7.12 **(a)** What is the sheet resistance of a 0.5-μm-thick copper line with a resistivity of 1.7 μohm-cm?

(b) What is the resistance of a line that is 50 μm long and 0.5 μm wide?

(c) What is the capacitance of this line to the substrate if a 1-μm thick "low-K" dielectric is utilized with $\varepsilon = 2\varepsilon_o$?

(d) What is the *RC* product associated with the 50-μm line?

CHAPTER 8

Packaging and Yield

The low cost normally associated with integrated circuits results from mass production in which many wafers, each containing a large number of IC dice, are all processed together. There may be tens to thousands of dice per wafer and 25 to 200 wafers per lot. After wafer processing is completed, however, the dice must be tested, separated, and assembled in packages that are easy to handle and to mount in electronic systems. The testing and assembly operations substantially increase the cost of the final product.

In this chapter, we first present an overview of testing and die separation. Then we discuss IC assembly, including die attachment, wire bonding, and a survey of the various types of packages used with integrated circuits.

The ultimate cost of the integrated circuit is related to the total yield of assembled and tested devices. In the early stages in the development of a new process or circuit, we are lucky if a few functional dice are found per wafer. Late in the life of a process with a mature circuit design, yields of 60 to 80% are not uncommon. A discussion of the dependence of yield on defect density and die size concludes this chapter.

8.1 TESTING

Following metallization and passivation-layer processing, each die on the wafer is tested for functionality. Special parametric test dice are placed at a number of sites on the wafer. At this stage, dc tests are used to verify that basic process parameters fall within acceptable limits. To perform the tests, a probe station lowers a ring of very fine, needle-sharp probes into contact with the pads on the test die. Test equipment is connected to the circuit through the probes and controlled by a computer system. If the wafer-screening operation shows that basic process and device parameters are within specification, functional testing of each die begins.

Under computer control, the probe station automatically steps across the wafer, performing functional testing at each die site. Defective dice are marked with a drop of ink. Later, when the dice are separated from the wafer, any die with an ink spot is discarded. It has become impossible to exhaustively test complex VLSI devices such as

microprocessors. Instead, a great deal of computer time is used to find a minimum sequence of tests that can be used to indicate die functionality. At the wafer-probe stage, functional testing is primarily static in nature. High-speed dynamic testing is difficult to do through the probes, so parametric speed tests are usually performed after die packaging is complete.

The ratio of functional dice to total dice on the wafer gives the *yield* for each wafer. Yield is directly related to the ultimate cost of the completed integrated circuit and will be discussed more fully in Section 8.7.

8.2 WAFER THINNING AND DIE SEPARATION

As mentioned in Chapter 1, wafers range from approximately 350–1250 microns in thickness. Large-diameter wafers must be thicker in order to maintain structural integrity and planarity during the wide range of processing steps encountered during IC fabrication. However, many applications require much thinner dice; a thickness of approximately 275 μm is commonly used by many manufacturers. Flash memory cards, credit cards, and medical electronics are just three examples of applications that use dice that have been thinned to only 125 μm. The thinning is done using back lapping and polishing processes similar to those for CMP described in Chapter 6.

Following initial functional testing, individual IC dice must be separated from the wafer. In one method used for many years, the wafer is mounted on a holder and automatically scribed in both the x- and y-directions using a diamond-tipped scribe. Scribing borders of 75 to 250 μm are formed around the periphery of the dice during fabrication. These borders are left free of oxide and metal and are aligned with crystal planes if possible. In <100> wafers, natural cleavage planes exist perpendicular to the surface of the wafer in directions both parallel and perpendicular to the wafer flat. For <111> wafers, a vertical cleavage plane runs parallel, but not perpendicular, to the wafer flat. This can lead to some separation and handling problems with <111> material.

Following scribing, the wafer is removed from the holder and placed upside down on a soft support. A roller applies pressure to the wafer, causing it to fracture along the scribe lines. Care is taken to ensure that the wafer cracks along the scribe lines to minimize die damage during separation, but there will always be some damage and loss of yield during the scribing and breaking steps.

Today, diamond saws are used for die separation. A wafer is placed in a holder on a sticky sheet of Mylar as shown in Fig. 8.1. The saw can be used either to scribe the wafer or to cut completely through the wafer. Following separation, the dice remain attached to the Mylar film.

8.3 DIE ATTACHMENT

Visual inspection is used to sort out dice that may have been damaged during die separation, and the inked dice are also discarded. The next step in the assembly process is to mount the good dice in packages.

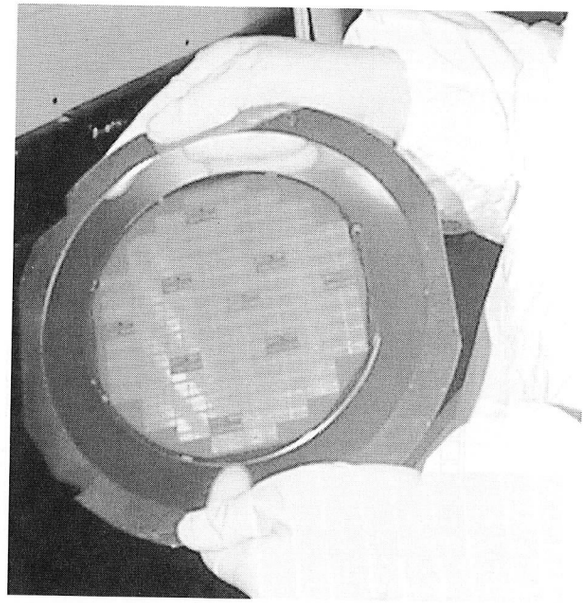


FIGURE 8.1
150-mm wafer mounted and ready for the dicing saw.

8.3.1 Epoxy Die Attachment

An epoxy cement may be used to attach the die to a package or "header." However, epoxy is a poor thermal conductor and an electrical insulator. Alumina can be mixed with the epoxy to increase its thermal conductivity, and gold- or silver-filled epoxies are used to reduce the thermal resistance of the epoxy bonding material and to provide a low-resistance electrical connection between the die and the package.

8.3.2 Eutectic Die Attachment

The gold–silicon eutectic point occurs at a temperature of 370 °C for a mixture of approximately 3.6% Si and 96.4% Au. Gold can be deposited on the back of the wafer prior to die separation or can be in the form of a thin alloy "preform" placed between the die and package. To form a eutectic bond, the die and package are heated to 390 to 420 °C, and pressure is applied to the die in conjunction with an ultrasonic scrubbing motion. Eutectic bonding is possible with a number of other metal-alloy systems, including gold–tin and gold–germanium. A solder attachment technique is also used with semiconductor power devices.

8.4 WIRE BONDING

Wire bonding is the most widely used method for making electrical connections between the die and the package. The bonding areas on the die are large, square pads 100 to 125 μm on a side, located around the periphery of the die. Figure 8.2 shows a 2.5 × 2.5 mm die with 100 μm wire bond pads on a 125 μm pitch. Many manufacturers do not put pads in the corners, because of reliability concerns. In particular, packaging-induced die stress and die cracking tend to be highest in the corners.

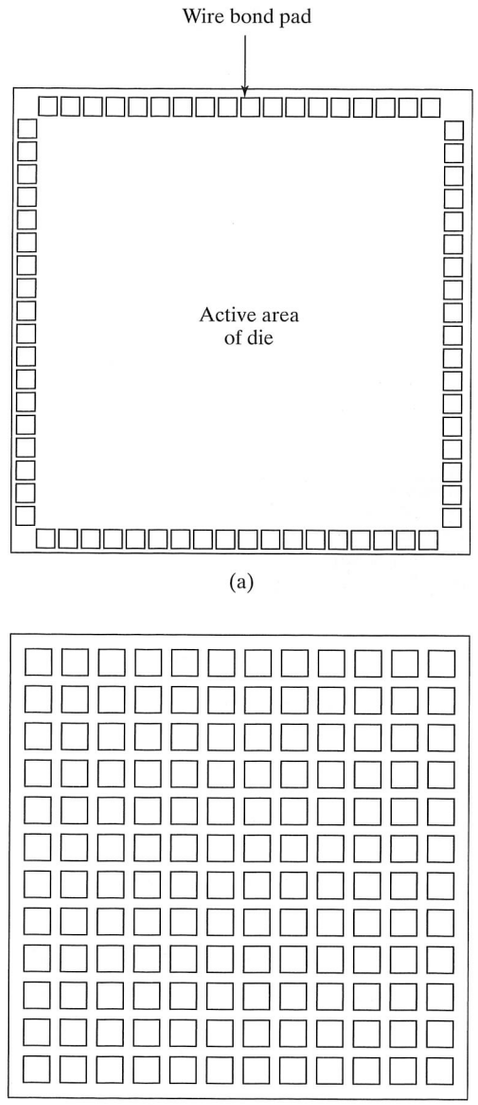
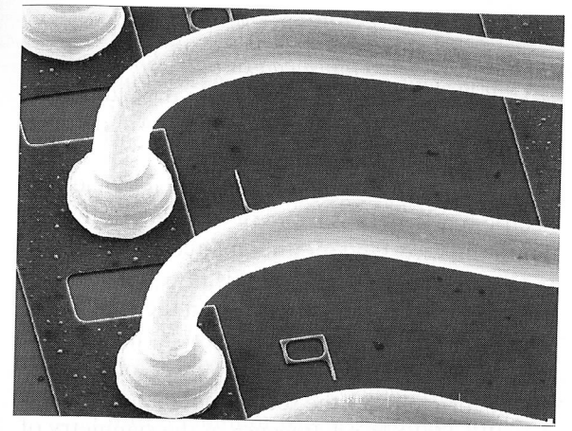


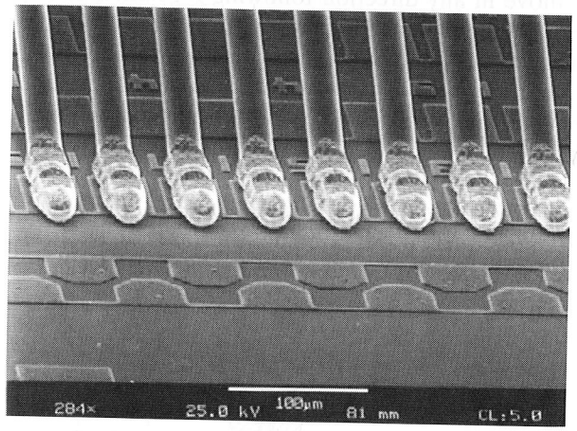
FIGURE 8.2

Die layout for (a) peripheral pads and (b) area array of pads.

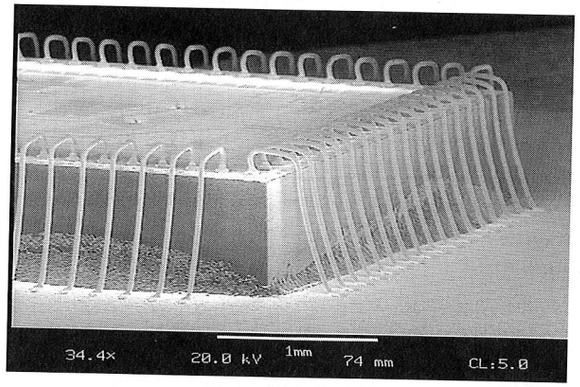
Thermocompression bonding was originally used with gold wire, and ultrasonic bonding is used with aluminum wire. A combination of the two is termed thermosonic bonding. Present high-density wire bonding can achieve a 50–70 μm pitch (see Fig. 8.3), and two or more staggered rows of wire bond pads are sometimes used to increase the number of possible pads/die. The ITRS goal is to reach a wire bond pitch of 40 μm by the year 2010. Fine wires interconnect the aluminum bonding pads on the IC die to the leads of the package.



(a)



(b)



(c)

FIGURE 8.3

(a) An SEM micrograph of gold ball bonding (b) SEM of high density gold ball bonding (c) SEM micrograph of wire bonded die. Courtesy of Kulicke and Soffa Industries, Inc. (K&S).

8.4.1 Thermocompression Bonding

The thermocompression bonding technique, also called *nail-head* or *ball bonding*, uses a combination of pressure and temperature to weld a fine gold wire to the aluminum bonding pads on the die and the gold-plated leads of the package. Figure 8.4 shows the steps used in forming either a thermocompression or thermosonic bond.

A fine gold wire, 15 to 75 μm in diameter, is fed from a spool through a heated capillary. A small hydrogen torch or electric spark melts the end of the wire, forming a gold ball two to three times the diameter of the wire. Under either manual or computer control, the ball is positioned over the bonding pad, the capillary is lowered, and the ball deforms into a "nail head" as a result of the pressure and heat from the capillary.

Next, the capillary is raised, and wire is fed from the spool as the tool is moved into position over the package. The second bond is a wedge bond produced by deforming the wire with the edge of the capillary. After the formation of the second bond, the capillary is raised and the wire is broken near the edge of the bond. Because of the symmetry of the bonding head, the bonder may move in any direction following formation of the nail-head bond. An SEM micrograph of a gold-ball bond is shown in Fig. 8.3(a).

A problem encountered in production of gold–aluminum bonds is formation of the "purple plague." Gold and aluminum react to form intermetallic compounds. One such compound, $AuAl_2$, is purple in color, and its appearance has been associated with faulty bonds. However, this compound is highly conductive. The actual culprit is a low-conductivity, tan-colored compound, Au_2Al, which is also present.

During thermocompression bonding, the substrate is maintained at a temperature between 150 and 200 °C. The temperature at the bonding interface ranges from 280 to 350 °C, and significant formation of the Au–Al compounds can occur at these

temperatures. Limiting the die temperature during the bonding process helps to prevent the formation of the intermetallic compounds, and high-temperature processing and storage following bonding should be avoided.

In addition to the foregoing problem, many epoxy materials cannot withstand the temperatures encountered in thermocompression bonding, and thermocompression bonding has been replaced by the ultrasonic and thermosonic bonding techniques discussed in the next two sections.

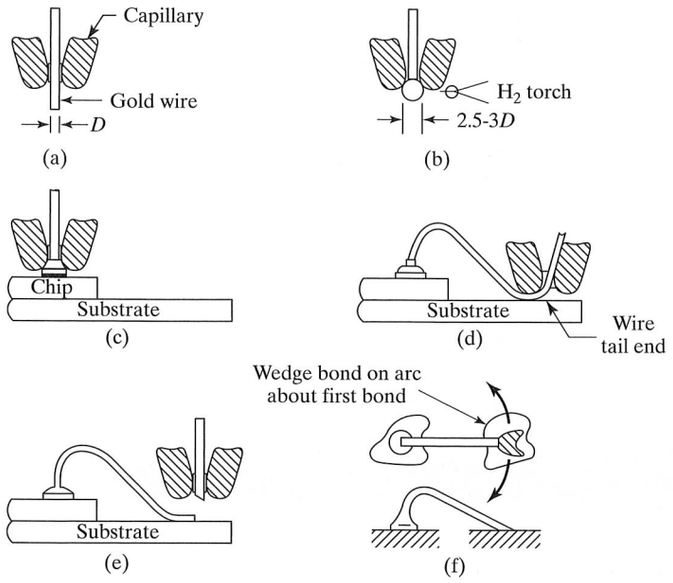
8.4.2 Ultrasonic Bonding

Oxidation of aluminum wire at high temperatures makes it difficult to form a good ball at the end of the wire. An alternative process called *ultrasonic bonding*, which forms the bond through a combination of pressure and rapid mechanical vibration, is used with aluminum wire. Aluminum wire is fed from a spool through a hole in the ultrasonic bonding tool, as shown in Fig. 8.5. To form a bond, the bonding tool is lowered over the bonding position, and ultrasonic vibration at 20 to 60 kHz causes the metal to deform and flow easily under pressure at room temperature. Vibration also breaks through the oxide film that is always present on aluminum and results in formation of a clean, strong weld.

As the tool is raised after forming the second bond, a clamp is engaged and pulls and breaks the wire at a weak point just beyond the bond. To maintain proper wire alignment in the bonding tool, the ultrasonic bonder can move only in a front-to-back motion between the first and second bonds, and the package must be rotated 90° to permit complete bonding of rectangular dice.

FIGURE 8.4

Thermosonic ball-wedge bonding of a gold wire. (a) Gold wire in a capillary; (b) ball formation accomplished by passing a hydrogen torch over the end of a gold wire or by capacitance discharge; (c) bonding accomplished by simultaneously applying a vertical load on the ball while ultrasonically exciting the wire (the chip and substrate are heated to about 150 °C); (d) a wire loop and a wedge bond ready to be formed; (e) the wire is broken at the wedge bond; (f) the geometry of the ball-wedge bond that allows high-speed bonding. Because the wedge can be on an arc from the ball, the bond head or package table does not have to rotate to form the wedge bond. Reprinted with permission from *Semiconductor International* magazine, May 1982 [1]. Copyright 1982 by Cahners Publishing Co., Des Plaines, IL.



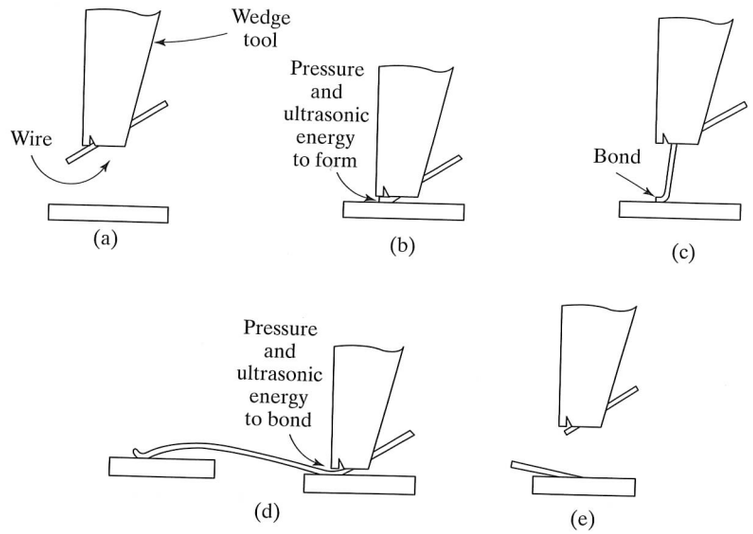


FIGURE 8.5

(a) In ultrasonic bonding, the tool guides wire to the package terminal; (b) pressure and ultrasonic energy form the bond; (c) and (d) the tool feeds out wire and repositions itself above the IC chip. The tool lowers and ultrasonically forms the second bond; (e) tool lifts, breaking the wire at the bond. Reprinted with permission from *Circuits Manufacturing*, January 1980. Copyright Morgan-Grampian 1980.

8.4.3 Thermosonic Bonding

Thermosonic bonding combines the best properties of ultrasonic and thermocompression bonding. The bonding procedure is the same as in thermocompression bonding, except that the substrate is maintained at a temperature of approximately 150 °C. Ultrasonic vibration causes the metal to flow under pressure and form a strong weld. The symmetrical bonding tool permits movement in any direction following the nail-head bond. Thermosonic bonding can be easily automated, and computer-controlled thermosonic bonders can produce 5 to 10 bonds per second.

8.5 PACKAGES

IC dice can be mounted in a wide array of packages. In this section, we discuss the round TO-style packages, dual-in-line packages (DIP), the pin-grid array (PGA), the leadless chip carrier (LCC), and packages used for surface mounting. Later in this chapter, we will look at flip-chip mounting, ball-grid arrays, and tape-automated bonding. The long-term trend is to attempt to achieve a package footprint that is similar in size to the semiconductor die itself. These packages are referred to as Chip Scale Packages (CSPs).

8.5.1 Circular TO-Style Packages

Figure 8.6(a) shows a round TO-type package that was one of the earliest IC packages. Different configurations of this package are available with 4 to 48 pins. The silicon die is attached to the center of the gold-plated header. Wire bonds connect the pads on the die to Kovar lead posts that protrude through the header and glass seal. Kovar is an iron–nickel alloy designed to have the same coefficient of thermal expansion as the glass seal. A metal cap is welded in place after die attachment and wire bonding.

8.5.2 DIPs

The DIP shown in Fig. 8.6(b) is extremely popular, because of its low cost and ease of use. Plastic and epoxy DIPs are the least expensive packages and are available with as few as 4 leads to more than 80 leads. In the postmolded DIP, a silicon die is first mounted on and wire-bonded to a metallic lead frame. Epoxy is then molded around both the die and the frame. As a result, the silicon die becomes an integral part of the package. Plastic DIPs have evolved into other forms of in-line packages, including single-in-line packages (SIPs) and zig-zag-in-line packages (ZIPs) shown in Fig. 8.6(c), as well as many surface-mount packages to be described in Section 8.5.5.

In a ceramic DIP, the die is mounted in a cavity on a gold-plated ceramic substrate and wire bonded to gold-plated Kovar leads. A ceramic or metal lid is used to seal the top of the cavity. Ceramic packages are considerably more expensive than plastic and are designed for use over a wider temperature range. In addition, ceramic packages may be hermetically sealed. A premolded plastic package similar to the ceramic package is also available.

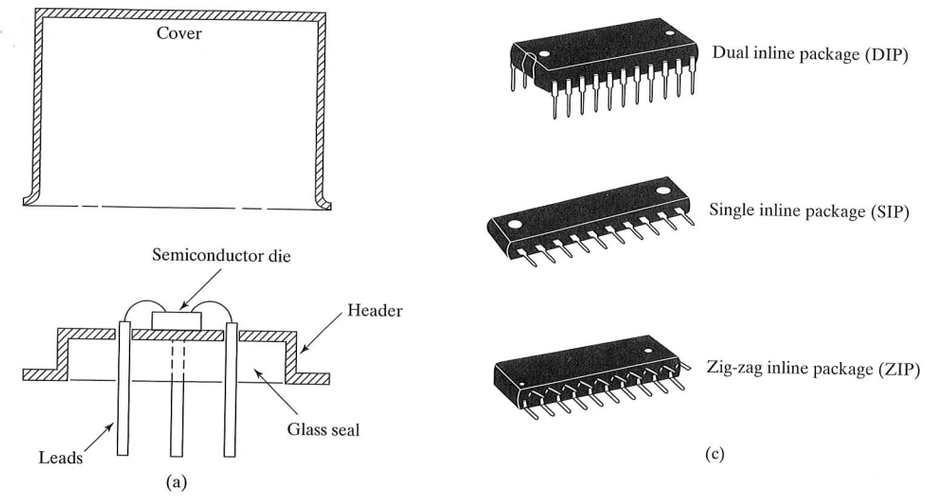


FIGURE 8.6

(a) TO-style package; (b) ball- and wedge-bonded silicon die in a plastic DIP. The die support paddle may be connected to one of the external leads. For most commercial products, only the die paddle and the wedge-bond pads are selectively plated. The external leads are solder-plated or dipped after package molding. Copyright 1981, IEEE. Reprinted with permission from Ref. [3]. (c) DIP, SIP, and ZIP packages.

8.5.3 Pin-Grid Arrays (PGAs)

The DIP package is satisfactory for packaging IC dice with up to approximately 80 pins. The *pin-grid array* of Fig. 8.7 provides a package with a much higher pin density than that of the DIP package. The pins are placed in a regular x–y array, and the package can have hundreds of pins. Wire bonding is still used to connect the die to gold interconnection lines, which fan out to the array of pins. Other versions of this package use the flip-chip process (which will be discussed in Section 8.6).

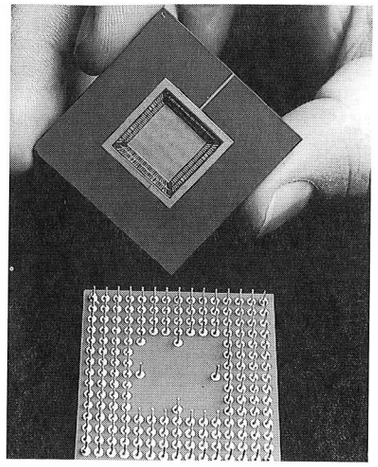
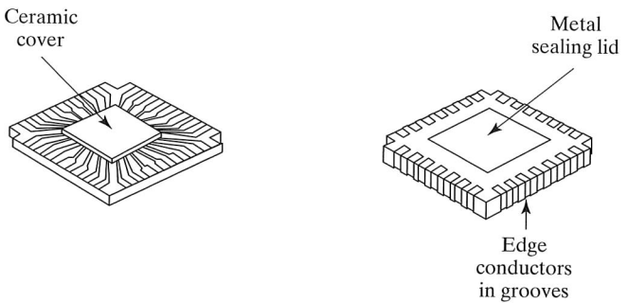


FIGURE 8.7
An example of a PGA in which the chip faces upward in the cavity.

8.5.4 Leadless Chip Carriers (LCCs)

Figure 8.8 shows two types of leadless chip carriers. In each case, the die is mounted in a cavity in the middle of the package. Connections are made between the package and die using wire bonding, and the cavity is sealed with a cap of metal, ceramic, or epoxy. The package in Fig. 8.8(a) has contact pads only on the top surface of the chip carrier. The chip carrier is pressed tightly against contact fingers in a socket mounted on a printed-circuit board. Another type of LCC is shown in Fig. 8.8(b). Conductors are formed in grooves in the edges of the chip carrier and are again pressed tightly against contact pins in a socket on the next level of packaging.

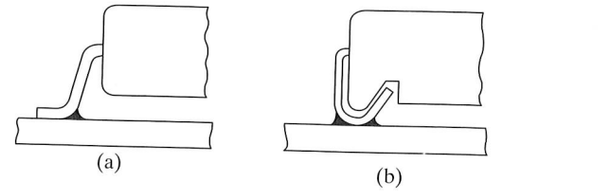


(a) (b)

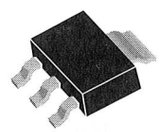
FIGURE 8.8
(a) Ceramic leadless chip carriers with top connections; (b) LCC with edge connections in grooves on the sides of the package.

8.5.5 Packages for Surface Mounting

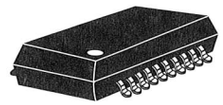
The TO-, DIP-, and PGA-style packages are made for mounting in holes fabricated in printed-circuit boards. Surface-mount packages do not require holes. The gull-wing package shown in Fig. 8.9(a) has short-lead stubs bent away from the package, whereas the leads of the J-style package of Fig. 8.9(b) are bent back underneath the package. Both styles permit soldering of the package directly to the surface of a printed-circuit board or hybrid package. Several versions are shown in Fig. 8.9(c)–(h). The leadless chip carriers described in the previous section are also available with leads added for surface mounting.



(a) (b)

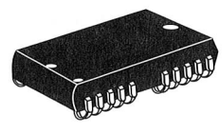


(c)
Small outline transistor (SOT)

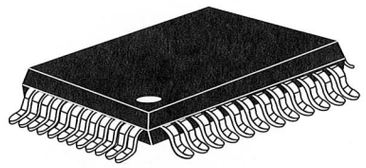


(d)
Small outline integrated circuit (SOIC)

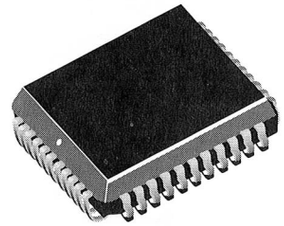
(e)
Thin small outline package (TSOP)



(f)
Small outline j-leaded (SOJ)



(g)
Quad flat pack (QFP)



(h)
Plastic-leaded chip carrier (PLCC)/
Quad J-leaded pack (QJP)

FIGURE 8.9
(a) Gull-wing and (b) J-lead surface-mount (c) Small outline transistor (d) Small outline IC (e) Thin small outline (f) Small outline J-leaded (g) Quad flat pack (h) Plastic-leaded chip carrier or Quad J-leaded packages.

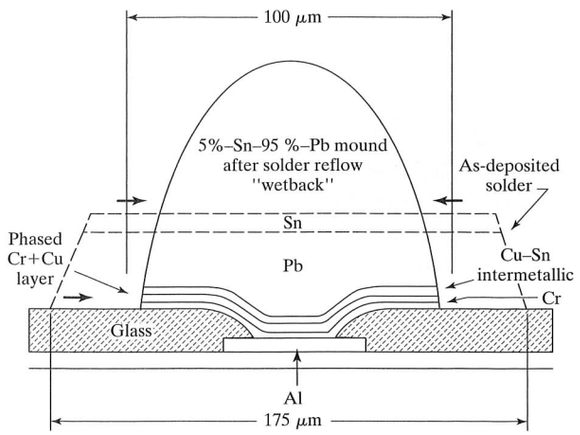
8.6 FLIP-CHIP AND TAPE-AUTOMATED-BONDING PROCESSES

As can be envisioned from the preceding discussion, the die-mounting and wire-bonding processes involve a large number of manual operations and are therefore quite expensive. In fact, the cost of assembly and test may be many times the cost of a small die. The one-at-a-time nature of the wire-bonding process also leads to reduced reliability, and failure of wire bonds is one of the most common reliability problems in integrated circuits. The flip-chip and tape-automated-bonding processes were developed to permit batch fabrication of die-to-package interconnections.

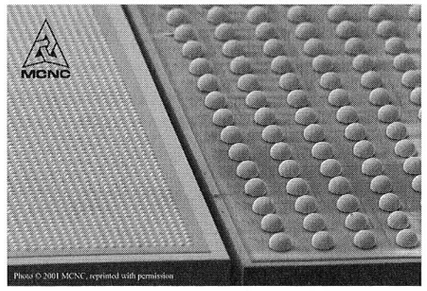
8.6.1 Flip-Chip Technology

The *flip-chip* mounting process was developed at IBM during the 1960s [4]. It took almost three decades for the semiconductor industry to adopt flip-chip technology. For today and the foreseeable future, however, flip-chip technology is an integral part of the ITRS plan for high-density chip interconnection. The first step is to form a solder ball (see Fig. 8.10) on top of each bonding pad. A sandwich of Cr, Cu, and Au is sequentially evaporated through a mask to form a cap over each of the aluminum bonding pads. Chrome and copper provide a barrier and a good contact to the aluminum pad. Gold adheres well to chrome and acts as an oxidation barrier prior to solder deposition. Lead-tin solder is evaporated through a mask onto the Au–Cu–Cr cap, occupying an area slightly larger than the cap. The die is heated, causing the solder to recede from the oxide surface and form a solder ball on top of the Au–Cu–Cr bonding-pad cap.

After being tested and separated, the dice are placed face down on a substrate. Temperature is increased, causing the solder to reflow, and the die is bonded directly to the interconnections on the substrate. Solder balls provide functions of both electrical interconnection and die attachment. Hundreds of bonds can be formed simultaneously using this technique, and bonding pads may be placed anywhere on the surface of the die, rather than just around the edge. In addition, the bond between the die and the substrate is very short. The main disadvantages of this technique are the additional



(a)



(b)

FIGURE 8.10

(a) Cross section through a solder ball before and after reflowing. Copyright 1969 by International Business Machines Corporation. Reprinted with permission from Ref. [4]. (b) Flip-chip Pb/Sn solder bumps in standard 250 μm pitch (right) and 50 μm. Courtesy of MCNC Optical and Electronic Packaging Group.

processing complexity, the higher thermal resistance between die and substrate, and the inability to visually inspect the completed solder joints. To enhance reliability, an under-fill material is used in the gap between the chip and substrate. The under-fill material is usually dispensed in liquid form and then cured. To be effective, the under-fill must achieve good adhesion at the interfaces with both the die and the substrate.

Two forms of solder-ball footprints are in common use at the die level. The first simply replaces wire bonds with solder balls around the periphery of the die. However, the real power of the flip-chip approach is achieved through the use of large arrays of solder balls placed over the full die area. For example, the die in Fig. 8.2(a) has only 72 pads, but an area array would achieve 144 interconnections using 125-μm pads on a 200-μm pitch as shown in Fig. 8.2(b).

The original IBM technique was a difficult and expensive process to implement, and these factors impeded its widespread adoption. More recently, a screen-printing approach has been used to deposit solder paste to form solder balls for low-volume applications. Gold stud bumping, shown in the photograph in Fig. 8.11, is yet another approach in which a gold wire is bonded to the chip pad, but the wire is removed after the first bond, leaving the gold bump on the pad. These chips may then be flipped over and mounted to form a flip-chip structure.

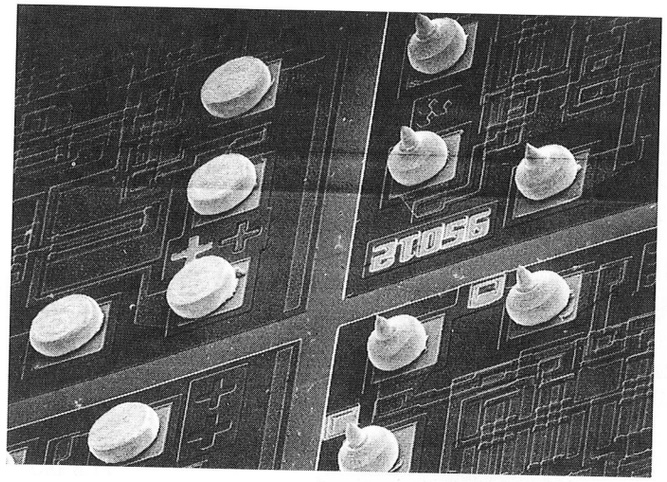
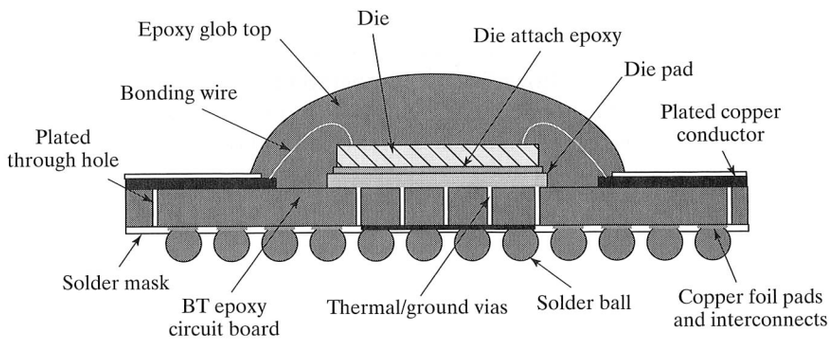


FIGURE 8.11

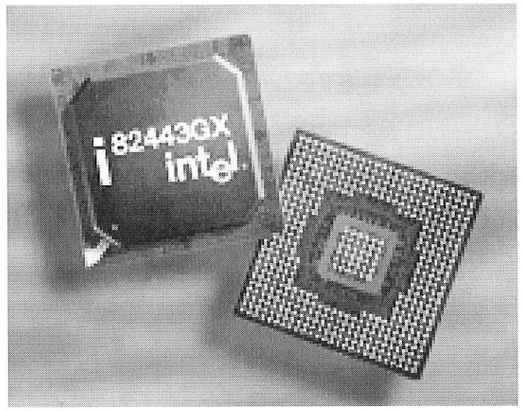
Bumps formed by modification of the wire-bond process. Courtesy of Kulicke and Soffa Industries, Inc. (K&S).

8.6.2 Ball-Grid Array (BGA)

Ball-grid arrays essentially apply the solder-ball approach to the package rather than to the chip as indicated in Fig. 8.12. Chips are mounted to the BGA substrate using either peripheral wire bonding or flip-chip technology, and then the chip is coated with some form of molding compound. Standard BGA array ball pitches are in the range of 1270 μm (50 mils). Fine-pitch ball-grid arrays (FBGA, or μBGA) are projected to decrease from a pitch of 500 μm in the year 2000 to 300 μm by 2010.



(a)



(b)

FIGURE 8.12

(a) Ball grid array cross section. (b) Intel microprocessor packaged using a BGA. Courtesy of Intel Corp.

8.6.3 The Tape-Automated-Bonding (TAB) Process

In *tape-automated bonding*, dice are attached to copper leads supported by a tape similar to 35-mm film. The film is initially coated with copper, and the leads are defined by lithography and etching. The lead pattern may contain hundreds of connections.

Die attachment requires a process similar in concept to the solder-ball technology discussed earlier. Gold bumps are formed on either the die or the tape and are used to bond the die to the leads on the tape. Figure 8.13 outlines the steps used to form a gold bump on a bonding pad [5]. A multilayer metal sandwich is deposited over the passivation oxide. Next, a relatively thick layer of photoresist is deposited, and windows are opened above the bonding pads. Electroplating is used to fill the openings with gold. The photoresist is removed, and the thin metal sandwich is etched away using wet or dry etching. The final result is a 25-μm-high gold bump standing above each pad. As in the flip-chip approach, bonding sites may be anywhere on the die.

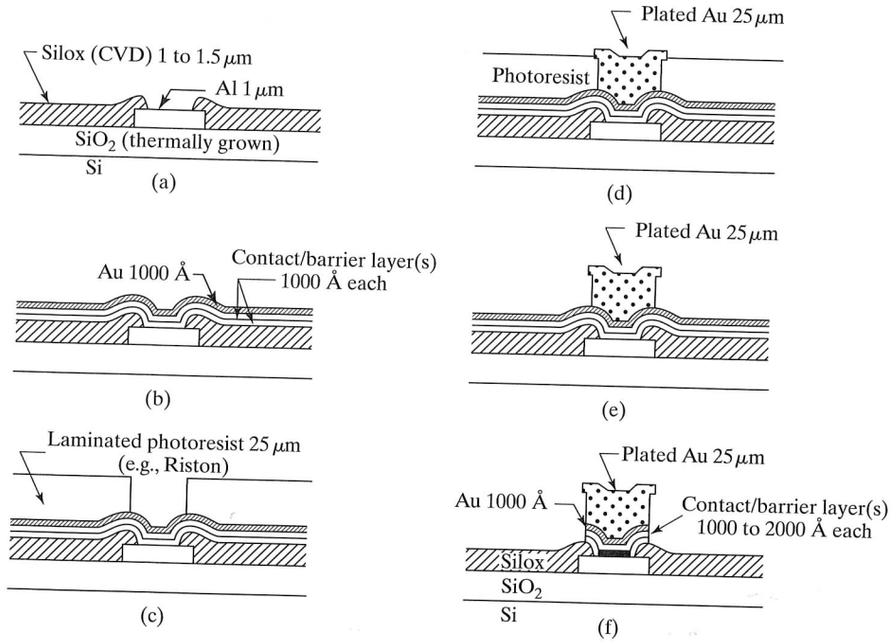


FIGURE 8.13

Process sequence for making gold bumps on aluminum metallurgy devices. (a) The wafer is cleaned and sputter-etched; (b) a contact/barrier layer (which also serves as a conductive film for electroplating) is sputter-deposited with a layer of gold for oxidation protection; (c) a thick-film photoresist (25 μm) is laminated and developed; (d) gold is electroplated to a height of approximately 25 μm to form the bumps; (e) the resist is stripped; (f) the sputter-deposited conductive film is removed chemically or by back-sputtering. Reprinted with permission of *Solid State Technology*, published by Technical Publishing, a company of Dun & Bradstreet, from Ref. [5].

The mounting process aligns the tape over the die, as in Fig. 8.14. A heated bonding head presses the tape against the die, forming thermocompression bonds. In a production process, a new die is brought under the bonding head and the tape indexes automatically to the next lead site.

TAB-mounted parts offer the advantage that they can be functionally tested and burned in once the dice are attached to the film. In addition, the IC passivation layer and gold bump completely seal the semiconductor surface.

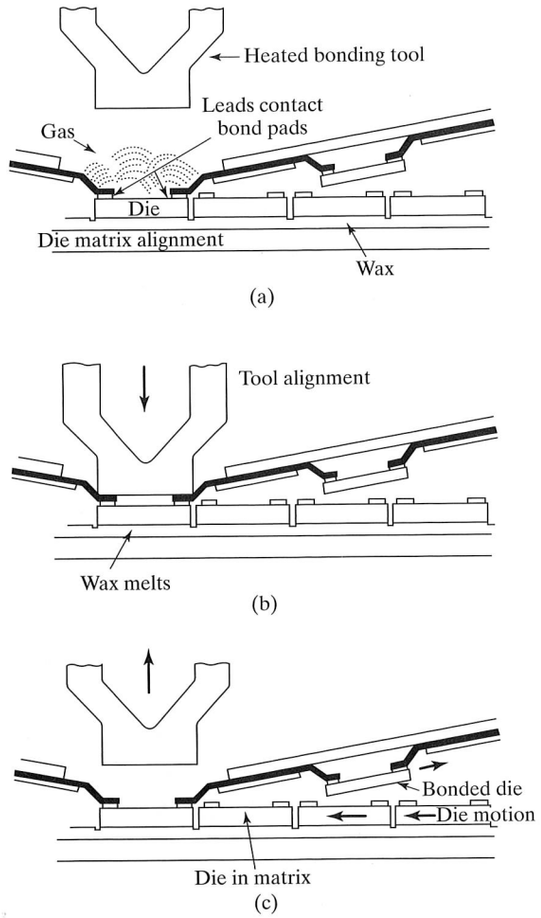


FIGURE 8.14

Tape-automated-bonding procedure. (a) Preformed leads of film are lowered into position and aligned above bonding pads on the die, which is held in place with a wax; (b) bonding tool descends and forms bond with pressure and heat; heat melts the wax, releasing the die; (c) tool and film are raised, lifting the bonded die clear so a new die can be moved into position and the process can be repeated. Reprinted with permission from *Small Precision Tools Bonding Handbook*. Copyright 1976.

8.6.4 Chip Scale Packages

Many of the technologies described, including wire bonding, TAB, and flip-chip mounting, are evolving to the Chip Scale Package (CSP) in which the goal is to achieve a package whose area is no larger than the die itself. Figures 8.15(a–b) present drawings of two possible approaches to the CSP, one based upon wire-bonded die and the second employing a form of TAB. Another approach referred to as chip-on-board attempts to eliminate the package altogether by mounting the bare die directly on the printed circuit board or flexible substrate as shown in Fig. 8.15(c). Either wire bonding or flip-chip mounting may be used, and the final structure is encapsulated with an epoxy glob-top coating.

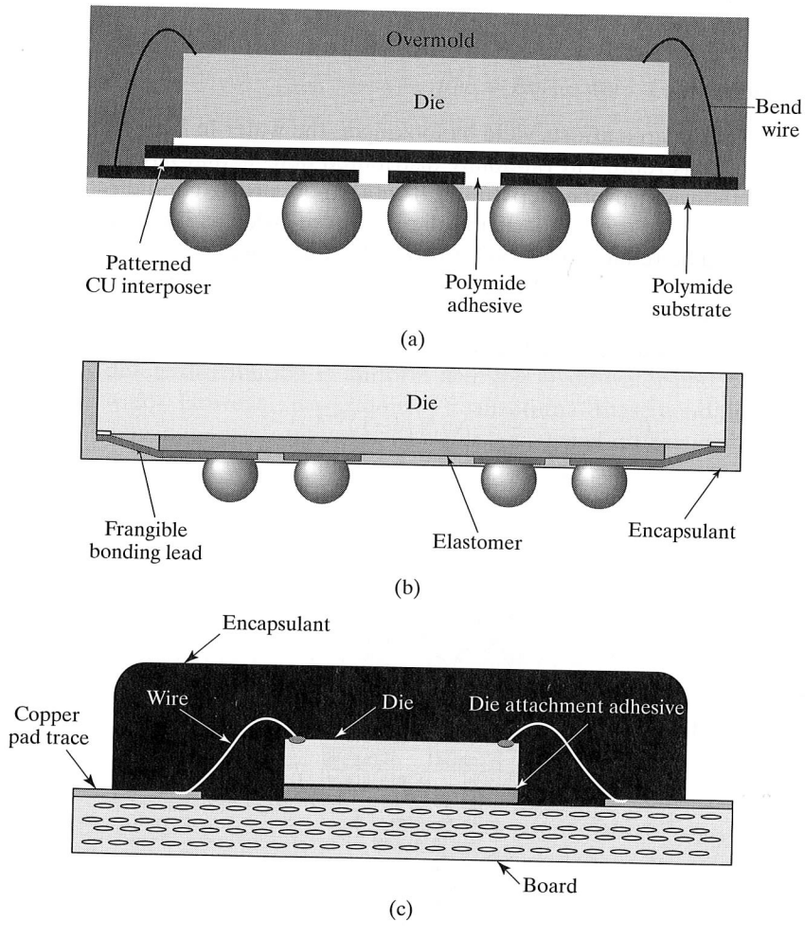


FIGURE 8.15

(a) Chip scale package using wire bonding. (b) Alternate form of a CSP. (c) Chip-on-board packaging.

8.7 YIELD

The manufacturer of integrated circuits is ultimately interested in how many finished chips will be available for sale. A substantial fraction of the dice on a given wafer will not be functional when they are tested at the wafer-probe step at the end of the process. Additional dice will be lost during the die separation and packaging operations, and a number of the packaged devices will fail final testing.

As mentioned earlier, the cost of packaging and testing is substantial and may be the dominant factor in the manufacturing cost of small die. For a large die with low yield, the manufacturing cost will be dominated by the wafer processing cost. A great deal of time has been spent attempting to model wafer yield associated with IC processes. Wafer yield is related to the complexity of the process and is strongly dependent on the area of the IC die.

8.7.1 Uniform Defect Densities

One can visualize how die area affects yield by looking at the wafer in Fig. 8.16, which has 120 die sites. The dots represent randomly distributed defects that have caused a die to fail testing at the wafer-probe step. In Fig. 8.16(a), there are 52 good dice out of the total of 120, giving a yield of 43%. If the die size were twice as large, as in Fig. 8.16(b), the yield would be reduced to 22% for this particular wafer.

An estimate of the yield of good dice can be found from a classical problem in probability theory in which n defects are randomly placed in N die sites. The probability P_k that a given die site contains exactly k defects is given by the binomial distribution:

$$P_k = \frac{n!}{k!(n-k)!} N^{-n}(N-1)^{n-k} \tag{8.1}$$

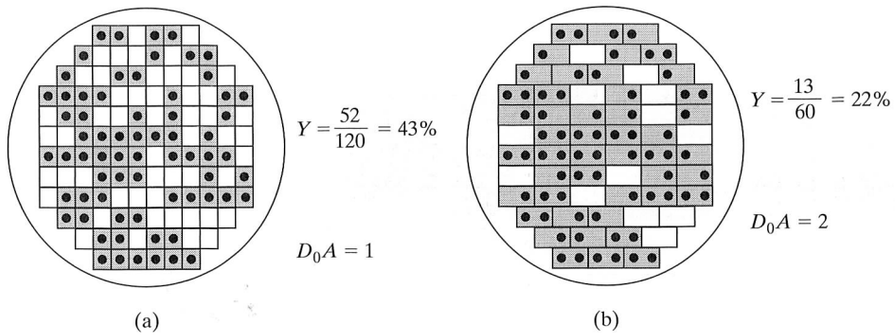


FIGURE 8.16

Illustrations of wafers, showing effect of die size on yield. Dots indicate the presence of a defective die location. (a) For a particular die size the yield is 43%; (b) if the die size were doubled, the yield would be only 22%.

For large n and N, Eq. (8.1) can be approximated by the Poisson distribution:

$$P_k = \frac{\lambda^k}{k!}\exp(-\lambda) \tag{8.2}$$

where $\lambda = n/N$. The yield is given by the probability that a die is found with no defects:

$$Y = P_0 = \exp(-\lambda) \tag{8.3}$$

The area of the wafer is equal to NA, where A is the area of one die. The density of defects, D_0, is given by the total number of defects, n, divided by the area of all the chips, and the average number of defects per die, λ, is given by

$$\lambda = n/N = D_0A, \quad \text{for} \quad \mathrm{D_0 = n/NA} \tag{8.4}$$

The yield based on the Poisson distribution then becomes

$$Y = \exp(-D_0A) \tag{8.5}$$

This expression was used to predict early die yield, but was found to give too low an estimate for large dice with $D_0A > 1$. Equation (8.5) implicitly assumes that the defect distribution is uniform across a given wafer and does not vary from wafer to wafer. However, it was quickly realized that these conditions are not realistic. Defect densities vary from wafer to wafer because of differences in handling and processing. On a given wafer, there are usually more defects around the edge of the wafer than in the center, and the defects tend to be found in clusters. These realizations led to investigation of nonuniform defect densities.

8.7.2 Nonuniform Defect Densities

Murphy [6] showed that the wafer yield for a nonuniform defect distribution can be calculated from

$$Y = \int_0^\infty \exp(-DA)f(D)\,dD \tag{8.6}$$

where $f(D)$ is the probability density for D. He considered several possible distributions, as shown in Fig. 8.17. The impulse function in Fig. 8.17(a) represents the case in which the defect density is the same everywhere, and substituting it for $f(D)$ in Eq. (8.6) yields Eq. (8.5). The triangular distribution in Fig. 8.17(b) is a simple approximation to a Gaussian distribution function and allows some wafers to have very few defects and others to have up to $2D_0$ defects. Application of Eq. (8.6) results in the following yield expression:

$$Y = \left[\frac{1-\exp(-D_0A)}{D_0A}\right]^2 \tag{8.7}$$

A uniform distribution of defect densities is modeled by $f(D)$ in Fig. 8.17(c) and predicts a yield of

$$Y = \left[\frac{1-\exp(-2D_0A)}{2D_0A}\right] \tag{8.8}$$

More complicated probability distributions have also been investigated, including the negative binomial and gamma distributions [7,8]. These result in the yield expression in Eq. (8.9)

$$Y = \left[1 + \frac{D_0A}{\alpha}\right]^{-\alpha} \tag{8.9}$$

in which α represents a clustering parameter which ranges from 0.5 to 10. The 1999 ITRS is basing future defect density requirements on Eq. (8.9) with a clustering parameter of 5 [9]. To achieve a yield loss of less than 20% due to random defects, the ITRS projects the required defect density to be less than 0.1/cm^2 by the year 2010 with a critical defect size of less than 30 nm.

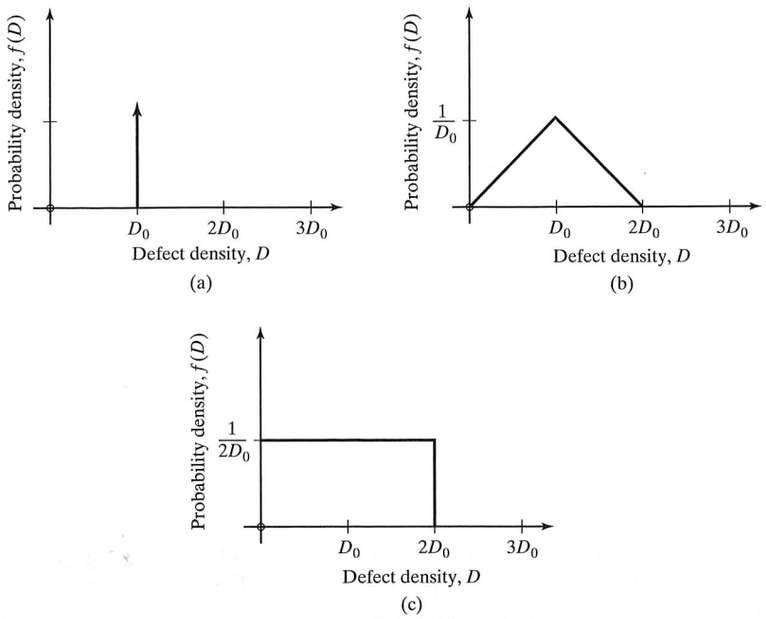


FIGURE 8.17

Possible defect probability density functions. (a) Impulse, where every wafer has exactly the same number of defects; (b) a triangular approximation to a Gaussian density; (c) a uniform density function.

Figure 8.18 plots the various yield functions versus D_0A, the average number of defects in a die of area A. Early yield estimates based on Poisson statistics are clearly much more pessimistic than those based on the other functions. However, the negative binomial yield model in Eq. (8.9) with $\alpha = 5$ is beginning to approach the exponential function. (See Problems 8.9 and 8.10.)

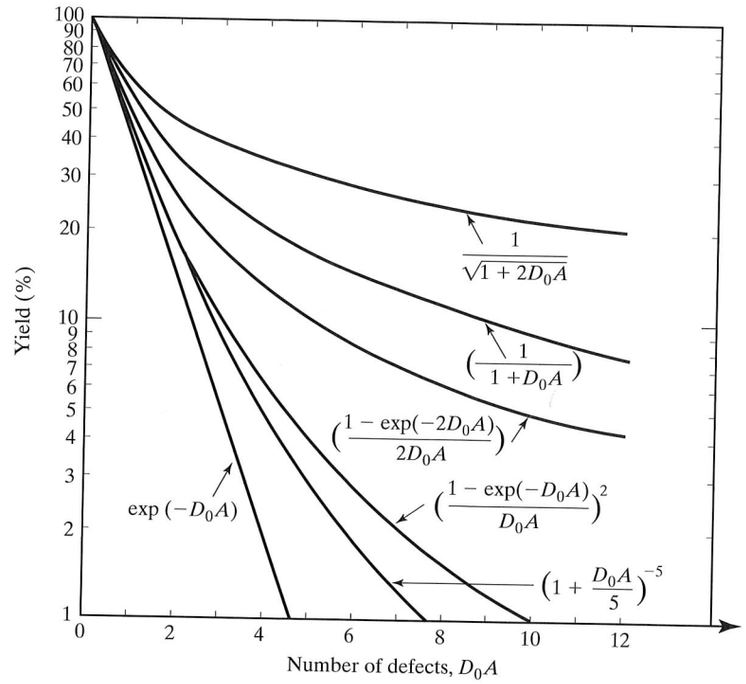


FIGURE 8.18

Theoretical yield curves for different defect densities. See Eqs. (8.5) through (8.9).

Example 8.1

A 150-mm wafer has a defect density of 10 defects/cm^2, and costs \$200 to process. The cost of assembly and testing is \$1.50 per die. (a) What is the total manufacturing cost for a 5×5 mm die in this process based on yield Eq. (8.7)? (The number of square dice per wafer is given approximately by $N = \pi\,(R - S)^2/S^2$, where R is the wafer radius and S is the length of the side of the die.) (b) The market price for this part is \$2.50. What must be the wafer yield needed for manufacturing cost to drop below the market price?

Solution: The area of the die is 0.25 cm, so the average number of defects per die is $D_0A = 2.5$. Equation (8.7) predicts a yield of 13.5%. The wafer has a radius of 75 mm and contains approximately 616 dice. So the average wafer will yield 83 good dice. The cost of the packaged dice will be $C = (\$200/83) + \$1.50 = \$3.91$. Getting the cost to the market price requires $\$2.50 = (\$200/ND) + \$1.50$. We must get $ND = 200$ good dice per wafer to break even, corresponding to a yield of $Y = 200/616 = 0.325$ or more.

SUMMARY

Following the completion of processing, wafers are screened by checking various processing and device parameters using special test sites on the wafer. If the parameters are within proper limits, each die on the wafer is tested for functionality, and bad dice are marked with a drop of ink.

Next, the dice are separated from the wafer using a diamond saw or a scribe-and-break process. Some die loss is caused by damage during the separation process. The remaining good dice are mounted in ceramic or plastic DIPs, LCCs, PGAs, surface-mount, or BGA packages using epoxy or eutectic die-attachment techniques.

Bonding pads on the die are connected to leads on the package using ultrasonic or thermosonic bonding of 15 to 75 μm aluminum or gold wire. Batch-fabricated flip-chip and TAB interconnection processes that permit simultaneous formation of hundreds or even thousands of bonds can also be used.

The final manufacturing cost of an integrated circuit is determined by the number of functional parts produced. The overall yield is the ratio of the number of working packaged dice to the original number of dice on the wafer. Yield loss is due to defects on the wafer, processing errors, damage during assembly, and lack of full functionality during final testing. The relationship between wafer yield and the size of an integrated-circuit die has been explored in detail. The larger the die size, the lower will be the number of good dice available from a wafer.

REFERENCES

[1] J. W. Stafford, "The Implications of Destructive Wire Bond Pull and Ball Bond Shear Testing on Gold Ball–Wedge Wire Bond Reliability," *Semiconductor International*, p. 82, May 1982.

[2] W. C. Till and J. T. Luxon, *Integrated Circuits: Materials, Devices and Fabrication*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[3] J. R. Howell, "Reliability Study of Plastic Encapsulated Copper Lead Frame Epoxy Die Attach Packaging System," *Proceedings of the Reliability Physics Symposium, IEEE* pp. 104–110, 1981.

[4] P. A. Totta and R. P. Sopher, "SLT Device Metallurgy and Its Monolithic Extension," *IBM Journal of Research and Development, 13*, 226–238, May 1969.

[5] T. S. Liu, W. R. Rodgrigues de Miranda, and P. R. Zipperlin, "A Review of Wafer Bumping for Tape Automated Bonding," *Solid State Technology, 23*, 71–76, March 1980.

[6] B. T. Murphy, "Cost-Size Optima of Monolithic Integrated Circuits," *Proceedings of the IEEE, 52*, 1537–1545, December 1964.

[7] R. B. Seeds, "Yield and Cost Analysis of Bipolar LSI," *IEEE IEDM Proceedings*, p. 12, October 1967.

[8] C. H. Stapper, "On Yield, Fault Distributions, and Clustering of Particles," *IBM Journal of Research and Development, 30*, 326–338, May 1986.

[9] *The International Technology Roadmap for Semiconductors*, The Semiconductor Industry Association (SIA), San Jose, CA: 1999. (http://www.semichips.org)

FURTHER READING

[1] G. G. Harman and J. Albers, "The Ultrasonic Welding Mechanism as Applied to Aluminum- and Gold-Wire Bonding in Microelectronics," *IEEE Transactions on Parts, Hybrids and Packaging, PHP-13*, 406–412, December 1977.

[2] K. I. Johnson, M. H. Scott, and D. A. Edson, "Ultrasonic Wire Welding–Part I: Wedge-Wedge Bonding of Aluminum Wires," *Solid State Technology, 20*, 50–56, March 1977.

[3] K. I. Johnson, M. H. Scott, and D. A. Edson, "Ultrasonic Wire Welding–Part II: Ball-Wedge Wire Welding," *Solid State Technology, 20*, 91–95, April 1977.

[4] C. Plough, D. Davis, and H. Lawler, "High Reliability Aluminum Wire Bonding," *Proceedings of the Electronic Components Conference, IEEE*, pp. 157–165, 1969.

[5] N. Ahmed and J. J. Svitak, "Characterization of Gold–Gold Thermocompression Bonding," *Proceedings of the Electronic Components Conference, IEEE*, pp. 92–97, 1976.

[6] L. S. Goldmann, "Geometric Optimization of Controlled Collapse Interconnections," *IBM Journal of Research and Development, 13*, 251–265, May 1969.

[7] K. C. Norris and A. H. Landzberg, "Reliability of Controlled Collapse Interconnections," *IBM Journal of Research and Development, 13*, 266–271, May 1969.

[8] J. E. Price, "A New Look at Yield of Integrated Circuits," *Proceedings of the IEEE, 58*, 1290–1291, August 1970.

[9] C. H. Stapper, Jr., "On a Composite Model to the IC Yield Problem," *IEEE Journal of Solid-State Circuits, SC-10*, 537–539, December 1975.

[10] C. H. Stapper, "The Effects of Wafer to Wafer Defect Density Variations on Integrated Circuit Defect and Fault Distributions," *IBM Journal of Research and Development, 29*, 87–97, January 1985.

PROBLEMS

8.1 Make a list of at least ten process or device parameters which could easily be monitored using a special test site on a wafer.

8.2 A simple microprocessor contains 115 flip-flops and hence 2^{115} possible states. If a tester can perform a new static test every 100 nsec, how many years will it take to test every state in the microprocessor chip? If the wafer has 100 dice, how long will it take to test the wafer?

8.3 **(a)** How many pads can be placed on a 10×15 mm die if a single row of pads is used? Assume the use of 100-μm pads on a 125-μm pitch with no pads in the corners.

(b) Repeat for 75-μm pads on a 100-μm pitch.

(c) How many solder balls would fit on the same die using an area array if the ball pads were 125-μm pads on a 200-μm pitch?

8.4 Compare the five yield formulas for a large VLSI die in which $D_0A = 10$ defects. Assume a clustering parameter of 1.0. How many good dice can we expect from 100- and 150-mm-diameter wafers using the different yield expressions? (The number of square dice per wafer can be estimated from $N = \pi\ (R - S)^2/S^2$, where R is the radius of the wafer and S is the length of one side of the die.) Assume $S = 5$ mm.

8.5 What is the wafer yield for the defect map in Fig. 8.16 if the die is four times the size of that in Fig. 8.16(a)? What is the yield predicted by Poisson statistics? Assume the data from Fig. 8.16 is best represented by Eq. (8.9). What value of clustering parameter best fits the data?

8.6 A new circuit design is estimated to require a die which is 5×8 mm and will be fabricated on a wafer 125 mm in diameter. The process is achieving a defect density of 10 defects/cm^2, and the wafer processing cost is $250.

(a) What will be the cost of the final product if testing and packaging adds \$1.60 to the completed product?

(b) The circuit design could be partitioned into two chips rather than one, but each die will increase in area by 15% in order to accommodate additional pads and I/O circuitry. If the testing and packaging cost remains the same, what is the cost of the two-chip set? Base your answers on Eq. (8.8). (See Problem 8.4 for the number of dice per wafer.)

8.7 **(a)** Repeat Problem 8.5 for a defect density of 5 defects/cm^2 and a wafer cost of \$150.

(b) Repeat Problem 8.5 for a defect density of 5 defects/cm^2 and a wafer cost of \$300.

8.8 A die has an area of 25 mm^2 and is being manufactured on a 100-mm-diameter wafer using a process rated at 2 defects/cm^2. A new process is being developed which allows the die area to be reduced by a factor of 2. However, because of the smaller feature sizes, the new process costs 30% more and is presently achieving only 10 defects/cm^2.

(a) Is it economical to switch to this new process?

(b) At what defect density does the cost of the new die equal the cost of the old die?

(c) Based on your judgment, would you recommend switching to the new process even if it is not now economical? Why?

(d) At what die size is the cost the same in either process? Use Eq. (8.8) for this problem.

8.9 What is the limit of the yield distribution in Eq. (8.9) as the clustering parameter approaches infinity?

8.10 Compare the predictions of yield equations 8.5 and 8.9 for D_0A ranging from 1 to 10 with $\alpha = 5$ and $\alpha = 5{,}000$.

8.11 Suppose $D_0 = 0.1/\text{cm}^2$. What is the average number of defects on 150 mm, 200 mm, and 300 mm wafers?

8.12 Suppose that going from 100-mm wafers to 150-mm wafers changes the wafer processing cost from \$150/wafer to \$250/wafer, and the defect density remains constant at 10 defects/cm^2. Assume a die cost of \$1.00 to find the die sizes. Use Eq. (8.9) with a cluster factor of 2. Use a calculator or computer to find the answer by iteration.

8.13 What would be the die yield in Fig. 8.16(b) if the defect positions were the same but the die pattern was rotated by 90°? How many good dice with four times the area of that in Fig. 8.16(a) would now exist?

8.14 A Gaussian probability density function for defect density is given by

$$f(D) = \frac{2}{D_0\sqrt{\pi}} \exp -\left[\frac{2(D-D_0)}{D_0}\right]^2, \text{ for } 0 \le D \le 2D_0, \text{ and 0 otherwise}$$

Calculate the yield Y for various values of D_0A and compare your results to those of the triangular distribution given in Eq. (8.7). (You may want to use a calculator or computer to perform the iteration.)

8.15 The wafers shown in Fig. 8.16 actually have 120 defects placed randomly on the wafer. Obviously, some chips must have several defects. Use Eq. (8.1) to predict how many dice will have exactly 1, 2, 3, 4, and 5 defects.

8.16 What defect density is required to achieve a yield of 70% for a 10×15 mm die if the process is characterized by a cluster parameter of 5? (b) Repeat for 80% yield. (c) Repeat for 90% yield.

8.17 What defect density is required to achieve a yield of 75% for a 20×20 mm die if the process is characterized by a cluster parameter of 6? (b) Repeat for 85% yield.

CHAPTER 9

MOS Process Integration

In Chapter 9, we explore a number of relationships between process and device design and circuit layout. Processes are usually developed to provide devices with the highest possible performance in a specific circuit application, and one must understand the circuit environment and its relation to device parameters and device layout.

In this chapter, we look at a number of basic concerns in MOS process design, including channel-length control; layout ground rules and ground-rule design; source–drain breakdown and punch-through voltages; and threshold-voltage adjustment. Metal-gate technology is discussed, and the important advantages of self-aligned silicon-gate technologies are presented. Discussions of CMOS and silicon-on-insulator technologies complete the chapter.

9.1 BASIC MOS DEVICE CONSIDERATIONS

To explore the relationship between MOS process design and basic device behavior, we begin by discussing the static current–voltage relationship for the MOS transistor, as developed in Volume IV of this series [1]. The cross section of two NMOS transistors is shown in Fig. 9.1. In the linear region of operation, the drain current is given by

$$I_D = \bar{\mu}_n C_O (w/L)(V_{GS} - V_{TN} - V_{DS}/2)V_{DS} \tag{9.1}$$

for $V_{GS} \ge V_{TN}$ and $V_{DS} \le V_{GS} - V_{TN}$. $C_O = K_O\varepsilon_O/X_O$ is the oxide capacitance per unit area, $\bar{\mu}_n$ is the average majority-carrier mobility in the inversion layer, and V_{TN} is the threshold voltage. W and L represent the width and length of the channel, respectively.

One of the first specifications required is the circuit power-supply voltages, which set the maximum value of V_{GS} and V_{DS} that the devices must withstand. Once this choice is made, the only variables in Eq. (9.1) that a circuit designer may adjust are the width and length of the transistor. Thus, the circuit designer varies the circuit topology and horizontal geometry to achieve the desired circuit function.

Other device parameters are fixed by the process designer, who must determine the process sequence, times, temperatures, etc., which ultimately determine the device