Physical limits to the useful packaging density of electronic systems

by R. Fabian Pease
O.-K. Kwon

Increasing the density of electronic circuits and systems has been a major thrust for many years; the benefits are increased speed, reduced power-delay product, and reduced cost. Most of this effort has been directed toward the chip, but during the last decade system performance has been increasingly limited by packaging, and so emphasis has been shifting in that direction. Initially it was believed that heat dissipation was a serious fundamental limit, but advances in heat-sink technology have effectively eliminated that concern. One of the most serious problems is signal distribution. Although we can fabricate submicron metal lines, such lines are not normally practical as chip-to-chip interconnections because their resistance leads to undue signal delay and distortion; increasing their aspect ratio will increase cross talk. It is not clear what constitutes an optimal configuration, but for metals at room temperature a signal-line pitch of 30 to 40 μm appears practical. For low temperatures, and especially for superconducting lines, the pitch could be made very much finer, leading to greatly improved system density.

Introduction

The performance advantages of scaling down monolithic integrated circuits (ICs) were well established over ten years ago [1, 2]. These advantages were usually quoted in terms of speed and power-delay product, and implied a constant scale of integration; one study[1] was carried out to investigate the effects of scaling to constant chip size, and it was apparent that there were problems with interconnection and, at least in the case of bipolar circuitry, with power dissipation. In practice, what happened was that the chip size increased as minimum dimensions were scaled down; the net result has been that bipolar technology has become relatively unpopular for VLSI because of power dissipation. Consequently, many VLSI chips which usually employ MOS technology are limited in speed and density by their internal wiring [3].

The advantages of scaling down the dimensions of assemblies of chips (i.e., systems) have been less vigorously pursued in terms of both studies and implementation. There have been a number of reasons for this inertia:

[1] T. J. Riley and R. F. W. Pease, unpublished memorandum, AT&T Bell Laboratories, Murray Hill, NJ, 1972.

1. Most of the microelectronics industry has grown up with chip manufacturing being distinct from system assembly. This has led to the establishment of a strong investment in the technology of single-chip packages; traditionally, these are mounted on substrates [usually printed circuit (PC) boards]. Unfortunately, the spacing interconnections on traditional single-chip packages such as the dual-in-line package (DIP) has been 2500 μm (0.1 in.), although newer packages are offering line pitches as small as 600 μm. This large spacing, more than two orders of magnitude larger than the line pitch on the chip, obviously places a very severe constraint on the density of the final system.
2. Even when we break out of the straitjacket of single-chip packaging, our ability to fashion fine (less than 10-μm-wide) conducting lines on PC boards and on ceramic substrates has been regarded as doubtful and is indeed often impaired by the roughness of such substrates.
3. Conductors having large cross sections are needed to ensure low loss over the relatively long (10 cm or more) interconnections.
4. The accuracy required (say 2 μm) for aligning chips sufficiently accurately with a fine-line package was felt to be beyond the resources usually devoted to assembly (and testing). However, because this is considerably more relaxed than the alignment tolerance used in the lithographic processing of the chips themselves, this problem appears to be one that can be alleviated by applying state-of-the-art technology.
5. There existed a myth that closer packing of the chips would lead to intractable problems of heat dissipation.
6. Distributing stable dc power in very dense systems was perceived to be a problem.

In the following sections we discuss these points under two general subjects, thermal management and electrical management.

Thermal management

This was felt to be a serious fundamental problem until a few years ago [4, 5]; indeed, it was a relevant parameter in a figure of merit for packaging proposed by Keyes in 1978 [6],

$$F = (27/4)^{1/3}(m^2/c_1^2 Q)^{1/3},$$

where m is the average length of a chip-to-chip interconnection in units of chip repeat distance, c_1 the velocity of propagation along interconnections, and Q the allowable power density. At that time 20 W/cm^2 was felt to be the upper limit of power density of the chip. In 1981, by combining a careful analysis of the problem with micromachining technology, it proved possible to design and build, within the IC chip itself, a heat sink which employed laminarly flowing liquid coolant (e.g., water) and which

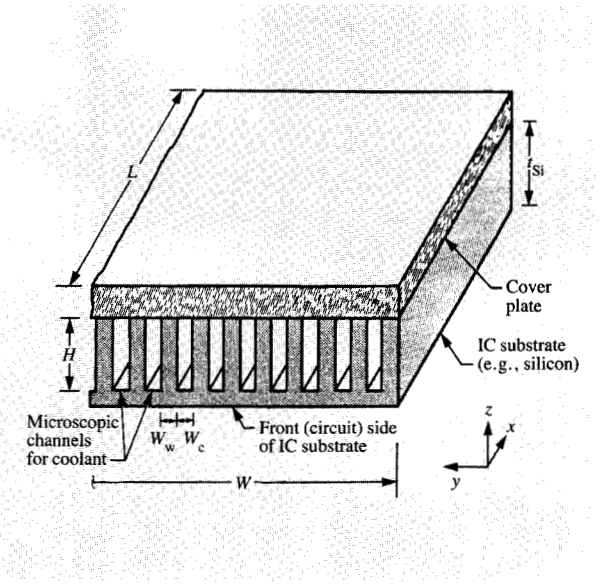


Figure 1

Schematic view of a microscopic heat sink incorporated into an integrated circuit chip. For a 1 × 1 cm^2 silicon chip, the optimal dimensions are approximately $W_w = W_c = 57$ μm and $Z = 350$ μm. The specific heat conductance is more than 10 W/K-cm^2 (from [7], reproduced with permission).

demonstrated a specific thermal resistance from junction to ambient of 0.07 K-cm^2/W [7] (**Figure 1**). In practice it was possible to dissipate over 1400 W in 1 × 1 cm^2 of silicon for a maximum temperature rise of 100 K [8]. What is significant about that design is the compactness of the total heat sink. Its volume is contained within the volume of the chip and so, unlike heat sinks used for individual power devices and for X-ray targets, no extra area is needed to spread out the heat, and the only extra volume required is for the supply and removal of liquid coolant.

One serious practical problem with the heat sink described above is that the use of liquid coolant is not regarded as appropriate for many applications. However, the use of liquid coolant in the IBM thermal conduction module (TCM) [8] has demonstrated its acceptability for large systems. In smaller systems an acceptable approach might well be to use a closed-circuit liquid system in conjunction with a forced-air heat exchanger at some distance from the high-speed circuitry.

A second practical problem is that the original configuration combines the IC and the heat sink in one piece of silicon. From the thermal point of view this has the significant advantage of eliminating the interface between chip and heat sink. Such interfaces have been a problem because of the difficulty of achieving a void-free, low-stress, thermally conducting bond. However, there are clearly

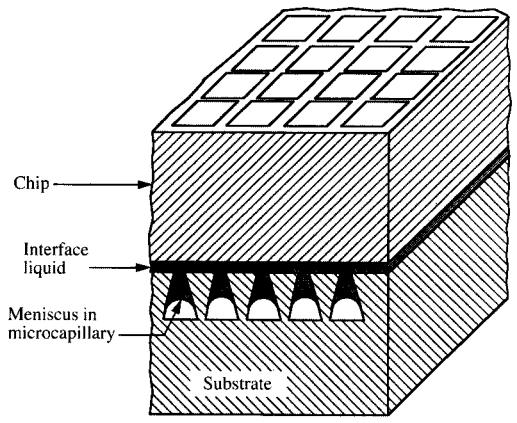


Figure 2

Schematic overview of re-entrant microcapillary attachment. Surface tension in the interface fluid creates an attachment suction pressure which ensures excellent thermal conductance across the interface with very low stress (from [9], reproduced with permission).

manufacturing advantages to decoupling the fabrication of the heat sink from that of the IC. One way of overcoming this problem is to achieve the bond through the use of surface tension in a thin interfacial liquid. As with the microchannel heat sink, a key ingredient of this approach is the use of new micromachining technology. A schematic view of the arrangement is shown in **Figure 2.** Machined into one surface is an array of fine (2-to-5-μm-wide) re-entrant grooves that serve as reservoirs for any interfacial liquid in excess of the quantity needed to just fill up the interfacial voids between the two solid surfaces. The attachment force is brought about by the hydrostatic suction pressure, P, throughout the liquid. This suction pressure arises from the pressure drop across the surface of the meniscus and is given by

$$P = T/R, \tag{1}$$

where T is the surface tension of the liquid and R is the radius of curvature of the meniscus.

Assuming a small contact angle, the radius of the meniscus is given by half the width of the groove at the level of the meniscus; thus the attachment pressure is controlled by the groove width. At first glance this might appear to suggest that we should make the grooves as narrow as possible, but this capability is limited in practice; if the grooves are made narrower than any interfacial voids, these voids will become empty because a meniscus stretched across their width will not be able to support the pressure drop P generated by the meniscus (with smaller radius) in the groove. This is standard capillary action **[Figure 3(a)]**. Obviously, empty or gas-filled interfacial voids are undesirable; their presence reduces the area over which the surface tension suction pressure acts, thus reducing the total attachment force, and the thermal conductivity of the gas in the void is almost certainly less than that of the interfacial liquid. Thus, in practice, the minimum width of our reservoir grooves should be limited to the maximum width of the interfacial voids. Hence, for a given surface tension, the lower limit to attachment force is set by the roughness of the mating surfaces.

A further requirement of the groove geometry is that the mouth should be narrower than the base. In this way the liquid is trapped at the surface, thus ensuring that the arrangement is stable. For example, if one groove starts to fill up (at the expense of its neighbor), its meniscus will increase in radius, tending to lower the local suction pressure and thus counterbalance the inequality.

The details of the fabrication of the grooves are described elsewhere [9]. Two methods have been successfully adopted. In the first, for the attachment of two polished silicon surfaces, the grooves were about 20 μm deep, 4 μm wide at the base, and 2 μm wide at the surface. Fabricating the grooves involved a combination of orientation-dependent etching of a (110) Si wafer and carefully controlled electroplating. Experimental results indicated that the attachment pressure was 0.37 atmospheres and the specific thermal conductance about 50 W/K-cm^2 (i.e., a thermal resistance of only 0.02 K/W for 1 cm^2).

The second technique was developed after it became clear that there was a need to attach surfaces that were considerably rougher than the polished silicon used originally **[Figure 3(b)]**. The surface is planarized by spin-coating it with a 15-μm-thick polymer film, which is then patterned into grooves about 8 μm wide. Copper is then sputter-deposited onto the polymer to coat the walls of the grooves with about 1 μm, thus providing a thermally and electrically conducting path. The desired re-entrant shape can result directly from the copper deposition process or from a subsequent nickel plating step, as in the original configuration. Experimental results obtained using this configuration include values for thermal resistance between 0.05 and 0.08 K/W-cm^2 (i.e., very attractive). Some electrical contact resistance measurements have also been carried out on the second configuration (with gold evaporated onto the mating surfaces). Values for specific contact resistance range from 5 to 15 mΩ-cm^2 (see footnote 2); these low values are very encouraging from the point of view of the possibility of supplying electrical power across the interface.

To summarize the state of thermal management: The fundamental limit is no longer an issue; no one seriously

contemplates dissipating 1 KW/cm^2, especially when the extra volume occupied by the heat sink is negligible. The most serious drawbacks are practical ones associated with the use of liquid coolants. The prospects for greatly improved gas cooling are not good because of the very low volumetric heat capacity of gases at atmospheric pressure and the difficulty of using gases at, say, 100 atmospheres [10]. However, the trend to avoid the use of liquid cooling by pushing the upper temperature limits of device operation may be unwise on grounds of reliability. Alternatively, the ability to dissipate 100 W/cm^2 and suffer only a 10 K temperature rise at the junction opens up many possibilities for high-speed systems. Some serious effort might be well invested in developing means to reliably handle liquid coolants for systems less massive than those employing the TCM. Other possibilities include the use of heat pipes [11], but it is not clear that for this application they offer a significant advantage over conventionally pumped liquid systems. Finally, it might be pointed out that the microchannel heat sinks are equally effective when liquid nitrogen is used as the coolant; thus the prospects are good for employing a closed-circuit liquid-nitrogen-cooled system in place of the more usual immersion.

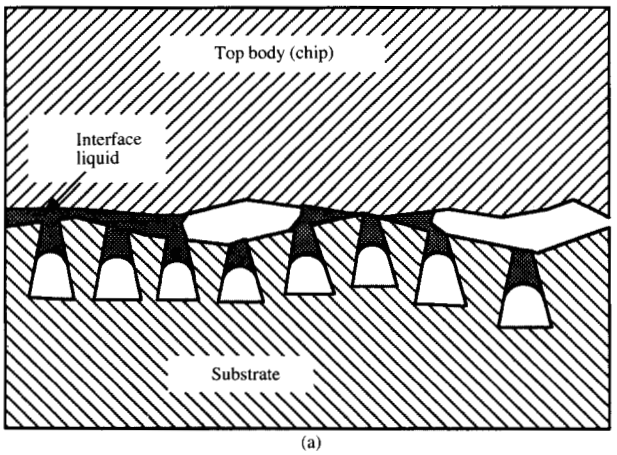
Signal distribution

In the Introduction a number of points were raised about the difficulty of scaling down system dimensions. The first point, that the coarse lead pitch on single-chip packages limits system density at the board level, is now becoming recognized and, notably at IBM, there is extensive work on multi-chip packages [12]; other projects are under way at other institutions [13, 14].

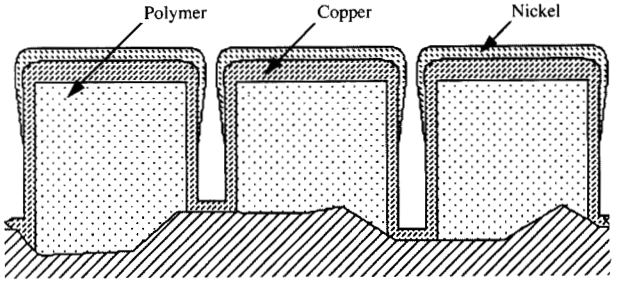
The second point, that the roughness of substrates limits our ability to fashion fine conductors, raises the issue of the choice of substrates. Presumably epoxy–glass PC boards are less expensive than ceramic boards, and the latter would be less expensive than silicon "boards." Furthermore, ceramic and PC boards permit the use of a relatively large number of levels of interconnection. For example, the multi-layer ceramic (MLC) substrate used in the TCM has at least 16 signal levels. However, it could be argued that a two-level interconnection system employing a 2-μm metal pitch, which is possible on silicon, should allow a greater density of interconnection than does the much coarser 16-level system of the MLC. There is, however, more to the argument than just the physical ability to fabricate a high-density interconnection substrate, and this brings us to the third point.

The length of chip-to-chip interconnections is typically much greater than that of the average intra-chip interconnection, and so we have to be particularly mindful of the limits imposed by the resistance of the former [3, 15].

[2] A. F. Paal, B. W. Langley, S. C. Douglas, D. B. Tuckerman, and R. F. W. Pease, submitted to *IEEE Trans. Components, Hybrids, & Manuf. Technol.*



(a)



(b)

Figure 3

(a) Effect of interface voids that are wider than the reservoir grooves. Capillary action causes the voids to empty, leaving less area exposed to the liquid with its surface tension suction pressure and relatively high thermal conductivity. (b) Microcapillary structure built up on a rough surface through the use of spin-casting of a polymer to achieve planarization. The copper surrounding the etched ribs provides both electrical and thermal conduction. The nickel ribs are capped by electroless nickel plating (from [9], reproduced with permission).

Some workers contend that the optimal approach is to use unterminated lines (assuming MOS receivers) and deliberately employ lossy fine lines to damp out ringing [16]. For the highest speeds, however, the use of terminated lines and bipolar technology, at least for the drivers, appears to be the favored approach; this is the case we shall study.

The general constraints are that if we try to increase interconnection density by shrinking only the lateral dimensions of the conductors, we will be limited by the resistance of the conductors. If we shrink the spacing between the conductors, cross talk will increase and set a limit; if we try to offset that problem by reducing dielectric thickness, the characteristic impedance of the lines will decrease and at some point become impractically low. Also, if we try to reduce resistance per unit path length by building up the thickness of the conductors, cross talk will again increase and eventually set a limit.

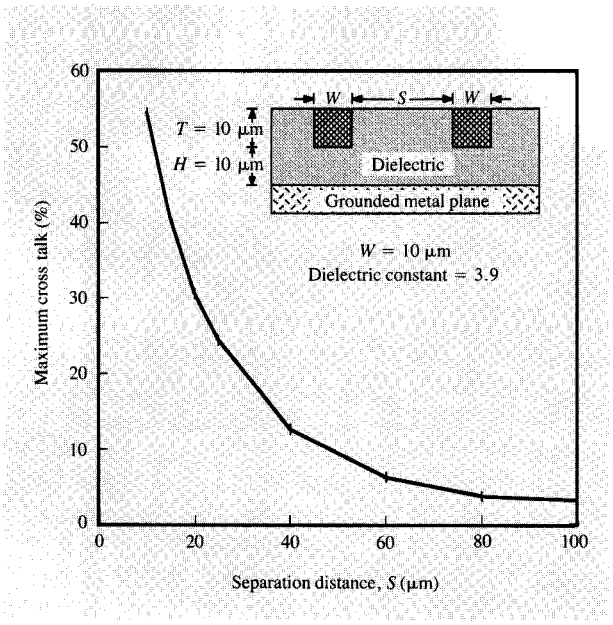


Figure 4

Relationship between cross talk and electrode separation for the microstrip structure shown in inset (from [26], reproduced with permission).

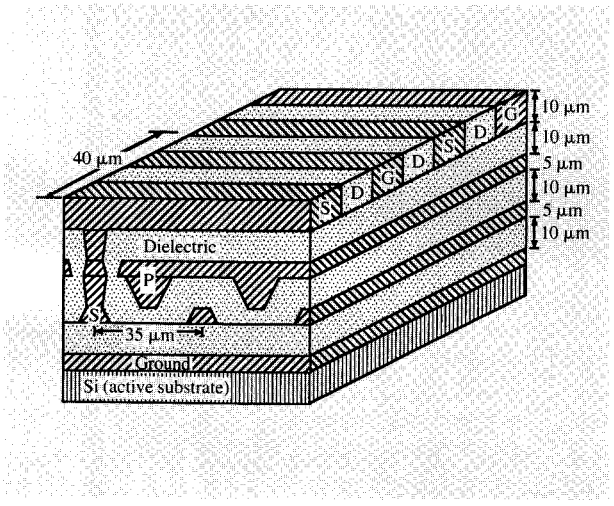


Figure 5

Hypothetical multi-layer interconnection structure used as vehicle for computing transmission-line parameters. The topmost layer (upper signal level) contains x-directional signal lines (S) with alternately shielded ground lines (G), at a 40-μm period. The lower signal level contains y-directional interconnection lines, at a 35-μm pitch. The conductor layer (P) between the two signal levels is a power plane.

We start by looking at some simple scaling arguments; we then review some calculations on a specific structure, and finally look at some relevant experimental results.

Suppose we scale down by a factor S a given multi-chip assembly whose packing density is set entirely by the inter-chip connection system made up of terminated lossy striplines driven predominantly in the TEM mode (the loss is assumed small enough for this to be valid). We scale down the dimensions by the same factor both vertically and laterally, but leave the dielectric constant unchanged. Then, from electro- and magneto-statics, the per-unit-length values of inductance L and of capacitance C will be approximately unchanged, so that both lossless characteristic impedance Z_0 and the phase velocity of the signal are unchanged on scaling (again assuming that the loss term is small). This invariance applies to both odd- and even-mode values, so that the cross talk term $(Z_{0e} - Z_{00})/(Z_{0e} + Z_{00})$ [17] is unchanged with scaling. Thus, the remaining question is whether the loss remains low enough to be consistent with the above assumptions. If the frequencies of interest are high, so that the skin depth is much less than the inner conductor thickness, the perimeter of the inner conductor is the critical term, and the resistance per unit length increases as S. Thus, the scaled stripline will have the same total resistance along its length; hence, in this case there is every incentive to scale down dimensions because the phase velocity is unchanged and the propagation delay should diminish as S. If the frequencies are very low, causing the skin depth to greatly exceed the inner conductor thickness, the resistance per unit length will increase as S^2, indicative of how far we can usefully scale down in the first case, since at some point the conductor thickness will become comparable with the skin depth—even at quite high frequencies.

Thus there are two extreme cases for scaling to constant package complexity:

1. Skin depth ≪ conductor thickness, in which case we can scale down cross-sectional geometry until our assumptions break down.
2. Skin depth ≫ conductor thickness, in which case total line resistance increases by S and may well set a lower limit to cross-sectional geometry.

The analogy with chip scaling can be taken further, as we can also scale to constant substrate size. Under this circumstance we cannot avoid the increased resistive loss that accompanies any reduction in the perimeter of the conductors; hence, as in the on-chip case, limits set by scaled-down interconnections will certainly become more serious. Unfortunately, we have found that for the multi-chip case, this is the more practical set of rules; by the same argument we should look at this case in some detail.

The vehicle chosen for our study was a line 10 cm long. For normal metals at room temperature, a cross-sectional

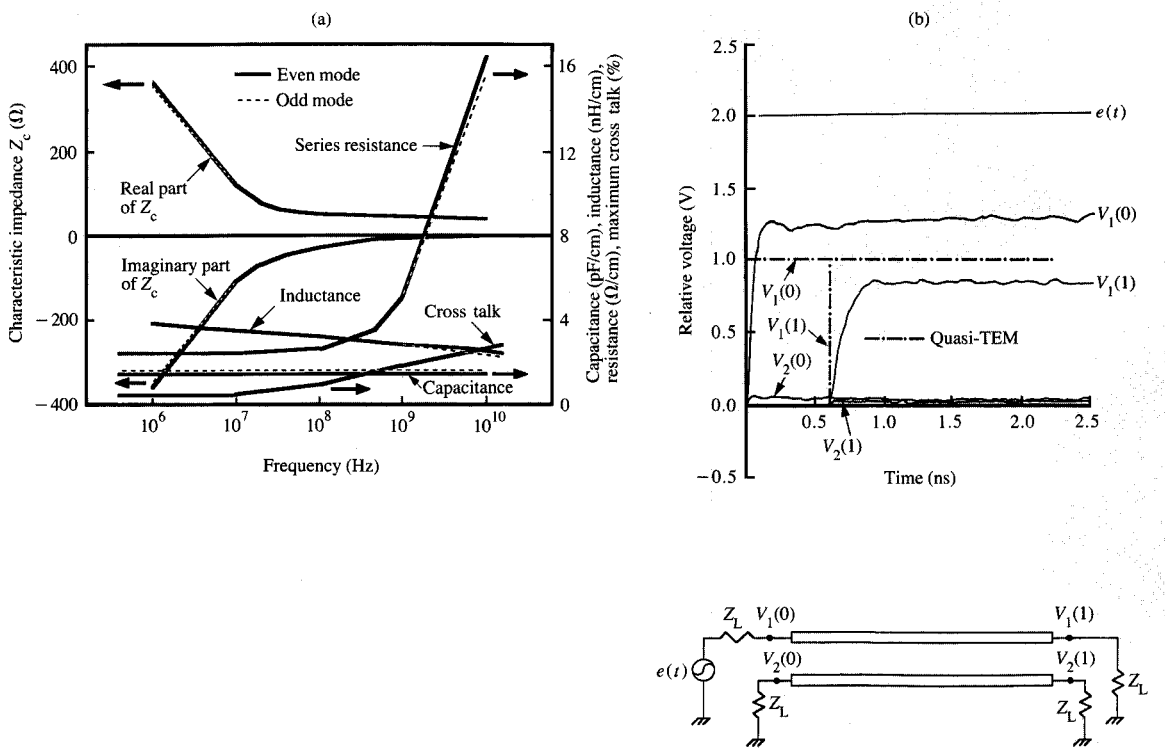


Figure 6

(a) Transmission-line parameters for the upper signal level of the structure of Figure 5. (b) Simulated step-function response for the 10-cm-long lines of the upper signal level of the structure of Figure 5. The lower diagram is a key to identify the notations used. The load impedance Z_L is assumed to be the characteristic impedance at 1 GHz. Note the very low cross-talk terms $V_2(0)$ and $V_2(1)$.

area of at least 50 μm^2 is needed to achieve sufficient conductance at low frequencies, and a perimeter of at least 50 μm to allow conductance at high frequencies, where the skin depth might be as low as 1 μm (copper at 10 GHz). The microstrip structure of **Figure 4** was the first example; the top surface contributed little to the high-frequency conductance because of the weak electric fields there. The results for the cross talk are also shown in Figure 4; we can see that to keep the cross talk within a reasonable level (say 5%), the spacing must be much greater than the conductor width. Indeed, we found that to achieve less than 5% cross talk it was more economical to insert alternating ground lines between signal lines. As a result we choose to study (by computer modeling) the multi-layer structure shown in **Figure 5.** Its second and fourth levels contain signal lines; its first and third levels contain, respectively, ground and power lines to allow adequately low resistance for distributing the high powers for the study.

The results of the study (for the uppermost signal level) are summarized in **Figure 6.** The characteristic impedance of the line corresponds to a distributed RC circuit at low frequencies and tends to 40 Ω at high frequencies, where the reactive terms dominate (as in a lossless line). The series loss increases and the inductance per unit length decreases due to the reduction of skin depth. Less clear is why the cross talk also increases at high frequencies; we believe that this can be explained physically by the increasing importance of the inductive coupling at high frequencies. Also shown are simulated waveforms for the step-function response for the top level (the results for the lower level are quite similar). Experimental results both at Stanford using a copper/polyimide structure and at Lawrence Livermore National Laboratory using SiO_2 as a dielectric agree well with these simulations.[3]

[3] A. Bernhardt, unpublished report, Lawrence Livermore National Laboratory, Livermore, CA, 1986.

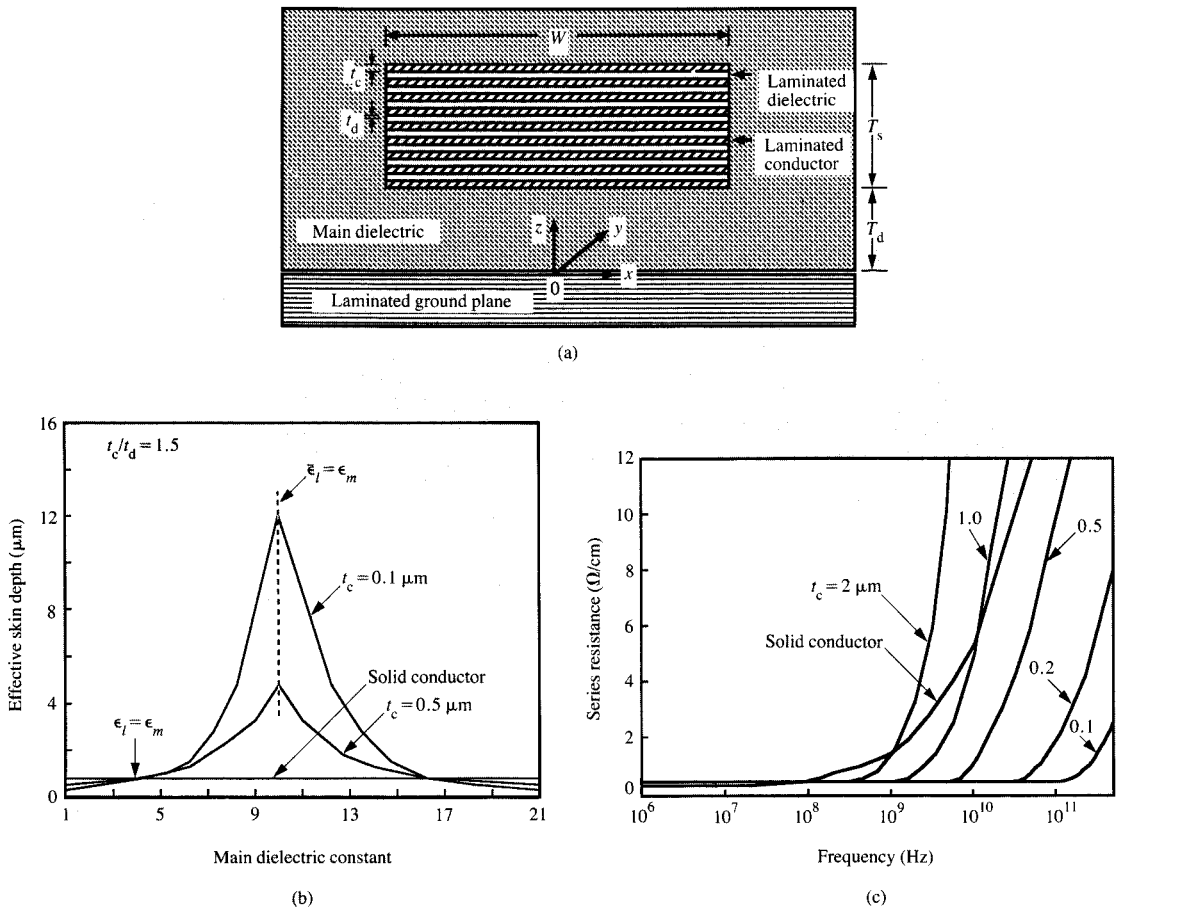


Figure 7

(a) Cross-sectional view of transmission-line structure using laminated conductors and dielectrics. (b) Effective skin depth versus laminated conductor thickness t_c for a microstrip line with laminated conductors as a function of main dielectric constant ϵ_m of the material between the signal lines and the ground plane at 10 GHz. The dielectric constant of the laminated insulator is 4; the ratio of conductor lamination thickness t_c to dielectric thickness t_d is 1.5; the average dielectric constant ϵ_l of the laminated stack is 10.0. Note that when the two dielectric constants are equal there is no advantage to this scheme; and that when the main dielectric constant is equal to the average dielectric constant of the laminated dielectric, the effective skin depth can be maximal for a given conductor lamination thickness. (c) Series resistance versus frequency for different conductor lamination thicknesses. It is assumed that the width of laminated conductors W is 50 μm, that the thickness of a laminated signal line T_s is 25 μm, and that the cross-sectional dimensions of the solid conducting line are the same as those of the laminated line.

Our conclusions from this study are that this structure may not be optimal, but it represents one configuration that should allow high-speed (at $c/2$) chip-to-chip communication over lengths typical of multi-chip modules. Thus, a signal-line pitch of 30 to 40 μm should be possible, representing an appreciable improvement over standard practice.

We can also see that the fundamental limit to further scaling down is the resistance of the lines. There are a few ways to alleviate this. One is to tolerate high resistance by using unterminated lines [14]. As mentioned earlier, this does not look attractive for the highest-speed applications (because of dispersion), and the dimensions so far quoted have been no finer than 50 μm.

A second approach is to reduce the loss at high frequencies by employing laminated conductors. At first sight, this does not appear attractive because the outer conductors (those nearest the opposing conductor) shield the inner ones **[Figure 7(a)]**. However, as originally pointed out by Clogston in 1951 for coaxial lines [18], it is possible to arrange the material properties of the dielectrics to achieve a

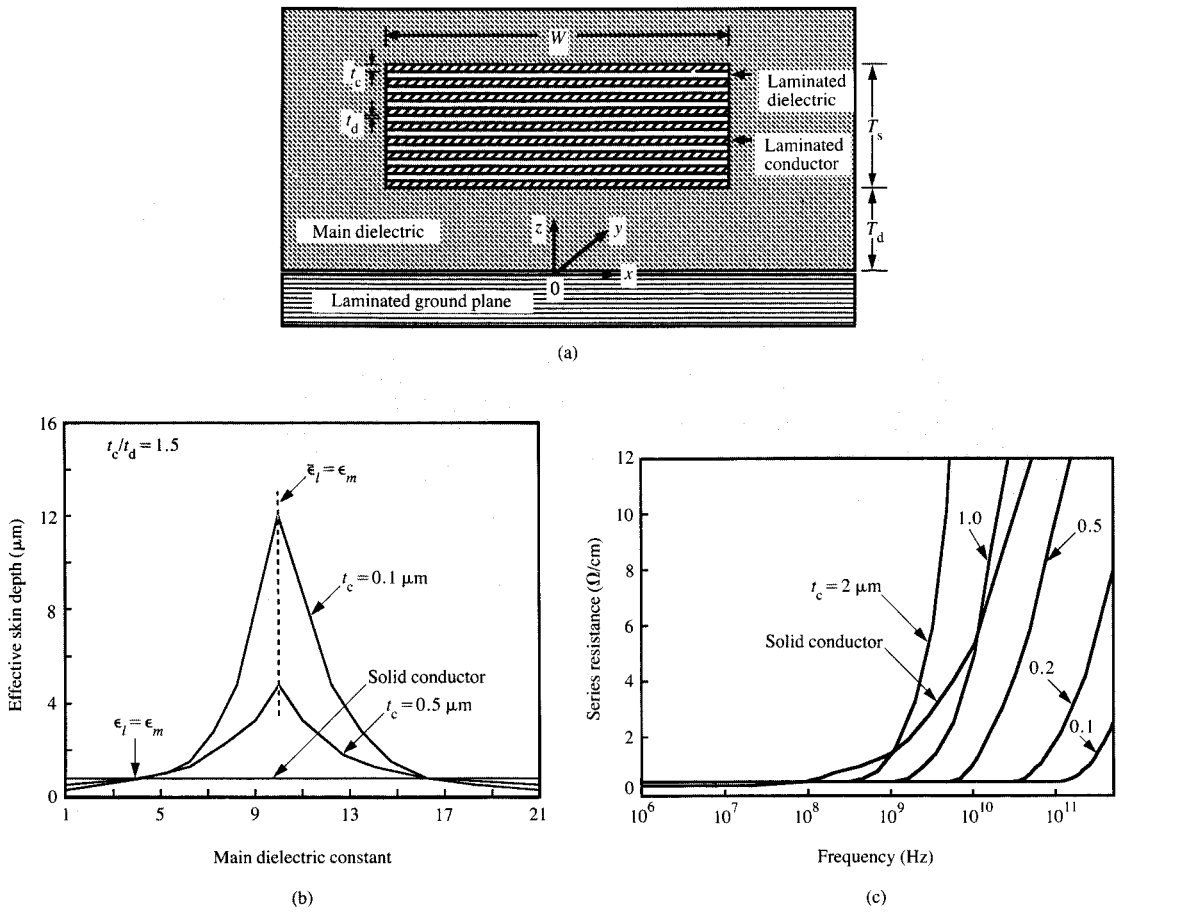


Figure 7

(a) Cross-sectional view of transmission-line structure using laminated conductors and dielectrics. (b) Effective skin depth versus laminated conductor thickness t_c for a microstrip line with laminated conductors as a function of main dielectric constant ϵ_m of the material between the signal lines and the ground plane at 10 GHz. The dielectric constant of the laminated insulator is 4; the ratio of conductor lamination thickness t_c to dielectric thickness t_d is 1.5; the average dielectric constant ϵ_l of the laminated stack is 10.0. Note that when the two dielectric constants are equal there is no advantage to this scheme; and that when the main dielectric constant is equal to the average dielectric constant of the laminated dielectric, the effective skin depth can be maximal for a given conductor lamination thickness. (c) Series resistance versus frequency for different conductor lamination thicknesses. It is assumed that the width of laminated conductors W is 50 μm, that the thickness of a laminated signal line T_s is 25 μm, and that the cross-sectional dimensions of the solid conducting line are the same as those of the laminated line.

Our conclusions from this study are that this structure may not be optimal, but it represents one configuration that should allow high-speed (at $c/2$) chip-to-chip communication over lengths typical of multi-chip modules. Thus, a signal-line pitch of 30 to 40 μm should be possible, representing an appreciable improvement over standard practice.

We can also see that the fundamental limit to further scaling down is the resistance of the lines. There are a few ways to alleviate this. One is to tolerate high resistance by using unterminated lines [14]. As mentioned earlier, this does not look attractive for the highest-speed applications (because of dispersion), and the dimensions so far quoted have been no finer than 50 μm.

A second approach is to reduce the loss at high frequencies by employing laminated conductors. At first sight, this does not appear attractive because the outer conductors (those nearest the opposing conductor) shield the inner ones **[Figure 7(a)]**. However, as originally pointed out by Clogston in 1951 for coaxial lines [18], it is possible to arrange the material properties of the dielectrics to achieve a

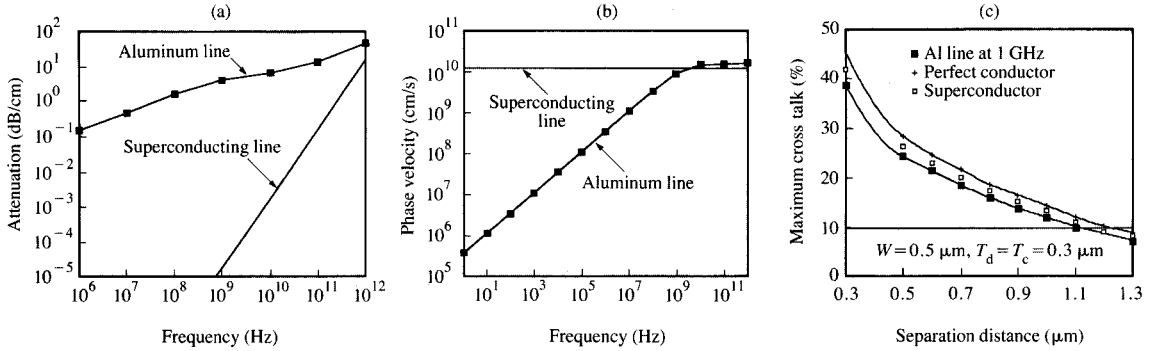


Figure 8

Calculated values of (a) attenuation and (b) phase velocity for superconducting and normally conducting microscopic transmission lines, with a signal-line width of 2 μm, a thickness of 0.5 μm, and an interdielectric thickness of 1 μm at 77 K. Note the expected improvement in performance of the superconducting lines. (c) Calculated maximum cross talk in percent for coupled, superconducting, and normally conducting microstrip lines with a signal-line width of 0.5 μm, a thickness of 0.3 μm, and an interdielectric thickness of 0.3 μm at 77 K. Note that the maximum cross talk for the case of superconducting lines with a pitch of more than 1.7 μm is less than 10%.

substantial increase in effective skin depth. In **Figures 7(b)** and **7(c)**, we show how the effective skin depth increases as a function of the dielectric constant between the conducting laminations, and how this translates into reduced loss over the relevant frequency range.

A third approach is to lower the operating temperatures to 77 K; for normal metals such as copper, the resistivity drops by about a factor of 6 to about 0.3 μΩ-cm. This allows us to reduce linewidth sixfold for the high-frequency case (skin depth ≪ conductor thickness) so that the signal-line pitch can be reduced to 6 μm. To retain the same resistance per unit path length at the other extreme, skin depth ≫ conductor thickness, the reduction factor is only $\sqrt{6}$.

The advent of materials that exhibit superconductivity at or above liquid nitrogen temperature (77 K) is obviously significant in this regard [19]. At first sight it might appear that we could eliminate all losses, so that the minimum conductor width would once more be set by our ability to fashion the conductors. However, assuming that these new materials behave as do the traditional superconductors, there are basic limits to the miniaturization of the striplines because of inductive effects, because of the penetration depth of the normal current before the superconducting region is reached, and because of the limiting current density of these materials. These effects were originally considered by Kautz, and more recently by Kwon et al. [20–22]. The results are summarized in **Figure 8,** from which we can see that the signal-line pitch can be reduced to about 2 to 3 μm. Because losses are so much less, this pitch is also valid for lines much longer than 10 cm. Of course, practical realization of these lines is far from straightforward, mainly because of materials considerations. For example, the satisfactory achievement of high values of limiting current density ($>10^5$ A/cm^2) at present appears to depend on crystal orientation. Furthermore, it is not clear that achieving suitable electrical contact to these materials will be straightforward. Nonetheless, the results of Figure 8 make it clear that the reward for developing a superconducting interconnection technology would be significant, especially for lines longer than 10 cm.

There are two matters relating to electrical management that have not yet been covered: power distribution and chip-to-substrate contacts. The problem of delivering raw power does not seem to be fundamental. The structure of Figure 5 employs dedicated planes for power and ground, and for a power density of 100 W/cm^2 this translates to 20 to 30 A/cm^2 (depending on the power supply voltage). For a 1 × 1 cm^2 area, this means that to maintain the maximum current density below 2 or 3×10^5 A/cm^2, the plane must be 1 μm thick, assuming that power is delivered from one edge; by the same reasoning, a 10 × 10 cm^2 area requires the power plane to be 10 μm thick. Alternatively, we could deliver power at more frequent intervals.

More serious is the noise that can appear on the power lines as a result of large switched currents interacting with the inductance of the leads that normally comprise the chip-to-substrate connections; this subject is well reviewed by Katopis [23]. One approach to ameliorating these effects might be to employ local (to the chip) active voltage regulation; this approach has not been seriously advanced in the past, probably because of perceived problems with the extra power dissipation. However, with high-performance

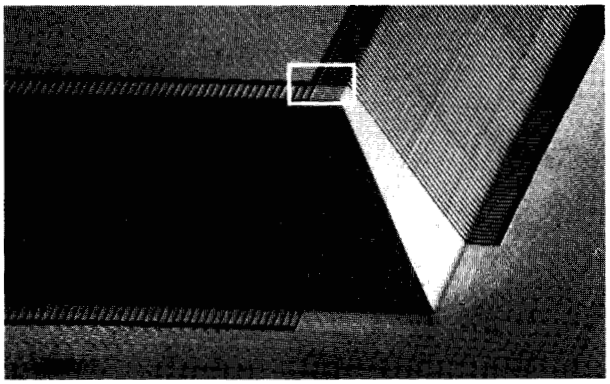


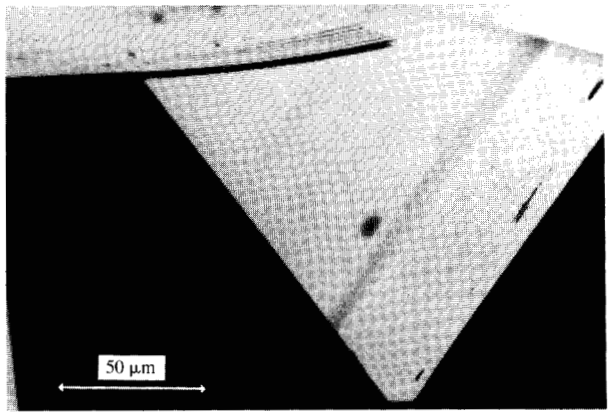
Figure 9

SEM view of laser-patterned conductors used for interconnecting high-speed chips on a silicon substrate. Note the good linewidth control on the beveled edge of the chip (from [25], reproduced with permission).

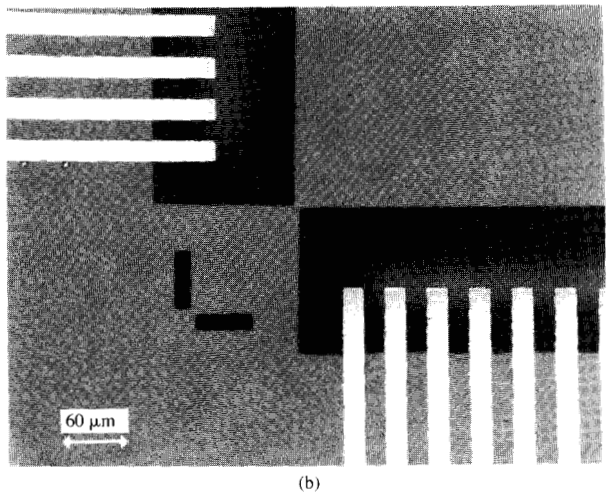
microscopic heat sinking, this is an approach that bears looking into. A second approach is to attack the inductance of the chip-to-substrate contact. The solder bump (e.g., IBM's "C4") approach is one quite effective technique that significantly reduces the inductance below the values normally associated with wire bonding (about 1 nH). Quite recently, significant progress has been reported in reducing the geometry of chip-to-substrate contacts; the result is a configuration that allows the transmission-line structure to reach onto the chip. One example is to use a laser pattern generator to fashion coplanar transmission lines from the substrate, up the beveled edge of a chip and onto its IC pattern (**Figure 9**). The signal-line pitch can be less than 35 μm, and the inductance is about 20 pH [24]. A further advantage of this scheme is that any misregistration between chip and substrate circuitry can be rectified by the laser pattern generator which is used; thus it is not necessary to rely on mechanical placement for achieving registration. An alternative approach might be to use an array of metallized SiO_2 microcantilevers as contacts. In an attempt to do so, such an array was constructed on 35- or 40-μm centers to match the substrate circuitry and arranged to overhang a moat centered on the chip perimeter [25]; in the absence of a mounted chip, the stress in the cantilevers caused a bowing upward by about 5 μm (**Figure 10**). When a chip was mounted on the region within the moat, there was sufficient flattening force to bring about satisfactory contact conductance for signals (>1*S*); permanent contact could be effected by cold welding or by laser fusing. The calculated inductance was about 10 pH. This technique does require accurate mechanical placement, but has the advantage of being reversible.

Conclusions

The packing density (and probably the speed [26]) of most of today's electronic systems is set by the density of chip-to-chip interconnections, by the density of chip-to-substrate contacts, and by the desire to retain air cooling. By employing microscopic heat-sink technology and associated liquid cooling, heat-removal problems should be considerably reduced. The problems of chip-to-chip interconnections could be ameliorated by the more widespread adoption of multi-chip packages. Within such packages, the packing density of the interconnections could be improved to a signal-line pitch (our figure of merit) of about 35 μm with normal conductors at room temperature, to a pitch of about 6–10 μm at liquid nitrogen temperature, and to a pitch of 2–3 μm if superconductors could be



(a)



(b)

Figure 10

(a) Metallized microcantilevers used for chip-to-substrate contacting. Note the upward bow. (b) Plan view of the same structure as in (a). The repeat distance is 40 μm (photography courtesy of J. L. Bravman and S. Hong).

employed. By applying new micromachining techniques, such as laser-induced patterning and orientation-dependent chemical etching, we can significantly reduce the inductance and increase the density of chip-to-substrate contacts.

Acknowledgments

The authors would like to acknowledge many stimulating conversations with Professors Bert Auld, John Bravman, and Malcolm Beasley, and with Drs. David Tuckerman, Anthony Bernhardt, and Bruce McWilliams of Lawrence Livermore National Laboratory. The preparation of this paper and much of the work reported herein was supported by the Semiconductor Research Corporation.

References

1. R. H. Dennard, F. H. Gaensslen, E. J. Walker, and P. W. Cook, "1μm MOSFET VLSI Technology: Part II—Device Designs and Characteristics for High-Performance Logic Applications," *IEEE Trans. Electron Dev.* **ED-26,** 325–333 (1979).
2. B. Honeisen and C. A. Mead, "Fundamental Limitations in Microelectronics—I. MOS Technology and II. Bipolar Technology," *Solid State Electron.* **15,** 819–829, 981–987 (1972).
3. A. K. Sinha, J. A. Cooper, Jr., and H. J. Levinstein, "Speed Limitations Due to Interconnect Time Constants in VLSI Integrated Circuits," *IEEE Electron Dev. Lett.* **EDL-3,** 90–92 (1982).
4. W. Anacker, "Josephson Computer Technology: An IBM Research Project," *IBM J. Res. Develop.* **24,** 107–112 (1980).
5. R. W. Keyes, "Limitations of Small Devices and Large Systems," *VLSI Electronics, Microstructures Science,* Vol. 1, N. Einspruch, Ed., Academic Press, Inc., New York, 1981, Ch. 5, pp. 185–230.
6. R. W. Keyes, "A Figure of Merit for IC Packaging," *IEEE J. Solid-State Circuits* **SC-13,** 265–266 (1978).
7. D. B. Tuckerman and R. F. W. Pease, "High Performance Heat Sinking for VLSI," *IEEE Electron Dev. Lett.* **EDL-2,** 126–129 (1981).
8. S. Oktay and H. C. Kammerer, "A Conduction-Cooled Module for High-Performance LSI Devices," *IBM J. Res. Develop.* **26,** 55–66 (1982).
9. D. B. Tuckerman and R. F. W. Pease, "Microcapillary Thermal Interface Technology for VLSI Packaging," presented at the IEEE U.S./Japan VLSI Symposium, Maui, September 1983; see also A. F. Paal and R. F. W. Pease, oral presentation at the IEEE International Electronic Manufacturing Symposium, San Francisco, September 1986.
10. D. B. Tuckerman, *Heat-Transfer Microstructures for Integrated Circuits,* Ph.D. Dissertation, Stanford University, Stanford, CA, 1984.
11. A. Basilius, "Heat Pipes in Electronic Packaging," *Electron. Packag. & Production,* p. 126 (November 1978).
12. E. E. Davidson, "Electrical Design of a High Speed Computer Package," *IBM J. Res. Develop.* **26,** 349–361 (1982).
13. W. Johnson and R. C. Jaeger, "Hybrid Silicon Wafer-Scale Packaging Technology," *Conference Digest,* IEEE International Solid State Circuits Conference, Anaheim, 1986, pp. 166–167.
14. H. J. Levinstein, C. J. Bartlett, and W. J. Bertram, Jr., "Multi-Chip Packaging Technology for VLSI-Based Systems," *Conference Digest,* IEEE International Solid State Circuits Conference, New York, 1987, pp. 224–225.
15. H. B. Bakoglu and J. D. Meindl, "Optimal Interconnection Circuits for VLSI," *IEEE Trans. Electron Dev.* **32,** 903–909 (1985).
16. E. N. Fuls, C. J. Bartlett, and W. J. Bertram, "A High Density Interconnect Technology for VLSI-Based Systems," *Extended Abstracts,* GaAS and VLSI Packaging Workshop, Research Triangle Park, NC, September 1987, pp. 16–17.
17. I. Catt, "Crosstalk (NOTSE) in Digital Systems," *IEEE Trans. Electron. Computers* **EC-16,** 743–763 (1967).
18. A. M. Clogston, "Reduction of Skin Effect Losses by Use of Laminated Conductors," *Proc. IRE* **39,** 767–782 (1951).
19. M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, and C. W. Chu, "Superconductivity at 93 K in a New Mixed-Phase Y-Ba-Cu-O Compound System at Ambient Pressure," *Phys. Rev. Lett.* **58,** 908–910 (1987).
20. R. L. Kautz, "Picosecond Pulses on Superconducting Striplines," *J. Appl. Phys.* **49,** 308–314 (1978).
21. R. L. Kautz, "Miniaturization of Normal-State and Superconducting Striplines," *J. Res. Nat. Bur. Stand.* **84,** 247–259 (1979).
22. O. K. Kwon, B. W. Langley, R. F. W. Pease, and M. R. Beasley, "Superconductors as Very High-Speed System-Level Interconnects," *IEEE Electron Dev. Lett.* **EDL-8,** 582–585 (1987).
23. G. Katopis, "Delta-I Noise Specification for a High-Performance Computing Machine," *Proc. IEEE* **79,** 1405–1415 (1985).
24. D. B. Tuckerman, "Laser-Patterned Interconnect for Thin-Film Hybrid Wafer Scale Circuits," *IEEE Electron Dev. Lett.* **EDL-8,** 540–543 (1987).
25. R. F. W. Pease, J. C. Bravman, O. K. Kwon, S. Hong, S. C. Douglas, B. W. Langley, and A. F. Paal, "High Performance Packaging for VLSI Systems," *Proceedings,* 1987 Advanced Research in VLSI Conference, Paul Losleben, Ed., MIT Press, Cambridge, MA, 1987, pp. 279–292.
26. B. A. Wooley, M. A. Horowitz, R. F. W. Pease, and T. S. Yang, "Active Substrate System Integration," *Proceedings,* 1987 IEEE International Conference on Computer Design, New York, pp. 468–471.

Received November 16, 1987; accepted for publication February 10, 1988

R. Fabian Pease *Stanford Center for Integrated Systems and Department of Electrical Engineering, Stanford University, Stanford, California 94305.* Dr. Pease received his B.A. in natural sciences in 1960 and his M.A. and Ph.D. in electrical engineering from Cambridge University in 1964; his doctoral research included the design, construction, and use of a high-resolution scanning electron microscope. After spending one year as a Research Fellow at Trinity College, Cambridge, he joined the faculty at the University of California, Berkeley, and continued research on scanning electron microscopy. In 1967 Dr. Pease joined Bell Telephone Laboratories, Murray Hill, New Jersey, where he worked first on the digital encoding of television signals and then on electron-beam and X-ray lithography. During that time he supervised a group responsible for developing technology for using the Bell Laboratories electron-beam exposure system EBES and sensitive electron-beam resists for mask-making and for direct write on the wafer. Since 1978, Dr. Pease has been a Professor of Electrical Engineering at Stanford University; he is conducting a research program into the generation and applications of microstructures. The projects include focusing of low-voltage electron and ion beams, nonconventional resists, direct-beam "write" technology, heat transfer in microstructures, and new approaches to interconnecting VLSI chips.

Oh-Kyong Kwon *Corporate Research and Development, Texas Instruments, Inc., P.O. Box 225474, Dallas, Texas 75265.* Dr. Kwon received his B.S. degree in electronic engineering in 1978 from Hanyang University, Seoul, Korea, and his M.S. and Ph.D. degrees in electrical engineering from Stanford University in 1985 and 1988,

respectively. From 1978 to 1980, Dr. Kwon served in the Army as a Research Assistant on the R.O.K.-U.S. Army joint research team in Seoul, Korea. In 1980, he joined the Research and Development Center of Gold Star Electric Co., Korea, where he designed facsimile systems. He is currently with Corporate Research and Development, Texas Instruments, Inc., Dallas. Dr. Kwon's research is focused on studies of high-density interconnections, involving the use of normal and superconducting conductors and optical interconnections; and on new approaches to system-level integration. Dr. Kwon is a member of the Institute of Electrical and Electronics Engineers and the Electrochemical Society.