Multilevel interconnections for wafer scale integration

J. F. McDonald and A. J. Steckl
Rensselaer Polytechnic Institute, Center for Integrated Electronics, Troy, New York 12181

C. A. Neugebauer and R. O. Carlson
General Electric Corporation, Schenectady, New York 12301

A. S. Bergendahl[a)]
IBM, Essex Junction, Vermont 05452

(Received 11 April 1986; accepted 16 June 1986)

Multilevel interconnections (MLIC) are the key to successful wafer scale integration (WSI). In this paper, we present a review of various approaches used in the implementation in MLIC to both monolithic WSI and wafer scale hybrid packaging. Electrical interconnections using both thin (micron dimensions) and thick (tens of microns dimensions) dielectric and metal films are discussed from a performance standpoint. The advantages of the thick film approach for realizing transmission line performance is indicated. Techniques for implementing thick electrical interconnections such as plate-up, via etching and lift-off, are described. This treatment includes silicon-die-on-wafer, silicon-die-on-ceramic, and GaAs-on-silicon wafer hybrids. Fabrication approaches which have been explored by several research groups are compared. Some projections concerning yield of the wafer interconnections are calculated. Finally, a brief examination of some potential approaches to three-dimensional stacking of wafers is made, including two-sided population of wafers.

I. INTRODUCTION

The ever-decreasing cost per function of integrated circuits has been achieved by a combination of factors: decreasing device dimensions, novel device/circuit design techniques, and increasing chip size. The chronological development of this increasing functionality per chip is illustrated in Fig. 1. It is clear that device density increases have historically greatly outpaced chip area growth. This trend, however, may not continue as further density improvements using conventional processing technology are becoming increasingly expensive and difficult to achieve.[1] In this context, wafer scale integration (WSI) provides an opportunity to continue this dramatic trend in level of integration. For this reason, some individuals now refer to WSI as the "post-shrink" technology of the future.

In the view of the authors, wafer scale integration implies the use of the entire wafer area to achieve a complete system function, which is conventionally implemented with a multitude of individually packaged chips interconnected on a board. Due to this substantial increase in the levels of integration, WSI has the potential of greatly reducing the cost per function, achieving better performance, higher reliability, and greatly reduced overall size.

These advantages are briefly discussed below.

WSI promises lower cost/function because of a much higher level of integration and because of high yields made possible by fault tolerance.

WSI promises higher performance because it leads to the highest possible packaging density of functions in a system and thus minimum interconnection lengths and signal delays.

WSI promises highest reliability because practically all interconnections in a system are by Al metallization on silicon, which is inherently more reliable than other interconnection techniques such as wire bonds, solder connections, or the like. An additional dramatic increase in reliability can also be expected with the advent of self-reconfigurability, which makes possible failure tolerance transparent to systems operations.

WSI appears to offer an approach to dramatically higher functional density for both a developing integrated circuit (IC) technology development environment, such as GaAs, and in a mature IC technology environment, such as silicon.

WSI appears to have significant application potential, both for regular structures such as memory and irregular structures such as random logic.

There are, of course, many issues which need to be resolved before the WSI approach can succeed. A major area of study in conjunction with WSI is the multilevel interconnection network necessary for integrating the various chips

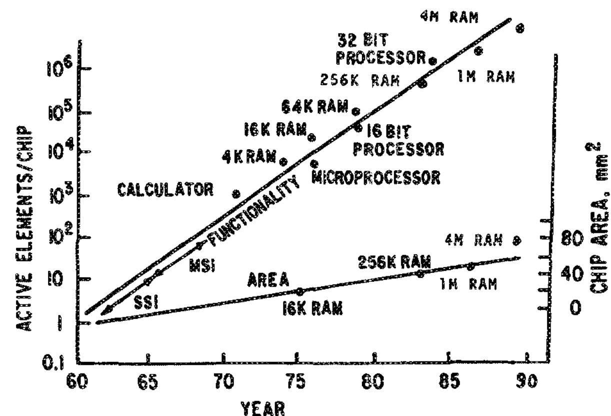


FIG. 1. Past progress in chip size vs VLSI device density indicating that in the past a dramatic increase in functionality has been achieved without increasing chip size rapidly.

 © 1986 American Vacuum Society

or components into a single working system. A very extensive network will be required for two reasons: (a) the complexity of the overall system; (b) the necessity to be able to interconnect non-near-neighbor (or even very distant) components. The second item is the result of the less than unity yield expected in WSI. Indeed, the optimal yield strategy for WSI combines a minimum amount of redundancy with a maximum chip area. The chip yield, or the probability of producing a chip with no defects, is a function of the product of the defect density (D) and the chip area (A). Using one of the conventional yield models based on the gamma distribution of defects results in a yield equation given by

$$Y=[1/(1+D_0A_c/\alpha)]^\alpha, \quad (1)$$

where α is an empirical parameter with a value close to unity. The resultant yield is plotted in Fig. 2 as a function of chip area for three defect densities: 1, 5, and 10 defects/cm^2. A defect density of 1/cm^2 can be expected for a very mature, well-developed technology, while D_0 of 5/cm^2 and higher should be expected when the feature size and process technology used is pushing the state of art. For each value of D_0, the corresponding yield curve has an effective "knee" beyond which the yield falls off rapidly with increasing chip area. In very large scale integrated (VLSI) circuits, the chip area is dictated by the feature size and the number of effective gates in the circuit. In WSI, however, the chip area is an adjustable parameter. One possible approach to choosing the optimum chip area is to set it equal to the value at the knee in the yield curve. If a greater value is chosen, then the yield will be greatly reduced. On the other hand, if a much smaller value is used, only slight improvements in yield can be expected while substantial increase in the interconnect network will be necessary. As can be seen in Fig. 2, for D_0 of 1/cm^2 an optimum chip area of 0.75 cm^2 is obtained in this fashion. The corresponding yield level is 55%. For the higher defect densities of 5 and 10/cm^2, optimum A_c of 0.2 and 0.13 cm^2 are obtained. The corresponding yield levels are 50% and 45%, respectively. These yield values, in the neighborhood of 0.5, indicate that a redundancy factor of ~2 is required for implementing WSI circuits.

In a sense, WSI can achieve a higher level of integration than VLSI through the use of redundancy and a careful partitioning strategy. This, of course, is a moving target. As shown schematically in Fig. 3, a given level of integration

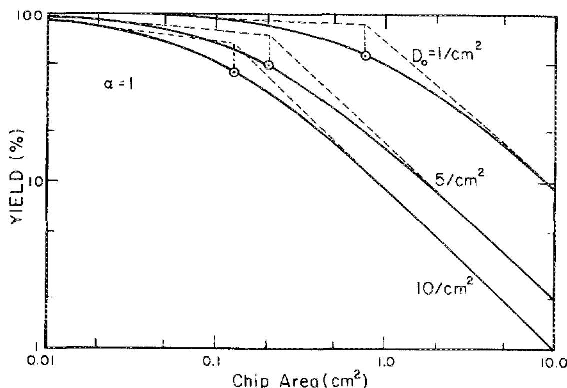


FIG. 2. Plot of chip yield vs chip area for various chip defect densities.

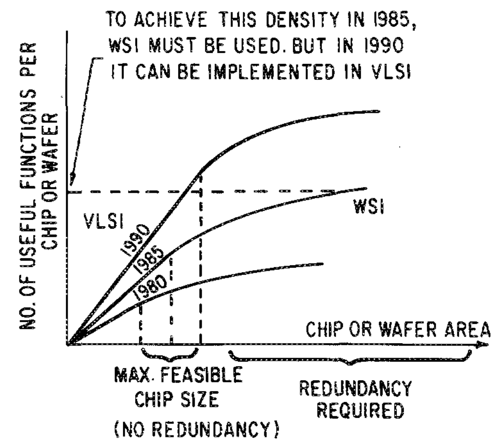


FIG. 3. Curves showing maximum chip size achievable with and without redundancy vs number of functions realized per chip for three different time periods.

achievable today only with WSI may be feasible with a conventional VLSI approach in 1990. At that time, however, the WSI approach, will allow for even greater levels of integration, possibly not achievable with VLSI until 1995. Therefore, one can interpret the value of WSI as a *technology accelerator* allowing us to implement larger scale systems before conventional approaches make that possible at roughly the same circuit density.

Two distinct implementations of WSI are emerging: monolithic and hybrid. In monolithic WSI the chip-to-chip interconnections are fabricated on the active wafer. In the hybrid or "reconstituted" wafer approaches, a separate "wafer"-like substrate is used for the interconnection network. Dies are then mounted upon this structure. The substrate can be a regular Si wafer, or made of ceramic or even metallic materials. Wafer-to-wafer or chip-to-wafer connections are necessary in the hybrid approach. However, both approaches have to deal with more or less the same multilevel interconnection technology issues. Furthermore, in some cases, the two strategies might be combined to provide a form of three-dimensional stacking of components.

WSI has a long and controversial history. Its origins are associated with the first efforts to produce VLSI.[2] These studies have continued for over two decades[3,4] with a recent resurgence of interest[5-8] especially with the development of repair strategies for programmable interconnections.[9,10] In the past, however, WSI technology seems to have been developed during brief, intense efforts (often poorly documented) designed to address deficiencies of other approaches. Today there is evidence of a more patient approach directed at solving long-term problems using multilayer interconnection strategies.

II. ELECTRICAL PERFORMANCE CONSIDERATIONS

Future electronic packaging needs will increasingly emphasize the speed and density of interconnections. Conventional boards and packages rely heavily on interconnections of dimensions much larger than those encountered inside integrated circuits. The dielectric constants of board materials such as ceramics are also substantially larger. A great deal of performance is therefore sacrificed in these external signal paths as they are currently implemented. External

paths, however, are not the only cause for loss of performance. Internal signal paths within the chip can also limit the performance of these circuits. Frequently, the surface density of active devices in a circuit which is limited to only two layers of interconnections can be as low as 20%. Since the fraction of the semiconductor substrate area in conventional board level packaging can be as low as 5%, the combined effect of both internal and external connections is substantial.

In addition, if one examines the paths typically found in multiple board assemblies mounted on edge connectors, then another cause for loss of performances is identified. Wires must be routed from positions possibly remote from the edge connectors on two boards with some length of connection on the backplane. Although the volume density for such stacked board systems is fairly high, the use of edge connectors imposes limitations on the signal paths which prevent the exploitation of this density.

Much of the current technology is motivated by the mechanical requirements for assembling and repairing multiple chip and board systems. Therefore, future performance improvements can be obtained only by adapting this circuit partitioning and repair strategy to finer pitch, dense wiring, or abandoning the repairability requirements altogether in the interest of speed. Perhaps, circuit redundancy and self-repair approaches will replace the need for mechanical repair, but these must be compatible with the desire for dense circuitry exhibiting high speed.

As a result of research and development efforts, unloaded gate switching times are quickly progressing toward the 1 ps range. Packaging technology must also advance or the performance promised by this device research will be completely lost in large systems due to the burden of interconnections.

Multilevel metallization for dense wiring offers an approach which addresses some of these needs. Both internal and external chip wiring can be viewed as part of the same multilayer interconnection structure. By scaling external wiring to dimensions comparable to those of internal wiring, a large improvement in chip density is possible. By introducing extra layers of such wiring some of the internal chip connections can even be shortened through the use of external paths leading to improvements in device density (with some increase in the number of bond pads). Finally, by three-dimensional stacking of such multilayer structures and implementing feedthrough wiring between the structures, the limitations of edge connector technology can be alleviated.

Since the fine pitch, dense wiring sought for interchip interconnections can be fabricated by conventional semiconductor technology, it is convenient to implement this wiring on a wafer substrate in order to be compatible with wafer handling equipment. Silicon and ceramic wafers are commonly employed in practice. We can then properly refer to the fine pitch wiring as wafer scale interconnections.

The longer external interconnections in this environment impose demands on both yield and electrical performance. Fortunately, the solution for the electrical requirements is compatible with the yield requirement. Both the metal and dielectric film thicknesses must be somewhat larger than typically found for internal chip wiring.

The electrical performance requirements are dictated by the fact that such long signal paths are normally operated as transmission lines in high-speed systems. This permits signal propagation at the maximum speed feasible for the given insulator's dielectric constant.

Ignoring dielectric losses and the skin effect (which expels electrons from the interior of a conductor at high frequencies), the high-frequency content of a positive traveling step wave discontinuity in a slightly lossy line is attenuated by the factor

$$A = \exp\{-R/2Z_0\}, \tag{2}$$

where R is the total resistance of the line, and Z_0 is the characteristic impedance of a lossless line of the same cross-sectional geometry.

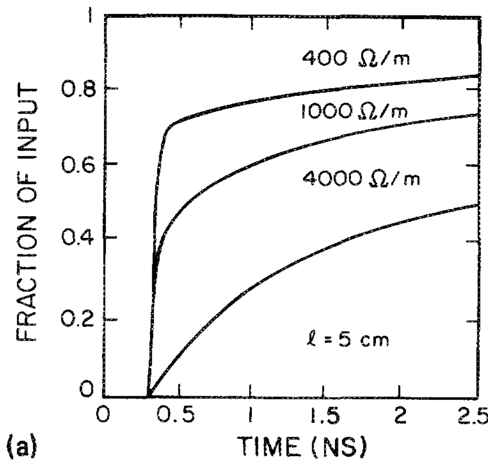
This high-frequency attenuation produces a degradation of step wave rise times which grow more severe with increasing line length. In order to minimize rise time degradation, the ratio $R/2Z_0$ must be much less than one. For a stripline geometry and for lines wide enough to ignore fringing fields, this requirement becomes

$$td \gg \rho \epsilon_0 \sqrt{\epsilon_r}\, l c_0, \tag{3}$$

where l is the length, t and d are the metal and dielectric thicknesses, respectively, ρ is the metal resistivity, ϵ_r is the relative dielectric constant of the insulator, and ϵ_0 and c_0 are the dielectric constant and speed of light for free space. The inequality in Eq. (3) can be satisfied for wafer length connections as long as 20 cm provided td is greater than $(4.6\ \mu\text{m})^2$ for a polyimide dielectric and aluminum lines of resistivity $2.67\ \mu\Omega/\text{cm}$. In practice, a metal thickness of $5\ \mu\text{m}$ and a dielectric thickness of $10\ \mu\text{m}$ can be used to obtain adequate transmission line performance out to a maximum of distance $l = 20$ cm.

Figure 4(a) shows the unit step response for 10 cm lines with various metal thicknesses. The lowest curve is for an SiO_2 dielectric thickness and metal width resulting in $L = 175$ nH/m and $C = 180$ pF/cm and metal thickness producing 4000 Ω/m, while the other curves are for metal thickness producing 1000 and 400 Ω/m, respectively. These curves do not include the effect of dielectric loss.

The curves illustrate that the line's step wave response consists of two parts. The first section is the electromagnetic wave arrival time delay, while the second part consists of an abrupt step followed by slow charging behavior due to the line resistance and capacitance. Thicker lines exhibit a larger step followed by less charging than exhibited by thinner lines. For lines of thicknesses comparable to normal integrated circuit dimensions ($0.5\ \mu\text{m}$), the abrupt step essentially disappears and only the slow charging is observed. Figure 4(b) shows the line response at various distances along a transmission line with metal resistance per unit length of 400 Ω/m on $10\ \mu\text{m}$ of a polyimide dielectric. The transition from the abrupt step present at shorter distances to the slow charging behavior at long distances is clearly seen. The frequency transfer function response for the line corresponding to Fig. 4(a) is shown in Fig. 5. The low-frequency response is unity decreasing to the level given by Eq. (2) for thick



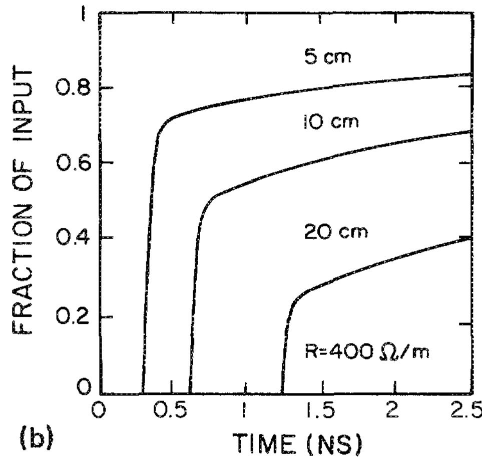


FIG. 4. (a) Line unit step response with self-inductance per unit length of 175 nH/m, self-capacitance per unit length of 180 pF/m, and three metal thicknesses resulting in line resistances per unit length of 4000, 1000, and 400 Ω/m. Line length is 5 cm. (b) Line unit step response with self-inductance per unit length of 175 nH/m, self-inductance per unit length of 180 pF/m, and metal thickness giving a resistance per unit length of 400 Ω/m for line lengths of 5, 10, and 20 cm.

films at a few hundred megahertz. The skin effect produces the upper break frequency near 10 GHz for the thicknesses described here. Of course, shorter length lines satisfy the inequality in Eq. (3) with thinner layers and with thinner lines the skin depth becomes significant only at higher frequencies. Hence, if additional layers are required, the shorter wires could be assigned to the thinner layers thereby providing wider bandwidths.

In addition to the question of line loss, the wafer interconnections must be designed to minimize coupling. Figure 6(a) shows the variation of mutual inductance and mutual capacitance of a line versus the dielectric thickness. Unfortunately, as shown in Fig. 6(b), we cannot make the dielectric thickness arbitrarily small because the self-inductance and capacitance also vary with this parameter, and hence, this affects the line's characteristic impedance. To keep Z_0 approximately at 50 Ω, a design of 5 μm of metal thickness with 10 μm metal width and 10 μm spacing combined with a dielectric thickness of 10 μm is used resulting in mutual inductance per unit length of 0.4 nH/cm and a mutual capacitance per unit length of 0.2 pF/cm.

Higher values of Z_0 result in lower current surges during

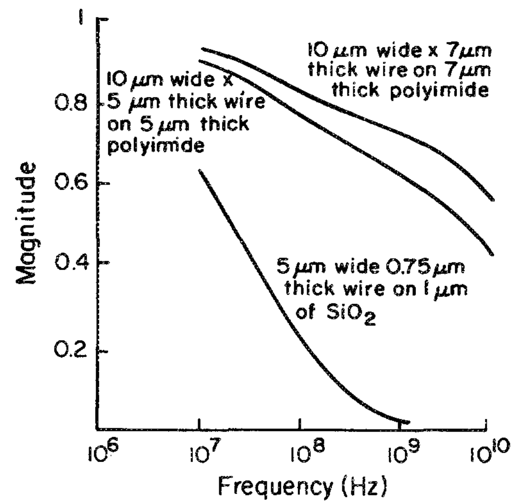
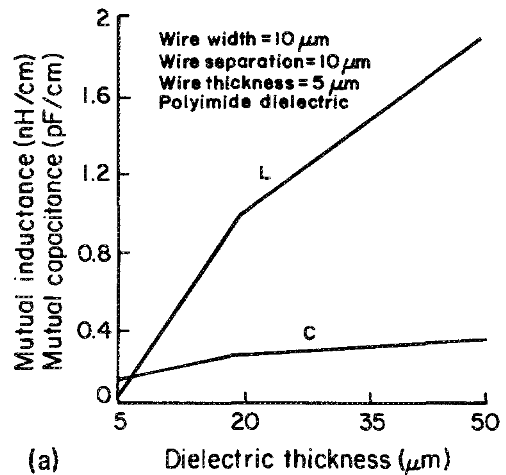


FIG. 5. Line frequency response for 10×7 μm metal wire on 7 μm of polyimide (top curve), 10×5 μm metal wire on 5 μm of polyimide (middle curve), and 5 μm×0.75 metal wire on 1 μm of SiO_2 (lower curve). Line length is 5 cm.

line charging and would permit larger losses in the line before the onset of signal distortion. However, it is difficult to make Z_0 much larger than about 100 Ω in these conductor geometries.



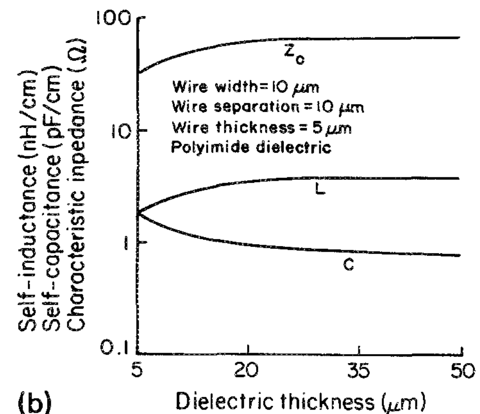


FIG. 6. (a) Mutual inductance (in nH/cm) and mutual capacitance (in pF/cm) for two rectangular wires 5 μm thick of width 10 μm and spacing 10 μm vs dielectric thickness. (b) Self-inductance (in nH/cm) and self-capacitance (in pF/cm) and characteristic impedance (in Ω) for a 5 μm thick, 10 μm wide rectangular wire vs dielectric thickness.

TABLE. I. High-density thin-film hybrid interconnects for high-frequency circuits.

(Dimensions in μm, resistances in Ω/cm)					
User Metal	Mosaic (Ref. 23) Aluminum	Raytheon (Ref. 24) Aluminum	RPI/GE (Ref. 12) Aluminum/ Copper	IBM (Refs. 25 and 26) Copper	AUGAT (Ref. 27) Copper
Thickness	1	1	5	5.5	25
Width	11	25	10	8	100
Pitch	22	100	25	25	250
Resistance	27	11	3–4	4	0.1
Dielectric	SiO_2	SiO_2/ sapphire	Polyimide	Polyimide	Polyimide glass
Dielectric K	3.8	5.5 eff.	3.5	3.5	3.8
Thickness	1	450	8	6	60
Deposition	Sputter/ CVD	Sputter/ CVD	Spun on	Spun on	Foil
Substrate	Silicon	Sapphire	Silicon	Alumina	Copper
Advantages	Line density Usual silicon processes	High Z_0 Usual silicon processes	Line density Silicon flat, good heat conductor	Line density Power and ground in alumina	Moderate complexity
Disadvantages	High line resistance High distortion and delay	High crosstalk Sapphire polish	Complex processing	Complex processing	Low line density

III. WSI MULTILEVEL INTERCONNECTIONS

A variety of approaches to multilayer wafer scale interconnections are being pursued. These include both "thin" films (micron-scale dimensions) and "thick" films (mil-scale dimensions) involving aluminum or copper as the metal and SiO_2, polyimide, or even Teflon as the dielectric. Some of these approaches are summarized in Table I, along with their respective advantages and disadvantages.

Obviously, the greater the film thickness compared to its width, the lower the series resistance of the lines linking chips mounted on a hybrid substrate. High series resistance means lossy lines, so that normal LC (L inductance, C capacitance) lines become RLC lines and even RC lines. As operating frequencies move toward 100 MHz and beyond, the signal rise times decrease toward 1 ns or less. For such fast rise times, even a few centimeters of path length of a lossy line can distort the signal and add significant delay time for signal transmission. Hence, the requirement of using copper or aluminum lines with polyimide dielectric, allowing the line thickness to increase from 1 to about 5μm.

Table I shows the density advantages of a polyimide dielectric with 5μm copper or aluminum. The most aggressive of these are at least temporarily held back by the greater fabrication complexity of line formation and yield problems for these schemes as opposed to the others in the table.

Thin-film polyimide dielectric layers in conjunction with thin-film copper conductors offer the greatest advantage because of the low dielectric constant and chemical stability of polyimides at the temperature required for the various soldering and joining steps in package construction, the high bulk electrical conductivity of copper (as opposed to aluminum, for instance), and most importantly, the possibility of using copper plate-up for metal deposition. This multilayer system will use thin films which remain mechanically stable and adherent even during temperature cycling, in spite of severe expansion coefficient mismatch with the substrate. This is in contrast to a thick polyimide foil, which would lead to severe substrate bending or delamination. The conductor linewidth can be as small as 10 μm, with 15 μm separation (or even smaller). The goal of 400 lines/cm wirability can thus be achieved in only two layers (X and Y). When completed with one power supply layer and one ground plane layer, it appears possible to achieve the entire interconnection network with only four conductor layers. It should be noted that this is possible here because of the high electrical conductivity of copper and the ability to fabricate thin polyimide insulating layers (which lowers crosstalk between conductor runs), in contrast to the more familiar ceramic multilayer approach, which is typically characterized by 15–20 mil linewidths and spacings and as many as 22 layers.

Multilayer test patterns of 25, 50, and 75 μm lines and spacings were fabricated using copper plate-up. Copper deposition was always preceded by a 1000 Å thick sputtered Cr film for better adhesion. A 1000 Å film of sputtered copper is then followed by a photoresist step, covering the Cu film except where it is desired to increase its thickness. Electroplating in copper solution now provides up to 10 μm thick copper conductors (additive plating). Polyimide (using DuPont No. PI 2555, or GE SPI-1000) is applied by spin-on techniques to fill the area between the conductors, and then eventually to cover the conductors themselves. Typical thicknesses and levels of planarization are indicated in Fig. 7.

Via holes in the polyimide layers are either formed by wet etching before the final cure step, or by reactive ion etching

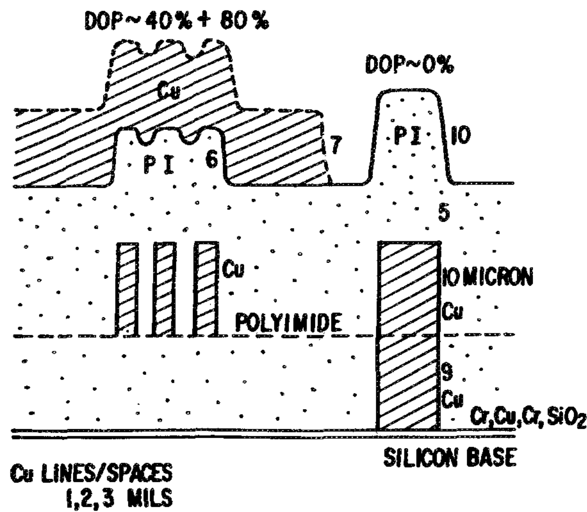


FIG. 7. Coverage of spun-on and cured polyimide over narrow and wide copper lines built up on a silicon substrate. All vertical dimensions are in microns. The three bars of Cu on the left-hand side are 2 mils wide on 4 mil centers. The isolated bar on the right-hand side is 3 mils wide. The top layer of polyimide (above the dotted line is 10 μm thick). The degree of planarity (DOP) is the difference between the metal step and the polyimide step divided by the metal step. The 3 mil wide bar is sufficiently wide that its DOP is 0% while the three narrower bars exhibit a 40% DOP at the ends of the cluster, and an 80% DOP in the middle of the cluster. Narrower lines have even higher DOP especially if spun on and cured in several steps.

(RIE), using an aluminum film as a stencil mask. Nearly straight side walls are obtained by the latter etching technique which, however, are then sloped more gently for improved step coverage after a short barrel etch.

Figure 8 shows Cu wires over 10 μm of polyimide. The sloped vias are 2×2 mils. Figure 9 shows some artwork patterned in this fashion (with 50 μm Cu lines on a wafer coated with polyimide) for the purpose of forming a wafer scale hybrid.

IV. WAFER TRANSMISSION MODULE USING LIFT-OFF

The wafer transmission module (WTM) is one of the possible wafer scale hybrid (WSH) embodiments of the concept

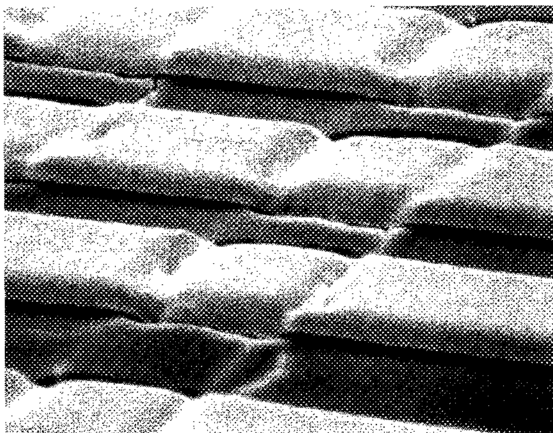


FIG. 8. Cu lines over 10 μm thick polyimide. Vias are 2×2 mils^2. 1000$\times$ magnification.

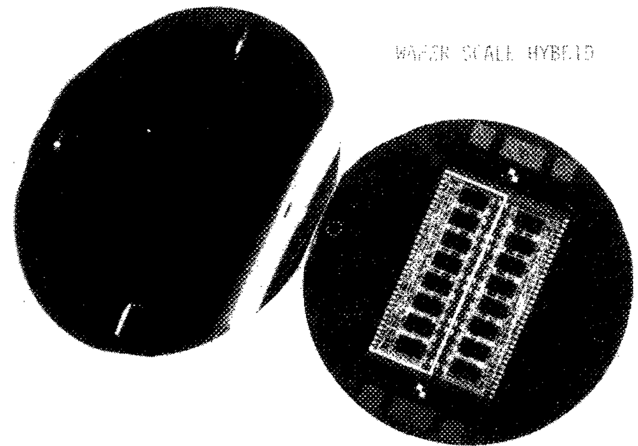


FIG. 9. Wafer scale hybrid approach. Interconnect network for ~1 megabit memory (16–64 K chips) on Si package.

of a multilayer metallization structure with wafer scale interconnections. Metal and dielectric thicknesses are sufficiently large to satisfy Eq. (2) and the structure incorporates ground and power planes as well as bypass capacitor layers. Figure 10 shows a cross section of the WTM. The integrated circuit dies are mounted onto the WTM structure and connected to its wiring either by wire bonding or flip bonding. A wire bonding layout for a 4 in. wafer is shown in Fig. 11. It contains 37 die sites of area (7 mm $\times$ 7 mm).

Figure 12 shows a routing diagram for two layers of horizontal and vertical wiring. At 20 μm pitch, the wiring capacity of the two layers is about 0.5 km in total length. Many digital systems of interest can be implemented on this scale with as little as 30 to 50 m of wire, but more complex architectures can demand 100 m or more.

The thick films required for wafer length transmission line behavior can also improve fabrication yield. However, the thick films also pose certain difficulties for fabrication. The thick topographic features demand an interlayer dielectric with an extraordinary planarizing capability. Furthermore, a multilayer structure of 5 or 10 μm thickness per layer can lead to a total thickness of 80 μm or more for the complete multilevel structure. Thermal expansion coefficients for the metal and dielectric must match extremely well, otherwise the thermal stress will produce cracking or delamination. In addition, lithography alignment for layer to layer on such thick films can be difficult. Hence, a fabrication process is required which will be very tolerant of small alignment errors. Finally, via formation through such thick layers is difficult even using sloped vias.

A modified thick-film, lift-off process has been developed to address the aforementioned problems.[12] The via fabrication problem is handled by incorporating an entire set of lift-off steps to fabricate stud or pillar vias. The directional deposition typically employed for lift-off can readily fill 10 μm deep vias without sealing them up provided the top opening is at least 5 μm $\times$ 5 μm. On the other hand, smaller (1 μm $\times$ 1 μm) openings that are more likely due to pinholes will seal before filling up completely, thus greatly reducing interlayer short circuits.

Use of polyimide (GE SPI-100 in this case) for the dielec-

FIG. 10. Cross-sectional view of the wafer transmission module (WTM) (not drawn to scale).

tric (spun and cured in several 1–2 μm layers) provides the desired degree of planarity when used with lift-off. Use of pillar vias and embedding metal in lift-off trenches keeps the surface essentially planar, thus assuring nearly perfect planarity after subsequent coatings of polyimide. The optical clarity of many types of polyimides enhances the ability to obtain layer-to-layer registration. Additionally, the near perfect match of thermal expansion coefficients of polyimide and aluminum suggests that most of the desired characteristics are met by this combination of materials although copper is strong enough to substitute for aluminum in this process. The only drawback of polyimide is its lack of hermeticity. In spite of this, silicon nitride has been found to form an adequate hermetic encapsulant when deposited by plasma enhanced chemical vapor deposition (CVD).[13]

The processing sequence for the thick-film lift-off is illustrated in Fig. 13. This process is a scaled-up version on thin-film methods that have been used successfully for many years at IBM.[14]

Each of the eight processing layers (including the four stud via layers) shown in Fig. 10 are fabricated using the thick-film, lift-off process shown in Fig. 13. First, a 10 μm planarizing dielectric layer is spun on and cured in several layers to avoid or reduce pinhole effects. After this, first a release layer (R) of molybdenum or soluble organic film is deposited on the wafer and then a thin shield layer of aluminum. Finally, the pattern definition resist layer is spun on. After lithographic resist patterning, the thin aluminum layer in the exposed regions is etched. An anisotropic dry etch opens a deep vertical-walled trench or well with a slight metal shield overhang. Next, the thick metal layer is directionally deposited. Because of the shield overhang and the directional deposition, the metal in the trench or well is not connected to the top layer metal. A final lateral etch of the release layer releases the top metal, permitting it to be "lifted off" under mild agitation.

A scanning electron microscopy (SEM) photograph of a typical 5 μm × 5 μm cross-section thick-film metal line (embedded in a coplanar layer of polyimide) produced by this lift-off process is shown in Fig. 14. This figure illustrates the advantages of lift-off in maintaining near perfect planarity for every layer of multilayer structure. The metal is deposited into the RIE trench until it nearly fills it. The formation of each dielectric layer is by several spin-coat-and-cure operations assuring extremely flat surfaces. The use of vacu-

FIG. 11. Layout of a 4 in. diameter WTM supporting 37 dies of 7 mm × 7 mm area.

FIG. 12. Sample X–Y routing for the WTM showing roughly 2000 connections of various lengths.

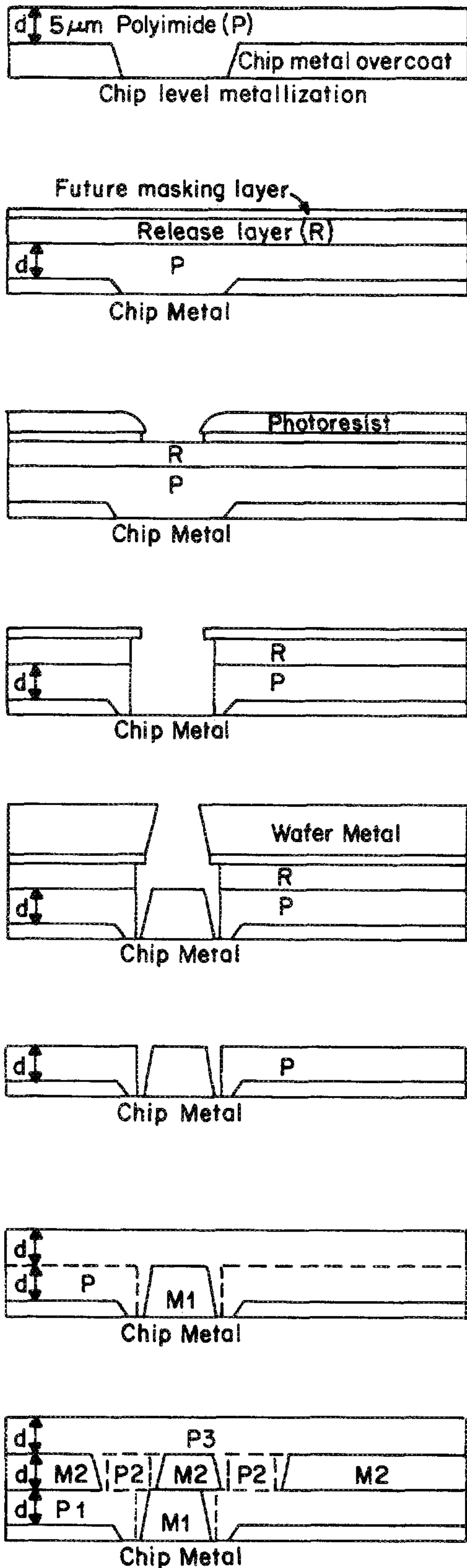


FIG. 13. Processing sequence for the thick-film lift-off process.

um systems for metal deposition avoids many of the contaminants found in wet etch and electroplating, leading to better yield and improved conductivity. This is offset to some extent by the additional complexity and cost associated with the lift-off process. Deposition of the thick metal films by rf

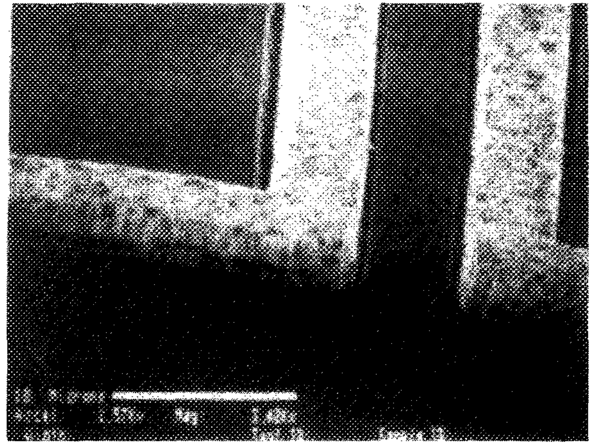


FIG. 14. SEM micrograph of 5 μm $\times$ 5 μm metal line embedded in polyimide and fabricated by thick-film lift-off.

heated evaporation and curing thick spin-on polyimide in several thin layers is a slow process. In spite of this, the yield required by more aggressive wafer interconnection design rules is perhaps best met through thick-film lift-off.

The reactive ion etch of SPI-100 must include both O_2 and CF_4 due to the presence of Si in this particular polyimide. This Si could oxidize in a pure O_2 plasma leading to stalagmites of SiO_2 coated polyimide or "grass" at the bottom of the trench. Additional Al coating of these stalagmites can occur due to ion erosion of the edge of the stencil shield or mask layer.[28] The silicon is present from a proprietary adhesion promoting agent in SPI-100. Figure 15 shows a sample of a 20 μm wide etched trench formed in DuPont 2550 polyimide (which does not contain Si, but shows evidence of Al coated "grass"). Figure 16 shows an experimental prototype

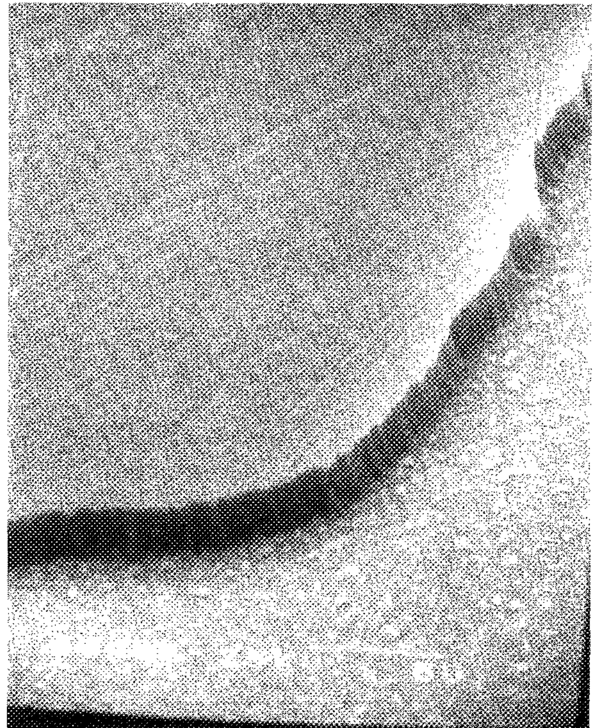


FIG. 15. Plasma etched DuPont 2550 polyimide in $CF_4/4O_2$ showing highly anisotropic profile obtained with an Al mask.

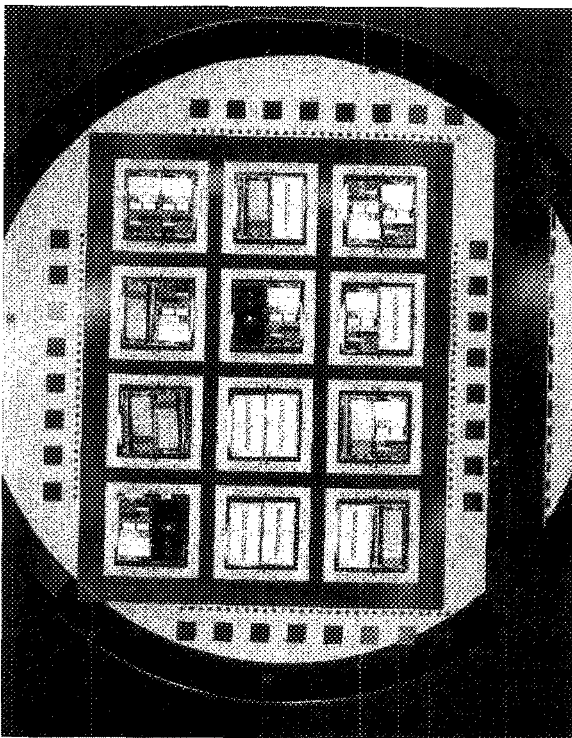


FIG. 16. An experimental sample of a wafer transmission module.

of the wafer transmission module fabricated by the process described in this section.

V. YIELD STUDIES

The yield of the WSI/WSH wiring process can be evaluated to a first approximation as follows. Assume that the wafer wiring is to be performed using lift-off under normal IC clean room conditions and methods appropriate for 2 μm lithography in semiconductor manufacturing are used. If these conditions remain constant for all wafers in a batch, we can assume a constant defect density for each batch; and if the defects occur at random, a Poisson distribution gives the

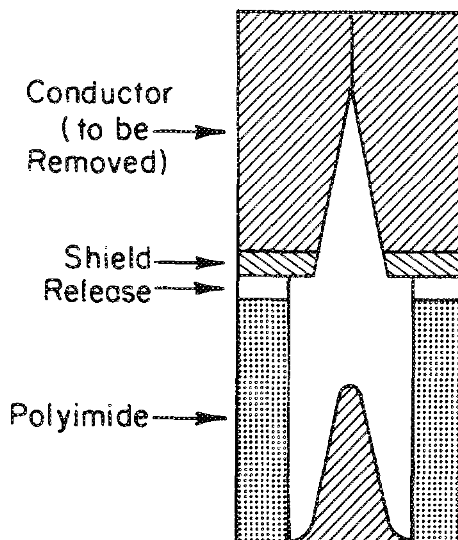


FIG. 17. Illustration of mechanism whereby small pinholes in the polyimide resulting from defects in the shield layer seal themselves up before forming an interlayer short.

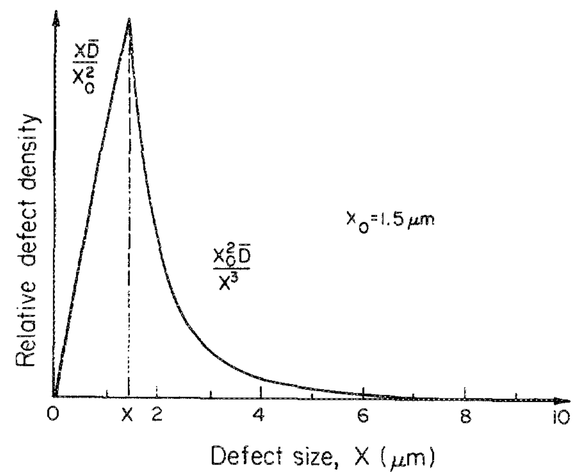


FIG. 18. Probability density of defect sizes for a given average defect density D and for a given lithographic resolution X_0.

probability that a number of defects will appear in a batch. The yield Y_b of the batch of wafers is then given by Stapper[15-19] as

$$Y_b = \exp(-\lambda_b), \tag{4}$$

where λ_b is the average number of faults per batch.

The faults in a batch may be due to opens or shorts in the lines or pinholes in the dielectric between the metal layers:

$$\lambda_b = \lambda_o + \lambda_s + \lambda_p. \tag{5}$$

The number of defects caused by pinholes can be neglected for this process. Each dielectric layer is spun and cured on the wafer in several thin layers making it virtually impossible for even two pinholes to align themselves and produce a fault. Most pinholes are smaller than 1 μm and could only occur in the resist or shield layers during fabrication. These will cause deep, narrow wells in the polyimide dielectric during the vertical dry etch. The directionally deposited Al will then grow toward the center of the pinhole as the layer becomes thicker, sealing it before enough Al can be deposited in it to create a short to the layer to be formed above as shown in Fig. 17.

The use of solid studs and a polyimide dielectric layer enhances the planarity of the surface used for the next metal layer. Metal thickness variations along the surface and possible breaks in the lines because of steps on the wafer surface are reduced at every level of the structure. Hence, λ_p is very small.

The evaluation of λ_s and λ_o depends on the average critical areas for shorts and opens A_s and A_o, and the defect densities for these failure mechanisms. Critical areas are regions where defects actually cause opens or shorts whereas defects in noncritical areas do not cause faults. Stapper shows that the average critical areas are functions of the defect size X and the probability distribution of defect sizes shown in Fig. 18. The curve shows a peak at defect size X_0 which is the minimum reproducible feature size. Below X_0, the lithographic equipment loses resolution and the curve falls linearly. Above X_0, the probability of a particular defect size diminishes with the cube of the defect size. The average critical area for opens is given by Stapper as

TABLE II. Yields (per complete layer) with $D_o = D_s = D$.

	Conventional VLSI				WSI
Wire width (μm)	3	3	5	5	10
Wire separation (μm)	3	22	5	20	15
Channels/layer	11785	2828	7071	2828	2828
Yield ($D = 0.1$ defects/cm^2)	19%	86%	25%	78%	97%
Yield ($D = 0.3$ defects/cm^2)	0.67%	65%	1.5%	48%	93%
Yield ($D = 0.5$ defects/cm^2)	0.02%	49%	0.09%	30%	88%

$$A_0 = \left[A_w \frac{X_0^2}{2w(2w+s)}\right], \qquad (6)$$

where w is the width of the lines and s is their separation. A_w is the usable area of the wafer for wiring. The average critical area for shorts A_s may be obtained by exchanging s and w in the above formula. The λ_o and λ_s can then be calculated as the product of the appropriate average critical area and the appropriate average defect density D_o and D_s.

Table II shows the expected yield for a complete interconnection layer for typical defect densities of 0.1 to 0.5 defects per square centimeter per layer. A 7 cm square wiring area on a 4 in. wafer is assumed. This would be equivalent to 70 m in length of 10 μm width wire. This table includes pinhole defects for the 3 and 5 μm wire widths which are the principal failure mode. The 3 and 5 μm wire width configurations do not include a ground plane. The use of a ground plane in these two configurations would greatly aggravate the already poor yields but would be necessary to obtain the best electrical performance.

This high yield for the WSI wiring is mainly due to the fact that IC processing conditions can be guaranteed for the wafer, including high-resolution lithographic equipment and good clean room conditions. Furthermore, by using larger dimensions (10 μm wire width and 15 μm wire separations), the wafer wires are immune to faults caused by small defects, and those defects which have a large size occur less often in a controlled environment.

In an exciting paper by J. Turnbull[20] at IBM's Burlington, Vermont facility an analysis has been presented for implementing an IBM 3081 sized architecture (800 000 gates) using WSI, Rent's rule, and the work of Heller, Hsi, and Mikhail to estimate wiring space demand.[21] In this analysis, they also used a defect density of 0.1 cm^2 and a two-wafer partition with a wire pitch of 8 μm and a predicted yield of 60%.

The defect densities assumed in the construction of Table II will have to be reduced further as the wiring demands increase. As the die gate capacity increases, the number of pin connections to wafer scale wiring may increase as suggested by Rent's rule.

VI. ADVANCED PACKAGING CONCEPTS

Very large system requirements may not be optimally addressed by simply using larger diameter wafers. This would result in excessively long interconnections in some cases. Instead, three-dimensional stacking of wafers or population of

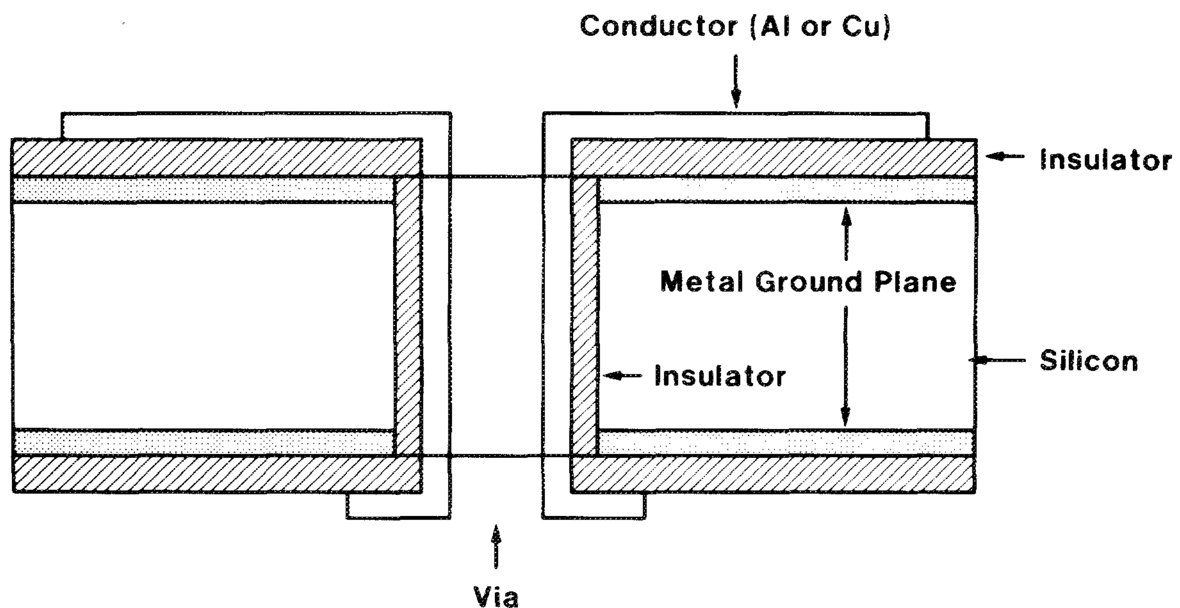


FIG. 19. Cross section of via between two microstrip lines. Vias made in 15–20 mil Si wafers using E51 25 laser resistor trimming system. Laser is 1.06 μm Q-switched Nd:YAG laser.

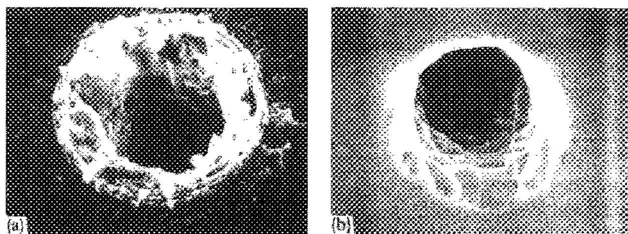


FIG. 20. (a) As laser drilled, via is quite rough with "spattered" silicon at the mouth of the via. (b) Chemical etching using 3:2:0.5 mixture of nitric:acetic:hydrofluoric acid removes the damaged region very effectively and leaves a rather smooth via surface (see H. H. Glascock, J. A. Loughran, and H. F. Webster, GE Report No. 86 CRD010).

both sides of a wafer may be preferable. To effectively utilize these concepts, via formation through the wafer substrate becomes a necessity (see Fig. 19). This permits a greater number of dies to be in close proximity than would be the case of a single surface, with short connections fed through these substrate vias. To preserve the transmission line characteristics of these connections, the substrate vias must have very low resistance. This means that the via should contain a high-conductivity metal as opposed to an aluminum-diffused silicon or recrystallized-amorphous silicon via.

To form these backside vias, a procedure is required to open the via through standard thickness wafers. For silicon substrates experiments were conducted with a *Q*-switched (Nd:YAG) laser with a peak power of 20 kW. The laser was used to "drill" holes in a 15 mil thick silicon wafer giving a hole with 1–3 mil taper diameter. The spattered silicon debris around the hole is partially oxidized because the drilling is done in air. It is thus readily removed by a chemical etch which also smoothes the rough bore and removes the microcracks around the periphery. This makes the wafer very strong and resistant to breakage. To form an electrical path through the hole, it must be metallized on the inside. Two different processes can be used to produce a metal surface layer. One is by sputtering metal down the hole, taking advantage of the conical shape of the hole. The other, which we prefer, is by a new method which we call laser sputtering.

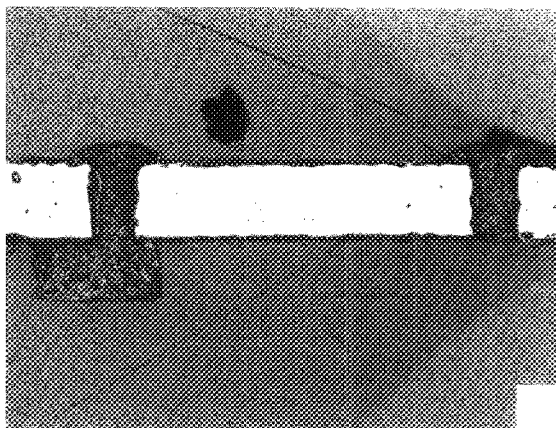


FIG. 21. Cross section of two via holes in a 15 mil thick silicon wafer formed by laser drilling, etching, copper sputtering, and then filled in a gold plating solution.

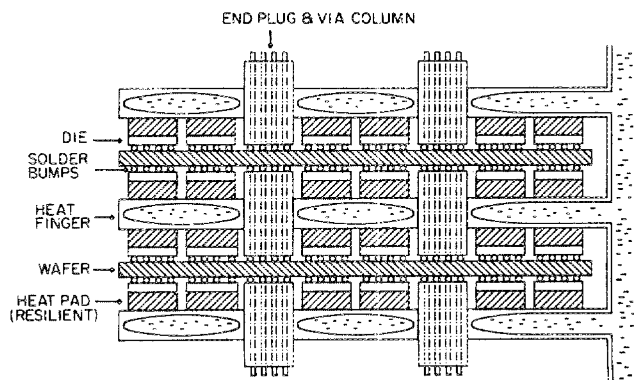


FIG. 22. Pictorial representation of a possible three-dimensional stacked structure employing heat pipes in between wafers with signal vias through the removal surfaces and contacts of low resistance and controlled impedance. Wafers are populated in double sided fashion by flip-chips.

Here a foil of gold or gold alloy is positioned to cover the holes at the wide diameter side. The laser beam is now beamed through the hole at the short diameter, hitting the foil at the other side and evaporating some metal, which coats the inside of the conical hole. In both cases, the thickness of the metal coatings must be increased by electroplating of Cu or Au, or electroless deposition of Cu. Since plating always leaves a residual hole, it is now filled with solder. Solder paste is screened on the wafer, using either 60 Sn/40 Pb or 92.5 Pb/6.5 Sn/2.5 Ag, and reflowed. Figure 20(a) shows the rough via hole through the Si wafer after laser drilling. Figure 20(b) shows the same hole after etch smoothing using a 3:2:0.5 mixture of nitric:acetic:hydrofluoric acid. Figure 21 shows a cross section of a gold plated via hole using Temperex solution for 2.8 h over sputtered Cu. Heat removal for stacked wafers would require interposing heat pipes or thermal conduction fingers in between the wafers. Signal vias must also be provided through these heat removal surfaces in order to permit short signal path lengths between the wafers. These must also employ contacts of low resistance and controlled impedance to preserve transmission line quality. A pictorial representation of such a structure is shown in Fig. 22.

VII. CONCLUSIONS

The multilayer wafer scale interconnection structure offers a solution for the high-frequency and density limitations imposed by current package and PCB level interconnect. Thick-film plate-up and lift-off fabrication techniques have been discussed and compared. Extremely encouraging projections have been made for wiring yield for moderate wiring demands and assuming defect densities typical for 2 μm lithography and manufacturing but applied to 25 μm pitch wiring using 5 μm thick metal on 10 μm thick polyimide dielectric fabricated by lift-off. Samples of this multilayer interconnection structure have been fabricated and characterized.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to the Semiconductor Research Corporation (SRC), and IBM for supporting the resident researcher status of one of the authors

(Dr. Bergendahl) at the R.P.I. Center for Integrated Electronics (CIE) during the period of this research. Thanks are also due to the many corporate sponsors of the CIE, especially the General Electric Corporation for sponsoring the core research grant on WSI at CIE. A special acknowledgment is due to the doctoral students Randolph H. Steinvorth, Gregory F. Taylor, and Captain Brian Donlan who worked to produce the first prototype WTM and WSI test structures at CIE and who were responsible for many of the figures used in this paper. At the GE Corporate Research Center the copper/polyimide studies relied heavily on the expert skills of J. G. McMullen to whom we express our appreciation.

[a] SRC Researcher in Residence at the Center for Integrated Electronics, RPI.

[1] W. P. Noble and W. W. Walker, IEEE Circuits Devices **1**, 45 (1985).
[2] R. L. Petritz, Trans. Metall. Soc. AIME **236**, 235 (1966).
[3] D. F. Calhoun and G. Wolfe, in Proceedings of the 1973 Electronics Component Conference, 1973, pp. 7–14.
[4] J. F. McDonald, E. H. Rogers, K. Rose, and A. J. Steckl, IEEE Spectrum **21**, 32 (1984).
[5] Y. Egawa, T. Wada, Y. Ohmori, N. Tsuda, and K. Masuda, IEEE Trans. Electron Devices **27**, 1612 (1980).
[6] I. Catt, Wireless World **87**, 57 (1981).
[7] D. L. Peltzer, VLSI Design Magazine **4**, 43 (1983).
[8] J. Grinberg, G. R. Nudd, and R. D. Etchells, IEEE Trans. Comput. **C-33**, 69 (1984).
[9] H. Stopper, Paper No. FAM 18.5, in Proceedings of 1985 International Solid State Circuits Conference, 1985, pp. 268–269.
[10] J. I. Raffel, A. H. Anderson, G. H. Chapman, S. L. Garverick, K. H. Konkle, B. Mathur, and A. M. Soares, in Proceedings of the International Symposium on Circuits and Systems, 1983, pp. 781–784.
[11] C. A. Neugenbauer, in Proceedings of the IEEE International Conference on Computer Design, 1984, pp. 115–120.
[12] G. F. Taylor, B. J. Donlan, J. F. McDonald, A. S. Bergendahl, and R. H. Steinvorth, in Proceedings of the IEEE Custom Integrated Circuit Conference, 1985, pp. 55–58.
[13] H. Eggers, H. Fritzsche, and A. Glasl, in Proceedings of the 1985 VLSI Multilevel Interconnect Conference, V-MIC 85, 1985, pp. 163–169.
[14] R. Larson, IBM J. Res. Dev. **24**, 268 (1980).
[15] C. H. Stapper, F. M. Armstrong, and K. Saji, Proc. IEEE **71**, 453 (1983).
[16] S. M. Hu, Trans. IEEE Electron Devices **69**, 4 (1984).
[17] C. H. Stapper, IBM J. Res. Dev. **29**, 87 (1985).
[18] T. E. Mangir, Proc. IEEE **72**, 690 (1984); **72**, 1687 (1984).
[19] M. B. Ketchen, IEEE Circuits Devices **1**, 24 (1985).
[20] J. R. Turnbull, in Proceedings of the Joint IEE/IEEE Southampton Workshop on WSI, 1985, Paper 5, pp. 1–14; *Wafer Scale Integration*, edited by J. Jesshope and W. Moore (Bristol and Boston, 1986), pp. 24–30.
[21] W. R. Heller, C. George Hsi, and W. F. Mikhail, IEEE Design Test Magazine (1984).
[22] J. Kilby, in Proceedings of the SRC Workshop on Wafer Scale Integration, 20–21 September 1984.
[23] H. Stopper, Paper No. FAM 18.5, in Proceedings of 1985 International Solid State Circuits Conference, 1985, pp. 268–269.
[24] E. T. Lewis, IEEE Trans. Components Hybrids Manuf. Technol. **2**, 441 (1979).
[25] C. W. Ho, D. A. Chance, C. H. Bajorek, and R. E. Acosta, IBM J. Res. Dev. **26**, 286 (1982).
[26] C. W. Ho, *VLSI Electronics: Microstructure Science* (Academic, New York, 1982), Vol. 5, Chap. 3, pp. 103–143.
[27] S. Lebow, in Proceedings of the 30th Electronic Components Conference, 1980, pp. 307–309.
[28] R. D. Verhaar and W. G. M. v. d. Hoek, in *Proceedings of the 1983 Microcircuit Engineering Conference* (Academic, New York, 1983), pp. 405–414.