2 BASIC INTEGRATED CIRCUIT MANUFACTURING

A. STARTING MATERIAL

Silicon is one of nature's most useful elements. Silicon is the material most commonly used for the manufacturing of semiconductors. Silicon, as a pure chemical element, is not found free in nature. It exists primarily in compound form with other chemical elements. In all of its various forms, silicon makes up 25.7% of the earth's crust, and is the second most abundant element in the Periodic Table Of Elements. It is exceeded only by oxygen. Silicon occurs chiefly as a compound of silicon and oxygen called an oxide or as a compound of silicon and salts called a silicate.

Silicon in the form of an oxide most commonly occurs as silicon dioxide, SiO_2, generally called silica, or sand. Other common forms of silicon dioxide are quartzite, quartz, rock crystal, amethyst, agate, flint, jasper, and opal. Many of the previous forms contain minute quantities of other elements that give the forms color. Some of these minerals are known as semi-precious gem stones.

Granite, hornblende, asbestos, feldspar, clay, mica, etc., are but a few of the numerous silicate minerals. Silicon, as sand, is one of the main ingredients of glass. Silicon is an important component in steel, aluminum alloys, and other metallugurical products. Silicon carbide is one of the more useful abrasive materials.

1. Purification

Silicon for semiconductor applications is taken from quartzite, the rock form of silicon dioxide. Quartzite has trace levels of other elements that must be removed in the purification process. The purifying of quartzite consists of reacting the quartzite with some form of carbon material. The carbon may be in the form of coal, coke, or wood. The process is performed at a temperature of approximately 2000°C. This chemical reaction produces metallurgical grade silicon (MGS) that is 98% pure, which is not good enough for semiconductor use, so must be further purified. This silicon is reacted with very strong hydrochloric acid (similar to strong swimming pool acid) to form a new liquid called trichlorosilane. The liquid is then purified by fractional distillation (similar to the distillation process to make whiskey). The resulting, highly purified trichlorosilane liquid is converted to polycrystalline electronic grade silicon (EGS) by the Siemens' process. The Siemens' process changes the liquid into a solid polycrystalline silicon (usually called polysilicon) rod. These process steps are summarized in Figure 2-1.

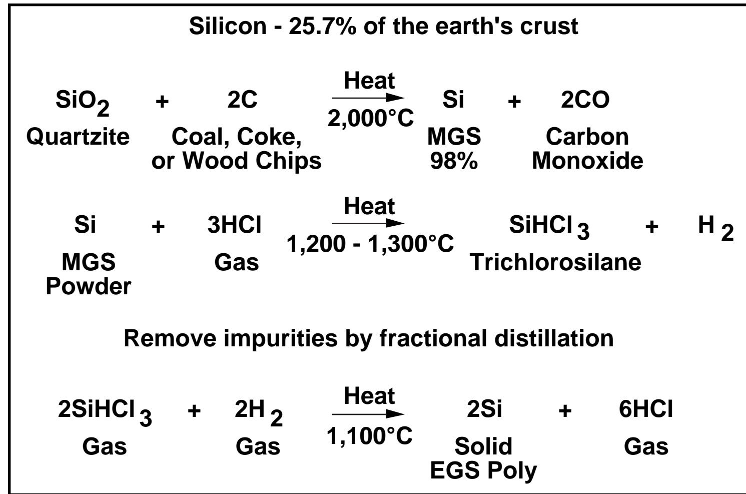


Figure 2-1. Polysilicon Creation

2. Czochralski Crystal Growing

The next process step converts the very pure silicon from a polysilicon crystal form into a single crystal or monocrystalline form (Figure 2-2). This process is known as Czochralski crystal growing, often called Cz, the abbreviation for Czochralski.

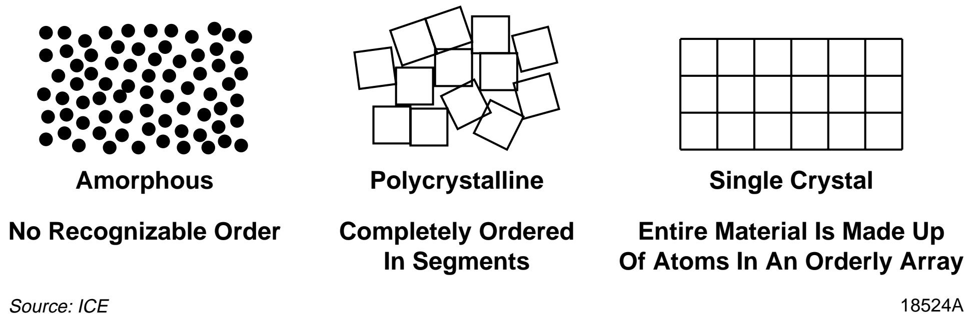


Figure 2-2. Types of Silicon Structures

The Cz process involves a special crystal growing furnace (Figure 2-3) wherein the pure polysilicon is placed into a very pure quartz crucible, along with the dopant material to make the wafer N-type or P-type. The process chamber is evacuated and purged with very pure argon. The crucible is heated to the melting point of silicon, 1420°C. A previously loaded monocrystalline silicon "seed" of the desired crystal orientation is lowered into the molten silicon and then slowly withdrawn as the "seed" and crucible rotate in opposite directions. This process causes the molten silicon to freeze out onto the "seed" crystal, forming a monocrystalline silicon ingot. The dopant species added to the polysilicon material determines the crystal's electrical characteristics.

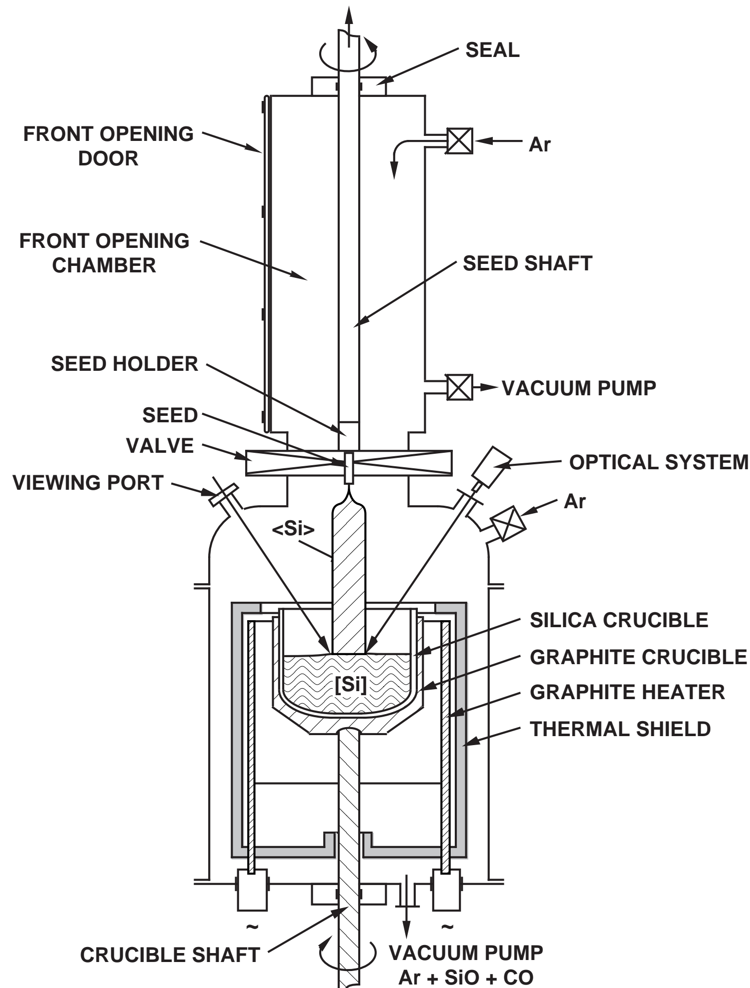


Figure 2-3. Czochralski Silicon Crystal Grower

The growing process is very automated and yields quality crystals. The growth process is very slow, typically 0.5 inch per hour for 150mm diameter crystals. Because of slow growth rates, the manufacturing process consumes large quantities of electricity.

After the growing process is completed, the silicon ingot is evaluated for both electrical and mechanical parameters. Generally, the yield is high and the usable silicon efficiency is 80 to 90 percent of the original crucible charge. This helps keep the cost down.

The diameter of the ingot during crystal growing cannot be controlled to the tolerance required by wafer handling equipment. To meet the diameter tolerance, the ingot is centerless ground to the exact diameter. Another part of the grinding process is to grind a major flat and minor flat or flats for crystal plane designation. The major flat is used for wafer orientation in wafer handling equipment and as a visual indicator to the manufacturing personnel.

3. Sawing Crystal Into Wafers

The next sequence of process steps involves sawing the ingot into the individual wafers and edge grinding the outer edge of the wafer circumference to a controlled shape. The sawing concept is illustrated in Figure 2-4. The edge grinding process removes the sharp edges of the wafer and dramatically reduces silicon particles when the wafer is handled.

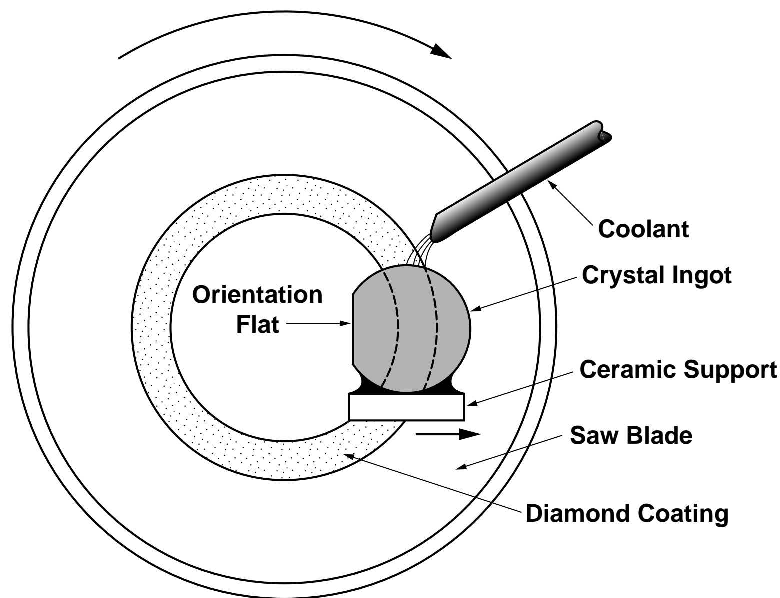


Figure 2-4. Inner-Diameter (ID) Sawing of Silicon Ingot

4. Wafer Preparation

The wafer is double-side (both sides simultaneously) lapped to reduce the surface roughness and then chemically etched to further smooth both sides of the wafer surface. A final surface polish is done on the designated side of the wafer. The designated side is determined by the ground wafer flat(s). The final surface must be very smooth (like a mirror) and free of surface defects and imperfections. A final chemical cleaning process will remove the polishing materials, particulants and any other potentially contaminating materials. These process steps are illustrated in Figure 2-5. Electrical and mechanical evaluation completes the processing.

The wafers are packaged in an ultra-clean environment and sealed in the storage-shipping containers. They are ready for use in the fabrication process.

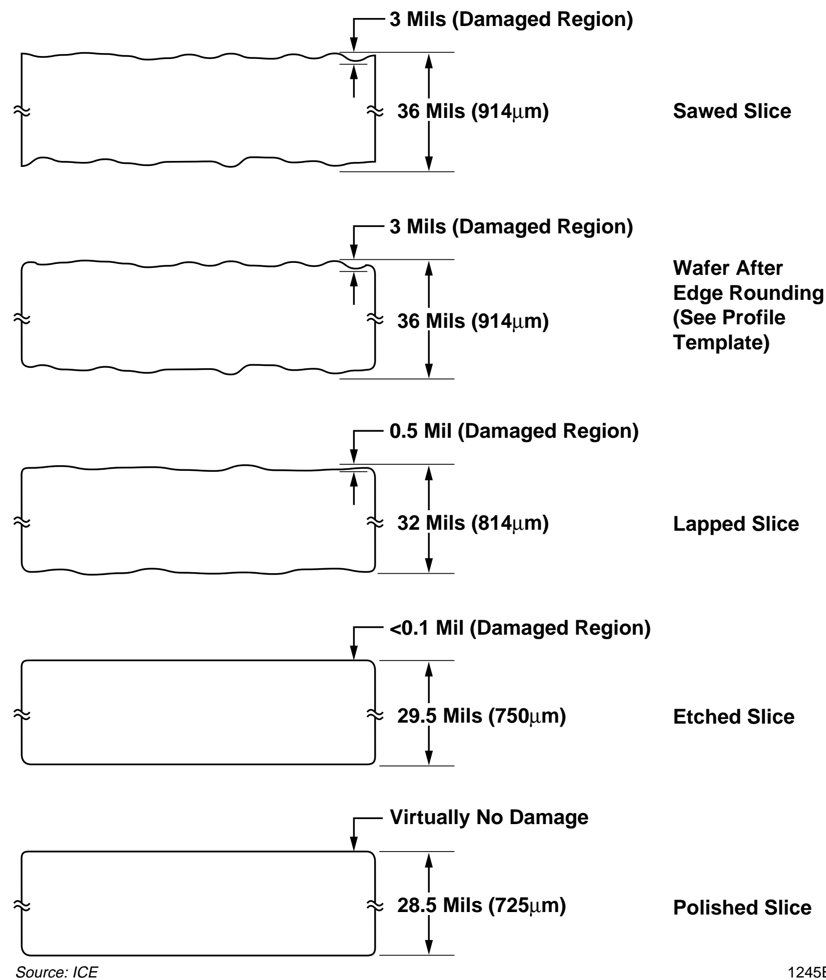


Figure 2-5. Wafer Preparation (for 150mm)

B. WAFER CLEANING

Contamination control during IC manufacturing is a major factor for Yield, Cost, Reliability, and Quality. Therefore, the total manufacturing cycle *MUST BE* controlled continuously to meet these needs. This means there is control over the environment, materials, chemicals, people, equipment and the process interactions with all of the above.

The physical size of contamination can vary over a large range. Figure 2-6 illustrates some of the contaminating materials and sizes. Contamination in any form, films or particles, larger than 10 percent of the smallest line width is considered detrimental. This is illustrated in Figure 2-7. In general, anything on the wafer surface that is not designed to be there is considered contamination.

An extremely critical part of the manufacturing sequence is the cleaning of the wafer surface after certain process steps and prior to other process steps.

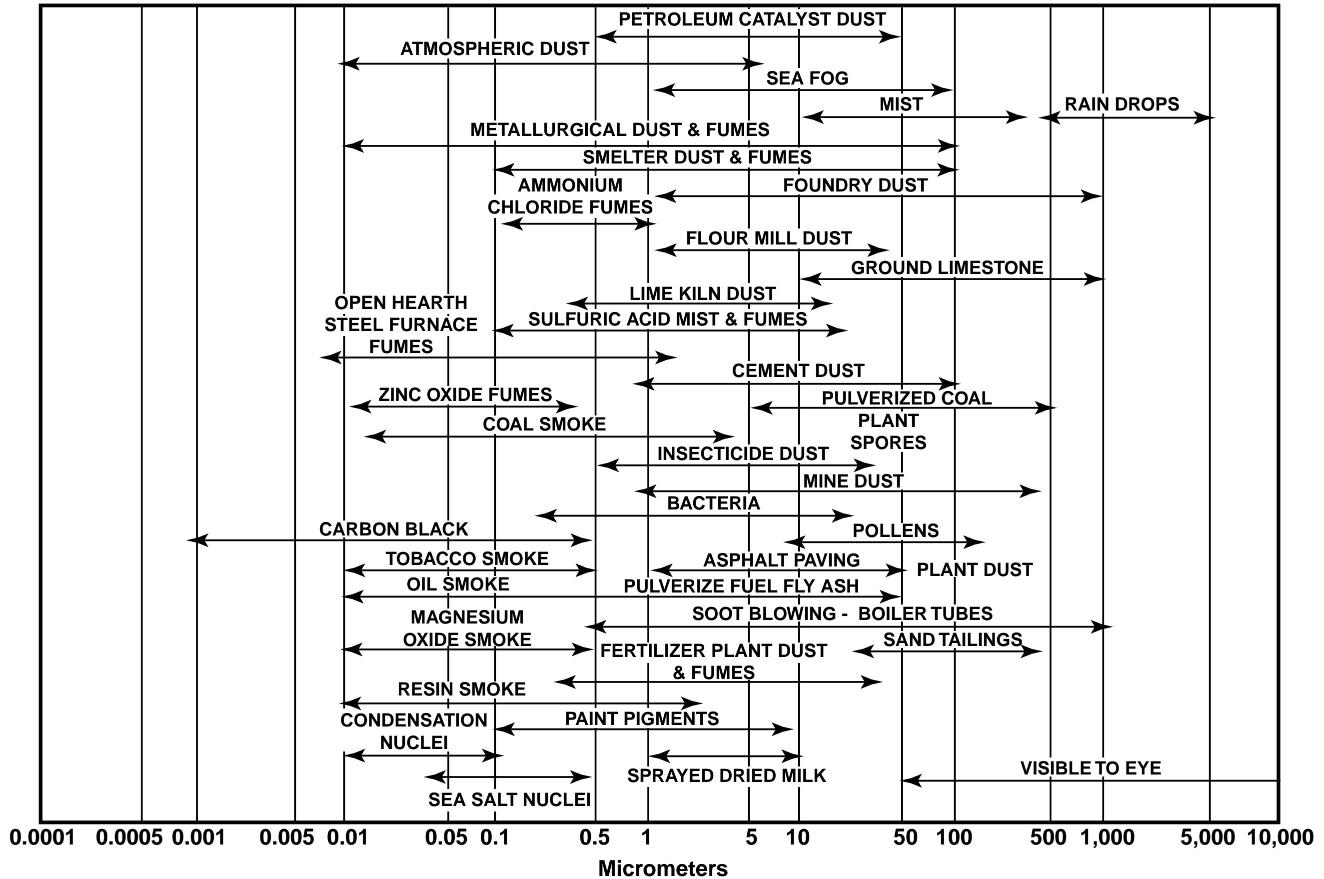


Figure 2-6. Sizes of Airborne Contaminants

Cleaning processes are required before the wafers are introduced into any elevated temperature process. A cleaning process is necessary to remove any form of surface contamination that may create a defective transistor within an IC die or produce instability in the circuit during the lifetime of the IC. The purity of the wafer surface is essential.

Contaminating materials on the wafer surface can lead to some of the following problems.

1. Prevent or mask effective cleaning or rinsing.
2. Prevent or mask effective concentration of dopants to be introduced into the silicon, whether by diffusion or ion implantation.
3. Cause poor or no adhesion of deposited layers.
4. Cause undesired chemical reactions and lead to decomposition of materials.
5. Alter the silicon crystal structure causing undesired electrical parameter changes.
6. Lead to long-term instability of electrical parameters.
7. Cause film degradation or catastrophic device failure.

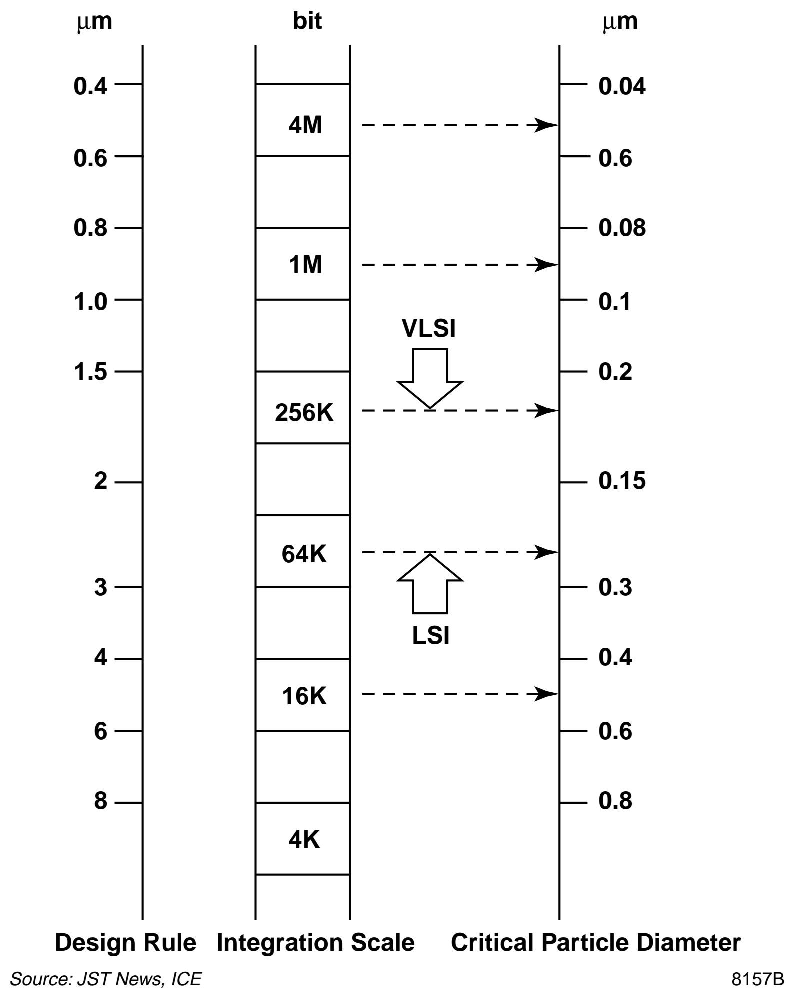


Figure 2-7. Correspondence Between Scale of Integration (Design Rule) and Size of Particles to be Removed

Many chemical cleaning processes have been evaluated for cleaning silicon wafer surfaces. These processes must be divided into two categories: those processes used prior to the metallization process and those used after the metallization process because of chemical attack on metal.

The selected process must remove a variety of contaminates and leave nothing on the surface as the result of the chemical reaction. This requirement has lead to the use of hydrogen peroxide (H2O2) as the preferred cleaning solution. The peroxide is a 30 percent concentration and unstabilized chemically. The unstabilized form is necessary because of purity.

With peroxide as the starting solution, two different chemical cleaning processes have evolved. One is known as the RCA clean. The other cleaning process is known by several different names: Piranha, Caro, and Sulfuric/Peroxide. These cleaning processes are summarized in Figure 2-8.

RCA TYPE CLEANING	SULFURIC ACID/PEROXIDE CLEANING (PIRANHA) (see note below)
6:1:1 of H_2O; 30 W/W% H_2O_2; and 29 W/W% NH_4OH (as NH_3) followed by a mixture of 6:1:1 of H_2O; 30 W/W% H_2O_2; and 37% HCl. Keep both mixtures at 75° - 80°C for 15 minutes. UPDI rinse to 15 meg·ohm; spin/rinse/dry.	CLEAN A H_2SO_4/H_2O_2 clean — 10 minutes UPDI/quick dump rinse — to specified resistivity 50:1 etch (H_2O:HF) — 60 seconds UPDI/quick dump rinse — to specified resistivity UPDI rinse to 15 meg·ohm Spin/rinse/dry
	CLEAN B H_2SO_4/Ammonium persulfate — 10 minutes UPDI/quick dump rinse — to specified resistivity 50:1 etch (H_2O:HF) — 60 seconds UPDI/quick dump rinse — to specified resistivity UPDI rinse to 15 meg·ohm Spin/rinse/dry

Note: Stabilize H_2SO_4 acid at 80°C before adding H_2O_2. Add H_2O_2 to make a 4:1 (H_2SO_4:H_2O_2) solution. Solution temperature will rise to approximately 180°C after adding H_2O_2.

Source: ICE 14759A

Figure 2-8. Wafer Cleaning Techniques Prior to Thermal Processing

The RCA clean is a two-step process with a high-purity DI water rinse following each step. The first step uses a mixture of one part peroxide (H_2O_2), one part ammonium hydroxide (NH_4OH), and five parts DI water. The solution is heated to 80°C. The wafers are immersed in the cleaning solution for 10 - 15 minutes and then are immediately placed in the second solution. The second solution is a mixture of one part peroxide, one part hydrochloric acid (HCl), and six parts DI water. This solution is also heated to 80°C. The wafers are left in this solution for 10 - 15 minutes followed by a DI water rinse and are dried. This process is most often used before a high-temperature processing step.

The two steps are sometimes called the SC-1 and SC-2 cleans where SC-1 is the DI water, peroxide, ammonium hydroxide solution and SC-2 is the DI water, peroxide, hydrochloric acid solution.

The sulfuric acid and hydrogen peroxide solution, often written as H_2SO_4-H_2O_2, is most often used to strip photoresist from the wafer surface.

The sulfuric acid-hydrogen peroxide solution is formulated by heating the sulfuric acid to 80 - 100°C and adding the peroxide just prior to the wafer cleaning step. The solution contains three to five volumes of sulfuric acid to one volume of peroxide. Since both sulfuric acid and hydrogen peroxide are strong oxidizing agents, mixing them together causes an exothermic chemical reaction (liberates heat) raising the solution temperature to 150 - 180°C, depending on the chemical ratios. This sudden increase in temperature aids in the chemical cleaning process. The wafers are immersed in the solution for 10 - 15 minutes followed by a DI water rinse and are dried.

C. DIELECTRIC FORMATION

A dielectric is a material that is a poor conductor of electricity or does not conduct at all. Dielectric layers are often called insulators. Dielectric layers play a critical role in the manufacturing and operation of semiconductors. They are used:

- for insulation between conducting layers (e.g., for devices with more than one level of metal and to separate the gate from the silicon in an MOS transistor),
- to protect the surface of a completed die,
- to mask off portions of the surface of the wafer during some manufacturing operations, and
- between the plates of capacitors in ICs.

Various forms of dielectrics related to semiconductors include silicon dioxide, silicon nitride, and silicon oxynitride. The most common is silicon dioxide.

1. Thermally Grown Silicon Dioxide

Silicon has a very unique property when exposed to any source of oxygen. A chemical reaction that forms silicon dioxide (SiO_2) will occur. Silicon dioxide is a very stable dielectric or insulating material. Silicon will react with oxygen in the air at room temperature to form a very thin layer of silicon dioxide. This layer will be 20 to 30Å thick wherein the thickness stops the chemical reaction. The process of reacting the silicon wafer with some source of oxygen is referred to as thermal oxidation. In other words, silicon at the interface between the oxide formed and the silicon wafer is being consumed by the chemical reaction of forming silicon dioxide. These silicon atoms are permanently changed into the compound, silicon dioxide. The effects of thermal oxidation are shown in Figure 2-9.

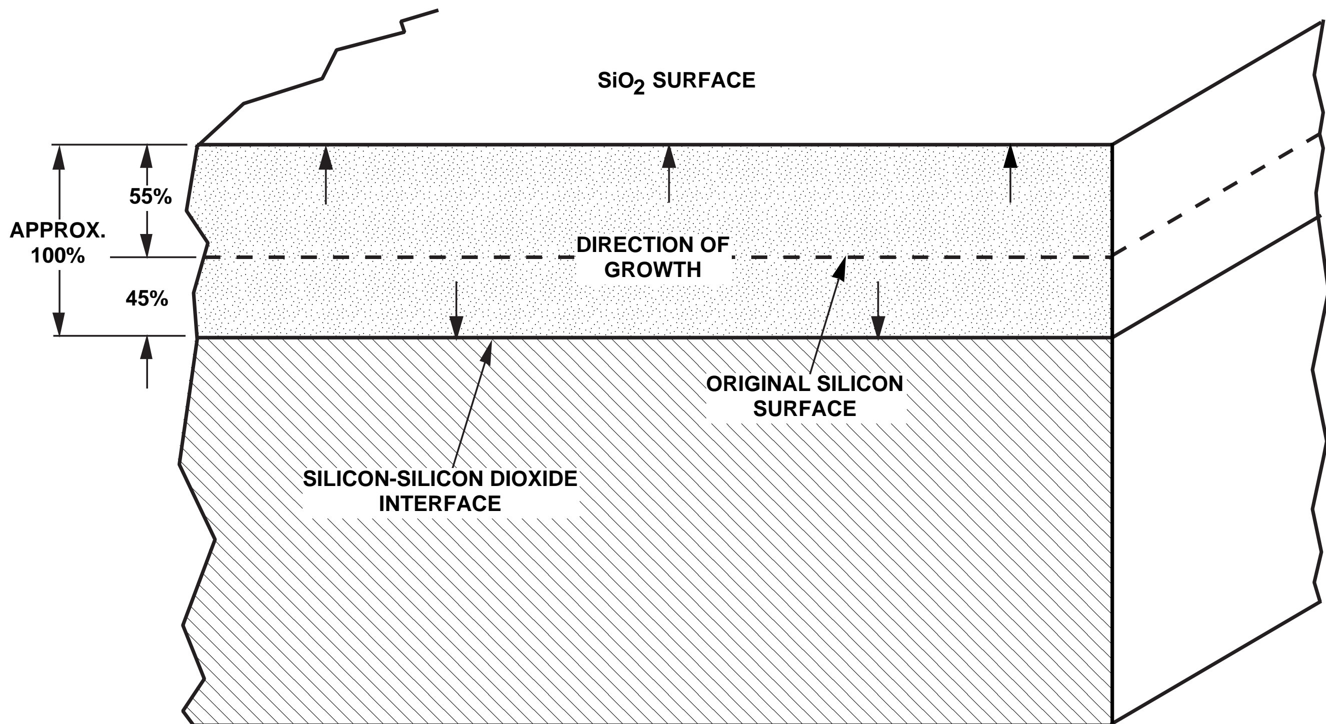


Figure 2-9. The Effects of SiO_2 Growth

a. Atmospheric-Pressure Oxidation

In semiconductor manufacturing, thermal oxidation is a well understood process. The process is done in the temperature range of 750°C to 1150°C to increase the reaction rate. The rate of oxidation also depends on the source of oxygen and the pressure inside the process chamber. The process parameters of time, temperature, oxygen source and pressure are determined by the overall process requirements and circuit design considerations.

The thermal oxidation oxidation rate is affected by the process engineer's choice of dry oxygen (pure O_2) or some form of water (H_20) vapor. Generally, if a thin oxide (<1000Å) is required, dry oxygen is used. Thicker oxides use water vapor. Oxidation using water vapor formed from the chemical reaction of gaseous oxygen (O_2) and hydrogen (H_2) is known as pyrogenic steam oxidation. This process has replaced the older process of using deionized water. The previous types of thermal oxidation are done at atmospheric pressure and illustrated in Figure 2-10.

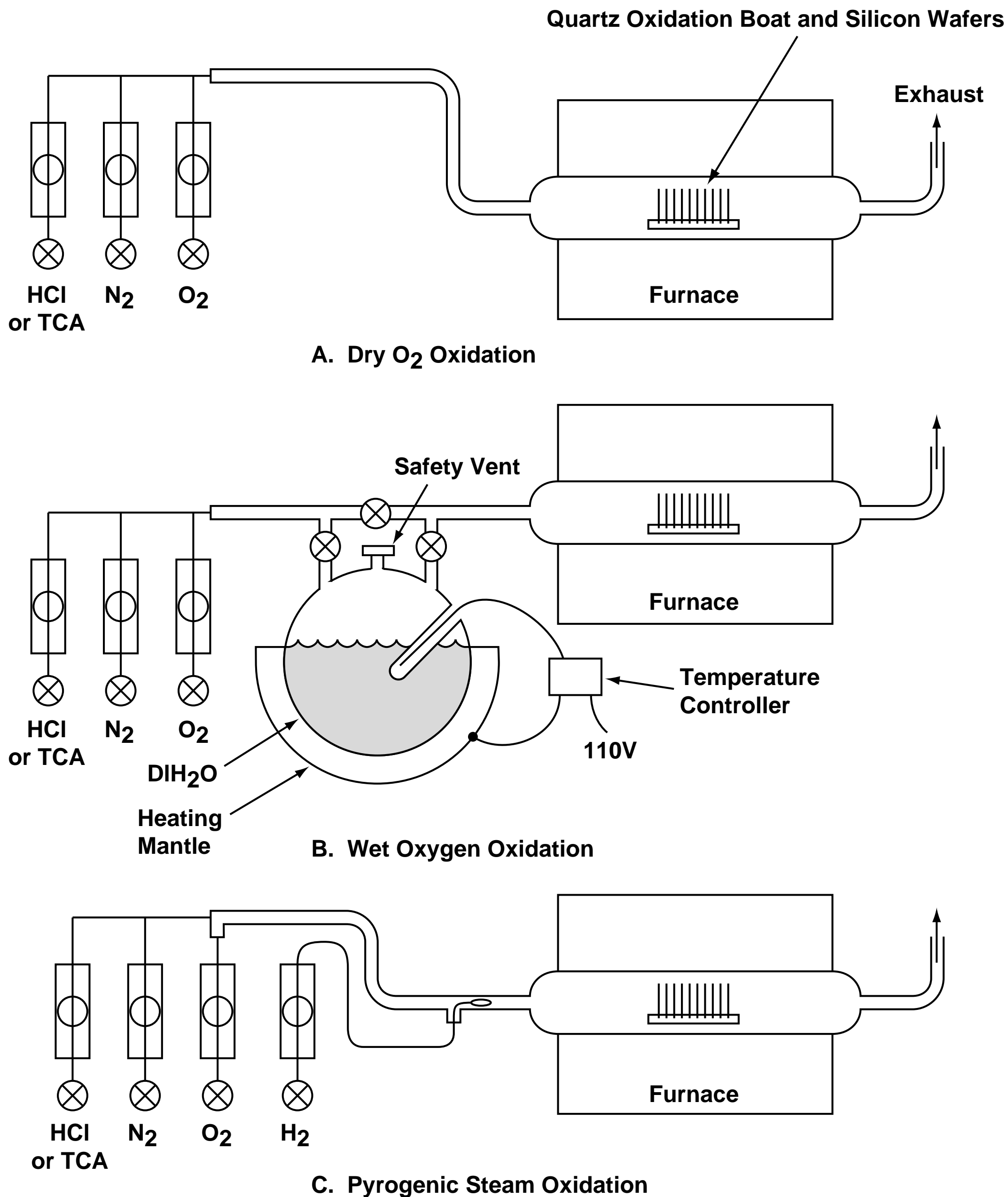


Figure 2-10. Oxidation Systems

b. High-Pressure Oxidation

As feature sizes of integrated circuits continue to decrease further into the submicron size, more shallow junctions are required and process temperatures are therefore being forced to lower levels. This has brought about a renewed use of high-pressure oxidation systems. These systems originated in about 1970, but have been slow to evolve into manufacturing.

The process chamber on this system is fitted with a High-Pressure Vessel. This allows the system to be pressurized from one to twenty-five (25) atmospheres (14.7 psi to 367.5 psi). This allows the process engineer to lower temperature by raising the pressure of the process chamber and has allowed some new degrees of process choices. In addition, this equipment has shown a reduction in oxidation defects when compared to the atmospheric-pressure process.

The disadvantages of high-pressure oxidation are the higher cost of the equipment, the equipment uses twice the floor space, and the safety considerations are stringent.

The advantages today more than offset the disadvantages because of the need for lower temperature processing.

2. CVD Silicon Dioxide

Silicon dioxide can be formed by manufacturing techniques other than thermal oxidation. One is known as Chemical Vapor Deposition (CVD). This process is done within a confined volume or separate process chamber. The process reactions bring about a chemical decomposition of certain elements or compounds by using heat, light, pressure/vacuum, and/or plasma energy to form a stable solid.

There is a clear distinction to be made between thermal oxidation to form silicon dioxide and CVD depositions to form silicon dioxide. Thermal oxidation reacts some form of oxygen with the silicon wafer at elevated temperatures. Thus, the silicon wafer is the source of silicon. When silicon dioxide is formed from a CVD process, both the silicon and the oxygen are brought to the process chamber from external sources. Thus the silicon wafer *is not* part of the chemical reaction. The silicon wafer is coated with the results of the CVD chemical reaction, silicon dioxide. Because of this difference, each type of silicon dioxide, thermal and deposited, has a unique set of physical characteristics. The characteristics determine which type of oxide is used at a given manufacturing step.

There are many silicon compounds available for use in the CVD process. However, silane (SiH_4) and dichlorosilane (SiH_2Cl_2) are the more commonly used materials. Oxygen for CVD can be pure dry oxygen, or can come from decomposing some compound of oxygen. Commonly used compounds are carbon dioxide (CO_2), nitrous oxide (N_20), and nitrogen dioxide (NO_2).

a. Uses of CVD-Deposited Silicon Dioxide

CVD-deposited oxides are used at several different places in the manufacturing sequence. The following lists some of the more common uses.

1. To increase oxide thickness of thermal oxides
2. Capacitor dielectric
3. Dielectric over polysilicon

4. Dielectric over metal
5. Buffer oxide layer to match mechanical requirements
6. Masking oxide layer
7. Final passivation

Deposited oxide is not normally used as a gate oxide for MOS transistors. Gate oxide is formed by some thermal oxidation technique.

b. CVD Equipment

The physical hardware for CVD-deposited oxide can be configured many different ways. CVD reactors can be classified by energy source, pressure, and chamber design. This is illustrated in Figure 2-11. The primary purpose of this process is to deposit a uniform film of silicon dioxide that has a certain composition. The LPCVD (Low-Pressure CVD) option is becoming the hardware of choice to meet this requirement. The system is shown in Figure 2-12.

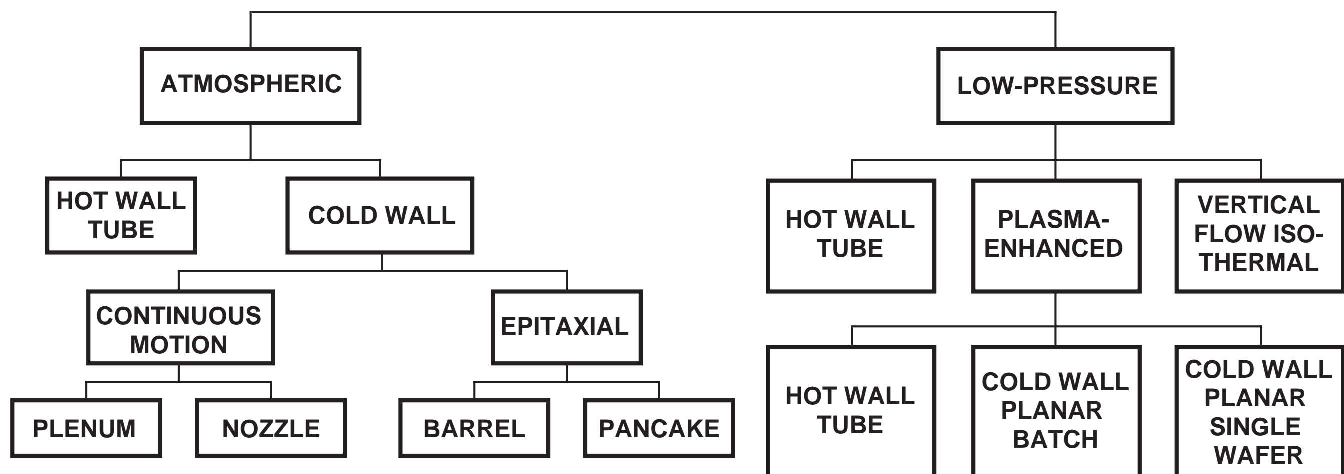
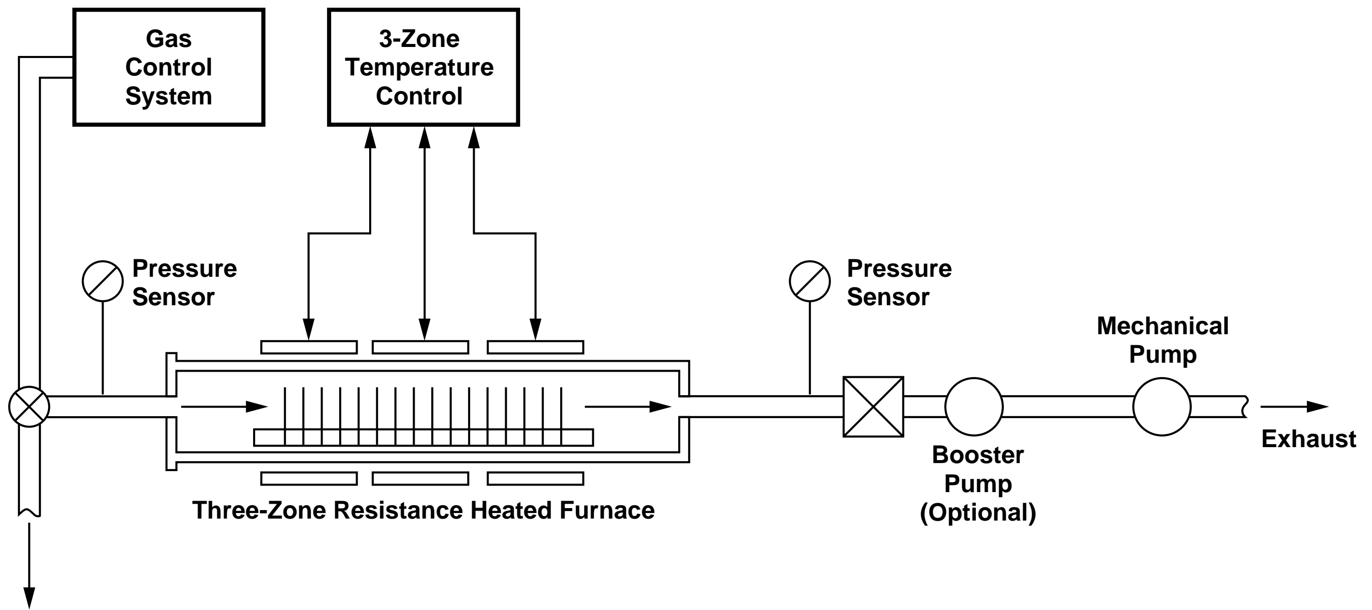


Figure 2-11. CVD Reactor Types

It has been found useful for some deposited silicon-dioxide films to have other elements added during the deposition. Two of the more common elements added are phosphorus and boron. These elements can be added individually to the silicon dioxide and sometimes they are both included in the silicon-dioxide film at the same time during the deposition process.

In addition to silicon-dioxide films deposited by CVD technology, silicon nitride (Si_3N_4), silicon-oxynitride ($Si_xO_yN_z$), and various forms of polysilicon are deposited. Each of these films has certain chemistry considerations along with temperature and pressure requirements. Typical reactions for CVD are tabulated in Figure 2-13.



Source: ICE 1313D

Figure 2-12. Block Diagram of a Low-Pressure Chemical Vapor Deposition System

PRODUCT	REACTANTS	DEPOSITION TEMPERATURE (°C)
SILICON DIOXIDE	$SiH_4 + CO_2 + H_2$	850 – 950
	$SiCl_2H_2 + N_2O$	850 – 900
	$SiH_4 + N_2O$	750 – 850
	$SiH_4 + NO_2$	650 – 750
	$Si(OC_2H_5)_4$	650 – 750
	$SiH_4 + O_2$	400 – 450
SILICON NITRIDE	$SiH_4 + NH_3$	700 – 900
	$SiCl_2H_2 + NH_3$	650 – 750
PLASMA SILICON NITRIDE	$SiH_4 + NH_3$	200 – 350
	$SiH_4 + N_2$	200 – 350
PLASMA SILICON DIOXIDE	$SiH_4 + N_2O$	200 – 350
POLYSILICON	SiH_4	600 – 650

Source: VLSI Technology 14777

Figure 2-13. Typical Reactions for CVD Depositions

D. PHOTOLITHOGRAPHY

1. Overview

In the photolithography process sequence, the wafer is covered with a layer of light-sensitive material (photoresist), which is then selectively exposed to light. The selective exposure is accomplished by shining the light through a quartz plate (mask or reticle) with a patterned opaque material on it (Figure 2-14).

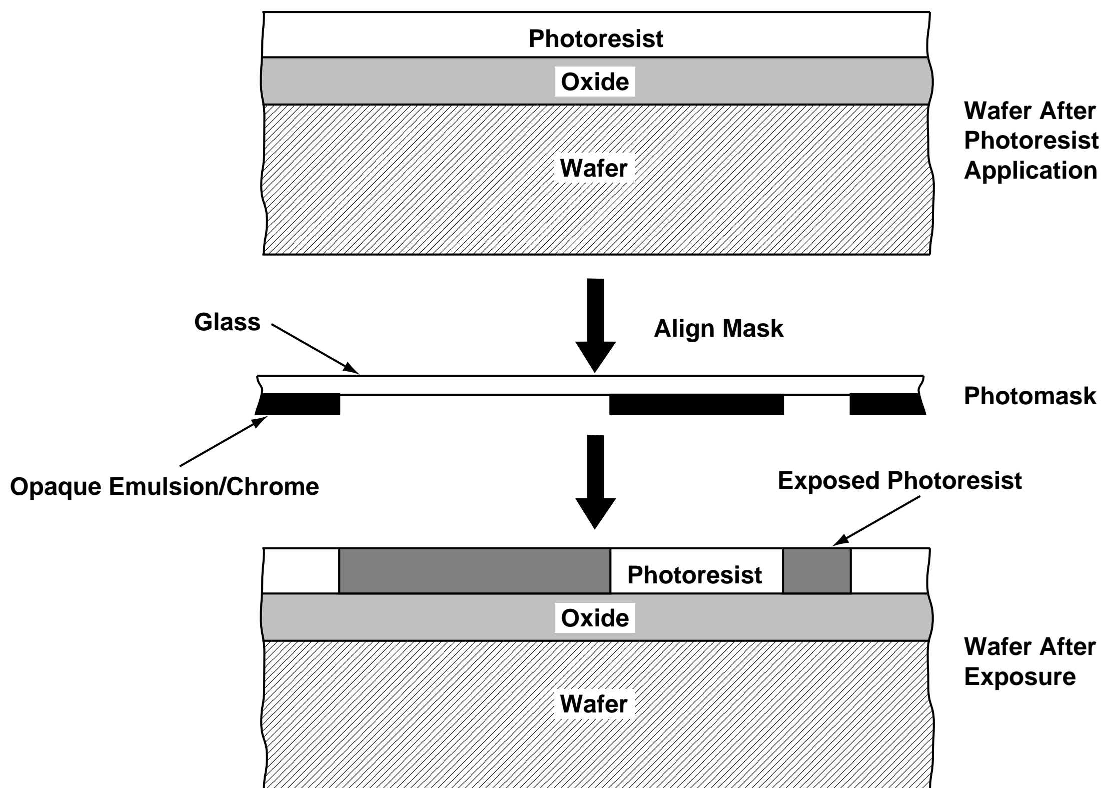


Figure 2-14. Alignment and Exposure

The exposed photoresist is washed away and the remaining, unexposed photoresist is hardened by baking. The portions of the layer below the photoresist not covered by the hardened photoresist is removed and then the photoresist is removed (Figure 2-15).

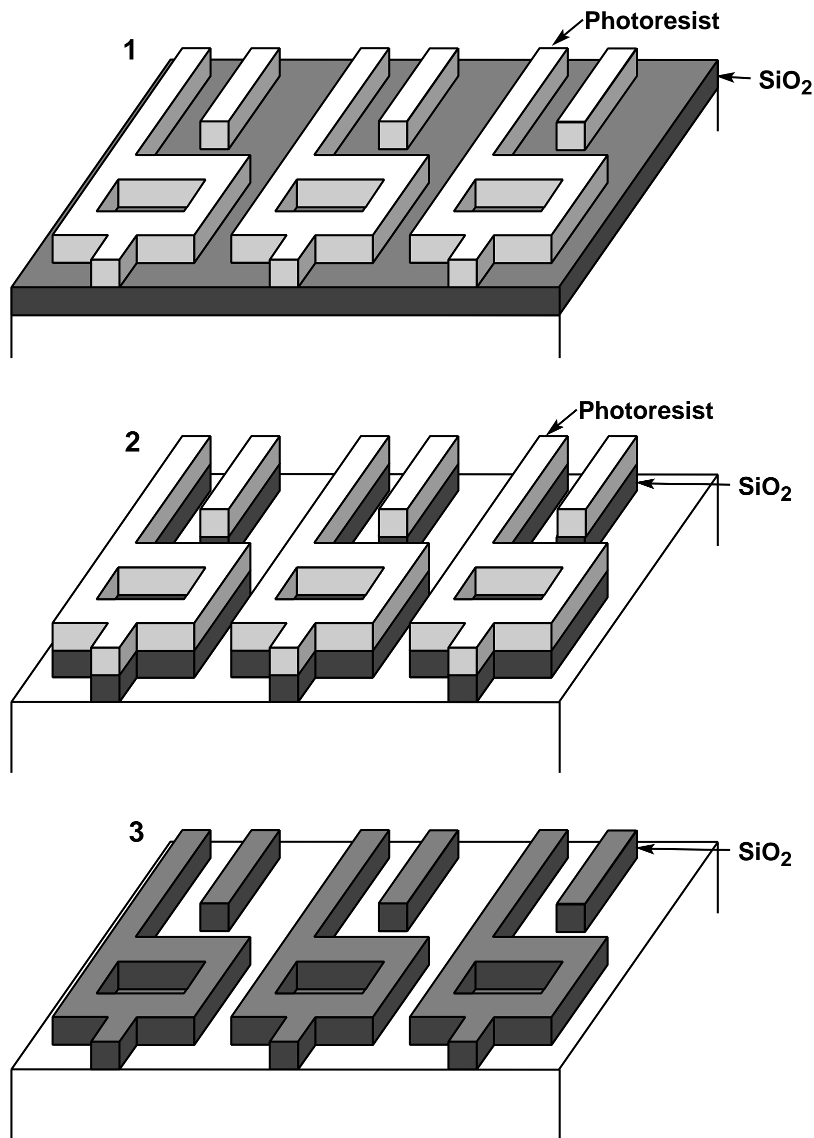


Figure 2-15. Photolithography Using Positive Photoresist

The following is an in-depth analysis of photolithography.

2. Introduction

The photolithography process encompasses all of the patterning operations necessary to transfer an image from one medium to another. The term, photolithography, implies the use of light as part of the transfer process. Frequently the term, microlithography, is used to mean the photolithography process for transferring feature sizes in the submicron range.

The photolithography process involves the transfer of geometric images created by a circuit designer to a photosensitive film applied to the surface of the silicon wafer. The sized geometrical features are the physical designs of the various circuit elements, i.e., transistors, resistors, capacitors, etc., that will make up the electrical circuit design.

Since the early 1970's the concept of making the physical size of the transistor smaller to increase the transistor performance has been the evolutionary force behind the semiconductor industry. This has placed a constant pressure on the photolithography process to print smaller feature sizes.

Figure 2-16 illustrates the critical aspects of photolithography. There are many patterns that must be transferred to the wafer, each sized properly and registered (aligned) to one another in the correct sequence. The transfer starts with the circuit designer creating the electrical circuit, aided by the power of the computer. After the electrical circuit is created, the circuit schematic is converted into the physical sizes and shapes of the circuit elements that are arranged (layout) in a surface plane and the electrical connections are made to each circuit element.

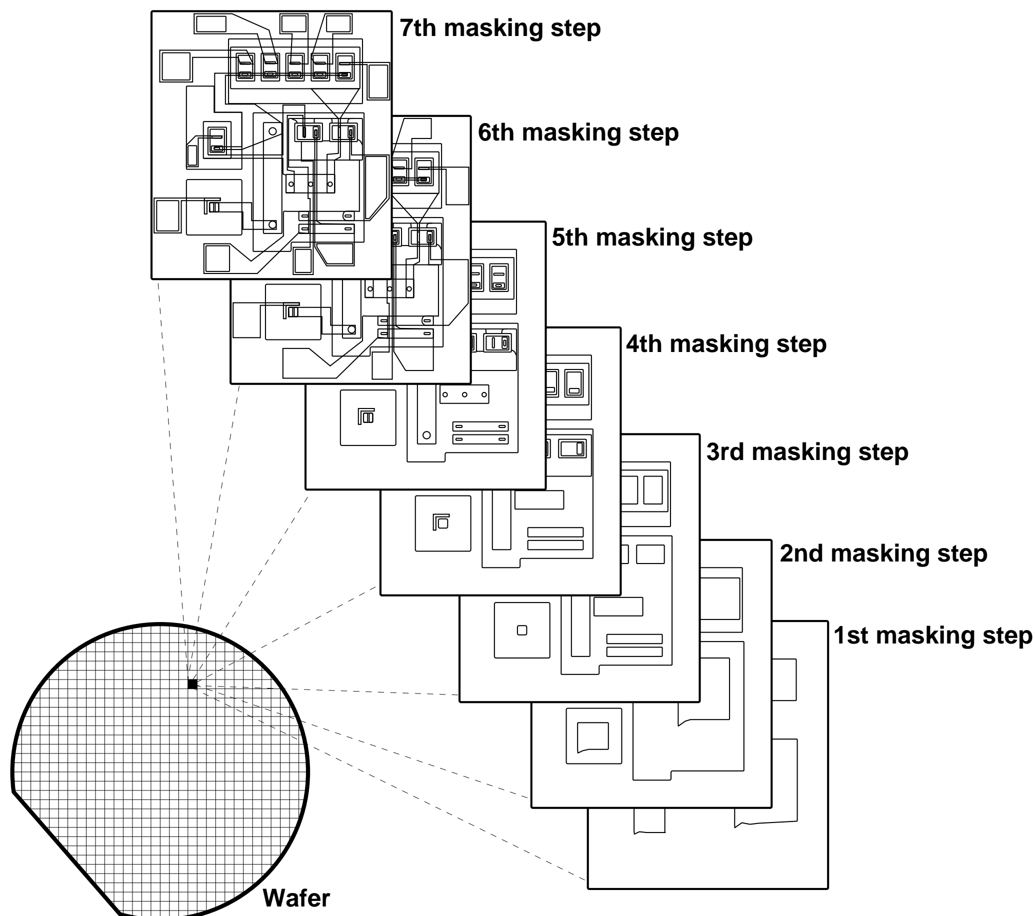


Figure 2-16. The Layers Transferred to a Wafer During a Seven-Mask Process

3. Masks

A photomask is a quartz plate with one layer of patterns for all of the ICs on a wafer on it (Figure 2-17). A reticle only has the patterns for a few ICs on it (Figure 2-18). The patterns are formed with opaque substances such as emulsion or chrome.

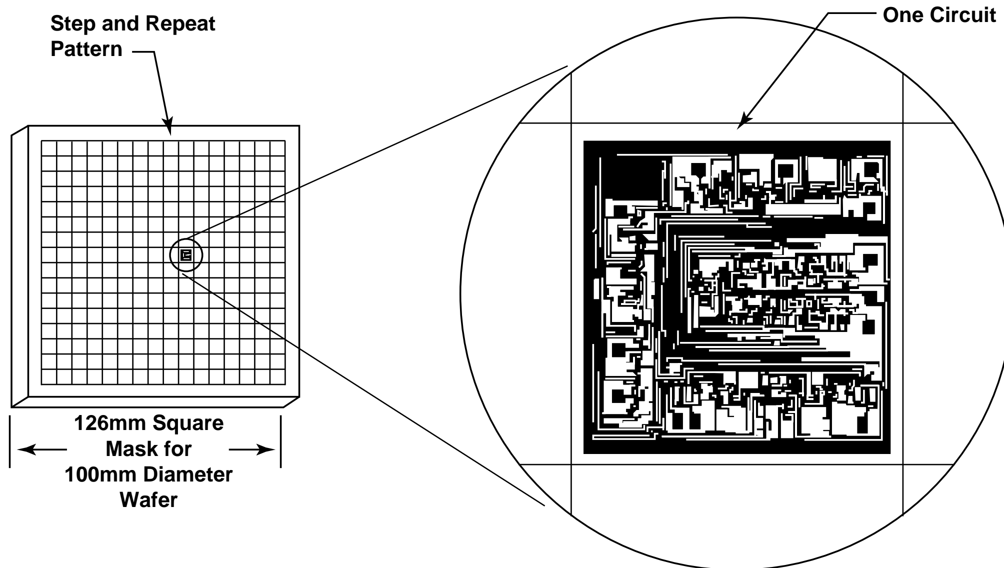


Figure 2-17. Example of a Photomask

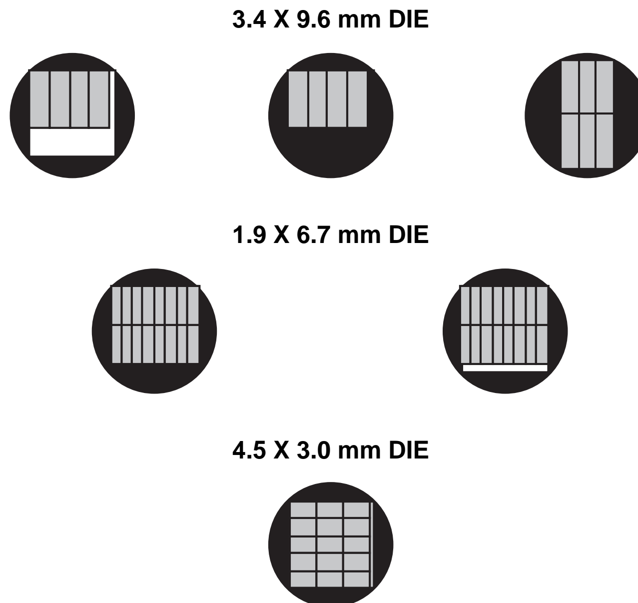


Figure 2-18. Die Placement in Fixed Reticle of 21mm

Emulsion images on the glass substrate are difficult to clean. Typically they are used for a limited number of exposures and then discarded. On the other hand, chrome images on the glass substrate can be cleaned, inspected, and reused many times.

a. Reticles

The design process is illustrated in Figure 2-19. The information generated from the design process is used to manufacture a reticle for each layer of the circuit. The reticle is an intermediate-sized representation of the circuit used in the image transfer process. The reticle can be sized as large as ten times (10X) the actual circuit size down to the actual circuit size (1X).

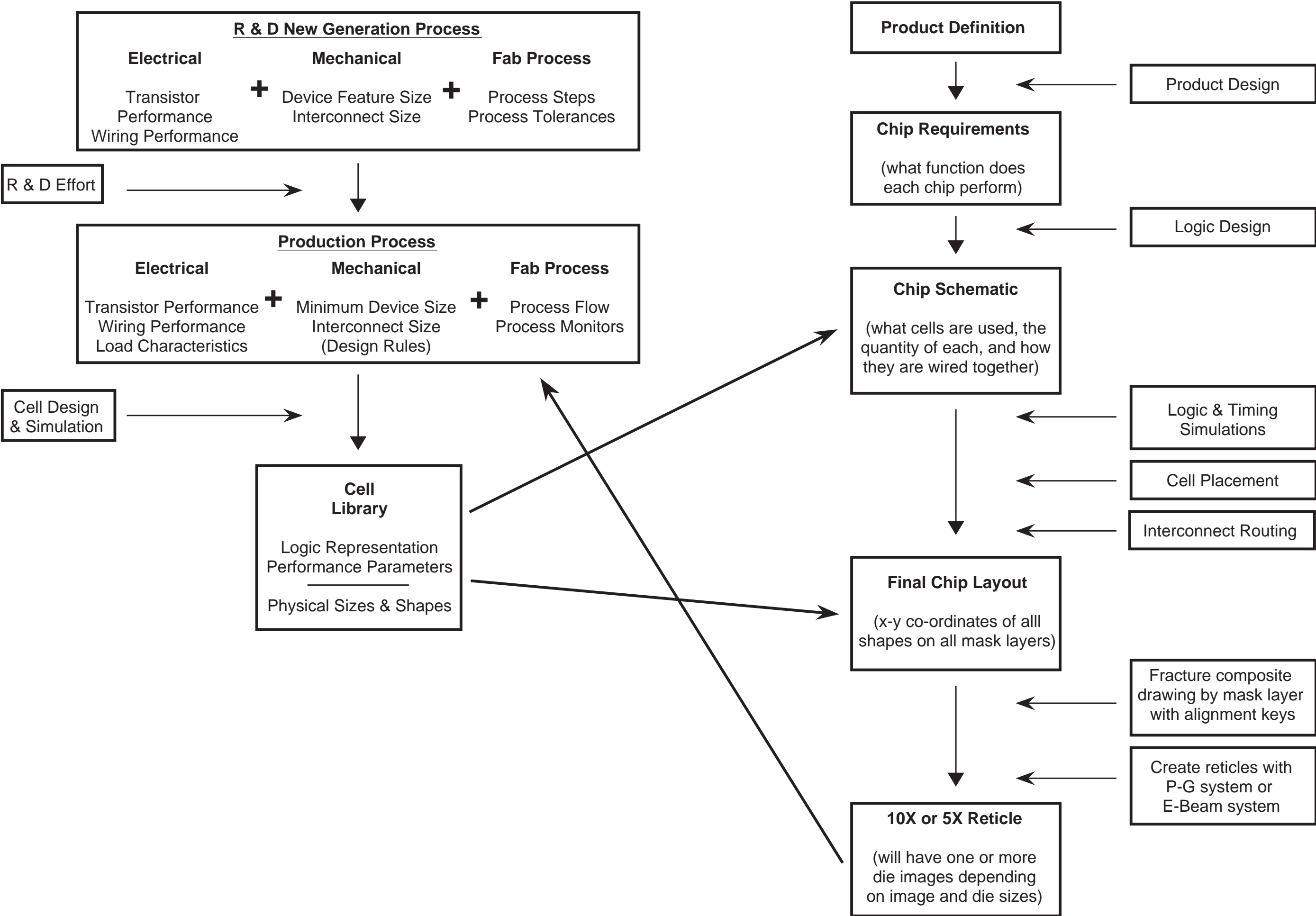


Figure 2-19. The Design Process

The manufacturing procedures and equipment used for reticle generation require the best precision and reproducible imaging technology known. The quality of the reticle can impact the yields of the actual IC manufacturing.

The manufacturing of the reticle can follow a variety of paths. The alternative routes for reticles' manufacturing are shown in Figure 2-20.

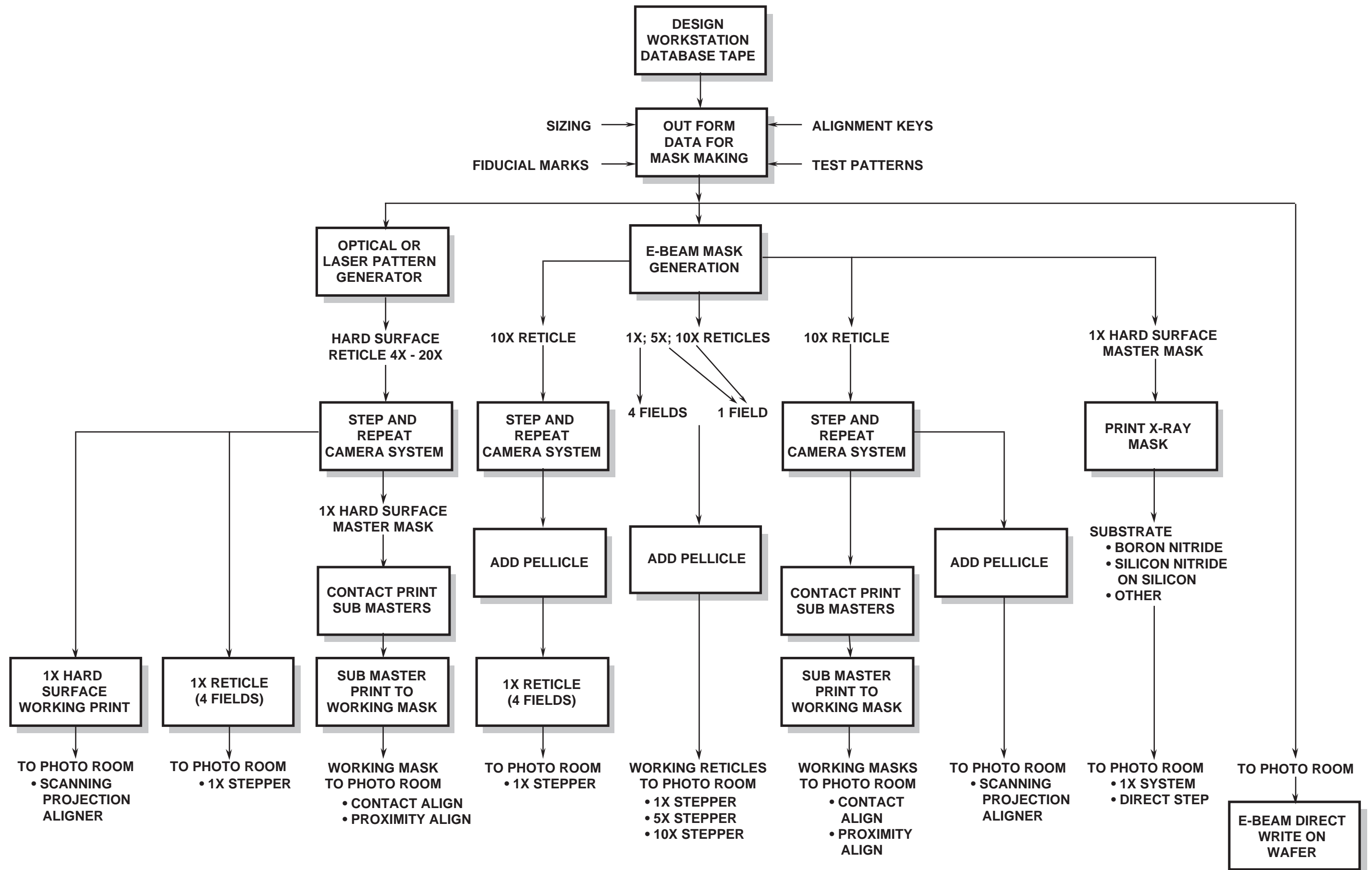


Figure 2-20. Microlithography Roadmap

Reticles for Direct Step on Wafer are made in 10X, 5X, 4X, or 1X size, depending on the type of stepper equipment chosen. A reticle's size is determined by the magnification of the stepper lens and usable field size of the stepper lens. This will determine how many dice can fit within the usable field area. Figure 2-18 illustrates different die sizes fitting into a 21mm field size.

b. Photomasks

The photomask is created by a series of additional process steps beyond the making of a 10X reticle for each layer. Refer to Figure 2-20. At the contact print to submaster stage the mask polarity must be determined (depends on whether negative or positive photoresist is used). The submaster to working print stage can print either emulsion or chrome on glass substrates. The choice is determined by the wafer alignment equipment. A photomask is shown in Figure 2-17.

c. E-beam

Also shown in Figure 2-20 are two other alternative technologies to form the designer's requirements into a photoresist layer on the silicon wafer. One alternative is to use an e-beam system to write the pattern directly on the photoresist-coated wafer. This approach eliminates the reticles or mask-making stage but has a very slow throughput.

d. X-ray

The other technique uses x-ray energy to expose the photoresist-coated wafer. The mask for this technique is very expensive and difficult to manufacture. However, x-ray steppers have excellent resolution capability and will probably be needed by the end of the 1990's.

4. Photolithography Sequence

Figure 2-21 illustrates the manufacturing sequence for photolithography. These photolithographic steps encompass all of the patterning operations, including wafer priming, coating, align, expose, develop, etch, and photoresist removal. The characteristics of the light-sensitive photoresists determine the basic process technique. As was shown in Figure 2-16, the photolithographic process is repeated several times. Because of this it is often referred to as the hub of the IC fabrication process.

a. Photoresist, Negative and Positive

There are two kinds of photoresist commonly used: negative and positive. The chemical behavior of each resist is illustrated in Figure 2-22. The negative resist responds to the radiation (UV light) in a manner that prevents the developer solution from removing the exposed resist. The image formed in the resist is the same as the clear area on the mask. The unexposed resist is removed by the developing process. Positive photoresist has the opposite response to the radiation. The areas of the photoresist that are exposed are removed by the developer solution. Thus, the unexposed resist remains and forms the image on the surface of the wafer.

The chemistries of positive and negative photoresist are very different. Positive photoresist is developed with a mild alkaline (basic) solution and the negative resist requires a solvent (xylene) for developing.

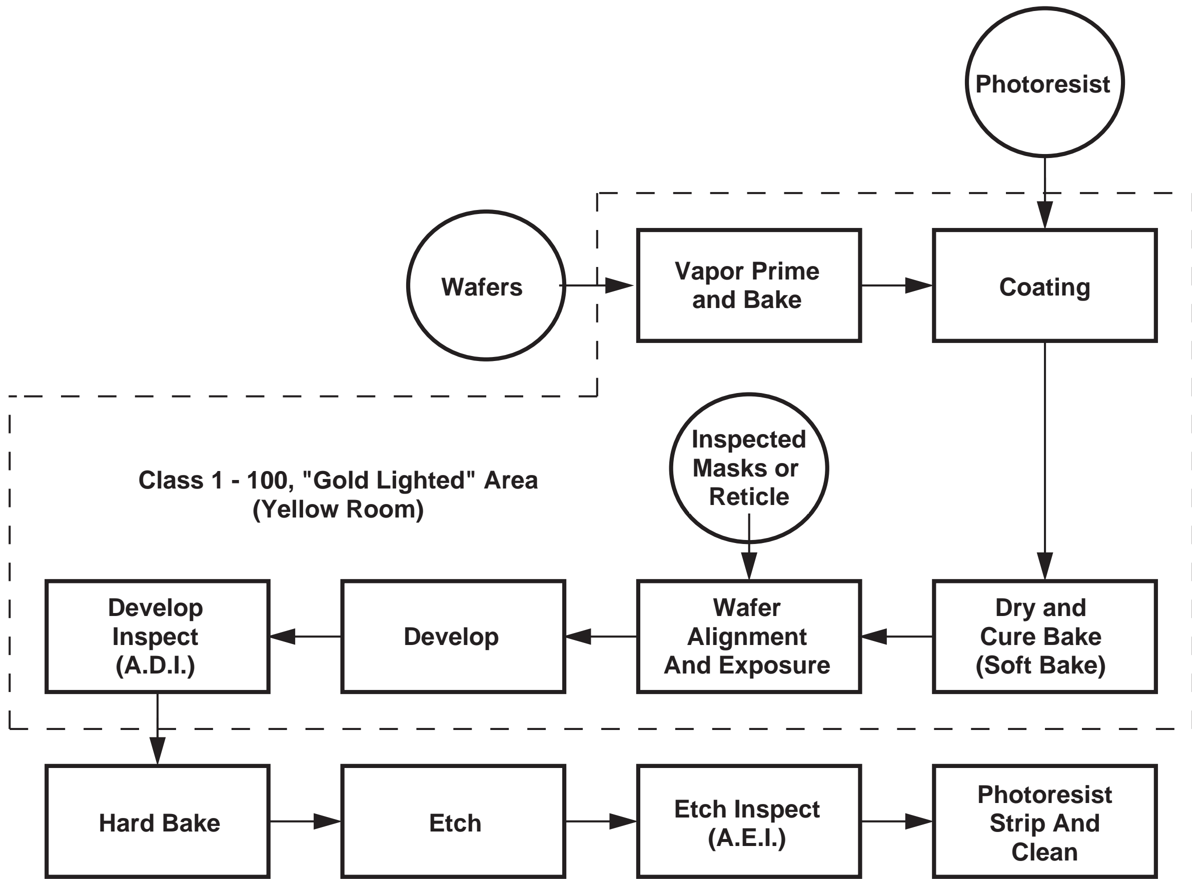
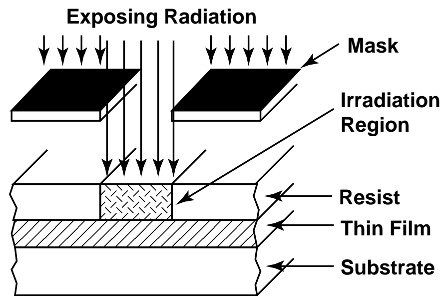


Figure 2-21. Photolithography Process Flow Chart

i. Exposure Wavelengths

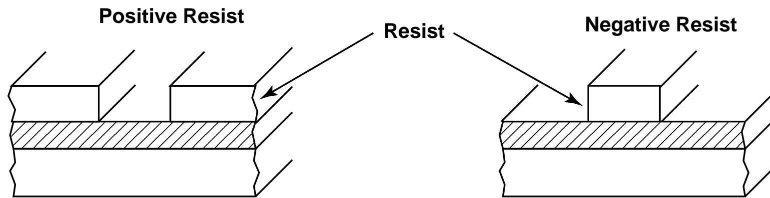
A portion of the electromagnetic spectrum is illustrated in Figure 2-23. The blue-violet region is referred to as the ultraviolet wavelengths. Also shown is the relationship between wavelength (l) and the frequency (f) for this form of energy.

The light-sensitive responses of both resists are similar in that both are exposed by light in the blue-violet wavelength (190 - 450nm). These wavelengths are commonly found in mercury arc lamps and similar bulbs. For this reason, the process area for photoresist must have these wavelengths filtered out during manufacturing to prevent unwanted exposure. Yellow filters or lights are used to illuminate the work area for these processing steps.



Actinic: **The property of radiant energy, especially in the visible and ultra-violet spectral regions, by which chemical changes are produced.**

Developing



Etching and Stripping

Source: ICE

7417A

Figure 2-22. Characteristic of Negative and Positive Resists

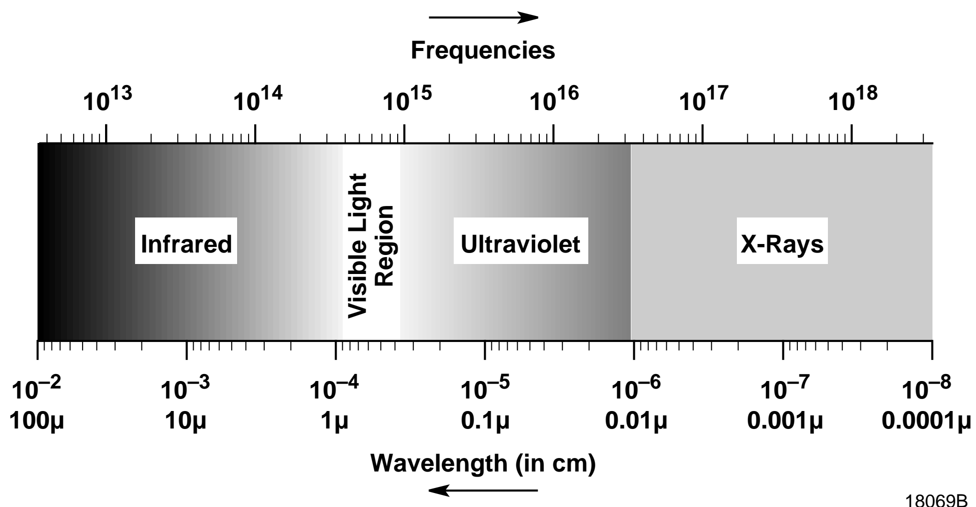


Figure 2-23. Part of the Electromagnetic Spectrum

ii. Photoresist Parameters

Photoresist is typically characterized by several parameters that affect its performance. These are:

1. adhesion
2. etch resistance
3. resolution
4. photo sensitivity
5. step coverage

b. Wafer Preparation Before Photoresist Application

i. Cleaning

The photoresist process is extremely sensitive to any form of contamination on the surface of the wafer prior to applying the photoresist. To assure the best surface possible, various cleaning techniques are often used prior to the priming process. The cleaning process may be some type of wafer scrubbing (brush or high pressure), chemical cleaning with agitation, or chemical cleaning only.

ii. Priming

The surface of the silicon substrate can have various layers depending on where in the manufacturing cycle the current operation resides. Each layer on the silicon surface has a different effect on photoresist adhesion. To provide better adhesion, a process step called priming is performed. The use of a priming solution increases the adhesion of the photoresist to the surface. Primers may be applied by immersing the substrates in the priming solution, spraying the solution on, or by passing a priming vapor over the surface of the wafer. Some primers have to be baked before subsequently coating the wafers with photoresist. Other priming techniques require that the wafer surface be dehydrated at elevated temperatures prior to applying the priming solution.

c. Photoresist Application

Photoresist may be applied to the surface of the wafer using a variety of techniques. These techniques include dipping, spraying, brushing, roller coating, and spin coating. Spin coating is the method most often used in the fabrication of semiconductors.

The spin coating process applies a resist layer to the silicon substrate as uniform as possible for the required thickness. The uniformity of the coating is very important. As the wafers progress further through the manufacturing cycle, the topology of the surface continues to change in thickness above the silicon surface. The variation in the vertical heights causes some variation in the resist thickness uniformity.

The coating process involves holding the wafer on the motor spindle by vacuum. The resist is dispensed onto the wafer surface. The wafer is slowly spun at a low RPM to spread the resist over the entire wafer surface. After a few seconds of low RPM, the spindle will rapidly accelerate to a much higher RPM (4,000 to 6,000) for final thickness control. The wafer is decelerated back to a low RPM wherein a resist thinner solution is dispensed around the outer 2mm of the edge of the wafer to remove the bead of resist that builds up from the spinning process. Removing this bead eliminates a source of contamination from the process.

The wafer is transferred to a soft bake heat source to remove the solvent from the resist. This leaves a firm resist layer ready for the align and expose process. A cross section of the structure is illustrated in Figure 2-24.

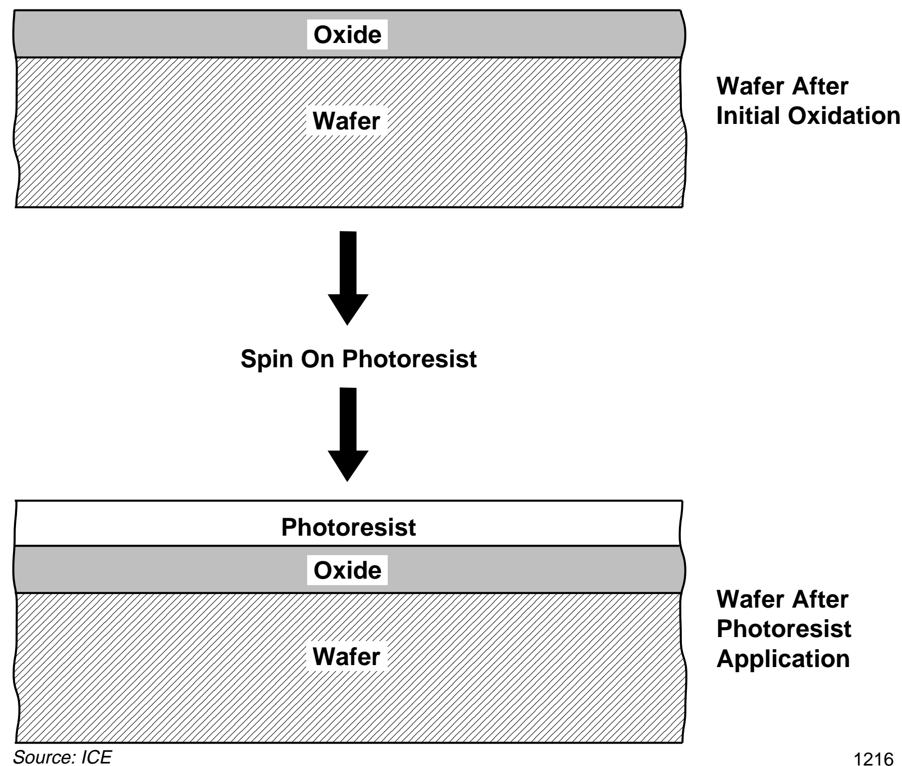


Figure 2-24. Photoresist Application

d. Alignment/Exposure

The process of forming an image in the photoresist-coated surface has undergone considerable change. Early wafer align/expose systems typically worked on the principle of "global" or "blanket" exposure, in which the entire wafer was exposed at once. As resolution requirements kept pushing the feature size smaller and smaller and overlay registration (lining up one layer to preceding layers) had to become more accurate, alternatives were investigated.

The align/expose process was illustrated in Figure 2-14. There are several types of align and expose equipment (Figure 2-25). The oldest alignment system is the contact aligner and is normally used in older fabs manufacturing older products. The contact printer concept is illustrated in Figure 2-26.

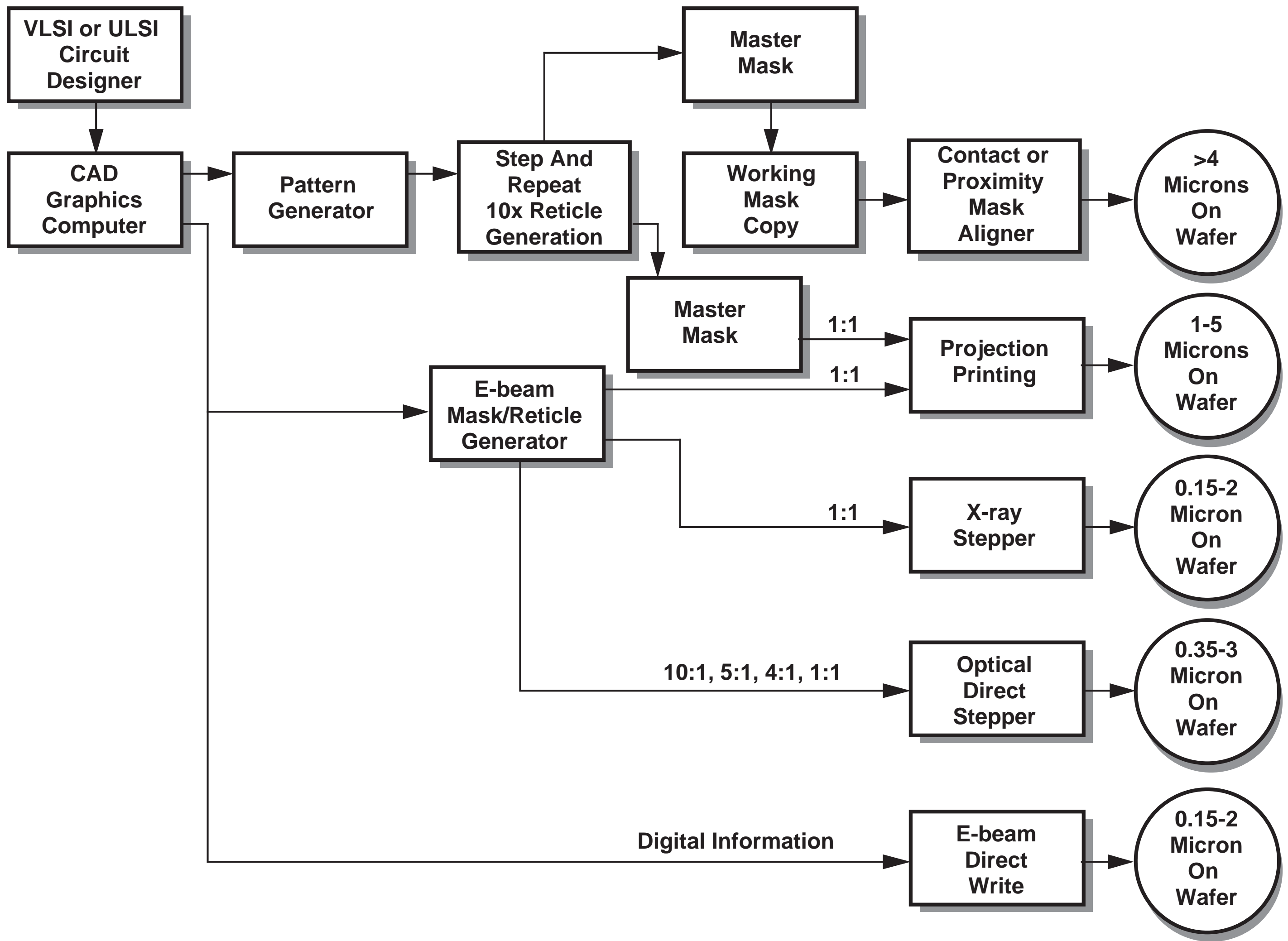


Figure 2-25. Wafer Patterning Systems

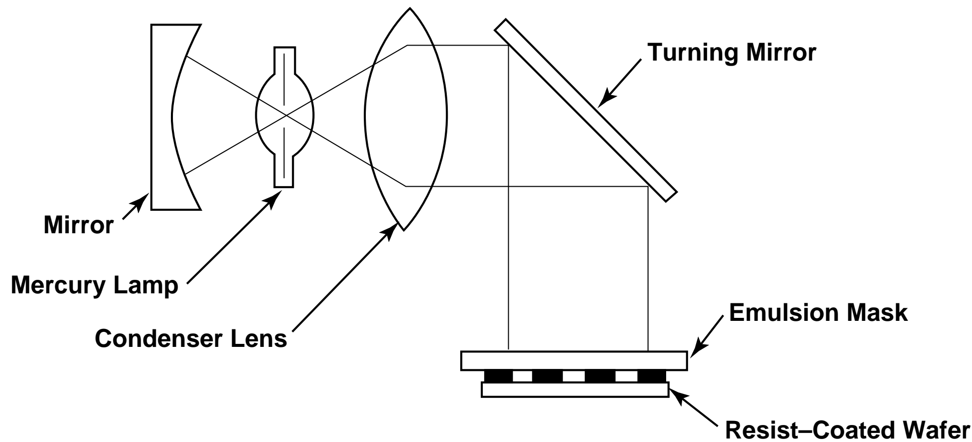


Figure 2-26. Contact Printer Optical Configuration

i. Contact Aligner

The contact aligner can achieve very good resolution but has poor overlay registration and a high defect density. The high defect density is the result of the mask and wafer making contact during the exposure. This contact between the mask and wafer can create physical defects in both the mask and wafer. As the mask is used to align and expose additional wafers, the defect density will continue to increase. The detrimental effect is shown in Figure 2-27.

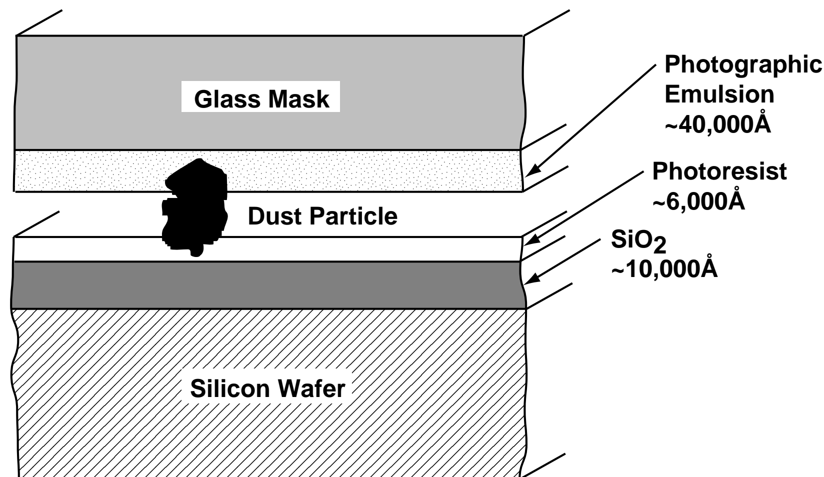


Figure 2-27. Dust Particle Scratching Photoresist and Photomask

ii. Proximity Printing

An early alternative to contact printing was proximity printing. The light source in this equipment is colliminated enough to allow a 10 to 20µm separation between the mask and the photoresist surface. This equipment concept is illustrated in Figure 2-28. An improvement in defect density was realized but the resolution was typically limited to 4µm or greater.

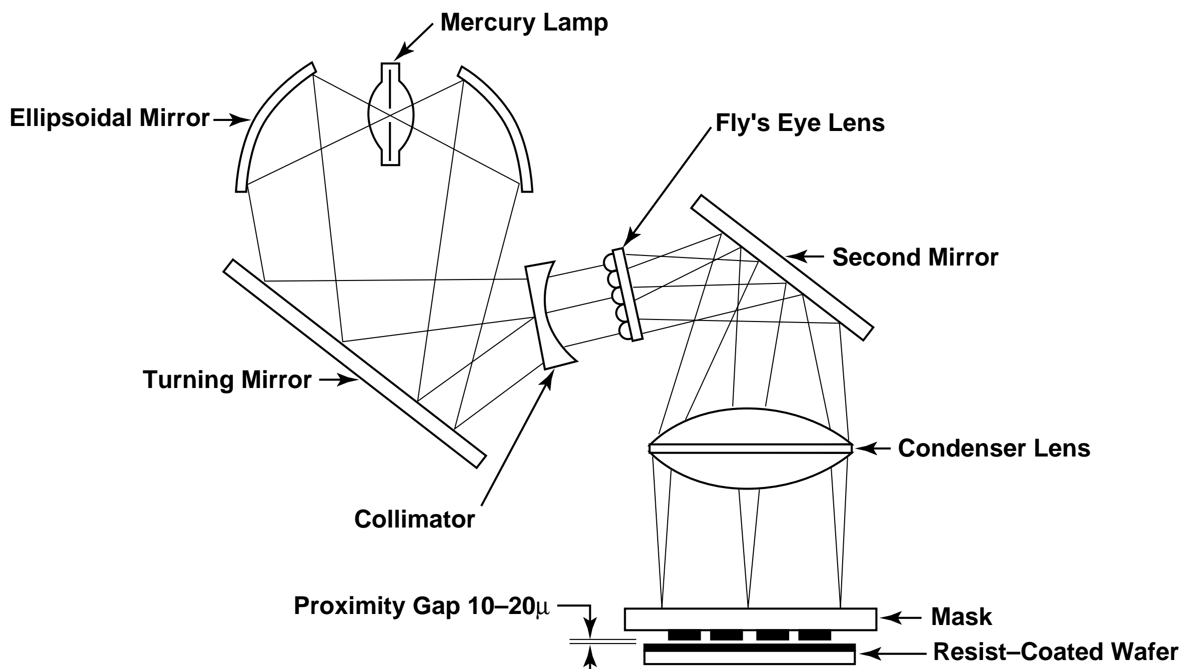


Figure 2-28. Proximity Printer Optical Configuration

iii. Projection Alignment

The next approach to improving the align/expose process was the introduction of the projection aligner. In this system there is no contact between the wafer and the mask. Rather than blanket exposure, where the entire mask is exposed with a flood of light, the exposure is accomplished with an arc of light. The wafer and the mask move through the equipment coincidentally on parallel planes. The system is illustrated in Figure 2-29.

The projection aligner needs better temperature and humidity control, a cleaner cleanroom, and reduced vibration in the photo room. The aligner also places more stringent requirements on the mask and photo process to achieve an improved defect level.

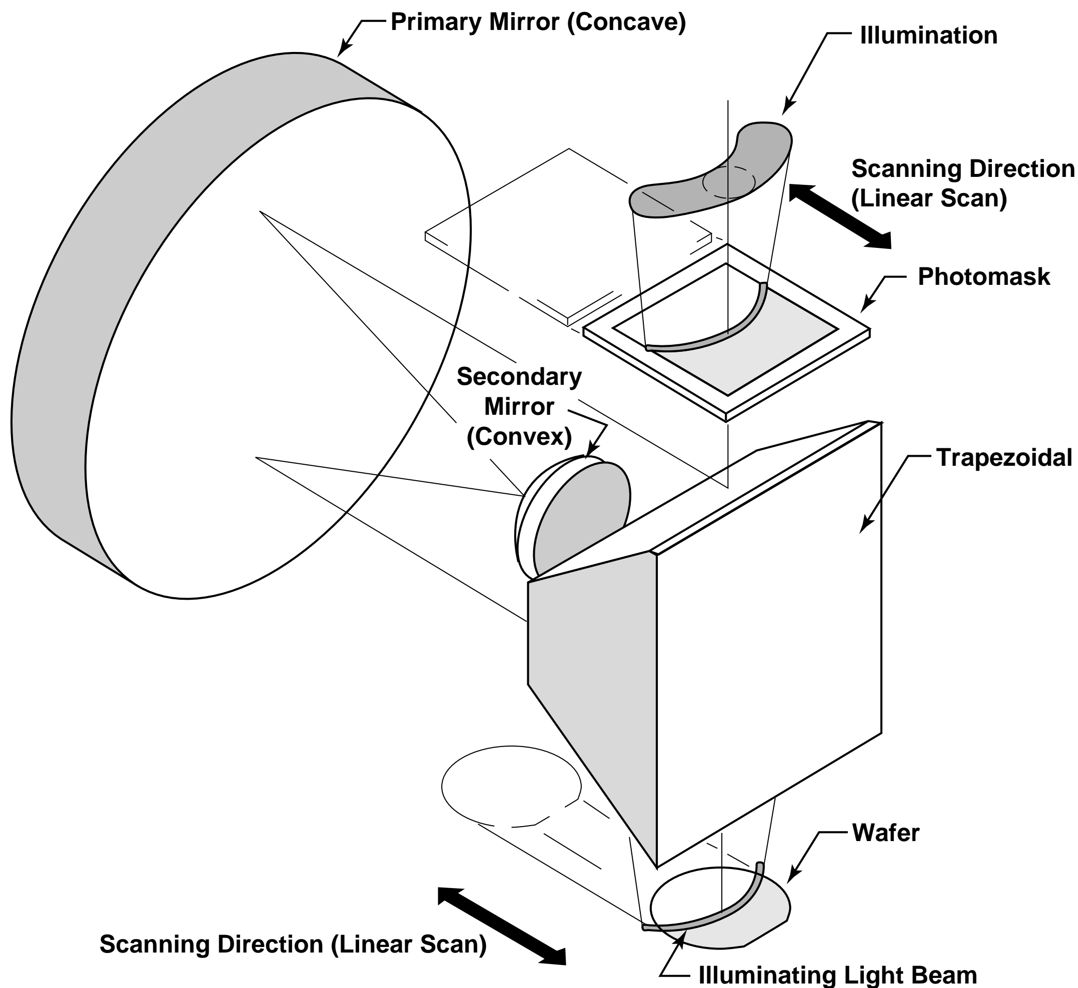


Figure 2-29. Principle of Scanning Projection Aligner

Over several generations of equipment, the projection aligner has become one of the most widely used alignment systems in IC manufacturing. The equipment capability has been extended to slightly below one micron.

With the concept of "mix and match" alignment systems, the projection aligner will continue to be used for several more years as long as the overlay registration can keep pace with the circuit designers' requirements.

iv. Direct Step on Wafer

As feature size progressed below the three-micron size, another align/expose system was introduced to IC manufacturing. This system is known as the Direct Step on Wafer (DSW or Stepper).

When originally introduced, the system used the 10X reticle for the masking process. This eliminated the requirement for the global mask.

The DSW system ushered in a new technology that offered exciting new possibilities for both resolution and overlay registration. Taking advantage of the precision developed for mask making and with the working image confined to a small place in the center of lens, optical distortion can be minimized. This tremendous gain in optical performance came at the expense of lower throughput. The DSW is illustrated in Figure 2-30.

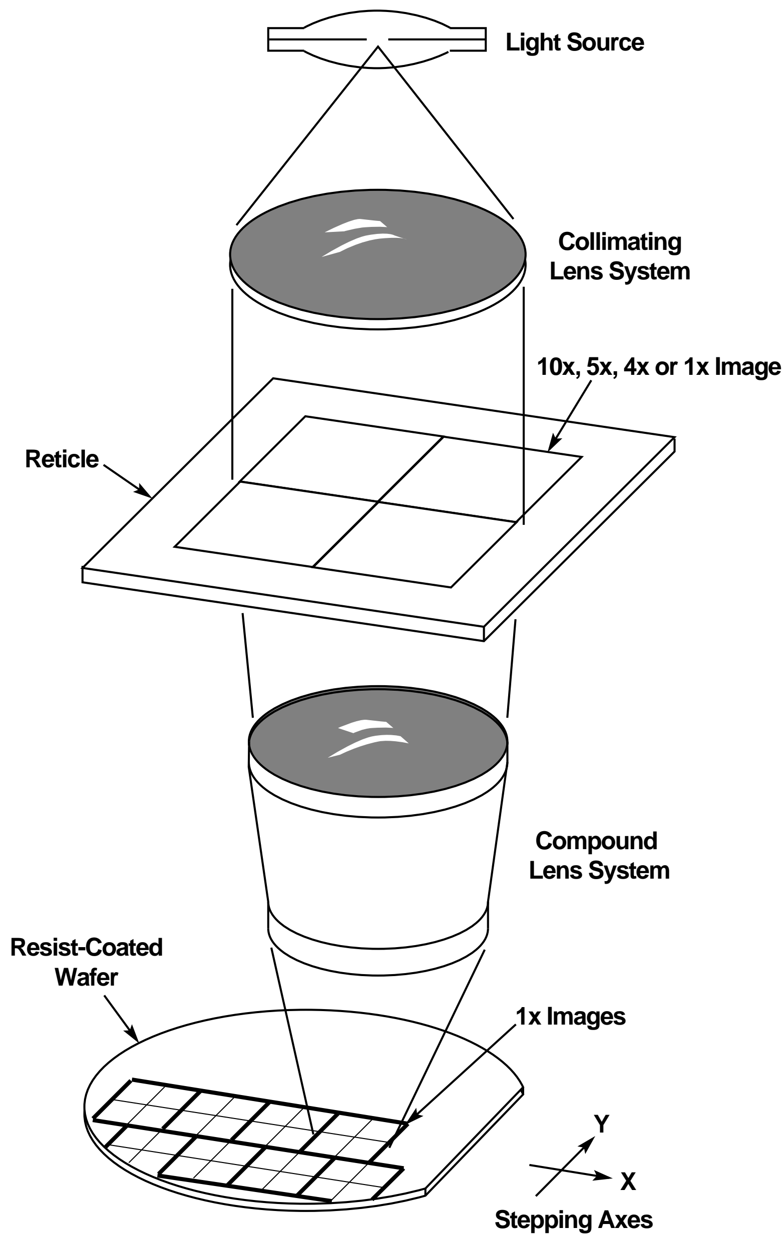


Figure 2-30. Direct Step on Wafer (DSW) Aligner Optical Configuration

- **Mix and Match**

DSW has rapidly gained acceptance but the disadvantage of lower production rates has required larger capital investments. To balance between these two conflicting requirements, the concept of "mix and match" evolved as a reasonable compromise.

In mix-and-match lithography, the scanning projection aligner is used for the less critical levels, and the DSW systems are designated for the more critical mask layers. The net result is better overall yield and a more cost effective capital investment.

To further address the throughput issue, stepper technology found alternative solutions to throughput by reducing the lens magnification. This led to the 5X, 4X and 1X DSW technology. Each of these magnifications has its own compromise but at a slightly improved throughput rate as they decline from 10 to 1.

The issue of lens size vs. die size was illustrated in Figure 2-18. The direct step sequence is illustrated in Figure 2-31 and a comparison of stepper reticle sizes is illustrated in Figure 2-32.

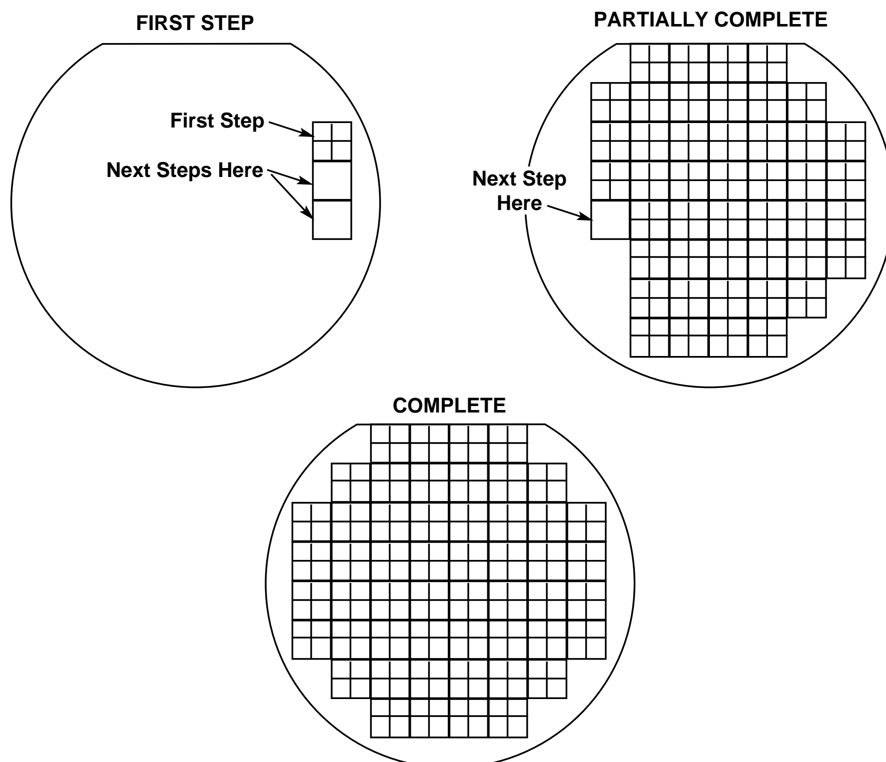


Figure 2-31. Direct Step on Wafer (DSW)

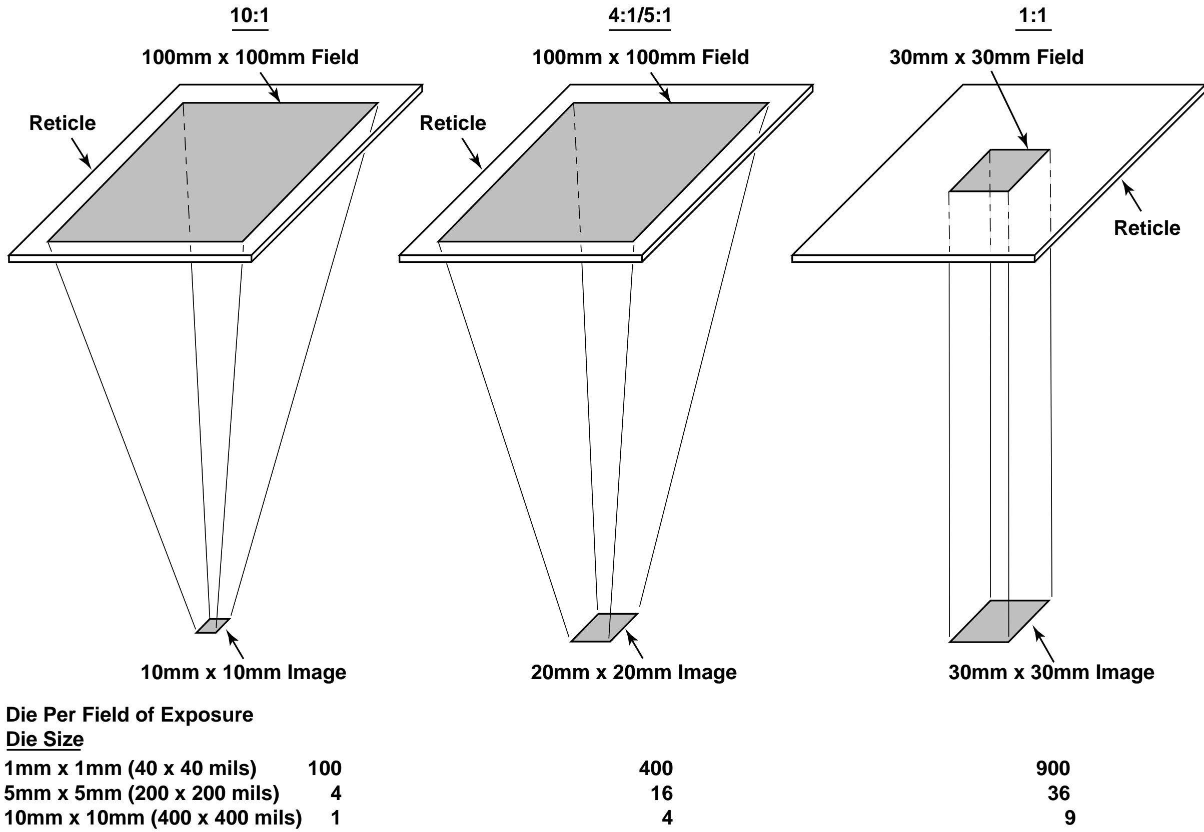


Figure 2-32. Comparison of Stepper Reticle Sizes

- **Pellicles**

The cleanliness of the wafer fabrication area has a significant impact on the defect density of the align/expose process. To improve the defect density of this operation, the reticle is commonly placed in a fixture that has a thin membrane, called a pellicle, stretched over the surface of the fixture. The pellicle then serves as a protective cover over the reticle and de-focuses particles. This prevents the particles from printing. The pellicle is illustrated in Figure 2-33.

The performance of the lens system of DSW equipment continues to improve and the photoresist chemistries continue to be improved to respond to shorter wavelengths of light. The collective result is the capability of manufacturing products, particularly DRAMs, with feature sizes less than 0.5um. As research continues, it may be possible for an optical system to extend to 0.15µm features. If this becomes a reality toward the latter part of this decade, then implementation of e-beam and x-ray will be postponed.

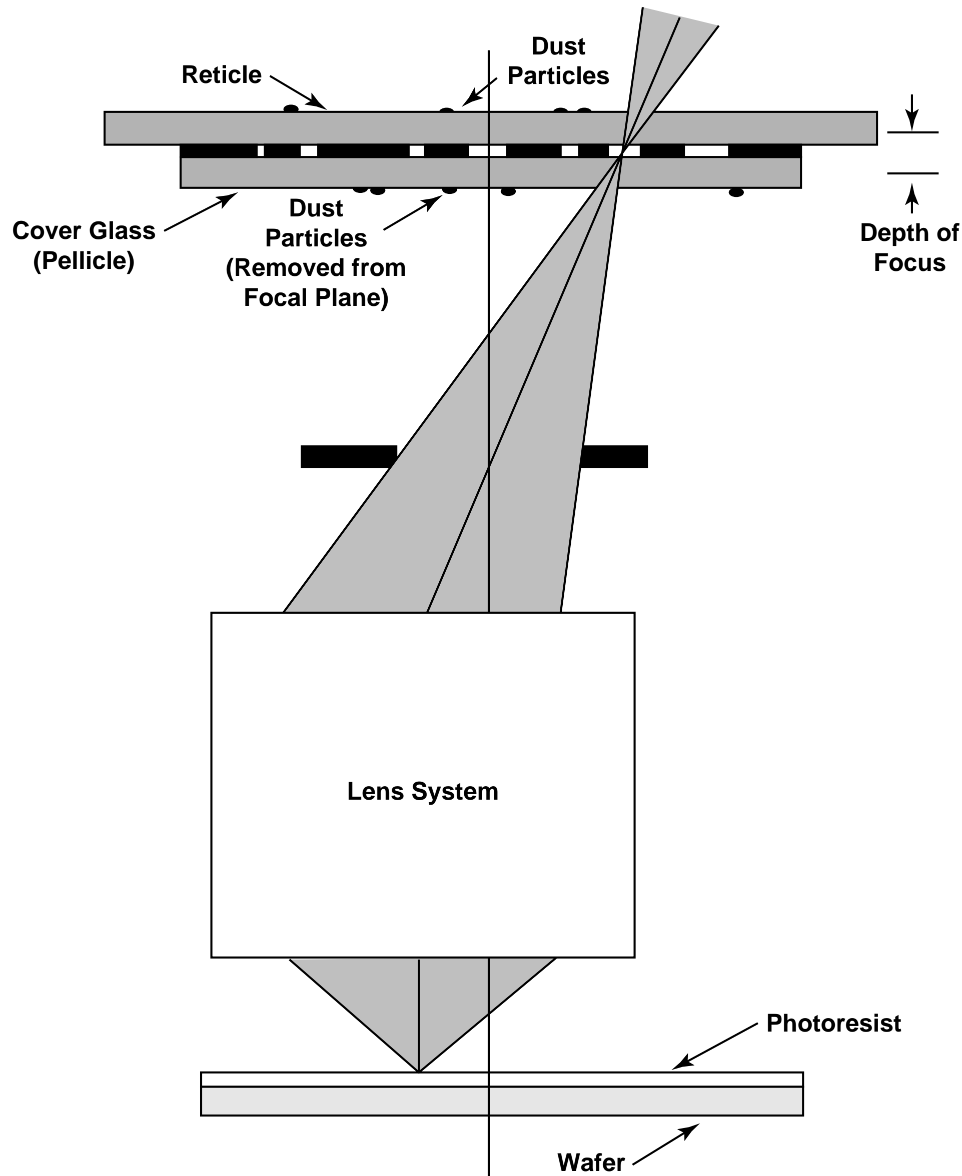


Figure 2-33. Pellicle Protection Mechanism

v. E-beam

As previously shown in Figure 2-20, the electron-beam (e-beam) system has dominated reticle and mask manufacturing. However, direct electron-beam exposure of the resist-coated silicon wafer without a mask or reticle has been very slow to gain production use. The main problems associated with using e-beam wafer exposure systems are related to the poor sensitivity of the photoresists available, the beam current, the electron-electron interaction at the resist-substrate interface, the rate the beam can be moved and the relative high system cost per wafer exposed. The e-beam concept is illustrated in Figure 2-34.

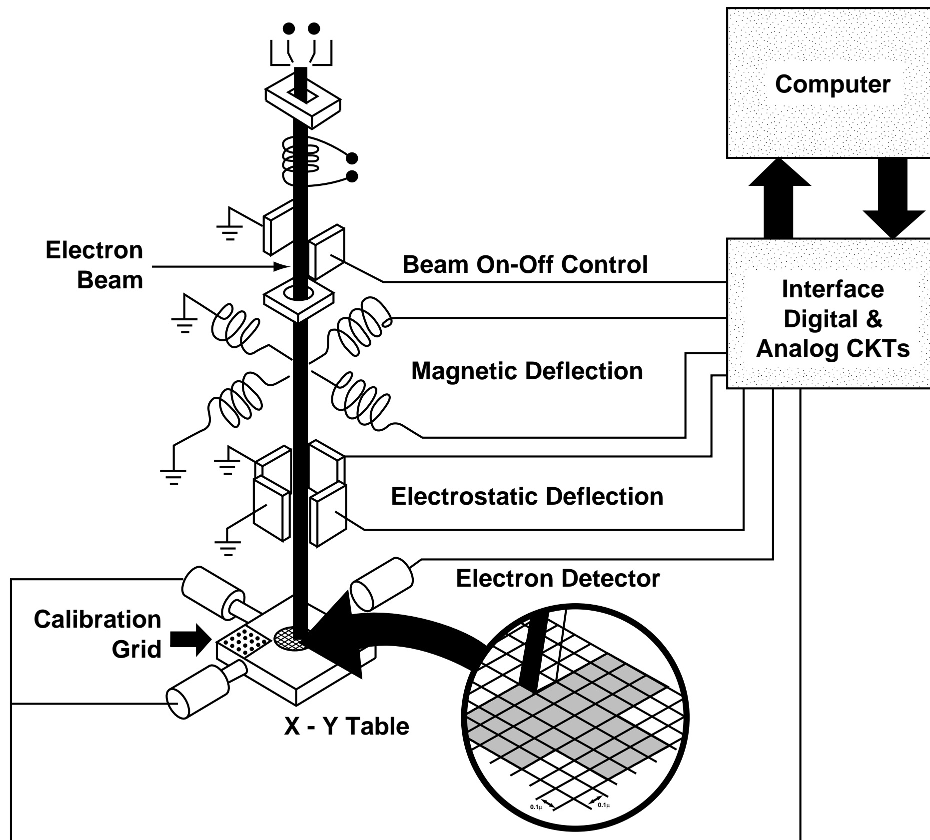


Figure 2-34. E-Beam System

Other disadvantages in the system for e-beam direct-write on the resist-coated silicon wafer are the poor thermal characteristics and the erosion rates in dry etching of the e-beam resist. Both of these parameters are very important in sub-micron lithography.

E-beam systems have been designed using either raster scanning or vector scanning of the beam.

The greatest improvement for the e-beam system will probably come from more compatible photoresist chemistries. This will provide better etch performance. It should be realized that an e-beam system used for reticle and mask masking is not as demanding as writing directly on the wafer. Wafer surfaces are not flat like reticles and masks and therefore create variation in resist thicknesses and scattering effects. Further, the underlying films are considerably thicker making the etching process effects on the resist more severe.

The direct-write e-beam process is essentially a serial process in that each image is formed within a die and then from die to die. Thus, the throughput is very low. This low throughput makes the process very expensive. However, for small volumes, engineering prototypes, or the quick time to market category, the cost may not be the most important consideration. For these needs, the e-beam system is a good choice for align and expose.

vi. X-ray

X-ray energy is another possibility as an align/expose system for imaging on wafers coated with photoresist sensitive to x-rays. X-rays have very short wavelengths thereby making possible improved resolution. Because of the very short wavelength relationship to improved resolution, x-ray technology continues to be actively pursued as an imaging technology for the future.

- **X-ray Sources**

The energy for an x-ray system comes from either high-energy sources like the synchrotron ring or electron impact low-energy sources. The synchrotron ring puts out a wide spectrum of radiation from the infrared region down to the very short wavelengths of x-rays. Beamlines are used to access this form of radiation.

The impact source of x-rays also delivers a broad spectrum of wavelengths by directing an electron beam at a target material. The x-ray radiation given off by the electron-beam impact matches the physical constants of the material and is generally below the intensity of the synchrotron source.

X-ray exposure has been used for full field proximity, 1:1 projection, and step and repeat. As in optical technology, x-ray stepping can take advantage of greater uniformity radiation in a smaller field size and greater intensity of energy. Full field x-ray is optimized for throughput at the expense of overlay registration and pattern resolution. The impact concept is illustrated in Figure 2-35.

- **X-ray Masks and Reticles**

A serious limitation to the x-ray method of image transfer is in the mask or reticle creation. The penetration power of the x-ray energy imposes severe restrictions on the materials used for reticles or masks (Figure 2-35). This causes the tooling cost for the mask or reticle to be three to five times the cost of optical masks or reticles. At this stage of development the slight difference in the spectral output between the synchrotron (0.05 to 0.5nm) and the electron impact (0.8 to 2.2nm) has effect on this tooling cost.

vii. Summary

A comparison of lithography technologies is shown in Figure 2-36. A comparison of throughput, capital cost, feature size and overlay registration is summarized in Figure 2-37.

The image transfer process is constantly undergoing change and refinements. Some of these activities are found in multilayer resist patterning, contrast enhancement techniques, the use of anti-reflective coatings, chemical amplification of the resists and adding chemical dyes to the resists. These techniques and others to follow in the future will allow further improvements to be made. The net result will be smaller feature sizes, higher levels of integration, and more complex integrated circuits.

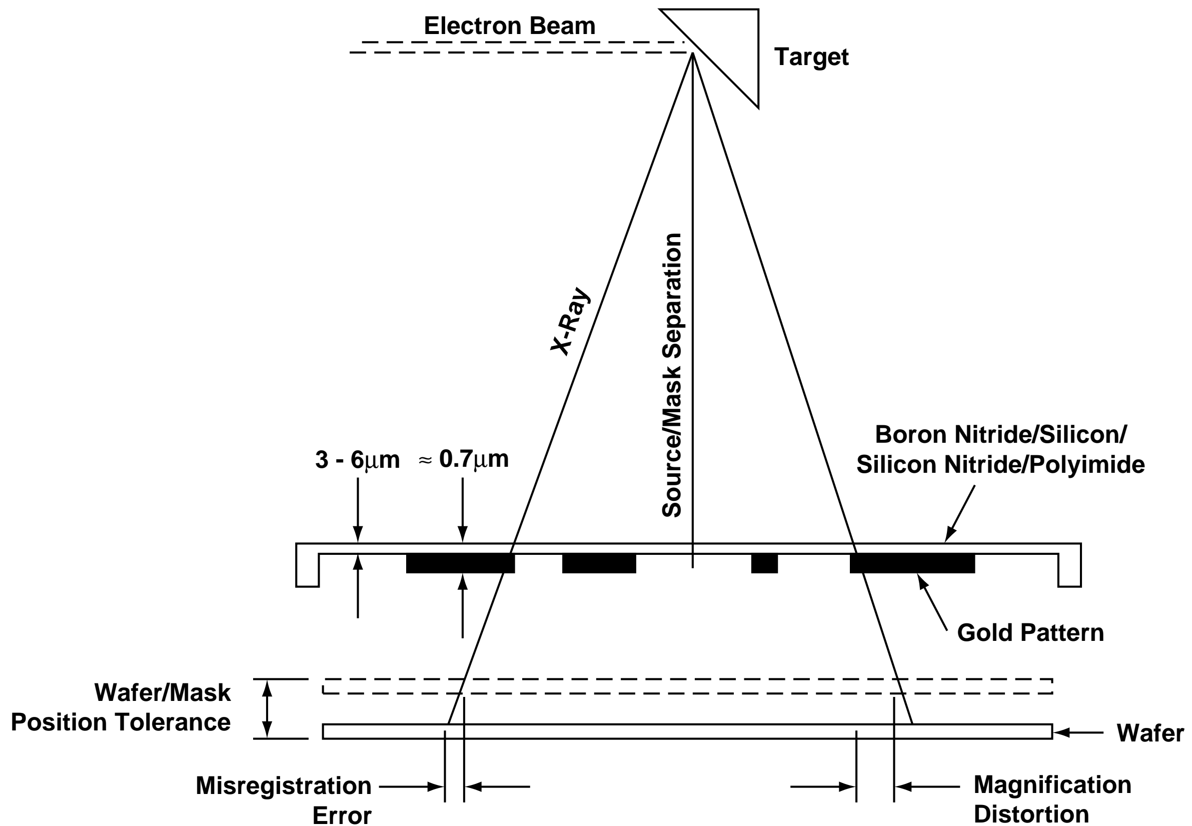


Figure 2-35. X-Ray Lithography

After the align/expose process has been completed, some types of photoresists require a post-exposure bake to further stabilize the film. Otherwise, the wafer is ready for the develop process.

e. Photoresist Developing

i. Positive Resist

The developer solution reacts chemically with the photoresist to form the image in the photoresist layer. If the resist layer is positive resist, an alkaline developer solution will react only with the exposed resist. This chemical reaction dissolves the exposed resist, leaving the unexposed resist. The chemical reaction of the develop process is stopped by flooding the wafer with D.I. water. The water neutralizes the alkaline solution to stop the chemical reaction.

TECHNOLOGY	ADVANTAGES	DISADVANTAGES
Contact Printing	Inexpensive Low maintenance High throughput	Damages photomask High defects
Proximity Printing	Inexpensive Minimal mask damage High throughput Good depth of focus	Regular adjustment required Limited resolution
Scanning Projection	Good resolution Long mask life Low defects	Expensive Regular adjustment required Temperature sensitive Vibration sensitive Global alignment only
Optical Direct Step on Wafer	Very good resolution Low defects Long mask life Good registration	Very expensive Special mask required Distortion/stepping errors Environmental control required
Electron Beam on Wafer	Excellent resolution No mask required Very low print defects Excellent pattern flexibility	Very expensive Low throughput Special Resist Required Very Complex System
X-Ray	Excellent resolution High throughput	Special mask required Special resist required Global alignment and step & repeat

7779A

Figure 2-36. Comparison of Lithography Technologies

MACHINE	MACHINE RATE/HOUR	COST ($)	FEATURE SIZE (μm)	OVERLAY REGISTRATION (μm)
Contact Aligner	60	50K	4.0	2.0
Proximity Autoalign	60	150K	2.5	1.5
P & E Autoalign	60	750K	1.0	0.8
Stepper I 5X	20	1.0M	1.0	0.5
Stepper II 5X	20	1.6M	0.5	0.15
Stepper III 5X	20	2.6M	0.35	0.08
Stepper IV 4X	25	3.0M	0.35	0.08
Step and Scan 4X	30	3.2M	0.35	0.08
X-Ray Stepper	25	5.0M	0.3	0.1
X-ray Stepper II	30	5.2M	0.2	0.08
E-Beam	5	4.0M	0.15	0.08

14658A

Figure 2-37. Photolithography Equipment Summary (All Rates Normalized to 150mm Wafer)

ii. Negative Resist

Negative resist requires a solvent solution for developing the image. Xylene is the common developer for negative resist. The exposure process causes the negative resist to polymerize. The polymerization of the resist renders the resist insoluble in the developer. Thus, only the unexposed resist is dissolved and washed away. The chemical reaction of the Xylene requires another solvent to stop the action. N-butyl acetate is commonly used to neutralize and stop the developing action and stabilize the system.

iii. Developing Methods

There are several methods used to develop the photoresist. The particular method chosen is determined by the requirements of feature size and the parameters of the resist system. The more commonly used methods are immersion (batch develop), spray, and a modified spray system known as puddle develop.

- **Batch Develop**

Batch develop is the oldest method. The wafers are placed in carrier that is not affected by the developer chemical. The carrier is immersed in the developer solution for a given period of time. The wafers are then removed from the developer solution and placed into a second solution to stop the develop process and rinse the wafers.

The batch develop process has gone through many changes and refinements over the years but has been gradually replaced by one of the spray methods: spray for negative resist and puddle for positive resist.

- **Spray (Puddle) Develop**

The spray technique has the advantage of processing one wafer at a time with fresh developer solution, improving the uniformity of the process and allowing the handling of the wafers from cassette to cassette. This provides excellent productivity.

The conventional spray system used for negative photoresist must be modified for positive resist. The spray system is an adiabatic process. The change in spray pressure between the nozzle and atmosphere causes too much of a temperature change for consistent develop results with positive resist. To overcome this sensitivity to temperature variation, the puddle technique dispenses the developer solution onto the wafer with a very minimal pressure change. (It is like removing a nozzle from a hose). The solution is allowed to flow over the entire surface of the wafer quickly. The puddle of fluid is allowed to stand on the surface for the required develop time wherein a flow of D.I. water floods the surface to neutralize the developer solution. This process has also been refined by controlled temperature of the developer solution, heated wafer holder, multiple puddles, low-pressure spraying after the puddle step, etc. Consequently, consistent results are achieved with this method.

iv. Plasma Descum

A process used to further improve the surface to be etched after the develop process is called Plasma Descum. The Descum process is done by placing the wafer in a plasma system wherein oxygen is injected into the low-pressure process chamber with an applied R.F. field. The R.F. energy ionizes the oxygen and causes a chemical reaction between the resist and the ionized oxygen. This in effect etches the wafer and removes thin films of resist that may not have been completely removed in the develop process, providing a cleaner surface for the etch process.

v. Spin-Dry Process

Some form of the spin-dry process is used to remove the final rinse solution from the wafer for both the positive and negative develop processes. The wafers are ready for a visual inspection. The result of the develop process for positive resist is illustrated in Figure 2-38.

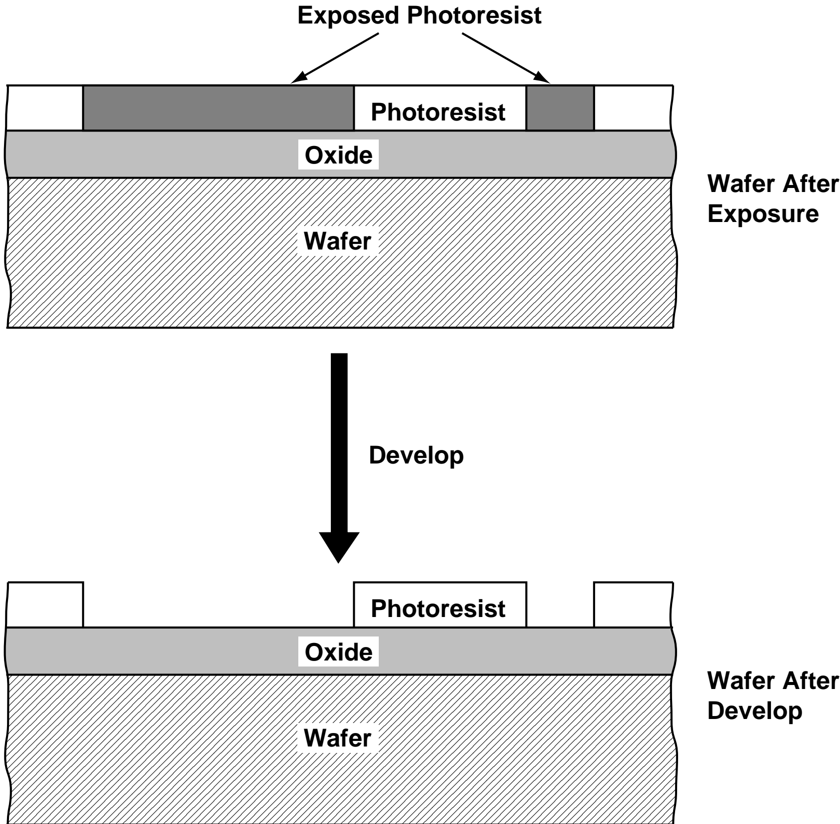


Figure 2-38. Photoresist Develop

vi. Inspection After Develop

Visual inspection of the photoresist-patterned surface is a difficult task. And, as the circuitry becomes denser, the inspection after develop (ADI) will become even more difficult. Therefore, the photoresist process will rely more on Statistical Process Control and less on the visual inspection.

The current trend in technology for monitoring the develop process after developing is to use programmed pattern recognition visual inspection equipment. This type of automation eliminates the eye fatigue problem and provides more consistent results. The automated visual inspection station sample inspects using a statistical plan. This provides immediate feedback for manufacturing control and aids in keeping the rework rate low.

The decision to rework a wafer (or wafers) after the develop inspection has a significant impact on the cycletime of the overall process. Good manufacturing practices dictate keeping reworks low at this inspection. The long-term goal is to be able to eliminate this inspection through continued improvement by using Statistical Process Control.

vii. Post-Develop (Hard) Bake

After the inspection is completed, the wafers are given a "hard bake" at a temperature of 130°C to 160°C to dehydrate the photoresist prior to etch or ion implant processes. This bake stabilizes the resist characteristics and makes the resist less sensitive to these hostile process environments.

f. Etch

The etching process removes the material not protected by the hardened photoresist. The result of the etching process was illustrated in Figure 2-15 for positive resist and is shown in Figure 2-39 for negative resist.

These figures are idealistic representations of the patterning process. In actual manufacturing, there are varying degrees of acceptable results. Figure 2-40 illustrates several examples.

The etching process is divided into two basic methods: isotropic and anisotropic. These terms describe the geometrical shape the film will have after the etching process.

i. Isotropic/Anisotropic

An isotropic etch process will etch laterally while etching the material vertically. The worst case is when the lateral rate is equal to the vertical etch rate. The anisotropic etch will etch only in the vertical direction. These concepts are illustrated in Figure 2-41.

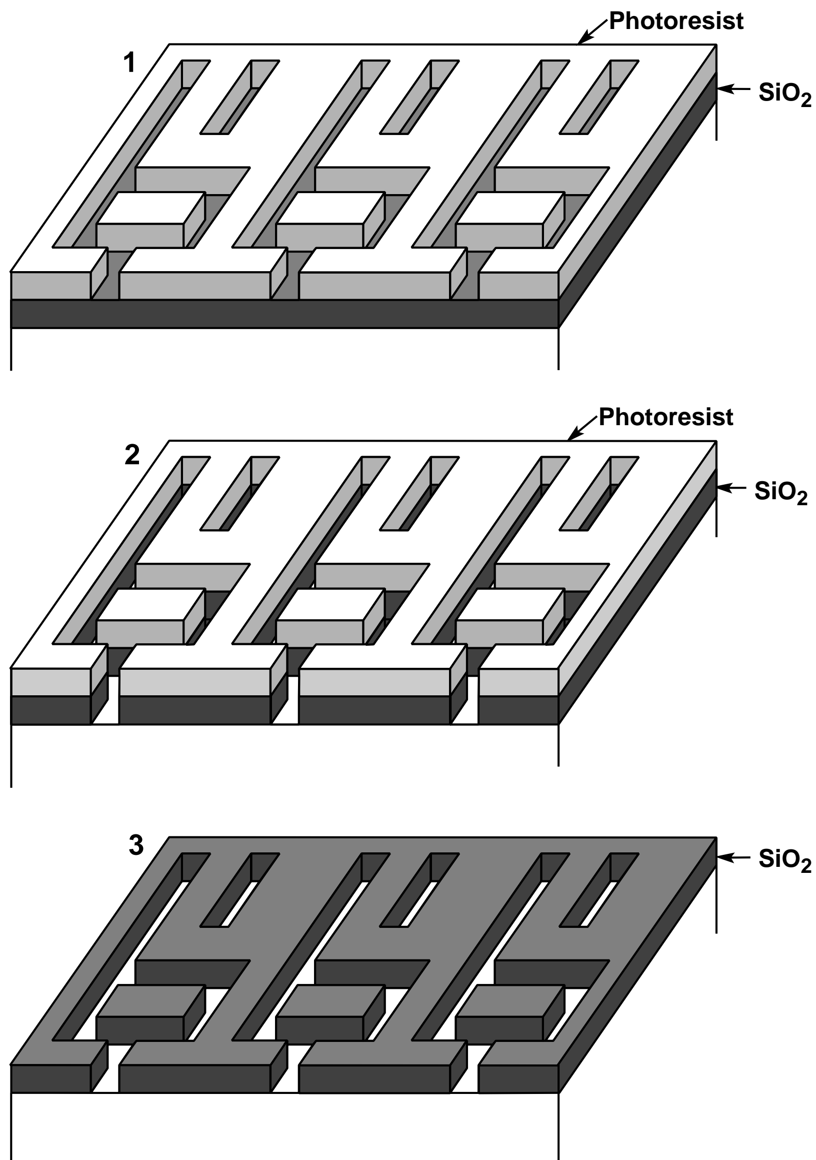


Figure 2-39. Photolithography Using Negative Photoresist

The etching process is essentially a subtractive process to remove the material not protected by the photoresist. There are several methods used to remove any given material. The more common methods are: wet chemistry, plasma dry chemistry, reactive ion etching dry chemistry, and ion milling.

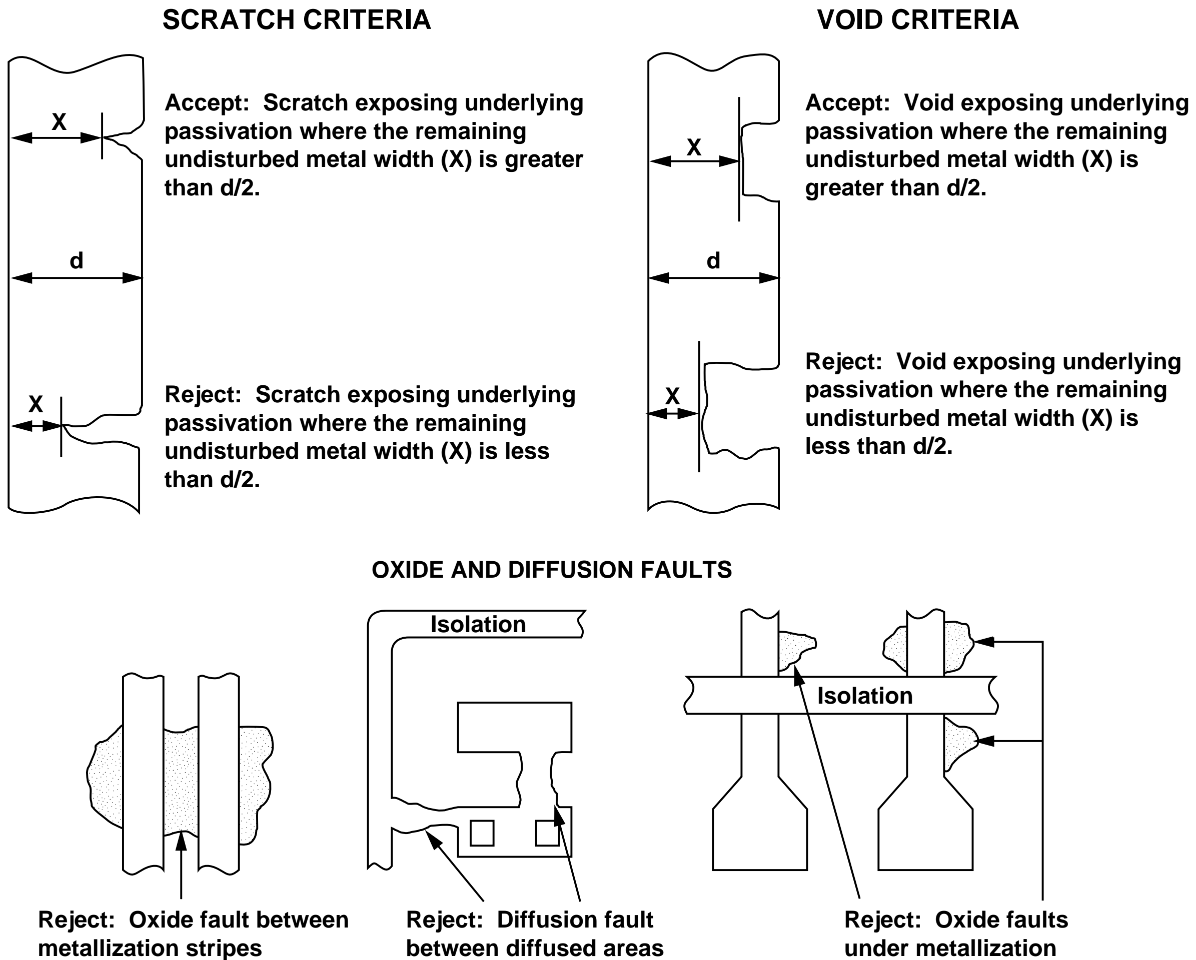


Figure 2-40. Major Visual Defect Examples

Wet chemical etching is the oldest etch method used. A chemical solution that rapidly etches the layer to be removed while having minimal reaction with the photoresist is chosen. This interaction of the etchant to photoresist vs. the layer to be etched is called selectivity. Therefore, the higher the selectivity of the etchant or etch system, the better.

ii. Dry vs. Wet Etch

Feature size has a significant impact on the choice of wet vs. dry etching technique and equipment (Figure 2-42). Generally, feature sizes greater than three microns can be resolved with wet etching. Feature sizes less than three microns usually require dry etching.

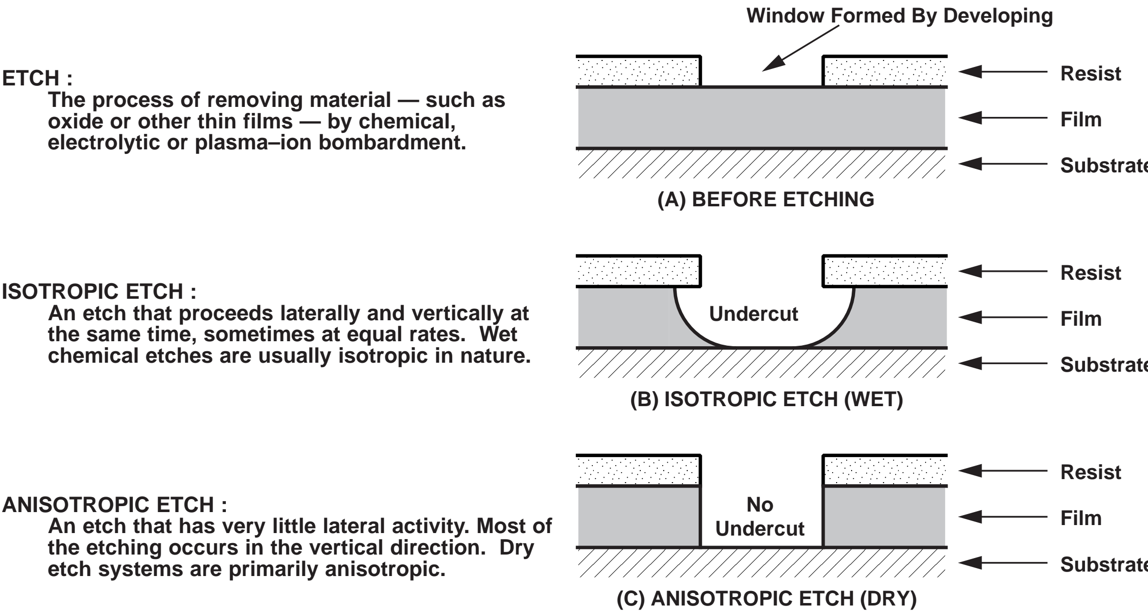


Figure 2-41. Vocabulary - Etch Process Section

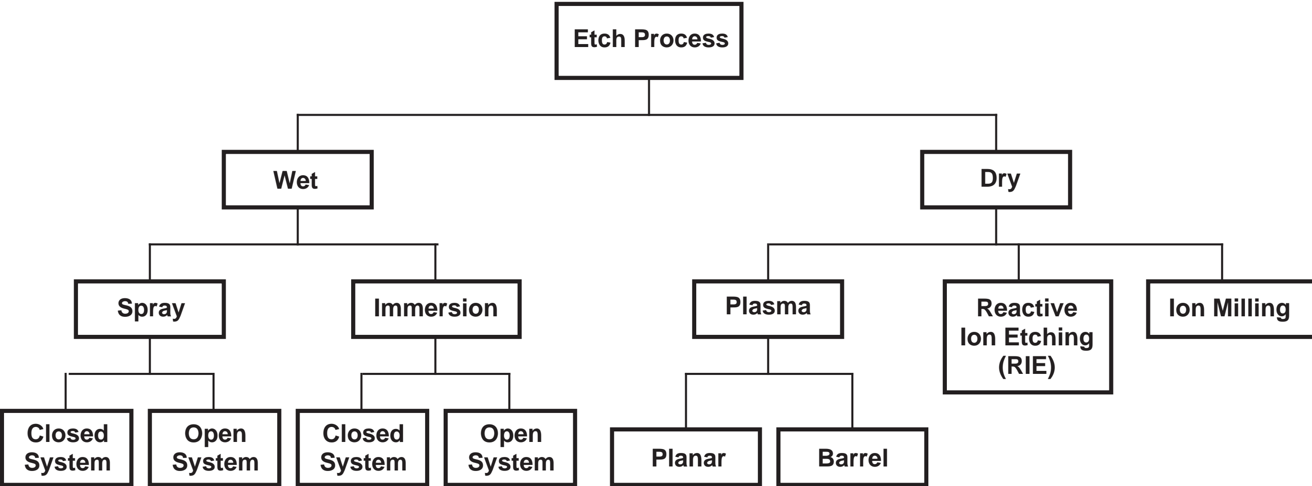


Figure 2-42. Etch Methods

iii. Wet Etch

Wet chemical etching of oxides is done with Buffered Oxide Etch (BOE). This etchant is composed of ammonia fluoride (NH_4F) and hydrofluoric acid (HF). The concentration of the HF is varied to affect the etch rate of the solution.

Silicon-nitride films are etched with hot phosphoric acid (H_3PO_4). Generally, the temperature is in the 150°C - 180°C range. Because photoresist does not stand up well at this temperature, thin layers of silicon dioxide are used on top of the silicon nitride as the etch mask.

Aluminum films are etched with a mixture of phosphoric acid, acetic acid (HAC), and nitric acid (HNO_3) diluted with D.I. water. The etch is used in the temperature range of 20°C - 80°C.

Monocrystalline and polycrystalline silicon are etched with a mixture of HF, HNO_3, and HAC. The temperature of the solution is maintained in the 20 - 30°C range.

Wet chemical etching can be done by either immersion in a given chemical solution or by spraying the chemical onto the surface. After the etching process is completed, the wafers are rinsed with D.I. water and dried in a spin-dryer.

Unfortunately all of the wet chemical etches leave the edge of the film removed with some degree of slope.

iv. Dry Etch

The choice among the use of the various dry etch methods is feature-size driven. The more critical the tolerance, the more likely RIE will be chosen over plasma methods and hardware.

Dry etch processing is performed under some degree of vacuum (Figure 2-43). Dry etch systems require pumping systems for low-pressure operation and gas flow control devices for the various etchants along with an R.F. power supply and control electronics. These etch systems are more complex and require more maintenance than the simpler wet etch stations.

Barrel-type reaction chambers were initially used for plasma etching. This design was acceptable for removing photoresist but had poor uniformity characteristics for dielectric etching. The plasma barrel-type etcher is shown in Figure 2-44.

The next generation design was the planar parallel-plate plasma reactor. This improved uniformity of etching dielectric layers and led to the lower operating pressure reactive ion etcher (RIE). This is shown in Figure 2-45.

Ion milling systems are still being perfected. The selectivity of the system is poor and the capital cost is high.

g. Photoresist Removal

At the completion of the etching process, the photoresist is removed. The pattern has now transferred from the mask or reticle into the resist and from the resist into the layer below the resist (Refer back to Figures 2-15 and 2-39). This basic process is repeated for all of the number of layers a particular design requires.

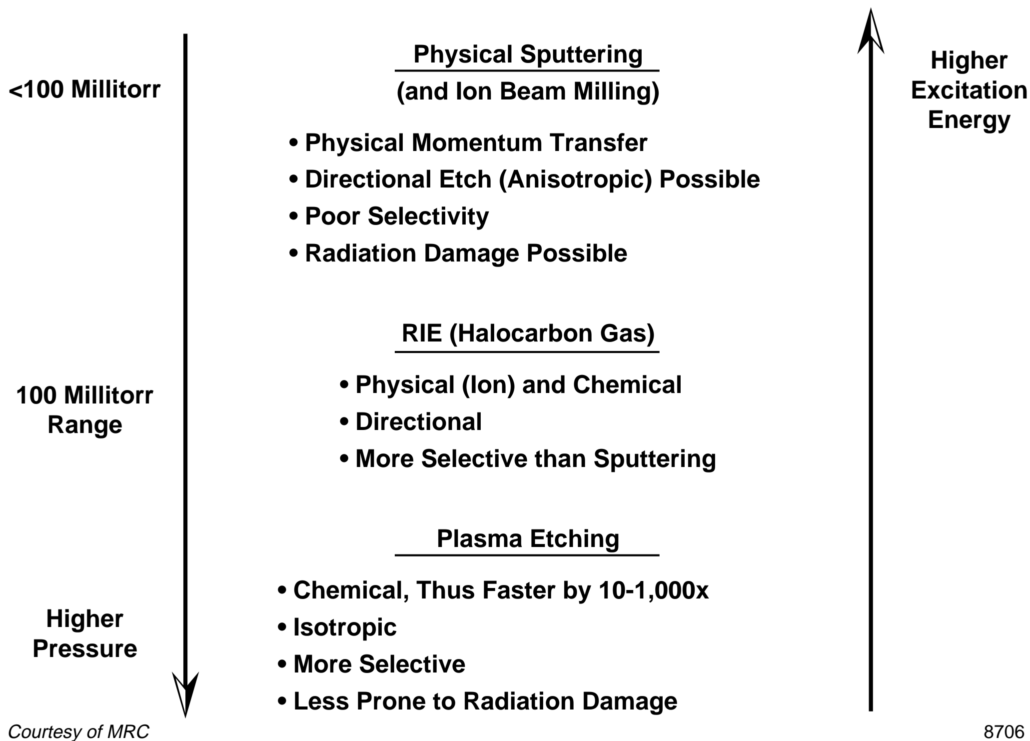


Figure 2-43. The Dry Etching Spectrum

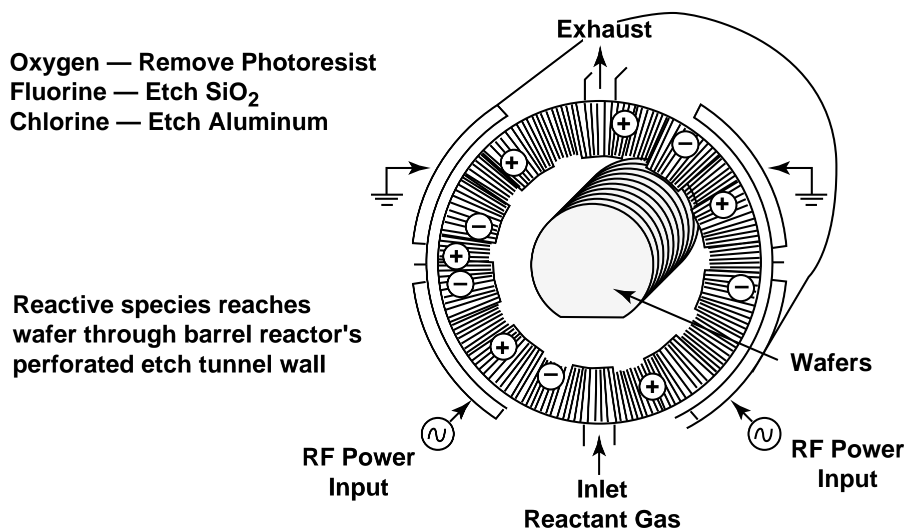


Figure 2-44. Diagram of Barrel-Type Plasma Etcher

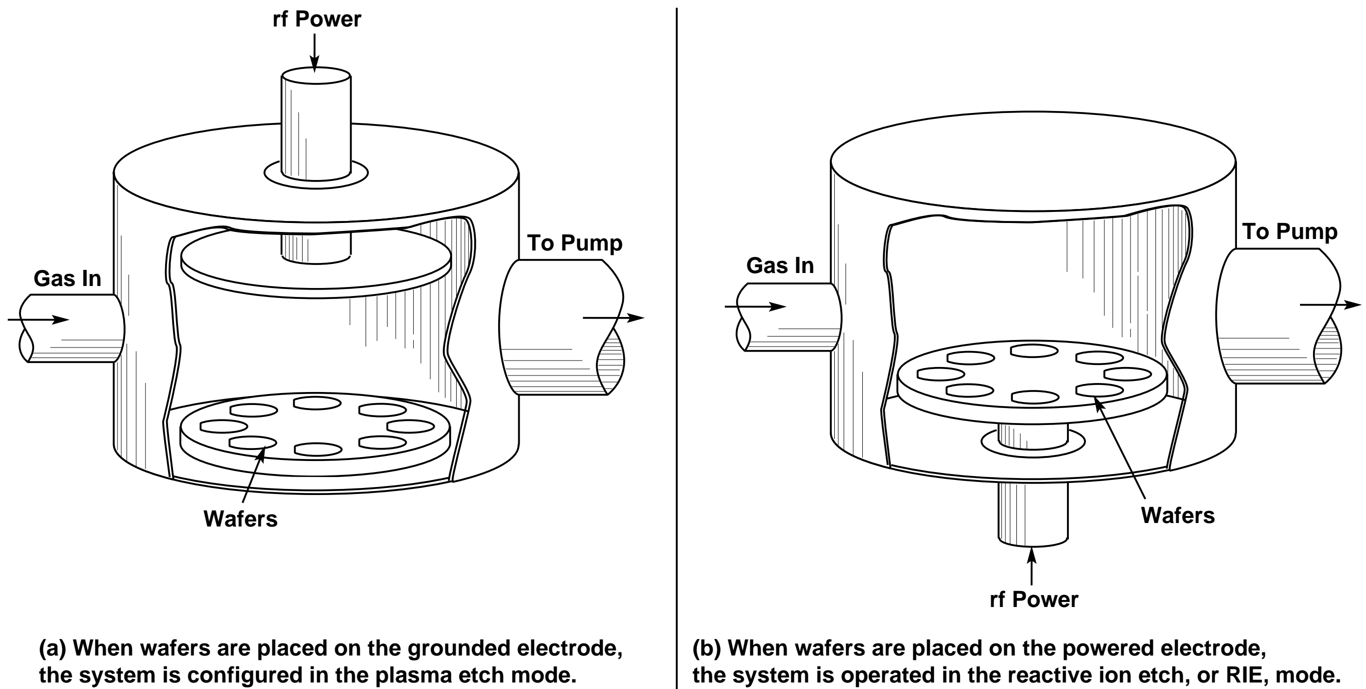


Figure 2-45. Planar Type Reactor

The actual removal of the photoresist can be done in several different ways.

1. Wet chemistries
2. Dry chemistries
3. Heat
4. Light energy and gases.

Each removal technique has advantages and limitations.

Photoresist used for layers prior to metallization can be removed with wet chemistries such as sulfuric acid (H_2SO_4) and an oxidizer like peroxide or ammonium persulfate, sulfuric acid with chromic acid, organic acids, and solvents.

Dry chemistry removal usually uses oxygen in the presence of R.F. energy. The reaction generates gaseous oxides of carbon and water. These species are pumped away by the vacuum pump. Unfortunately, there are a few metallic ions present in the resist system that do not get removed with this cleaning method. However, this can be overcome with a wet chemical clean. In the case where a metal conducting layer is the layer under the photoresist, organic acids can be used as the follow-up clean.

Heating the photoresist as a cleaning process is not used because of the residue remaining. A technique that is gaining acceptance is to use ultraviolet energy in the presence of ozone (O_3).

After the photoresist has been stripped (removed), the wafer is given some type of visual inspection. Like other visual inspections, pattern recognition inspection equipment and sampling techniques are used.

E. JUNCTION FORMATION

Recalling from the materials section, dopant atoms are added to the silicon lattice when the silicon ingot is grown to provide the electrical characteristics of the wafer (N-type or P-type). During the process of producing ICs on the wafer, atoms of the same polarity as the wafer or of the opposite polarity must be introduced into the wafer in selected regions. This alteration of the dopant levels is done by solid-state diffusion or ion implantation.

1. Diffusion

Diffusion is the term used to describe the movement of atoms, molecules, or particles from a location of high concentration to a new location of lower concentration. An example of the process of diffusion can be visualized by watching a small drop of red coloring being dropped into a glass of clear water.

Initially, as the drop strikes the clear water, the red color is highly concentrated. After a short period of time the deep red color begins to lessen and the color starts to spread out over a larger volume. As more time elapses, the color continues to spread out and the red color starts to change to pale red and on into a pink. In a given time the pink will almost disappear in the now nearly clear water.

The process just described represents diffusion as a function of time and temperature. The rate an atom, molecule or compound diffuses from a high concentration to a lesser concentration is related to temperature and time. The parameter that ties the diffused rate to temperature for a given time is known as the diffusion coefficient or diffusivity.

The dopant atom must displace a silicon atom in the crystal structure of the silicon to become electrically active. The process of diffusion is used in semiconductor processing to introduce a controlled amount of a chosen dopant into selected regions of a semiconductor crystal. The diffusion process used to accomplish this substitution is divided into two distinct steps.

1. Predeposition
2. Redistribution or drive-in

a. Predeposition

The appropriate diffusion hardware is chosen to allow the predeposition process to introduce a carefully controlled amount of the desired dopant into the semiconductor crystal. The variables in the process and selection of the dopant species are: the rate the species diffuses into the silicon at a given temperature, the amount of dopant the silicon will accept into the crystal structure at a given temperature, and the amount of silicon dioxide required to mask the dopant at the given process temperature. The required transistor parameters determine how many dopant atoms should be in the silicon and how far the dopant concentration should be below the surface. This selective doping process is shown in Figure 2-46. The illustration is also indicating that the dopant source is available during the entire predeposition cycle.

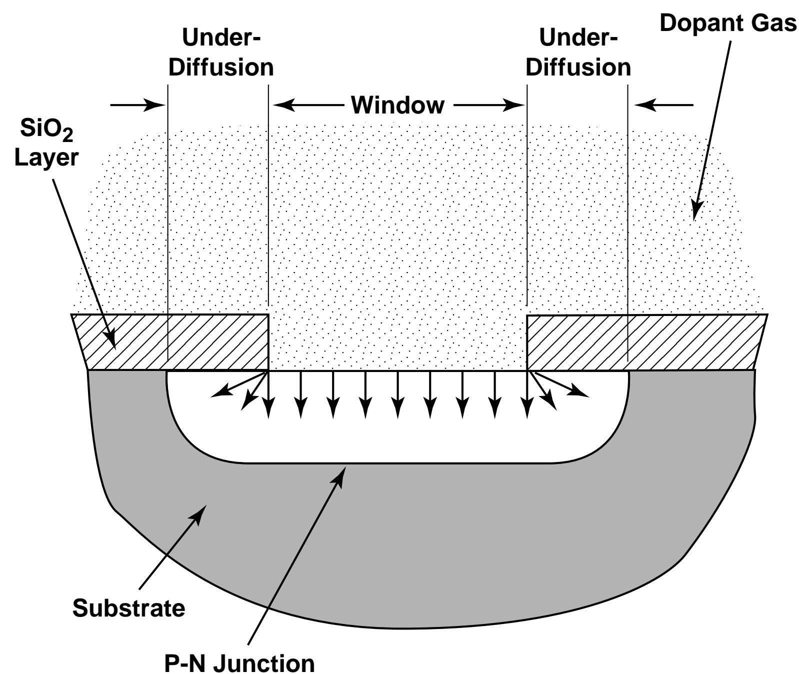


Figure 2-46. Selective Doping Using SiO_2 as a Mask

The dopant species can be brought to the wafer surface several different ways. Liquid chemical sources can be introduced into the process tube as controlled chemical vapors or applied directly to the surface prior to loading the wafers onto the quartz diffusion boat. Gas sources can be precisely metered into the process tube. Solid diffusion sources made from high-temperature materials can be loaded onto the quartz boat with the patterned side of the wafer facing the source material. Each of these methods has its own set of advantages and limitations.

The predeposition process will form a thin layer of doped silicon dioxide on the wafer surface during the predeposition cycle. This oxide has incorporated into its structure the dopant. This doped oxide layer is generally removed prior to the redistribution process. The removal is done by wet chemical methods or a sequence of a low-temperature thermal oxidation for a brief time followed by the wet chemical method.

If the source is made available for an extended period of time, more and more dopant atoms will be added to the silicon lattice. As this diffusion process is allowed to continue longer and longer, the concentration gradient goes in deeper and deeper. If left to continue, eventually the silicon wafer would be completely doped uniformly throughout the thickness of the wafer. This is illustrated in Figure 2-47.

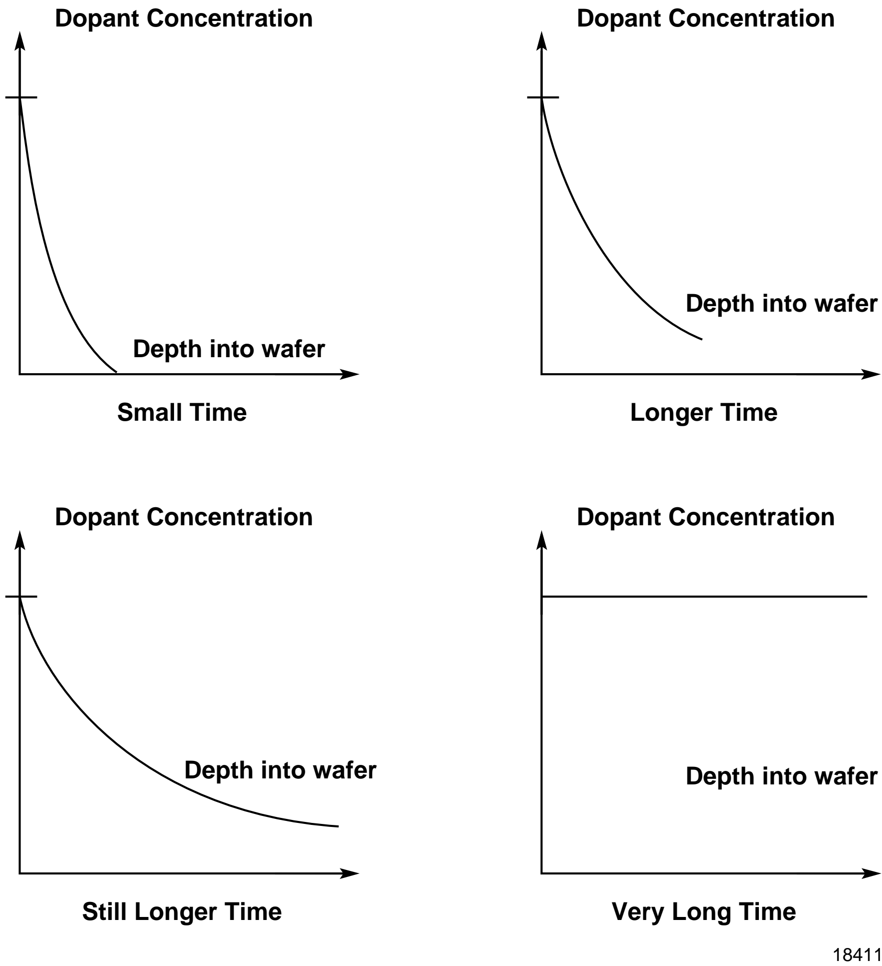


Figure 2-47. Profile of Dopant Present in a Wafer as a Function of Time

To control the device parameters, the predeposition process embeds a precise quantity of dopant atoms into the silicon using a dedicated high-temperature furnace. At the completion of the predeposition process, the wafers are removed from this furnace and moved to a different furnace for the second step of the process.

b. Drive-In

The redistribution or drive-in process is a diffusion step in which no additional dopant is introduced into the wafer. This process step is done in a controlled atmosphere of nitrogen, nitrogen and oxygen, steam, or combinations of these. At some part of the redistribution cycle, an oxidizing atmosphere is used to regrow a protective oxide layer over the diffused regions. The process variables of time, temperature, and ambient gases are controlled. These variables control the final junction depth, the thickness of the oxide in the diffused region, and the exact dopant profile in the wafer.

The predeposition and redistribution process flow charts are shown in Figures 2-48 and 2-49. In wafer fabrication, both processes are performed in furnace equipment similar to that shown in Figure 2-50. Wafers that are ready for predeposition are cleaned to remove any contamination from storage or prior operations.

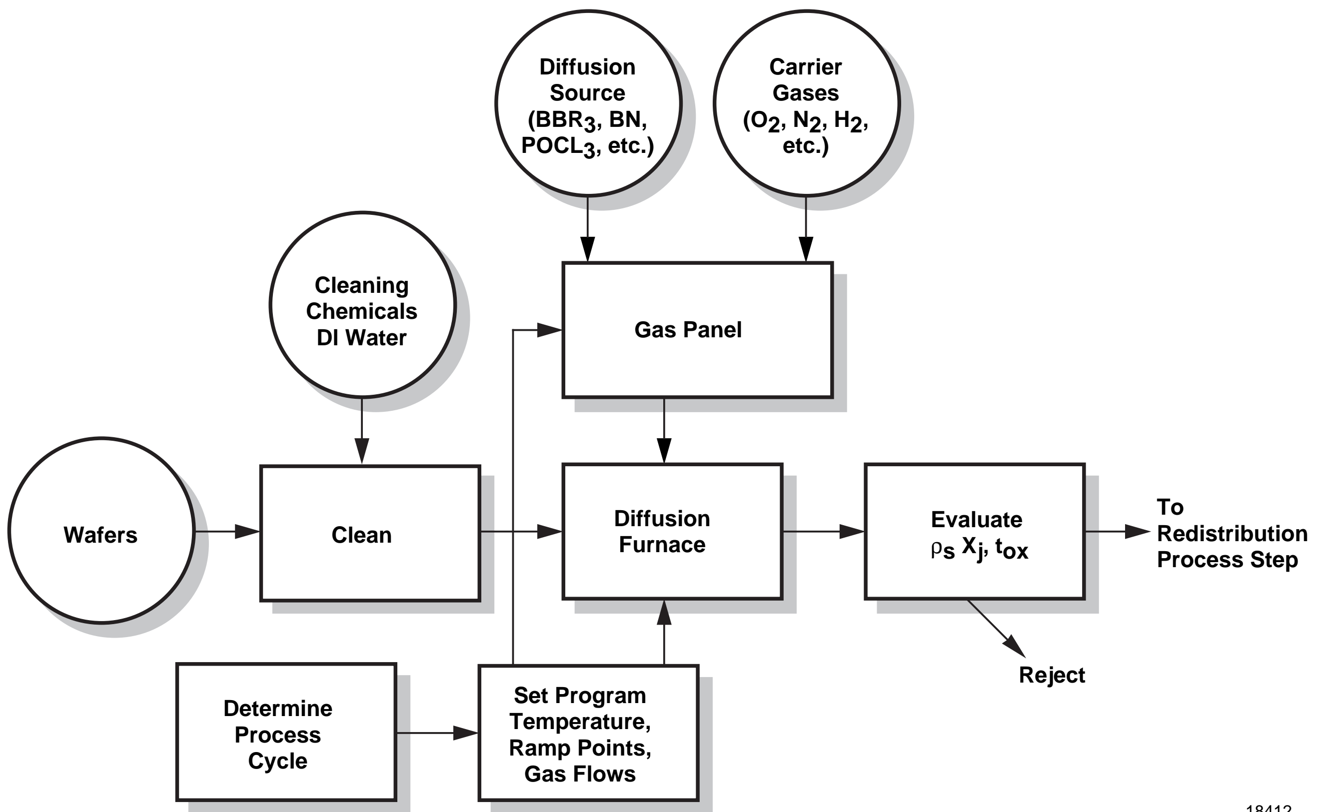


Figure 2-48. Diffusion Process Flow Chart for Predeposition

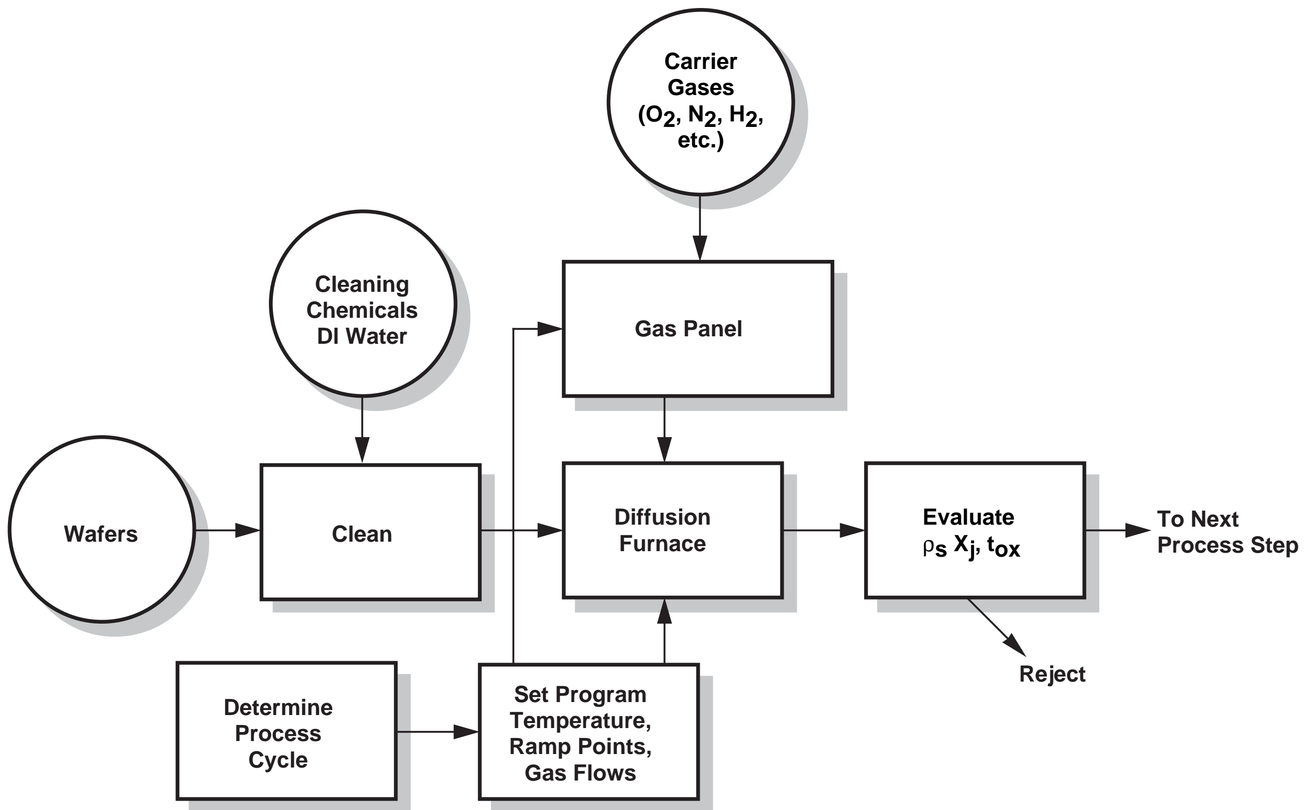
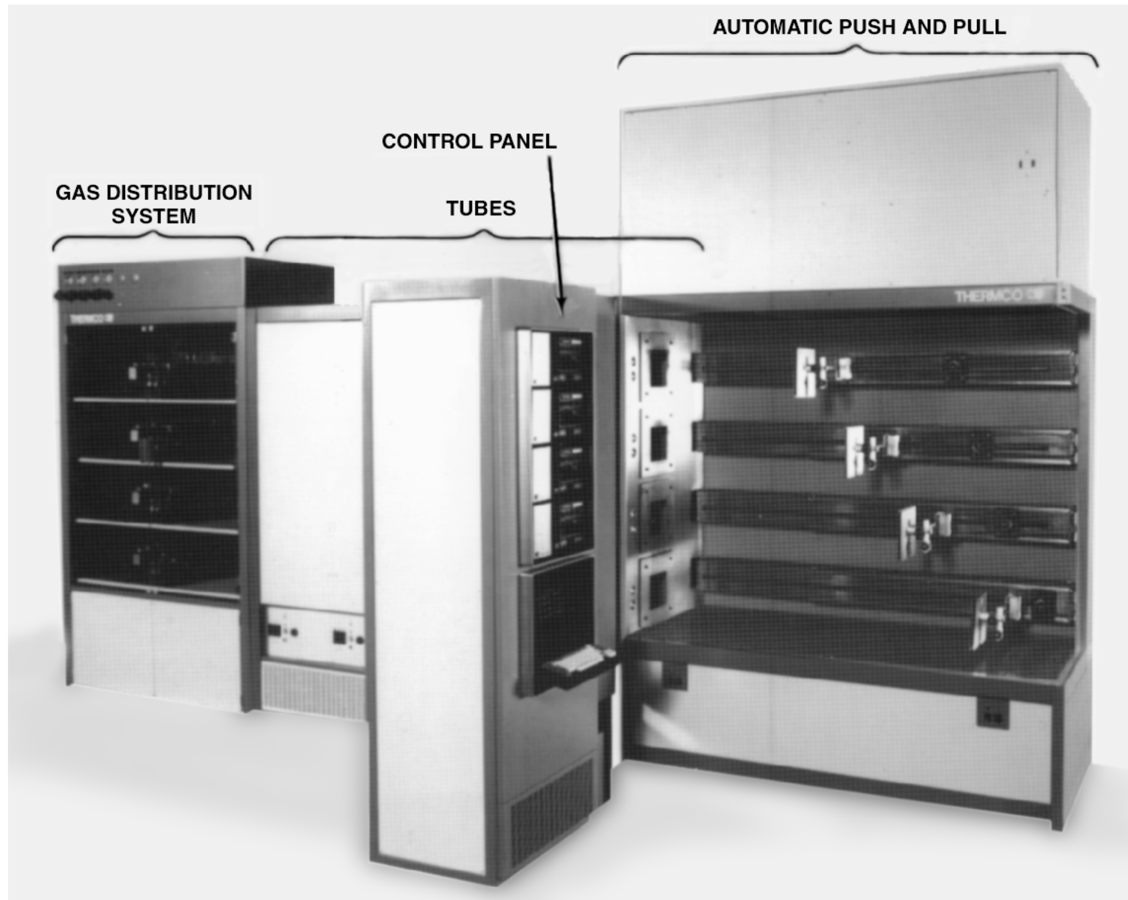


Figure 2-49. Diffusion Process Flow Chart for Redistribution

c. Process Control Measurements

Process control measurements are performed after both predeposition and redistribution (drive-in). Because of the selective nature of these process steps, "test wafers" are used in both the predep and redistribution process steps for these process control measurements. Typically, the sheet resistances of the test wafers are recorded after the predep step and the test wafers are placed on the redistribution boat. After the redistribution process is completed, the new sheet resistances, oxide thicknesses, and junction depths are measured and recorded. These measurements are assumed to be representative of the actual product wafers. For correlation purposes, the product wafers have test patterns that are measured at the completion of the process. These readings are then compared to the test wafers' readings. The correlation is good.



Courtesy of Thermco 5749A

Figure 2-50. Oxidation/Diffusion Systems

i. Resistivity

The resistivity measurement made on the test wafer is illustrated in Figure 2-51. This type of resistivity measurement is often measured in terms of ohms per square, $\mathbf{R_s}$ or $\mathbf{R_\square}$. For the diffusion process, this means that the test wafer used is doped with the opposite type carriers, i.e., a P-type wafer is used for N-type diffusions and an N-type wafer is used for P-type diffusions. The ohms per square concept is further illustrated in Figure 2-52. In this figure, the measuring current, **I**, is being forced through thickness, **Xj**, the diffused region. As long as the lateral dimensions are large and the thickness, **Xj**, is small compared to the probe needle spacing, s, the sheet resistance, $\mathbf{R_s}$, can be determined. In Figure 2-52, $\mathbf{L_1}$ equals $\mathbf{L_2}$, and $\mathbf{R_s}$ is the ratio of the measured voltage, **V**, to the forcing current, **I**, multiplied by the constant, 4.532. In practice, the current, **I**, is decade multiples of 4.53. This makes the measured voltage, **V**, a direct reading of resistance in ohms per square.

This reading has sensitivity to light. Therefore, the readings are generally taken in a dark enclosure.

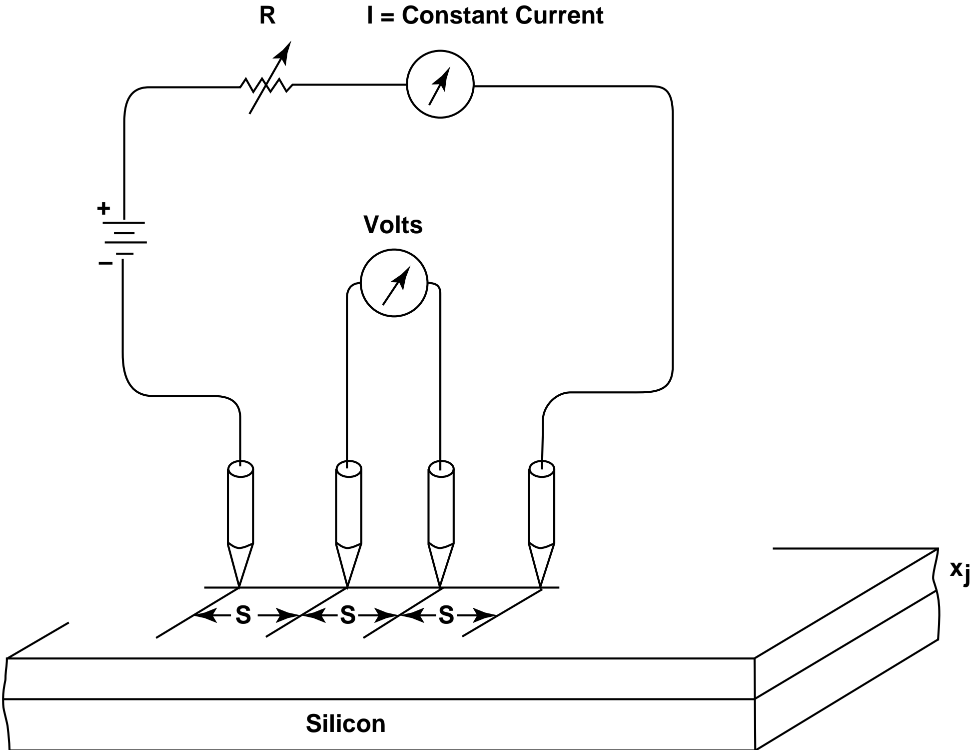


Figure 2-51. Resistivity Measurement (ρ)

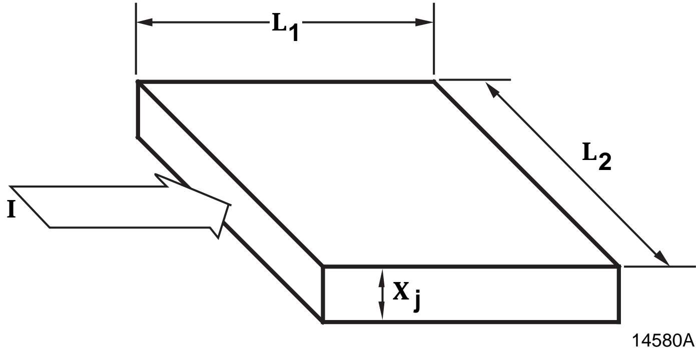


Figure 2-52. Illustration for Ohms per Square

ii. Junction Depth

The junction depth measurement, **Xj**, is made on a test wafer by lapping the wafer at a very shallow angle. The angle lapping concept is illustrated in Figure 2-53. After lapping, the wafer is etched in a chemical that stains the P region darker than the N region. This difference in color will reflect monochromatic light such that the interference fringes can be counted. Knowing the wavelength of the monochromatic light and the number of fringes allows the junction depth to be determined. Figure 2-54 shows the results of junction staining and monochromatic light on a bipolar transistor. Figure 2-55 illustrates an MOS transistor with the source and drain junctions delineated.

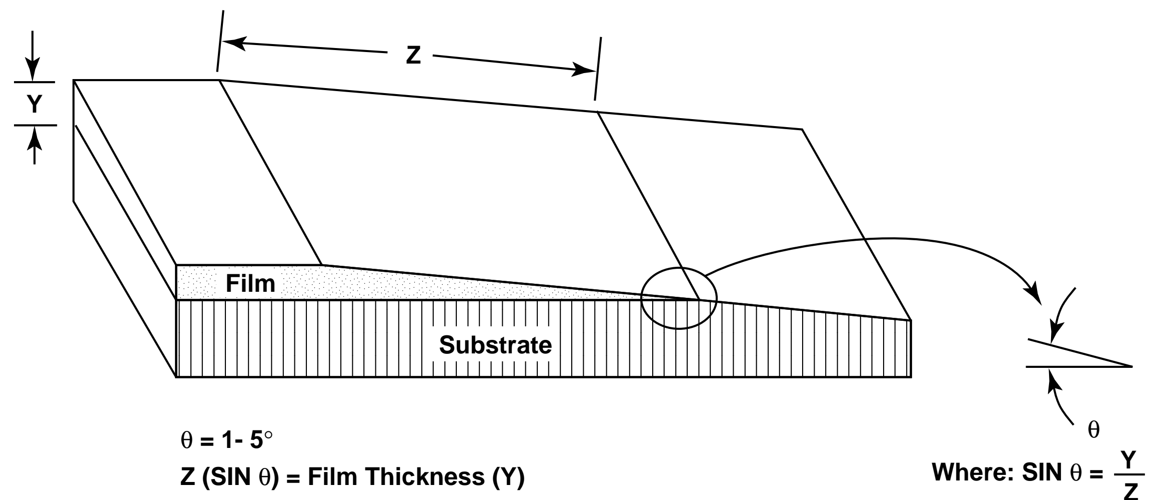


Figure 2-53. Angle Lapping Measurement Method

2. Ion Implantation

An alternative method for introducing dopant atoms into silicon emerged in the late 1960's. This new development is called ion (an ion is a charged atom) implantation. Essentially, ion implantation is the process of shooting ions of the desired dopant species into the wafer through openings in the oxide or hardened photoresist.

In contrast to diffusion, ion implantation is a low-temperature technique. It provides a flexibility not available with diffusion. The ion implant process takes ions of a desired dopant, accelerates them using an electric field, and scans them across the wafer surface to obtain a uniform predeposition. Scanning is the process of the beam moving back and forth across the surface of the wafer. This predep concentration is similar to the predeposition described in the diffusion section, except the dopant does not displace silicon atoms in the silicon lattice after the implant is completed. More on this in a later paragraph.

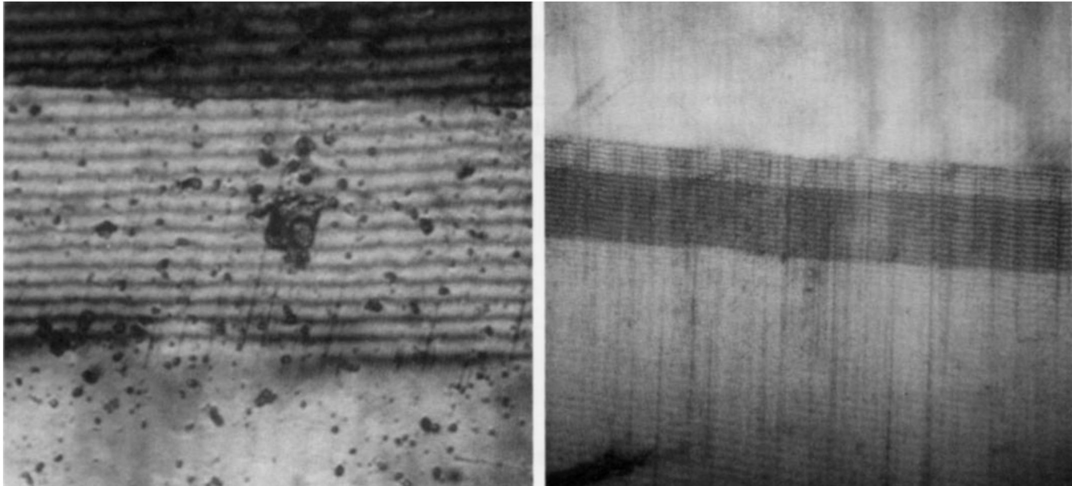


Figure 2-54. Monochromatic Light Examinations

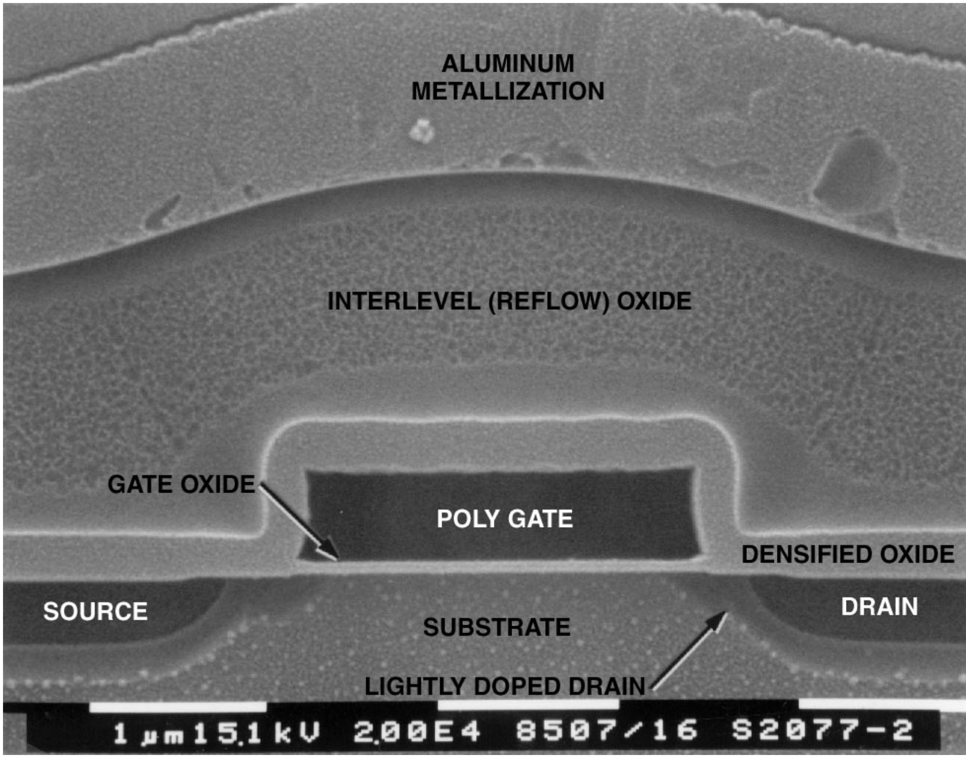


Figure 2-55. Silicon Gate MOS Cross-Section

The first requirement of an ion implantation system is the ability to generate ions of the desired species. A gaseous source containing the dopant atom is fed into a vacuum chamber and ionized by coming into contact with a hot filament. The correct ions are separated from any other ion generated by bending them a set angle using an electromagnetic field. This separation by atomic mass allows only the desired ions to enter the acceleration tube. The selected ions are accelerated using an electric field and strike the target wafer at some predetermined high-energy level. The ion beam passes through an electro-optical beam shaping region of the accelerator tube to control the diameter of the beam and causes the beam to be scanned. It is a scanned beam that strikes the wafer surface and penetrates inward to some depth. The basic concepts are summarized in Figure 2-56. A diagram of an ion implanter is shown in Figure 2-57.

1. **Generation of ions in some kind of source region.**
2. **Separation of ions by mass to prevent undesired species from reaching the target.**
3. **Acceleration of the selected ions to high energies.**
4. **Bombardment of the sample (located in a target chamber) by the high-energy ions.**

1422A

Figure 2-56. Basic Concepts of Ion Implant

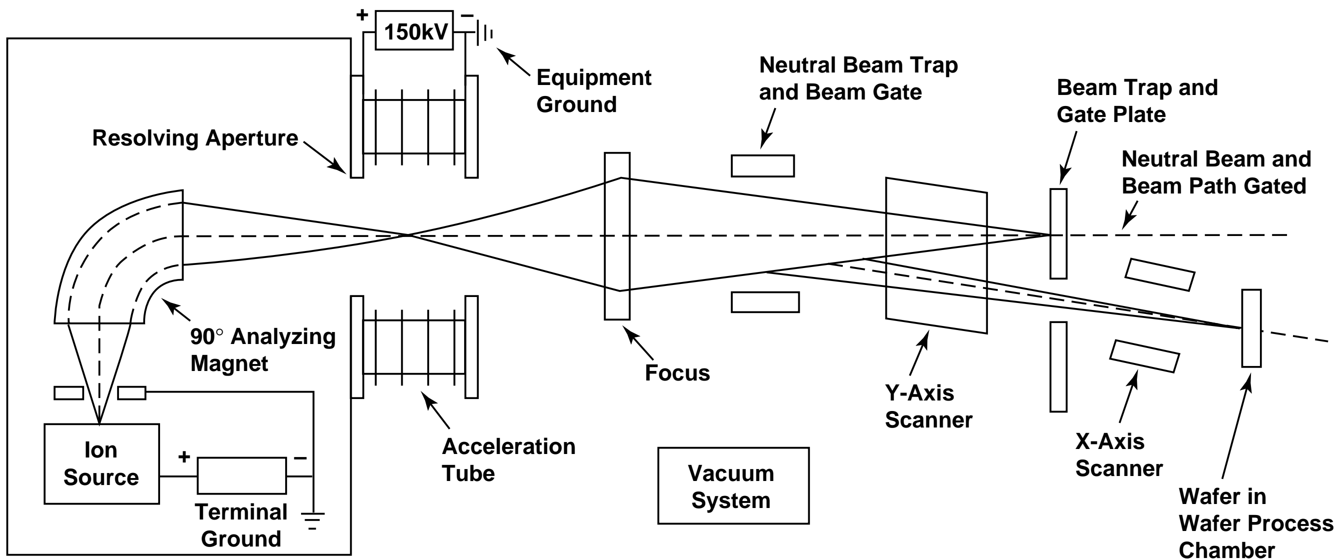


Figure 2-57. Configuration of a Typical 200KeV Ion Implanter

The ion implanter is a large, complex, and expensive piece of fab equipment. It contains several vacuum systems, a gas distribution system, high-voltage power supplies, complex cooling systems, and computer-controlled electronic systems. For all of this, the fab gets a control of the dopant process that can not be obtained any other way. The ion implanter has played such an important role in semiconductor fabrication that MOS technology would probably not dominate the IC industry without it.

Once the doping species has been selected, the two variables that can be controlled are the dose (the number of ions that strike the wafer surface per unit area) and the acceleration energy (which determines the depth of the implant into the wafer). The dose is controlled by counting the ions as they pass a detector, and the acceleration energy is controlled by changing the voltage on the accelerator tube. The ability to control both dose and energy provides a powerful and unique tool for semiconductor manufacturing.

The results of the implant process is a silicon lattice that has been disrupted and damaged. The path of an implanted ion is shown in Figure 2-58. A heat treatment (anneal) must be performed to reorder the silicon lattice back to monocrystalline and cause the implanted dopant atoms to be substituted for silicon atoms in the silicon crystal structure. Without this heat cycle no transistor (or IC) action would occur.

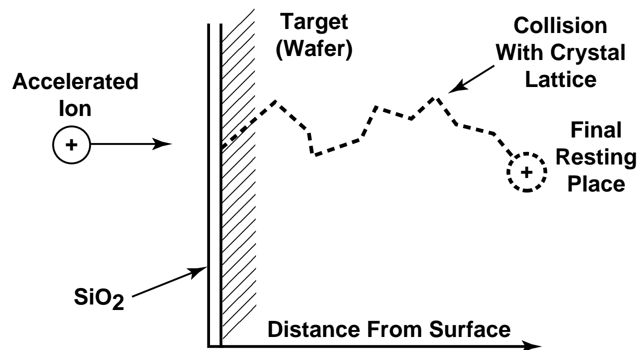


Figure 2-58. Path of an Implanted Ion

Evaluation of the implant process can be performed in a manner similar to that performed for the diffusion process. However, the test wafer used for sheet-resistance evaluation must be annealed before the sheet resistance can be measured.

Other evaluation alternatives are to use a special film on a glass disc and measure the optical properties of the film before and after implant, or use the therma-wave measurement system.

The therma-wave system measures the crystal damage of the implant and correlates the measurement to the implant dose and energy. This technique is the only measurement made immediately after the implant prior to the anneal. This provides immediate "feed-back" for process control.

The process engineer has more flexibility in masking ion implants than in diffusion. Diffusions are normally masked with oxides or oxide/nitride combinations. Implants can be masked by these, photoresists, and metal layers. Since implant is a low-temperature process, photoresist is a very common masking material. The engineer determines what thickness of resist is needed for a given implant species, dose, and energy. An example is shown in Figures 2-59 and 2-60.

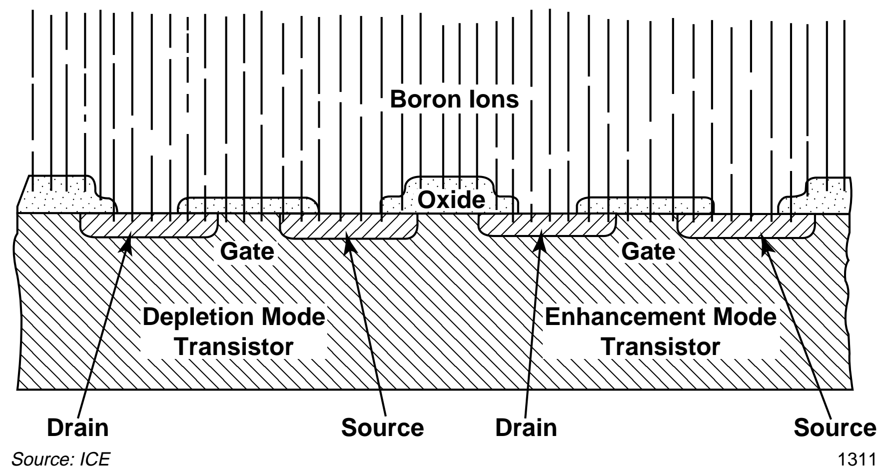


Figure 2-59. Step 1, The Adjustment of Thresholds

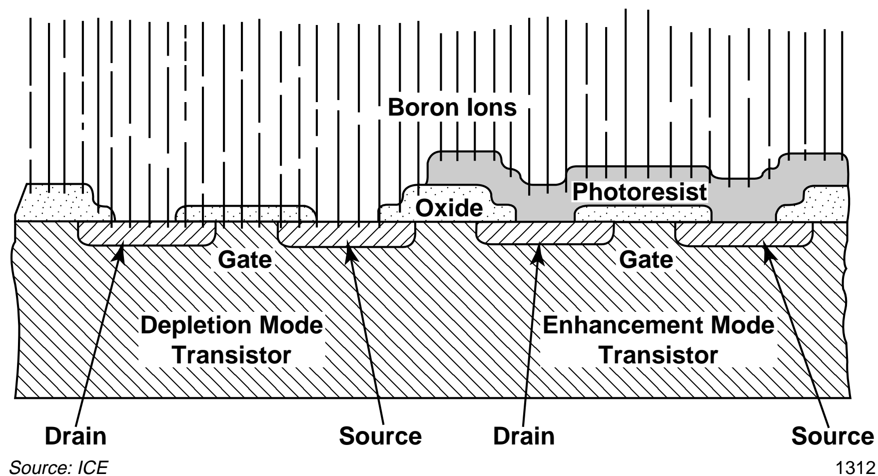


Figure 2-60. Step 2, Fabrication of the Depletion-Mode Transistors

Future uses of the implanter will expand it to other fab roles. In addition to the dopant role, the implanter can modify material characteristics, like gettering, etch rates, and grain boundaries to name a few.

F. EPITAXIAL DEPOSITION

The term epitaxial is derived from Greek, meaning to build upon. Epitaxial deposition, in general, is the deposition of a layer of single-crystal silicon on a single-crystal (monocrystalline)wafer. The deposited layer is a crystallographic extension of the substrate in terms of atomic order (i.e., it has the same crystal structure). Thus, the substrate could be considered the "seed" that is necessary to promote the single-crystal deposition.

Epitaxial deposition is a chemical vapor deposition (CVD) process. The subject of chemical vapor deposition was covered in the dielectric section (page 2-9). However, the original use of CVD started with the deposition of single-crystal silicon in the late 1950's and has played a major role in the industry since.

Epitaxy (or epi) has played a major role in the evolution of bipolar transistors and bipolar integrated circuits. In recent years both MOS discrete transistors and ICs have started to use epi as a key part of their structures.

Silicon CVD processes (epitaxial and polysilicon depositions) can be categorized by temperature range, pressure, and reactor wall temperature. The categories are summaries in Figure 2-61. In the case of polycrystalline silicon (poly, polysilicon) the crystallography of the substrate is not replicated during the deposition (e.g., poly on silicon dioxide). It is also possible that the deposition could yield an amorphorus film (e.g., silicon dioxide) that has no crystal structure. These types of films are shown in Figure 2-62.

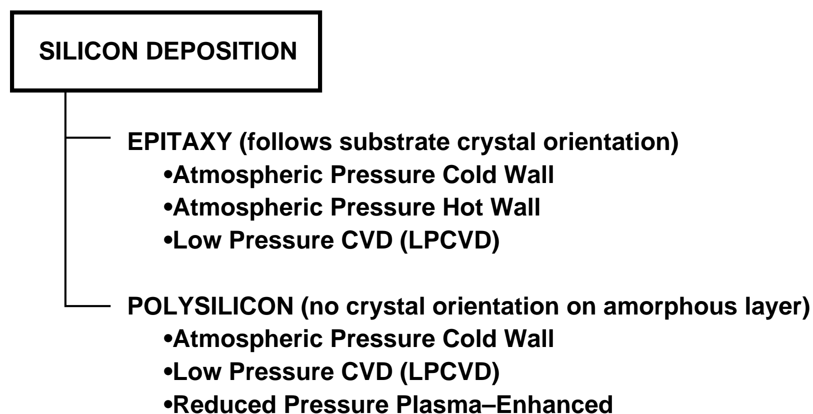


Figure 2-61. Silicon Chemical Vapor Deposition Processes

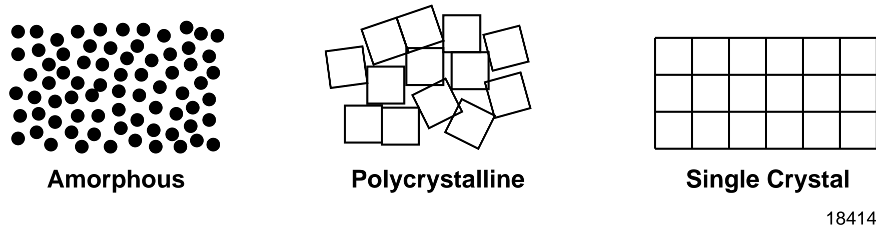


Figure 2-62. Types of Deposited Films

The primary purpose of the silicon epi process is to deposit high-quality, single-crystal films. During the deposition process dopant atoms are added to give the film the desired (N or P) electrical characteristics. The deposition time can be varied to control the thickness of the epi layer. The combination of the above provides the process with tremendous flexibility. Pictorial representations are shown in Figures 2-63 and 2-64.

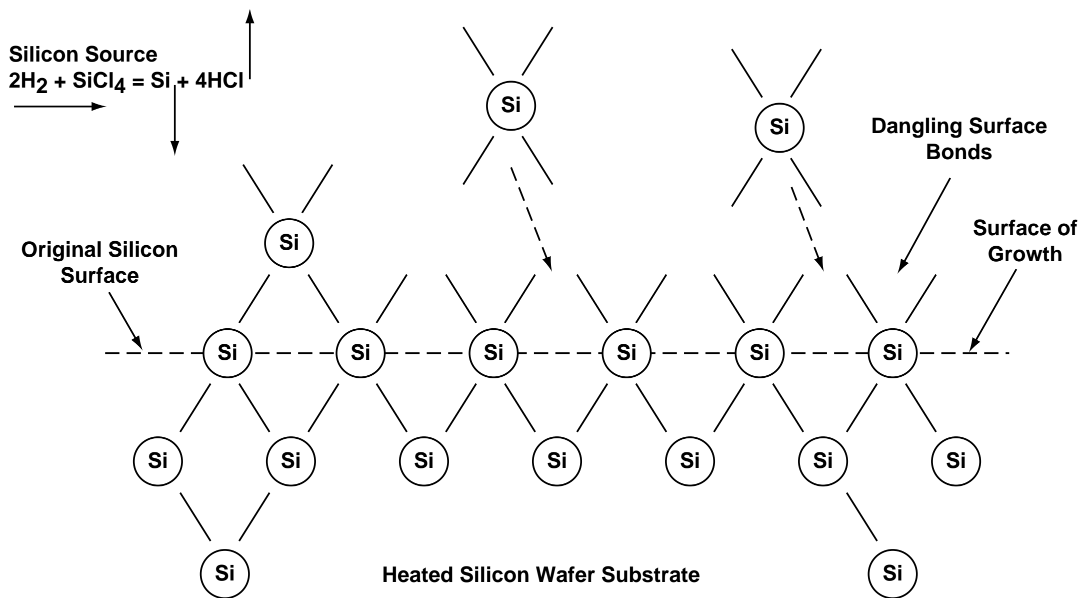


Figure 2-63. Pictorial Representation of Epitaxial Growth

Cleanliness of the epi process is critical to the quality of the film and the subsequent yields. This cleanliness must be considered in all aspects:

1. Wafer cleaning prior to the deposition process
2. Cleanliness (quality) of the chemicals used for the cleaning
3. Cleanliness of the gases going into the reaction chamber
4. Cleanliness of the inside of the reaction chamber
5. Cleanliness of the environment in the epi area
6. Cleanliness of the wafer handling.

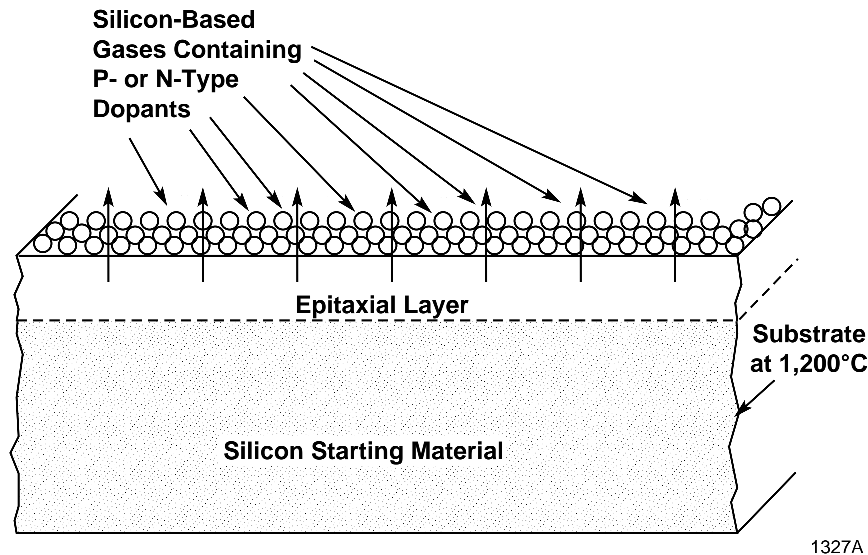


Figure 2-64. Epitaxial Growth

The evaluation of the epi layer measures the following.

1. Thickness
2. Resistivity
3. Surface quality
4. Crystallographic defects
5. Carrier lifetime
6. Dopant profile.

Thickness is measured with FTIR (Fourier Transform Infrared Spectroscopy) and resistivity is measured on a test wafer with the 4-pt. probe. Carrier lifetime and dopant profile are measured by C-V techniques. Surface quality and crystallographic defects are inspected by microscope.

The results of the epi process provide a wide range of epi thicknesses and resistivities on a variety of substrates. In addition, the deposited layer can be reasonably doped homogenously (evenly) as the deposition is done. For some devices this is an advantage over diffusion or implanting because it gives a graded doping.

The thickness can range from 0.6µm for high-performance devices to several hundred microns for discrete power products. The application of the product dictates the epi requirements.

G. POLYSILICON DEPOSITION

Polysilicon deposition is another application of CVD technology for thin films. Unlike epi, polysilicon is not required to follow the crystal orientation of the substrate (e.g., poly on silicon dioxide, which is amorphous).

Thin-film polysilicon has many important applications in the semiconductor industry. Polysilicon was the key ingredient leading to the "self aligned" MOS technology. Heavily-doped polysilicon has become the most widely used gate-electrode material for MOS products, both discretes and ICs. In addition, it serves as an interconnect, capacitor plate(s), doping source, and can be oxidized to form a stable layer of SiO_2. Polysilicon is utilized in these roles because of its compatibility with subsequent high-temperature processing, its excellent interface with SiO_2, its high reliability as a gate electrode material, and its ability to be deposited over steep topography with good conformal coverage. Lightly-doped polysilicon films are used as resistors in static memory products and to fill trenches in DRAMs.

Thin films of polysilicon are made up of small single-crystal grains of about 1000Å separated by grain boundaries. The film that will be a polysilicon layer can be either amorphorus or polycrystalline "as deposited." Subsequent heat cycles at elevated temperatures will cause an amorphous film to become polycrystalline. This "as-deposited" undoped film has an extremely high resistivity.

The resistivity of poly can be changed by doping the film with phorphous, boron, or arsenic. Because of the variation in grain size and grain boundary effects, heavily doping the film only reduces the sheet resistance to 10-30 Ω/❑ (ohms/sq.). In between these extremes, ion implantation can control the resistivity for various resistor values or doping applications.

- **LPCVD MID-TEMPERATURE (600 - 700°C)**
  - **— Vertical-Stacked Wafers**
  - **— Resistance Heating**
  - **— Pressure (100 - 500 Millitorr)**
  - **— In-Situ Doping**
- **PECVD LOW-TEMPERATURE (LESS THAN 300°C)**
  - **— Horizontal Wafer Position**
  - **— Vertical-Stacked Wafers**
- **ADVANTAGES OF LOW PRESSURE PROCESSING**
  - **— Uniformity**
  - **— Lower Temperature**
  - **— Less Wafer Warpage**
  - **— Lower Cost**
- **DISADVANTAGES**
  - **— Pyrophoric Gas**
  - **— Toxic Doping Gases**
  - **— Exhaust Scrubbing**

9525C

Figure 2-65. Low-Pressure Polysilicon Deposition

The deposition of polysilicon is generally done using LPCVD (Low-Pressure CVD) equipment operating in the 600-630°C range. This allows reasonable load size per run and acceptable productivity. Silane (SiH_4) is the silicon source material.

PECVD (Plasma-Enhanced CVD) reactors emerged in the late '80's and early '90's. PECVD systems can be operated at lower temperatures making them attractive for submicron technology. Figure 2-65 summarizes LPCVD deposition. Figure 2-66 illustrates the concept of the system.

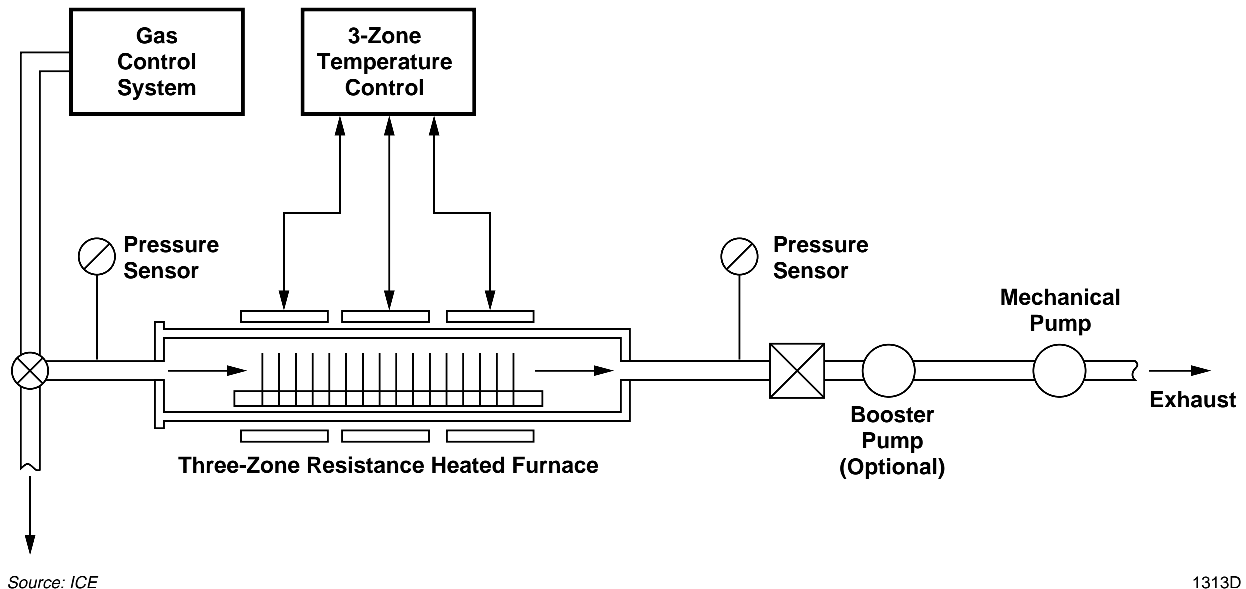


Figure 2-66. Block Diagram of a Low-Pressure Chemical Vapor Deposition System

Earlier, it was indicated the polysilicon could be doped to control the resistivity. The doping process for polysilicon can be a separate process after the deposition or it can be incorporated into the deposition. There are advantages and limitations to each process technique.

H. METAL DEPOSITION

Metal (usually aluminum or aluminum with a small amount of other materials) is used to connect the individual components (diodes, transistors, resistors, and capacitors) in an integrated circuit. A layer of metal is deposited on the entire top surface of a wafer. Through photolithography and etch, selected portions of the metal are removed. The remaining aluminum serves as the conductors (wires) between the various components of each IC.

Noyce's revolutionary concept of wiring all the transistors at the top surface of the silicon rather than having individual discrete dice on a substrate altered the direction of the semiconductor industry. The prior art of the Kilby patent integrated components on a common substrate but still used loops of very fine wire to interconnect the components. Noyce's new method of wiring circuits all within the die area of the circuit opened up endless new possibilities. This section will cover the metal technology used for interconnects.

1. Process Sequence

The metallization process sequence is shown in Figure 2-67. This figure illustrates that an earlier photo-patterning process has removed the silicon dioxide from the areas where ohmic contact will be required. A layer of metal is deposited over the surface and the photo-patterning process is repeated to delineate the metal interconnect. This simplification of the metallization process shows the genius of Noyce. The process concept is the same for interconnecting two transistors or seventy million transistors. Further, this fundamental concept can be extended to more than one level of interconnect as will be shown later.

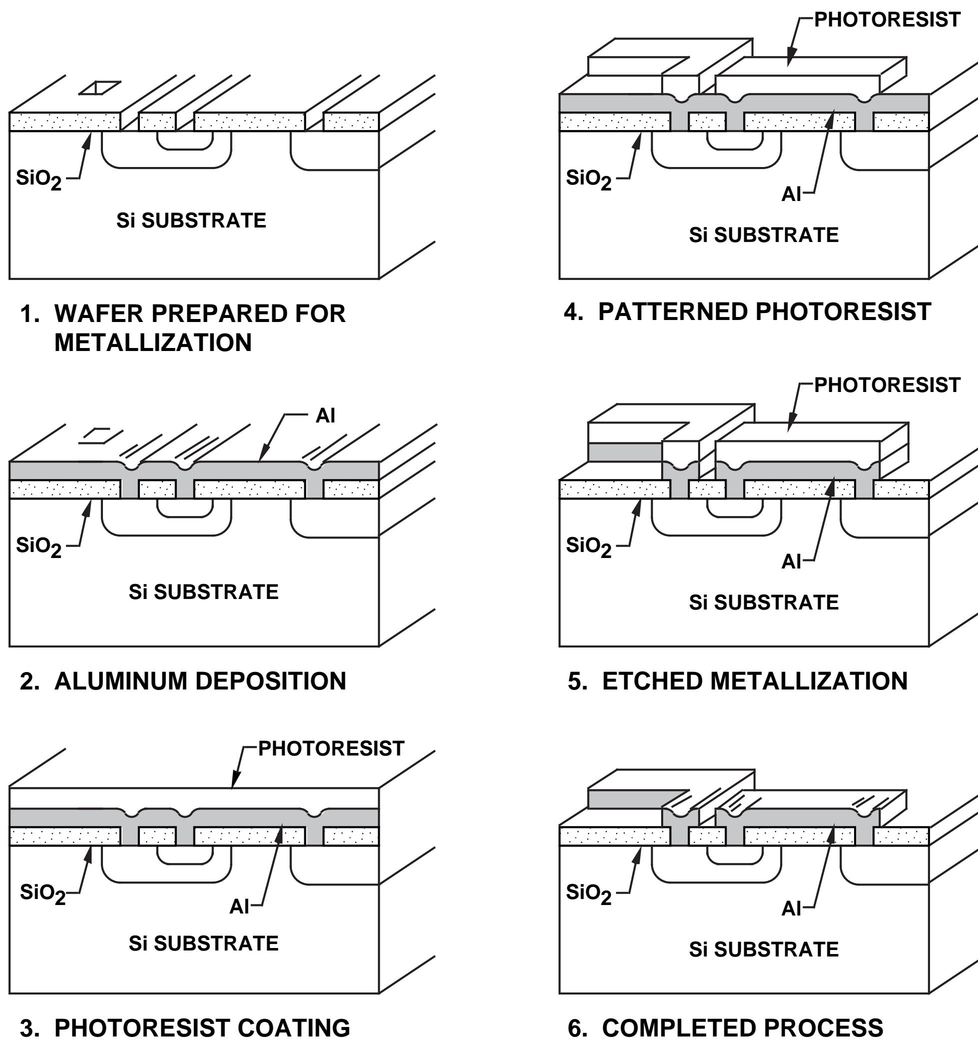


Figure 2-67. Metallization Process Sequence

The metal deposition process flow is shown in Figure 2-68. The wafer(s) receive a wet chemical clean just prior to loading into the deposition system. The deposition system is equipped with a vacuum system for controlling the internal partial pressure and system electronics to control the deposition process. The resulting film is measured for film thickness, uniformity, and step coverage. Figure 2-69 diagrams a metallization system.

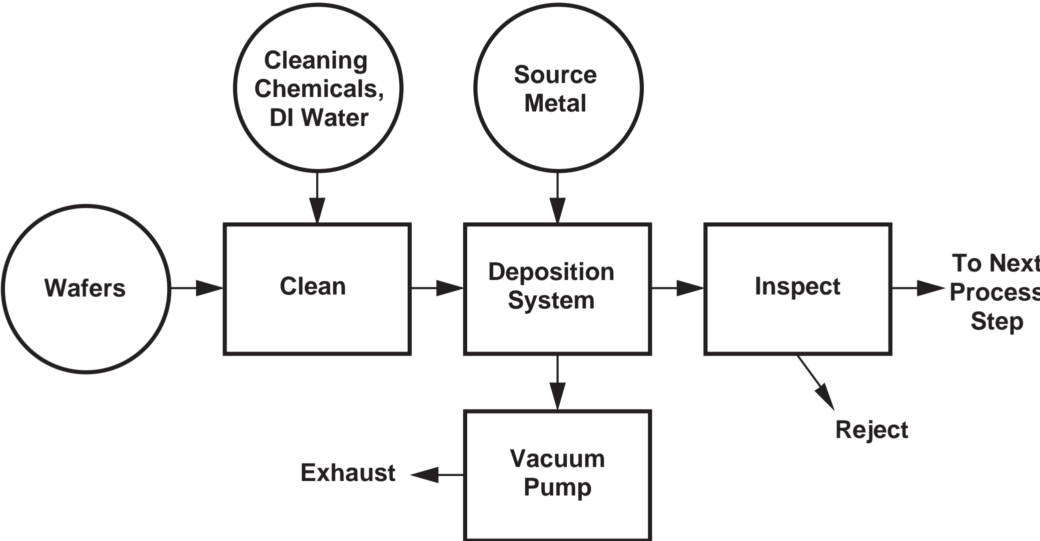


Figure 2-68. Metal Deposition Process Flow Chart

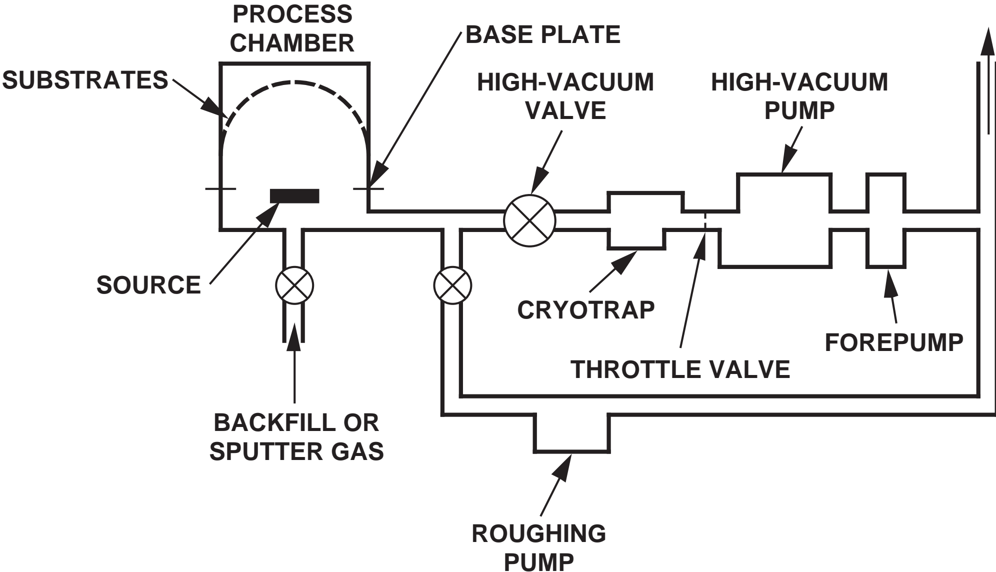


Figure 2-69. Metallization System

2. Materials

The basic materials requirements to interconnect the components within an IC are listed in Figure 2-70. Unfortunately, no individual metal can meet all of the requirements. Aluminum is the most often used metal but its use is being challenged by newer circuit requirements.

- **Low electrical resistivity**
- **Ohmic and low contact resistance**
- **Stable contact formation to silicon or other metal(s)**
- **High temperature stability**
- **Excellent adhesion and low stress**
- **Good electromigration resistance**
- **Good corrosion resistance**
- **Controlled oxidation properties and stability in an oxidizing environment**
- **Ease of formation**
- **Ease of fine line pattern transfer**
- **Smooth surface features**

18419

Figure 2-70. Conductor Material Properties Requirements for VLSI

When pure aluminum is used to interconnect shallow p-n junction devices, a reliability problem can occur. The silicon atoms in the substrate diffuse into the aluminum metallization. Over extended time and temperature, sufficient silicon is depleted to cause a metal short through the p-n junction. This concept is illustrated in Figure 2-71. The problem can be minimized by adding a small percentage of silicon into the aluminum.

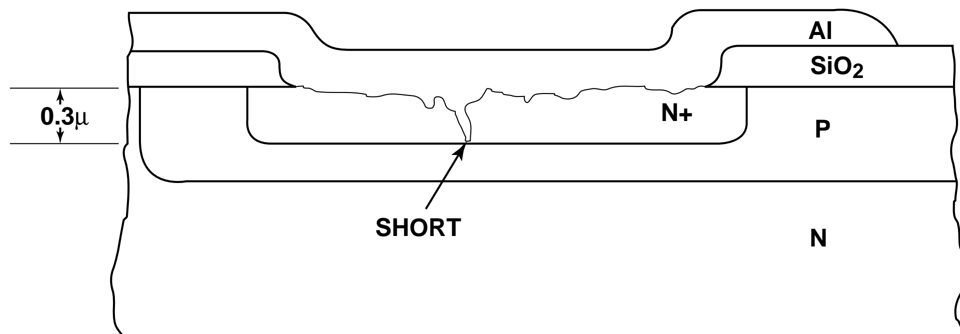


Figure 2-71. Aluminum/Silicon Dissolution

3. Methods

a. E-beam, Filament Evaporation

The methods for depositing metal films are tabulated in Figure 2-72. The semiconductor industry has evolved through each of the methods except Ion Beam deposition, which has not been implemented to any great extent yet. Sputtering is the most commonly used method while LPCVD/PECVD is being used for selected films of tungsten. Diagrams for filament evaporation and electron-beam (e-beam) are shown in Figure 2-73 and Figure 2-74, respectively.

Evaporation
- **Hot Filaments**
- **Resistive Filaments, Heated Crucible**
- **R.F. Heated Crucible**
- **Electron-Beam**

Sputtering
- **Planar Diode**
- **Planar Triode**
- **Planar Magnetron (Diode or Triode)**
- **R.F. Diode**

Chemical Vapor Deposition
- **LPCVD**
- **PECVD**

Ion Beam Deposition

18420

Figure 2-72. Methods of Metal Deposition

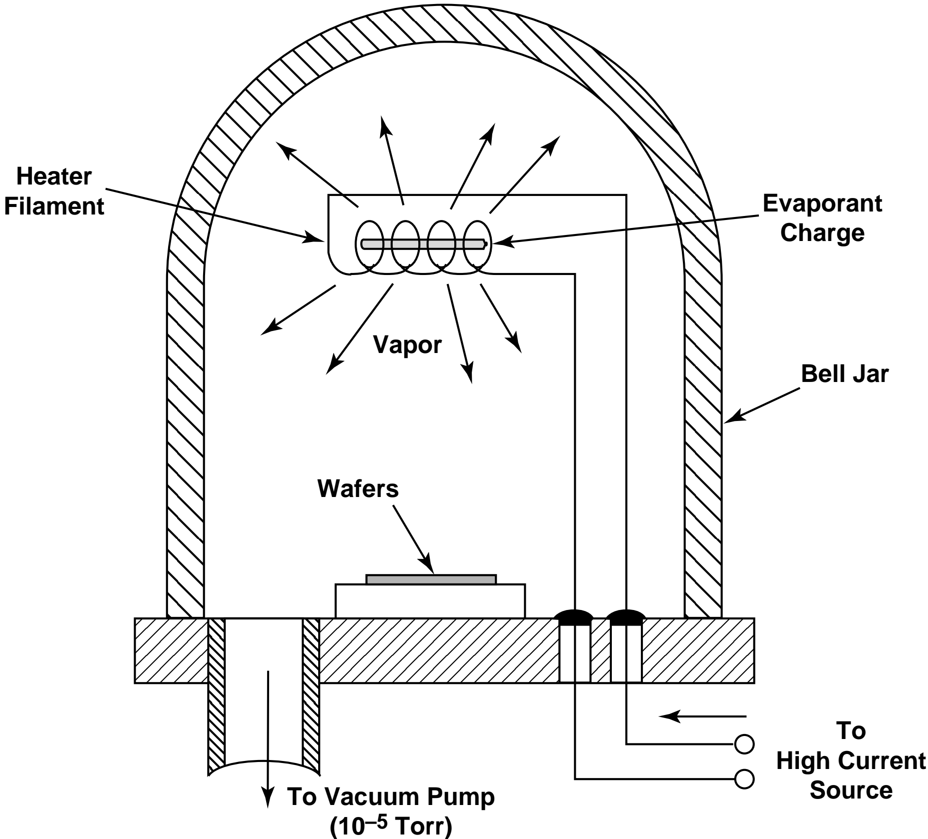


Figure 2-73. Vacuum Evaporation System

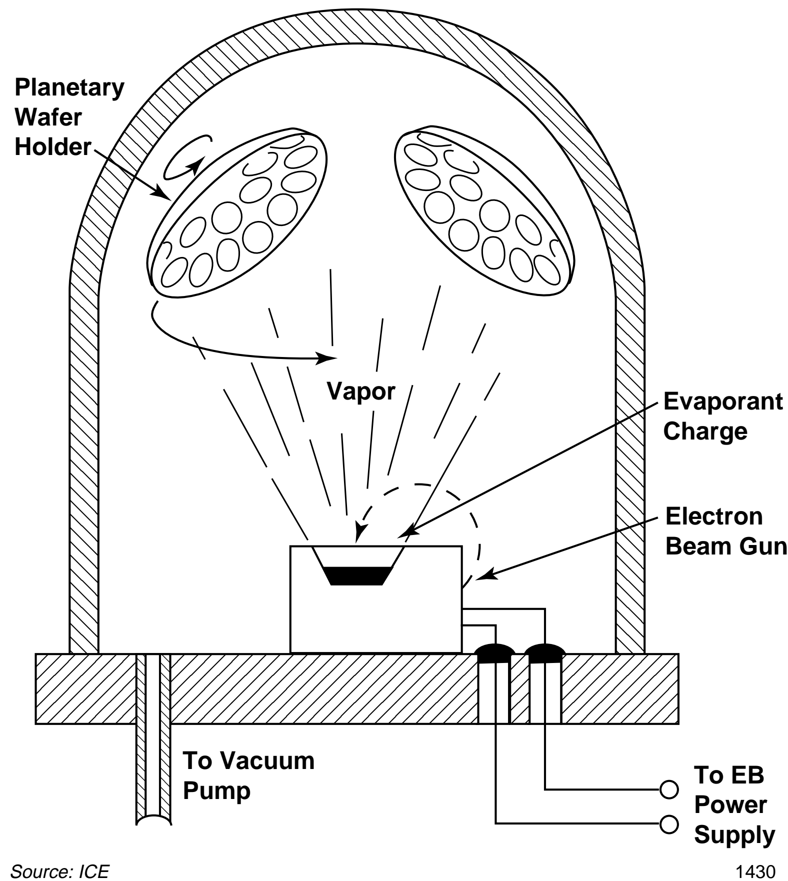


Figure 2-74. Electron Beam Evaporation

The electron-beam system requires a separate e-gun for each material to co-deposit the film. This increases the cost of the deposition system.

b. Sputtering

Low-pressure sputtering evolved as an improved alternative to the e-beam deposition system. A sputtering system is diagrammed in Figure 2-75. The main advantage of sputtering is the physical nature of the process. Argon is the ionized atom most commonly used to bombard the target material. This physical bombardment causes the target material to be deposited from many different angles, which provides a more uniform deposition, regardless of target composition. The resulting step coverage is improved as shown in Figure 2-76.

Many refinements have been made on the hardware of sputtering deposition systems. These improvements have made it the method of choice for the majority of wafer fabs. Various ways of configuring the sputtering system were noted in Figure 2-72. One of the more popular concepts is shown in Figure 2-77.

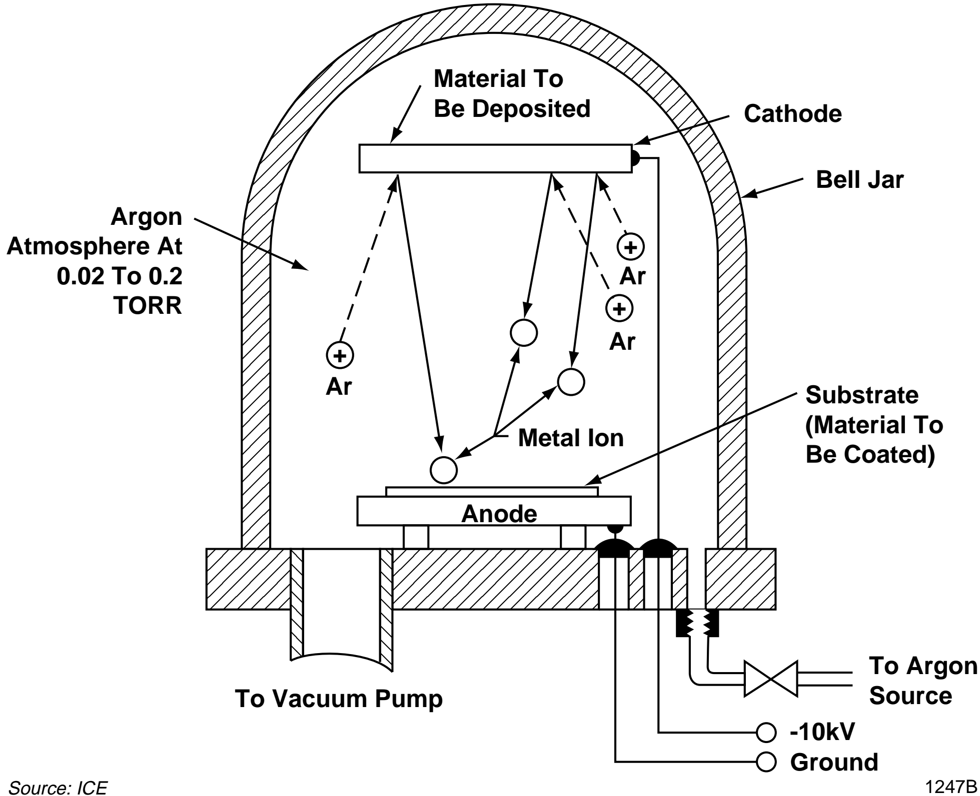


Figure 2-75. Low-Pressure Sputtering

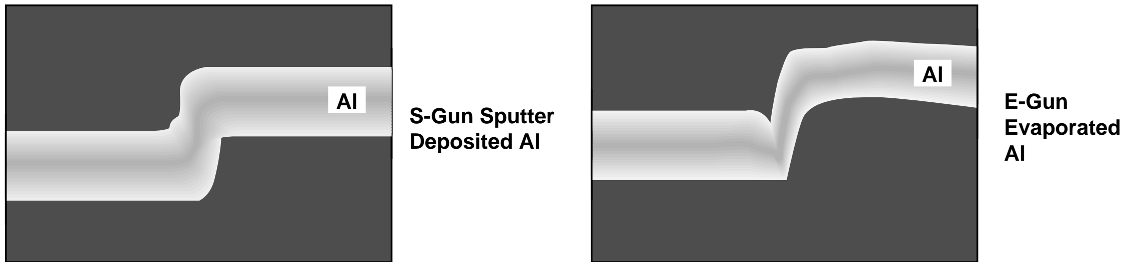


Figure 2-76. Aluminum Step Coverage Comparison

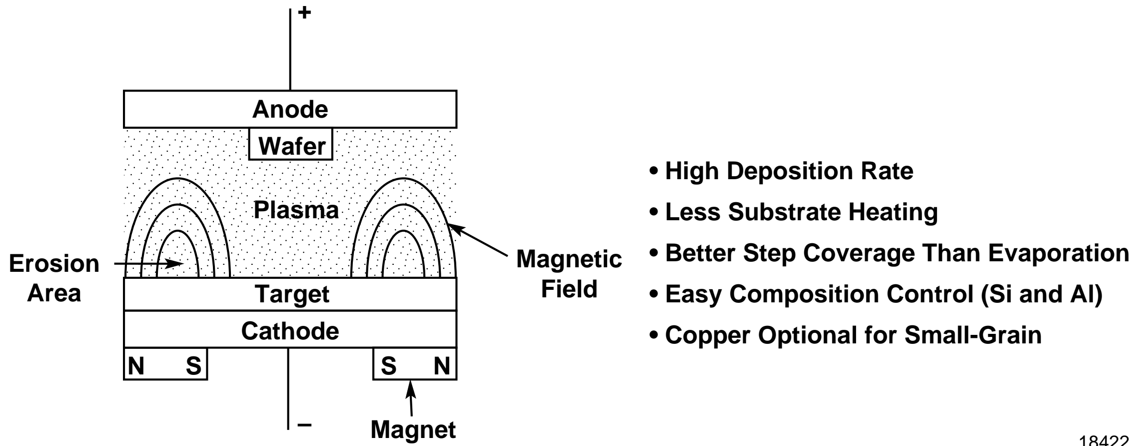
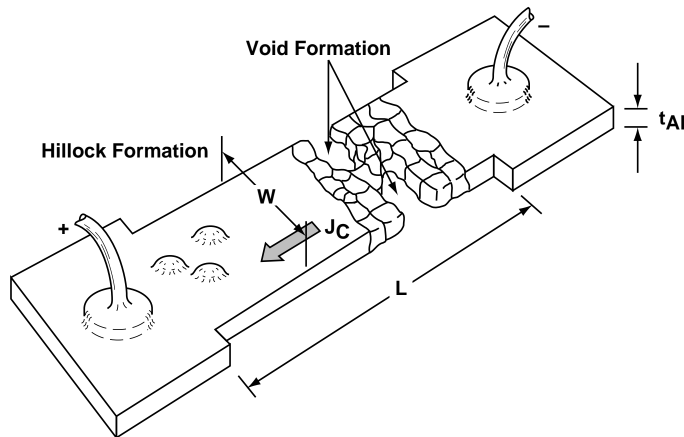


Figure 2-77. Planar DC Magnetron Sputtering

4. Electromigration

Another reliability problem associated with the interconnect metallization is electromigration (similar to blowing a fuse, see glossary for definition). The details are beyond the scope of this book but the general concept is illustrated in Figure 2-78. From the drawing it can be observed the problem is sensitive to the thickness of the film. Thickness uniformity over topological changes is known as step coverage. Poor step coverage contributes to the problem of electromigration. To further reduce electromigration, a small percentage of copper is added to the silicon-aluminum metal. Semiconductor manufacturers evaluate their products on an "on-going" basis for this characteristic.



Variables

- J ~10^5 amps/cm^2 (Al), ~10^6 amp/cm^2 (Al-4% Cu)
- t_{Al} = 6000Å to 12,000Å (0.6μm to 1.2μm)
- W = 1μm to 50μm
- Grain size 0.3 to 8μm
- With and without surface passivation

Source: Technology Associates 15076A

Figure 2-78. Electromigration Description

5. Alloy (Sinter)

The successful patterning of the interconnect metal does not guarantee that adequate ohmic contact has been made to the silicon. An "alloy" or "sintering" process step is usually used to insure a low-resistance contact between the interconnect metal and the silicon. The alloy step is done in a furnace at a temperature in the range of 400-475°C. The alloy temperature, time and ambient will vary from process to process, but the limits are determined by the metal composition and junction depth considerations in conjunction with the metallurgical phase-diagram. Sometimes this alloy step is referred to as an anneal step.

6. Multilayer Interconnect

As the integration density of ICs has increased, interconnecting the various components has an impact on the die size (interconnect takes up space). To realize the maximum number of interconnects per unit area, multiple layers of interconnects are used. This is illustrated in Figure 2-79.

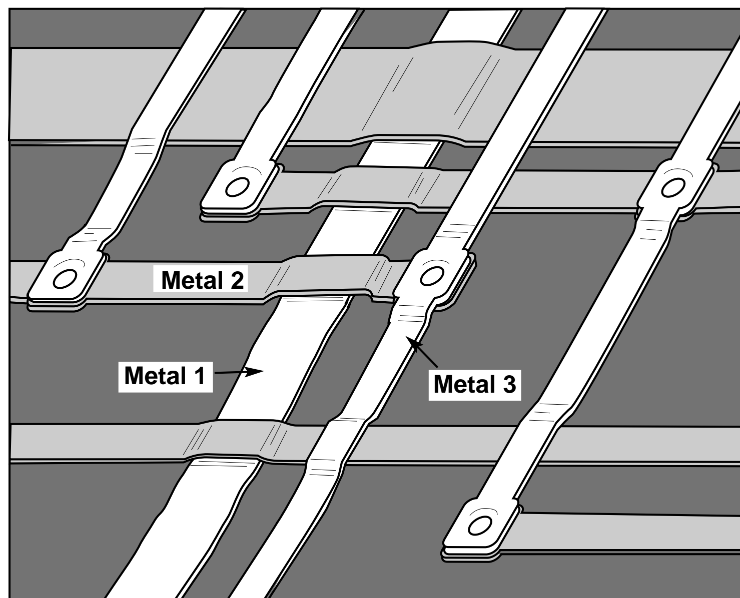


Figure 2-79. Three-Layer Metallization on IC Chip

Multilayer interconnect has placed new demands on the metallization capabilities of existing processes. Research has produced several alternative materials, processes, and equipment.

a. Barrier Metals

The use of a barrier metal at the silicon-to-metal and metal-to-metal interfaces has reduced the junction shorting/leakage problem and improved ohmic contact resistance consistency. Barrier metals are also used on top of the metal layers to assist the various considerations in patterning and etch. Further, the barrier improves the interface between layer 1 and layer 2, layer 2 and layer 3, etc. A barrier metal is illustrated in Figure 2-80.

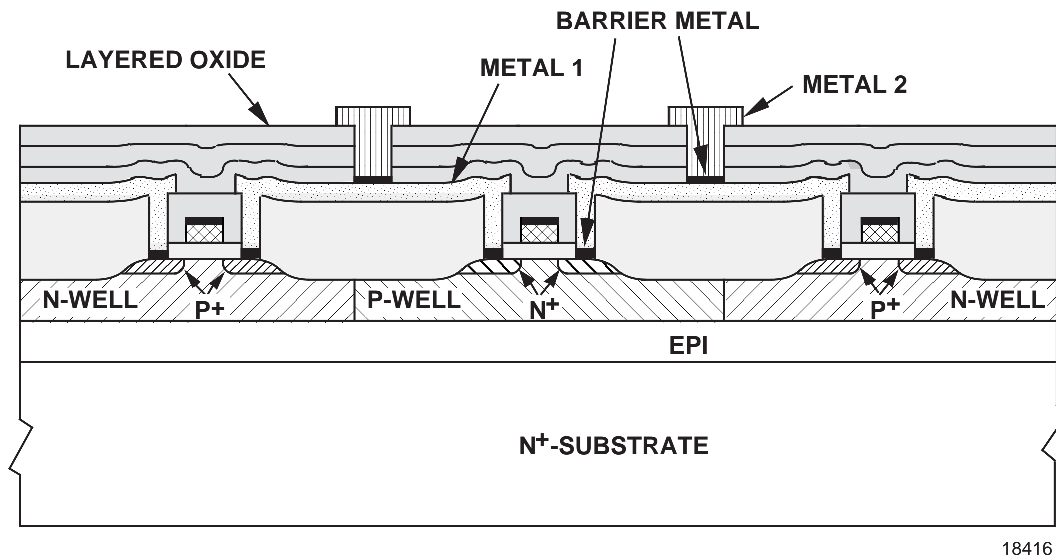


Figure 2-80. Double-Layer Metal With Barriers

b. Vias

In conjunction with barrier metal technology, the capability to fill an ohmic contact area with another metal was developed. Frequently this fill metal is tungsten. Tungsten can be deposited to conform to the topological changes and it is easy to dry etch. This is illustrated in Figures 2-81 and 2-82.

c. Dielectric Isolation

Another part of the multilayer metal technology is the dielectric material used to separate the layers of metallizations. Various types of CVD and spin-on dielectrics are used. To smooth out the topography of the surface of these dielectrics, various etch backs and mechanical polishing are used.

Generally, after the first layer of metal has been completed, the "alloy" or "sintering" process step is done and the wafers' test patterns are tested.

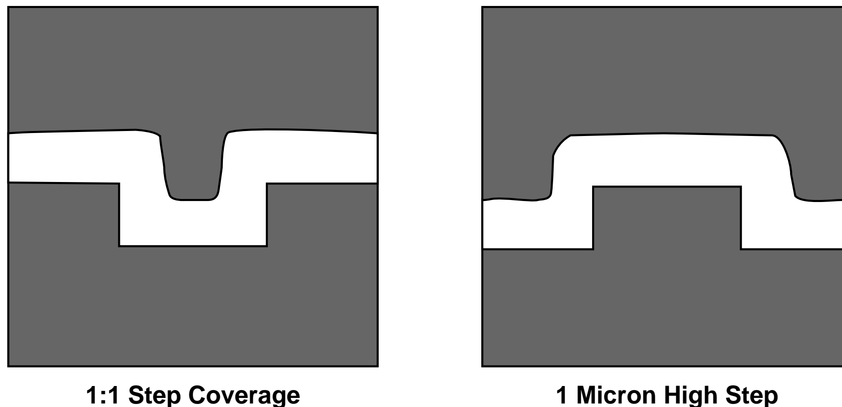


Figure 2-81. CVD Tungsten Thin Film

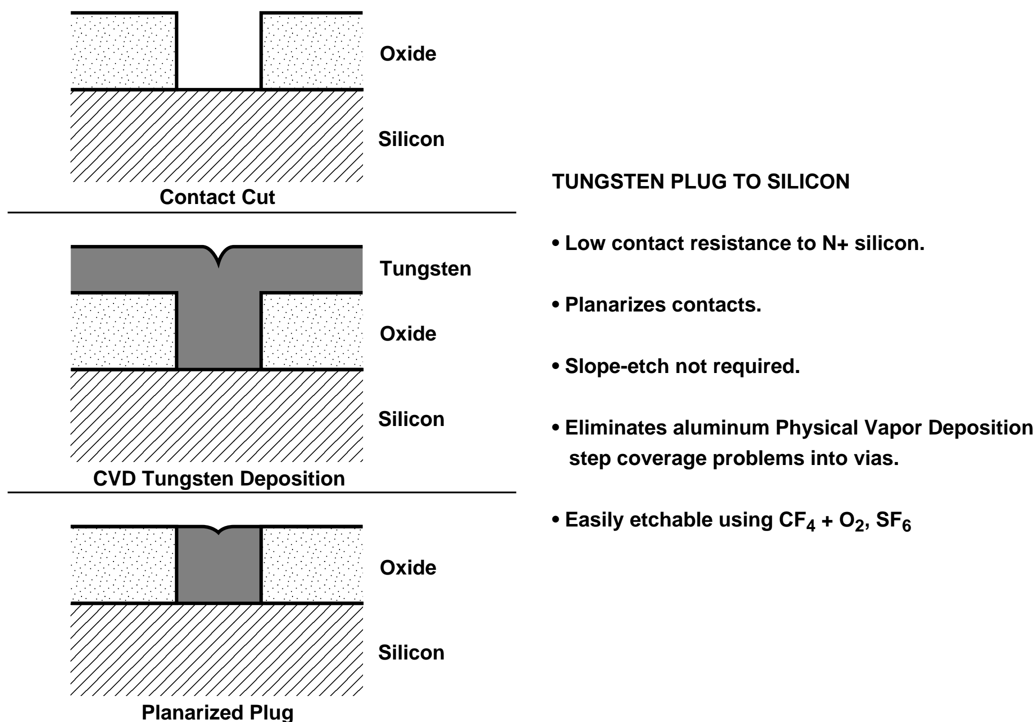


Figure 2-82. For Via Fill to Silicon

I. PROCESSES

1. CMOS

Figures 2-83, 2-84, and 2-85 provide a graphical representation of the IC manufacturing cycle. Figures 2-86, 2-87, and 2-88 provide graphical cross-sections of the CMOS structure as the circuit travels through the manufacturing cycle. Each group of process steps are narrated relative to Figures 2-83, 2-84, and 2-85.

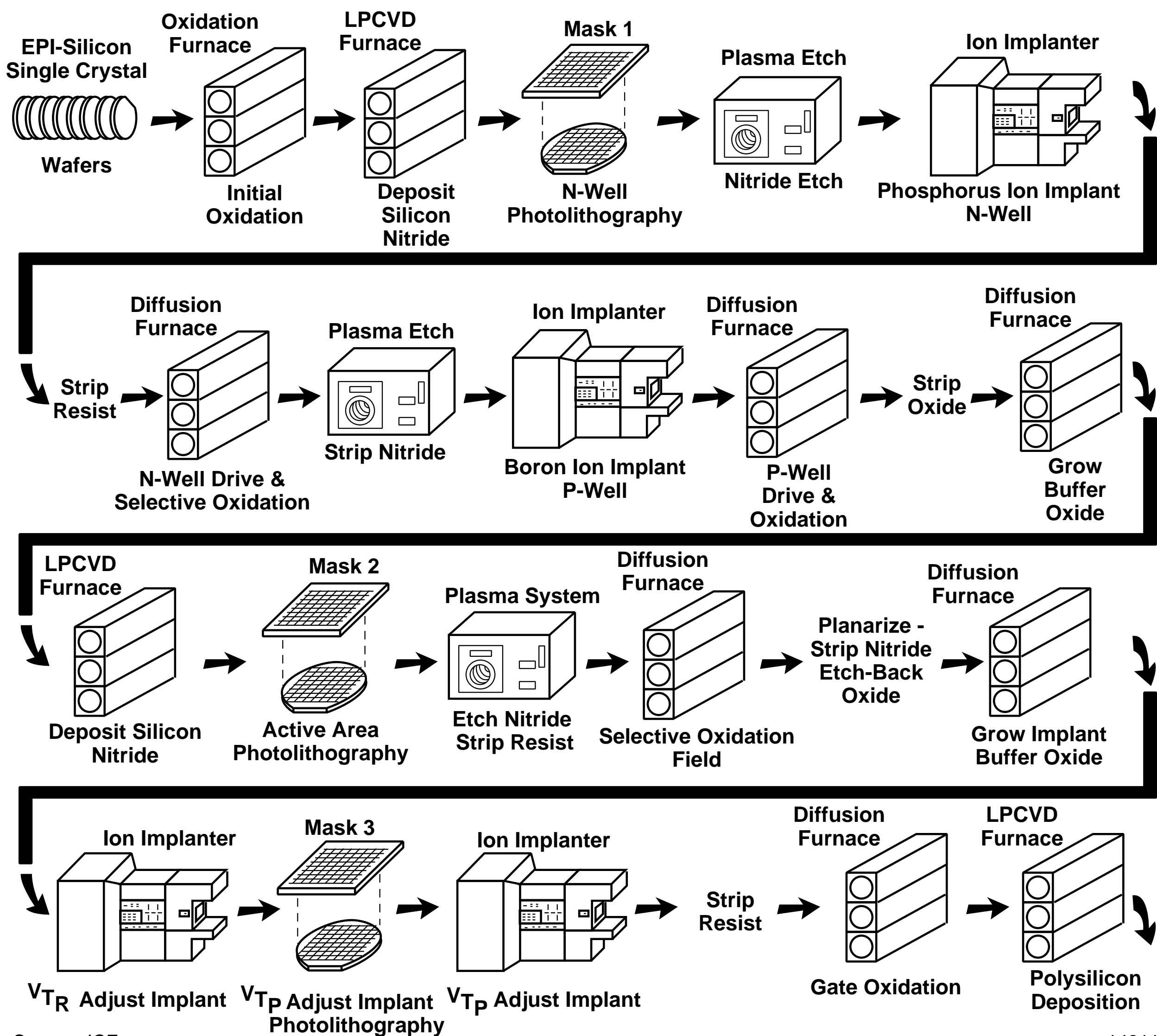
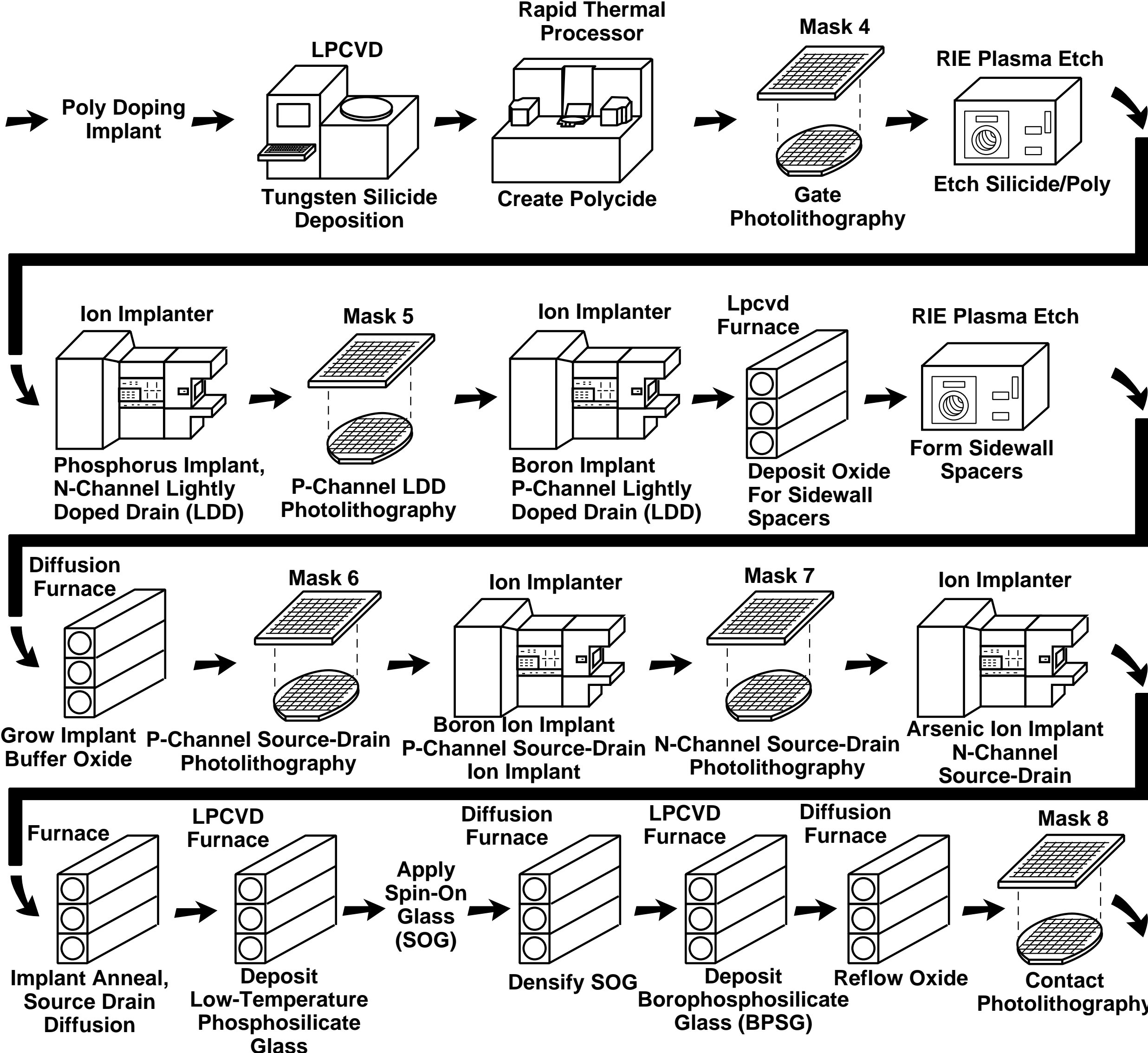


Figure 2-83. Twin-Well Silicon-Gate CMOS Manufacturing Sequence (1 of 3)



Source: *ICE* 14812

Figure 2-84. Twin-Well Silicon-Gate CMOS Manufacturing Sequence (2 of 3)

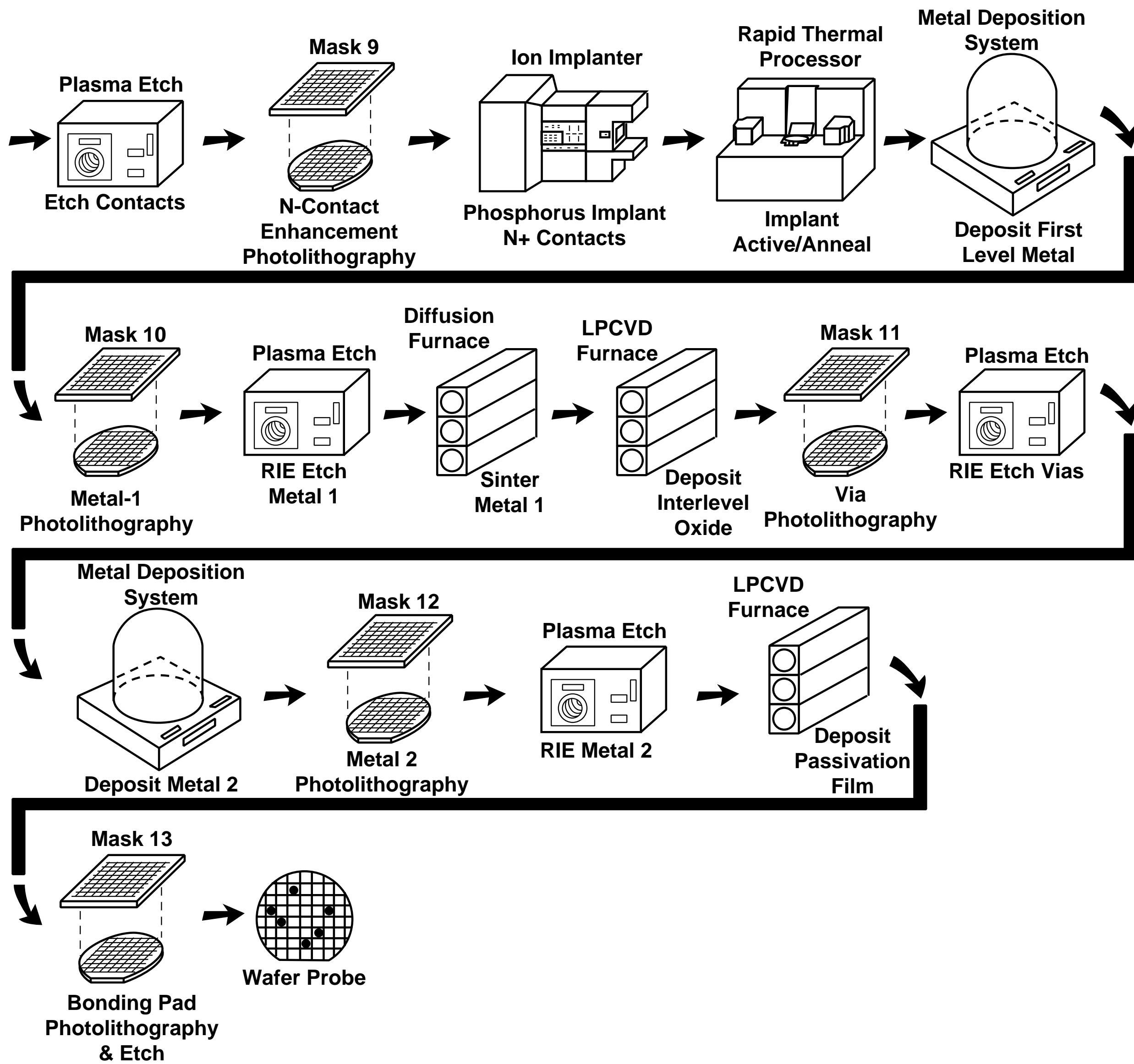
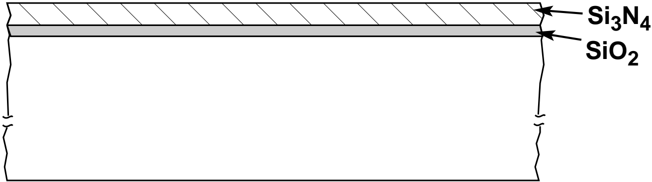
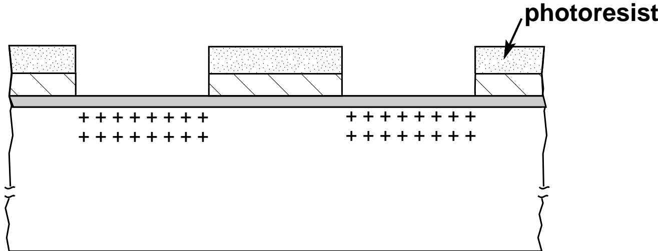


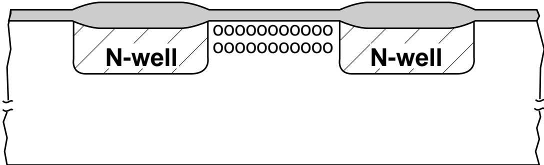
Figure 2-85. Twin-Well Silicon-Gate CMOS Manufacturing Sequence (3 of 3)



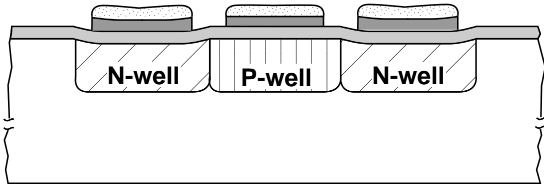
1. Starting material with SiO_2 and Si_3N_4.



2. N-well lithography, etch, phosphorus implant, resist strip, N-well drive and oxidation, strip nitride



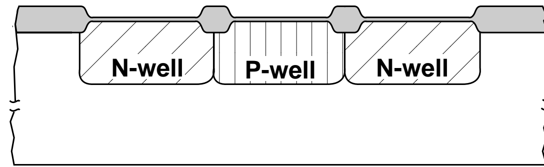
3. Boron P-well implant, P-well drive/oxidation, strip all oxide, grow thin oxide and deposit nitride



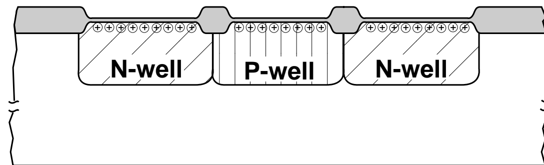
4. Active area lithography and etch nitride, strip photoresist

17950

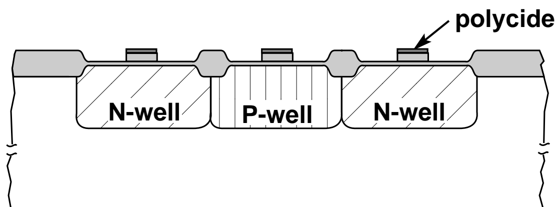
Figure 2-86. Twin-Well Silicon-Gate CMOS Process Flow (1 of 3)



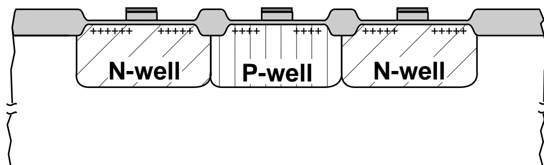
5. Grow field oxide, strip nitride, oxide etch back, grow thin oxide



6. V_{TN} lithography, V_{TN} implant, strip resist, V_{TP} lithography, V_{TP} implant strip resist, RTA implants, strip thin oxide



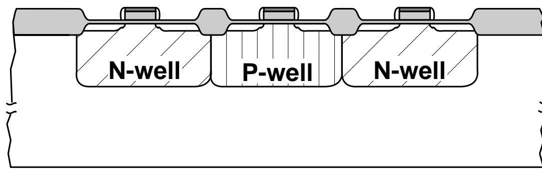
7. Grow gate oxide, deposit polysilicon, implant polysilicon with arsenic, deposit tungsten silicide, anneal, gate lithography, etch polycide, strip resist



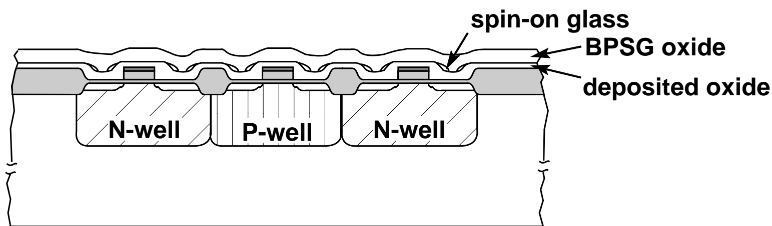
8. LDD N^-channel lithography, LDD N^-implant, strip resist, LDD P^-channel lithography, LDD P^-implant, RTA implants

17951

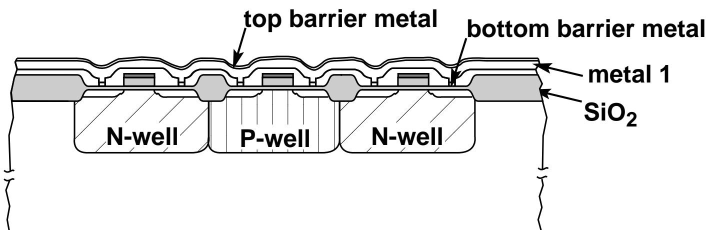
Figure 2-87. Twin-Well Silicon-Gate CMOS Process Flow (2 of 3)



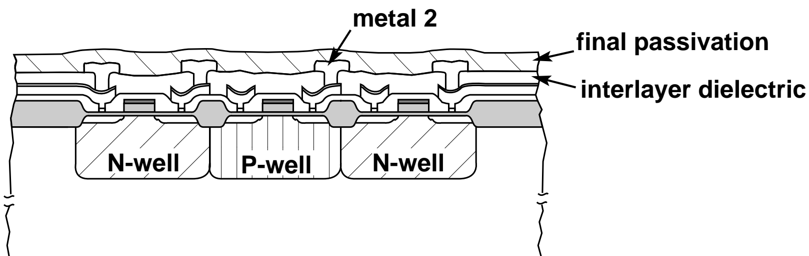
9. Deposit spacer oxide, RIE etch back spacer oxide, N^+ S/D lithography, N^+ S/D implant, strip resist, P^+ lithography, P^+ S/D implant, strip resist, anneal implants



10. Deposit oxide, apply spin-on glass, deposit BPSG oxide, reflow oxides, etch back reflow oxide



11. Contact lithography, bottom barrier metal deposition, aluminum deposition, top barrier metal deposition



12. Metal 1 lithography, deposit interlayer dielectric, via lithography, deposit metal 2, metal 2 lithography, anneal, deposit final passivation

17952

Figure 2-88. Twin-Well Silicon-Gate CMOS Prcoess Flow (3 of 3)

2. NMOS

A simple NMOS IC process is shown in Figure 2-89.

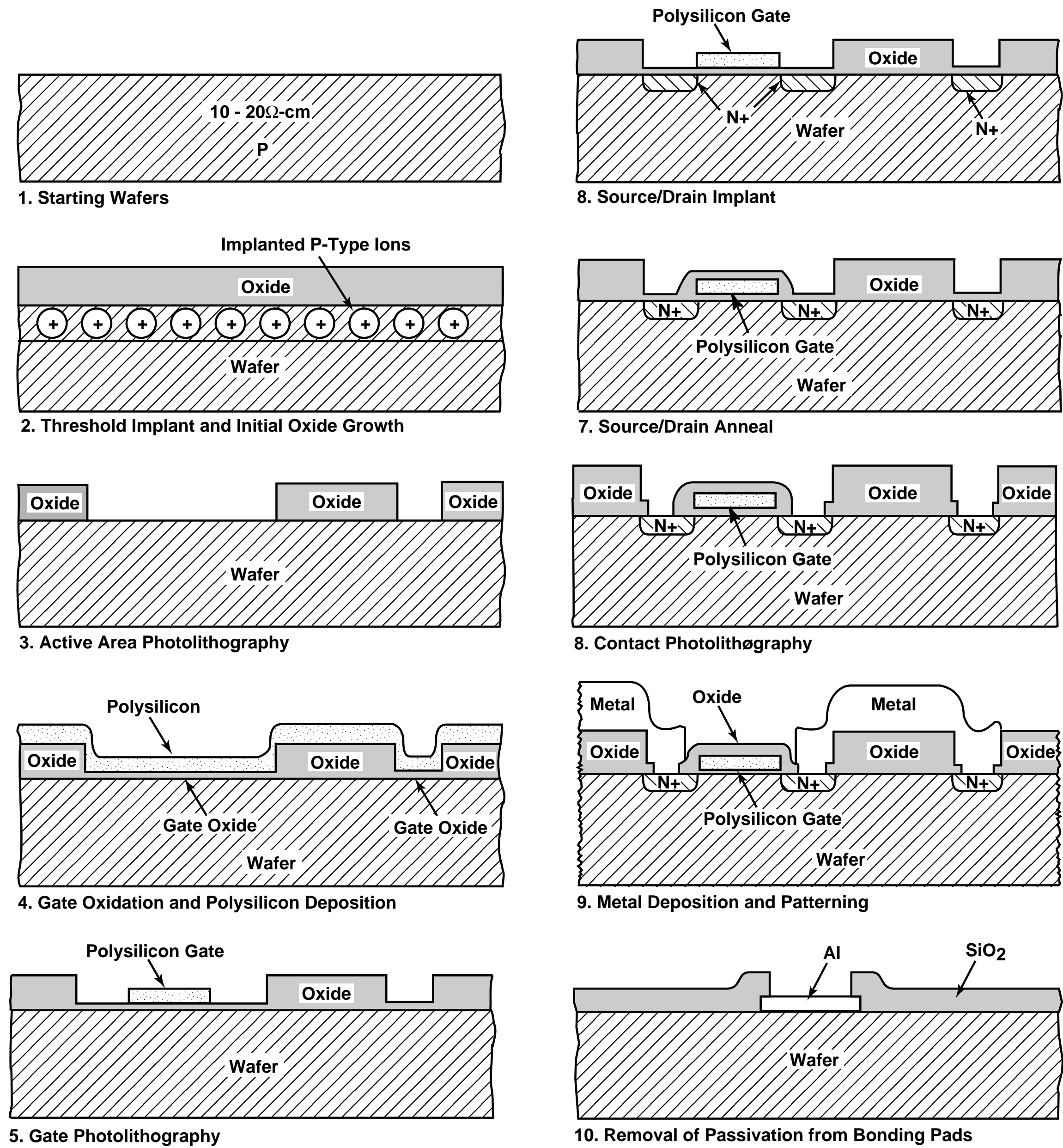


Figure 2-89. NMOS Process

3. Bipolar

A simple bipolar process is shown in Figure 2-90.

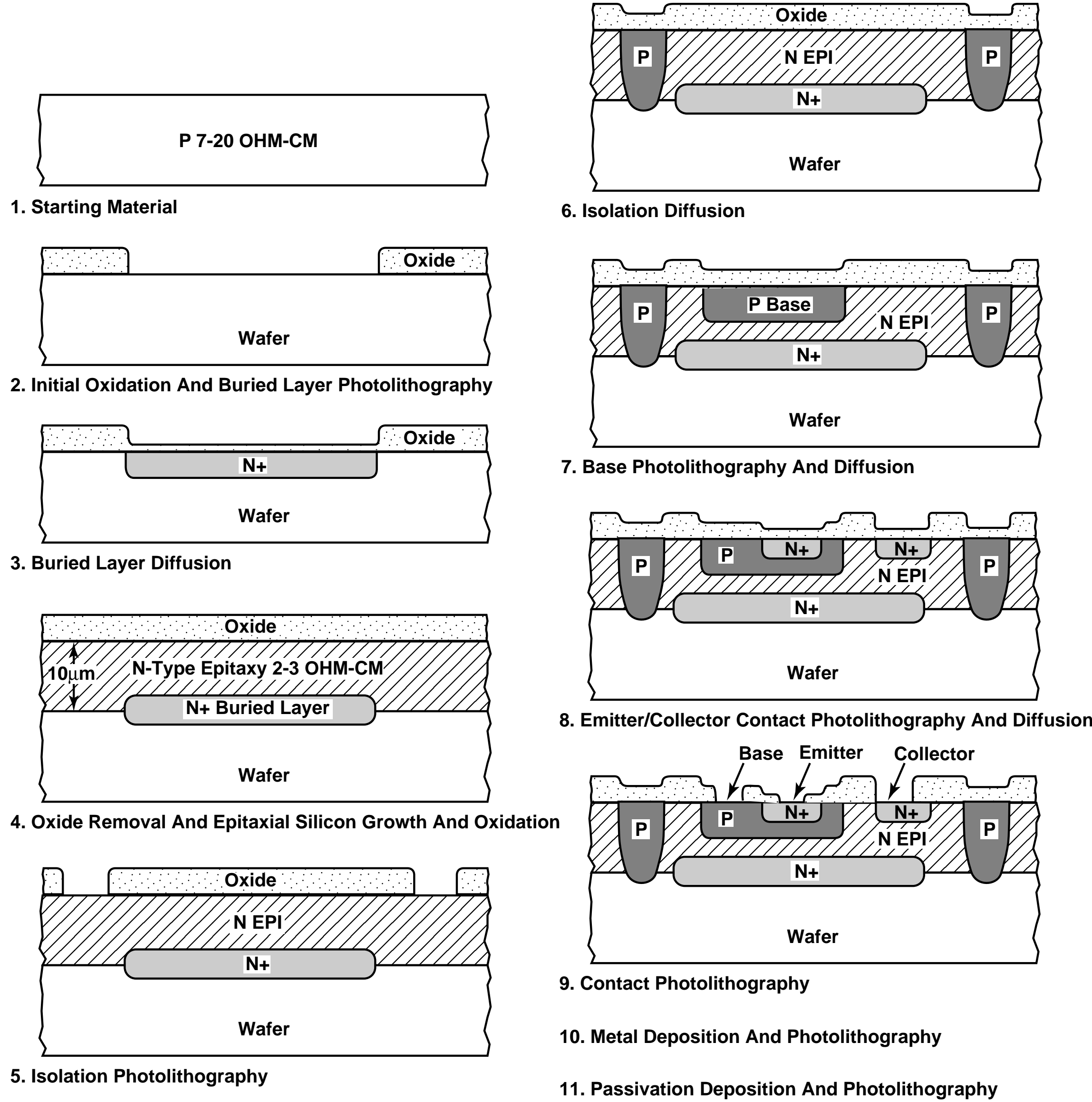


Figure 2-90. Bipolar Process Flow

4. Bipolar (ECL)

An advanced Bipolar ECL IC process is shown in Figure 2-91.

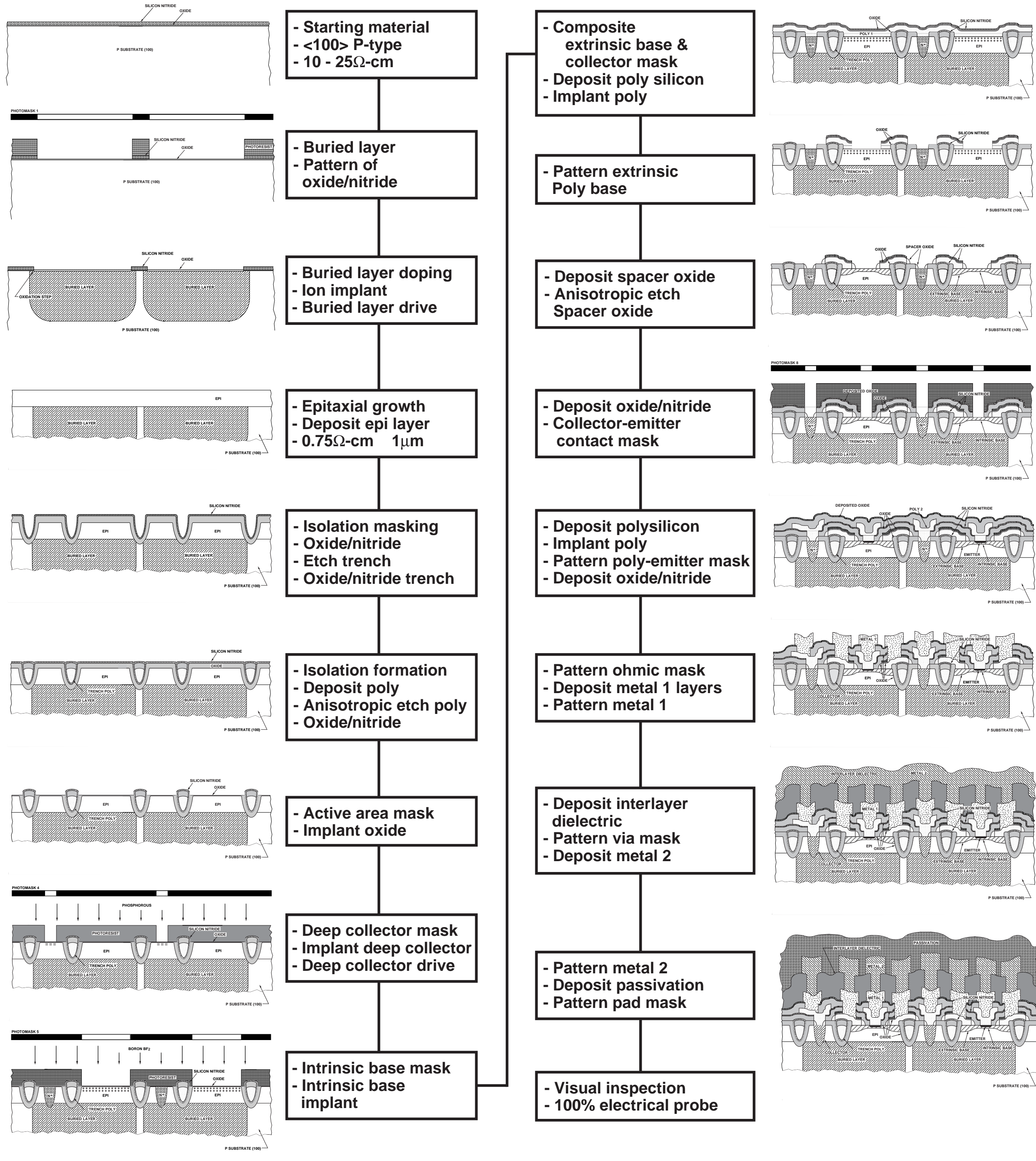


Figure 2-91. Advanced Bipolar Digital Process Summary

J. ASSEMBLY

After all the wafer fabrication steps have been completed, the good dice within the wafer are ready to be separated and assembled into final product form. The assembly or packaging process will place the electrically good devices (chips, die, and dice are other terms used to identify the individual circuits) in a package, interconnect the device to the package's leads and provide some form of final sealing.

Since the early days of the semiconductor industry, the packaging process has often been considered to be a lower level of technology than the wafer fab process. As to why this label emerged is purely speculation, but because of the high-labor content per package, the volume assembly of semiconductors was relocated to the Asia-Pacific region. This region continues to dominate the volume assembly.

As the semiconductor industry moved from the SSI (small-scale integration) era through MSI and LSI (medium and large) into the VLSI (very large-scale integration) levels of integration, packaging technology has changed dramatically. The changes have included automation, advances in materials, and new package designs.

The IC assembly sequence is outlined in Figure 2-92. The wafer is sawn into individual dice (separated). Each good die is attached to a package substrate (as shown) or leadframe. The bonding pads around the perimeter of the die are connected (usually wire bonded) to the internal terminations of the external leads. Finally, the package is completed by sealing the pieces of the housing together or by encapsulation with molding compound. This sequence will convert a tested wafer into a packaged product. The basic concepts of packaging are illustrated in Figure 2-93. Packaging requirements can be translated into package functions and constraints, which are shown in Figure 2-94. These functions and constraints along with the package design factors dictate the actual package design. The design factors consider the chip performance requirements, the density of chip, materials issues, assembly methods, and cost. These factors are expanded in Figure 2-95.

The cost factor has played a major influence on packaging. As previously stated, labor cost was responsible for assembly moving to the Asia-Pacific region. Material costs for plastic packages are less expensive than for ceramic or metal packages. Plastic package assembly methods have been automated to a higher level than either ceramic or metal packages. All of these cost factors favor the plastic package and have elevated it to be the highest volume type of IC package produced.

The major segments in the assembly process are outlined in Figure 2-96. These generic outlines can be further divided into flows for plastic packages and ceramic packages. These flows will be illustrated after reviewing the die separation, die bonding, and wire bonding processes.

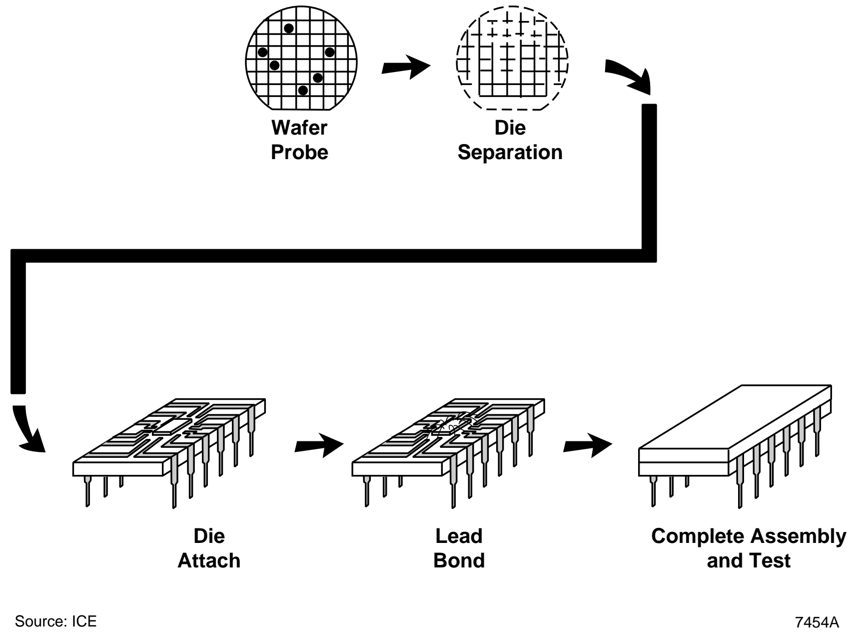


Figure 2-92. IC Assembly Sequence

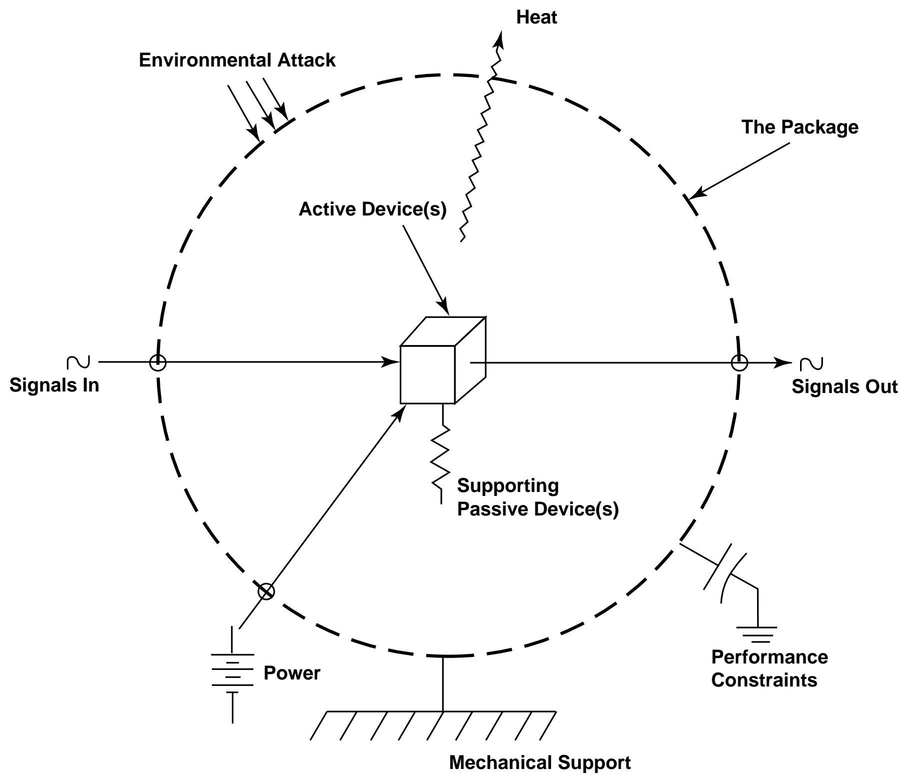
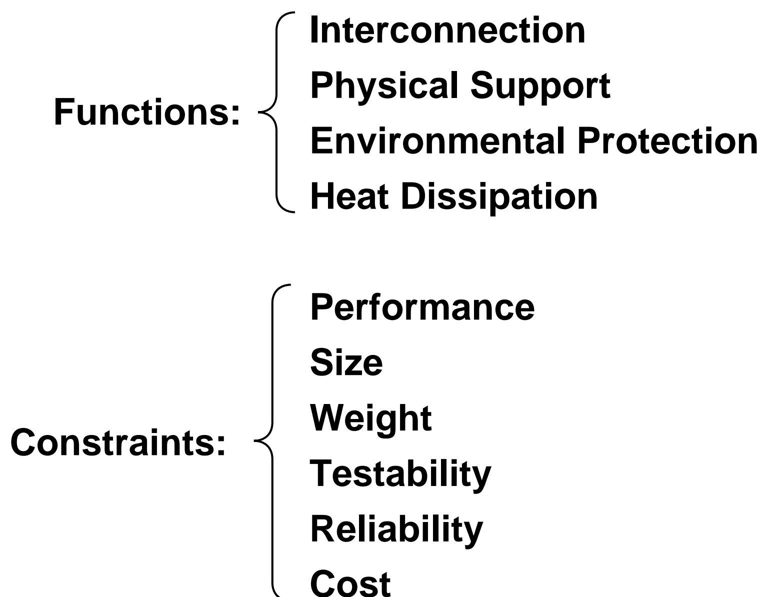


Figure 2-93. The Basis for Electronic Packaging



Source: ICE 7044

Figure 2-94. The Electronic Package

PERFORMANCE
- SIGNAL RISE TIME
- PROPAGATION DELAY
- SYSTEM IMPEDANCE
- ALLOWABLE LRC
- SWITCHING TRANSIENTS

DENSITY
- CHIP SIZE(S)
- SIGNAL I/O's
- GEOMETRY CONSTRAINTS EACH LEVEL
  - — BOND PAD SIZE AND PITCH
  - — PACKAGE PAD SIZE AND PITCH
  - — SUBSTRATE/BOARD
    - — LINE WIDTH AND PITCH
    - — NUMBER OF LAYERS

MATERIALS
- PACKAGE (Al_2O_3, BeO, METAL, PLASTIC)
- SUBSTRATE/BOARD
- THERMAL EXPANSION MATCH BETWEEN LAYERS

ASSEMBLY
- DIE BOND
- FIRST LEVEL INTERCONNECT
- LID SEAL/ENCAPSULATE
- PACKAGE ATTACH
  - — SURFACE/THROUGH HOLE/OTHER
- LEAD ATTACH
- CAVITY UP/DOWN

COST
- PACKAGE MATERIALS
- CAPITAL COSTS
- LABOR
- YIELD

7133B

Figure 2-95. Package Design Factors

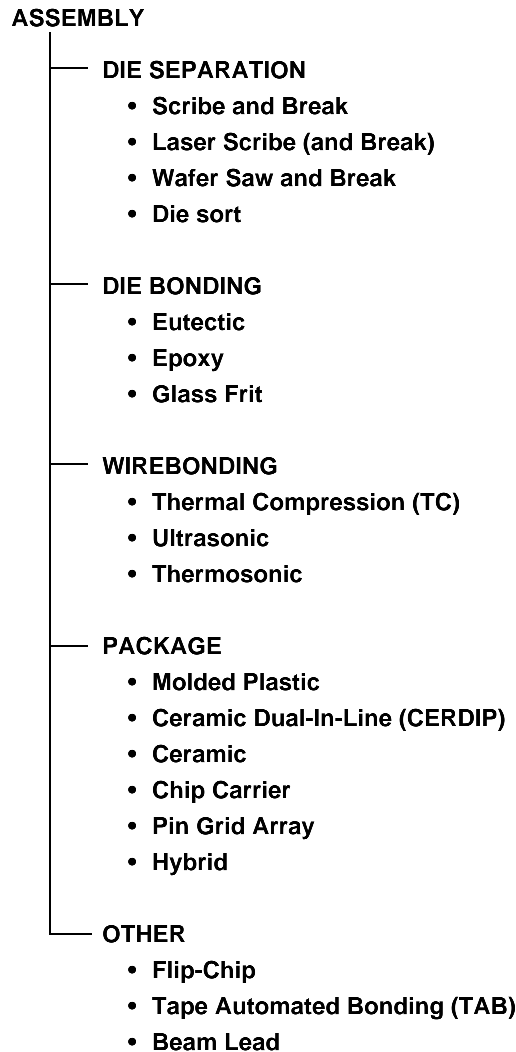


Figure 2-96. Assembly Processes

1. Die Separation

Sawing is the most preferred die separation technique. Cassettes of wafers can be placed into the automatic film mounter wherein each wafer will be properly oriented and placed onto an adhesive-type film that is secured to a rigid frame. This frame can be either metal or plastic. The mounted wafer is transferred to another cassette for transporting to saw.

The handling system in the saw transfers the frame onto the saw chuck. A pattern recognition system orients the frame properly and the correct sawing parameters are selected. The parameters have been previously loaded into the saw's computer memory. The saw will saw the wafer completely thru the wafer thickness in both directions, clean the slurry from the wafer surface and transfer the frame into an outgoing cassette. The dice are held in place by the adhesive on the tape. Visual inspection of the sawing operation is performed per the details of the SPC program.

The wafer saw is illustrated in Figure 2-97. An example of dice after sawing is shown in Figure 2-98. If the wafers are not sawn completely through their thickness, then a breaking process is used to separate the dice. This is done by applying pressure to the backside of the wafer.

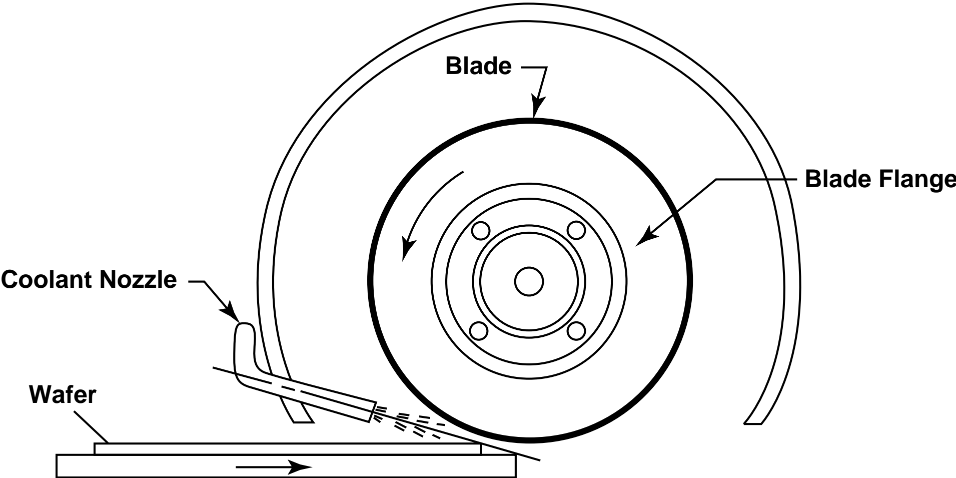


Figure 2-97. Wafer Saw

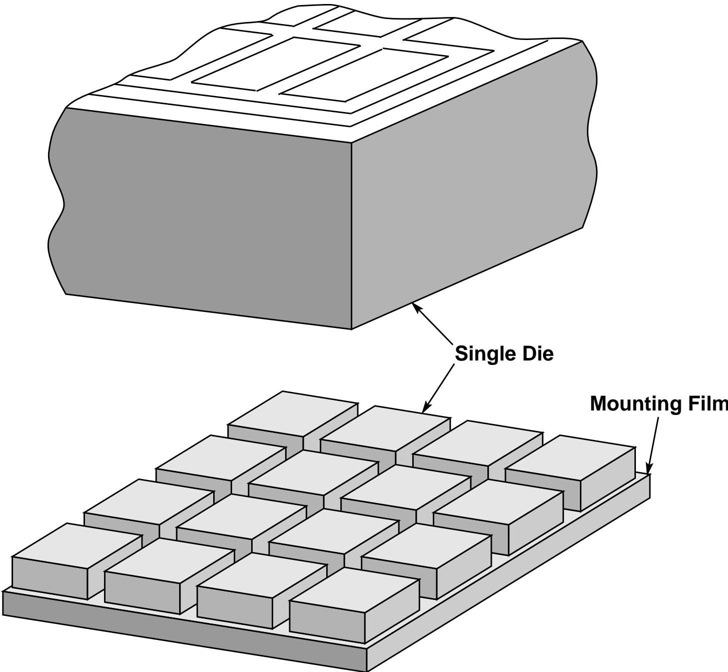


Figure 2-98. Examples of Wafer Saw Cut

2. Die Attach (Bond)

The sawn wafers are transferred to the die attach operation, where the dice are picked up from the adhesive film one at a time and attached to the package substrate or leadframe.

The automated, high-speed die bonder picks up the dice with a collet assembly connected to a vacuum line. The system handles the die only on the edges. This prevents damaging the top surface. The collet assembly transfers the die to the package substrate or leadframe. This method is shown in Figure 2-99.

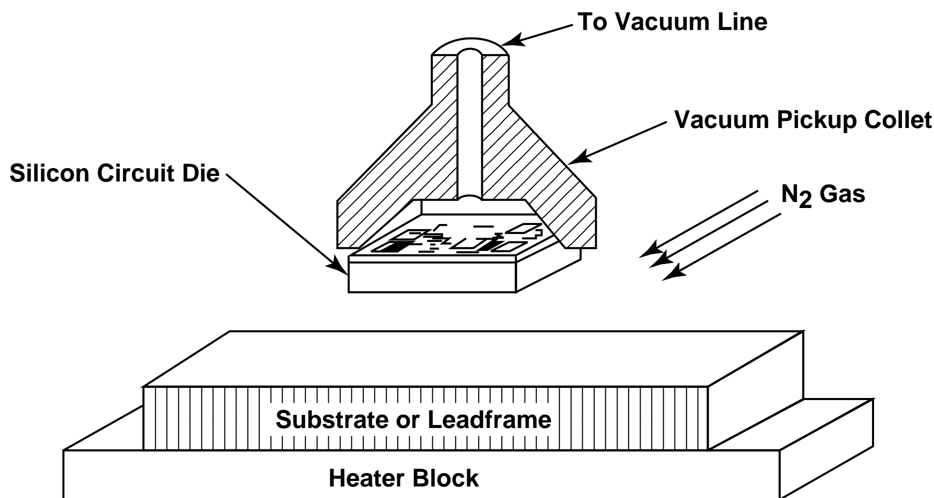


Figure 2-99. Die Bonding

If the die bonder is not equipped with pattern recognition capability and an operator is orienting the die, the operator selects un-inked dice and manually orients the bonder so it places them in the correct position on the package substrate or leadframe.

If the die bonder is equipped with pattern recognition, it automatically picks up only un-inked dice and orients them correctly before placing them on the package substrate or leadframe.

a. Eutectic Attach

There are several different materials and methods used for die attach. For a eutectic attach a thin film of gold is deposited on the back side of the wafer at the end of the wafer fab process. This gold layer is alloyed into the silicon to form a gold-silicon eutectic. The gold-backed die is attached to a package substrate or leadframe having either a gold- or silver-plated area. This attach method is common in ICs needing a good thermal path, mechanical strength, and electrical contact. Many bipolar ICs use this system.

b. Epoxy Attach

The most common attach material is epoxy. It can be formulated with or without silver. If the attach only requires a mechanical connection with modest thermal demands, then a straight epoxy is used. The epoxy will be silver-filled if the attach requires either an electrical connection or a lower impedance thermal path. The majority of MOS products use straight epoxy. An appropriate heat cycle cures the epoxy.

c. Glass Frit Attach

The third alternative is to die attach with glass frit material. This material is generally used for die attach in ceramic packages that require a high-temperature hermetic seal. The glass frit system generally uses a silver-filled mixture. A die attach technology comparison is tabulated in Figure 2-100.

DIE ATTACH TECHNIQUE	CONDUCTIVITY	DIE ATTACH OR CURING TEMPERATURE	MAXIMUM POST ATTACH TEMPERATURE	SHELF LIFE	COMMENTS
Si - Au EUTECTIC	• THERMALLY AND ELECTRICALLY CONDUCTIVE	>377° C	≈475° C	INDEFINITE	• VERY EXPENSIVE • REQUIRES DIE SCRUBBING • POTENTIAL FOR VOIDS ON LARGE DIE SIZES • ESTABLISHED PROCESS
EPOXY	• THERMALLY CONDUCTIVE ONLY • THERMALLY AND ELECTRICALLY CONDUCTIVE (SILVER OR GOLD)t (BEARING)	120° C TO 175° C	300° C TO 400° C	8 MONTHS TO 12 MONTHS - - - - - - - - - - - - POT LIFE FOR TWO COMPONENT MATERIALS: 4 HOURS TO 4 DAYS	• INEXPENSIVE • MAY REQUIRE TWO COMPONENT SYSTEM (RESIN AND HARDENER) • NO SCRUBBING • ESTABLISHED PROCESS
GLASS FRIT	• THERMALLY AND ELECTRICALLY CONDUCTIVE (SILVER BEARING)	70° C TO 120° C DRYING FOLLOWED BY 420° C TO 450° C FUSION	≈500° C	9 MONTHS	• MODERATELY EXPENSIVE • ONE COMPONENT SYSTEM • NO SCRUBBING • RELATIVELY NEW PROCESS FOR SILVER-BEARING GLASS

10295B

Figure 2-100. Die Attach Technology Comparison

3. Wire Bond

Wire bonding is used to connect the bonding pads (usually aluminum) on the die to the posts (bonding areas) on the package substrate or leadframe (Figure 2-101). The wire is attached to the pad on the die by a bonding tool (needle, capillary) and is fed through the tool as the tool moves from the pad to the post. The tool attaches the wire to the post and severs the wire before repeating the cycle. There are three types of wire bonding: *thermocompresssion, ultrasonic*, and *thermosonic ball bonding.*

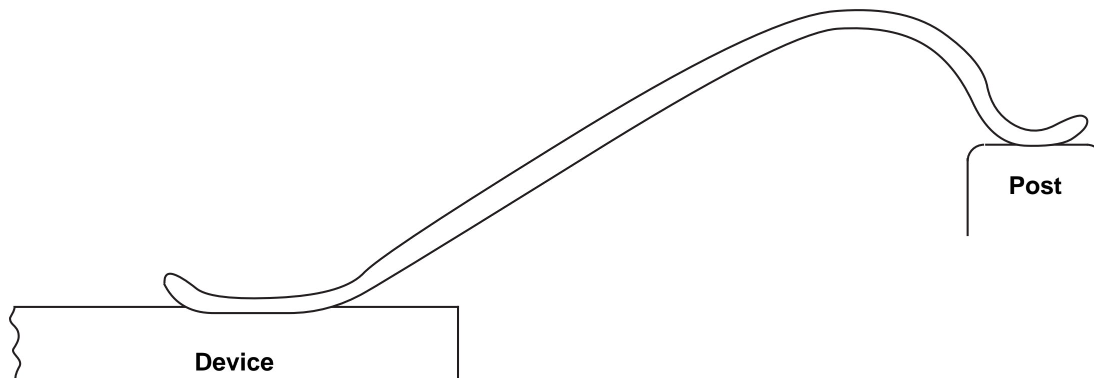


Figure 2-101. Graphic Representation of a Thermocompression Wire Bond

a. Thermocompression

In thermocompression bonding both the package substrate or leadframe and bonding capillary are heated. Pure gold or nearly pure gold wire is used.

The bonding mechanism positions the wire/capillary assembly over the aluminum bonding pad on the die. A force is applied to the wire through the capillary. The combination of the applied force and the heat causes the gold/aluminum interface to blend together to form the bond. The capillary feeds out wire as it moves to the post, where a second bond is made in a similar manner. The thermocompression wedge bond at the bonding pad is shown in Figure 2-102. A graphic representation of the complete wire bond was illustrated in Figure 2-101.

b. Ultrasonic

Some packages require a single-metal system at the die. Since the majority of the ICs produced use aluminum or aluminum alloys for the bonding pads, aluminum wire is needed for wire bonding. This need has been satisfied through the use of ultrasonic energy for the bonding process.

Ultrasonic bonding of aluminum wire is accomplished by feeding the wire through the heel of the bonding capillary (similar to what is used for thermocompression bonding). The wire resides in a groove on the bottom of the capillary. The capillary/wire is positioned over the bonding pad. The capillary applies a force on the wire and ultrasonic energy is transmitted through the capillary to the wire/bonding-pad interface to effect a metallurgical bond. The substrate is not heated. The capillary on the bonding pad is shown in Figure 2-103. The ultrasonic bonding sequence is illustrated in Figure 2-104.

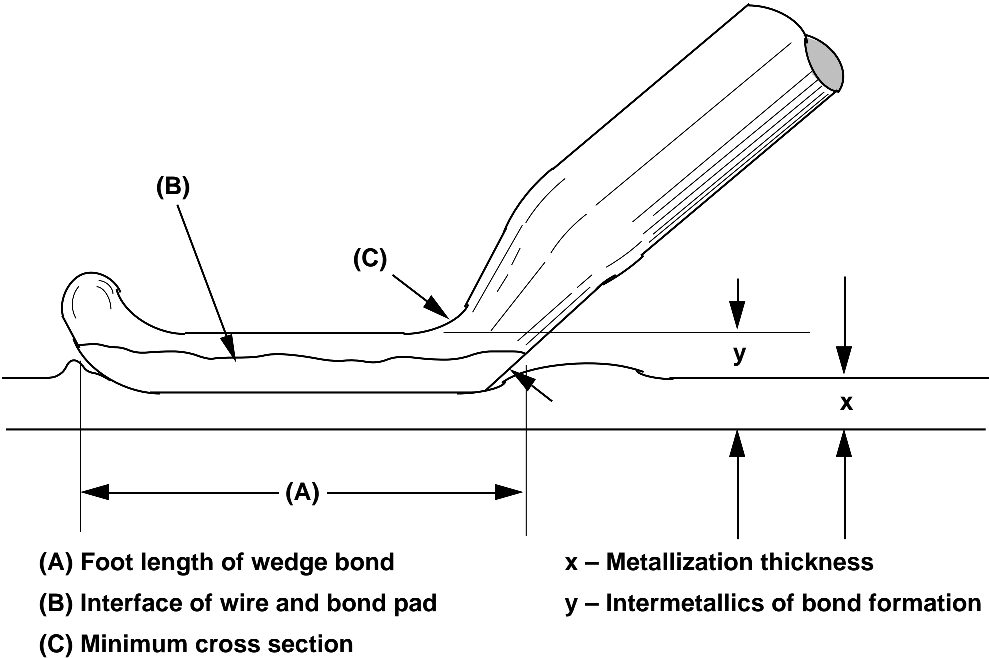


Figure 2-102. Cross Section of Wedge Bond

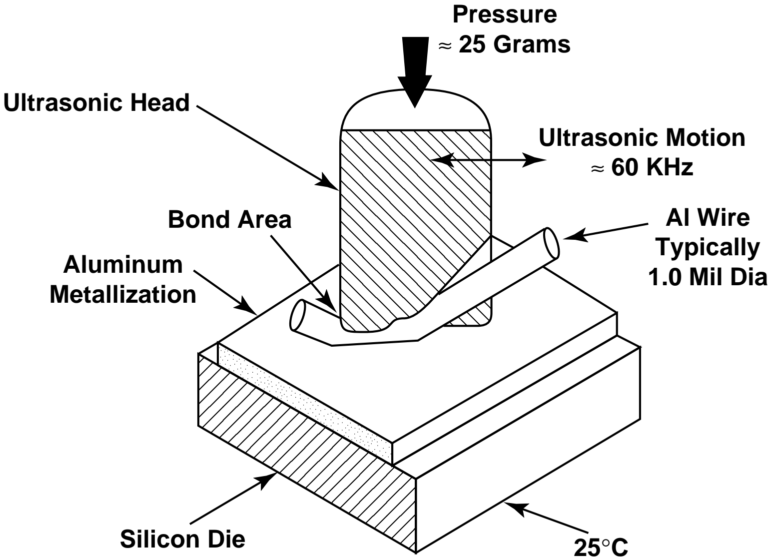


Figure 2-103. Ultrasonic Wire Bonding

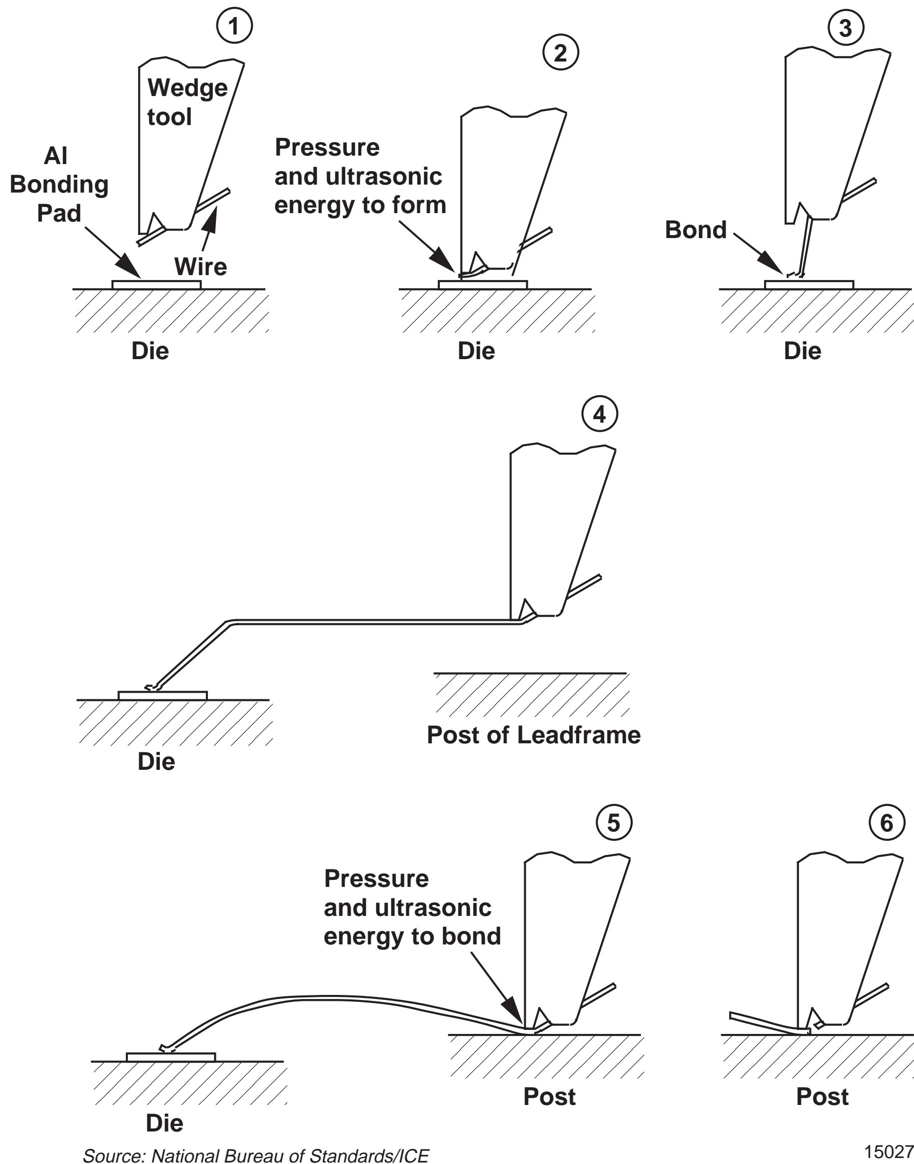
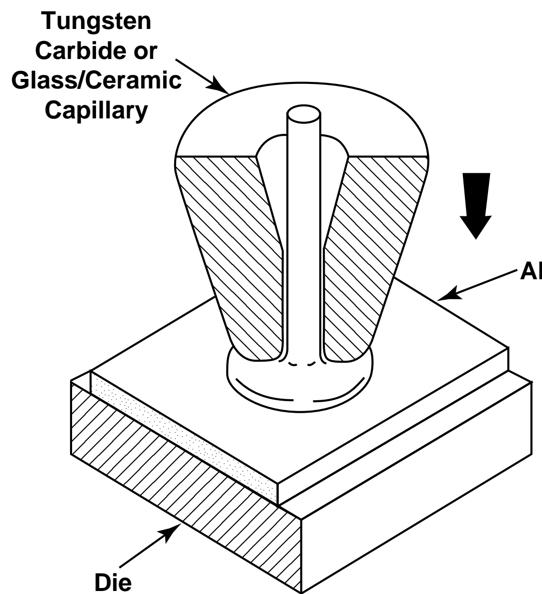


Figure 2-104. Ultrasonic Bonding Sequence

c. Thermosonic Ball Bond

The third type of wire bonding system is called thermosonic ball bonding. Unlike thermocompression and ultrasonic bonding capillaries, the ball-bond capillary has the gold wire fed vertically through a center bore from the top exiting out the bottom. This is illustrated in Figure 2-105. At the beginning of a bond cycle, the wire sticks out from the capillary. It is heated and the melted wire forms a ball (Figure 2-106, No. 2). The ball is positioned on the bonding pad. Pressure is applied to the capillary and ultrasonic energy is transmitted through the capillary. The package substrate or leadframe is heated. The substrate heat is transmitted to the die. The combination of heat, pressure and ultrasonic energy creates a metallurgical bond between the gold bonding wire

and the aluminum metallization. The capillary is moved to the package bonding post and properly positioned. The wire is now exiting from one side of the tip of the capillary. Pressure and heat cause the wire to be attached to the post. The second bond shape is partially controlled by the physical shape of the inside bore of the capillary. The inside chamfer (radius) causes the wire to be thinned enough during the second bond such that a slight upward motion at the end of the bonding cycle breaks the wire. The inside chamfer is shown in Figure 2-107. A completed ball bond is shown in Figure 2-108.



Wire: 0.0007in. to 0.002in. Au
Substrate: 175 – 225°C
Needle: Cold
Gas: N_2 at 2 CFH
Weight: 50 Grams

Source: ICE 1827C

Figure 2-105. Ball Bonding

The ball bonding system has excellent control of the wire loop. Loop control is an important parameter in packaging. The packaging trend continues toward thinner package profiles. The thinner the package the greater the demand on loop control. An example of loop control is illustrated in Figure 2-109.

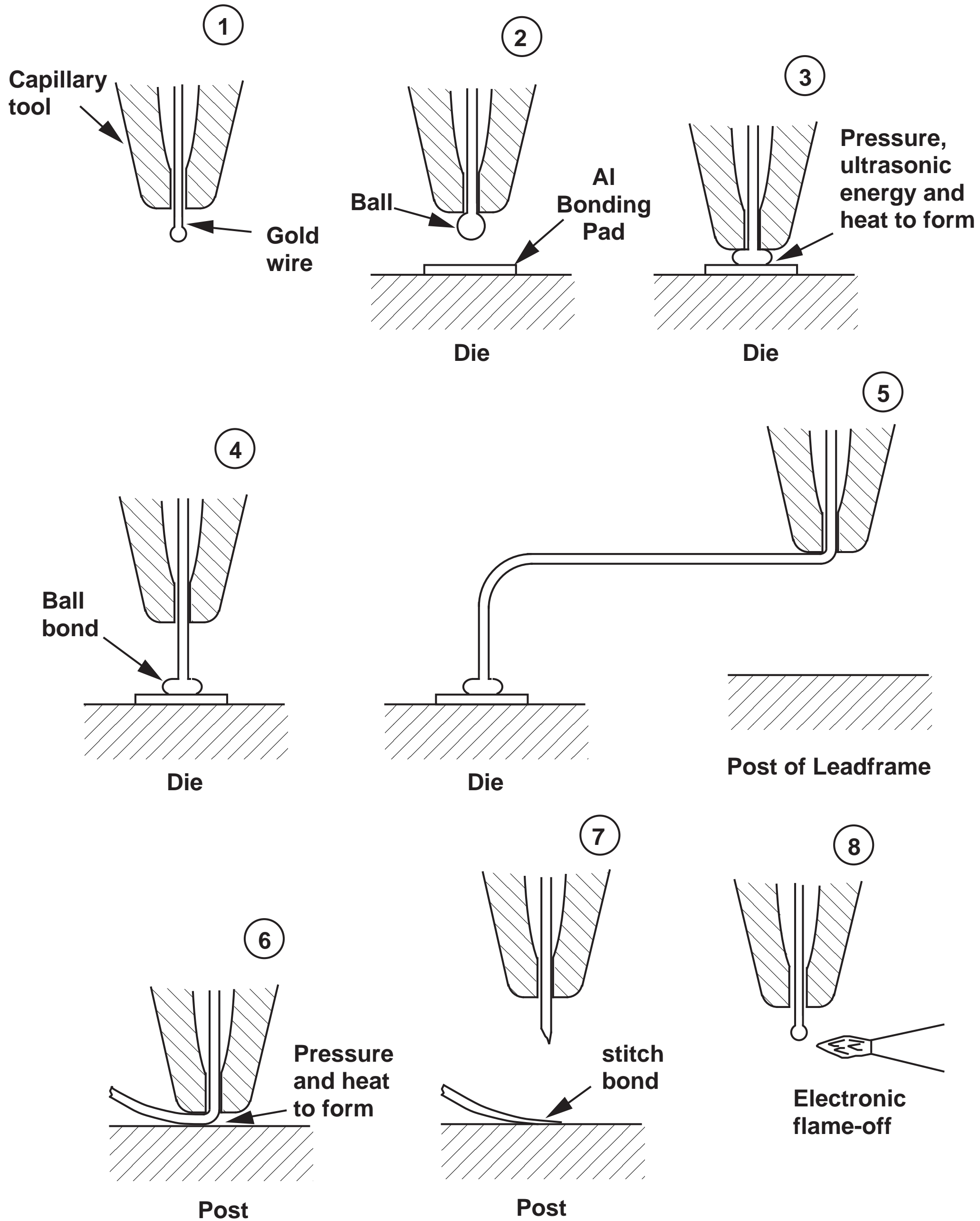


Figure 2-106. Thermocompression/Thermosonic Bonding Sequence

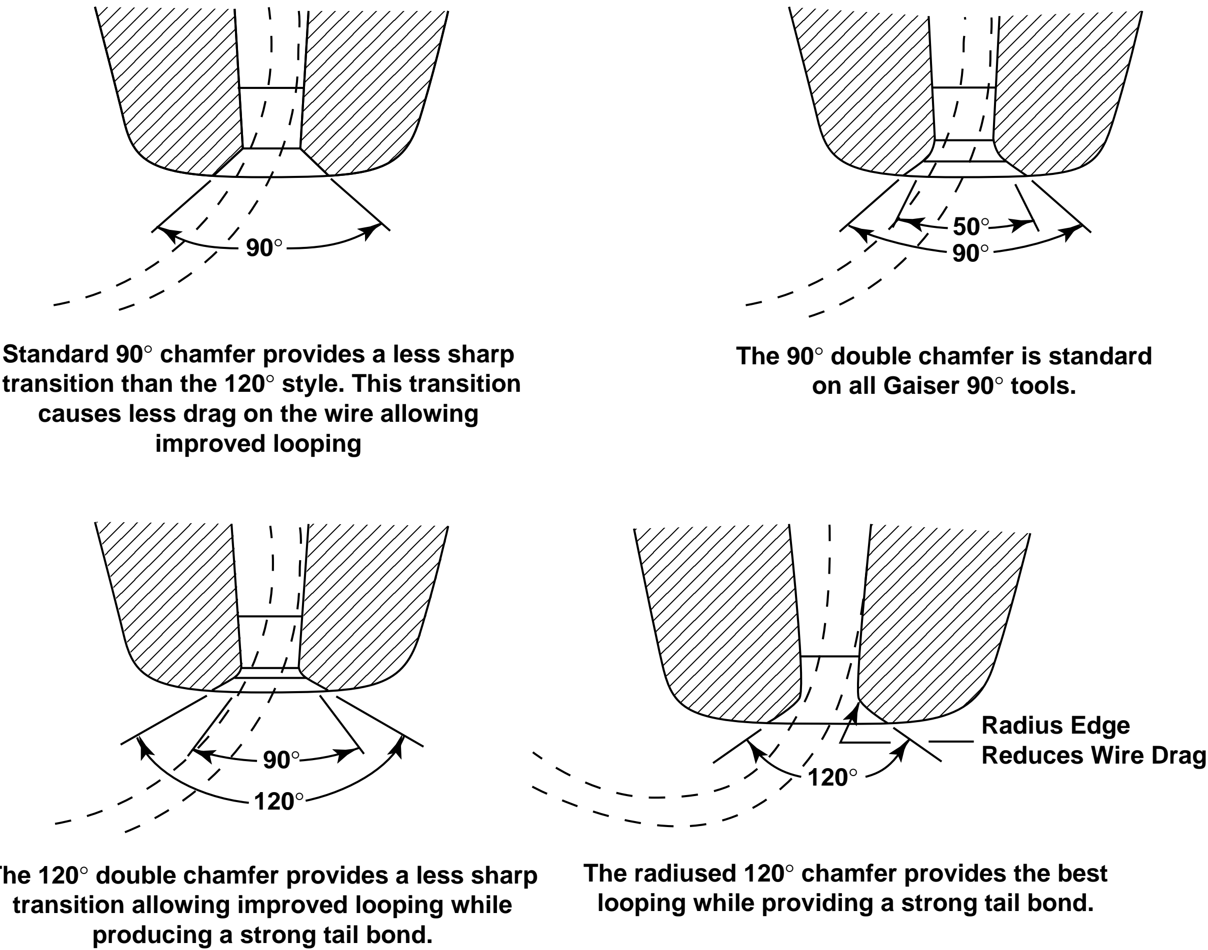


Figure 2-107. Inside Chamfer Discussion

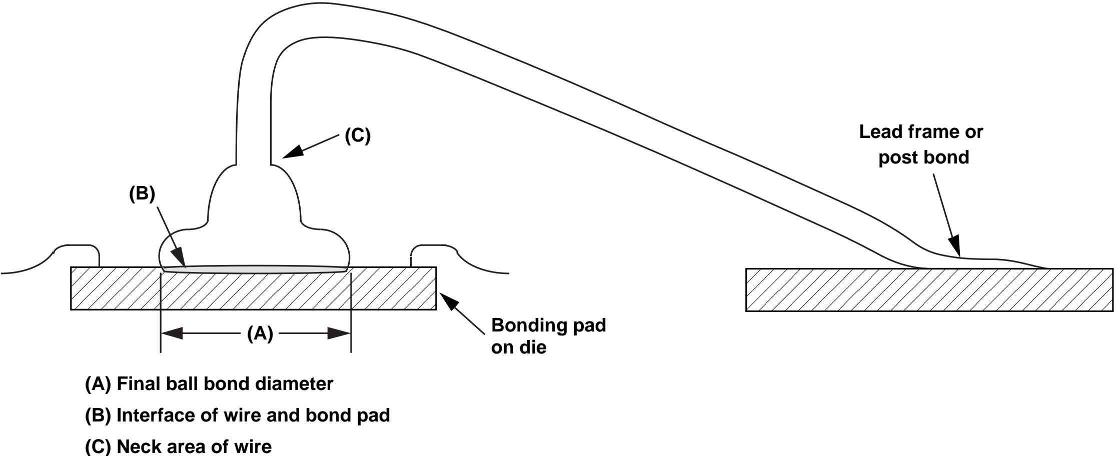


Figure 2-108. Cross Section of Ball Bond

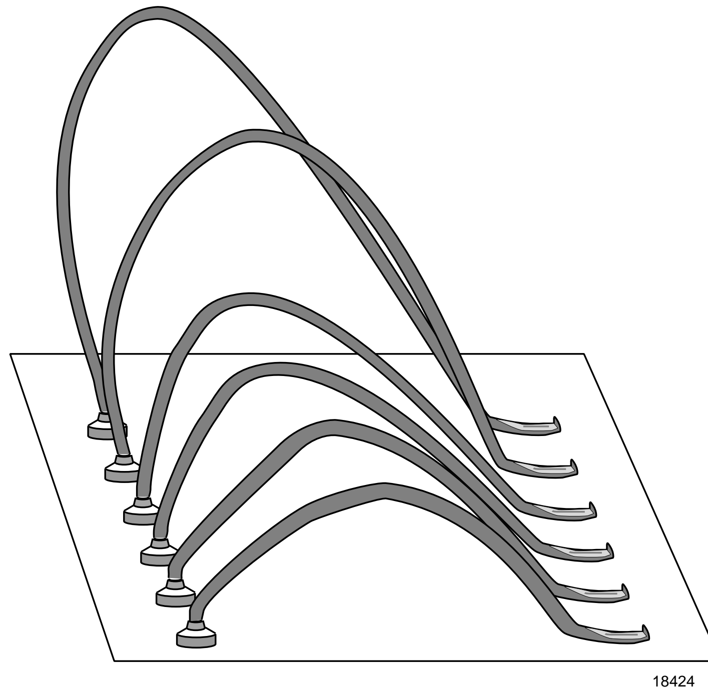


Figure 2-109. Loop Control of Wire Bonds

- **Lift at bonding pad metal**
- **Lift at leadframe**
- **Break at heel of bond (leadframe)**
- **Break at ball (neck - down)**
- **Center wire break**

15020

Figure 2-110. Wire Bond Failure Categories

The wire bonding process is controlled by having a good Statistical Process Control program. Part of this program will be a sample visual inspection, wire pull testing, and ball shear testing. The pull test data should be categorized for any bond failure. Generally the five categories shown in Figure 2-110 are used. Each of these categories must be monitored closely, particularly the lifted bond. The ball shear test should serve to guard against lifted bond failures that would not be detected by the conventional wire pull test.

Wire bonding technologies summarized in Figure 2-111. This figure also shows the better productivity of the thermosonic bonding system.

One of the future challenges facing the wire bond method of lead attach is the need to bond to smaller bonding pads placed closer together. Part of the spacing issue centers on the use of staggered bonding pads, often referred to as "zero pitch" (Figure 2-112).

WIRE BOND TECHNIQUE	PROCESS PARAMETERS	WIRE	AUTOMATED ALIGNMENT WITH PATTERN RECOGNITION	WIRE BONDS PER SECOND (TYPICAL)	AUTOMATIC LOOP CONTROL	COMMENTS
ULTRASONIC	- AMBIENT TEMPERATURE - ULTRASONIC GENERATOR - 60 Khz - 2 TO 5 WATTS - BONDING FORCE 20 GRAMS TO 150 GRAMS	Al OR Al/Si 1 MIL TO 2 MIL FOR IC 5 MIL TO 20 MIL FOR POWER	YES	2 - 3	YES	- ROTARY BONDING HEAD AVAILABLE - ALL PARAMETERS FULLY PROGRAMMABLE (X, Y, Z, θ) - HIGH RELIABILITY APPLICATION
THERMOCOMPRESSION	- 250°C TO 300°C - BONDING FORCE 20 GRAMS TO 150 GRAMS	Au 0.7 MIL TO 2 MIL	YES	3 - 6	YES	- ALL PARAMETERS FULLY PROGRAMMABLE (X, Y, Z, θ) - ELECTRONIC FLAMEOFF - COMMERCIAL APPLICATIONS - CAN PROMOTE METALLIC INTERDIFFUSION IF BOND TEMPERATURE IS TOO HIGH
THERMOSONIC	- 175°C TO 225°C - ULTRASONIC GENERATOR - 60 Khz - 0.5 TO 2 WATTS - BONDING FORCE 20 GRAMS TO 150 GRAMS	Au 0.7 MIL TO 2 MIL	YES	4 - 8	YES	- ALL PARAMETERS FULLY PROGRAMMABLE (X, Y, Z, θ) - ELECTRONIC FLAMEOFF - COMMERCIAL APPLICATIONS - ALUMINUM WIRE THERMO-SONIC BALL BONDING IN FIELD EVALUATION

10296B

Figure 2-111. Wire Bond Technology Comparison

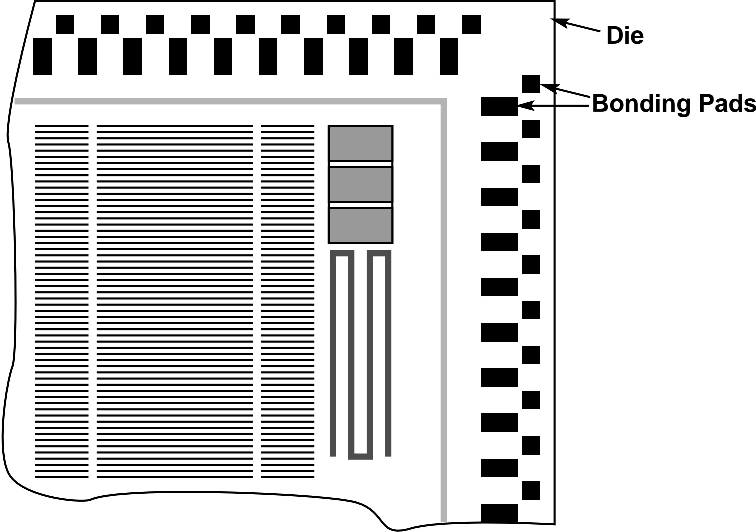


Figure 2-112. Staggered Bonding Pads

4. Plastic Packages

a. Molding

The Plastic Dual-In-Line Package (PDIP) shown in Figure 2-113 has been the industry workhorse since it came into existence in the early 1960's. The PDIP package has served the industry well, rising up to an 80 percent market share in 1981 before beginning to decline in the late 1980's. The assembly process flow chart for the PDIP is shown in Figure 2-114.

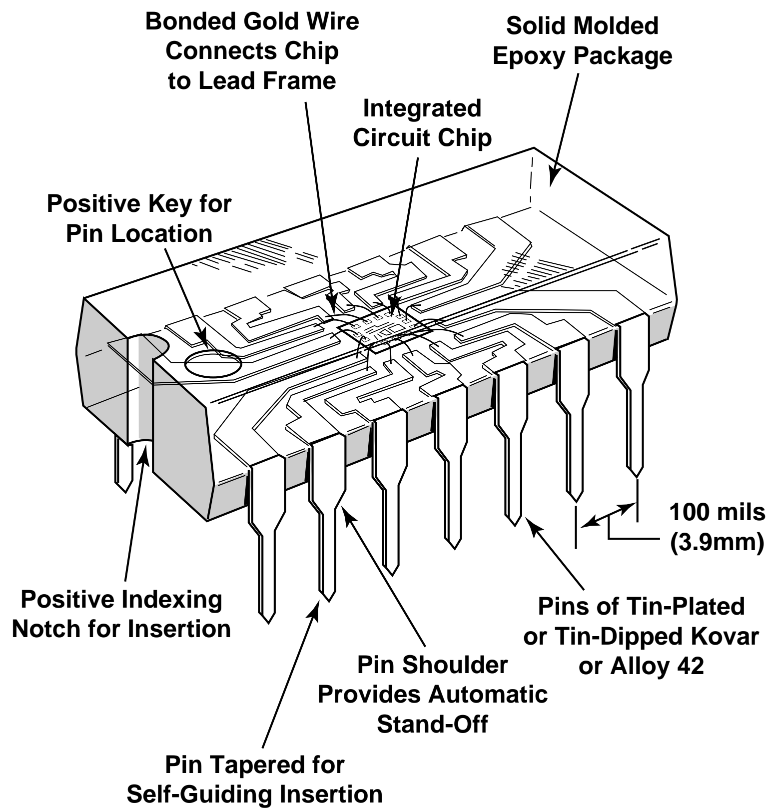


Figure 2-113. Plastic Dual In-Line Package

The plastic assembly method utilizes mass production techniques by using strips of leadframes. This is shown in Figure 2-115. The leadframe strips are loaded into a magazine. The magazine is shown in Figure 2-116. This allows mass handling throughout the assembly process. After die bond and wire bond, the magazine can be interfaced with the various molding process methods. A large block cavity mold is shown in Figure 2-117. The other molding method molds one strip at a time.

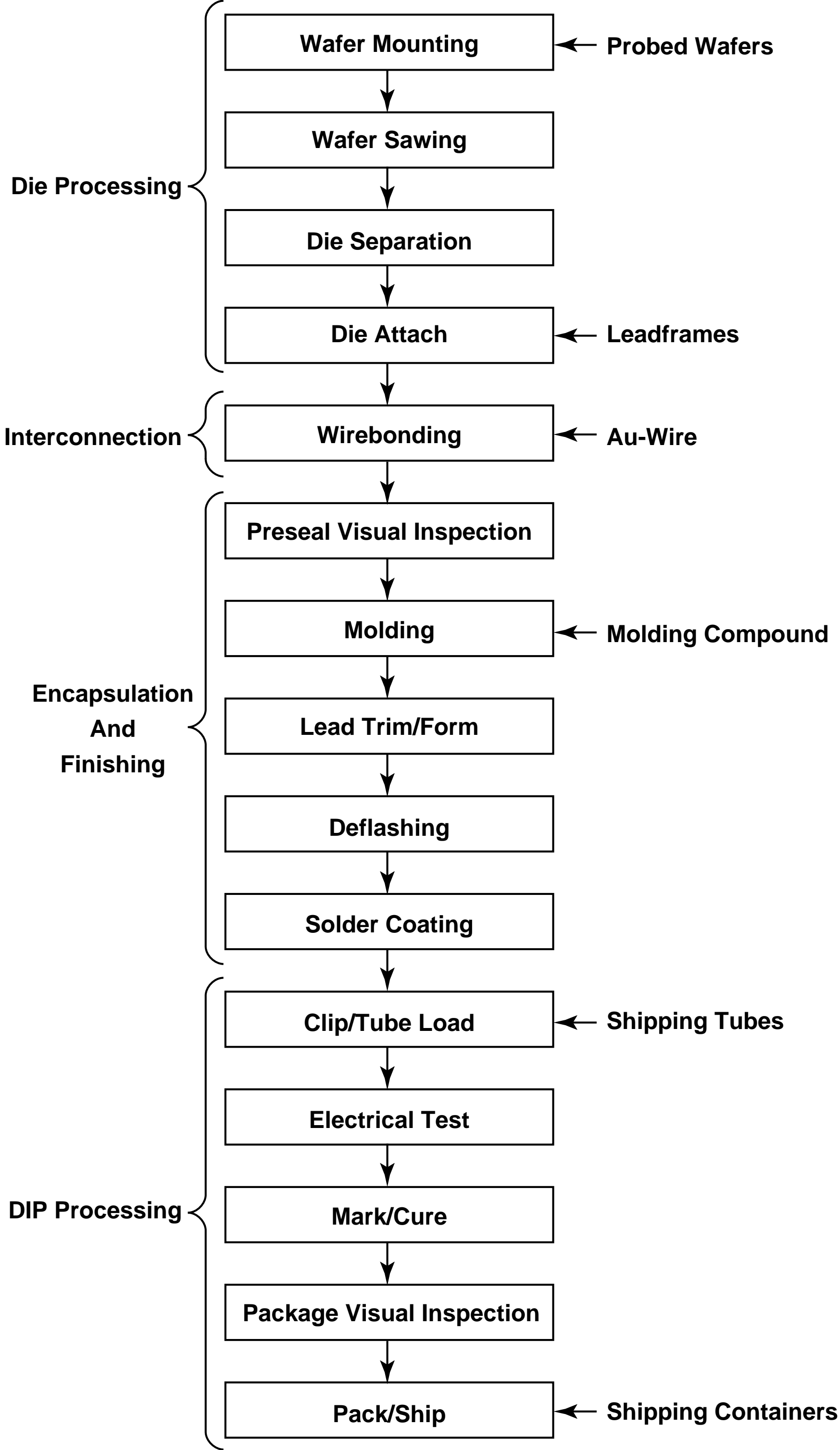
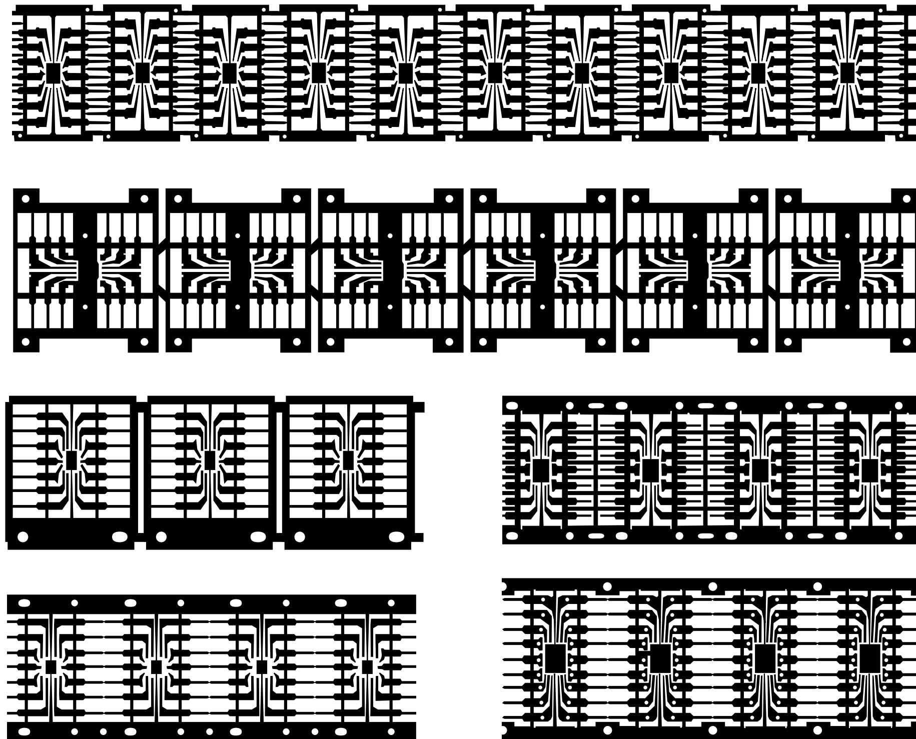


Figure 2-114. Assembly Process Flow Chart for Plastic DIP



Source: Copeland & Jenkins, Ltd. 6609

Figure 2-115. Leadframes for Plastic DIP

b. Lead Trim and Form

After the molding operation, the magazines are placed into a large oven to cure the molded material. The strips are transferred to a deflash operation to remove any excess molding materials from the leads of the devices. The next process feeds the strips into the lead trim-and-form tooling. The output from this operation transfers the completed device into a handling tube. Often this handling tube is referred to as a "rail." At this stage the tubes feed a handler that transfers the device onto a "wave soldering" pallet. After the lead soldering operation, the device is returned to the "rail" or tube. The tubes of devices are ready for final package testing.

The previously described processes have been adapted to all of the various plastic package designs, i.e., SOIC, PLCC, PQFP, etc. This has resulted in a cost effective assembly manufacturing.

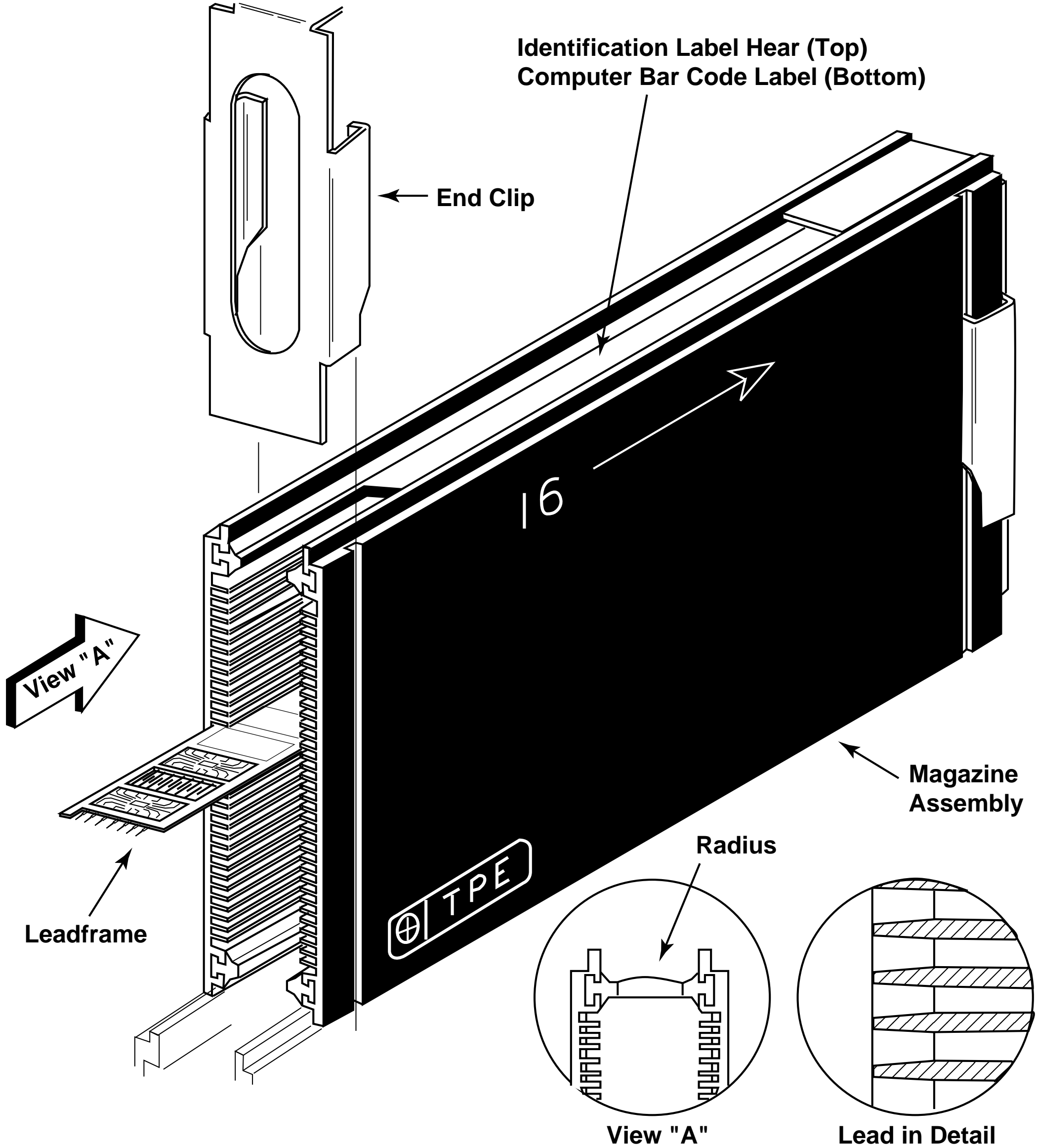


Figure 2-116. Typical Leadframe Magazine

5. CERDIP

The Ceramic Dual-In-Line (CERDIP) package is the most popular ceramic package. An assembly flow chart is shown in Figure 2-118. This is a hermetic package. Many of the handling techniques used in the plastic package assembly have been adopted into ceramic assembly. Because of the hermetic seal, the CERDIP leadframe is lead formed and clipped into individual units after the leadframe is manufactured.

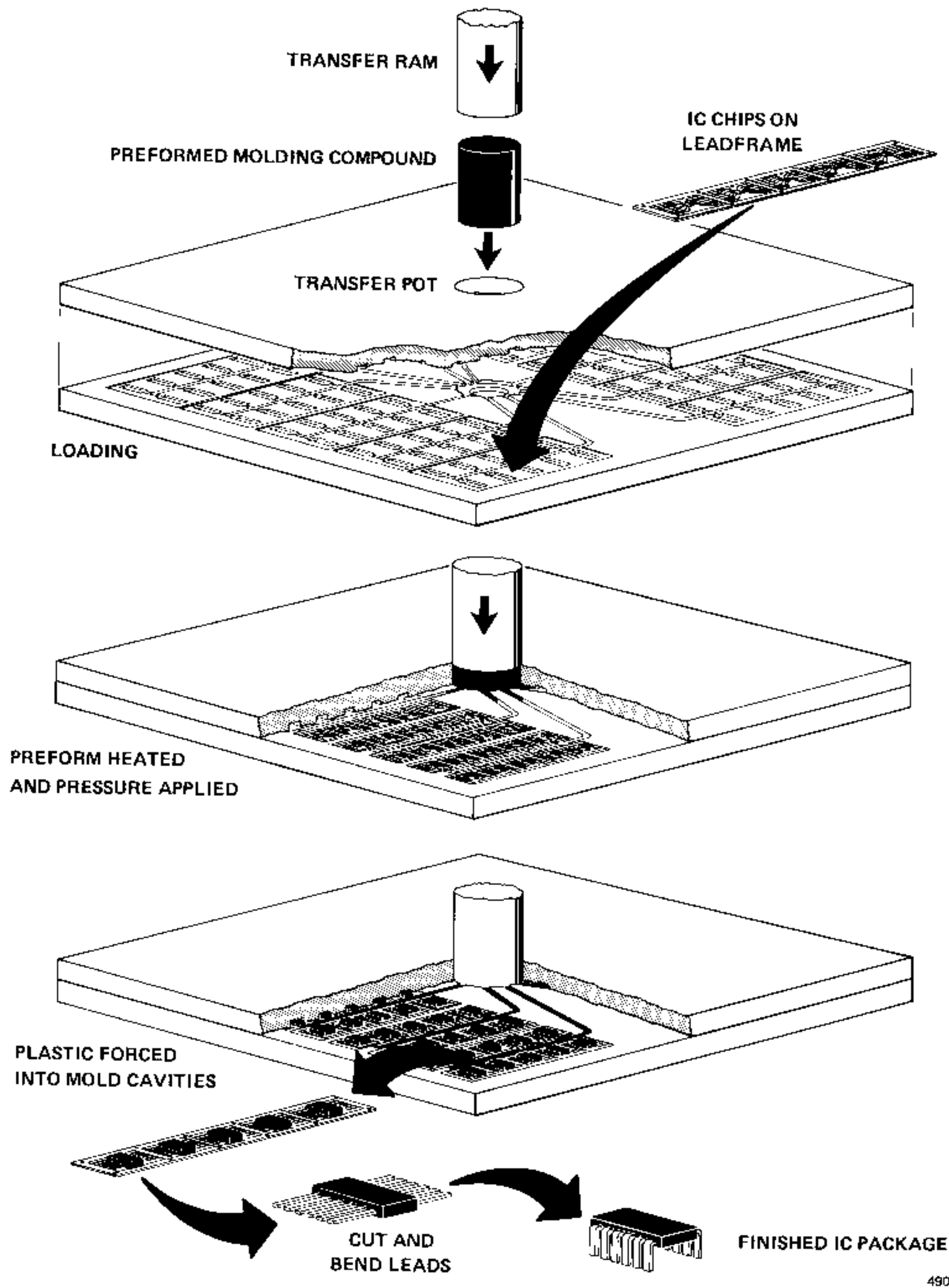


Figure 2-117. Transfer Molding Procedure

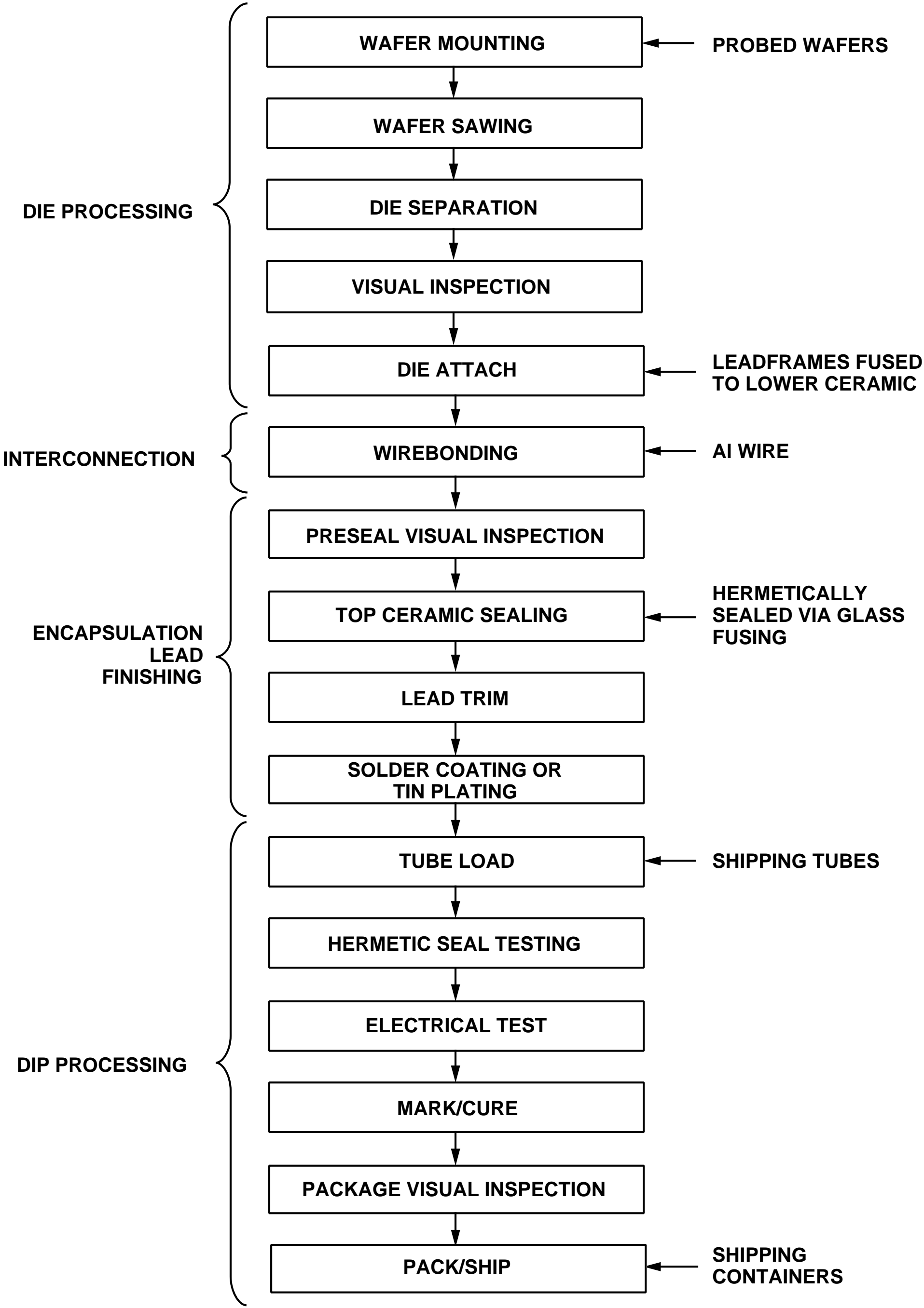


Figure 2-118. Assembly Process Flow Chart for CERDIP

A bottom ceramic with a glass glaze is placed on a heated stage and the leadframe is aligned and embedded into the soft glaze (Figure 2-119, No. 2). The die is oriented and eutecticly attached to the lower ceramic.

The units are then wire bonded and sealed (Figure 2-119, Nos. 2, 3). A glass frit is used to attach the top and bottom of the package together. The melted glass frit provides a hermetic seal when it cools. After the leads of the units are plated, they are trimmed (Figure 2-119, No. 5).

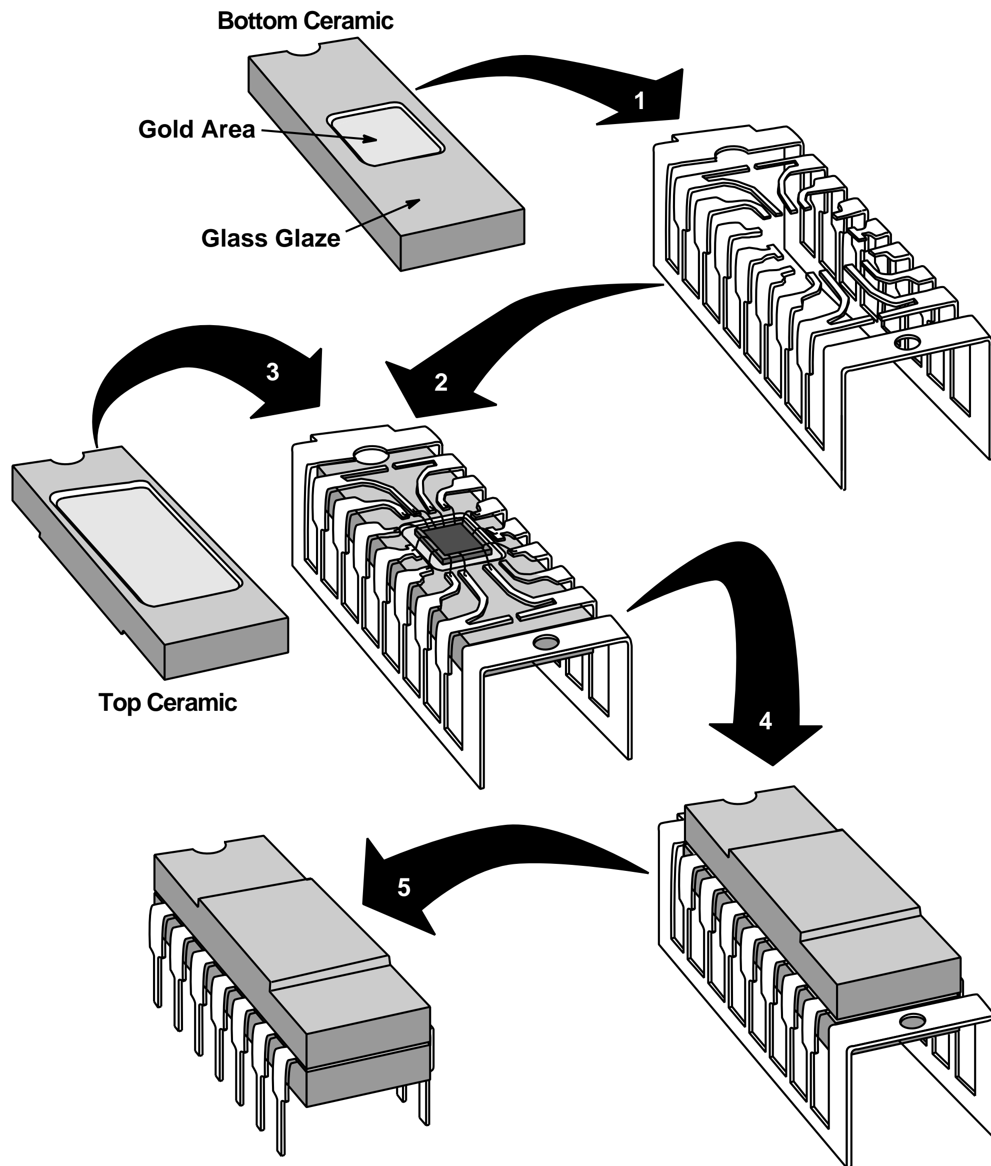


Figure 2-119. CERDIP Assembly Sequence

The final testing of the packaged device is done by automated handlers interfaced with the appropriate tester. In many test areas, the marking equipment is configured in-line with the testing to complete the manufacturing cycle. An in-line Q.C./Q.A. function is integrated into the test area.

There are many other package types used for ICs. The assembly processes for these packages are modifications of the plastic and ceramic assembly methods already presented.

K. PROBE AND TEST

1. Wafer Testing

a. Test Patterns

The electrical evaluation of an IC in wafer form requires two different types of testing. The first electrical test is performed on *"test pattern"* structures. The other electrical evaluation is the testing of the IC to a given electrical specification.

The *"test pattern"* structures are arrangements of individual test patterns, each designed to evaluate specific process attributes. For example, there will be enhancement mode and depletion mode transistors of various geometrical sizes to represent the circuit elements, resistivity structures, critical dimension and overlay registration structures, defect density structures, interconnect test structures for metal step coverage, contact resistance, electromigration, stress migration, and parasitic circuit structures.

All of these test structures' electrical characteristics are measured, recorded, and correlated to the various real-time test wafer measurements taken after each critical process step is completed.

The maturity of the product, the specific technology, the type of photo equipment, and the engineering management will determine where the test structures are arranged on the wafer and the statistics used to accept individual wafers or the wafer lot. Generally the *"test patterns"* are sampled statistically on the wafer and wafers in a lot.

If a sample fails the electrical criteria, the lot will be judged by engineering for further action. Some of the possibilities are:

1. Test additional samples
2. Test all wafers
3. Reject wafers failing the criteria
4. Reject the lot

b. Probe Testing

The wafers receive a 100 percent wafer probe evaluation after the *"test patterns"* are evaluated. This testing is product specific to data sheet or customer requirements. This test of the wafer has been primarily a room temperature test. However, there is a trend to raise the substrate temperature to 50°C. Higher temperatures will cause temperature-sensitive parameters to change, allowing a more critical evaluation.

Each whole die on the wafer is tested. The probe machine moves the wafer to a location under a set of probe needles. The probe needles are lined up with the bonding pads of the first die to be tested. The needles are lowered onto the die (or the die is raised to come into contact with the needles) and that die is tested. The needles separate from the die and the wafer moves to a position where the needles are lined up with another die. The cycle is repeated until all of the whole dice on the wafer are tested. The dice that fail the test are identified, usually with a drop of ink.

After the 100 percent wafer testing is completed, the wafers are sent to the assembly process for final packaging.

2. Final (Product) Testing

The final testing of the product after packaging tests the product to the data sheet specification or to a customer-specific test requirement. A variety of final tests are performed to check for specified electrical characteristics.

Automation is used in a majority of the final test activities. Only when a package is new to the industry or customized and does not fit existing hardware will hand insertion to a socket be used. Bent leads or poor lead finish on a package can cause severe problems for final test. Thus, automation is necessary to prevent these and other problems.

3. Burn-in

"Burn-in" typically follows Final Test. The use of *"burn-in"* evolved from semiconductors used in military systems in the early 1960's. The concept has stayed with semiconductor manufacturing in varying degrees. The reason for applying voltage to a semiconductor device at an elevated temperature for some period of time is illustrated in Figure 2-120. This is considered a *"stress test."* From a reliability point of view the elevated temperature with voltage applied for a given time period will cause marginal devices to fail. This removes a potential *"field failure."* The screening is intended to improve the statistical level of reliability by removing the *"infant motality"* devices.

This *"burn-in"* screen is used in many different ways by the semiconductor industry. Within a given company the *"burn-in"* conditions vary over products and technologies.

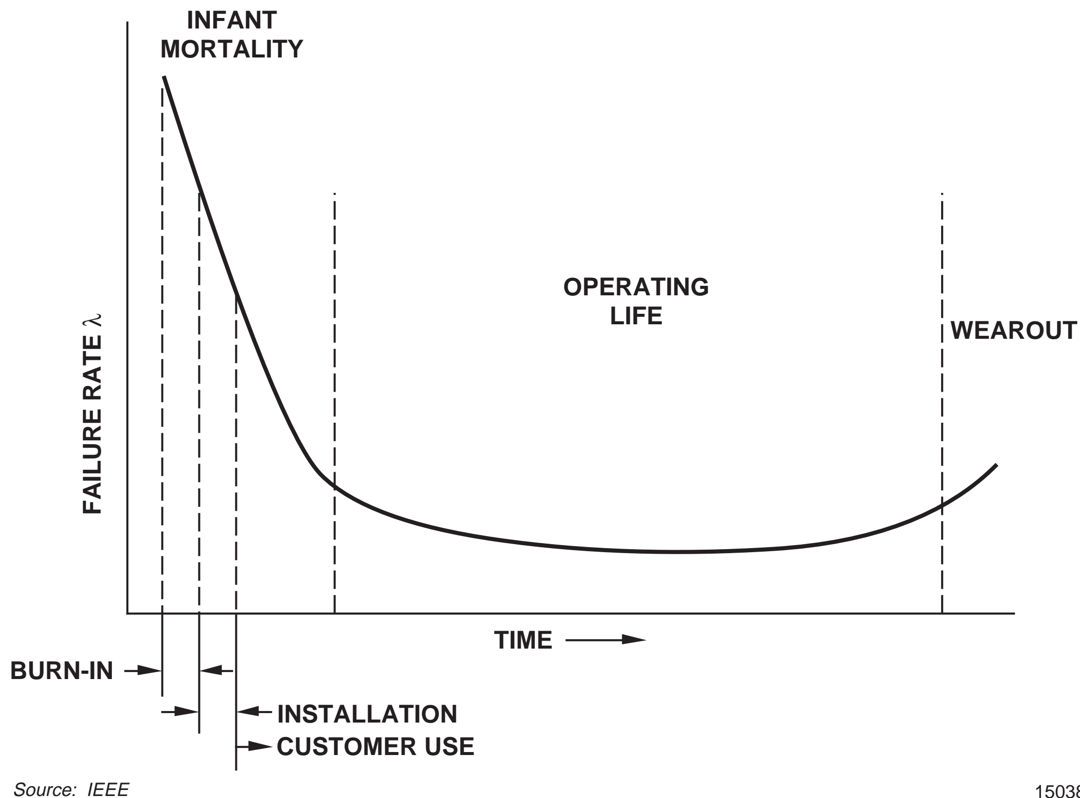


Figure 2-120. Bathtub Curve Prediction of Reliability

L. SUMMARY OF SEMICONDUCTOR MANUFACTURING

The impact of integrated circuits on the electronics industry has been phenomenal. Not only are they changing the physical appearance and modes of operation of electronic equipment, but are also causing major changes in the entire structure of the electronic industry. Methods of transacting business are changing. The supplier-user interface is changing. The decision to custom design ICs or use standard or semicustom devices becomes a major consideration for many electronic companies.

The IC manufacturing process roadmap is illustrated in Figure 2-121. The logic designer can either work for a systems company or be part of the IC design team within the semiconductor manufacturer. The circuit designer generally works for the semiconductor manufacturer. The circuit designer translates the logic designer's requirements into a semiconductor circuit design.

The circuit designer will convert the electrical schematic of the circuit into the physical size of each component, i.e., transistor, diode, resistor, capacitor, etc., that makes up the circuit. The designer uses a workstation (CAD) to accomplish the design and do the many different simulations required for design verification.

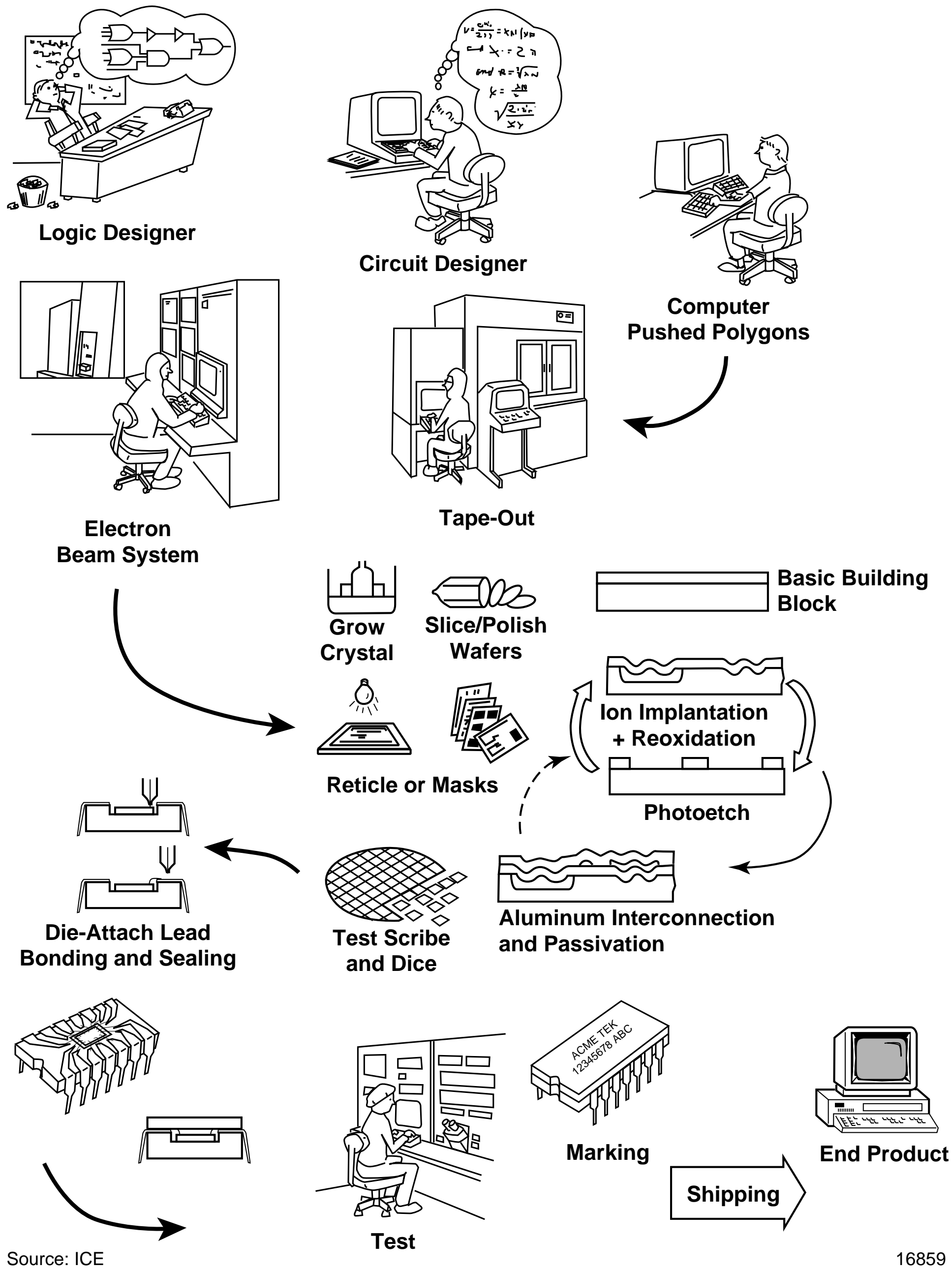


Figure 2-121. Integrated Circuit Manufacturing Process

The geometrical layout is the final output of the workstation in the form of a database tape. The database tape is the input information for the electron-beam system. The electron-beam system uses the input data to create the reticles or masks required in IC manufacturing. The number of reticles or masks is determined by the actual manufacturing process cycle. This is often referred to as the "Fab Process." The newer Fab Processes use between fourteen and twenty-four reticles or masks.

The central region of Figure 2-121 depicts the wafer fab process. After the wafer fab process is completed, the wafers are tested electrically to the required specifications and then forwarded to the assembly process.

The assembly process will package each electrically good die. After packaging is completed, the device will be given a final electrical test, burn-in as required, and shipped to the end user. The end user creates the electronic system by bringing together all the necessary electrical components, mechanical hardware and the final package of the product.

There is no doubt the world is in the silicon age. Silicon devices in the form of discrete products or ICs touch our lives in many ways everyday. And — since silicon is the second most abundant element in the earth, there doesn't appear there will ever be a shortage of the raw material.