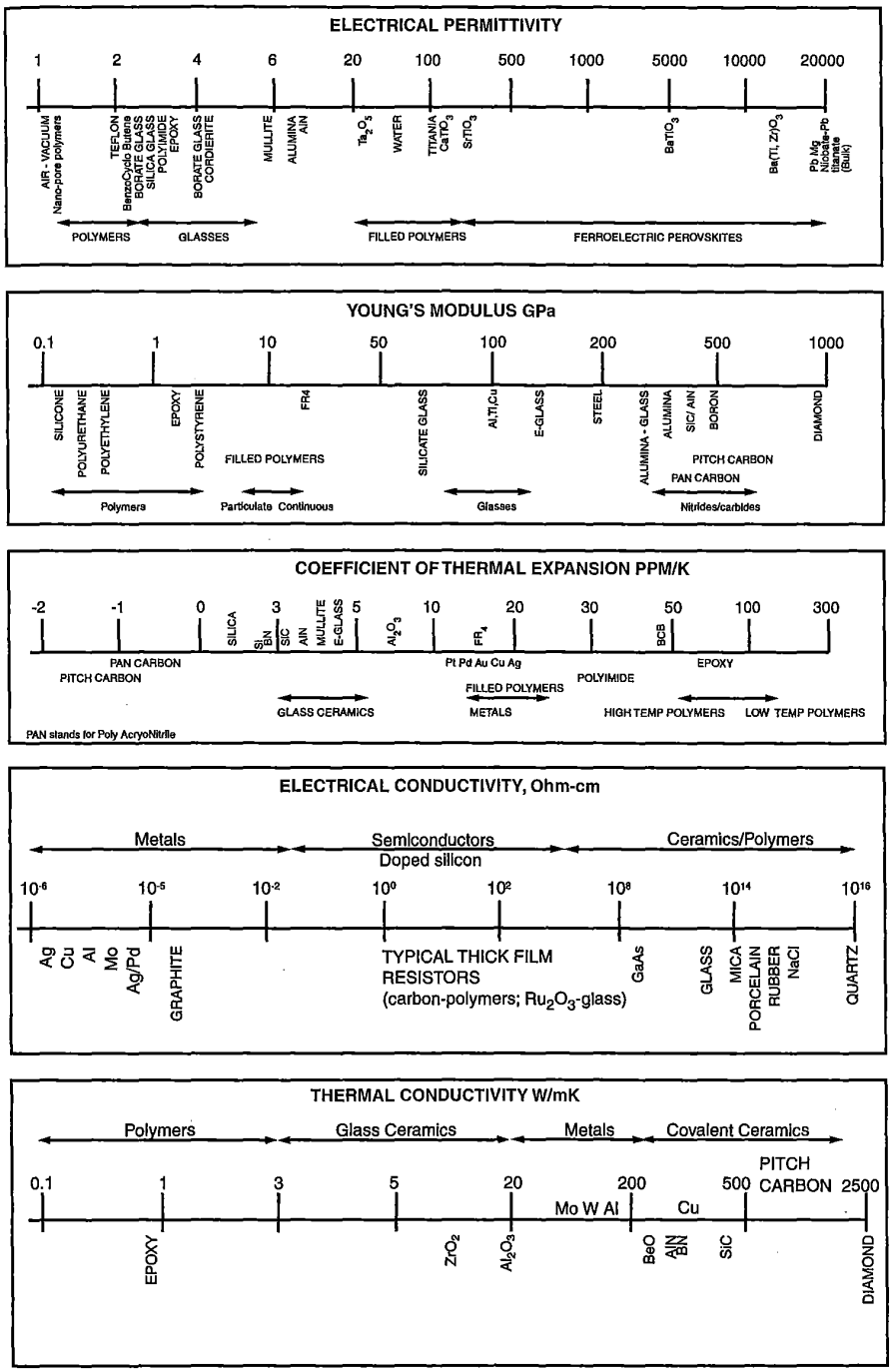


FIGURE 18.3 Key packaging materials and their properties spectrum.

Samsung, EX1028, p. 365

number of nuclei form a quasi-continuous band. The outermost energy band or valence band provides the basis for the prediction of the properties of a solid. If the band is partially filled, the electrons can pick up enough energy from an electric field to be able to "jump" to a nearby higher energy level in that band, where they can easily be available for conduction. This is the typical situation for metals. For some metals, all the innermost bands should be filled. A typical example is the 3*s* band in magnesium. However, in these elements, the band above the valence band overlaps with it. The overlap between 3*s* and 3*p* bands creates available energy states in these lattices. These substances which have completely filled shells in the form of single atoms, but have overlapping bands in the lattice, are called semimetals. The electrons in the conduction band are referred to as free electrons and can be considered a classic gas of conduction electrons moving freely through a fixed lattice of ion cores.

For insulators, the valence and conduction bands are separated by a large energy gap. The electrons cannot jump easily to the conduction band. They stay in the valence band, where they are tightly bound to the nearby nucleus and hence immobile for conduction. The schematic depiction of valence and conduction bands and band gaps in metals, semiconductors and insulators is illustrated in Figure 18.4.

Electrical conductivity is limited by the collisions between "electrons" and "imperfections" in the lattice of the conductor. These collisions will cause the electrons to lose their energy and momentum. The result is "Joule heating" which also manifests as an electrical resistance. Impurities, lattice defects, grain boundaries and the thermal vibrations of the lattice are the typical imperfections in a lattice. This explains why the resistance in almost all metals increases with temperature. When mixing different metals to form an alloy, foreign atoms are deliberately added to the metal, increasing the number of imperfections, thereby increasing the resistance compared with that of the pure metals.

For insulators, the uppermost occupied band, or valence band, is completely filled, and there are no electrons in the next band. In this case, the properties depend on the forbidden energy gap between them. The forbidden energy gap between the valence band and the conduction band is so large that neither thermal energy, nor an applied electrical field, is able to excite the electron into the conduction band. For semiconductors, this

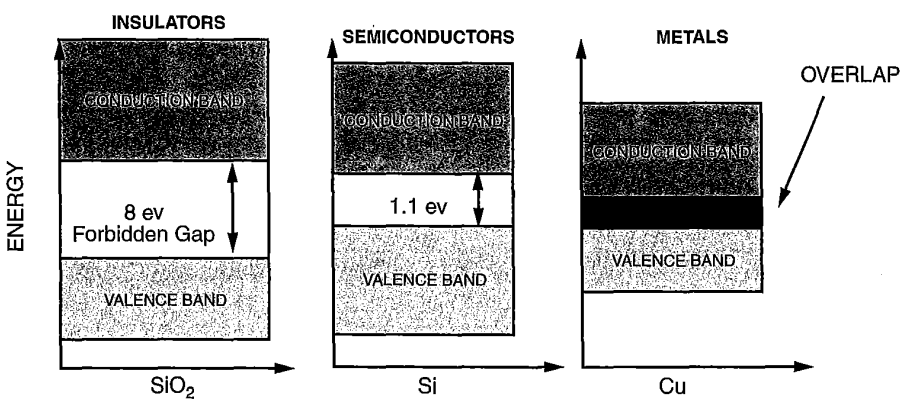


FIGURE 18.4 Band gaps in materials.

Samsung, EX1028, p. 366

forbidden gap is much smaller. This means that some of the electrons have enough thermal energy to jump into the conduction band. The number of excited electrons in an intrinsic semiconductor is exponentially dependent on the absolute temperature. This partly explains the very large temperature dependence of the electrical conductivity for these materials.

Semiconductors are doped with other elements to improve their conductivity and have a means to control their current carriers. When pentavalent atoms are added to a pure semiconductor, there is an excess electron that is not covalent bonded to any atom. These are called N-type semiconductors, where electrons are the majority charge-carriers. Similarly, when a trivalent atom is added to form a P-type semiconductor, there is a deficiency of electrons, creating a vacancy or hole. The movement of holes from one atom to another constitutes charge motion.

Polymers and ceramics are good electrical insulators, quartz (silica) being known as the best insulator. Selection of insulators for packaging is more based on their dielectric constant and dielectric loss, as discussed in the next section.

Permittivity and Loss Tangent

When an electric field is applied across a dielectric or an insulator, the positive charges are displaced towards the negative end of the electric field and vice versa. This displacement induces polarization, and hence a higher electric flux density inside the material, compared to the field created in a vacuum by the same electric field. Electrical flux density is a measure of the electric field strength and charge distribution inside the material. Relative permittivity or dielectric constant is the ratio of the flux density in the material to that in the vacuum. The increased polarization or flux density also manifests as higher charge storage capacity or capacitance inside the material.

$$\text{Capacitance} = \frac{\text{Charge stored in a material}}{\text{Applied voltage}}$$

$$\text{Permittivity} = \frac{\text{Capacitance of a material}}{\text{Capacitance of vacuum}}$$

Macroscopically, polarization can be estimated by measuring the capacitance of the material. The dielectric constant is the most important material property in microelectronic packaging, because in electrical interconnections, the propagation delay is related to the dielectric constant of the propagation medium. The velocity of propagation (v), or its inverse, the cable-time delay, is dependent only on the dielectric constant (ε_r):

$$v = \frac{c}{\sqrt{\varepsilon_r}}$$

where c is the velocity of light. Polarization occurs due to different mechanisms, as described in Table 18.1. With increasing frequency, certain polarization mechanisms relax, and hence the dielectric constant decreases as the resonant frequency for that mechanism is reached. The typical variation of permittivity with frequency is shown in Figure 18.5.

The dielectric constant values for important materials are compiled in Figure 18.3. Many polymers have dielectric constants in the range of 2–4. Electronic polarization is

Samsung, EX1028, p. 367

TABLE 18.1 Classification of polarization mechanisms.

Polarization	Mechanism	Frequency Response	Materials
Electronic	Electron cloud displacement	At all frequencies (even optical) up to 10^{15} Hz	All materials; significant in inert gases
Ionic	Displacement of ions	Up to lattice vibration frequency 10^{13} Hz	Ionic crystals
Orientation	Dipole orientation	Intermediate frequencies (audio/radio) up to microwave 10^{6} Hz	Polar liquids
Space charge	Accumulation of charges at interfaces	Very low frequencies 10^{3} Hz	Electrodes, multiphase materials

the chief mechanism in these materials. Inorganic solids with dielectric constants in the range of 5–10 show some ionic polarization, too. Certain ionic crystals show polarization even in the absence of a field (spontaneous polarization). These are called ferroelectric materials. The polarization results from the asymmetric crystal structure in these ionic compounds, where certain ions are slightly displaced from their electrically neutral positions. Ferroelectric behavior is seen below a certain temperature, referred to as the Curie temperature, above which the material becomes paraelectric. The transformation from paraelectric to the ferroelectric phase is associated with a crystallographic change from a symmetric to asymmetric phase. In case of barium titanate, the transformation occurs from a cubic to tetragonal phase, leading to ferroelectricity. The ionic displacement can generally occur only in certain crystallographic directions. The dipoles exist as spontaneously polarized regions called domains. The domains can have the same possible crystallographic directions as the dipoles that constitute them. The existence of domains and the domain size are guided by the minimization of total free energy.

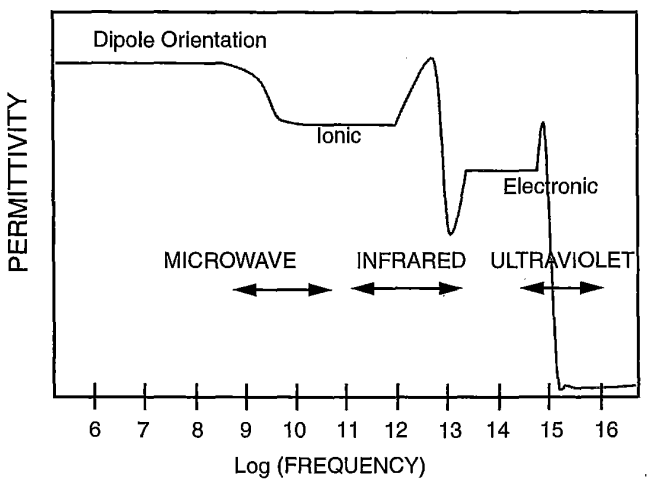


FIGURE 18.5 Polarization relaxation phenomena in dielectrics.

Samsung, EX1028, p. 368

Total free energy = polarization energy of all domains + domain boundary energy

The domains have a high degree of polarizability. Hence, they easily reorient in the direction closest to that of the electric field, leading to a high permittivity or capacitance.

Ferroelectric materials show dielectric constants on the order of 1,000–20,000 in bulk form. The dielectric constant of ferroelectrics is dependent on the grain size, and hence, most of the nanostructures and submicron ferroelectric thin films may not show such high-dielectric constants. The dielectric constant of $BaTiO_3$ is found to be highest when the grain size is about 1 micron. For grain or crystal sizes below 100 nm, the dielectric constant is not expected to be more than a few hundred.

Dielectric materials for electronic applications can be classified as low-K or high-K type. Materials that form the medium separating transmission lines should have low dielectric constant, and hence, polymers and low-K ceramics are favored. Devices for decoupling capacitors, filters, etc., need high-K materials. More details on the device and electrical characteristics can be obtained in Chapter 11.

When a sinusoidal *electromotive force* (EMF) is applied at low frequencies, the polarizations inside a dielectric material develop completely before the field reverses. The time-variant polarization is equivalent to an alternating current, which leads the EMF exactly by 90 degrees. The polarizability is greatest at those frequencies at which the dipoles can oscillate in resonance to the field. However, at higher frequencies, when the polarization can no longer keep up with the EMF, the current can only lead with a phase difference of $90 - \delta$. In other words, the current has a component that is in-phase with the voltage, hence resulting in a dissipation of energy. Physically, the energy loss can be viewed as resulting from the molecular friction that opposes the molecular motion, leading to an energy loss during polarization. Dielectric loss is a measure of the electrical energy dissipated per cycle. The losses are characterized by including a resistive element along with the capacitance. The loss angle δ is the ratio between resistive and capacitive current. Dielectric loss is an important electrical parameter for nonperfect dielectric materials, particularly in high-frequency applications. Assuming we have the circuit model of Figure 18.6 with the parallel resistance R_p as the only parasitic, the loss angle may be written as:

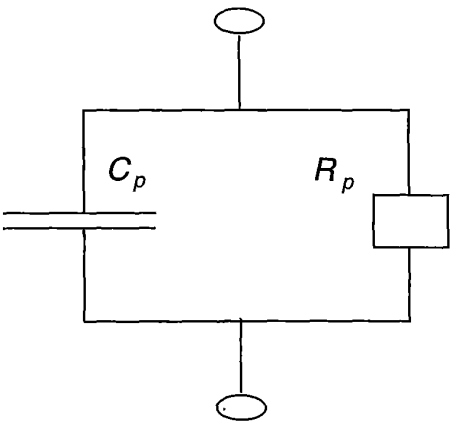


FIGURE 18.6 Circuit model for dielectric loss.

Samsung, EX1028, p. 369

$$\tan \delta = \frac{\left[\frac{1}{R_p}\right]}{[wC_p]} = \frac{1}{wR_pC_p} = \frac{1}{Q}$$

where R_p = the ohmic resistance
$1/wC_p$ = the capacitive impedance at frequency w
Q = the ratio between stored energy and dissipated energy in a capacitor

The polarization and loss can be more generically expressed using a complex dielectric constant. The real portion of the dielectric constant is responsible for the polarization of the material, because it represents that portion of the dielectric constant, which results from the oscillations being in-phase with the field. The imaginary portion of the dielectric constant is responsible for the energy loss or the power loss. Tan δ is related to the complex dielectric constant with the real (ε') and imaginary parts (ε'')

$$\varepsilon = \varepsilon' + j\varepsilon''$$

by

$$\tan \delta = \frac{\varepsilon''}{\varepsilon'}$$

Dielectric loss is a significant contributor to the voltage transmission loss (V_{out}), particularly at higher frequencies, or with longer and narrow conductor or interconnects. This can be described using the following equation:

$$V_{out} = V_{in} \exp\left(-\pi \tan \delta \frac{fl}{v}\right)$$

where f is the frequency, l is the length of interconnect, and v is the velocity of signal propagation in the medium. Hence, tan δ is an important material parameter to be considered in designing high-frequency structures. It is important to note that there are also other contributors to losses, such as resistance loss and skin-effect loss. The typical losses for dielectrics are shown in Table 18.2. It can easily be seen that polymers suffer high losses, and hence inorganic materials are a favorable choice for high-frequency applications.

TABLE 18.2 High frequency losses of typical dielectrics (ceramics, glasses and polymers).

Material	Quartz	BeO	Titanates	Alumina	AlN	Cordierite	PTFE	Cyanate Ester	Polyimide	BT	FR-4
Dielectric constant	3.75	6.6	20–10,000	9.2	8.3	4.9	2.05	3.8	4	4	4.5
Tan $\delta(10^{-4})$	1.5	2.0	1.5–3	2	3–10	10	4	90	150	150	260
	←					→	←				→
	Ceramics						Polymers				

PTFE, polytetrafluoroethylene; BT, bismaleimide triazine.

Samsung, EX1028, p. 370

18.2.2 Thermal Properties

Thermal Conductivity

The amount of heat transferred through a material per unit of time, denoted as heat flux Q, is proportional to the temperature gradient (dT/dx). The ratio of heat flux and temperature gradient is termed thermal conductivity:

$$\frac{Q_x}{A} = -K_x \frac{dT}{dx}$$

A is the area of conduction medium and x is the direction of heat flow.

This is analogous to Ohm's Law for electrical conduction. Thermal conductivity is a tensor with nine components, because it relates the flux (a vector) to the divergence of temperature. It is expressed as W/mK. Heat flows from a high temperature to low temperature region, and therefore, a minus sign in the above equation makes the conductivity a positive number. The anisotropy in the lattice structure, or the anisotropic distribution of different phases in a composite, correspondingly results in an anisotropic thermal conductivity.

Thermal conductivity is a very important property for dissipation of heat from an IC. The heat evolution from an IC may range from 1 watt for portable products to as much as 150 watts per chip for high-performance computer applications. The thermal conductivity is a measure of the ease with which heat can flow through a material. High thermal conductivity materials allow for easy heat dissipation from the chip, and hence, control the chip temperature.

There are two contributions to heat transfer by conduction. In nonmetallic materials, heat transfer occurs through lattice vibrations or oscillations. The vibrational energy levels are quantized, implying that only certain modes are allowed. The lattice vibration quanta, which also describe the field of ionic/atomic displacements, are called phonons. Similar to electrons, phonons can be described in terms of band diagrams. However, unlike electrons, the number of phonons is not conserved. Phonons are created by rising temperature and eliminated by lowering the temperature. The hot end of a body possesses more phonons that drift to the cold end.

If the interatomic bonds are visualized as springs, it can be easily seen that the transfer of lattice vibrations is easier in a stiffer lattice with smaller atoms. Phonons are more easily scattered in lattice with atoms of different size. For example, covalent solids (SiC) with smaller atoms tend to conduct heat better than ionic solids (ZrO_2). The heat conduction is also faster in simpler lattices because of minimal scattering, for example diamond. Impurities and defects scatter the phonons easily and adversely affect the conductivity. Aluminum and silicon nitride, which conduct heat faster in pure and single-crystal state, become poorer conductors because of the defects and impurities that are added to aid sintering. Polymers and glasses are relatively poor conductors because of their amorphous structure, resulting in enormous scattering of phonons. The other major contribution of heat conduction comes from electron movement and is very similar to electrical conduction. The total contribution to the thermal conductivity (K) can be semi-empirically expressed as:

$$K = A \frac{\sqrt{\text{bond stiffness}}}{\sqrt{\text{atomic mass}}} + B \exp\left(\frac{-E}{RT}\right)$$

where the term incorporating bond stiffness represents the contribution of phonon

Samsung, EX1028, p. 371

transfer, and the activation term quantifies the electronic contribution. Thermal conductivities vary from 0.1–2000 W/m K for different materials. The materials and their properties are covered in Figure 18.3.

The transport process of electrical current and thermal energy is the same for metals. This indicates a close relationship between the electrical and thermal conductivities. This relationship, called Wiedemann–Franz Law, states that the ratio of thermal conductivity to electrical conductivity is proportional to the temperature, and that the ratio is independent of the metal. The relationship is written as:

$$\frac{K}{\sigma} = T\frac{\pi^2}{3}\left(\frac{k}{e}\right)^2$$

where k is the Boltzmann constant and e is the electron charge. Inserting the fundamental constants we obtain:

$$\frac{K}{\sigma T} = 2.45\ 10^{-8}\ \frac{\text{Watt Ohm}}{°\text{C}^2}$$

Coefficient of Thermal Expansion

The dimensional change that occurs during heating or cooling of a material is characterized by its coefficient of thermal expansion (CTE). CTE can be defined as the dimensional change that occurs per unit rise in temperature per unit length:

$$\text{CTE} = \frac{dl}{l\ dT}$$

As described before, an electronic package consists of different materials held together, processed and operated over a range of temperatures. CTE mismatch between different materials induces high stresses, and is hence a severe concern for thermomechanical reliability. The film stress at the interface between two materials with CTE mismatch ($\Delta\alpha$) and temperature gradient (ΔT) can be estimated as:

$$\sigma_f = \frac{\Delta\alpha\ \Delta T\ E_f}{(1 - \nu_f)}$$

where E_f is the Young's modulus of the film. High stresses from thermal mismatch at interfaces cause interfacial fracture, delamination of films, and fatigue fracture, and are a major concern for thermomechanical reliability and yield. When the interface is strong enough to withstand the stress inside the film, the stress will cause the beam to bend. From the analysis of Stoney, the curvature is given as:

$$\sigma_f = \frac{E_s t_s^2 K}{6(1 - \nu_s)t_f}$$

where t_s and t_f are the thickness of the substrate and film, respectively, K is the curvature of the composite and E_s is the substrate modulus. ν_f and ν_s refer to the Poisson's ratio of the film and substrate.

The dimensions of CTE are 1/°C. CTE values typically range from -1×10^{-6} to 200×10^{-6}. Hence, the exponent is excluded by using the units ppm/°C. Most materials expand on heating, resulting in a positive CTE. Certain materials such as graphite and

Samsung, EX1028, p. 372

lithium aluminosilicates (spodumene) have negative CTE in some directions. CTE behavior can be understood from the interatomic potential and its variation with the interatomic spacing. At higher temperatures, atoms have higher vibrational amplitude. Because of the asymmetry in the interatomic potential, the mean atomic spacing increases with temperature and results in thermal expansion.

The crystal structural anisotropy in many materials result in anisotropic thermal expansion properties. It is generally observed that stiffer bonds (covalent ceramics such as SiC, AlN) and loosely packed structures (silica) lead to a lower thermal expansion coefficient. Polymer chains are held by relatively weaker bonds and show high coefficient of expansion. Above the glass transition temperature, the CTE of polymers suddenly rises to higher values. CTE values for most electronic polymers are obtained from *thermomechanical analysis* (TMA). The CTE values for different materials are illustrated in Figure 18.3.

Composites and glasses are generally tailored to tune their CTE for specific applications. An important success story in this regard is the design of a glass ceramic that has CTE exactly matching that of silicon to eliminate stresses in the solder joints that connect the chip and the glass ceramic. Polymers for underfill applications are typically loaded with silica in order to reduce the CTE mismatch between the IC chip and the PWB, thereby improving the efficiency of underfills.

EXAMPLE 18.1

Consider a package as shown in Figure 18.7. We have:

$$\delta = d\,\Delta T\,\Delta\alpha$$

where δ is the relative displacement, d is the distance to the neutral point and ΔT is the temperature difference and $\Delta\alpha$ is the difference between the thermal coefficients of temperature of the integrated circuit and the substrate. In this case, the shear strain is:

$$\gamma = \delta/h = d\,\Delta T\,\Delta\alpha/h$$

where h is the stand-off height.

Assume that we have a Si-chip with a thermal expansion of coefficient of 2.6 ppm/°C assembled with eutectic soldering flip chip technique on an FR-4 board with a thermal expansion of coefficient of 16 ppm/°C. The stand-off height after soldering is 50 μm. The chip edge length is 20 mm and the temperature varies from −40 to +125°C. Calculate the total shear strain during the temperature cycling in this interval for the solder joint. Assume the ideal case where all the deformation is taken up by the solder joint.

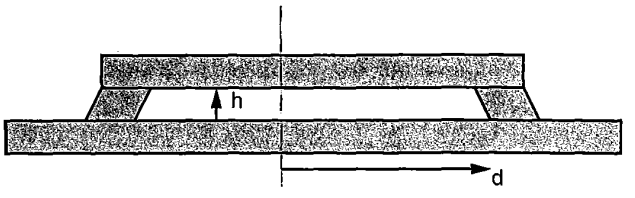


FIGURE 18.7 Schematic of a Si-chip interconnected to a substrate using solder interconnect (Example 18.1).

Samsung, EX1028, p. 373

Solution

$$\Delta\alpha = 16 - 2.6 = 13.4 \text{ ppm/°C} = 13.4 * 10^{-6} \text{ ppm/°C}$$

$$\Delta T = 125 - (-40) = 165 \text{ K}$$

$$d = 20/2 = 10 \text{ mm}$$

Therefore,

$$\delta = 10 * 165 * 13.4 * 10^{-6} = 0.02210 \text{ mm} = 22.11 \ \mu\text{m}$$

The shear strain is:

$$\gamma = \frac{22.11}{50} = 0.44$$

Glass Transition Temperature

The glass transition temperature (T_g) characterizes the transition of an amorphous material from a brittle state to a rubbery state. Crystalline materials do not show this transition. Glass transition is manifested by drastic changes in many of the material's physical properties, such as volume and modulus as shown in Figure 18.8. Polymers are subjected to severe thermomechanical stresses and dimension instabilities, related to warpage, above the glass transition temperature. Hence, T_g is an important material property for electronic packaging. The thermomechanical reliability of a package can be severely enhanced by using a high T_g polymer. The glass transition temperature is characterized from thermomechanical analysis (TMA) and *dynamic mechanical analysis* (DMA) by monitoring the dimensional, or modulus, changes with temperature.

Molecular motion is the most important factor in determining T_g. Molecular chains tend to fluctuate because of thermal energy. The fluctuations are characterized by a distribution of relaxation times that are temperature-dependent. Below T_g, the molecular motion is hampered (relaxation times are very high) and the volumetric changes lag behind the temperature, resulting in an excessive free volume. Above T_g, the molecular chains possess sufficient thermal energy to fluctuate between different conformations

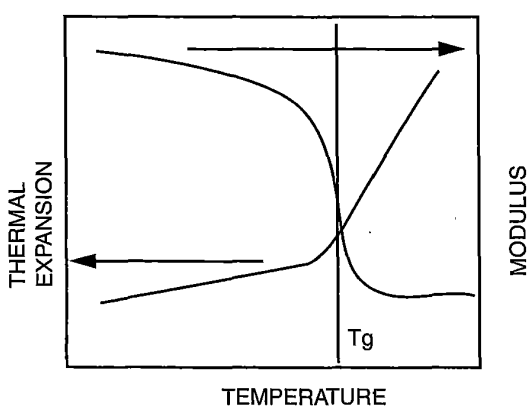


FIGURE 18.8 Glass transition temperature phenomena in polymers.

Samsung, EX1028, p. 374

(rotation of side groups, etc.), thereby wiggling independent of other segments. Some materials show two distinct relaxation times, one corresponding to the main chain and the other to the side-chains. The relaxation of main chains gets frozen at a higher temperature (α-transition) compared to that of the side chains (β-transition). Hence, these materials show two glass transition temperatures. Though glass transition temperature is rate-dependent, it is often represented as a second-order thermodynamic transition.

For a molecule with chain length much greater than 25–30 carbon atom segments, T_g is independent of molecular weight. The main polymer structural parameters that determine the T_g are:

1. The elements and bond types that determine the intrinsic mobility of the main chain
2. Shape and rigidity of the side groups that affect the chain flexibility via the side-group steric hindrance
3. The free volume between the side groups
4. The side-group regularity and the symmetry along the chain which assist efficient chain packing leading to microcrystalline regions that reduce the free volume and increase T_g
5. Side groups that provide sites for the formation of covalent, ionic or coordination cross-links

Linear polymers and long flexible chains with low hindrance to motion have very low T_g. Bulky side groups decrease the mobility of the chain, thus raising the T_g. For example, aromatic side groups drastically raise T_g compared to aliphatic side groups. The presence of polar groups and hydrogen bonding raise the T_g. These are some of the tools a polymer chemist can play with to modulate the T_g.

Among the commercially important polymer materials, polyimide has the highest T_g (above 350°C), while silicone has the lowest. The most widely used polymers in packaging are epoxies, which have T_g in the range of 120–200°C. The glass transition temperature directly depends on the degree of cure.

18.2.3 Mechanical Properties

Materials in electronic system packages are always subjected to large forces. The forces may be caused by flexure and impact during fabrication or actual use, or from the internal thermal gradients and differential expansion properties at the interface with other materials. These forces can result in stresses that can be large enough to cause mechanical failures such as die cracking, interfacial delamination, distortion, warpage, etc. The thermomechanical design of a package cannot be thorough without a complete knowledge of the mechanical behavior of packaging materials.

Young's Modulus

Materials deform in response to an applied force. The deformation may be permanent or temporary, time dependent or time independent, and is classified accordingly. The force deformation relationships are expressed in terms of stresses and strains. Stress is the force per unit area. Strain is the dimensionless deformation (deformation per unit length). These are tensors and need nine components for their complete description. Materials typically

Samsung, EX1028, p. 375

have a linear stress–strain curve until a critical stress is reached (elastic limit), after which plasticity results. Young's modulus is the stress to generate unit strain inside a material, and is generally measured as the slope of the linear portion of the stress–strain curve. The stiffness of a material is its ability to resist elastic deformation or deflection upon loading, and is characterized by Young's modulus and material geometry. Brittle materials such as ceramics fail within the elastic limit and have no plastic deformation. Metals and polymers show considerable plastic or nonrecoverable deformation.

Young's modulus is directly related to the interatomic bonding. The range of Young's moduli of different materials is shown in Figure 18.3. Materials with stronger bonds (covalent) show higher moduli. Transition metals have higher moduli (200 GPa) than alkali metals because of the partial covalent character of their bonds. The second and third transition metals have even higher moduli (even up to 600 GPa) but have very high densities. The anisotropy in the crystal structure results in anisotropic mechanical properties, as seen in graphite.

Polymers have low Young's modulus and elastic limit, because the polymer chains are bonded by relatively weaker secondary bonds. The value of the modulus is dependent on the nature of secondary bonding, presence of bulky side groups, branching in the chains and cross-linking. Unbranched polyethylene has a modulus of 0.2 GPa, while polystyrene with bulky side groups has a modulus of 3 GPa and three-dimensionally cross-linked rubber has a modulus of 3–5 GPa. Rubber-like polymers referred to as elastomers, do not follow Hooke's Law (linear stress–strain relation). However, they behave elastically in the sense that they return to original dimensions after the removal of the load. The high elastic strains are caused by the uncoiling (unkinking) of the polymer chains. Alignment of polymer chains is a method of increasing stiffness. Most deformation that occurs in polymers is nonrecoverable and plastic. The nonlinear portion of the stress–strain curve is not well characterized, making stress analysis of polymer-based materials relatively complex. Stress–strain behavior depicting the typical elasticity, elastic limit, plasticity and fracture stress in metals, polymers, ceramics and glasses is shown in Figure 18.9.

The limitations of Young's modulus in available materials can be easily seen in the above discussion. Polymers have low modulus while ceramics with high modulus are brittle, and stiffer metals are generally heavy. In composite materials, an attempt is made to increase modulus without increasing brittleness and weight. The most popular com-

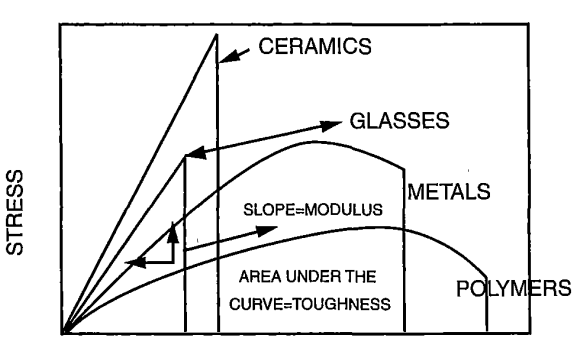


FIGURE 18.9 Stress–strain behavior of materials.

Samsung, EX1028, p. 376

posite materials in electronic packaging are silica glass fibers embedded in an epoxy resin matrix for making printed wiring boards. A simple estimate for the modulus of a composite can be made from the rule of mixtures. If a composite has continuous fibers in a direction, the rule of mixtures predicts the modulus in that direction of the composite as:

$$E = V_f E_f + V_m E_m$$

This equation generally overestimates the values and is valid for straight fibers with good adhesion between the matrix and the fiber. The modulus in the other direction can be estimated as:

$$\frac{1}{E} = \frac{V_f}{E_f} + \frac{V_m}{E_m}$$

When the fibers are short and discontinuous, a series model is used to predict the modulus. For chopped fibers, the modulus lies in between those predicted by the parallel and series models.

Mechanical failures are said to occur when the material fails under a load. If the failure leads to breaking into two or more pieces, it is referred to as fracture. Fracture occurs at loads far less than the load required to break the interatomic bonds. This is due to the presence of flaws and defects inside the material. Fracture mechanics describe the progression of flaws (cracks) through a stressed material as a function of time, geometry and environmental conditions. The plastic deformations in metals limit the crack-extension and therefore, metals have higher toughness compared to ceramics. Under cyclic loading, the failures occur at a still lower load level, referred to as fatigue fracture. At high temperatures, fracture is assisted by creep and is referred to as creep fracture.

Fatigue is the most common failure mode in IC board interconnections and is discussed here and in Chapter 22 (Fundamentals of Packaging Reliability). When an active IC device dissipates heat to its surroundings during operation, differential thermal expansion generates stresses in the interconnection structure. These stresses produce instantaneous elastic and plastic strains in the material joint. Fatigue fracture occurs by crack initiation and propagation. Materials with fine grain size have a better fatigue resistance. Solders after solidification are typically heat-treated to obtain finer microstructures that improve the fatigue resistance. Solders with nanograin size may hence gain popularity. The mechanical properties of solders change strongly over the normal temperature range of operation. This is because of the low melting point of solder. Below 0°C, the solder is brittle, whereas above room temperature it is plastic. Above room temperature, mechanical built-in stresses that result from thermal mismatch with the other materials will relax and vanish over time. However, cyclical plastic deformations change the grain structure, weaken the solder fillet and can lead to fatigue. The time for solder fracture depends on relative deformation (strain), temperature and frequency of deformation. A simplified relationship is given by the Coffin-Manson's formula:

$$N^{0.5} \times \delta_p = \text{constant}$$

where N is the number of cycles until fracture and δ_p is the relative deformation amplitude. In practice, the relations are more complex. This can be explained by the temperature changes that give rise to the deformations, and the fact that material is plastic part of the time and elastic part of the time.

Samsung, EX1028, p. 377

EXAMPLE 18.2

Assume that the lifetime after the temperature cycling of a eutectic solder joint can be explained using the Coffin-Manson relationship:

$$N = k(\gamma_p)^{\alpha}$$

where the constant k is approximately equal to 400 cycles and α equal to -0.5. γ_p is the total plastic shear strain. Calculate the lifetime in terms of the number of temperature cycles to failure under ideal case, for the case in Example 18.1 when all the elastic strain is neglected.

Solution

The lifetime in terms of cycles to failure in this case is:

$$N = 400 * 0.2211^{-0.5} = 851 \text{ cycles}$$

EXAMPLE 18.3

Real Service Lifetime Estimation

Consider that the above case is valid for solder joints in a car engine. Two cycles corresponds to one day's use of the car (start in the morning from the home and turn off at the office for one cycle, start in the afternoon at the office and turn off at the home for the second cycle). Estimate the real service life of the car before it fails.

Solution

The real service lifetime is:

$$851/2 = 425 \text{ days before it is time for repair.}$$

EXAMPLE 18.4

In a low-cycle fatigue test of a solder joint, the logarithmic fatigue cycle as a function of logarithmic plastic strain is given by Figure 18.10. Assume that the solder joints obey the Coffin-Manson frequency-modified equation:

$$N_f * \nu^{(K-1)} = \theta * (\Delta\varepsilon)^{\eta}$$

Determine the constants θ and η such that the frequency is equal to 0.3 s^{-1} and $K = 0$.

Solution

Logarithm of the Coffin-Manson frequency-modified equation yields:

$$\text{Log } (N_f * \nu^{(K-1)} = \log \theta + \eta \log(\Delta\varepsilon)$$

As $K = 0$, we have:

$$\text{Log } N_f = \log (\theta * \nu) + \eta \log(\Delta\varepsilon)$$

Therefore, the intercept is equal to log $(\theta * \nu)$, and the slope is equal to η.

In this case, we have:

$$\log (\theta * \nu) = 0$$

$$\theta * \nu = 1$$

Since $\nu = 0.3$, we have $\theta = 3.33$ cycles/s.

Samsung, EX1028, p. 378

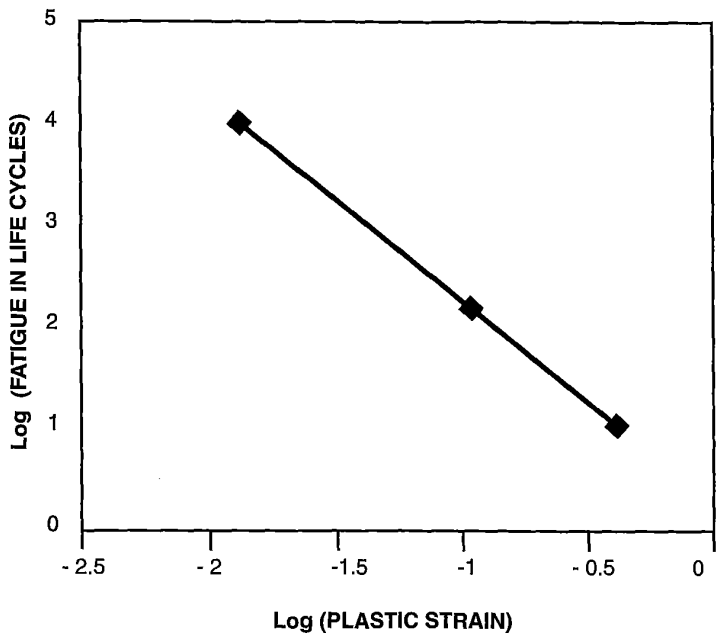


FIGURE 18.10 Fatigue cycles for failure versus plastic strain on a log-scale (Example 18.4).

Simple mathematical consideration gives $\eta = -2$. Hence, the final equation becomes:

$$N_f = (\Delta\varepsilon)^{-2}$$

18.2.4 Chemical Properties

Surface Tension and Wetting

All materials in the solid or liquid state have energy associated with their surfaces. The energy arises from the unsaturated bonds on the surface. The unit of surface tension is energy per area (J/m^2), or force per length (N/m). This surface energy depends on the surface characteristics of the material (contamination, absorbed molecules, etc.) and the surface energy of the surrounding medium. The surface energy of the solid is denoted by γ_s, and the surface energy of the liquid is denoted by γ_l. However, at the interface between solder and solid, the surface energy will depend on the mutual interaction. The surface energy at this interface is therefore denoted by γ_{sl}. The degree of wetting by the molten solder will depend on the relative magnitudes of the surface energies for the solder and the substrate metallization. Wetting is relevant in metal joining processes such as soldering. In a practical soldering process, the solid metal and the liquefied solder will be enclosed by gas. The combined surface energies of the exposed solid surface, the exposed liquid surface and the solid-liquid interface will tend to be minimized. The minimization of the free energy can be described by Young's equation:

Samsung, EX1028, p. 379

$$\gamma_{sl} + \gamma_l \cos\theta = \gamma_s$$

In the case of a partial wetting condition (Figure 18.11), the contact angle θ between the drop and the solid reveals a lot of information. A lower contact angle implies better wetting. In the case where the solid metal is contaminated by oxides or other organic residue materials, the surface energy of the solid will be much less than the surface energy of the solder. This will cause the solder to minimize its surface by contracting into a spherical shape on top of the solid. Various fluxes are added to the solder in order to clean the solid surface from its contaminants and hence improve the wetting.

Adhesion

Adhesion between dissimilar surfaces such as metals/polymers or ceramic/polymers is generally caused by weak chemical forces such as van der Waals forces. The van der Waals forces lack the stoichiometry and directionality of the chemical forces formed by covalent and ionic bonds. During the formation of a van der Waals bond, the electron distribution is merely perturbed. Pure physical forces such as mechanical interlocking can also assist adhesion. This type of adhesion is dependent on surface morphology and surface area. Metals and polymers are typically roughened in order to increase their adhesion. Adhesion can also be based on other types of chemical forces such as acid-based interactions. The strength of interfacial bonds between polymers and inorganic substrates can be appreciably enhanced when polymers are acidic and substrates are basic or vice versa. The interaction has two contributions: an increased thermodynamic work of adhesion, resulting from large exothermic reactions at the interface; and an increased tensile strength, resulting from electrical charge injection into the polymer from the substrate.

Epoxy is widely used for building layers on a substrate because of its good adhesion to metals and ceramics and its low cost. Many reliability problems at the interfaces arise from the lack of proper adhesion. In order to appreciate the adhesion problem, a typical example is discussed here. Adhesion between copper and epoxy is critical in the manufacture of multilayered printed wiring boards and the encapsulation of semiconductor chips. Copper in its pure form exhibits poor adhesion to polymer substrates because natural copper oxide is mechanically weak and difficult to be wetted by the adhesives. To improve adhesion, surfaces are treated with oxidizing agents. To improve the bond durability in acidic environments, coupling agents such as cyclic imidazole and benzotriazole derivatives are used. These compounds have both active protons to form a com-

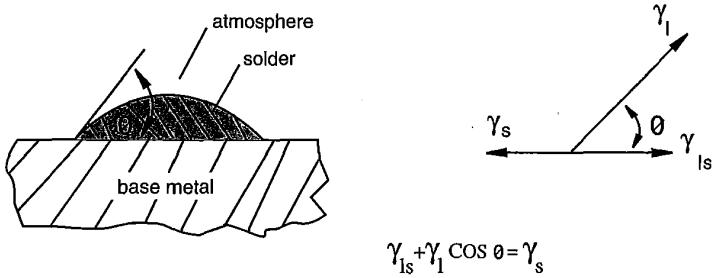


FIGURE 18.11 The definition of the contact angle θ.

Samsung, EX1028, p. 380

plex with the cuprous hydroxide on the copper surface, and functional groups such as amine and carboxylic acids to interact with the epoxy resins.

18.3 MATERIALS PROCESSING

An electronic package is a multilayered structure with thickness of layers ranging from 2 microns to more than 100 microns. Conventional ceramic and PWB packaging processes that are typically used to build layers of 20–100 microns are classified as thick-film processes. The main processes are doctor blading and lamination of green sheets or epoxy/glass cloths and screen-printing. These processes are used to make the single-chip packages or multichip or multilayered substrates. Thin-film processes are used to build the subsequent dielectric layers, conductor and passive patterns. The typical thin-film processes are polymer solution coating, plating, sputtering, evaporation, *chemical vapor deposition* (CVD), etc., that resemble the ones that are used in IC fabrication. The term "thin film" refers to the technology used to form the structure rather than the actual thickness of the structure. A generic classification of thick-film and thin-film processes is shown in Figure 18.12.

18.3.1 Thick-Film Processes

Ceramic

Ceramics are generally regarded as high-performance materials because of their hermiticity, high reliability, low CTE and low losses. Typical properties of ceramics used for packaging are shown in Table 18.3. Single-chip ceramic packaging exists in various forms such as *dual-in-line packages* (DIPs), chip carriers, flat packs and pin grid arrays. These packages are generally made either by the green sheet or dry-pressing process. The typical properties of related ceramics are classified in Table 18.3. For providing complex circuitry with many interconnections, particularly with multichip modules, a single layer of ce-

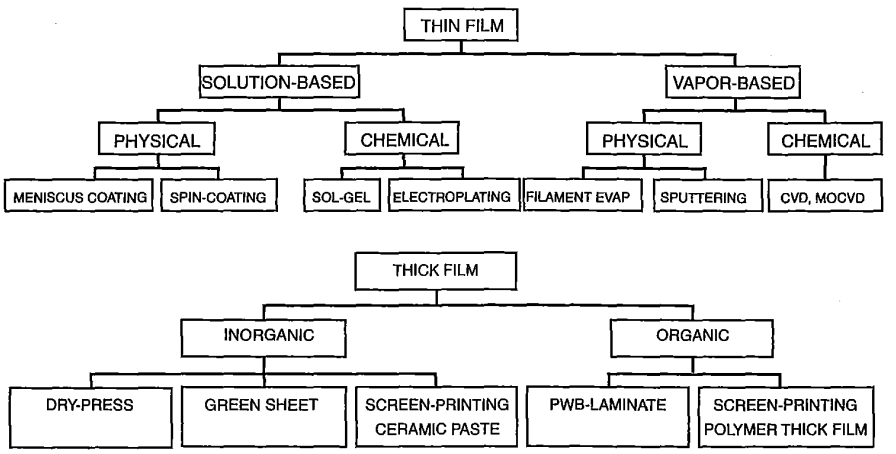


FIGURE 18.12 A general classification of thick- and thin-film processes.

Samsung, EX1028, p. 381

TABLE 18.3 Properties of some important ceramic materials used for hermetic packaging of ICs.

Ceramic	Relative Permittivity	Thermal Conductivity (W/m K)	Thermal Expansion Coefficient (ppm/°C)	Loss Tangent tan δ (10^{-4})	Elastic Modulus GPa
Al_2O_3	9.8	20	7	2	350
AlN	9	230	4.1	3–10	380
SiC	40	270	3.7	500	380
BeO	6.8	260	5.4	4–7	345
LTCC	4–7	5	3–5	2	150

ramic will not be sufficient. Hence, multilayered substrates are employed for this. A multilayered ceramic consists of a cofired stack of ceramic green sheets on which metal wiring is printed and vias are punched for interlayer connections. Metallization of the green sheet is accomplished by extruding the metal paste through a nozzle as the nozzle traverses a metal mask in contact with the green sheet. This process is also used to fill the vias. A schematic of the whole process of making cofired multilayered substrates is illustrated in Figure 18.13. The tapes shrink by 12–20% during sintering, which has to be accounted for by the circuit features such as vias on the green sheet. The firing temperature dictates the metal wiring used for the circuitry. The sintering temperature of the powder has to be less than that of the melting temperature of the metal used for conductors. Pure alumina sinters at above 1500°C and hence requires W/Mo (tungsten/molybdenum) metals for the wiring. These are generally referred to as *high-temperature cofired ceramics* (HTCC) and are losing popularity in the industry because of their high temperature processing. In addition, alumina has a high dielectric constant and high CTE, further limiting its applications. Multilayered capacitors are made from a similar process using $BaTiO_3$.

Glass ceramics are crystalline materials that are made starting from glasses. They possess valuable properties of both crystals and glasses. They retain their strength at higher temperatures like crystals, and the compositions can be controlled to tune the properties, such as CTE and dielectric constant. Glass ceramics can be sintered at relatively lower temperatures (<1000°C) and hence, more conductive Ag-Pd pastes are widely used for the interlayer circuitry. These materials are referred to as *low-temperature cofired ceramics* (LTCC). Glass ceramics typically consist of glass forming compounds, mixed with alumina or silica. The glassy phases melt at low temperatures, completely wet the alumina and aid in sintering. Alumino-silicates of lead, lithium or boron, forming as high as 40 volume percent of the total dielectric material, are typically added to alumina, thereby lowering the sintering temperature to 850°C. They offer rugged, hermetic and reliable packaging with the ability to have multiple layers and high wiring density. Furthermore, miniaturization and improved device performance can be achieved by integrating passive components such as resistors, capacitors and inductors. Multichip modules for high-speed supercomputer applications are typically made from these materials and processes. Components made from LTCC are receiving wide attention for high-frequency/RF applications in the telecommunications industry.

Samsung, EX1028, p. 382

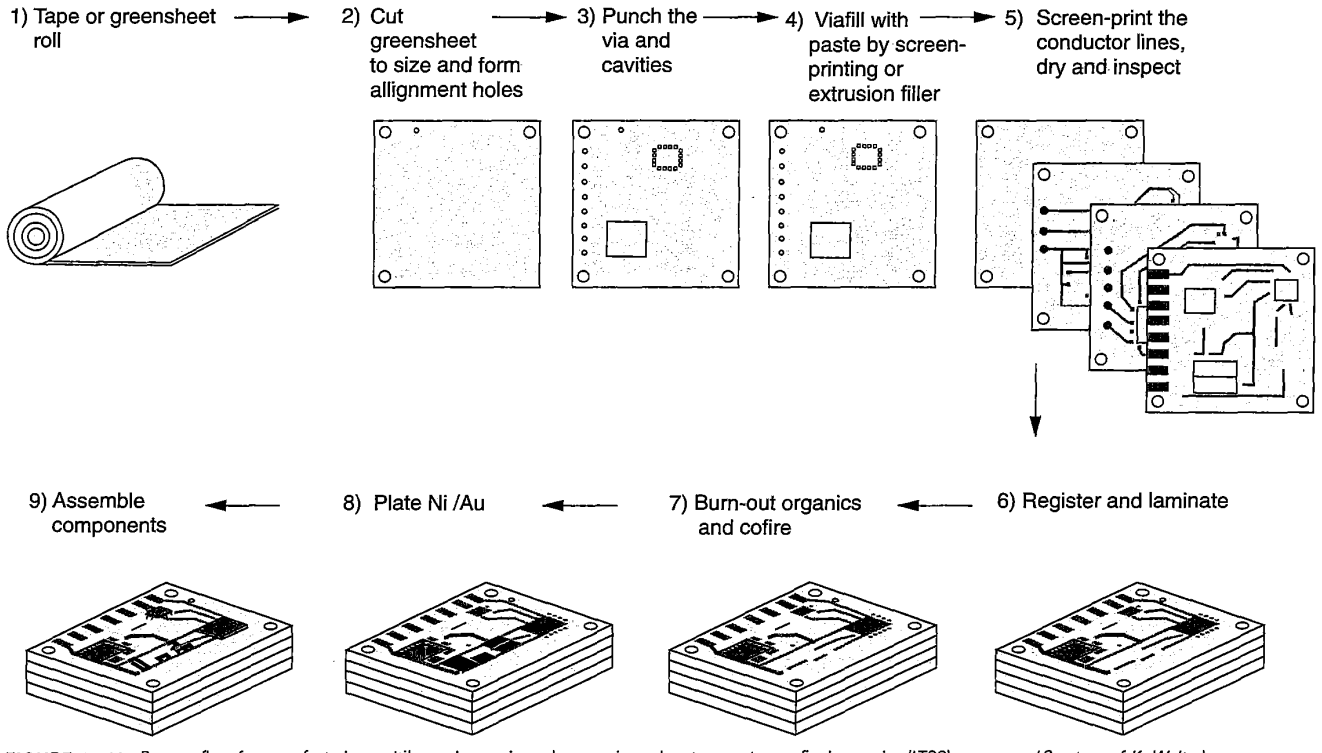


FIGURE 18.13 Process flow for manufacturing multilayered ceramic packages using a low-temperature cofired ceramics (LTCC) process. (*Courtesy of K. Wolter*)

MEMS23206

Samsung, EX1028, p. 383

Ceramic processing is generally synonymous with powder processing, because the raw material that makes the final product is in powder form. Parts with smaller and simpler shapes are made by flowing the powder into a metal die set with the desired shape and pressing the powders to make a powder compact. The powder is mixed with binder and lubricants to assist the pressing process. For making large, thin and flat sheets that are not suitable to be made by dry-pressing, the ceramic powder is dispersed in a suitable solvent to achieve sufficiently low viscosity (good flowability) and cast to give a desired shape. The solvent can be aqueous (water-based) or nonaqueous (methyl ethyl ketone/ethanol, etc.). Aqueous dispersions are gaining popularity because they are more environmentally friendly. However, nonaqueous suspensions dry faster, result in less defects such as bubbles, and are less sensitive to minor changes in the formulation. The powder + solvent + additives together constitute what is generally referred to as slip, slurry or suspension. The slip is spread on a flat surface and the solvents are allowed to evaporate. One process of making sheets, known as tape casting, consists of flowing the dispersion under a blade onto a moving substrate. The blade levels the slurry, making it flat. The opening under the blade guides the thickness of the cast sheet.

The cast film is a pack of loosely held particles, and therefore does not have enough strength for subsequent handling. Various polymer additives are added to provide enough temporary strength for further processing. Polymers, which provide the necessary plasticity and strength, are termed as binders. Popular binders are polyvinyl butyral for nonaqueous systems and acrylate-based polymers for aqueous systems. The additives result in a smooth surface and also help the layers stick to each other when they are stacked and laminated together in a press. The thick film, when peeled off from the substrate, is referred to as a green sheet. Tape casting is widely used for large-scale fabrication of ceramic substrates and multilayer structures.

The flow behavior of the slurry has to be controlled for getting a good quality tape. Some of the suspension qualities that a ceramic engineer needs to ensure are good dispersion at high solids loading in the solvent (typically more than 50 volume percent), good wettability to the substrate, and few or no defects such as bubbles. After drying, the tape should be easily peeled off from the substrate.

In order to attain bulk ceramic properties, the green sheet must be heat-treated at high temperatures, typically ranging from 850–1800°C (850°C for glass ceramics, 1500°C for pure alumina and 1800°C for covalent ceramics such as AlN and SiC). This process is referred to as sintering or densification. The removal of binder (pyrolysis) by oxidation of carbonaceous materials further complicates the heat treatment process. During sintering, the particles coalesce to form a continuous network and porosity is removed. The driving force that causes sintering is reduction in surface energy, or correspondingly, the surface area. The surface energy manifests as a compressive force between particles in contact, hence causing their coalescence. The compressive force (sintering force or sintering stress) can be related to the surface energy in terms of the curvature of the neck (point where the surface meets the particle contact), contact angle at the neck, length of the interparticle contact, etc. The coalescence needs to be assisted by the movement of atoms from the particle contacts (grain boundaries) to the surface, causing the growth of interparticle contacts and reduction in porosity. Since long-range atomistic movements occur rapidly at high temperatures, sintering is a high-temperature process. The atomic movement is due to diffusion in crystalline solids or viscous flow seen in amorphous solids. Formation of liquid-phase expedites the sintering and lowers the sintering tem-

Samsung, EX1028, p. 384

perature. Atomic movements are significant at temperatures of 0.7 T_m (melting point), and sintering is generally done at these temperatures.

Sintering is a complex process where multiple phenomena take place. This is further complicated in liquid-phase sintering, where some additional processes take place including grain sliding and grain rearrangement, and dissolution of small grains into the liquid, which then precipitate on coarse grains, etc. However, the following things should be evident to the reader:

1. Smaller particles sinter faster and at lower temperatures. However, even for the smallest particles, temperatures above 850°C are quite common.
2. The addition of glasses lower the sintering temperature of alumina by forming a low-melting point and low-viscosity phases that aid faster sintering.
3. The elimination of porosity manifests as shrinkage in the final product.

Screen-Printing

In hybrid technology, chips are interconnected on a substrate, typically ceramic, with the capability of including passive components such as resistors, capacitors and insulators as an integrated part of the substrate. An example is shown in Figure 18.16. Thick-film hybrid technologies have packaging density between integrated circuit technology and printed wiring board technology. Screen-printing is a widely used thick-film process for applying films of viscous pastes on a substrate. High temperature thick-film hybrid technology generally uses sintered Al_2O_3 (alumina) as the substrate material. Conductors, resistors and insulators are deposited as thick layers (100 microns) by screen-printing. Conducting patterns are obtained by screen-printing thick-film pastes of metals (Ag-Pd), while pastes with a higher resistivity (ruthenate, Ru_2O_3) are used for resistors. Insulators are made with dielectric pastes. Capacitors are screen-printed in a three-stage process: (1) a conductive layer as the bottom electrode, (2) a thin layer of the dielectric material, and (3) a second conductive layer as the top electrode.

The thick-film pastes can be ceramic or polymer-based. Ceramic pastes are made up of active particles (metals or metal oxides) in a matrix of glass particles, organic filler materials and solvents. The active particles yield the required electrical properties to the film, for example Ag/Pd for conductivity, Ru_2O_3 for required resistivity, etc. The glass frits provide the mechanical integrity by bonding the film to the substrate. Glass frits are typically borosilicate glasses with silicate network modifiers such as PbO, CdO, Bi_2O_3, ZnO, BaO and Al_2O_3 to change physical characteristics of the film such as the melting point (which also reflects as sintering temperature), viscosity, CTE, etc. After each printing step, the hybrid substrate is dried to volatilize the solvent and the organic vehicles by evaporation. The boards are then heat-treated (sintered) at a high temperature (700–1200°C) to make up a homogenous mixture of active fillers in a glass matrix. Polymer thick-film pastes are cured at a relatively low temperature and are not stable at higher temperatures.

Screens and stencils provide the basic tools for patterning a substrate with conductors, resistors, dielectric, solder paste and other materials. When properly controlled, screen-printing is a fast and reliable method for applying controlled quantities of solder paste onto a printed wiring board. The speed of the process is essentially unaffected by the size of the board or the number of solder joints, as complex patterns can be dispensed

Samsung, EX1028, p. 385

as easily as simple ones. However, it does have limitations. One of the most important limitations of screen-printing is that it does not tolerate a casual approach to process control. Every element of the process, as mentioned below, must be carefully monitored:

- Solder-paste parameters, like viscosity and metal content
- Printing parameters, like screen characteristics, printer setup, and squeegee speed
- Room ambient conditions, like temperature and humidity

The basic screen-printing cycle is illustrated in Figure 18.14. The key components are:

- The screen—a mask with openings at locations where paste is to be dispensed
- The solder paste, which is applied to the top surface of the screen
- The squeegee—a rubber blade that travels along the screen pushing paste through the openings
- The board, which is held securely is place by a suitable fixture

During printing, the screen or stencil openings fill with paste. Paste is transferred to the substrate as the screen snaps back. After printing, gravity and paste rheology control spreading during pre-drying.

The squeegee pushes a bead of paste along the surface of the screen. As the paste reaches an opening in the screen, the downward pressure exerted by the squeegee forces paste through the opening and onto the board. The amount of paste that remains on the

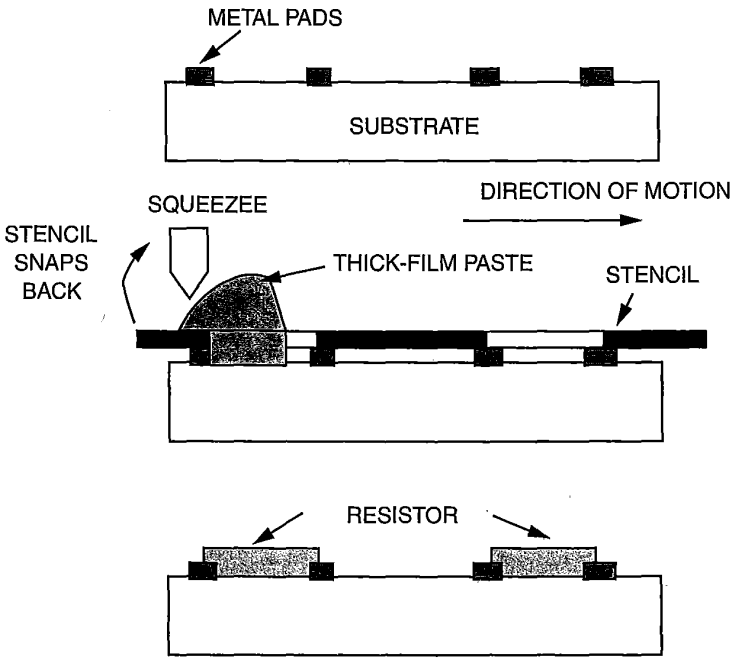


FIGURE 18.14 Screen-printing process for printing resistors in between metal pads.

Samsung, EX1028, p. 386

board after the screen retracts is a complex function of paste rheology and its wetting characteristics to both board and screen.

Organic

With just a few exceptions, organic materials are all excellent insulators. These materials have gained widespread use in the electronics industry because of their low cost, good dielectric properties, reasonable mechanical properties and ease of processing. Table 18.4 shows some common organic materials used for electronics packaging purposes.

Most IC packaging materials are made from thermoset-type polymer materials. A characteristic feature of these types of material is that they do not melt upon heating once they are cured. During the curing, the thermoset polymer forms links or chemical bonds between adjacent chains forming a three-dimensional network. The mechanical properties depend on the molecular units making up the network, and the length and density of the cross-links. The level of cross-linking can vary from material to material as well as on curing processes. Materials with high cross-link densities are hard, rigid and often brittle. Heating to high temperatures can soften materials with lower cross-link density, but they do not melt like a thermoplastic. Thermosetting resins are usually isotropic.

Earlier IC plastic packages were made by compression molding and potting. Potting involves positioning of an electrical circuit in a container and pouring the molten encapsulant into the cavity. Present IC packages are typically made by a transfer molding process. Components are formed in a closed mold from a thermosetting material that is conveyed under pressure in a hot, plastic state from an auxiliary chamber called transfer-pot, through runners and gates into the closed cavity or cavities. The four key pieces of this equipment are a preheater to make the plastic melt and flow, a press to apply pressure and shape the material, a die mold to contain the cavity, runners and gates, and a cure oven for cross-linking the thermosetting polymer.

For system-level and multichip packages, a multilayered organic laminate is used, which is generally referred to as printed wiring board (PWB). The starting material for a PWB consists of laminated layers of binder and reinforcement. Common binders are

TABLE 18.4 Properties of important organic materials in electronics.

Polymer/Polymer Composite	Relative Dielectric Constant	Thermal Expansion Coefficient [ppm/°C]	Approximate Processing Temp [°C]
Epoxy-Kevlar (*x*-*y*) (60%)	3.6	6	200
Polyimide-quartz (*x*-axis)	4.0	12	200
FR-4 (*x*-*y* plane)	4.7	16	175
Polyimide	3.5	50	350
Benzocyclobutene	2.6	35–60	240
BT (Bismaleimide triazine)	3.5	50	220
Poly norbornene	2.47	80	250
Cyanate ester-based	2.8	62	220
Teflon™ (DuPont Co.)	2.2	20	400

Samsung, EX1028, p. 387

of epoxy and phenolic type. Woven glass fibers and paper are widely used as reinforcements. Quartz and aramid (Kevlar) are also used to a limited extent. However, they are costly and difficult to process. Glass/epoxy laminates, referred to as FR-4, are the most common PWBs in use today. FR-4 has low stiffness (20 GPa) and high CTE (18 ppm/°C) and hence may not be suitable for many future applications involving multilayered thin-film structures and direct-chip attach. Stiffer boards with lower CTE will be required for this. Hence, alternate fillers such as carbon-cloth may be considered. Stiffer and high temperature polymers such as liquid-crystal polymers, for example, poly benzoxazole, may also be pursued as the matrix materials.

FR-4 stands for flame-retardant. Contrary to the earlier materials, these do not burn during wave soldering or in unfavorable conditions. Recently, it has been demonstrated that flame retardants in PWBs can cause significant toxicity in human bodies. Therefore, efforts are being made to find a replacement. So far, most alternatives seem to be phosphor-based materials. Details are discussed in Chapter 21.

Epoxy resin is generally made from ethylene chlorohydrin and bisphenol-A. When epoxy is heated under pressure, polycondensation takes place between these materials and the cross-linking of the polymer starts. The laminate is made from many layers, typically 7–10, of woven glass impregnated with the resin. Each layer is partially cured or B-staged. A photomicrograph of a woven cloth is shown in Figure 18.15. They are placed layer upon layer, in a lamination press, with one layer of copper foil on one or both sides, depending on what sort of wiring board is being made. High pressure and high temperature in the lamination process cause the resin to melt and create a uniform material. At that time, the rest of the polymerization process to the C-stage takes place, and we get a complete glass/epoxy material that is highly resistant to chemicals and climatic exposure. The most common thickness of the laminate is 1.6 mm, with 18- or 35-μm Cu foil, ½ or 1 ounce, which is the weight per square foot of the copper foil. The copper foils are made by rolling or electrodeposition.

For one- or two-layer wiring boards, a laminate with copper foil in large panels is the starting material for the wiring board producer. For production of multilayer boards, the producer of the wiring boards must also do a laminating process.

18.3.2 PWB Processes

The simplest type has only one layer of copper metal foil (single-sided boards) for conductors on one side of the board. Conductor patterns are formed by lithography, using a screen-printed resist (for coarse patterns) or UV exposure (photolithography for finer

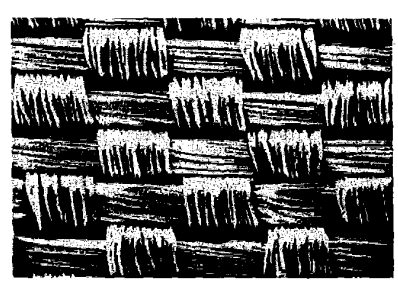


FIGURE 18.15 Woven glass fiber for printed wiring board reinforcement.

Samsung, EX1028, p. 388

patterns). This simple step is referred to as "print and etch" process. A solder coating is deposited for preventing copper oxidation during soldering.

Two-sided boards have copper conductor patterns on both sides. Surface mounted components are mounted on one side, whereas hole-mounted components are mounted on the opposite side, with leads passing through the vias. Hence, the starting board has copper-clad lamination on both sides. The vias are formed by drilling holes that connect components from one side of the board to the other side. Drilling can be mechanical or laser-assisted. In laser drilling, the photons have to be absorbed by the material. Localized heating leads to the removal of the material. The most commonly used lasers are Nd:YAG and CO_2. The holes are then electroplated. There are two versions of electroplating, namely panel plating or pattern plating. In pattern plating, copper is electroplated only on the desired conductor pattern and in the holes. This is done by defining the chemical-plated copper using a photoresist. The photoresist should prevent the deposit of the metal under the photoresist pattern. On the other hand, in panel plating, the pattern is defined by photolithography after electroplating. The photoresist pattern limits the etching of electroplated copper to the selected areas. Pattern plating is dominant, because finer features are obtained with less wastage of copper. These two approaches are also called subtractive (panel plate) or additive (pattern plate) processes.

Multilayered boards are the most complex version of PWB packaging. They are obtained by laminating several double-sided boards covered by an inner and outer layer. Key features of this process are:

1. The conductor patterns are defined on each laminated layer. Interconnections are obtained with vias. The vias are classified as through-hole or buried-holes, depending on whether they appear on the outer surface or not.
2. The epoxy of one board has to adhere well to the copper of the other board. So, copper is roughened using a microetch process in order to mechanically anchor to the epoxy.
3. Drilling of via holes through the board for subsequent hole-plating is a critical step, as it forms an electrical contact between the hole plating and inner-layer conductors. Often, the epoxy softens during drilling due to frictional heating and creates an insulating layer on the walls of the holes (smearing). The smeared insulating layer is etched (desmearing process) with plasma or strong oxidizers such as permanganate.

18.3.3 Thin-Film Processes

Increased integration demands more layers in thick-film technologies. Although MLCs provide a large number of layers (60–100), there are inherent disadvantages. Vias that pass through many layers have high inductances and result in large switching noise. In addition, a certain compatibility between the wire spacing on the film and the connector spacing on the chip is essential. The density of pins inevitably increases with improved chip performance, and this will demand smaller wire spacing on the dielectric film. Thick films offer limited wiring density, so their ability to package highly integrated, high-speed chips is limited. This led to the development of thin-film packages in which lines are made of high-conductivity metals and are defined by lithographic methods similar to those used in IC fabrication. A lower dielectric constant is essential for the thin-film

Samsung, EX1028, p. 389

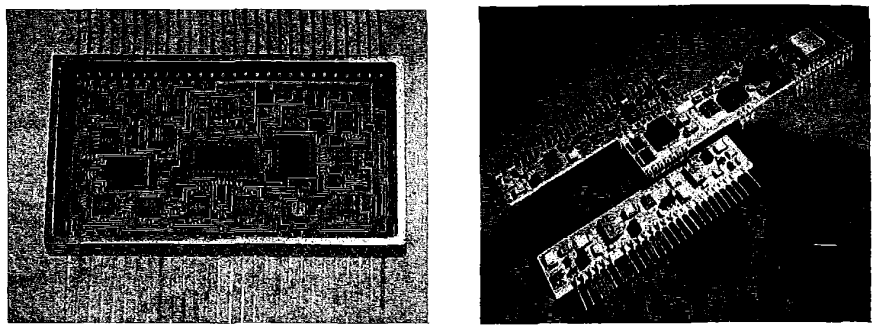


FIGURE 18.16 A thin-film (left) and thick-film hybrid circuits.

dielectrics to reduce cross-talk and increase the pulse-propagation velocity. Thin-film packaging is a sequential build-up process, while thick-film packages involve parallel processing steps where the individual green sheets can be inspected for defects.

A combination of thin-film and thick-film packaging provides more design flexibility, resulting in improved electrical performance optimization. Wiring channels can be optimized between the thin- and thick-film layers. Use of thick-film layers, along with thin-film layers, reduces the complexity of thin-film layers and increases its yield. In thin-film hybrid technology, the conductors, resistors and capacitors are processed on the substrate similar to thick-film technology. However, the materials and deposition techniques used are different. Materials used for thin-film packaging have granularity smaller than 1 micron, and therefore photoresists made from polymer films, amorphous inorganic structures from sputtered materials, etc. are used. The typical feature sizes are in the order of 25 microns, so photolithography is primarily employed for conductor-pattern definition and via generation instead of screen-printing. An example of a thin-film hybrid is shown in Figure 18.16. Glass and alumina are the most common substrate materials for the high temperature depositions. There is an increasing trend to use PWBs, because the related materials and processes are cheaper. However, the available materials, systems and processes are limited for thin films on laminates. The substrate property requirements such as CTE mismatch, thermal conductivity, inertness, chemical stability, surface smoothness, and uniformity are more stringent for thin-film processes. The processes are classified in Table 18.5. Typical electrical properties that can be achieved with these materials and

TABLE 18.5 Materials and processes for thin-film technology.

Function	Material	Process
Conductor	Copper, Aluminum	Vapor deposition, sputtering, electroplating, photolithography
Resistor	NiCr alloys, Ta_2N, Ta, TaO_2N	Sputtering
Insulator	Polymers (epoxies, polyimide)	Solution-coating (spin, meniscus etc.)
Capacitor	$BaTiO_3$, TiO_2, Ta_2O_5, SiO_2	Sputtering, evaporation MOCVD, hydrothermal techniques, etc.

MOCVD, metal organic chemical vapor deposition.

Samsung, EX1028, p. 390

the tolerance values can be found in Chapter 11 on passive components. Conductors are made by deposition of a metal thin film, with a typical thickness of 1 μm, most often gold or aluminium. This is done by evaporation or sputtering deposition.

Traditional thin-film circuits are made with one layer of conductor lines and one resistor layer. Recently, the more complicated multilayer thin-film hybrid technology has increased in use. With multilayer technology, higher packaging density with increased complexity can be achieved. Electrical characteristics can also be improved by adding separate ground and power supply conductor layers, giving better shielding and the possibility to obtain controlled characteristic impedance in high frequency applications, etc.

Most thin-film hybrid circuits use ICs in the form of bare silicon chips, which are mounted with adhesive and connected by wirebonding. Usually, the complete circuit is packaged in a hermetic metal or ceramic package.

Thin-film technology offers a substantial increase in packaging density compared to thick-film hybrid technology. It offers excellent electrical characteristics and high reliability. This is a specialized technology with relatively few suppliers. Thin metal films on hybrid substrates and semiconductor wafers are normally deposited by vacuum evaporation, sputtering or electroplating. Vacuum evaporation of metals is typically followed by a photolithography process.

Physical Vapor Deposition (PVD)

Vacuum Evaporation Vacuum evaporation takes place in a chamber evacuated to 10^{-6} Torr (Figure 18.17). There are three main reasons to perform the deposition in vacuum:

1. Increase the mean free path of the evaporated particles, so that the deposition is efficient and uniform.
2. Reduce the vapor pressure, and the temperature required to achieve a particular deposition rate.
3. Remove atmosphere and other contaminants that might otherwise contaminate the deposited film.

The metal that is to be deposited is placed in a boat made up of a high temperature metal, such as W or Mo. The boat is heated by passing a high current through it (resistance heating), until the metal evaporates and its molecules spread out upwards in all directions in the vacuum chamber. The molecules hit the substrate that is located on the substrate holder above. A different way to heat the metal is to bombard it with an electron beam that is accelerated in an electric field with the boat as the anode. Al, Cu, and Au are some of the metals that are frequently deposited by these processes. The adhesion to the bottom substrate is improved by interlayers of Ti, Cr, or nichrome, which oxidize easily and form a chemically compatible interface with the bottom ceramic.

Sputtering Sputtering is a low-pressure process where a "target" is bombarded with energetic positive ions. When the ions hit and give off their energy, particles are ejected from the target. These "sputtered" particles hit the substrate that is to be covered. The plasma may be a noble gas such as argon, at pressure of the order of 0.1–10 Pa (10^{-3}–10^{-1} Torr). The gas is ionized in a strong electrical field of 1000–2000 V between the target and the anode. A glow discharge is created in the argon gas when it is decomposed into positive ions and free electrons. The ions are accelerated toward the cathode (target),

Samsung, EX1028, p. 391

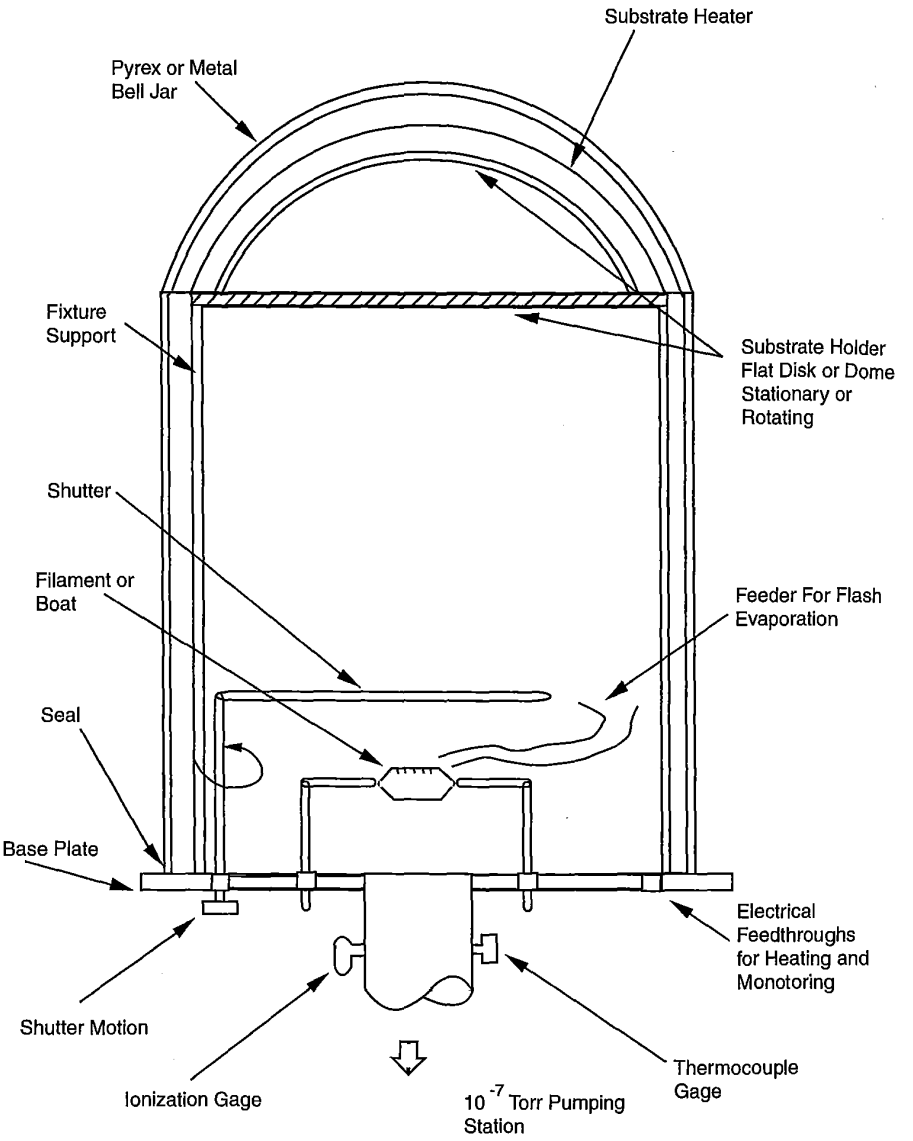


FIGURE 18.17 Vacuum evaporation.

where they give off their charge by receiving electrons, and are again neutral. Target material is torn off by the energy released, and deposits on the substrate. Energetic secondary electrons interact with neutral argon atoms, causing them to ionize again. Typical deposition rates are 100–1000 Å per minute. Film thickness by sputtering, as well as by vacuum deposition, is limited to a few μm. Electrically insulating target materials are not suitable for DC-sputtering, because they will be electrically charged

Samsung, EX1028, p. 392

and then disturb the glow discharge. In such cases, AC radio frequency is used on the target holder electrode. Other sputtering configurations such as magnetron sputtering use permanent magnets to provide a magnetic field parallel to the target surface. The field serves as a trap for secondary electrons near the target, preventing them from reaching the substrate. This has a two-fold advantage. The substrate is cooler because fewer electrons bombard it and secondly, the deposition rate is enhanced due to a higher electron–atom collision rate in the region near the target.

Sputtering is mostly used for depositing metals. However, it is widely pursued for various resistor materials such as Ta_2N (10–125 Ohm/sq), TaSi (100 Ohm/cm), CrSi (360 Ohm/sq), TaN and NiCr (35–100 Ohm/sq). Capacitor films of Ba (Zr, Ti)O_3 (ε_r = 145), alumina (40 nF/cm^2) and inductor films of sputtered copper (50-micron wire) on polymer coated aluminum core were also demonstrated.

Chemical Vapor Deposition (CVD)

In CVD, chemicals in vapor phase react to form a solid film on a surface. The typical sequence for this process consists of:

- Diffusion of reactants to the surface
- Adsorption of reactants onto surface sites
- Surface reaction on the surface, usually assisted by catalysis
- Desorption of the by-products
- Diffusion of by-products from the surface
- Incorporation of the solid product onto the microstructure of the growing film

Though CVD involves standard thin-film growth processes such as nucleation of the film product, growth and coalescence, the central feature that distinguishes it is the heterogeneous reaction over the substrate on which we intend to grow the film. The reaction is generally endothermic (assisted by external heat addition). The energy needed for this growth comes by heating the substrate, and the technique is referred to as thermal CVD. The need for lowering the deposition temperature led to the development of plasma-assisted CVD. In this process, interactions involving charged particles produced in plasma alter the reaction path by lowering the activation energy and hence form products at relatively lower temperatures. Compound thin films are typically grown with *metal organic precursors* (MOCVD). This process is primarily developed for growing compound (GaAS) semiconductors for optoelectronic applications. It is being widely explored for other packaging applications. Process optimization, monitoring and control are more difficult with CVD processes. However, materials such as silicides, nitrides and fluoropolymers that may not be deposited by other processes are only possible through CVD.

Solution-Based: Physical

Most of the polymer deposition techniques are guided by the interplay between density, viscosity and surface tension of the liquid, and gravity. Higher viscosity and surface tension make the films thicker, while a higher density and gravity make the films thinner. The typical thickness obtained is of the order of 10 microns. Only three techniques are mentioned here.

Samsung, EX1028, p. 393

Spin-Coating Spin-coating can typically yield thicknesses from 2–20 microns. The thin-film coating is obtained by rotating the substrate at a high speed (300–10000 rpm) after dispensing the solution on the substrate. Empirically, it is seen that the thickness is inversely proportional to the square root of rpm.

$$\text{Film thickness} = \frac{\text{Constant}}{\sqrt{\text{Angular velocity}}}$$

The favored substrates are small round wafers. Spin-coating is a standard IC fabrication process. The technique can offer excellent thickness control and reproducibility. However, it results in a lot of waste. Thicker films are produced with viscous solutions at lower speeds. However, the uniformity is affected in this process. Hence, multiple coatings are generally preferred for thicker films. Polyimide is the best common example of a spin-on material.

Meniscus Coating In meniscus coating, a liquid polymer solution is pumped out of a narrow slit on the top of a tube over which the substrate slides. Material may be collected under the tube and recirculated into the center of the tube. So it is more efficient in terms of waste, compared to spin coating. Process parameters include distance from substrate to tube, substrate velocity, surface tension of the solution, viscosity, solvent evaporation rate, etc. Film thickness mainly depends on the substrate speed. The process is illustrated in Figure 18.18.

Dip-Coating Dip-coating involves a vertical motion of the substrate after being dipped in a reservoir. The thickness is typically found to vary as the square-root of viscosity, withdrawal speed, and is inversely proportional to the square-root of density and gravity.

$$\text{Film Thickness} = \frac{\text{Constant}\sqrt{\text{Viscosity speed}}}{\sqrt{\text{Density}}}$$

Solution-Based: Chemical

Wet-chemical techniques offer a potentially superior method for low-cost thin-film processing. Since synthesis occurs at lower temperatures, problems such as interdiffusion and

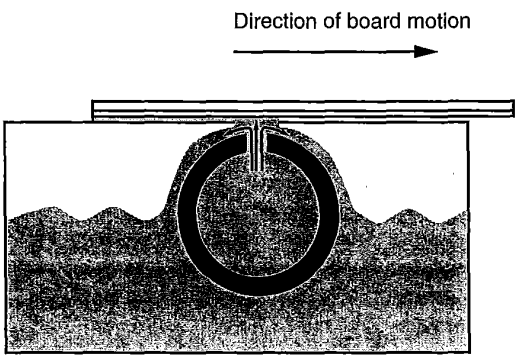


FIGURE 18.18 Illustration of the meniscus coating process.

Samsung, EX1028, p. 394

interfacial reactions that are detrimental to device performance can be eliminated. The energy consumption of these processing routes is lowest, compared to vapor-deposition techniques. Most of these processes are still in the development stage and are not widely used for conventional electronic packaging.

Sol-Gel Deposition Sol-gel deposition allows for the deposition of films with a high degree of chemical homogeneity at relatively low temperatures. The process starts with the formation of a homogenous solution ("sol") of raw materials. Most frequently, the raw materials are organometallic compounds such as metal alkoxides that are dissolved in alcohol to give a homogeneous solution. "Sol" also refers to a mixture of solid nano-sized colloidal particles in a liquid. The solution (sol) is subsequently gelated by reaction with water (hydrolysis) or exposure to atmosphere. The gelation forms a polymeric network or a colloidal network. Many factors influence the gelation: the nature and concentration of the catalyst, the amount and nature of solvent, and the sequence of mixing. The gel is amorphous and weak because it has continuous pores and trapped organics (water, hydroxyl groups, etc.). Thin-film coatings are typically obtained from polymer solution deposition techniques such as dip-coating or spin-coating. Heat treatment and densification of the gel form the final film. The high surface area of the dried gels results in very high reactivity, which, in turn, results in a relatively low-temperature process compared to conventional ceramic processes that involve sintering of micron-sized particles. It is demonstrated to yield thin-film capacitors and low-K dielectrics. Nanoscale composites from sol-gel processes have many superior properties and are going to dominate various future applications.

Hydrothermal Deposition Hydrothermal treatment involves dissolution of reactants and precipitation of products in hot, pressurized water. It is a standard technique to form fine powders with superior physical and chemical properties. The raw materials are typically the same as those used in the sol-gel process. A subsequent hydrothermal treatment of the sol can assist in the formation of thin films at lower temperatures. Water serves as a pressure-transmitting medium. More importantly, the presence of water enables reactions to take place at lower temperatures. The reactions are carried out in closed vessels, typically under strong alkaline conditions. The bath conditions such as temperature, alkalinity, etc. can be predicted from the theoretical phase stability diagrams of the relevant systems. The phase diagrams provide a thermodynamic basis for the formation of films. The film growth is dependent on the type of substrate. Substrates with structures closer to the films allow easier nucleation and growth of defect-free films. A combination of electrochemical and hydrothermal techniques is also being investigated by some researchers.

Electroless and Electroplating Electroless plating is a metal deposition process, usually in an aqueous solution medium, which proceeds by a chemical exchange reaction between the metal complexes in the solution and the particular metal to be coated. Since the deposition occurs by a chemical driving force, the process does not require an external current. Electroplating, on the other hand, is a process of depositing an adherent metallic coating onto a conductive object (cathode) immersed in an electrolytic bath composed of a solution of the salt of the metal to be plated. The deposition occurs by passing a

Samsung, EX1028, p. 395

DC current through the electrolyte, using a metal anode and the conductive substrate (cathode). Anode is generally made of the metal to be deposited.

Plating is a cheap and low-temperature process, and so is widely used for thin-film metallization. Metal plating is mainly discussed here, particularly with reference to copper. Plating is typically done in two steps. In the first step, copper is chemically deposited using a catalyst-assisted electroless plating process. Electroless plating plays an important role in metallization of insulators, because it renders a conducting surface for the electroplating step that follows.

Activation for Chemical Plating Metals such as Cu, Au, Ag, Pt and Pd initiate electroless plating by acting as catalysts for electroless copper deposition. The most common activator is Pd, which is typically deposited in a two-step process. In the first step, the panel is dipped into a solution containing Sn^{2+} ions, which adsorbs on the surface and increases its sensitivity. Sn^{2+} ions are hence termed as sensitizers. The second step takes place in an acidic solution of palladium chloride which transforms to metallic Pd as:

$$Sn^{2+} + Pd^{2+} \rightarrow Sn^{4+} + Pd$$

In the plating process, Pd catalyses the deposition of copper. The same treatment can also be done in a single step. Sn-Pd colloids, which act as "catalyst centers," are directly adsorbed on the substrate surface by immersion in a mixture of $PdCl_2/SnCl_2$ solution. However, Sn (IV) ions are also adsorbed on the substrate in this process, and they act as inhibitors to the reaction. To increase the activation, these ions are removed as much as possible by an activator. The adherence of catalyst is determined by its chemisorption or physisorption to the polymer matrix.

Electroless (Chemical) Plating of Copper The panel is dipped into a reducing bath containing Cu^{2+} ions, for example, in the form of dissolved $CuSO_4$. Formaldehyde, HCHO, is the common reducing agent. In this bath, Cu^{2+} is reduced to Cu that covers the whole surface, including the vias. At the same time formaldehyde is oxidized into acetic acid. The plated thickness is approximately 0.5 μm. The purpose is to create an electrically conducting surface for the subsequent electroplating step. A low pH is desired, because it minimizes the attack of caustic solution on dielectric polymers. The formulation of the chemical plating bath is very critical. The composition of the bath is complex, and if the composition, temperature, and "bath loading" (the ratio between plated area and bath volume) are outside the acceptable values, Cu^{2+} may be reduced to metallic Cu in the bath itself. This will then ruin the bath.

Electrolytic Plating of Copper The panel is dipped into an electrolyte that contains Cu^{2+} ions, such as $CuSO_4$ dissolved in H_2SO_4. The panel forms the negative electrode (cathode), and a metallic copper plate forms the positive electrode (anode) of an electrolytic cell. A schematic set-up is shown in Figure 18.19. At the anode, copper is dissolved:

$$Cu \rightarrow Cu^{2+} + 2e^{-}$$

The reaction at the cathode is the following:

Samsung, EX1028, p. 396

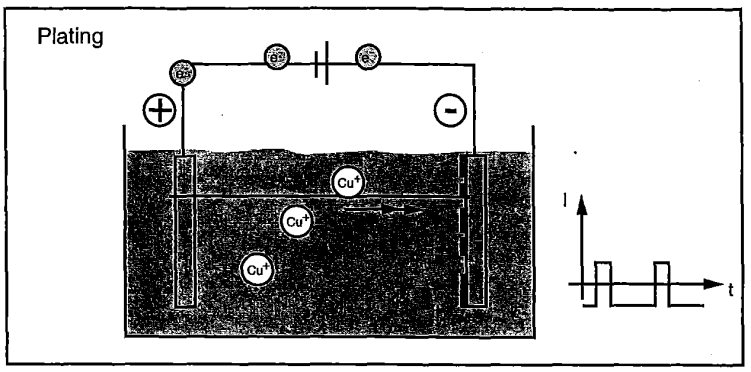


FIGURE 18.19 A schematic of electroplating.

$$Cu^{2+} + 2e^- \rightarrow Cu$$

Thus, metallic copper is deposited on the panel. Approximately 25–30 μm Cu is normally plated in order to get good coverage in the via holes.

Tin/Lead Plating for Etch Masking The Sn/Pb pattern serves as an etch mask and also aids in soldering. The panel is connected to the cathode of an electrolytic bath containing Sn^{2+} and Pb^{2+} ions. The anode is metallic Sn/Pb alloy. The electrolyte is based on fluoroboric acid, HBF_4. The ratio between the concentration of the ions in the bath and on the anode is such that the deposited layer of metal on the panel will be approximately the eutectic mixture 63Sn/37Pb (percent by weight). The normal thickness is about 7 μm.

Electrochemical deposition provides large-area fabrication, with very low capital costs compared to the alternate vacuum deposition methods, and thus is a viable technology for future low-cost electronic packaging.

18.3.4 Photolithography

Photolithography is probably the single most important process enabling the semiconductor and electronic industry in general. It is used for transfer and definition of fine patterns that are not amenable by screen-printing. This technique has been refined by the semiconductor industry, where accuracy and resolution below 0.2 μm can be achieved. In packaging, the conductor line widths are typically far higher.

The pattern is generated on a CAD or a similar system and is transferred onto a photographic film, called the photomask or artwork. The film is typically a black-and-white plastic foil. However, a glass plate is used when high precision is required. The pattern on this film is made by exposure with a computer-controlled photoplotter, laser- or a computer controlled electron beam, for sub-micron resolution. For transferring the pattern, a thin layer of photosensitive material, called photoresist, is deposited on the substrate or laminate surface. This is typically done by spray coating, screen-printing, laminating in the form of a dry film, or by spinning. After the photoresist is applied, the

Samsung, EX1028, p. 397

mask is aligned with respect to the prior patterning on the substrate. The photoresist on the surface is then illuminated through the photo mask, creating exposed and nonexposed areas according to the mask pattern.

Photoresists can be classified as negative or positive, depending on whether light initiates cross-linking in the polymer making the illuminated portion difficult to dissolve in the developer (negative resist), or light breaks the molecules, making the illuminated portion easier to dissolve in the developer (positive resist). With a negative photoresist, we remain with the hardened resist in the portions that are exposed through the photomask. The unexposed portions are dissolved in a suitable solvent. Hence, negative resists have the same resist pattern remaining as the white areas on the photomask. Correspondingly, positive resists have the resist pattern remaining at the dark areas on the photomask. Positive resists generally give the best definition, but they are more costly than the negative ones. Negative resists have sensitizers or photo-initiators, which absorb light and initiate a chemical change in the resin to produce an insoluble film. In addition to the resin, sensitizers and solvents, various additives that enhance the performance of the resist, such as stabilizers, antioxidants, adhesion promoters and coating aids, are present.

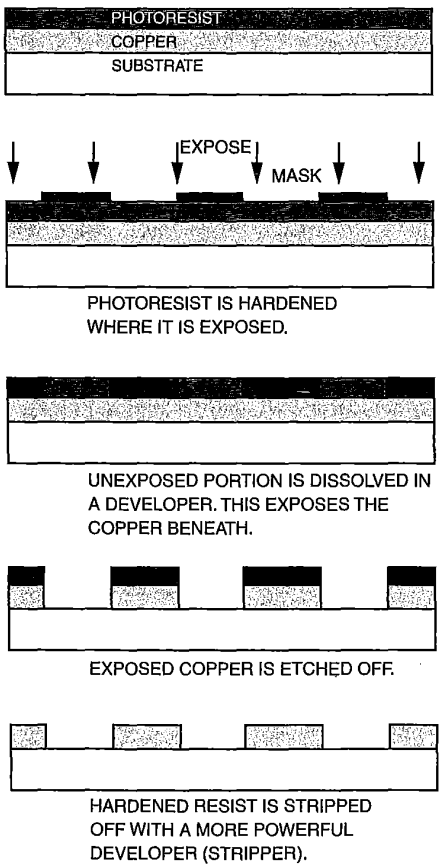


FIGURE 18.20 Process flow for photolithography using a negative photoresist.

Samsung, EX1028, p. 398

Frequently, these photoresists are supplied as dry-films and are laminated using a vacuum lamination process. The dry-film composite system offers several advantages, such as no solvent evaporated during coating, high yield, elimination of costly coating equipment and simplification of the process. Resists of various thicknesses are available. The patterns are required to be as thick as the circuitry in order to limit the metal pattern during additive plating.

The cured resist is resistant to many chemicals like acidic etchants, and thereby protects the material underneath from attack. Thus, when metal or a dielectric layer are to be etched, the photoresist may be used as an etching mask. During the etching, the exposed material will be etched, whereas the material covered by the resist remains unattacked. In the case of metal plating, the resist will inhibit the formation of a plated metal layer on the covered areas. This results in plating only onto the exposed areas. When the etching or plating process is completed, the rest of the photoresist is dissolved in a suitable solvent. The process flow is schemtically shown in Figure 18.20.

18.4 SUMMARY AND FUTURE TRENDS

Most of the existent materials cannot provide the suitable combination of properties to perform the required functions and suitable processability. New materials that improve the cost/performance of a package are continuously developed at both academic and industrial levels. This section discusses some of the recent trends in developing novel materials.

18.4.1 Interconnections

Most of the details about packaging interconnections and assembly are discussed in Chapter 9 (Fundamentals of Chip Assembly). Lead and other heavy metals, halogens, etc. are classified as toxic elements because of the serious health problems associated with their use. The environmental concerns of using lead-free solders are discussed in Chapter 21 (Microsystems and Environment). There is a strong drive to replace lead in the solders with other elements, and yet retain the advantages of Pb-Sn alloys. Alternatives to lead-based solders and associated technical problems are discussed in this section.

Early attempts to introduce alternate materials were plagued with process incompatibilities associated with their use. There are two approaches to lead-free solders: lead-free metallic solders and conductive polymers. The options for lead-free solders rely on tin as the base metal with smaller amounts of other metals—such as bismuth, copper, silver or zinc—added to enhance the performance. Tin, which is considered to be one of the least toxic metals, will most likely endure as the base metal, since it is relatively inexpensive, sufficiently available, and possesses desirable physical properties. The reason for choosing tin-based alloys is that tin interacts very strongly with a wide range of metals, forming strong metallurgical bonds; this property is critical to accomplish the fundamental purpose of soldering-joining two surfaces. Tin by itself is unsuitable as a solder since it "whiskers," migrates under electrical field, has a high melting temperature and forms a brittle grain structure at cold temperatures. The question then would be "with what to alloy tin?"

Among the other metals, cadmium, mercury, lead, and thallium must be excluded from consideration because of their intrinsic toxicity. The alkali metals have low melting

Samsung, EX1028, p. 399

temperatures, but due to their extreme reactivity, this group of elements must also be excluded. Lanthanum, strontium, radium, polonium and thallium are rare-metal elements; it is impossible to use them in large quantities in solder alloys due to their limited supply. Therefore, the most recent work is focused on the tin-based lead-free alloys.

When developing new solders, there are many different aspects that have to be considered, the most fundamental property being its melting temperature. High processing temperatures will cause thermally-induced defects in both boards and components. In practice, the solder alloy melting temperature (liquidus) should not be higher than 200°C in most electronic assembly products. On the other hand, the melting point must not be too low because the typical electronic product in service operates at 50–80°C, and boards may be tested at temperatures of up to 125°C. Furthermore, unlike pure metals, most alloys melt over a range of temperatures, called plastic range. The replacement materials need to be stable for all the methods of soldering: printing and reflow wave solder, BGA solder balls, flip chip die preparation, preforms, and hand soldering. They should also be compatible with new techniques such as no-flow underfill. Soldering processes require a narrow plastic range. Table 18.6 lists some tin-eutectic alloys and their melting temperatures.

When choosing alternative metals, consideration must also be given to their health risks as well. Recent studies in the U.S. and in Europe came to the following conclusions concerning the toxicology of lead and some alternative metals:

- Cd and Pb are extremely toxic elements and should not be used.
- Sb is toxic and should not be considered as a major alloying element. In Europe, this material is considered potentially carcinogenic.
- Ag and Cu are relatively harmless.
- Sn and Zn are essential elements in the human diet. However, they may be toxic at sufficiently high exposures. Zinc oxidizes easily and therefore has high corrosion potential when incorporated in paste form with added acid or basic fluxes. In addition, the alloy drosses excessively during wave-soldering.
- Bi is a relatively benign metal with a history of medicinal use. However, it is a by-product of lead and hence may be avoided. It has poor compatibility with many fluxes. It produces a tenacious oxide, requiring very active flux and/or a nitrogen reflow environment. This results in high solder defects and poor mechanical properties at elevated temperatures.

Some other important technical aspects are wettability and sufficient mechanical strength. For electronics and heat exchangers, electrical and heat conductivity are also of importance. Another vital aspect, especially for electronics, is that the melting point must be low to prevent damaging the construction or components during soldering. It is necessary to bear in mind that the alloying elements must be available in sufficient amounts for an increasing demand, and not be environmentally dangerous or toxic. They must also have a reasonable price.

In Figure 18.21, the numbers indicate weight percent. Additional information along with some alternative basis metals and their alloys is shown in Table 18.6. It can be seen that Sn-Ag, Sn-Au, Sn-Cu, Sn-Sb, Sn-Zn, Sn-Bi, and Sn-In binary systems, or the derived multicomponent alloy systems, have been focused on. There is still a wide range to chose from, although all these alloys are not commercially available. The question of which

Samsung, EX1028, p. 400

CONFIDENTIAL BUSINESS INFORMATION

TABLE 18.6 Alternative metals and alloys to lead, pricing, melting point and suitability for replacement.

Element	Toxicity	Typical Composition (% by weight)	Alloy price (US $/kg)	T_{melt} (°C)	Remarks	Suitable for Lead Replacement
Pb	High	Sn-37Pb	3.6	183	Well-established alloy Low price	
Ag	Low	Sn-3.5Ag	10.1	221	Good thermal fatigue res. Fast Cu dissolution	Yes
Cu	Low	Sn-4Ag-Cu	10.8	216–219	Good thermal fatigue properties at high temperatures	Yes
		Sn-3.8Ag-0.7Cu	10.5	217–219		
		Sn-4Ag-1Cu	10.8	216–219	Favorably priced	
		Sn-0.7Cu	5.3	227		
Bi	High	Bi-42Sn	7.0	138	Suitable for low temperature applications Sn/Bi melts at 138°C Obtained as a by-product of lead extraction	Yes
		Sn-2Ag-7.5Bi-0.5Cu	8.2	198–212		
		Sn-3.4Ag-4.8Bi	10.0	200–216		
Sb	Medium	Sn-5Sb	5.1	232–240	High melting point Creation of needle-shaped Ag_3Sn phases Similar properties as Sn/Ag/Cu	Yes
		Sn-25Ag-10Sb	39.3	230–235		
		Sn-2Ag-0.8Cu0.5Sb	8.0	217–220		
In	Medium	In-48Sn	132.5	118	Expensive Limited availability Corrodes in combination with humidity Very soft	No
		In-3Ag	246.8	143		
		Sn-20In-2.8Ag	58.1	189		
Zn	Low	Sn-9Zn	5.0	199	Problem of oxidation, Strong dross formation	No
Au	Low	Au-20Sn	8403.3	280	Much too expensive	No

KE1668(

Samsung, EX1028, p. 401

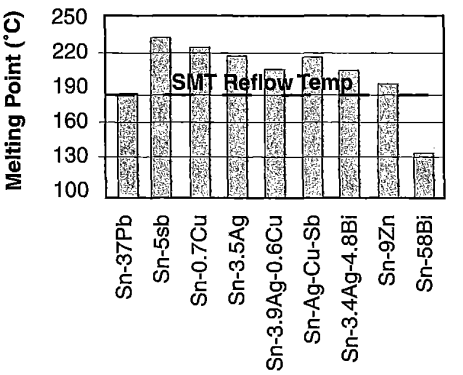


FIGURE 18.21 Melting points of candidate lead-free solders.

lead-free solder is the best is often asked, but is difficult to answer, since there is no absolute drop-in replacement for tin-lead with identical melting temperature, cost, wetting and strength properties.

For low-cost electronic assembly (like consumer electronics and computing devices), the research has been narrowed down to a few binary eutectic alloys such as Sn-58Bi, Sn-52In, Sn-3.5Ag and Sn-0.7Cu. They have been extensively evaluated as potential replacements for tin-lead solders based on performance, cost, and availability. These lead-free alloys mostly consist of tin, unlike the 60/40 tin-lead combination. Only small additions of copper or silver are needed to produce a eutectic alloy. A eutectic alloy has the same liquidus temperature as solidus temperature, and quickly fixes components on the board surface. If the properties of the alloy result in formation of liquid over a range of temperatures, components can float on the board surface and easily become misaligned.

The most convenient way to separate the available lead-free alloys is to consider their melting temperature. Most of them fit into one of the following categories: low melting temperature (below 180°C), melting temperature equivalent to the tin-lead eutectic (180–200°C), mid-range melting temperature (200–230°C), and the high temperature alloys (230–350°C). It should be noted that the eutectic temperature of simple binary alloys is relatively easy to establish and has been known for some time. By considering the service temperature that the alloy will experience, together with any maximum temperature limitations of the components or other items to be joined, an alloy of suitable melting point can be found. Because of the higher melting point of new lead-free solders, reflow temperature has to be higher by 20–30°. This may have some detrimental effect on the components and machinery, and hence requires further study. Sophisticated electronic products such as routers, servers, mainframes with large ball grid arrays and build-up microvia boards would be less tolerant to reflow temperatures that are higher by 30°C.

Organic-Based Electrical Interconnects

Polymer materials are generally nonconductive. Metallic fillers are added to make them conductive. The conductive properties of adhesives are generally explained by the percolation theory. The percolation threshold is defined as the point at which the conductive particles start to form continuous conductive paths through the adhesive. Electrical conductivity increases dramatically when the volume fraction of a metallic filler reaches a

Samsung, EX1028, p. 402

critical value. It is believed that all particles contact each other and form electrical paths inside the material. Depending on the amount of fillers, the adhesive is classified as nonconductive (has no filler), anisotropic conductive (is conductive only in one direction, using a fraction of conductive particles well below the percolation threshold) and isotropic conductive (the fraction of conductive particles is well above the percolation threshold). The volume-fraction at percolation depends on the form and size of the conductive particles, as well as the geometry of the adhesive joint but is typically in the order of 15–25%.

The main advantages of using polymer-based interconnects are low die stress because of the low modulus of the adhesives compared to solders and low processing temperature. Although these adhesives have been improved to near solder electrical conductivity and better stencil yields, their usage has been limited to niche low temperature assemblies such as optical interconnects and low-melt substances. Their inability to replace solders with SMT assembly is due to poor reliability and assembly yield and very high cost. In addition, conductive adhesives are brittle due to high filler loading, and cannot provide self-centering to improve assembly yields, unlike high surface tension solders. Nonreworkability, and silver migration failure between the tightly pitched leads, also contributes to unsuitability of isotropic adhesives for mass solder replacement. Silver-filled epoxies cost about 5–13 times more than solders, which further limits their applicability.

Nonconductive Adhesive

The concept of making electrical contacts with nonconductive adhesive is relatively new but has many points of resemblance with ordinary electrical contacts. In both cases, the conducting surfaces of the two parts to be electrically connected must be maintained in intimate contact with each other. The force responsible for this contact is called contact force, and is caused by elastic deformations of one or more parts in the contact assembly. This elastic deformation or energy is supplied during the initial creation of the contact. In an ordinary contact, this energy is supplied in the form of mechanical work, whereas in the use of a nonconductive adhesive, there will be an addition of chemical energy transformed into elastic energy during the cure of the adhesive. The adhesive does not by itself contribute to the electrical conduction. However, the contact area has a metallic or quasi-metallic surface which, when sufficiently thin, permits conduction by electron-tunneling.

Anisotropic Conductive Adhesive (ACA)

Conventional ACA is an adhesive consisting of conductive particles dispersed in an adhesive matrix. These particles can be pure metals such as gold, silver or nickel, or metal-coated particles with plastic or glass cores. The particles typically range from 3–10 microns in diameter. The volume fraction of particles is well below the percolation-threshold. Below the threshold, there are no continuous conductive paths among the particles in the x-y plane of the film. Hence, the interconnections are electrically isolated. Assembly involves deposition of the adhesive on the substrate, with subsequent placement of the component. Pressure is applied on the component during a 5 to 15 second snap-cure, followed by a complete oven cure. During the connection process, a large number of particles are trapped between the opposing contacts on the chip and the substrate (Figure 18.22). These particles create an electrical bridge and thereby provide the elec-

Samsung, EX1028, p. 403

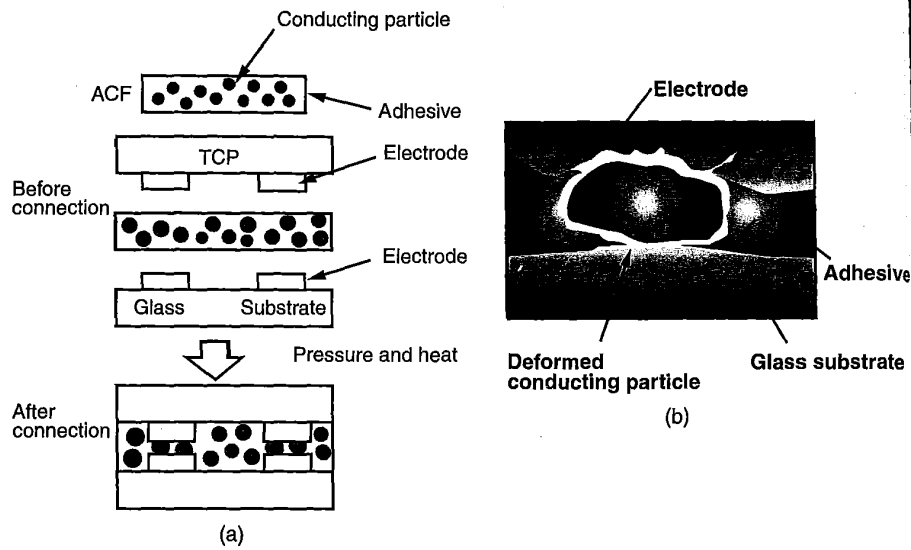


FIGURE 18.22 (a) Procedure of interconnection using ACF; (b) cross section of interconnection using ACF.

trical connection only in the z-direction (direction in which the pressure is applied). The adhesive matrix provides mechanical rigidity and contact pressure as well as protection.

Anisotropic adhesives are mostly used to attach LCD display drivers since solder reflow temperatures would destroy the LCD. They are not likely to ever replace solder interconnects because of the applied pressure requirement and very poor yield and reliability. Coplanarity is extremely critical and affects assembly yield. The cured epoxy is vulnerable to moisture sensitivity and hence fails by cracking and swelling.

The low processing temperature of *anisotropic conductive adhesives* (ACA) is one of the primary reasons for its widespread use for LCD driver attachment. In addition, it can be fine-pitched, because there is less chance of solder bridging or smearing of conductive adhesive. These anisotropic conductive materials are continuously improving, both from a manufacturing point of view, as well as in reliability. Before the component is mounted, an adhesive film or paste is applied uniformly over all the contact areas. Since there is no need for precise placement of the ACA, the mounting process is greatly simplified.

Isotropic Conductive Adhesive (ICA)

The isotropic conductive adhesive is an epoxy, typically filled with silver particles. Silver is used because it provides good electrical conductivity even when oxidized. Other metals like copper, nickel, and even gold, are also used to some extent. As the adhesive is conductive in all directions, a lot of care must be take to avoid short-circuiting between neighboring pads. However, there is no need for high contact pressure, as in the case of nonconductive and anisotropic conductive adhesives. Another important consideration when using adhesives for interconnect is the fact that, unlike solder, they lack the ability to align the component during the connection process. This is because the surface energy

Samsung, EX1028, p. 404

of the adhesive is only of the order of a few percent, compared to that of the solder, and wetting by adhesive is generally nonselective.

The present limitations of ICA technology are the high initial contact resistance, unstable contact resistance and inferior impact strength. The resistance of conductive path consists of the bulk resistance of metallic particles and the contact resistance. The contact resistance is larger than the bulk resistance and is found to increase dramatically at elevated temperatures and humidity environments. Metal oxide formation resulting from galvanic corrosion was responsible for the increase in contact resistance of an ICA with non-noble metals. Various corrosion inhibitors were investigated and shown to improve the stability of contact resistance. The inhibitors adsorb on the metal surface and separate the metal from corrosive environments. They also react with oxygen and consume the oxygen, hence preventing the corrosion. The impact resistance of the ICAs can be improved by rubber-modified epoxy resins. Figure 18.23 shows the schematic of an ICA.

18.4.2 Low-Dielectric Constant Dielectrics

It has been shown before that the incorporation of low-K dielectric materials between the metal lines increases the signal propagation speed and reduces the capacitative coupling and cross-talk between the interconnects. Many low-K materials are presently under investigation, ranging from fluorinated silicon oxides to inorganic, and even more exotic, materials like porous xerogels and nanofoams.

Organic polymers are promising candidates for *interlevel dielectrics* (ILD), since they display mostly electronic polarization and therefore have a low dielectric constant. The most popular interlevel dielectrics are spin-on polyimides with dielectric constants in the range of 3.0–3.3. Fluorinated polyimides have been developed in order to obtain lower dielectric constants. Though they possess good planarizing capabilities, they have several disadvantages, such as moisture absorption, low break-down potential, increased leakage currents, poor adhesion and corrosion of metal components. Other novel aromatic pol-

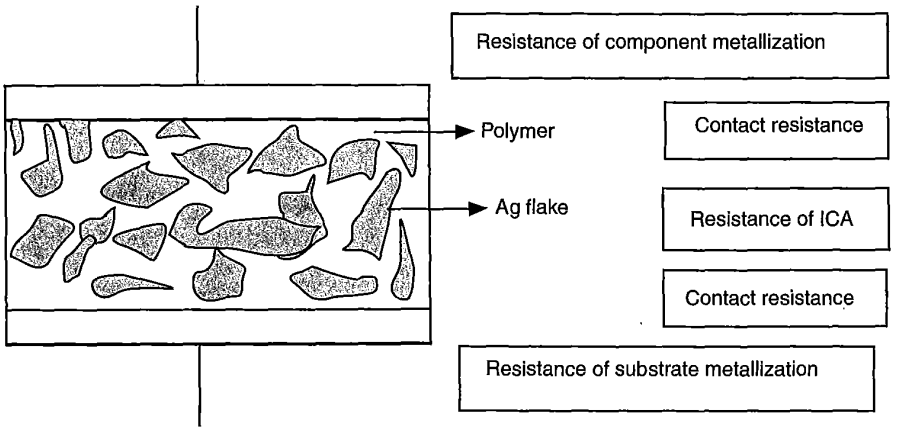


FIGURE 18.23 Schematic of an ICA.

Samsung, EX1028, p. 405

ymer systems include fluorinated polyarylene ethers (FLARE), benzocyclobutene (BCB), SiLK (silicon like low-K material) and parylene AF-4. BCB is a thermosetting polymer with dielectric constant in the range 2.6–2.8 and has exceptional spin-on planarization and low moisture absorption. SiLK is a spin-on aromatic material with a dielectric constant of 2.7 and thermal stability in the excess of 450°C. Parylene AF-4 is vacuum deposited, rather than spun-on, and exhibits thermal stability, minimal moisture water absorption, a low dielectric constant of 2.28 and good adhesion.

There are also a number of organic/inorganic hybrids, such as methyl silsesquioxane (MSQ) and carbon-doped silicon dioxide (black diamond). These materials, provide the thermal stability and strength of inorganic materials and also exhibit a lower dielectric constant due to lower polarizability of the Si-C bond, compared to the Si-O bonds in silicon dioxide. MSQ is a spin-on dielectric material with a dielectric constant of 2.7 and has good adhesion and thermal stability in excess of 500°C. Fluoropolymers have dielectric constants ranging from 1.8–2.5 and seem promising for future circuits. PECVD of fluorocarbons is a possible solution to increase the thermal stability by producing cross-linked films. Plasma deposition offers several advantages, such as formation of dense pinhole free films, large variety of possible monomers, enhancement of adhesion by suitable pre- and post-treatments, excellent gap-fill abilities and compatibility with the subsequent processing.

Among the inorganic materials, the borosilicate glass + silica system is of particular interest. It has dielectric constant ranging from 3.9–4.2. The CTE can be matched to that of silica by suitable additives. This material has been widely used for manufacturing multilayered and cofired glass ceramic substrates with dense wiring. Further lowering of dielectric constant can be obtained by the creation of nanosized pores in the material. Nanopores were deliberately introduced by using hollow silica microspheres or the incorporation of a second phase that can be decomposed and removed during heat treatment. It should be noted that incorporation of open space in a microstructure can lower the mechanical strength and thermal conductivity and make the material more prone to moisture absorption. Porous sol-gel silica films, and polymer–ceramic composite with controlled porosity, are some of the other systems that have been pursued with some success. The materials with the lowest dielectric constants are aerogels and xerogels, which have a porous matrix of either silicon dioxide or polymer. These materials have dielectric constants less than 1.5.

18.4.3 Board Materials

The major building blocks for the microsystems packaging are ultra-high density wiring on substrates, a reliable flip chip technology and embedment of passives within the dielectric layers on the substrate. The high-density interconnection on substrates is typically fabricated by a sequential deposition of alternate layers of copper metallization and polymer dielectric on a base substrate, as discussed before. The CTE mismatch between the dielectric layers and the substrate, coupled with the curing strain, the sequencing and asymmetry of the building layers, etc. induce severe stresses and warpage in the substrate. Analytic models show that a stiffer substrate will be needed to prevent the warpage. Covalent ceramics (carbides or nitrides) are some of the stiffest-known materials other than diamond.

Samsung, EX1028, p. 406

The other cornerstone for future technologies is the flip chip process involving interconnection of unpackaged integrated circuits directly onto low-cost organic substrates. A chief concern is the thermal expansion coefficient (CTE) mismatch between the silicon die and the substrate. A silicon die has an approximate coefficient of thermal expansion (CTE) of 2–3 ppm/°C, while the package has a CTE of approximately 18–20 ppm/°C. Though underfill materials are used to enhance flip chip reliability, it is clear that the final solution is to develop a substrate without any CTE mismatch. Typical low CTE materials used for packaging are glass ceramics, kevlar–polyimide composites and metal alloy cores sandwiched between organic materials (copper-invar-copper or copper-molybdenum-copper). Metal–metal composites (Al-Si) and metal–ceramic composites (Al-SiC) have been developed for various packaging applications. However, many researchers find that loading of SiC or Si in aluminum to lower the CTE below 6 ppm/°C is a major processing challenge. If materials can be tuned to have exact CTE as silicon but also have other positive attributes of FR-4 such as low-cost, ease to machine and drill, they can easily replace FR-4. Polymers with low-CTE (or negative CTE) fillers are being evaluated for this purpose. However, in order to attain the high modulus, it may be imperative to use other stiffer metal or ceramic matrix materials.

The other characteristic of substrate materials that significantly limit the processing is its T_g. Present packaging substrate materials such as FR-4 are continuously being improved by modifying the resin to obtain a higher T_g. Cyanate ester, or other high temperature polymers, are being developed for this. Significant changes in this regard are replacement of low-T_g bisphenol-type epoxies with the high T_g novolac-type or cyanate ester resins. Materials are also being modified to lower the high frequency losses. The requirements of CTE, modulus, T_g and dielectric loss seem to be more important than others. However, an ideal substrate material requires the following combination of properties:

- Close CTE match to silicon
- High modulus
- Low-dielectric constant
- Low loss tangent
- Low moisture uptake
- Low cost
- Temperature stability > 250°C
- Good thermal conductivity
- Compatibility with processes to incorporate microvias
- High adhesion (Cu/substrate and polymer to substrate)
- Good dimensional stability
- Flatness or planarity

18.4.4 Underfill Materials

Underfill encapsulation is used to reduce the shear stress of the solder joints between the chip and the substrate. The stresses are generated by the mismatch in the coefficient of

Samsung, EX1028, p. 407

thermal expansion. The encapsulants not only provide dramatic fatigue life enhancement with minimal impact on the manufacturing process, but also extend use to a variety of organic and inorganic substrate materials. Typical results show 10–100 fold improvement in fatigue life, compared to an unencapsulated package. Therefore, the underfill encapsulation has been key to the development of flip chip DCA technology.

More than 90% of the current underfill encapsulants dispense liquid encapsulants on one or two edges of the assembled flip chip package. This allows capillary action to draw the underfill into the gap between the chip and substrate of the assembled package. The main disadvantage of this process is that each step in the process—flip chip fluxing, reflowing, solder bumping, deflux cleaning and the process of underfilling and curing the encapsulants—is separated, which results in low production efficiency. No-flow underfilling process was invented to dispense the underfill materials on the substrate or the semiconductor devices first, then perform the solder bump reflow and underfill encapsulant cure simultaneously. Therefore, the no-flow process not only eliminates the strict limits on the viscosity of underfill materials, process temperature and package size, but also improves the production efficiency. The successful no-flow underfill material should meet the following primary requirements:

1. Minimal curing reaction at temperatures below the solder reflow temperature (210–230°C)
2. Rapid curing reaction after maximum solder bump reflow temperature
3. Good adhesion of underfill to the chip, substrate and solder joints
4. Lower shrinkage of the material during curing, lower CTE, and reasonable modulus to minimize the thermal stress resulting from the curing process and consequent cooling
5. Self-fluxing capability, passivating the substrate conductor oxides prior to the solder reflow (Reworkability is another major objective, particularly for wafer-level material development.)

To successfully rework the flip chip package, the cleaning of the underfill residue from the substrate is the key issue. The most common approach is to develop thermally or chemically degradable resins. Chemically degradable resins include functional groups that can cleave easily when reacted with a strong acid during the repair process to provide reworkability. Thermally degradable resins have in-situ polymerizable thermoplastic resins and thermally cleavable resins containing cross-linking sites.

To develop a no-flow underfill process for flip chip solder joint interconnects, several latent catalysts have been studied. Typically, metal chelate catalysts were reacted with epoxy resins (cycloaliphatic type epoxy), cross-linkers (anhydrides) or hardeners, and other additives such as adhesion promoters, silica fillers, self-fluxing agents and surfactants to form the low-cost high-performance underfills.

18.5 HOMEWORK PROBLEMS

1. Explain the following trends:
   - SiC has higher thermal conductivity than zirconia.
   - AlN has higher modulus compared to alumina.

Samsung, EX1028, p. 408

- Polymers have high CTE compared to metals and ceramics.
- Glass alumina has lower permittivity and lower CTE compared to pure alumina.
- Glass ceramics have lower thermal conductivity compared to pure ceramics.
- Polymers have lower density, higher moisture absorption compared to metals and ceramics.
- Metals have higher toughness compared to ceramics.

2. A 50-micron thick polymer film is deposited on a board (2 mm thick) and cured at 170°C (Figure 18.P2). At this temperature, the polymer is in a stress-free state. The board is now cooled to room temperature. Calculate the stress at the polymer–board interface (refer to Table 18.P2). Calculate the warpage that occurs in the board. How much should the modulus of the board be, so that it is planar at least by 0.1% ($dy/L < 0.001$)? Assume that the T_g of the board and polymer is 120°C. How much should the adhesion strength be so that the film would not delaminate? Refer to some elementary books on mechanics of composites to get the relevant equations.
3. A capacitor is to be made from a dielectric having a breakdown strength of 2×10^6 V/m and relative dielectric constant of 1000. The electrodes are metal plates fixed to the sides of the dielectric (0.2 mm thick). Due to the slight warpage in one of the electrode plates, 30% of the electrode area is separated by an air-filled gap of 1 micron. The remaining 70% is in intimate contact with the dielectric. The breakdown field in the air is 3×10^6 V/m. Discuss the effect of air-gap on the capacitance. What is the breakdown voltage for this device?
4. Find the specific capacitance of a layer of composite with dielectric constant 60 and thickness of 2 microns. If the required capacitance is 40 nF/cm^2, calculate the dielectric constant required for this film. If you can choose a thin-film vapor deposition technique to get a 0.2 micron, what dielectric constant would be sufficient? List three materials available in this dielectric constant range that can be deposited using a simple PVD process. On what basis would you pick up the right candidate from these three materials? When would you prefer to use a MOCVD?
5. Calculate the resistance across a copper wire of 0.05 mm diameter and 5 cm long. Calculate the time a signal takes to travel from one end of the wire to other end when the wire is in a medium having dielectric constant of 2.5. How does it get affected when the medium has dielectric constant of 1.5?
6. To make a slurry of ceramic for making a multilayered substrate by tape casting, if 50 grams of glass ceramic are added (density 3.5 g/cc) to 6.0 g of water with an appropriate

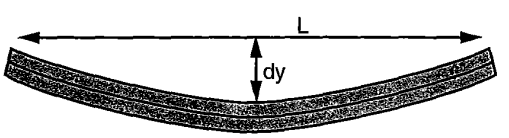


FIGURE 18.P2 Geometry of warpage during thin-film build-up on a substrate (Homework Problem 2).

TABLE 18.P2 Useful thermomechanical data for Homework Problem 2.

	CTE (ppm/°C)	Modulus (GPa)
Properties of board (below T_g)	18	20
Properties of polymer (below T_g)	60	2.5
Properties of board (above T_g)	8	5
Properties of polymer (above T_g)	200	.5

Samsung, EX1028, p. 409

dispersant, what would the solids loading be? What would be the disadvantages faced during subsequent processing, if more water had to be used for making the dispersion? How would you minimize the solvent, and hence increase the solids loading?

7. If a chip radiating heat at 10 W/cm^2 cannot operate at temperatures above 50°C, and assuming that all heat should conduct through a heat sink attached to the top of a chip, and is a 0.2-cm thick slab, calculate the thermal conductivity of the heat sink. Assume that the room temperature is 20°C. Other than heat conduction through a solid, what are the ways to remove the heat?
8. Calculate the required carbon fiber content in a polymer matrix, so that the modulus of the composite is 200 GPa. Assume that the polymer modulus is 2.5 GPa and the modulus of carbon fibers is 600 GPa. Assume continuous fibers and use simple rules of mixtures for parallel and series models. Refer to some elementary books on mechanics of composites to get the relevant equations.
9. Nanocarbon fillers and carbon fiber cloth are used to reinforce the epoxy resin to make a stiff substrate. Carbon nanofillers are first added to the epoxy, which is then impregnated into a carbon cloth to make the laminates. By using rules of mixtures, determine the composite Young's modulus. Density of carbon fibers is 2.12 g/cc; density of epoxy is 1.3 g/cc. Modulus of carbon fibers is 600 GPa and epoxy is 2.5 GPa. Weight of a 10 cm^2 composite is 20 g; weight of carbon cloth = 140 g/cm^2. The number of cloths in the composite is 3.
10. How is the contact angle relevant in understanding wetting? For better wetting, is a lower or higher contact angle required? How would you lower the contact angle and improve wetting? Find out one technique to measure contact angle.
11. FR-4 printed wiring board (PWB) is a composite material made of layers of woven glass fiber impregnated with epoxy. Assuming that the thermal expansion of glass fiber and epoxy are 4 ppm/K and 50 ppm/K, estimate the thermal expansion of the board in the (x-y) plane as well as in the normal direction (thickness). How would you expect this to change above the glass transition temperature of the epoxy.
12. Most organic materials will absorb a few percent or more of humidity in a humid environment. How would you expect this to change the effective dielectric constant of the material and the dielectric loss?
13. Give the names of three different plastics that are used extensively in electronics. Describe one important use for each of these materials in electronics.
14. Define and explain the concept glass transition temperature for a plastic material. It is recommended to make use of a graphic presentation to illustrate the explanation. Give the typical glass transition temperature for epoxy and polyester.
15. Explain the process technology for transfer of patterns by screen-printing and stencil-printing as a production technology for electronics. Explain the principle of subtractive process by wet-chemical etching. Name an etchant used for copper etching.
16. Describe a commonly used manufacturing technology for a single layer through-hole printed wiring board. This is best done by outlining a flow chart with a supplemental text for each process step. Similarly, explain a commonly used manufacturing technology for double-sided printed wiring boards with surface mount devices (SMD) on both sides of the board.

18.6 SUGGESTED READING

1. ASM International: *Electronic Materials Handbook,* Vol. 1, 1989.
2. P. Bendale, et al. "Preparation of Barium Titanate Films at 55 C by an Electrochemical Method." *Journal of the American Ceramic Society,* Vol. 76, No. 10, 1993, 2619–2627.
3. R. C. Buchanan. *Ceramic Materials for Electronics*. New York: Marcel Dekker, 1986.

Samsung, EX1028, p. 410

4. C. Coombs. *Printed Circuits Handbook,* 3rd Edition. New York: McGraw Hill, 1988.
5. F. Fowkes. "Acid-Base and Electrostatic Contributions to the Adhesion of Polymers to Metals and Ceramic Substrates." *Materials Research Society Symposia Proceedings,* Vol. 40, Pittsburgh, PA: Materials Research Society Publications, pp. 239–240.
6. R. E. Hummel. *Electronic Properties of Materials.* New York: Springer-Verlag, 1992.
7. D. Johnson. "Sol-gel Processing of Ceramics and Glass." *American Ceramic Society Bulletin,* Vol. 64, No. 12, 1985, 1597–1602.
8. S. Komarneni. "Nanocomposites." *Journal of Materials Chemistry,* Vol. 2, No. 12, 1992, 1219–1230.
9. J. Liu. *Conductive Adhesives for Electronics Packaging.* British Isles: Electrochemical Publications Ltd., 1999.
10. D. Pollock. *Physical Properties of Materials for Engineers.* Boca Raton: CRC Press, Inc., 1982.
11. D. Riemer: "Analytical Engineering Model of the Screen Printing Process." *Solid State Technology,* August/September 1988.
12. D. Shanefield. *Organic Additives and Ceramic Processing.* Boston: Kluwer Press, 1995.
13. S. Sivaram. *Chemical Vapor Deposition, Thermal and Plasma Deposition of Electronic Materials.* New York: Van Nostrand Reinhold, ITP Inc., 1995.
14. E. Slamovich and I. Aksay. "Structure Evolution in Hydrothermally Processed (<100°C) $BaTiO_3$ Films." *Journal of American Ceramic Society,* Vol. 76, No. 1, 1996, 239–247.
15. P. Townsen, D. Barnett, and T. Brunner. "Elastic Relationships in Layered Composite Media with Approximation for the Case of Thin-Films on a Thick Substrate." *Journal of Applied Physics,* Vol. 62, No. 11, 1987.
16. R. Tummala, E. Rymaszewski, and A. Klopfenstein, eds. *Microelectronics Packaging Handbook.* New York: Chapman and Hall, 1997.
17. L. Vilps and D. Kerps. "Simple Stress Formula for Multilayered Thin Films on a Thick Substrate." *Journal of Applied Physics,* Vol. 55, No. 3, 1536–1537.
18. J. H. Vincent and G. Humpston. "Lead-free Solders for Electronic Assembly." *GEC Journal of Research,* Vol. 11, No. 2, 1994, 76–89.
19. R. Klein Wassink. "Soldering in Electronics." *Electrochemical Publications,* 2nd Edition, 1989.
20. C. P. Wong and D. Baldwin. "Novel No-Flow Underfills for Low-Cost Flip Chip Applications: Materials and Processes." *Future Circuits International,* Issue 3, 1998, 67–70.
21. M. Yoshimura and W. Suchanek. "In Situ Fabrication of Morphology-Controlled Advanced Ceramic Materials by Soft-Solution Processing." *Solid State Ionics,* Vol. 98, 1997, 197–208.
22. H. Yun, et al. "Adhesion Improvement in Copper/Epoxy Joints." *Journal of Materials Science,* Vol. 27, 1992, 5811.

ACKNOWLEDGMENTS

The authors wish to thank Professor Janet K. Lumpp of the University of Kentucky for reviewing the chapter and providing useful suggestions. The editor gratefully acknowledges the contribution of Professor Janet K. Lumpp of the University of Kentucky for critically reviewing this chapter.

Samsung, EX1028, p. 411

GLOSSARY

accelerated stress test A test conducted at a stress, e.g., chemical of physical, higher than that encountered in normal operation, for the purpose of producing a measurable effect, such as a fatigue failure, in a shorter time than experienced at operating stresses.

accelerator An organic compound which is added to an epoxy resin to shorten the cure time.

active components Electronic components, such as transistors, diodes, electron tubes, thyristors, etc., which can operate on an applied electrical signal so as to change its basic characteristics, i.e., rectification, amplification, switching, etc.

active trimming The process of trimming components such as resistors while the circuit is under power. Such components are fabricated directly onto the substrate of a hybrid or multichip module, and the trimming is usually performed using a laser beam.

additive plating Processing a hybrid circuit substrate by sequentially plating conductive, resistive, and insulative materials, each through a mask, thus defining the areas of traces, pads, and elements.

additive process A process in which conducting material is added to specific areas of a substrate. Groups of tracks, individual tracks, or portions of tracks can be built up to precise thicknesses by iterating the process multiple times with selective masking.

advanced statistical analysis program (ASTAP) The "Advanced Statistical Analysis Program" is the IBM circuit analysis simulation program. It performs DC, time domain, and frequency domain simulations. Statistics can be applied to all simulations to predict operating tolerances. Among its many features is a transmission line analysis program. See also Spice.

alloy A solid-state solution or compound formation of two or more metals. Alternatively, a combination of metals resulting in a phase or phases containing some of each constituent.

alpha particle Decay product of some radioactive isotopes. It is a high-energy (mv range) helium nucleus capable of generating electron/hole pairs in microelectronic devices and switching cells, causing soft errors in some devices.

analog circuit A collection of components used to generate or process analog signals.

analog-to-digital (A/D) The process of converting an analog value into its digital equivalent.

anisotropic adhesive Special adhesives, which contain minute particles of conductive material. These adhesives find particular application with the flipped-chip techniques used to mount bare die on the substrates of hybrids, multichip modules, or circuit boards. The conducting particles are only brought in contact with each other at the sites where the raised pads on the die are pressed down over their corresponding pads on the substrate, thereby forming good electrical connections between the pads.

application specific integrated circuit (ASIC) An integrated circuit chip with personalization customized for a specific product. Personalization refers to wiring on the integrated circuit chip.

area array TAB Tape automated bonding where edge-located pads and additional pads on the inner surface area of a chip are addressed in the bonding scheme. This is practiced with extremely complex dice, VLSI etc. Also for use with ICs where peripheral pad pitch cannot be further reduced and all I/Os must be accommodated.

Samsung, EX1028, p. 412

array A group of elements (pads, pins) or circuits arranged in rows and columns on one substrate.

aspect ratio The ratio of the length of the hole to the diameter of the hole in a board.

assembly A hybrid circuit which includes discrete or integrated components that have been attached to the next level of package, usually a card.

assembly/rework Terms denoting joining and replacement process of microelectronic components. Assembly refers to the initial attachment of device and interconnections to the package. Rework refers to the removal of a device, including interconnections, preparation of the joining site for a new device, and rejoining of the new device. Rework is necessary for either repair or engineering change.

backbonding Bonding active chips to the substrate using the back of the chip, leaving the face, with its circuitry, face up. The opposite of backbonding is face down bonding.

back-end-of-the-line (BEOL) The portion of the integrated circuit fabrication where the active components (transistors, resistors, etc.) are interconnected with wiring on the wafer. It includes contacts, insulator, metal levels, and bonding sites for chip-to-package connections. Dicing the wafer into individual integrated circuit chips is also a BEOL process. The front-end-of-the-line (FEOL) denotes the first portion of the fabrication where the individual devices (transistors, resistors, etc.) are patterned in the semiconductor.

backplane The medium used to interconnect a number of circuit boards. Typically refers to a special, heavy-duty printed or discrete wired circuit board.

backside metallurgy (BSM) A metallization pad electrically connected to internal conductors within a multilayered ceramic package, to which pins are brazed.

ball grid array (BGA) An area array of solder balls joined to a SCM or MCM and used to electrically and physically connect the package to the next level of package, usually a printed circuit board.

ball limiting metallurgy (BLM) The solder wettable terminal metallurgy, which defines the size and area of a soldered connection, such as C4 and a chip. The BLM limits the flow of the solder ball to the desired area, and provides adhesion and contact to the chip wiring.

bandwidth The maximum pulse rate or frequency that can reliably propagate through a transmission line. For a data bus, bandwidth is commonly used to describe the maximum data rate, which is the single line pulse rate multiplied by the number of parallel bus bit lines.

bare Die An unpackaged integrated circuit.

BiCMOS A technology in which the function of each logic gate is implemented using low-power CMOS, while the output stage is implemented using high-drive bipolar transistors.

BIFET The combination of bipolar and FET transistors integrated together on the same piece of silicon for enhanced performance and cost.

binder Materials (organic) added to thick-film compositions and to unfired substrate materials to give sufficient strength temporarily for prefire handling.

BiNMOS A relatively new low-voltage integrated circuit technology in which complex combinations of bipolar and NMOS transistors are used to form sophisticated output stages, providing both high speed and low static power dissipation.

bipolar junction transistor (BJT) A family of transistors.

bipolar transistor Contains two p-n diodes, one between the base and the emitter, and one between the base and the collector. Minority carriers are injected in the base from the emitter across the base-emitter junction, travel through the base, and are attracted to the collector through the base-collector junction. These transistors consume more power than Field-Effect Transistors (FET), but also achieve higher performance.

block copolymer A copolymer compound resulting from the chemical reaction between n number of molecules, which are a block of one monomer, and n number of molecules, which are a block of another monomer. Example: stearine (rigid) with silicone (elastic).

BIST (built-in self-test) A test strategy in which additional logic is built into a component, thereby allowing it to test itself.

blind via A via that is only visible from one side of the substrate.

board The package element can be best defined as an organic printed-circuit card or board on which smaller cards or modules can be mounted. Its connections to the next higher level involve discrete wire or cables.

boiling Phase change and formation of bubbles in a superheated liquid.

bondability Those surface characteristics and conditions of cleanliness of a bonding area which must exist in order to provide a capability for successfully bonding an interconnection material by one of several methods, such as ultrasonic or thermocompression wire bonding.

braze A joint formed between two different materials by formation of liquid at the interface.

Samsung, EX1028, p. 413

brazing Joining of metals by melting a nonferrous, filler brazing metal, such as eutectic gold-tin alloy, having a melting point lower than that of the base metals. Also known as soldering.

BTAB The acronym for tape automated bonding when the raised bump for each bond site is prepared on the tape material, as opposed to the bump being on the chip.

bumped tape A tape for the TAB process where the inner-lead bond sites have been formed into raised metal bumps on the tape, rather than on the chip. This ensures mechanical and electrical separation between inner lead bonds and the non-pad areas of the chip (die) being bonded.

buried via A via used to link conducting layers internal to a substrate. Such a via is not visible from either side of the substrate.

burn-in The process of electrically stressing a device (usually an elevated temperature and voltage environment) for an adequate period of time to cause failure of marginal devices.

burn-off Removal of unwanted materials, typically organics from greensheets, or organic contamination from substrates.

cache memory A small, high-speed, memory (usually SRAM) used to buffer the central processing unit from any slower, lower cost memory devices such as DRAM. The high-speed cache memory is used to store active data, while the bulk of the data resides in the slower memory.

camber A term that describes the amount of overall warpage present in a substrate.

capacitance The electrostatic element that stores charge. In packaging systems, it is used in lumped equivalent circuits to represent part of a line discontinuity. It is also used in a distributed system to represent the electrostatic storage property of a transmission line. Because it delivers current in response to a change in voltage, another use is to filter powering systems.

card A printed-circuit panel (usually multilayer) that provides the interconnection and power distribution to the electronics on the panel, and provides interconnect capability to the next level package. It is also known as a daughter board. It plugs into a mother printed-circuit board.

card-on-board Packaging technology in which multiple printed-circuit panels (cards) are connected to printed-circuit panel (board) at 90° angles.

catalyst A substance that initiates a chemical reaction under different conditions (such as lower temperatures) than would otherwise be possible. The catalyst itself remains unchanged at the end of the reaction.

cell design, standard A semicustom product implemented from a fully diffused or ion implanted semiconductor wafer carrying horizontal rows of primary cells, interlaced with wiring channels (bays). Vertical wiring is supplied by additional processed layers, which may use the cell areas or lie in channels on an overhead layer. Channel widths may vary to suit particular chip logic, so that chip sizes are not fixed for all products of a family.

central processor (CP) Computer processor responsible for fetching, interpreting, and executing program instructions. Also called Processor Unit (PU) and Central Processing Unit (CPU).

ceramic Inorganic, nonmetallic material, such as alumina, beryllia, or glass-ceramic, whose final characteristics are produced by subjection to high temperatures. Often used in forming ceramic-substrates for packaging semiconductor chips.

ceramic ball grid array (CBGA) A ceramic package using ball grid array technology. See Ball Grid Array technology.

ceramic column grid array (CCGA) A ceramic package using ball grid array technology. See Ball Grid Array technology.

ceramic dual-in-line package (DIP) Dual-in-line package in ceramic. See Dual-in-line Package.

ceramic quad flat pack (CQFP) Quad Flat Pack in ceramic. See Quad Flat Pack.

cermet A solid homogenous material usually consisting of a finely divided admixture of a metal and ceramic in intimate contact. Cermet thin films are normally combinations of dielectric materials and metals.

channels Provide communications paths for input and output from the computer system.

characteristic impedance (Z_0) The voltage-to-current ratio of an electric signal propagating through an infinitely long transmission line. If L denotes the inductance per unit length and C denotes the capacitance per unit length, then $Z_0 = (L/C)^{0.5}$.

chemical vapor deposition (CVD) Depositing circuit elements on a substrate by chemical reduction of vapor of volatile chemical in contact with the substrate.

chemorheology The study of the processability or flow (rheology) and the chemistry of the polymer system. Processability parameters include, for instance, hold temperatures, injection speeds, and compaction pressures. The chemical aspect, on the other hand, involves the rate of

Samsung, EX1028, p. 414

reaction, the mechanisms, the kinetics, and the cessation of the chemical reaction at the end of the polymerization.

chip The uncased and normally leadless form of an electronic component part, either passive or active, discrete or integrated. Also referred to as a die.

chip carrier A special type of enclosure or package to house a semiconductor device. It has electrical terminations around its perimeter, or solderpads on its underside, rather than an extended lead frame or plug-in pins.

chip design, depopulated A gate array or standard cell array chip in which the wiring capacity (and hence chip area) is deliberately chosen to make automatic wiring possible only for those chips having some amount less than their maximum possible logic cell occupancy. This increases wafer productivity and circuit placement flexibility.

chip-on-board (COB) One of the many configurations in which a chip is directly bonded to a circuit board or substrate. These approaches include wirebonding, TAB, or solder interconnections, similar to the C4 structure. In low-end and consumer systems, chip-on-board generally refers to wirebonding of chips directly to board. See also Direct Chip Attach (DCA).

chip-on-chip (COC) A process in which unpackaged integrated circuits are mounted on top of each other. Each die is very thin and it is possible to have over a hundred die forming a 3D cube.

chip-on-flex (COF) Similar to chip-on-board (COC), except that the unpackaged integrated circuits are attached to a flexible printed circuit.

cladding Thin layer of a corrosion-resistant metal coating bonded to a metal core, usually by heating or rolling. Typical examples are steels clad with stainless steel, nickel alloys, or copper alloys. Copper cladding on both sides of invar is accomplished this way.

clock skew A cycle time adder caused by the amount of tolerance associated with the clock signal arrival times at all of the system latch inputs.

coated-metal core substrate A substrate consisting of an organic or inorganic insulation coating bonded to metal. Insulated surface or surfaces are used for circuit deposition.

coefficient of thermal expansion (CTE) The ratio of the change in dimensions to the change in temperature-per-unit starting length, usually expressed in cm/cm/°C. The acronyms TCE and CTE are synonymous.

Coffin-Manson equation A commonly used formula, first proposed by S. S. Manson and L. F. Coffin, relating to the fatigue lifetime of a metal to the imposed strain amplitude. Others have extended the formula to incorporate time and temperature dependent phenomena.

cofiring Processing thick-film conductors and dielectrics through the firing cycle at the same time to form multilayer structures.

cofired ceramic A substrate formed from multiple layers of "green" ceramic that are bonded together and fired at the same time.

colorant An inorganic or organic compound that is added to a polymeric resin to impart a desired color.

column grid array (CGA) A column grid array is an area array of solder columns joined to a SCM or MCM, and used to electrically and physically connect the package to the next level of package, usually a printed circuit board. A column grid array is used when the package performance requires a higher reliability than is provided with the similar ball grid array.

compliant bond A bond that uses an elastically and/or plastically deformable member to impart the required energy to the lead.

complimentary metal-oxide semiconductor (CMOS) This refers to logic in which cascaded field effect transistors (FET) of opposite polarity are used to minimize power consumption.

compression seal A seal between an electronic package and its leads. The seal is formed as the heated metal, when cooled, shrinks around the glass insulator, thereby forming a tight compression joint.

conduction Thermal transmission of heat energy from a hotter region to a cooler region, in a conduction medium.

conduction adhesive An adhesive material, usually epoxy, that has metal powder added to increase electrical conductivity, usually conductor added is silver.

conductive epoxy An epoxy material (polymer resin) that has been made conductive by the addition of a metal powder, usually gold or silver. Best common conductors are silver, copper, and gold. See also Superconductor, and Conductor Adhesive.

conductive ink technology A technique in which tracks are screen printed directly onto the surface of a circuit board using conductive ink.

conductor, electrical A class of materials that conduct electricity easily. They have very low resistivity, which is usually expressed in micro-ohm-cm. The best conductors include silver, copper, gold, and superconducting ceramics.

conductor, thermal A class of materials such as copper, aluminum, and beryllia, that conduct heat.

Samsung, EX1028, p. 415

conformal coating A thin nonconductive coating, either plastic or inorganic, applied to a circuit for environmental and/or mechanical protection.

connections The connections belonging to nets interconnection logic units on a given package level—connecting it to the next higher package level.

connectivity See Wiring Density.

contact angle The angle between the bonding material, usually a liquid-like solder, and the bonding pad. Also called Wetting Angle.

controlled collapse chip connection (C4) A solder joint connecting a substrate and a flip-chip, where the surface tension forces of the liquid solder support the weight of the chip and controls the height (collapse) of the joint.

controlling collapse Controlling the reduction in height of the solder balls in a flip-chip processing operation.

convection Transmission of thermal energy from a hotter to a cooler region through a moving medium, such as air over water.

coplanar leads (flat leads) Ribbon-type leads extending from the sides of the circuit package, all lying in the same plane for surface mount applications.

copolymer A compound, resulting from the chemical reaction and polymerization of two chemically different monomers. The resulting larger molecules contain repeating structural units of the original molecules.

cordierite A crystalline ceramic material of composition $2MgO\text{-}2Al_2O_3\text{-}5SiO_2$ that can be crystallized from glass of same composition, or sintered from powders.

coupled noise (same as cross talk) The electromagnetic and electrostatic linkages between two nearby conductors that allow one line to induce a signal on the other. See also Cross Talk.

coupler A chemical agent, frequently an organosilane, used to enhance the bond between a resin and a glass reinforcement.

coupling capacitor A capacitor used to block dc signals, and to pass high-frequency signals between parts of an electronic circuit.

crazing Fine cracks that may extend on or through layers of plastic or glass materials.

creep Nonrecoverable deformation proceeding at relatively low strain rates, less than about 10^{-6}/sec, usually associated with sufficiently high temperatures to allow significant rates of diffusion.

crossover The transverse crossing of metallization paths without mutual electrical contact. This is achieved by the deposition of an insulating layer between the conducting paths at the area of crossing.

cross talk Signals from one line leaking into another nearby conductor because of capacitance, or inductive coupling, or both (i.e., owing to the capacitance of a thick-film crossover).

crystallization Formation of crystalline phase out of amorphous material during high-temperature processing. Undesirable or uncontrollable crystallization is called divitrification.

cumulative distribution function (CDF) Distribution of a parameter as a fraction of the total number of measurements with respect to a statistic, e.g., "probits," or standard deviations relative to a particular statistical distribution.

cure To harden a material using heat, ultraviolet light, or some other process.

curing agent An inorganic or organic compound that initiates the polymerization of a resin.

curing cycle For a thermosetting material, commonly a resin compound such as a bonding adhesive, it is the combination of total-temperature profile to achieve the desired result; for example, the complete irreversible hardening of the material, resulting in a strong bond.

current carrying capacity The maximum current that can be continuously carried by a circuit without causing objectionable degradation in electrical or mechanical properties.

current slew rate The rate of change in current with respect to time (di/dt).

custom design A form of design in which the choice and arrangement of components and wiring on a package may vary arbitrarily within tolerances from a regular array.

cycle time Unit of time in which elements of the central processor complete their logical functions. Some elements will require more than one cycle to complete a function. See Cycles per Instruction.

cycles per instruction The number of cycles required to process an instruction.

decoupling capacitor A shunt-placed capacitance that is used to filter transients on a power distribution system.

deep sub-micron Typically taken to refer to integrated circuits containing structures which are smaller than 0.5 microns.

delamination Occurs when a composite material formed from a number of layers is stressed, thermally or otherwise, such that the layers begin to separate.

Samsung, EX1028, p. 416

delay equations A set of mathematical terms that are used to predict the propagation times between driving and receiving circuits that are interconnected through signal wires. These equations are usually derived from simulation data using numerical curve fitting techniques.

delta-I noise (ΔI) See Switching Noise

design limits The fail points that are incorporated into the hardware design rules that drive the computer-aided design system. See Noise Rules and Wiring Rules.

devitrification The undesirable formation of crystals in glass during firing. The desirable process is called crystallization.

dew point The temperature at which moisture at a given partial pressure becomes saturated, and when cooled below which, it condenses.

die Integrated circuit chip as cut (diced) from finished wafer. See chip.

die bond Mechanical attachment of silicon to substrate usually by solder, epoxy, or gold-silicon eutectic, including interface metallurgies on chips and substrates. The die bond made to the back (inactive) side of the chip with the circuit side (face) up.

die stacking A technique used in specialist applications in which several bare die are stacked on top of each other to form a sandwich. The die are connected together and then packaged as a single entity.

dielectric Material that does not conduct electricity. Generally used for making capacitors, insulating conductors (as in crossover and multilayered circuits) and for encapsulating circuits.

dielectric constant The term used to describe a material's ability to store charge when used as a capacitor dielectric. It is the ratio of the charge that would be stored with free space to that stored with the material in question as the dielectric.

dielectric layer (1) An insulating layer used to separate two signal layers. (2) An insulating layer used to modify the electrical characteristics of an MCM-D substrate.

dielectric loss The power dissipated by a dielectric as the friction of its molecules opposes the molecular motion produced by an alternative electric field.

differential scanning calorimetry (DSC) A technique for measuring the physical transitions of a polymer as a function of temperature, compared to another material undergoing a similar heating process but not undergoing any transitions or reactions. DSC uses a servo system to supply energy at a varying rate to both sample and reference, so as to keep their temperatures equal. A DSC output plots energy supplied vs. average temperature.

digital-to-analog (D/A) The process of converting a digital value into its analog equivalent.

diode A two-terminal device that only conducts electricity in one direction; in the other direction it behaves like an open switch. The term diode is typically taken to refer to a semiconductor device, although alternative implementations such as vacuum tubes are available.

direct access storage device (DASD) Computer storage hardware subsystem that uses magnetic recording on a rotating disk surface. Access to the information is accomplished with the use of a moveable arm, which positions one or more read/write heads along the radius of the disk to the desired track.

direct chip attach (DCA) A name applied to any of the chip-to-substrate connections used to eliminate the first level of packaging. See also Chip-on Board.

discrete component Individual components or elements, such as resistors, capacitors, transistors, diodes, inductors, and others, as self-contained entities.

discrete device Typically taken to refer to an electronic component such as a resistor, capacitor, diode, or transistor that is presented in an individual package. More rarely, the term may be used in connection with a simple integrated circuit containing a small number of primitive gates

distance to neutral point (DNP) The separation of a joint from the neutral point on a chip. This dimension controls the strain on the joint imposed by expansion mismatch between chip and substrate. The neutral point is usually the geometric center of an array of pads, and defines the point at which there is no relative motion of chip and substrate in the *X-Y* plane during thermal cycling.

doctor blade A method of casting slurry into a thin sheet by the use of knife blade placed over moving carrier to control slurry thickness.

double-sided substrate A substrate carrying active circuitry on both its topside and bottom-side, electrically connected by means of metallized through-holes, or edge metallization, or both.

DRAM (dynamic RAM) A memory device in which each cell is formed from a transistor-capacitor pair. Called dynamic because the capacitor loses its charge over time, and each cell must be periodically recharged if it is to retain its data.

driver The off chip circuit that supplies the signal voltage and current to the package lines. Also called an output buffer circuit.

Samsung, EX1028, p. 417

dry film photoresist Photoresist material that is processed dry, usually by lamination of prefabricated film.

dry pressing Pressing and compacting together of dry powdered materials with additives in rigid die molds under heat and pressure to form a solid mass, usually followed by sintering to form shapes.

dual-in-line package (DIP) A package having two rows of leads extending at right angles from the base and having standard spacing between leads and between rows of leads. DIPs are made of ceramic (Cerdip) and plastic (Pdip).

dynamic flex A form of flexible circuitry developed for applications where continued flexure is necessary. In contrast, static (flex) once installed, remains fixed.

dynamic random access memory (DRAM) Electronic information storage that employs transient phenomena, typically charge stored in a leaky capacitor. Refresh cycles are required to restore, and thus maintain the information. DRAM is the simplest and least expensive of electronic memories, but is also the least impressive performer.

effective inductance (LEFF) A simplified characterization of the goodness of an AC power distribution system. It consists of a single lumped inductance that, when multiplied by the total current slew rate, predicts the total switching noise across the circuit load.

electrically long transmission line One in which the delay from the near-end to the far-end is greater than one-half of the near-end signal's transition time. When this occurs, the reflections from the far-end do not distort the near-end signal during its transition time.

electrically short transmission line One in which the delay from the near-end to the far-end is less than one-half of the near-end signal's transition time. For this case, reflections from the far-end interfere with the near-end transition waveform, causing distortion that usually increases the net delay.

electroless plating Metal deposition, usually in an aqueous medium, which proceeds by an exchange reaction between metal complexes in the solution and the particular metal to be coated; the reaction does not require externally applied electric current.

electromigration (1) A process in which structures on an integrated circuit's substrate are eroded by the flow of electrons in much the same way as land is eroded by a river (also known as subatomic erosion). (2) The process of forming transistor-like regions in a semiconductor, using an intense magnetic field.

electron beam lithography An integrated circuit fabrication process in which fine beams of electrons are used to draw extremely high-resolution patterns directly into the resist without the use of a mask.

electroplating Deposition of an adherent metallic coating onto a conductive object placed onto an electrolytic bath composed of a solution of the salt of the metal to be plated. Using the terminal as the anode (possibly of the same metal as the one used for plating), a DC current is passed through the solution, affecting transfer of metal's ions onto the cathodic surface.

electrostatic discharge (ESD) Discharge of static charge on a surface or body through a conductive path to ground. An electronic component, or higher-level assembly, may suffer damage when it is included in the discharge path.

elongation The ratio of the increase in wire length at rupture, in a tensile test, to the initial length, expressed in percent.

emitter-coupled logic (ECL) Emitter-coupled logic is also known as current-switch (logic) circuits. In it, a current source feeds emitters of several transistors. The base of all but one acts as input terminals; the last base is connected to a reference voltage. Very popular circuit for high-performance applications, it is often combined with an emitter-follower output stage to further enhance its performance. It is then called SCEF, for current-switch emitter follower.

enameling A process that produces pore-free glass dielectric coating over a metal-core substrate.

encapsulation Sealing up or covering an element or circuit for mechanical and environmental protection.

end of life (EOL) The end of the useful operating life of a component or equipment determined by a "wear-out" or life terminating mechanism measured in units of time. EOL is usually specified as an objective in reliability calculations.

engineering change (EC) A change in design. An electrical design change is frequently implanted by cutting out or adding an electrical path to the manufactured hardware, e.g., laser deleting a line or adding a wire on a ceramic substrate.

entity A group of circuits separated from other circuits by a physical package boundary and associated input and output connections.

eutectic A term applied to the mixture of two or more substances with the lowest melting point possible between those components.

Samsung, EX1028, p. 418

etching The process of selectively removing any material not protected by a resist using an appropriate solvent or acid. In some cases the unwanted material is removed using an electrolytic process.

external resistance A term used to represent thermal resistance from a convenient point on the outside surface of an electronic package to an ambient reference point.

failure The temporary or permanent impairment of a device function caused by physical, chemical, mechanical, electrical, or electromagnetic disturbance or damage.

failure rate The rate at which devices from a given population can be expected (or were found) to fail as a function of time (e.g., %/1000 hr of operation).

fast wave propagation The transmission of energy along a signal line at the speed-of-light (velocity) expected for the dielectric structure. In the case of low-loss line the fast wave refers to the portion of the signal that travels at the velocity expected for the dielectric medium.

fatigue Used to describe the failure of any structure caused by repeated application of stress over a period of time.

ferroelectric A crystalline dielectric that exhibits dielectric hysteresis; an electrostatic analogy to ferromagnetic materials.

field effect transistor (FET) A transistor in which a voltage applied to a thin conductor over a thin insulator controls current flow in a semiconductor region (gate) or one polar type. This component originates and terminates in two regions of the opposite polar type located at either end of the gate region.

field replaceable unit (FRU) A component or subsystem of an electronic assembly, which may be replaced at the site of installation. A first or second level package is commonly a FRU for most computers.

filler A substance, usually ceramic or metal powder, used to modify the properties of fluids or polymers.

finite element modeling A computationally intensive numerical modeling tool in which the body is discretized into small regularly shaped elements.

first-incident switching The case that occurs when all of the receivers on a multi-drop net switch at the first time the signal arrives from the driver. Nets that are not first-incident are referred to as multi-reflection nets.

flame retarder An inorganic or organic compound added to a polymer mixture that causes the resulting plastic to self-extinguish after a flame is removed.

flat PAC An integrated circuit package having its leads extending from all four sides and parallel to the base.

flexible circuit carrier Printed circuits employing flexible substrates, processed by patterning copper onto thin flexible Kapton or polyimide films. Originally used only as a connector, now employed for multilayers.

flexible coating A plastic coating that is still flexible after curing.

FPC (flexible printed circuit) A specialist circuit board technology, often abbreviated to "flex," in which tracks are printed onto flexible materials. There are a number of flavors of flex, including static flex, dynamic flex, and rigid flex.

flexural strength Strength of a material measured by bending, typically used for brittle materials, such as glasses and ceramics, expressed in Mpa.

flip-chip A leadless, monolithic structure containing circuit elements, which is designed to electrically and mechanically interconnect to the hybrid circuit by means of an appropriate number of bumps, which are covered with a conductive bonding agent, located on its face. Alternatively, bonding of chips with connect pads, face down by solder connection. See also Controlled Collapse Chip Connection (C4).

floor planning A procedure in physical design which permits approximate shaping and placement of logic and memory circuit groupings on a package before final placement and wiring.

flow regime, laminar Flow where fluids layers are undisturbed and smooth.

flow regime, turbulent Flow where fluid particles are disturbed and fluctuate.

flux In soldering, a material that chemically attacks surface oxides so that molten solder can wet the surface to be soldered, or an inert liquid that excludes oxygen during the soldering process.

footprint The area occupied by a device mounted on a substrate.

fracture toughness A basic property of a material, or of an interface between dissimilar materials, describing its crack resistance in a mechanical or thermomechanical stress field.

free-space optical interconnect A form of optical interconnect in which laser-diode transmitters communicate directly with photo-transistor receivers without employing optical fibers or optical waveguides.

frit Glass composition ground up into a powder form and used in thick-film compositions as the portion of the composition that melts upon firing, to give adhesion to

Samsung, EX1028, p. 419

the substrate and hold the conductive composition together.

front-end-of-the line (FEOL) See back-end-of-the-line.

FR-4 Electronic Industries Association's designation for a fire retardant epoxy resin/glass cloth laminate. By common usage, the resin for such a laminate.

functional test A test strategy in which signals are applied to a circuit's inputs, and the resulting signals which are observed on the circuit's outputs are compared to known good values.

G-10 Is a grade of epoxy-impregnated glass cloth printed-circuit board material per NEMA (National Electrical Manufacturers Assoc.)

gallium arsenide (GaAs) A 3:5 valence high-speed semiconductor formed from a mixture of gallium and arsenic. GaAs transistors are approximately eight times faster than their silicon equivalents and use approximately one tenth of the power; but the material is difficult to manufacture and to work with.

gate array A semicustom product, implemented from a fully diffused or ion-implanted semiconductor wafer, carrying a matrix of identical primary cells arranged into columns with routing channels between them in the X and Y directions.

gate, logic Usually an electric circuit that combines information of its two inputs to form its output signal in accordance with the logic function it performs.

gate, sea of (see also gate array) A form of custom chip layout in which the wiring tracks required for interconnecting a fixed array of logic cells are disposed in rows and columns having widths measured in numbers of wiring tracks per row or column channel which vary to suit the local wiring demand, from point to point and from one logic product to another.

gate, structural Term used to designate a hinged frame which contains a number of boards and can swing out for servicing and/or access to the interior of a fixed frame.

glass Inorganic, nonmetallic and amorphous material obtained by melting oxide(s) into glass and retaining the structure by fast cooling.

glass − ceramic Inorganic, nonmetallic material obtained by controlled crystallization of glass into nonporous and fine microstructure.

glass + ceramic Inorganic, nonmetallic material obtained by admixing crystalline ceramic with glass and sintering composite.

glass ceramic Refers to family of glass − ceramic and glass + ceramic.

glass fabric Cloth woven of glass yarns, which are made of filaments.

glass transition temperature (T_g) In polymer or glass chemistry, the temperature corresponding to the glass-to-liquid transition, below which the thermal expansion coefficient is low and nearly constant, and above which is very high.

glazed substrate A glass coated ceramic substrate that effects a smooth nonporous surface.

glazing In the present context, glazing refers to the coating of a smooth, adherent layer of glass, such as a sealing glass, by melting the glass over metal or ceramic surfaces.

glob top A glob of encapsulant material surrounding a chip in the chip-on-board assembly process. The attached chip must have already passed pretest and inspection, because rework after final curing of the epoxy or silicone globs is virtually impossible.

global wiring Wiring interconnecting components mounted on a package (as opposed to the wiring inside the components). Also refers to that wiring independent of its detailed allocation to wiring tracks within a channel.

gold flash An extremely thin layer of gold with a thickness measured on the molecular level, which is either electroplated or chemically plated onto a surface.

green A term—unrelated to the actual material color, used in ceramic technology, meaning unfired. For example, a "green" substrate is one that has been formed, but has not been fired.

green ceramic Unfired, malleable ceramic.

green–sheet A composite organic-inorganic, flexible sheet ready for metallization, if desired, and lamination to form green substrates which upon removal of organics results in ceramic substrate. Green refers to the unfired state.

ground plane A conductive layer on a substrate, or buried within a substrate, that connects a number of points to one or more grounding electrodes.

guided probe A form of functional test in which the operator is guided in the probing of a circuit to isolate a faulty component or track.

guided-wave A form of optical interconnect, in which optical waveguides are fabricated directly on the substrate of a multichip module. These waveguides can be created using variations on standard opto-lithographic thin-film processes.

gullwing A common lead form used to interconnect surface mounted packages to the printed-circuit board.

Samsung, EX1028, p. 420

The leads, normally 100 to 250 Φm thick, are bent outward, downward, then again outward from the package body, providing feet for solder interconnection, and some degree of mechanical compliance.

hard glass Glasses having a high softening temperature (>700°), such as the borosilicate glasses used to seal feed through leads into metal packages.

heat flux The rate of flow of heat energy across or through a surface, measured in watts/cm^2.

heat sink The supporting member to which electronic components, their substrate, or their package bottom is attached. This is usually a heat conductive metal with the ability to rapidly transmit heat from the generating source (component).

hermetic Sealed so that the object is gas-tight. The test for hermicity is to fill the object with a test gas—often helium—and observe leak rates when placed in a vacuum. A plastic encapsulation cannot be hermetic as it allows gases to permeate.

high-level noise tolerance (NTH) The receiver noise tolerance that occurs when the input signal is in its UP state.

homogenous medium A signal propagation structure in which only a single dielectric is present.

hot-gas reflow The technique in which a heated gas, including air, is impinged on a site to be solder-reflowed, usually to form a solder interconnection.

hot-knife soldering The technique in which a heated blade (heated electrically or conductively) is used as a heat source for melting solder during package joining. The blade may be used to force mechanical contact throughout the joining process.

hybrid An electronic sub-system in which a number of integrated circuits (packaged and/or unpackaged) and discrete components are attached directly to a common substrate. Connections between the components are formed on the surface of the substrate, and some components such as resistors and inductors may be fabricated directly onto the substrate.

hybrid module A special carrier of hybrid microcircuits and other components interconnected as a unit, or as a component of an electronic subsystem. The module may be of single construction or made up of submodules, each usually with a compartment to house hermetically packaged hybrids and discrete passive component parts, such as transformers, axle-lead resistors, etc. Nonhermetic hybrid modules generally are parylene coated. In this book, hybrid module also refers to a module containing a combination of thick and thin films.

impedance The resistance to the flow of current caused by resistive, capacitive, or inductive devices (or undesired elements) in a circuit.

impregnation The process of coating a substrate, say glass cloth, with a resin solution and drying. The dried product is called prepreg.

inductance The electromagnetic element that stores flux lines. In packaging systems, it is used in lumped equivalent circuits to represent part of a line discontinuity. It is also used in a distributed system to represent the electromagnetic storage property of a transmission line. Because it induces an opposing voltage in response to a change in current, it causes package delta I (ΔI) noise.

inert atmosphere A gas atmosphere such as helium or nitrogen that is nonoxidizing or nonreducing to metals.

infrared reflow (IR) The technique in which primarily long wavelength light is used to heat solder joints to the melting temperature. Normally, a circuit board having prepositioned packages is transported through an IR reflow furnace.

inhomogeneous medium A signal propagation structure in which multiple dielectrics are present.

injection molded Molding by injecting liquefied plastic into a mold of desired shape.

injection molded card (IMC) Card for electronic packages made by injection molding of plastics into a mold cavity of desired shape.

input/output terminal (I/O) A chip or package connector (terminal) acting to interconnect the chip to the package or one package level to the physically adjacent level in the hierarchy. Usually refers to the number of contacts necessary to wire or interconnect an assembly. Pin out, connections, and terminals are other common words to describe the same. Care must be taken to differentiate between the total number of I/Os between levels, signal I/Os, I/Os used to distribute power, and reference I/Os.

insulation resistance (IR) The resistance to current flow when a potential is applied. IR is measured in megohms.

insulator metal substrate technology (IMST) A substrate, such as one made of porcelainized steel, which is not subject to size limitations and may have superior thermal dissipation characteristics. IMST refers to insulated-metal substrate technology of Sanyo. It is a single-sided aluminum core with epoxy coating and etched copper wiring.

insulators A class of materials with high resistivity. Materials that do not conduct electricity. Materials with

Samsung, EX1028, p. 421

resistivity values of over 10^6 Σ cm are generally classified insulators.

integrated circuit A microcircuit (monolithic) consisting of interconnected elements inseparably associated and formed in situ on or within a single substrate (usually silicon) to perform an electronic circuit function.

interchip wiring The conducting wiring path connecting circuits on one chip with those on other chips to perform a function.

interconnection The conductive path required to achieve connection from a circuit element to the rest of the circuit.

interface The boundary between dissimilar materials, such as between a film and substrate, or between two films.

internal resistance A term used to represent thermal resistance from the junction of a device, inside an electronic package, to a convenient point on the outside surface of the package.

interpenetrating polymer network (IPN) A polymer alloy made up of two or more crosslinked polymers. The networks interact with each other only through permanent physical entanglements rather than through covalent bonding as in a copolymer.

invar A trademark of International Nickel Co., Inc. for a very low thermal expansion alloy of nickel and iron.

ion implantation A process in which beams of ions are directed at a semiconductor to alter its type and conductivity in certain regions.

ion migration The movement of these ions within a material, or across the boundary between two materials, under the influence of an applied electric field.

isopak A unique pin-grid array consisting of Kovar pins sealed in glass-to-Kovar plate flush for chip bonding.

J-lead An IC package terminal lead configuration that resembles in crossection the letter J.

Josephson superconducting device Superconducting ceramics acting as Josephson devices, at very low temperatures, typically at 4°K.

keeper bar, tab A strip of dielectric material, such as polyimide, that remains attached to each row of outer-TAB leads following excise operation. It helps to reduce misalignment, non-planarity and damage during outer lead bonding.

kirkendahl voiding Voids induced in a diffusion couple between two metals with different interdiffusion coefficients.

known good die (KGD) IC semiconductor chips that have been tested before being packaged and are known to function as required.

kovar An alloy of iron (53%), cobalt (17%), and nickel (29%) with thermal expansion matching alumina substrate and certain sealing glasses. Most common lead frame and pin material.

laminate A material constructed from thin layers or sheets. Often used in the context of circuit boards.

lamination The process of consolidating sheets of prepreg under heat and pressure to form a solid product. Applied also to the consolidation of prepregs and precircuitized subcomposites to form a composite.

large-scale integration (LSI) Term used to designate chips with more than one thousand transistors.

laser diode A special semiconductor diode which emits a beam of coherent light.

laser soldering The technique in which heat to reflow a solder interconnection is provided by laser, usually a longer wavelength YAG or CO_2 laser. The joints are heated sequentially, then cooled rapidly.

lead frame A sheet metal framework on which a chip is attached, wirebonded and then molded with plastic.

leadless-chip carrier plastic (PLCC) A plastic package containing a chip that has terminal leads emanating from four sides. Each lead has a J-configuration and is designed for surface mounting to a printed-circuit board.

leaded-chip carrier (LLCC) A surface mounted package having metallized contacts as its periphery (rather than wire leads) which are soldered to metallized contacts on the printed-circuit board or substrate.

lift-off Patterning of metal by lift-off materials usually in a solvent.

line discontinuity A load point, consisting of a lumped equivalent circuit of resistance, capacitance, and inductance, anywhere on a transmission line that produces spurious reflections.

line loading Externally connected resistance, inductance, and capacitance, or combination of these on a transmission line.

line resistance Resistance of conductor lines in a package, measured in ohms per unit length or for a given cross section, ohms per square.

lines per channel The number of conductive lines between through holes in an organic board or ceramic substrate.

liquid crystal display (LCD) Display technology based on a liquid crystal materials whose light transmission is

Samsung, EX1028, p. 422

changeable by the application of an electrical field. LCD devices are used in numeric read outs and for flat screen television receivers.

logic design The process of determining the choice and interconnection of logic units (e.g., nands or nors) to accomplish logical functions in an overall digital system.

logic part A physical implementation of an interconnected and packaged group of logic circuits used in general more than once in a digital system.

logic primitive A basic logic function embodied as a single unit.

logic service terminal (LST) A terminal (on a package component) carrying logic signals, as opposed to one used only for electrical power.

loss-less transmission line A single path with total series resistance less than 10% of the characteristic impedance of the line. For this case, the signal level attenuation is approximately 105 (V_{out}/V_{in}) when the line is terminated in its characteristic impedance.

lossy transmission line A signal path with total series resistance that exceeds two times the characteristic impedance of the line. When this occurs, the fast-wave portion of the input signal is attenuated by more than "1/e" or 36% (V_{out}/V_{in}) when the line is terminated in its characteristic impedance.

low level noise tolerance (NTL) The receiver noise tolerance that occurs when the input signal is in its DOWN state.

low-fired cofired Similar in principle to standard cofired ceramic substrate techniques. However, low-fired cofired uses modern ceramic materials with compositions that allow them to be fired at temperatures as low as 650°C to 750°C. Firing at these temperatures in an inert atmosphere such as nitrogen allows non-refractory metals such as copper to be used to create tracks.

low-loss transmission Line A signal path where the total series resistance is greater than 0.1 times but less than two times the characteristic impedance of the line/ in this range. The signal level attenuation is between 10% and 63% (V_{out}/V_{in}) when the line is terminated in its characteristic impedance.

macro A collection of continuous cells defined as a group within a chip image.

manhattan distance (length) Wire length between terminals of a net or connection, measured in X or Y directions on a package wiring plane as on city blocks.

mask The photographic negative that serves as the master for making thick-film screens and thin-film patterns.

MCM-C A multi-chip module with a structure that consists of multi-layer cofired ceramic.

MCM-D A multi-chip module structure that consists of deposited thin film organic layers.

MCM-L A multi-chip module structure that consists of laminated organic layers gates in a device.

multilayer A printed circuit board constructed from a number of very thin single-sided and/or double-sided boards that are bonded together using a combination of temperature and pressure.

mean time to failure (MTTF) Applicable to individual parts or devices in reliability technology. It is the arithmetic average of the lengths of time-to-failure registered for parts or devices of the same type, operated as a group under identical conditions.

medium-scale integration (MSI) Refers to the number of logic.

memory adder An adder, to the basic cycle-per-instruction rate of the Central Processor (CP), due to requirements for data or instructions, not available in the CP when necessary.

metal migration An undesirable phenomenon whereby metal ions, notably silver, are transmitted through another metal across an insulated surface, in the presence of moisture and electrical potential.

metallization A film pattern (single or multilayer) of conductive material deposited on a substrate to interconnect electronic components.

metal-oxide semiconductor (MOS) A family of transistors where the controlling terminal is connected to a plate that is separated from the semiconductor by an insulating layer. This plate was originally made out metal (we now use polysilicon, or poly) and the insulator is an oxide, hence the "metal-oxide" appellation.

metallization layer A layer of conducting material on an integrated circuit that is selectively deposited or etched to form connections between logic gates. There may be several metallization layers separated by dielectric (insulating) layers.

metallized ceramic (MC) Ceramic (fired) substrate metallized with thick and thin films of metals.

micron A unit of length equal to a micrometer (μm).

microstrip line A signal line on the surface of a dielectric with air above it and a reference plane on the opposite side of the dielectric.

microstructure Structural features, such as crystal of phase boundaries, or defects in homogeneities within a solid, usually resolved at a high magnification.

Samsung, EX1028, p. 423

miner's rule An outgrowth of the cumulative damage concept, this formula predicts the fatigue lifetime of a structural element when the load history encompasses multiple types of various amplitudes of stress.

module A chip carrier on which the chip terminals are fed out by various means to terminals spaced to suit the spacing and dimensions of wires on the next higher level of package (i.e., card or board). It may also contain wiring planes and power planes interconnecting several of its chips, and thus be used as a card.

moire A technique to measure in-plane surface displacement, whereby a precision grid of fine lines attached to the surface is deformed relative to a stationary reference grid. The resulting fringe pattern can be analyzed to provide displacement values throughout the field.

mold release An organic compound added to a molding compound or powder that migrates to the mold surface to form a waxy layer between the plastic and mold metal, and to allow easy removal of the part from the mold.

MBE (molecular beam epitaxy) A technique for creating thin films on substrates in precise patterns, in which the substrate is placed in a high vacuum, and a guided beam of ionized molecules is fired at it, effectively allowing molecular-thin layers to be "painted" onto the substrate where required.

monolithic systems technology (MST) Ceramic package manufactured by screen printing Ag-Pd conductors onto alumina substrate swagged with pins and chips bonded with solder connection (C4).

Monte Carlo analysis A statistical analysis method whereby the resultant distribution is generated by using a random number generator to pick a large number of cases from the input distributions.

multichip module (MCM) A module or package capable of supporting several chips on a single package. Most multichip packages are made of ceramic.

multichip package An electronic package that carries a number of chips and interconnects them through several layers of conductive patterns. Each one is separated by an insulative layer and interconnected via holes.

multilayer ceramic (MLC) Ceramic substrate consisting of multiple layers of metals and ceramics interconnected with vias.

multilayer substrates Substrates that have buried conductors so that complex circuitry can be handled, using assembly processes similar to those used in multilayer ceramic capacitors. Fabricated either as a conventional MLC, or a cofired multilayer ceramic (CMC) hybrid structure, in high and low temperature versions.

multi-reflection switching A network that is not first incident. See first-incident switching.

N-type A piece of semiconductor doped with impurities that make it amenable to donating electrons.

near infrared reflectance analysis (NIRA) Infrared spectroscopy that covers the region from 0.75 μm to 2.5 μm. Regular infrared spectrophotometers use a glowing light source to provide light with wavelengths from 2.5 to about 15 μm.

negative resist A process where ultraviolet radiation passing through the transparent areas of a mask causes the resist to be cured. The uncured areas are then removed using an appropriate solvent.

NET A group of terminals interconnected to have a common DC electrical potential in a package.

noise In a digital system, noise is any undesirable parasitic effect that causes signal wave form distortion, excessive delay, or false switching. Common types of noise are: reflection, coupled and switching.

noise saturation A phenomenon in which switching noise does not increase as the number of simultaneously switching circuits increases beyond an observed number. This effect is caused by a negative feedback between the large noise generated and the current drawn by each of the switching circuits. An increase in circuit delay occurs when noise saturation occurs.

nonlinear dielectric A capacitor material that has a nonlinear capacitance-to-voltage relationship.

optical interconnects Composed of basic optoelectronic devices, these components and modules are used as circuit building blocks. The light-emitting diode (LED) converts electrical energy to light where junction electroluminescence occurs as a result of the application of direct current at low voltage to a suitably doped crystal when forward biased. The light from this source then is detected by the reverse-biased pn-junction photodiode and /or phototransistor. Light, of the proper wavelength, creates a current flow, a photocurrent, in the external circuit proportional to the effective irradiance on the device.

optical lithography A process in which radiation at optical wavelengths (usually in the ultraviolet range) is passed through a mask, and the resulting patterns are projected onto a layer of resist coating the substrate material.

optical mask A sheet of material carrying patterns that are either transparent or opaque to the wavelengths used in an optical-lithographic process. Such a mask can carry hundreds of thousands of fine lines and geometric shapes.

organic resist A material that is used to coat a substrate and is then selectively cured to form an impervious layer.

Samsung, EX1028, p. 424

These materials are called organic because they are based on carbon compounds as are living creatures.

organic solvent A solvent for organic materials such as those used to form organic resists.

organic substrate Substrate materials such as FR4, in which woven glass fibers are bonded together with an epoxy. These materials are called organic because epoxies are based on carbon compounds as are living creatures.

outer lead bonding The process of joining the outer leads of package, typically TAB (Tape Automated Bonding) to the next level of assembly (usually card or board). The inner leads on the tape are joined to the chip by the process known as inner lead bonding.

overcoat A thin film of insulating material, either plastic or inorganic (e.g., glass of silicon nitride), applied over integrated circuit elements for the purpose of mechanical protection and prevention of contamination.

overflow wire See wiring overflow.

overglaze A glass coating that is grown, or deposited, over another element, normally for physical or electrical protection purposes.

overlay One material applied over another material.

package crossing An interconnection that connects a terminal on one package with that on another.

package delay The time delays associated with interconnections between components that complete logical make-up functions. Values depend on materials and distance.

packaging level A member of a nested interconnected packaging hierarchy (e.g., chip, chip carrier, card, board in order of low to high level).

pad An area of metallization on a substrate used for probing or to connect to a via, plated through-hole, or an external interconnect.

pad-grid array package (for VHSICS) A package embodying a technology where solder-contact pads are not just around the package periphery (as with chip carriers), but cover the entire bottom surface in checkerboard fashion.

PGA (pad grid array or pin grid array) PGA may refer to a pad grid array or a pin grid array. A pad grid array refers to a packaging technology in which a device's external connections are arranged as an array of conducting pads on the base of the package. A pin grid array refers to a packaging technology in which a device's external connections are arranged as an array of conducting leads, or pins, on the base of the package.

parasitic effects The effects caused by undesired resistance, capacitance, or inductance inherent in the material or topology of a track or component.

passivation The formation of an insulating layer directly over a circuit or circuit element.

passive components (elements) Elements or components such as resistors, capacitors, and inductors which do not change their basic character when an electrical signal is applied. In contrast, transistors, diodes and electron tubes are active components.

passive trimming A process in which a laser beam is used to trim components such as thick-film and thin-film resistors on an otherwise unpopulated and unpowered hybrid or multichip module substrate. Probes are placed at each end of a component to monitor its value while the laser evaporates some of the material forming the component.

paste Synonymous with "composition" and "ink" when referring to screenable thick-film materials, usually consisting of metal or ceramic powders, dispersed in organic vehicles.

peel strength (peel test) A measure of adhesion between the conductor and the substrate. The test is performed by pulling or peeling the conductor off the substrate, and observing the force required.

permeability The property of a solid plastic material that allows penetration by a liquid or gas.

phase diagram State of metal alloy or ceramic over a wide temperature range. The phase diagram is used to identify phases as a function of composition and temperature.

phosphosilicate glass (PSG) Phosphorus-doped silicon dioxide (also known as P-glass). It is often used as a dielectric material for insulation between conducting layers, for inhibiting the diffusion of sodium impurities, and for planarization, since it softens and flows at high temperatures to create a smooth topography for subsequent metallization.

photolithography The generation of a pattern through a sequence of rubylith, photo-reduction, step-and-repeat, computer-aided design, or finally, state-of-the-art electron-beam technique. This procedure will generate a product (mask or otherwise) to become the primary tool in transferring an image onto a microelectronic substrate.

physical design The process of allocating chip of package components and their interconnections to their appropriate spatial locations or sockets in an overall system arrangement.

Samsung, EX1028, p. 425

pick-and-place The manufacturing process whereby chips are selected and placed on the correct substrate site in preparation for joining (or interconnecting) the chip to the substrate.

pin Round, cross-sectional electrical terminal and/or mechanical support. Used in plug-in type packages, either straight or modified as nail-head, upset, pierced, or bent variety. A pin's primary functions are: internally, to support to a wirebond or other joint; and externally, to plug into a second-level package connector.

pin-grid array (PGA) A package, or interconnect scheme, featuring a multiplicity of plug-in type electrical terminals arranged in a prescribed matrix format or array.

pin-through-hole (PTH) A team referring to the class of packages or modules that are soldered into plated through holes within the second-level package (printed circuit board).

pitch The center-to-center spacing, between pads, rows of bumps, pins, posts, exit leads, etc. Sometimes also the distance, as measured point, to corresponding point, between two adjacent images in a device matrix, on a semiconductor wafer or its photomask.

placement The manual or automatic placing of chip circuits, chip, chip carriers, and cards in their actual locations on corresponding images at a given package level.

planar motor Typically, a brushless DC servo motor of a flat planar configuration, constructed using printed-circuit assembly methods.

plasma An electrically conductive gas, composed of ionized atoms or molecules, used for dry-etching in the fabrication of devices.

plasma etching The action of an electrically conductive gas (composed of ionized gas or molecules), to remove an unwanted portion of conductive or insulative pattern.

plastic A polymeric material, either organic (e.g., epoxy, polyimide) or inorganic (e.g., silicon), used for conformal coating, encapsulation, or overcoating.

plastic ball grid array (PGBA) Ball grid array technology on a plastic carrier. See Ball Grid Array.

plastic quad flat pack (PQFP) See Quad Flat Pack.

plating A condensation of the word Electroplating or Electroless Plating, that describes the coating of a metal on plastic, or other surfaces, with metal that is electrolytically or chemically deposited from a bath.

plated through-hole (PTH) (1) A hole in a double-sided or multilayer board that is used to accommodate a through-hole component lead and is plated with copper. (2) An alternative name for the lead through-hole technique for populating circuit boards, in which component leads are inserted into plated through-holes.

polyimides A class of resin compounds containing the NH group which are derived from ammonia and are "imidized" from polyamic acids, at temperatures high enough to initiate and complete the imide ring closure. Polyimides are useful as organic dielectric interlevel layers in VLSI technologies. They are mostly thermosetting ring-chain polymers, whose useful characteristics include perfect planarity as a spun film; high temperature tolerance, excellent weathering and mechanical-wear characteristics; and a low dielectric constant, a decided advantage in reducing propagation delays in multilayer hybrid circuits (faster switching).

polysilicon layer An internal layer in an integrated circuit used to create the gate electrodes of MOS transistors. In addition to forming gate electrodes, the polysilicon layer can also be used to interconnect components. There may be several polysilicon layers separated by dielectric (insulating) layers.

porcelain A mixture of borosilicate glass with minor quantities of zirconia and other ingredients..

porcelain enamel technology (PET) The technology of coating glass on metal. See Porcelain.

positive resist A process where radiation passing through the transparent areas of a mask causes previously cured resist to be degraded. The degraded areas are then removed using an appropriate solvent.

power distribution The network of conductors throughout the package that supplies the operating voltages and currents to the circuits.

power plane A conducting layer in or on the substrate, providing power to the components. There may be several power planes separated by insulating layers.

prepreg Non-conducting semi-cured layers of FR4 used to separate conducting layers in a multilayer circuit board.

pressure contact Mode of interconnection where the contact points are not fully bonded (as in soldered), but maintain electrical contact by means of a continuously applied force (such as a spring or rubber).

printed-circuit board (PCB) A composite of organic and inorganic material with external and internal wiring, allowing electronic components to be mechanically supported and electrically connected.

probability density function (PDF) The normalized frequency of occurrence with respect to a particular statistical distribution, e.g., normal, binomial, Weibull, etc.

purple plague One of several gold-aluminum compounds formed when bonding gold to aluminum, and ac-

Samsung, EX1028, p. 426

tivated by re-exposure to moisture. High temperature purple plague is purplish in color and is very brittle, potentially leading to time-based failure of the bonds. Its growth is highly enhanced by the presence of silicon to form ternary compounds.

pyrolyzed (burned) A material that has gained its final form by the action of heat is said to be pyrolyzed.

quad flat pack (QFP) Ceramic or plastic chip carrier with leads projecting down and away from all four sides of a square package.

quad in-line package (QUIP) A DIP-like plastic package with leads coming out on 1.27 mm centers.

radial-spread coating Also known as Glob Top. A coating process whereby a calculated amount of resin is dispensed on top of a surface to be encapsulated. The surface can be either a chip or a circuit board. The resin fans out, reacts (by heat or on contact with air), and forms a solid protective coating.

radiation The combined process of emission, transmission, and absorption of thermal energy between bodies separated by empty space.

RAM (random access memory) A data-storage device from which data can be read out and new data can be written in. Unless otherwise indicated, the term RAM is typically taken to refer to a semiconductor device in the form of an integrated circuit.

reaction injection molding (RIM) A molding process where two or more streams of reactants are metered into a small mixing chamber, where turbulent mixing breaks up the fluids into finely interspersed striations for faster reaction. The mixture is then delivered to a mold to complete the polymerization.

reactive etching A process wherein a printed pattern is formed by reaction (chemical/plasma ion) removal of the unwanted portion of conductive of insulative pattern.

receiver The off-chip circuit that accepts the signal voltages and currents from the package lines. Also called an Input Buffer Circuit.

reflection noise Spurious voltage and current wavelets on transmission lines that are caused by series or shunt networks that disrupt the continuous nature of the characteristic impedance of the line. These wavelets initially travel in a direction that is opposite to the wave that stimulates them. Typical discontinuities that cause reflections are stubs, connectors, vias, missing ground planes and improper terminations.

reflow soldering 1) A surface mount technology process in which the substrate and attached components are passed through a reflow oven to melt the solder paste. 2) A method of soldering involving application of solder prior to the actual joining. To solder, the parts are joined and heated, causing the solder to remelt or reflow.

relative humidity (rh) The ratio of partial pressure of water in any gas at a particular temperature to the saturated vapor pressure of water in any gas at the same temperature, usually expressed in percent.

reliability The probability of survival of a component, or assembly, for the expected period of use. Expressed mathematically, $R = 1 - P(\text{failure})$ during the expected life.

Rent's rule An empirical relation, first recorded by E. Rent of IBM, which states that the number of used input/output terminals on a logic package is proportional to a fractional power of the number of subpackages interconnected in the package.

resin A term used for an organic polymer that when mixed with a curing agent crosslinks to form a thermosetting plastic.

resist A protective coating that will keep another material from attaching or coating something, as in solder resist, plating resist, or photoresist.

resistance The property of a conductor that opposes the flow of current by displacing energy as heat. In packages, it causes voltage and current losses in signal and power distribution systems.

resistivity (ρ) A proportionality factor that is characteristic of different substances, equal to the resistance that a centimeter cube of a substance offers to the passage of electricity. Expressed $R=\rho L/A$ where R is the resistance of a uniform conductor, L its length, A its cross-sectional area, and ρ its resistivity.

rheology The science dealing with deformation and flow of matter.

rigid flex Hybrid constructions which combine standard rigid circuit boards with flexible printed circuits, thereby reducing the component count, weight, and susceptibility to vibration of the circuit, and greatly increasing its reliability.

rosin flux A flux having a rosin base that becomes interactive after being subjected to the soldering temperature.

routing program An automatic program emboding algorithm with prescribed wiring.

rules design systems The use of sufficiently accurate mathematical expressions that are derived from a limited number of judiciously chosen circuit stimulation results using curve fitting techniques to properly design a hardware system in conjunction with a design aids (DA) pro-

Samsung, EX1028, p. 427

gram. Since these formulas are computationally fast, they can be applied to the design of every data system.

self-generated noise tolerance A set of differential pulse amplitudes and pulse widths for spurious signals that are generated by the simultaneous switching of internal circuits that can be impressed across an internal circuit's power terminals without falsely setting a downstream latch circuit.

schematic Common name for a circuit diagram.

screening The process whereby the desired film-circuit patterns and configurations are transferred to the surface of the substrate during manufacture by forcing a material through the open areas of the screen using the wiping action of a soft squeegee.

sealing Joining the package case header (or chip carrier base or substrate) with its cover or lid into a sealed unit. For hybrids, sealing connotes an important finishing operation in fabricating a hybrid microcircuit, signaling the stage when the assembly, in the form of a populated package, becomes a bona fide hermetic (or nonhermetic) device.

self stretching soldering technology (SST) The acronym for C4 solder joining, where two different solder alloys, or sized bumps, are used so that surface tension forces of the nonfunctional bumps are used to stretch or increase the height of the functional solder joints. Taller connections can withstand higher thermal cycle or power cycle strains.

semiconductor A special class of material that can exhibit both conducting and insulating properties.

sensor A transducer that detects a physical quantity and converts it into a form suitable for processing. For example, a microphone is a sensor which detects sound and converts it into a corresponding voltage or current.

sheet resistance The electrical resistance of a thin sheet of a material with uniform thickness, as measured across opposite sides of a unit square pattern. Expressed in ohms per square.

signal distribution The network of package conductors that interconnects the drivers and receivers.

signal wiring A conductive path carrying an electric signal.

silicon bumping The process of depositing additional metallization on a die's pads to raise them fractionally above the level of the Barrier Layer.

silicon efficiency The ratio sum total of area, of all silicon chips to the total packaging area, primarily at board level.

simultaneously switching driver A driver circuit that changes state in unison with other drivers on the same chip or nearby chips, thereby creating switching noises.

single-chip carrier An electronic package that connects single-chip terminals to second-level package by having a different number of terminals than the chip itself.

single-chip module (SCM) Module or package supporting one chip, as opposed to mulitchip which supports several.

single-in-line package (SIP) DIP-like package with a single line of leads, as opposed to two for DIP.

single-layer metallized package (SLAM) Ceramic leadless package without cavity, sealed by ceramic or glass to a ceramic cap.

sintering Heating a metal or ceramic powder, thereby causing the particles to bond together to form a monolithic body.

skin effect A high frequency effect that causes the resistance of a conductor to increase. The phenomenon occurs because the magnetic fields within the conductor force the current to flow on the outer surface or skin, as the frequency of the signal increases.

slow wave propagation Energy that travels at less than expected velocity for a dielectric structure because of series resistance in the signal line or return path.

slurry A thick mixture of liquid and solids. The solids are in suspension in the liquid.

small outline (SOP) Also called SOIC. Small outline integrated circuit package. It is a rectangular DIP-like package except that it is smaller, and leads on 1.27mm, 1.0mm, or 0.85mm spacing. It is meant for surface mounting.

small-scale integration (SSI) Refers to the number of logic gates in a device. By one convention, small-scale integration represents a device containing 1 to 12 gates.

soft error In memory device technology, a memory state error induced by a process that produces no permanent alteration of the physical condition of the device.

soft glass Glasses, typically high-lead content glasses, having low softening points that could be used to seal ceramic or metal lids to packages below 450°C. Also called solder glasses because of their ability to wet most metal surfaces.

softening point Refers to the temperature at which the log viscosity of glass is 7.6 poises, as defined and measured to ASTM specification.

Samsung, EX1028, p. 428

solder A low melting point alloy used in numerous joining applications in microelectronics. The most common solders are lead-tin alloys.

solder dam A dielectric composition screened across a conductor to limit molten solder from spreading further onto solderable conductors.

solder glasses Glasses used in package sealing that have a low melting point and tend to wet metal and ceramic surfaces.

solder mask A layer applied to the surface of the substrate that prevents solder from sticking to any metallization except where holes are patterned into the mask.

solder mask over bare copper (SMOBC) A technique in which the solder mask is applied in advance of the tin-lead plating. This results in lighter circuit boards because the tin-lead alloy is only used to plate the pads.

solderability The ability of a conductor to be wetted by solder and to form a strong bond with the solder.

soldering The process of joining metals by fusion and solidification of an adherent alloy having a melting point below about 300°C.

solid logic technology (SLT) Ceramic package technology practiced by IBM in the 1960s by firing Ag-Pd conductors on to dry-pressed and fired alumina substrate.

space transformer A package transforming a spatially dense set of chip connections to a less dense set of connection points on package.

spice The "Simulation Program for Integrated Circuit Emphasis" is the industry standard for circuit simulation. It contains many of the features inherent in ASTAP. See ASTAP.

sputter cleaning Bombardment of a surface with energetic argon or other noble gas ions to clean the surface of oxide films and residues that could interfere with subsequent electrical or mechanical contact layers. The bombardment knocks off (or sputters) surface atoms to render the surface clean.

sputtering The removal of atoms from a source by energetic ion bombardment. The ions are supplied by a plasma. The sputtering process is used to deposit films for various thin-film applications.

static flex Flexible wiring circuit carrier, which once installed, remains fixed.

steiner tree See Stub.

stencil A planar patterned mask used to transfer images on a surface. Usually metallized patterns on an insulated surface.

storage control element (SCE) Controls the data transfer paths and the interface between the channels, processor storage, and central processor. Also called System Controlled Element.

storage hierarchy The collection of memory elements (cache, main storage, etc.) and their controls that make up the memory for the processor.

stress corrosion Refers to the degradation of mechanical properties of brittle materials by crack propagation due to the acceleration of applied stress in the presence of corroding atmospheres such as water.

stripline A transmission line that is embedded within a single dielectric medium and sandwiched between two reference planes.

stub A short wire that interconnects input at a circuit with the main signal line.

stud The conductive path that runs vertically from one level of conductors to another in a multilayer substrate.

subtracting patterning The processing sequence generally followed in producing thin-film networks or circuits. Films are area-deposited (by vacuum evaporation, CVD, or sputtering) and the desired conductive, resistive, etc., pattern is etched into each layer through mask, using appropriate selective etchant fluids.

subtractive process A process in which a substrate is first covered with conducting material, then any unwanted material is subsequently removed, or subtracted

superconductor Material offering no resistance to the flow of current. In addition to metals, ceramics have been recently discovered to have this property.

surface mount device (SMD) A component whose packaging is designed for use with surface mount technology.

surface mount technology (SMT) A method of assembling hybrid circuits and printed wiring boards, where component parts are mounted onto, rather than into, the printed-wiring board, as in the mounting components on substrates in hybrid technology.

surface tension An effect of the forces of attraction existing between the molecules of a liquid. It exists only on the boundary surface.

switching noise An induced voltage on the power distribution system at the circuit terminals, resulting from the rapidly changing current resulting from the simultaneous switching of many drivers.

tape automated bonding (TAB) The process where silicon chips are joined to patterned metal on polymer tape (e.g., copper on polymide), using thermocompression

Samsung, EX1028, p. 429

bonding, and subsequently attached to a substrate or board by outer lead bonding. Intermediate processing may be carried out in strip form through operations such as testing, encapsulation, burn-in, and excising the individual packages from the tape.

tape ball grid array (TBGA) Ball grid array technology on TAB. See TAB.

temperature cycling An environmental test where the film circuit is subjected to several temperature changes from a low temperature to a high temperature over a period of time.

tensile strength The pulling stress that has to be applied to a material to break it, usually measured in MPa.

terminal A metallic connector or pad to a circuit within a chip or package that permits electrical interconnection to external circuits.

thermal coefficient of expansion (TCE) The ratio of the change in dimensions to the change in temperature-per-unit starting length, usually expressed in cm/cm/°C. The acronyms TCE and CTE are synonymous.

thermal conduction module (TCM) An IBM multichip (100 chips or more) module that is cooled by thermal conduction or pistons in contact with chips.

thermal conductivity The rate with which a material is capable of transferring a given amount of heat through itself.

thermal cycling A method to impose a cyclic stress on an assembly of microelectronic components by alternately heating and cooling in an oven. It is used to accelerate reliability testing of assemblies.

thermal fatigue Failure of a structural element from repeated temperature excursions, wherein the load develops from thermal expansion mismatch of dissimilar materials.

thermal gradient The plot of temperature variances across the bulk thickness of a material being heated.

thermal mismatch Difference in thermal coefficients of expansion of materials which are bonded together.

thermal network Representation of a thermal space by a collection of conveniently divided smaller parts, each representing the thermal property of its own part, and connected to others in a prescribed manner so as not to violate the thermal property of the total system.

thermal resistance (°C/W) The opposition offered by a medium to passage through it of thermal energy.

thermocompression bonding (T/C) A process involving the use of pressure and temperature to join two materials by interdiffusion across the boundary.

thermogravimetric analysis (TGA) A technique that measures material weight change as a function of increasing temperature.

thermoplastic A substance that becomes plastic (malleable) on being heated; a plastic material that can be repeatedly melted or softened by heat without change of its properties.

thermosetting The property of some organic materials to irreversibly polymerize and set or harden when heated to some appropriate temperature.

thermosonic bonding (T/S) A bonding process which use a combination of thermocompression (TC) bonding and ultrasonic bonding. It is done on what amounts to a gold-wire TC bonder with ultrasonic power applied to the capillary.

Thevénin equivalent Electrical model describing voltage-current behavior of an electrical network between any two of its nodes. In its simplest form, it can be a constant voltage source with a series impedance (resistance in case of direct current) or a constant current source shunted by an impedance. These impedances are often called source impedances.

thick-film A film deposited by screen printing processes and fired at high temperature to fuse into its final form. The basic processes of thick-film technology are screen printing and firing.

thick-film process A process used in the manufacture of hybrids and, to a lesser extent, multichip modules in which signal and dielectric (insulating) layers are screen-printed onto the substrate.

thin-film The film refers to a coating layer of thickness in the range of from a few (2–3) atomic layers to a few (1–5) microns (micrometers). The important feature distinguishing thin films from thick films, though, is not so much the difference in thickness as the method of deposition, which takes place by a variety of techniques such as chemical vapor deposition, evaporation, or sputtering.

thin-film packaging An electronic package in which the conductors and/or insulators are fabricated using deposition and patterning techniques similar to those used for integrated circuit chips.

thin-film process A process used in the manufacture of hybrids and multichip modules, in which signal layers and dielectric (insulating) layers are created using optolithographic techniques.

three-layer tape An interconnection medium used in tape automated bonding (TAB), where the tape is comprised of three layers of metallization (usually copper), with polymer and adhesive in between.

Samsung, EX1028, p. 430

through hole A hole connecting the two surfaces of a printed-circuit structure.

time-domain reflection A time varying voltage and current disturbance created at a discontinuity on a transmission line that travels in a direction opposite to its stimuli, thereby causing spurious line noises and signal distortions.

tin-lead plating An electroless plating process in which exposed areas of copper on a circuit board are coated with a layer of tin-lead alloy. The alloy is used to prevent the copper from oxidizing and provides protection against contamination.

tinned Literally, coated with tin, but commonly used to indicate coating with solder.

top side metallurgy (TSM) An acronym referring to the metallization on the top side of a substrate to which a chip is joined (such as C4 solder connection).

topography The surface condition of a film, bumps, craters, etc.

transfer molding An automated type of compression molding in which a preform of plastic (usually an epoxy-based resin) is poured from a pot into a hot mold cavity.

transfer utility grain point The point on a logic circuit's V_{out} vs. V_{in} transfer curve where the output voltage equals the input voltage. It determines the input signal swing at which noise will propagate and amplify through cascaded logic circuits.

transient mismatch Thermal mismatch between elements of a structure which, because of thermal lag, varies with time until reaching a steady-state value.

transistor-transistor logic (TTL) A Nand logic function is implemented by the switching of voltage changes on distinct emitter inputs of bipolar transistors sharing common base and collector voltages.

transmission line A conductor that is inductively and capacitively coupled to a nearby return path, forming a uniform distributed network with specific properties.

transmitted noise tolerance A set of pulse amplitudes and pulse widths for spurious signals at the input to a receiver circuit that will not falsely set a downstream latch.

tri-plate line Same as a stripline. See stripline.

two-layer tape A primary form of tape fabrication for tape automated bonding (TAB), starting with the metallic sputtering and subsequent pattern-plating on Kapton carrier tape. No adhesive is used in bonding copper to Kapton.

ultra large scale integration (ULSI) So far, an extreme in the circuit integration, used to indicate presence of one hundred million transistors (or more) on a single semiconductor chip.

ultrasonic bonding A process involving the use of ultrasonic energy and pressure to join tow materials.

vacuum deposition Deposition of a metal film onto a substrate in vacuum by metal evaporation techniques.

vapor phase reflow The technique for solder reflow to form package interconnections. The solder joint is heated by the heat of condensation of an inert vapor. The most common material of choice is perfluorocarbon.

vapor phase soldering A surface mount process in which a substrate carrying components attached by solder paste is lowered into the vapor cloud of a tank containing boiling hydrocarbons. This melts the solder paste, thereby forming good electrical connections.

very high speed integrated chip (VHSIC) Very high speed integrated circuit, originally referring to 1.0 um ground rules.

very large scale integration (VLSI) Level of integration with more than approximately ten thousand transistors on a single semiconductor chip. Upper boundary not well defined.

via An opening in the dielectric layer(s) through which a riser passes, or else whose walls are made conductive.

via, fixed A via built into a package on a predetermined grid, and in general, interconnecting both adjacent and nonadjacent planes.

via, programmable A via interconnecting adjacent wires on two adjacent wiring planes. Location doesn't correspond to the same grid locations as fixed vias.

via, segmented A fixed via interconnection. A predetermined subset of all wiring planes.

via, through A fixed via passing through all wiring planes.

viscosity The intrinsic property of a fluid that resists internal flow by offering counteracting forces.

voltage slew rate The rate of change in voltage with respect to time (dv/dt).

wafer Commonly, a slice of a semiconductor crystalline ingot used for substrate material when modified by the addition, as applicable, of impurity diffusion (doping), ion implantation, epitaxy, etc., and whose active surface has been processed into arrays of discrete devices or ICs by metallization and passivation.

wafer probing The process of testing individual integrated circuits while they still form part of a wafer. An automated tester places probes on the device's pads, applies power to the power pads, injects a series of signals

Samsung, EX1028, p. 431

into the input pads, and monitors the corresponding signals returned from the output pads.

wave soldering The technique for solder application and reflow in which a jet of liquid solder is directed at the two metallic points to be interconnected. The technique usually involves processing steps to apply flux and remove excess solder.

wearout The time following the stable failure-rate period during which the expected, or observed, failure rate of an item increases and exceeds a specific value.

welding Joining two metals by applying heat to melt and fuse them with or without a filler metal.

wetting The spreading of molten solder or glass on a metallic or nonmetallic surface, with proper application of heat and flux.

wire length, average The average length measured in logic unit pitches, of all connections in a given package level.

wireability The capability of a package to permit the interconnection of subpackages mounted on it and terminals attached to it, measured as the probability of wiring success. It is near one when sufficient wiring capacity, via availability and terminal access, are present.

wirebond A completed wire connection whose constituents provide electrical continuity between the semiconductor die (pad) and a terminal. These constituents are the fine wire; metal bonding surfaces like die pad and package land; and metallurgical interfaces between wire, and metals on both the chip and substrate.

wirebonding The method used to attach very fine wire to semiconductor components in order to interconnect these components with each other or with package leads.

wiring The manual or automatic prescription of routes (portions of tracks) for wires interconnecting package components or logic cells on chips. Also called Routing.

wiring assignment The manual or automatic prescription of particular pads, pins, connectors, or terminals to which corresponding wires are to be attached.

wiring capacity The total available length of wiring tracks in a package (before any wires are prescribed on the wiring image).

wiring channel A linear region on a package wiring plane containing space for at least one wiring track.

wiring demand The product of wiring connection count and average connection length, either locally in a limited region or total over an entire package.

wiring density Total wire length contained within a unit square, measured in inches per square inch or centimeters per square centimeter.

wiring overflow A wiring connection called for by logic design but not inserted in a proposed package wiring image during prior automatic wiring use of package.

wiring rules A set of electrical constraints that are used in conjunction with a design aids program to control the topological parameters of an interconnection network to assure proper functionality. Wiring rules are usually derived from electrical simulation results.

wiring track A linear extent of space in a wiring channel used to contain one (or more if collinear) conducting wires used to interconnect package components.

yellow wire Discrete wires that are yellow in color, interconnecting terminals on a package. Originally used in reference to all back panel wiring on early electronic assemblies.

zero-insertion-force connection (ZIF) A form of connector that allows the connector pins to be brought together under very low force, then wiped and pressed together during cam activation.

Samsung, EX1028, p. 432