P. A. Totta
R. P. Sopher

SLT Device Metallurgy and its Monolithic Extension

Abstract: The glass-passivated, face-down semiconductor chip joining technology employed in IBM's SLT (Solid Logic Technology), has become not only a fundamental element in the hybrid circuitry of System/360 but also the basis for later metallurgical designs. The "flip-chip," copper ball terminal, solder reflow technique is comprehensively reviewed and a discussion is given of its extension, through the use of ductile, all-solder terminals, to monolithic applications.

Introduction

The original design[1-3] for Solid Logic Technology (SLT) resulted in a family of hybrid microcircuits that has become the logic foundation of the IBM System/360 computer line; it embodies a number of features that have greatly influenced the subsequent development of IBM semiconductor metallurgy. This paper will describe the evolution of that metallurgy from its basis in SLT to more recent technology in SLT-related monolithic components. Emphasis will be on the metallurgical joint, or interconnection, between active semiconductor devices (chips) and their carrier modules (electrically inactive ceramic substrates bearing conductive lands). Although it can take any of several forms, this interconnection is a necessary part of any type of integrated or hybrid technology; as miniaturization increases and the demand for components continues to grow, its reliability and compatibility with automated manufacture become increasingly important. The discussion below, and the papers that follow,[4-7] will describe some of the work IBM has performed with those objectives in mind.

SLT design

The characteristics of SLT design and fabrication are illustrated in Figure 1; some of its more important features include the following:

(A) The module substrate is made from a high density Al_2O_3 and its surface bears screened and fired thick film ($>1\mu m$) resistors and electrode lands (circuit interconnections).

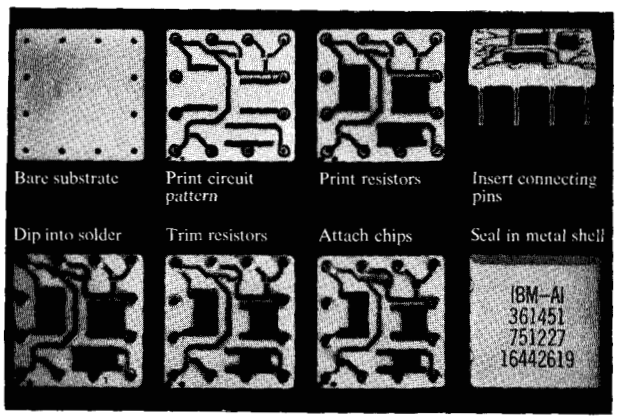


Figure 1 Sequence of fabrication steps used in manufacturing SLT hybrid logic circuits.

(B) The electrode lands are coated with high lead content solder (90%Pb-10%Sn) to improve the conductivity of the interconnecting circuit and to facilitate the attachment of "chip" semiconductor devices by reflow soldering. The high lead content of the solder ensures a melting point (300°C) well above that of the Sn-Pb eutectic (183°C) used to solder the module to the next level of packaging.

(C) The active devices, silicon transistors and diodes, are not conventionally sealed in hermetic packages to protect the semiconductor surface and the thin film ($<1\mu m$) metal interconnections. Instead, the face side of each "chip," a 0.027 in. square piece of single crystal silicon containing one or more transistors or diodes, is hermetically sealed under a thin film of glass.

P. A. Totta is at the IBM *Components* Division laboratory, East Fishkill, New York; R. P. Sopher is at the IBM Components Division laboratory, Burlington, Vermont.

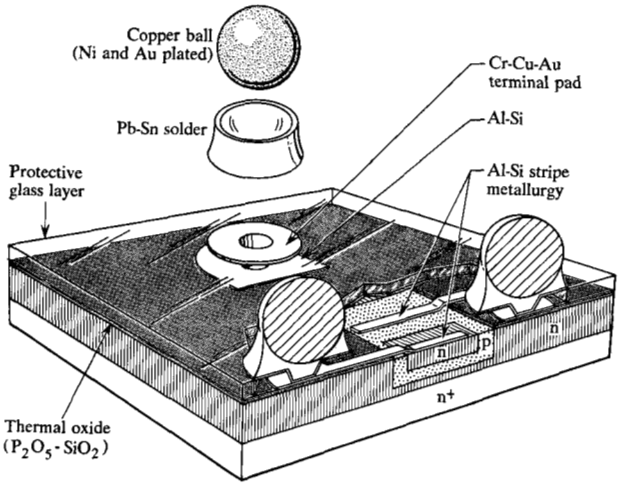


Figure 2 Basic metallurgy-glass design for all SLT transistor devices.

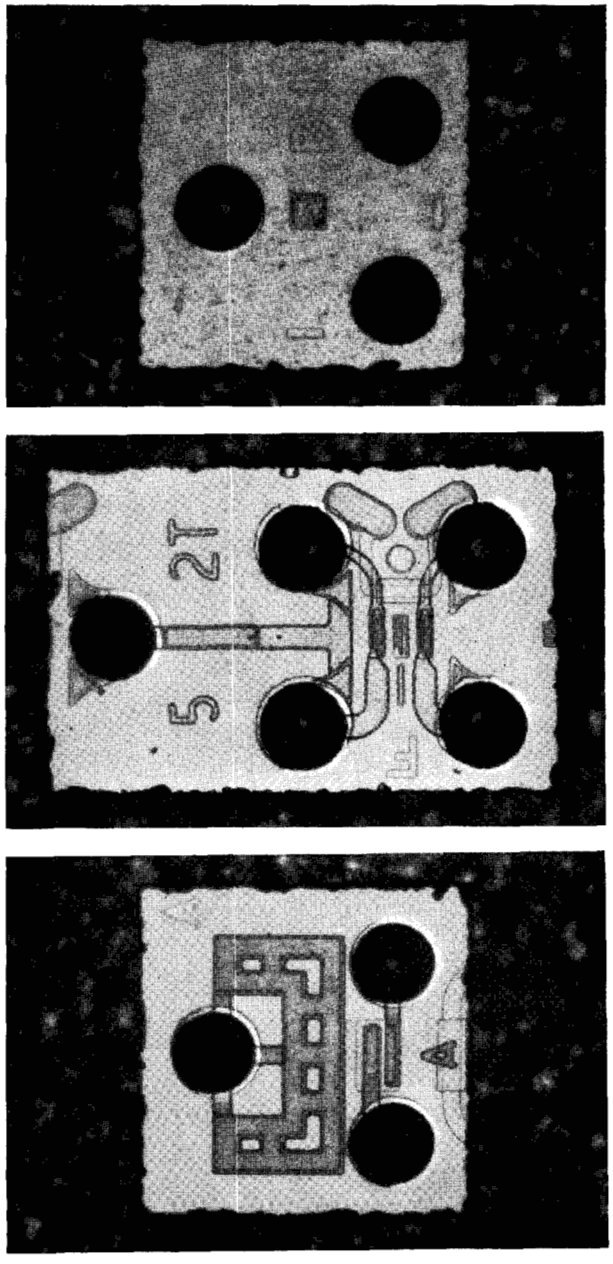


Figure 3 SLT chip transistors and diodes of various types used in all SLT hybrid integrated circuits.

(D) The chips are squarely cut, or "diced," from a previously glassed wafer of semiconductor material. Having very uniform dimensions, the chips may then be handled automatically during testing and chip placement. The chip edges, however, are bare or unpassivated, and the projections on the chip face must hold the device sufficiently high above the substrate surface that the solder on substrate lands will not electrically "short" the bare edges of the chip.

(E) The interconnection between chip and substrate is made by soldering the chip to the module in a face-down position. Three or more metallic projections (soldered spheres of Cu) on the face of the chip connect the thin film chip metallurgy to the thick film substrate conductors.

The conditions described above imposed a need for a unique chip metallurgy design. The final interconnection design is illustrated in Fig. 2 and embodies some specific departures from previous interconnection metallurgy: There are Al-Si alloy interconnection stripes under the glass, Cr-Cu-Au terminal pads to seal the terminal holes in the glass, and soft soldered Cu sphere terminals to provide stand-off capability. Some of the transistor and diode chips, containing the described metallurgy and used throughout all of IBM's hybrid module families, are shown in Fig. 3.

The combined chip and substrate design for SLT has resulted in a product that can be manufactured by a highly automated process and has led to some significant advances in the microcircuit field. First, of course, is the success of SLT itself. The millions of SLT modules that have now been made have exceeded all requirements in reliability, cost, and performance in System/360. Second, the knowledge gained from SLT has permitted the development of three other hybrid circuit families of higher circuit density for use in applications having specialized requirements. The variations on SLT that have evolved since the announcement of System/360 in 1964 are the following: ASLT (Advanced SLT), a two level module used in the very large computer systems; SLD (Solid Logic Dense), a "highly populated" version of SLT having two to three times the number of circuits per module; and ULD (Unit Logic Device), a smaller version with

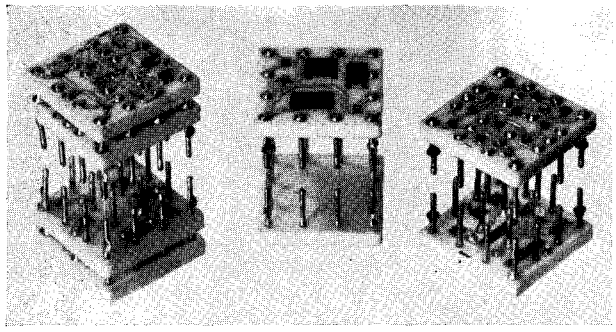


Figure 4 Comparison of ASLT (left), SLT (center), and SLD (right); bottom views are illustrated by mirror reflections.

clip connections instead of pins for use in the Instrumentation Unit of Saturn V for the Apollo moon flight. The commercial modules are compared with SLT in Fig. 4.

Third, and particularly important now, the investigation of some of the more difficult problems in SLT has resulted in an important byproduct—new thin film glass and metal technology for use in advanced design. Some of the SLT metallurgy and glass design features are directly adaptable to "flip-chip" monolithic circuit applications while others, not directly useful beyond the hybrid level, have provided the production and field experience that make newer technology possible.

The following sections discuss several considerations in SLT metallurgy that have been valuable in all three respects. The first two sections describe the device stripe metallurgy and terminal system, respectively; the third section describes the final fabrication, testing, and joining of the semiconductor chips. A final section describes the extension and modification of the technology to monolithic applications.

Al-Si stripe metallurgy

The earliest successful means for glass passivating a planar junction semiconductor wafer at IBM was the technique of fusing together a colloidal frit of glass particles on the wafer face, through uniform compaction on the surface by a centrifugal deposition technique.[8] The many kinds of glass which could be deposited by this method comprise a wide range of physical, mechanical, and chemical properties and a number of different melting points. Generally, the glass could be fired in nitrogen or a slightly oxidizing atmosphere, but it would usually exhibit its most desirable properties when fired in an oxidizing, as opposed to a reducing, atmosphere.

The choice of metals having good conductivity for stripe interconnections on devices, and able to withstand the oxidation and high firing temperature of glassing, was limited. Of those which were investigated, only aluminum seemed to offer the desired properties in a single-metal system. It could tolerate an oxidizing glass firing atmosphere with very limited oxidation and, therefore, maintain its conductivity. It required no supplemental metals to make ohmic contacts nor to provide adhesion to the underlying thermal SiO_2. Because it was a single-metal system, it could be etched to good definition. However, the comparatively low melting Al-Si eutectic temperature, 577°C, prevented the use of what would otherwise have been the more desirable glasses; these, unfortunately, had fusion points above 577°C. It was especially unfortunate that the glasses which best matched the coefficient of thermal expansion of Si also had the highest melting points. There was, however, one type of lead borosilicate glass* which could be fired just below the Al-Si eutectic temperature (in the range of 560-570°C) and did have the desired chemical and electrical properties. Upon cooling from the fusing temperature the glass is left in an undesirable state of tensile stress; however, if the film thickness is restricted to 1.5μm or less, cracking or crazing does not occur. This glass is the one which has been used predominantly for passivating SLT transistors.

Aluminum has been used for some time in the industry as the standard metallization for planar devices. For use on unpassivated chips in hermetic cans, the Al lands are usually sintered to the Si contacts at about 450° for a short period of time to achieve low resistance ohmic connections. However, in the SLT glass passivation process, the glass firing temperature is 100°C higher than the usual sintering temperature, and aluminum sintering is therefore achieved automatically during the fusing of the glass frit.

Because of the exposure of the Al-to-Si ohmic contact to the high glass firing temperature, a phenomenon was encountered in the development of SLT transistors which, at first appearance, suggested a liquid phase or eutectic Al-Si alloying problem in the contact holes. The nature of the problem is illustrated in Figs. 5 (top view) and 6 (sectional view). The cross section shows that a deeply penetrating zone of Al is approaching the emitter junction where it could short circuit the emitter-base junction. The zone is located only at the end of the contact hole where a long Al stripe over oxide approaches the hole; metal penetration is usually absent elsewhere along the contact hole.

The problem may sometimes be detected by visual observation through the glass, as in Fig. 5. When penetration is extensive, the zone appears as a slight, darkened depression at the oxide step into the contact hole; however, the depression depth does not conform to the zone depth, as is shown in Fig. 6. Some light-appearing, lateral penetrations of Al may also be seen extending beyond the edge of the etched land that lies between the thermal SiO_2 and the Si (Fig. 5). Later work showed that the same kind

* Drakenfield E1527: SiO_2, 29.1%; PbO, 51.1%; B_2O_3, 9.0%; Al_2O_3 3.2%; Na_2O, 1.7%.

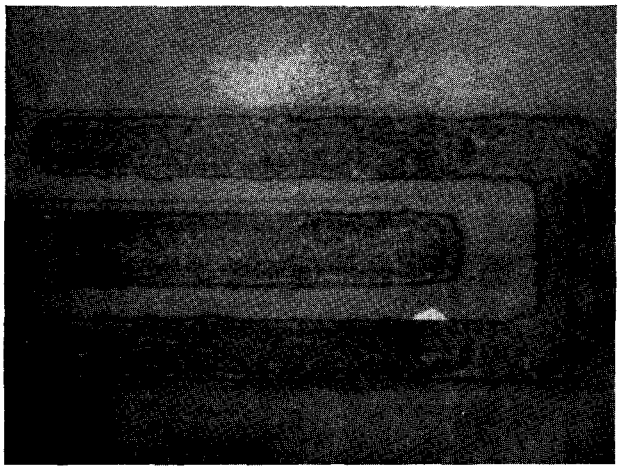
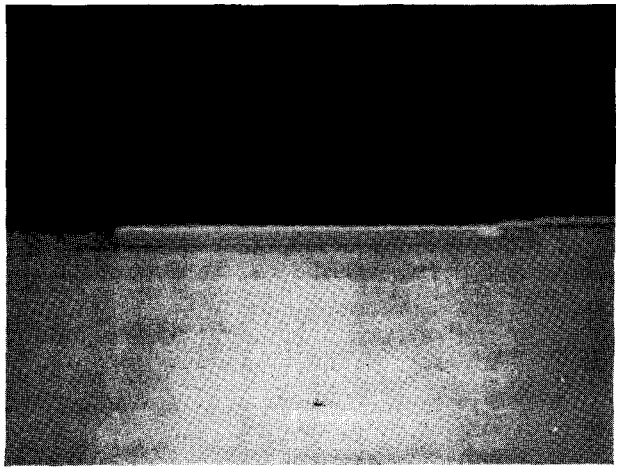


Figure 5 An SLT transistor having pure Al lands glassed with lead borosilicate at 568°C. Note the dark areas at the ends of contact holes where the Al conductor stripes first enter the holes. The bright triangle above the right end of the lower base contact hole is Al which has penetrated horizontally into the Si under the thermal SiO_2 (600×).

Figure 6 A longitudinal, 90° section of an Al emitter contact after glassing at 562°C. Note the deep, vertical penetration of Al into the emitter silicon at the oxide step-down, which approaches the junction (1300×).



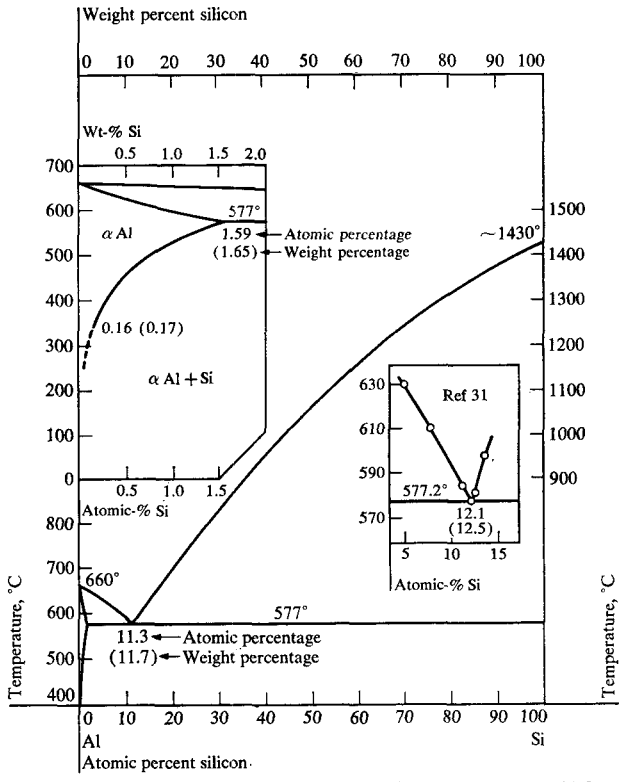
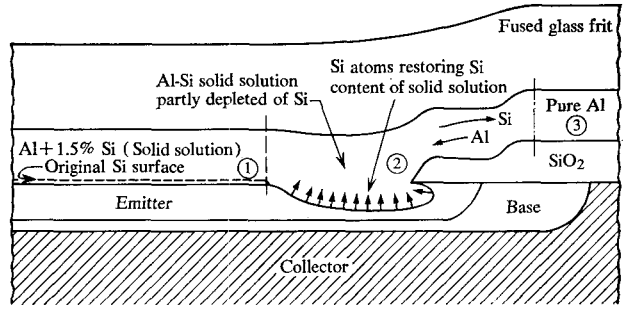


Figure 7 Aluminum-silicon equilibrium diagram. (After M. Hansen, in *Constitution of Binary Alloys,* McGraw-Hill Book Co., Inc., New York 1958.)

Figure 8 Hypothetical model demonstrating how solid state diffusion mechanisms may work to cause localized penetration of the Si by the Al at the oxide-to-silicon step during glassing.



of "alloying" did not require temperatures as high as 565°C. Similar penetrations may be seen at temperatures as low as 450°C, but as the temperature decreases, the penetrations are more shallow and thus more permissible; for deep junction devices such penetration can be tolerated.

To explain these "alloying penetrations," a hypothesis was formulated which fits the observations well and provides an acceptable solution to the problem: The hypothesis contends that the Al penetrations are not liquid phase alloying manifestations, such as might result from accidental excursions above the Al-Si eutectic temperature, since the penetrations may be seen at temperatures 100°C below the eutectic in both the base and emitter contacts. The explanation is based, rather, on the fact that there is a high solid solubility limit of Si in Al at the glass firing temperature (approximately 1.5% at 565°C; see Fig. 7), and the time at temperature is sufficiently long (about 20 minutes) that the interface between Al and Si will begin to follow equilibrium demands.

Figure 8 illustrates the mechanics of the hypothesis by showing a model cross section of an extended Al land, such as an emitter or base stripe, at a contact hole. The

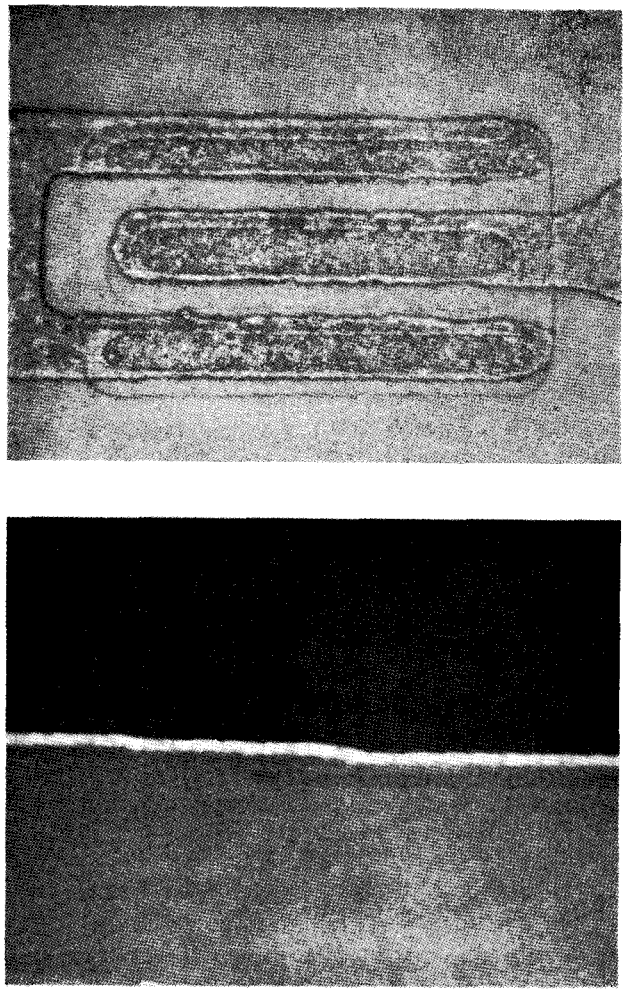


Figure 9 Result of using Al-1.5% Si alloy instead of pure Al for SLT contact stripes after glass firing at 560°C for 20 minutes: (a) top view through glass; (b) longitudinal section through emitter (SLT transistor).

figure suggests the kinetics of the diffusion process that takes place during the 20 minutes of glass firing at 565°C. In a time period early in the glassing cycle, the pure Al directly over the contact hole (segment 1) dissolves a very thin, planar layer of Si from the surface of the contact hole to satisfy the 1.5% solid solubility requirement at that temperature. Momentarily then, a solid solution exists over the contact hole and pure Al over the oxide (Segment 3). In the next period of time, Si from the (hypothetical) solid solution begins to diffuse into the pure Al, and Al diffuses in the reverse direction (Segment 2). If replacement Si from the crystal at the initial step-down point can go into solution more rapidly than Si can diffuse out of the solid solution further to the left, the result is a deep, wide "penetration" of Al into Si which is preferential at the stepdown as shown in Figs. 5 and 6.

The problem was solved in practice by pre-satisfying the 1.5% Si solid solubility requirement of the Al. An Al-Si alloy film is deposited[9] and etched for transistor contact stripes. The Al-1.5% Si alloy stripe has no tendency to diffuse into or penetrate the contact hole at the step-down (Fig. 9). The deposition of this alloy was initially achieved by codeposition from two sources; however this was difficult to control without sophisticated sources and rate monitors. Codeposition was eventually replaced by the flash evaporation of Al-3% Si alloy spheres. Silicon has a lower vapor pressure than Al; flash evaporation thus becomes necessary because slow evaporation of the alloy would cause preferential distillation of the Al. The Si content of the charge is intentionally higher than the 1.5% required in the evaporant film, because the technique does not efficiently transfer all of the Si from the preheated boat to the substrates.

At room temperature, the solid solubility requirement of Si in Al is, of course, much smaller than at the glassing temperature. Consequently, the excess Si precipitates throughout the film as a second phase of nearly pure Si. Visually, the precipitate gives the film a "peppered" appearance. Since there is so little solubility of Si in Al at room temperature, the apparent effect of the Si doping on the sheet resistance of the metal film after glassing is very small.

Terminal metallurgy

The thin film layer of glass, which seals the Al-Si contact "stripe" and the device junctions from the ambient atmosphere, must be penetrated in at least three places to make electrical connections to the emitter, base, and collector of a chip transistor. The penetration is achieved by chemically etching the glass through a photoresist mask with an HF-containing solution. Once the holes are opened, they must be hermetically sealed again with evaporated discs, or pads, of thin-film metal that are larger in diameter than the holes in the glass. Each pad must adhere tenaciously to the glass and the exposed spot of Al-Si metal film. This is necessary to effect a good seal, to make a stable, low resistance metal-to-metal contact, and also to provide a mechanically sound anchorage for the external terminal.

- *DC sputter cleaning of terminal via holes*

The chemical etching of holes in the fired frit glass, to expose the Al-Si contact stripe, is not effective in removing all the glassy residues in the bottom of the hole. X-ray diffraction reveals that beta SiO_2 remains on the surface of the Al-Si after glass hole etching. This residue is not removed by any usable chemical etchant and results in very high interface resistances. Also, in the chemical removal of photoresist used to mask the surface of the glass during hole etching, residues are left on the surface of the glass in

areas where the terminal pad is supposed to bond tenaciously. These latter residues greatly degrade the adhesion of Cr films and could cause mechanically unsound terminals.

The first of the problems, the high resistance terminal holes, and the second, the non-adherent terminal pads, have been solved together for SLT by dc sputter cleaning the substrates.[10] The Cr-Cu-Au terminal pad evaporation is accomplished by evaporating metal films through apertures in a thin Mo mask (0.004 in. thick). The holes in the metal mask are larger than, and concentric to, the holes in the glass film. It was found that by applying a large negative bias, of the order of 1000V, to the metal mask in a low partial pressure of argon gas (approximately 40 μm), the spot of glass and Al-Si visible through each aperture will be subjected to ionic bombardment (sputtering) of such intensity that the surface of the glass becomes lightly etched, and the glassy residue in the bottom of each hole is completely removed.

The primary target of ionic bombardment, the cathode, is actually the negatively biased Mo mask which provides electrons for Ar ion neutralization. However, the holes in the mask are so small, relative to the area of the Mo mask, that much ion bombardment and etching takes place in the mask holes as well as on the mask surface. The type of etching which results is much more direct and effective than the usual glow discharge cleaning of substrates in which substrate surfaces are mildly "scrubbed" by transient, low energy ions moving toward a more remote cathode. In SLT practice, about 500Å of glass (or glassy residue plus Al) are removed at each terminal site. The argon is then pumped out of the system, and the metal evaporation commences without breaking vacuum. The resultant metal-to-metal interface resistances are very low (of the order of 5 milliohms or less for a 2.3 mil diameter hole) and the adhesion of chromium (the adhesive metal layer in the pad) to glass is excellent.

• *Cr-Cu-Au terminal pads*

The terminal pad metallurgy or "ball limiting metallurgy" shown in Fig. 10 was chosen for SLT to embody a composite, multi-layered film of Cr, Cu, and Au. Each layer has a function in the composite.

The Cr, which is sublimated from a resistance heated source, is the first metal deposited through a metal mask after dc sputter cleaning. It is the adhesive film which attaches the terminal pad to the glass and to the aluminum in the bottom of the glass hole. In addition, the Cr layer, approximately 1500Å thick, is the primary diffusion barrier in keeping molten Pb-Sn solder from reacting with the metal film under the glass. The Cr film is under residual tensile stress and must be restricted in thickness to avoid cracking the already stressed, fired frit glass at the perimeter of the terminal pad.

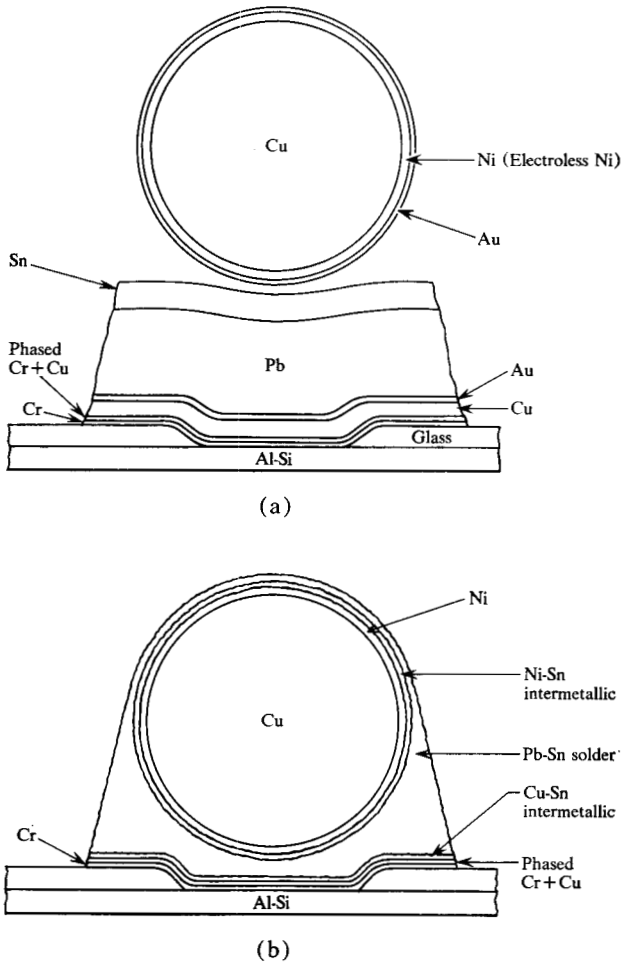
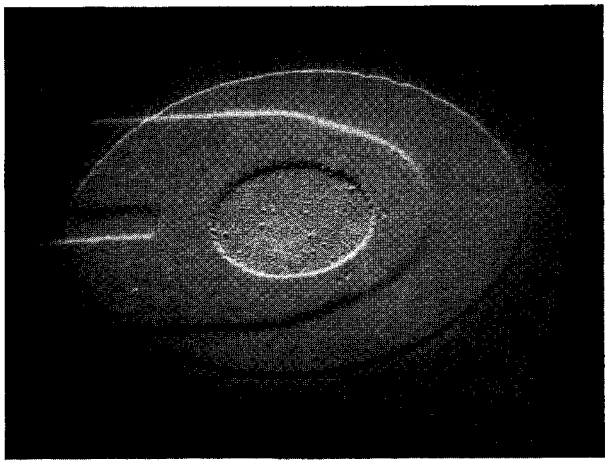


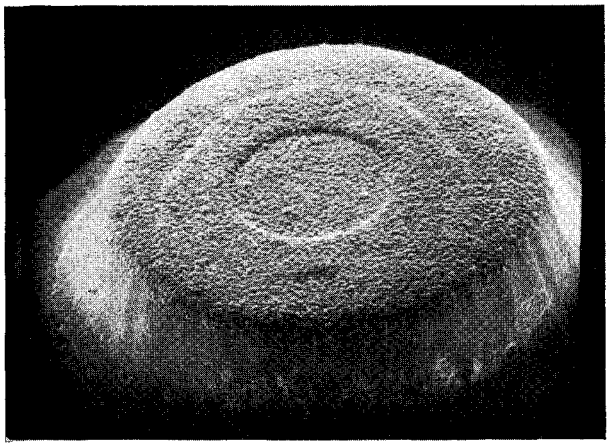
Figure 10 SLT ball terminal before and after Cu ball soldering.

The Cu is the solderable layer in the composite. The first 500Å of Cu are deposited or phased with the last 500Å of Cr. Cu interacts with the molten Pb-Sn solder by becoming wetted and forming Cu-Sn intermetallics during soldering and reflow operations. The entire film, a maximum of 5000Å thick in SLT, may be completely converted to Cu-Sn intermetallics in two to three solder reflow cycles. Were it not for the phased layer in which Cu is interlaced with the Cr, all of the Cu-Sn intermetallic layer would spall off and dissolve[11] in the Pb-Sn solder during the subsequent solder reflow cycle that joins the chip to the module. Complex modules (such as ASLT types) require several reflow soldering operations; therefore the terminal pad metallurgy must meet the conditions of multiple solder reflow, and these have been satisfied by the intentional use of overlapped Cr-Cu deposition.

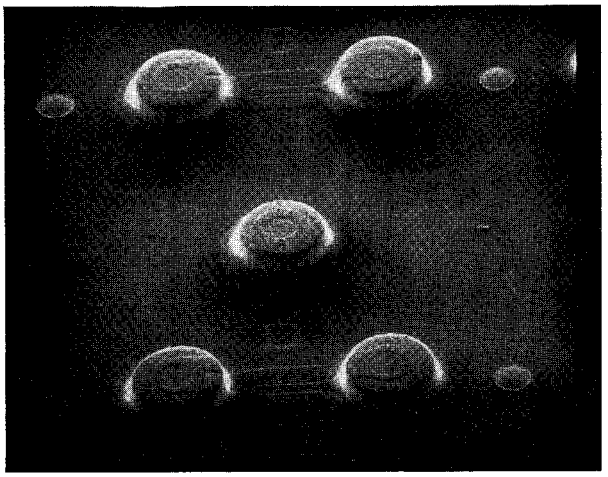
A thin Au film of 1000Å is deposited above the Cu to provide a film of noble metal which will protect the Cu



(a)



(b)



(c)

Figure 11 Scanning electron microscope views of evaporated terminal metallurgy: (a) the Cr-Cu-Au ball limiting land (550×, 45° angle); (b) evaporated Pb-Sn solder deposit (550×, 45° angle); (c) several preplaced solder pads (110×, 45° angle).

Figure 12 Copper ball loading of a pre-aligned metal mask.

surface from oxidation. The Au would not be necessary if the substrate were cooled thoroughly after ball limiting evaporation, before exposing the Cu to air. However, in manufacturing, it is more expedient to expose the system to air before the wafers are at room temperature; therefore the non-oxidizing Au film is necessary.

• *Solder preplacement*

Various methods were investigated for preplacing uniform volumes of Pb-Sn solder on the terminal pads of a silicon wafer. The methods included molten deposition by immersion in solder or by passing the wafer through a solder wave machine, silk screening, solder ball deposition through a metal mask, and solder electroplating of the Cu spheres. All of these techniques, however, led to unsatisfactory results. Molten application crazed the glass due to thermal shock; it also caused solder bridging between terminals. Silk screening was not uniform and involved undesired flux vehicles. Solder balls were difficult to load, and expensive. Solder plating did not provide enough solder for good fillets.

The method of preplacement of solder which has worked best for SLT is the vacuum evaporation[12] of the lead-tin alloy. The masked wafer, after the deposition of Cr-Cu-Au in one system, is transferred to another evaporator for Pb-Sn deposition. The Pb-Sn evaporator is arranged so that the substrates are in a small cone angle (~30°) over the source, for maximum deposition efficiency. While the total evaporant thickness in the Cr-Cu-Au evaporator is less than 1μm, the evaporant thickness in the Pb-Sn

evaporation is approximately 38μm (1.5 mils). The evaporation of such a thick film would be impractical except that the vapor pressure of Pb is quite high, and 95% of the solder is Pb. Although a charge of Pb-Sn alloy is used in the rf heated crucible, the fractional distillation of the Pb and Sn results in a two-layer deposit. The Pb is the very thick bottom film which is deposited first on the Cr-Cu-Au, and the Sn is a thin cap. The fact that the Pb is deposited directly on the Au has been found to enhance the adhesion of the solder because of the low-melting Au-Pb eutectic (219°C) that forms to a small extent during Pb-Sn evaporation. The Pb and Sn are homogenized in the solder joint during the subsequent soldering of the Cu sphere to the wafer. All metal evaporation masks are cycled through a cleaning step prior to the next Cr-Cu-Au deposition.

The uniformity of the Pb-Sn preplaced by the described evaporation technique is far better than any other method explored. The evaporant spots are extremely similar in volume and composition (Fig. 11). The scanning electron micrographs show that the evaporated spots of solder replicate the wafer surface so well that the terminal hole-in-glass can be seen perfectly above 1.5 mils of solder.

• *Cu ball terminals*[13]

After solder deposition, the wafer is aligned under a second metal mask having aperture holes slightly larger in diameter than those in the terminal pad mask. The surface of this mask is flooded with miniature Cu spheres which are very uniform in size (0.005 $\pm$ 0.001 in. diameter). These spheres have a plated layer of electroless Ni (1.25μm) and a thinner layer of electroless Au (0.25μm). The former reduces the amount of intermetallic formation on the ball, while the latter keeps the surface of the Ni plating solderable during storage. By a succession of vibratory motions and the repeated addition of Cu spheres (Fig. 12), the mask holes are completely filled; the superfluous spheres are swept away to be reused. The masked wafer is then put through a belt furnace which is carefully controlled in temperature profile and atmosphere. The main heating chamber contains pure hydrogen gas, which acts as the flux in reducing lead oxide and in promoting wetting of both the copper ball and the terminal pad. The wetting is of such excellent quality that the solder wets even the top of the ball while it flows and forms a good fillet between the ball and the terminal pad (Figs. 13 and 14). The

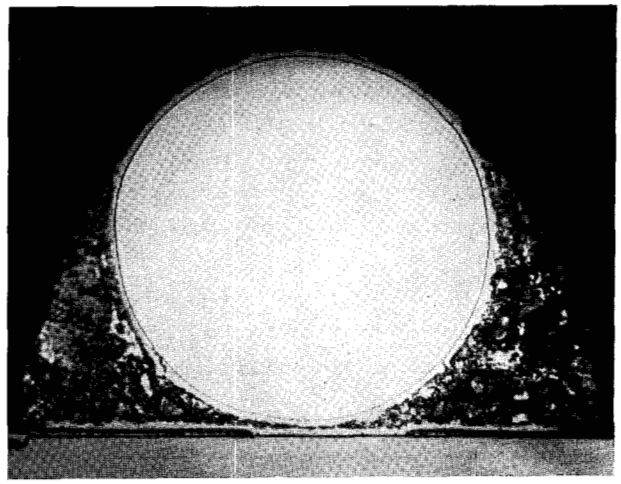
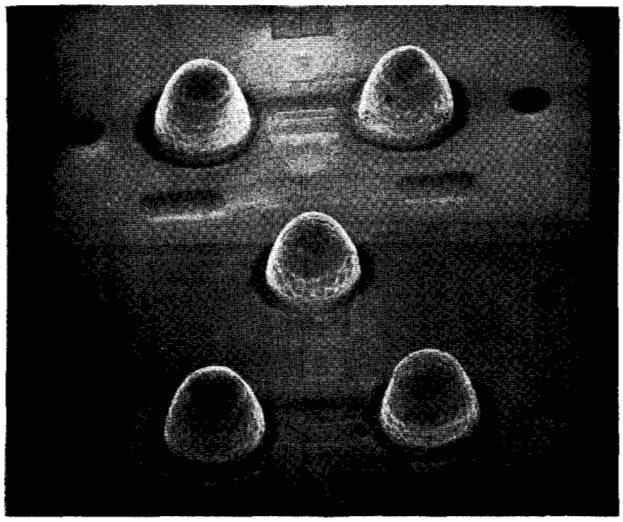
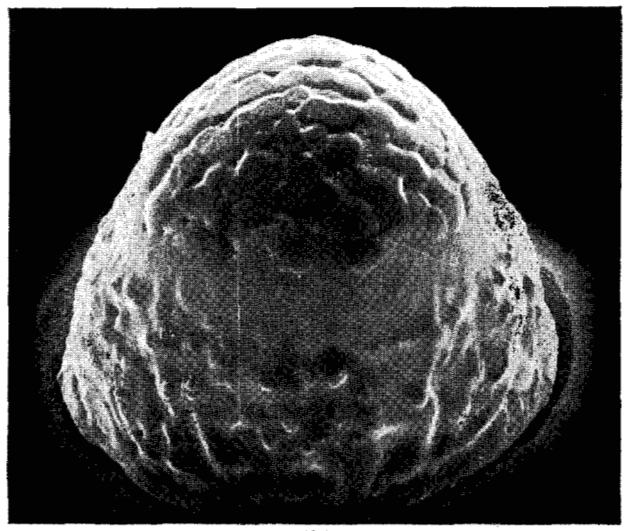


Figure 13 Photomicrograph of SLT copper ball terminal after ball soldering (540×).

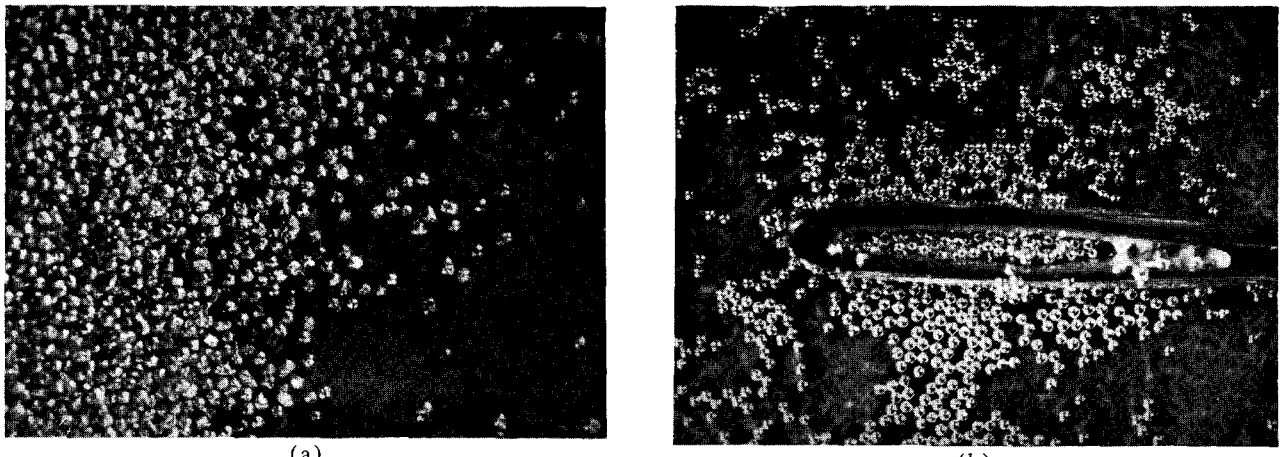
Figure 14 Scanning electron micrographs of wafer after copper ball attachment (a) several terminals on wafer (110×, 45° angle); (b) single terminal with Ni plated Cu ball completely covered with solder (550×, 45° angle).



(a)



(b)



(a) (b)

Figure 15 Copper ball manufacturing: (a) commercial spherical powder (as received); (b) after reprocessing.

solder on top of the ball was, at first, considered undesirable because of possible lead oxide formation and contamination of the automatic tester probes. To the contrary, it was later found the compliance of the Pb-Sn to the probes reduced contact resistance and resulted in higher final test yields.

• *Copper ball manufacture*

If one device wafer requires 3900 Cu spheres, an obvious manufacturing requirement was a source of uniform-size, excellent quality Cu spheres at a reasonable cost. A method was developed which is quite simple and economical, and which produces high-quality spheres at an excellent output rate.

IBM makes its Cu spheres for semiconductor use by reprocessing commercial spherical powder of the powder metallurgy type used in some kinds of liquid filters. This copper powder, of a "tough-pitch" Cu quality of composition, containing detectable amounts of copper oxide, is not sufficiently spherical to be used for the SLT application. It may, however, be purchased in quantity for less than one dollar per pound and is a good starting material for the ultimate product. A very narrow powder fraction of particle sizes which bracket the desired 0.005-in. ball diameter is screened out for use.

If this spherical Cu powder is blended with aluminum oxide powder of the micron-size kind used for metallographic polishing (nominally 10 parts of Al_2O_3 to 1 part of Cu powder), most of the Cu particles become completely surrounded by aluminum oxide particles. When this mixture is put through a hydrogen furnace at a temperature about 100°C above the melting point of Cu, several events take place. The somewhat irregular Cu spheroids are melted, and surface tension reshapes each mass into a nearly perfect sphere. The non-wettable alumina powder continues to keep most of the spheres separated. The hydrogen gas reduces the residual copper oxide impurities within the spheres while the remelting frees entrapped gaseous inclusion, thus producing higher purity, oxygen free copper particles, with few voids or inclusions. The cooling of the Cu-alumina mixture results in a mass of nearly perfect micro-spheres in a matrix of loose alumina powder. The alumina is easily separated from the Cu by screening, and is recycled. The Cu spheres, plus some Cu agglomerates or clusters, are again screened through powder metal sieves and separated into appropriate powder fractions. Then the desired range of sphere sizes is fed into a micrometer separator which consists of two rotating, gradually diverging rollers. The spheres are dropped through the rollers into bins which have a grading capability of 0.0001 inch.The proper size spheres, 0.005± 0.0001 in., are then electroless Ni and Au plated, counted (weighed), and stored. One million Cu spheres of this size weigh only 10 grams (Fig. 15).

Dicing, testing, and chip joining

The scribe-and-break technique of dividing a silicon wafer into chips tends to produce non-square edges when the plane of the crystals is ⟨111⟩. Also, the scribe technique tends to shatter the edges of the glass passivation films on the SLT wafers unless the glass is etched away in the kerf. From the beginning of SLT it appeared desirable to highly automate the testing and placement of chips; therefore, regularity in the size and shape of the chips was a necessity. Initially, ultrasonic cutting of the glassed Si wafers was used, in which metallic grids, of dimensions the same as the chips, propelled an abrasive slurry through the wafers to separate individual chips. Chip edges resulting from

using this technique met the SLT requirements, but the technique was very slow and expensive. Therefore, a more practical dicing technique, the slurry saw,[14] was devised.

The slurry saw consists of mandrel-mounted gangs of 0.002-in. thick, stainless steel, circular blades which are spaced from each other with precision ground discs of smaller diameter. These blades, which are toothless, rotate at approximately 12,700 rpm in an aqueous mixture, or slurry, of SiC particles and water. Each rapidly rotating disc transports abrasive particles to the Si with such effectiveness that a slit is abraded in the Si very quickly, with little or no fracture of glass or Si. The Si wafer, which is bonded to a plastic carrier with a thermoplastic cement, is aligned on a cutting block relative to the cutting pattern of the blades with the help of an optical comparator. With a single plunge cut, the wafer is cut into strips in one direction, rotated 90° and cut again (Fig. 16). The entire cutting operation takes approximately two minutes per wafer.

The separated chips are then put through a sequence of automatic handling steps. The loose chips are fed into machines which use special vibratory feeders, combined with orientation sensing devices to position each chip for as many as 40 electrical tests. As they are tested, at rates of 0.3 second per chip, the devices are automatically sorted into appropriate quality groups to be used for many kinds of logic modules.

The final chip handling step, chip joining, brings together tested substrates and tested semiconductor chips of several kinds, to form a logic module of one or more circuits. The chip handling equipment,[15] basically similar to that used for chip testing, automatically orients the chip and places it in the proper position on the pretinned substrate electrodes (Fig. 17). The consistent and controlled placement of chips by machine has played a major role in reducing the failure rate of SLT modules. Excessive pressure on the back of a chip can obviously cause glass and Si damage under terminals; machine placement has virtually eliminated such damage.

A small amount of water white rosin flux is deposited on the substrate electrode tips just prior to chip placement. The flux enhances the reflow soldering and acts as an adhesive to keep the chip properly positioned until it is soldered in place. Soldering takes place in a nitrogen atmosphere, unlike the previous ball soldering step in wafer fabrication, because the thick film, precious metal-oxide resistors on the substrate would decompose and degrade in a reducing atmosphere such as hydrogen or forming gas.

During the reflow soldering operation, at approximately 335°C in a belt furnace, both the 90%Pb-10%Sn solder on the substrate electrodes and 95%Pb-5%Sn solder on the chip are remelted. The solders on the substrate and the chip therefore begin to mix and homogenize. The joint may experience several subsequent reflow cycles after

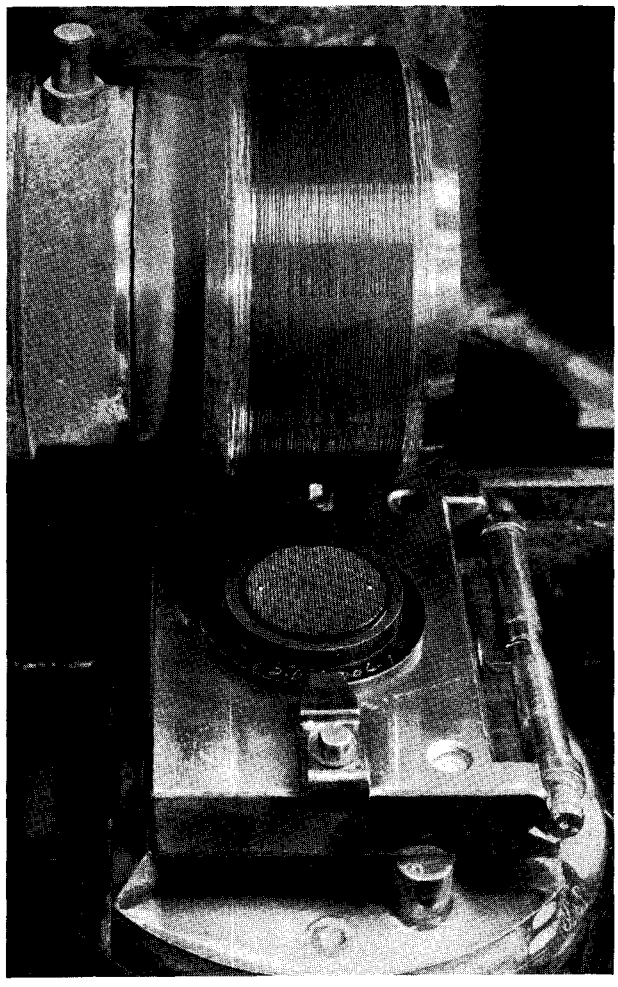
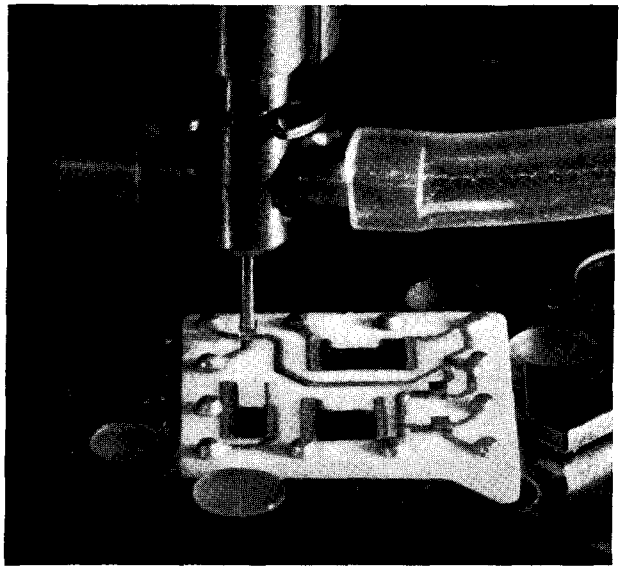


Figure 16 Slurry saw dicing.

Figure 17 Automated chip placement operation.



the initial chip joining cycle, since a module may be a part of a more complex module which is two-high, or the module may be repaired in a rework operation.

The joint strength of the soldered SLT chip is responsible for great mechanical stability in the electrical interconnection between chip and substrate. Approximately 300 grams of normal force are required to pull a three terminal SLT chip away from the substrate. Since a single, 1-mil-diameter thermo-compression bonded gold wire usually withstands less than 10 grams of load before breaking, it is apparent that the larger connection area of the solder fillet joint on SLT has gained a strong mechanical advantage over wire bonds, plus the manufacturing advantage that all joints on all chips on a module are solder reflowed at one time. The fracture in a tensile-loaded terminal joint usually occurs between the Cu ball and the chip terminal pad. Typical fracture analysis indicates that the primary course of fracture is through the solder, although a small amount of fracture surface may occasionally be found at the solder to Cu-Sn intermetallic interface. In special instances, when a module has been reflowed many times, the electrode paste-to-substrate interface will separate before the solder joint fails.

- *Reliability*

Chips joined to SLT modules have been subjected to many kinds of life tests including thermal cycle, vibration, shock, and centrifuge. The mechanical stability of the strong copper ball interconnection has been shown in these tests, and is also reflected by a failure rate in the field of 0.00004% per 1000 hours, which is comparable in magnitude to high reliability, "flight level" TC bonds but superior to most commercial-quality TC bonds. This is the actual failure experience of 270 billion ball-bond hours in field audited machines.[16]

A number of factors can be credited for the present overall reliability: automated chip placement, which consistently positions chips on substrates with controlled impact and pressure, as opposed to manual placement; dc sputter cleaning of wafers prior to terminal pad deposition for enhanced adhesion and lower interface resistance; and the continuous tightening of quality control in an active, volume production facility.

SLT device metallurgy applied to monolithics

Since the development of SLT device metallurgy for hybrid circuit technology, new materials, processes, and design innovations have appeared; the device metallurgy has consequently evolved toward new designs which have both improved the hybrid technology and found utility in monolithic applications.

Glassing is no longer restricted only to fired frit glassing technology, which requires relatively high temperatures and, usually, oxidizing atmospheres. RF sputtering[17] allows the deposition of any commercial glass or even amorphous SiO_2, Al_2O_3, or Si_3N_4. The deposition is made at comparatively low temperatures and in vacuum atmospheres containing low partial pressures of inert gases. The coefficient of thermal expansion of the glass can now better match that of silicon, and residual stresses in these films are often compressive rather than tensile.

Because of the lower glass temperatures and non-oxidizing atmospheres, metal stripe systems other than Al-Si alloy have now been considered for their merits. Even pure Al can be and is used, since the glass deposition temperatures are now low enough that the Al penetration problem observed in SLT is not appreciable. However, Al-Si alloy continues to be of interest for very shallow devices since the allowable penetration into emitter contacts during the necessary sintering step is greatly reduced.

- *Controlled-collapse interconnections*

A significant recent development in the design of chip-to-module interconnections is the "controlled-collapse" joining technique.[18] On both the chip and the module substrate, the lands can be designed[4] in such a way as to restrict solder wetting to the area around the terminal. This restricted wetting on both chip and module substrate allows the surface tension of the solder, in its molten state, to support the chip during reflow, thus preventing its collapse and the resultant chip edge shorting against the module lands. This double-sided "ball-limiting metallurgy" thus permits the use of all-ductile joints and eliminates the need for the solid, non-melting copper ball. The importance of this change lies in the dimensions of monolithic circuitry. The spacing between terminals on monolithic chips has become so small (as close as 3 mil pads on 6 mil centers) that much smaller Cu spheres would have been required. The increased handling difficulty of the smaller spheres, plus the variations in elevation among different individual contacts that could be tolerated in SLT because of the "tripod" arrangement but cannot in multiterminal monolithics, has made the all-ductile pad a very desirable element for monolithic technology.

To accommodate the pure solder joint of the "controlled chip collapse" system, however, some changes were required in the terminal metallurgy. The Cr-Cu-Au terminal pad remains essentially the same, although in some designs it was found that a thicker Cu layer was desirable for fatigue strength enhancement.[6] The compressive stresses in sputtered SiO_2 permit higher tensile stresses in the pad metallurgy, whereas fired frit glass was more restrictive as to metal film thickness and stress. Because of the restricted availability of solder on the module, much more solder is required on such a chip terminal than was necessary before.

Approximately 1.5 to 2.0 mils of solder evaporated through a metal mask was sufficient to provide a good

solder fillet for an SLT copper ball. Without the copper ball, and with a very restricted amount of solder available on the substrate electrode tips, a hemispherical volume of solder which is approximately 5 mils high became desirable for the chip. The maximum height of solder which can be evaporated through a 4-mil metal mask, without permanently locking the mask to the wafer by bridging the wafer and mask solder, is approximately the thickness of the mask. The frustum-shaped deposit becomes smaller in diameter as it gets taller. To provide additional solder when necessary, yet continue to use evaporation for solder preplacement, a solder "wet-back" technique is used. In its implementation, a special metal mask is put on the wafer after the Cr-Cu-Au evaporation which aligns an oversized, circular (sometimes oblong) hole in the mask concentrically with each circular pad of Cr-Cu-Au. The Pb-Sn solder may then be evaporated through this mask to the usual SLT thickness of 1.5–2.0 mils, or up to 4 mils where required. Solder is therefore deposited both on the terminal pad and on an annular surrounding area of glass. During a subsequent solder reflow step, in hydrogen, the glass is dewetted and the molten solder is drawn back by surface tension to the area of wettable metal of the ball limiting land (Fig. 18). The height of the resulting Pb-Sn mound is then adequate for testing, and for attachment of the chip to the module.

Figure 19 shows a multiterminal chip joined to a substrate with the new controlled chip collapse technology. Figure 20 shows a glass dam (one of several designs) which can be used to restrict solder wetting on a module electrode. A comparison of an early monolithic chip having pure solder terminals and an SLT diode chip with Cu ball terminals is shown in Fig. 21. The SLT concept of flip chip mounting with solder reflow interconnection joints, which has been used for discrete devices in hybrid circuits with great success, has therefore now been modified and developed to extend the technology to monolithic integrated circuits.[19] It is expected that monolithic circuits using this technology will exceed the performance of SLT in manufacturability, reliability and economy.

Acknowledgements

Although it is not possible to acknowledge all of the people who have contributed in some way to the chip metallurgy described in this paper, the authors wish to identify some who have contributed strongly in particular areas: L. Kuiper, Si doping of Al; F. Barson and J. Sturm, dc sputter cleaning of terminal holes; J. Langdon and R. Pecoraro, terminal pad concept; I. Hymes, ball terminal standoff; and R. Ross, Cu sphere fabrication. The authors also wish to credit J. Gajda and Mrs. A. Rifenberger for the metallography and R. Mullaney for the scanning electron micrographs that illustrate the paper.

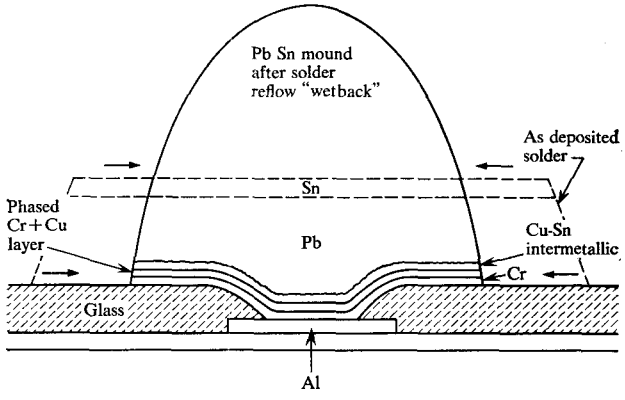


Figure 18 Solder wet-back scheme for solder pad formation on chips to be used in "controlled chip collapse" terminal connections.

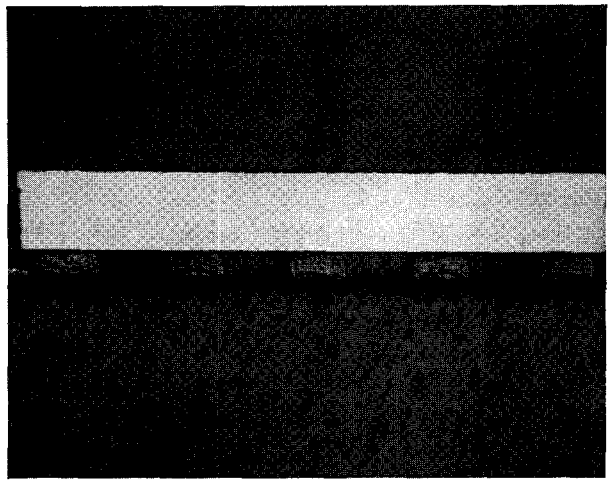


Figure 19 Multiterminal chip attached to module with limited-wetting electrodes on module (75×).

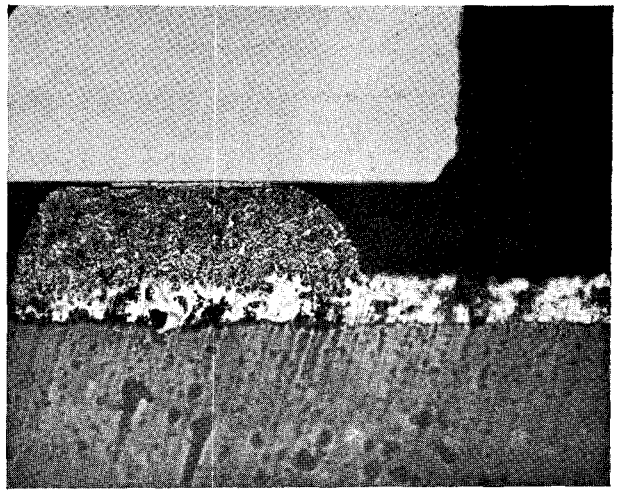


Figure 20 Enlarged view of a single solder interconnection showing a glass dam which restricts solder wetting to electrode tip (375×).

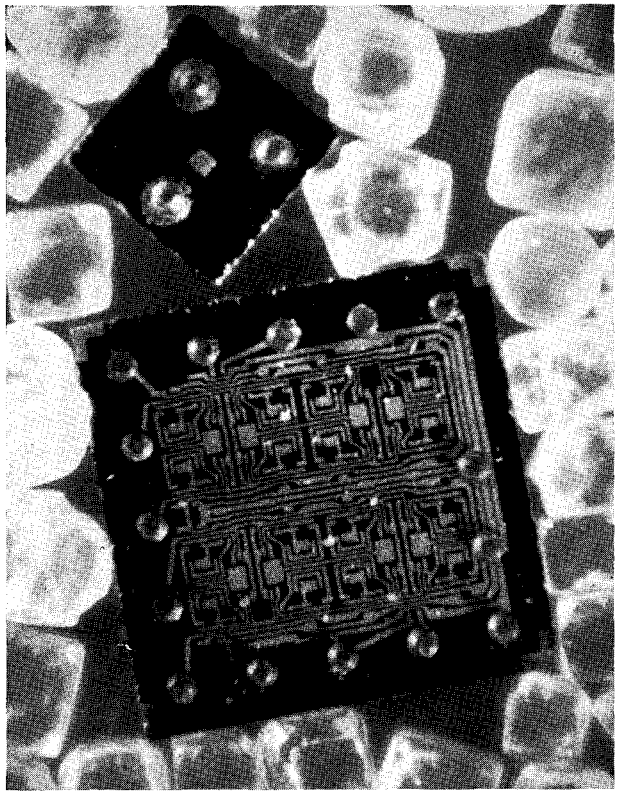


Figure 21 Comparison of an SLT diode chip with Cu ball terminals and a monolithic integrated circuit chip with solder terminals. (White crystals are table salt.)

References

1. J. L. Langdon, W. E. Mutter, R. P. Pecoraro and K. K. Schuegraf, "Hermetically Sealed Silicon Chip Diodes and Transistors," *IEEE Electron Devices Meeting,* Washington, D. C., Oct. 27, 1961.
2. J. A. Perri, H. S. Lehman, W. A. Pliskin and J. Riseman, "Surface Protection of Silicon Devices with Glass Films," *Electrochemical Society Fall Meeting,* Detroit, Mich., Oct. 2, 1961.
3. E. M. Davis, W. E. Harding, R. S. Schwartz, and J. J. Corning, "Solid Logic Technology: Versatile, High Performance Microelectronics," *IBM J. Res. Develop.* **8,** 102 (1964).
4. L. F. Miller, *IBM J. Res. Develop.* **13,** 239 (1969, this issue).
5. S. Oktay, *IBM J. Res. Develop.* **13,** 272 (1969, this issue).
6. L. S. Goldmann, *IBM J. Res. Develop.* **13,** 251 (1969, this issue).
7. K. C. Norris and A. H. Landzberg, *IBM J. Res. Develop.* **13,** 266 (1969, this issue).
8. W. A. Pliskin and E. E. Conrad, "Techniques for Obtaining Uniform Thin Glass Films on Substrates," *Electrochemical Technology,* 196, (1964).
9. B. L. Kuiper, U. S. Patent No. 3,382,568.
10. F. Barson and J. Sturm, U. S. Patent No. 3,410,774.
11. I. Ames and B. S. Berry, *IBM J. Res. Develop.* **13,** 286 (1969, this issue).
12. J. L. Langdon et al, U. S. Patent No. 3,401,055.
13. I. M. Hymes et al, U. S. Patent No. 3,303,393.
14. D. K. Seto, H. Wing, "High Speed Saw for Dicing Silicon Wafers," presented at *ASTME Meeting,* 1966.
15. W. Schuelke, "Modular Approach to System Design," *Automation* **14,** No. 4 77–83, (1967).
16. E. E. Platz, "Reliability of Hybrid Microelectronics," *1968 WESCON Symposium,* "Designing with Hybrid Microelectronic Circuits," August 21–22, 1968.
17. P. D. Davidse and L. I. Maissel, "RF Sputtering of Insulators," *3rd Int. Vacuum Congress, Stuttgart, Germany,* July, 1965; "Dielectric Thin Films through RF Sputtering," *Journ. Appl. Phys.* **37,** 574 (1966).
18. L. Miller, "Joining Semiconductor Devices with Ductile Pads," *1968 Hybrid Microelectronics Symposium, Society for Hybrid Microelectronics,* Chicago, Oct. 28–30, 1968.
19. B. Agusta, P. Bardell, and P. Castrucci, "Sixteen-Bit Monolithic Memory Array Chip," *IEEE Electron Devices Meeting,* Washington, D. C., Oct. 1965.

Received December 11, 1968