Dot-to-Dot Maskless Bulk Interconnect Process

Bill Hamburgen
HP Laboratories Palo Alto
HPL-2004-149
November 14, 2004*

3-D packaging, electroless plating, memory packaging, interconnect, roll-to-roll

Dot-to-dot (D2D) is a bulk, maskless process for creating electrical interconnections between adjacent features. It does not require the features to be aligned to a mask or other tool and is thus applicable even to complex curved and folded surfaces. The principle can be illustrated by a set of dimensionless points scattered on a plane. Each point is a circuit node electrically connected to a surrounding circular region of radius b. Where the separation between points is greater than 2b, no connection is made. Where the distance is less than 2b, adjacent circles join and establish a connection. Long wires linking multiple nodes are formed wherever a series of points are separated by less than 2b. While physically realizable D2D connections are built-up from finite-sized regions rather than points, the principles are the same. Since all connections on a large number of workpieces may be formed simultaneously, manufacturing costs should be extremely low. This paper is a comprehensive examination of D2D technology. It begins with an overview of stacked packaging and suggests a one-dimensional approach to attaining precise inter-layer alignment. Models for stacked module geometry and plating process capability are used to predict minimum attainable wiring pitch. Initial experiments using electroless nickel plating and a planar substrate support the fundamental viability of the D2D approach. Test patterns demonstrate wiring on pitches as low as 50 µm, with increasingly robust results at larger pitches. Future work should focus on better understanding joint formation and controlling spurious plating nucleation on actual substrate stacks.

* Internal Accession Date Only
Approved for External Publication

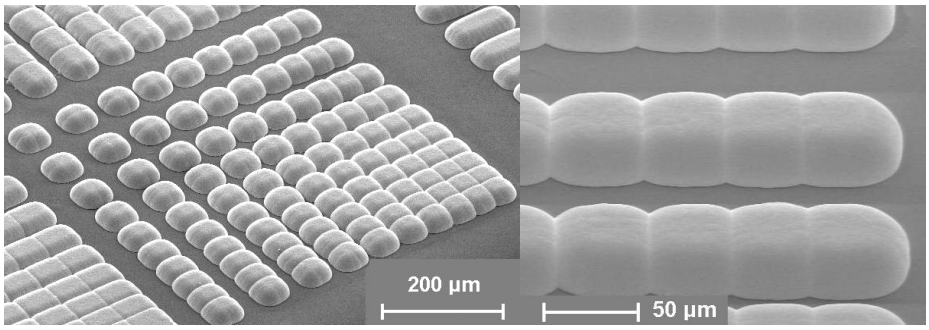
Dot-to-Dot Maskless Bulk Interconnect Process

Bill.Hamburgen@hp.com

HP Laboratories

Abstract

Dot-to-dot (D2D) is a bulk, maskless process for creating electrical interconnections between adjacent features. It does not require the features to be aligned to a mask or other tool and is thus applicable even to complex curved and folded surfaces. The principle can be illustrated by a set of dimensionless points scattered on a plane. Each point is a circuit node electrically connected to a surrounding circular region of radius b. Where the separation between points is greater than 2b, no connection is made. Where the distance is less than 2b, adjacent circles join and establish a connection. Long wires linking multiple nodes are formed wherever a series of points are separated by less than 2b. While physically realizable D2D connections are built-up from finite-sized regions rather than points, the principles are the same. Since all connections on a large number of workpieces may be formed simultaneously, manufacturing costs should be extremely low. This paper is a comprehensive examimation of D2D technology. It begins with an overview of stacked packaging and suggests a one-dimensional approach to attaining precise inter-layer alignment. Models for stacked module geometry and plating process capability are used to predict minimum attainable wiring pitch. Initial experiments using electroless nickel plating and a planar substrate support the fundamental viability of the D2D approach. Test patterns demonstrate wiring on pitches as low as 50 µm, with increasingly robust results at larger pitches. Future work should focus on better understanding joint formation and controlling spurious plating nucleation on actual substrate stacks.



Keywords: 3-D packaging, stacked multi-chip package, electroless plating, bumps, system in package.

Introduction:

This paper is a comprehensive and detailed examimation of Dot-to-dot (D2D) technology and its many potential variations. It targeted toward individuals who are familiar with the approach and are interested in some of the nuances of the technology. For a more terse and basic overview, the reader is directed toward a related conference paper [29].

The history of electronic packaging is a story of competition between linear and array approaches to interconnecting hierarchies of circuits. Today's most complex wired assemblies, such as multi-layer printed circuit boards (PCBs) and integrated circuits, are typically planar structures fabricated using mask-based lithography to define alternating layers of wiring and vias. The most aggressive system packaging technologies densely tile large numbers of chips onto multi-layer PCBs and attain planar circuit densities approaching what is possible on a silicon wafer. But high planar density is not enough, and there are many benefits to high volumetric density 3-D packaging [1], [27].

While over the long term mass stacking and interconnection of thinned wafers is likely to offer the highest volumetric packaging density [2], today sequentially stacking and wire bonding individual chips is the most common approach to creating a "system in package" [5]. An extreme example is a 1.4 mm thick package with nine 70 µm thick chips, six of them stacked and wirebonded [3]. But chip version management and yield issues impose limits on the applicability of wirebonded stacks, and alternate approaches are gaining increasing interest. Others have proposed built-up structures that incorporate chips and wiring on successive layers [4], [7]. Most such approaches require a high degree of sequential processing, and potentially suffer from yield issues similar to those of wirebonded stacks. To address these shortcomings, stacked ball grid array packages [12], and folded flex packages with several layers have been demonstrated [13], but these too appear unlikely to scale economically to large layer counts.

A fundamentally different approach involves creating individual substrates, each containing embedded thinned die and passive devices, and assembling them into a laminated stack and interconnecting the layers. This approach is quite similar to conventional multi-layer circuit board production, facilitating testing of individual layers prior to assembly and offering the advantages of a parallel rather than serial fabrication process. While conventional through-hole vias could be used to connect between layers, such vias use up valuable area on each layer of the assembly [1], [27]. The edge of the module, however, is essentially free real estate for inter-layer connections. Linear connections have historically offered a finer pitch than area array construction. When so, there is a threshold substrate size below which the peripheral approach will have a higher layer-to-layer connection count[1]. Also, as the substrate size shrinks, there is a reduction in the planar routing resources required to reach an edge, removing that disadvantage of peripheral connections. For these reasons, layer-to-layer connections in stacked laminated packages are typically patterned on the sawn edges of the module, where exposed traces can be accessed. Since the early 1990s, an increasing number of aerospace computer systems, array sensors, and memories have been built in this fashion [8], [28], and stacks of identical memory devices are beginning to find widespread commercial acceptance [6], [27]. It is not economical to apply conventional mask-based photolithography to the edge of a single module, and difficult to accurately stack multiple modules for mask-based processing. Instead lasers are typically

[1] For a square substrate with sides of length F, this length threshold is: $F = 4 P_a^2 / P_p$, where P_a is the area array pitch and P_p is the peripheral pitch.

used, either to pattern photoresist applied to the edge, or to directly ablate openings in a sputtered metal film. While modest volumes can be built in this fashion at moderate cost, it is unclear how acceptable the cost of sequential laser processing will prove in high volume applications.

The goal of the PIRM project at HP Labs in Palo Alto is to create write-once polymer memory at extremely low cost [14]. This material would be used as "electronic film" to provide archival storage for small mobile devices such as cameras and cell phones. Diode-fuse data storage elements [15], pull-up and addressing diodes and two layers of wiring – including crossovers – are patterned on a polyimide flex substrate with roll-to-roll (R2R) equipment using a revolutionary self-aligned imprint lithography (SAIL) process [21]. SAIL uses a single, multi-level level stamp to simultaneously pattern multiple device layers at sub-micron scales, eliminating alignment issues, even on large, unstable organic substrates.

While the SAIL process allows fabrication of individual layers at extremely low cost, this advantage can easily be squandered by the expense of laminating and interconnecting 40-80 or more layers to create a memory card. Addressing more than a gigabyte of data while meeting yield and reliability requirements will require hundreds of connections between the layers, and more such connections would enable more robust storage architectures. Conventional, sequential laser processing to create interconnects on the module edges was evaluated early in the PIRM program and found feasible, but a higher layer interconnect density at a much lower cost was clearly desirable to facilitate meeting PIRM's overall cost goals.

To minimize production cost, an approach was sought that would extend the economies of our R2R processing to the lamination and interconnection processes. Once modules are singulated, handling becomes expensive, so our aim was to process in roll format if possible, sheet format if not, and use strip or individual processing only as a last resort. To prepare the stack for interconnect formation, the layers are sequentially aligned, exercising high accuracy in one axis and only loosely controlling the second. The layers are sequentially stacked, laminated, and sawn at the module edges parallel to the alignment axis, exposing I/O traces from each layer. To facilitate economical handling, the saw does not penetrate the bottom layer of the stack, leaving the roll or sheet intact. Finally, the sawn edge is plasma cleaned to fully expose the trace ends, leaving clearly delineated pads that await connection to adjacent layers.

At this point the Dot-to-Dot (D2D) bulk interconnection process is applied. While in its most general form, D2D is applicable to surfaces and shapes of arbitrary complexity, it is most easily illustrated by considering a set of dimensionless points on a planar surface. Each point is activated, and grows or builds-up an electrically conductive circle of radius b. Wherever a distance of less than 2b separates points, their respective circles will join and become electrically connected. Where separation is greater than 2b, no connection is made. Long wires are created wherever a series of points are separated by <2b.

For the PIRM module edge interconnect, the activated regions consist of finite-sized pads formed by the trace ends. Columns of D2D connections must be formed between pads on adjacent layers, with no shorts between adjacent columns. We chose to use commercial adaptations of an EL-Ni plating process originally developed at Fraunhofer IZM and Technical University Berlin as a means of fabricating low-cost bumps and under-bump metallurgy [18]. Others have proposed using EL-Ni in place of electroplating to connect chip pads to redistribution wiring in a novel flip-chip package [10], and research in molecular-scale computing has led to a wide variety of approaches to maskless self-assembly of devices and interconnects [23]. We believe that D2D represents the first proposal for a generalized approach to using maskless anisotropic

conductor formation in a bulk process to create connections between adjacent electrically isolated features.

A D2D process for layer-to-layer interconnect must meet requirements for finished conductor strength, continuity and isolation, while accommodating substrate thickness variation, pad size and spacing tolerances, layer-to-layer misalignment, and variations in conductor build-up thickness. This paper will provide a series of models to capture the effects of these key variables. We will apply these geometric models to predict the highest wiring density that we are likely to attain, and perform a parametric study based on commercially available substrate technology and electroless nickel (EL-Ni) wafer plating processes [18], [19]. We'll show how D2D interconnect might be deposited on existing equipment for less than $.50 per memory card, and perhaps approach $.05 per card with a specialized roll-to-roll (R2R) plating line and suitable process development. This compares very favorably with laser-based edge patterning approaches.

While our D2D development has focused on the requirements of the PIRM memory card, we're interested in a broader range of system in package applications, in particular the possibility of building heterogeneous, multi-vendor systems. Work by others in stacked, multilayer wafer bonding and the more immediate needs of smart cards and wirebonded stacked-die packaging are pushing wafer thinning to well below 50 µm, a range where it becomes possible to consider embedding chips within openings in thin plastic films and interconnecting them. One can envision a manufacturing paradigm where 3-D stacks of active layers replace conventional planar PCB assemblies. Component and subsystem vendors would supply tested layers having standardized external power and signal trace locations. System builders could integrate many such layers – from multiple vendors when necessary - and use much of the infrastructure of the conventional circuit board industry to create dense, inexpensive 3-D packaging for extremely small systems. This is the larger vision that fuels our D2D development effort.

Layer-to-layer alignment, lamination and sawing:

Before interlayer wiring can be patterned, the layers must be aligned, laminated and sawn to expose the module edges. We will assume that incoming active substrate layers are supplied coated in a manner that planarizes both surfaces and protects the active devices. We also require that one or both surfaces be delivered with a B-staged (partially cured) adhesive applied.

Maintaining adequate alignment between layers is fundamental to attaining high inter-layer wiring density. Multilayer circuit boards are typically made by accurately registering to multiple fiducials on a given layer and punching corresponding alignment holes and/or slots. The layers are then assembled onto locating pins and the entire stack laminated with heat and pressure. There are several problems with this approach. The first is the registration error introduced by the aligning and punching operation, which is on the order of 125 µm per fiducial or 250 µm between layers [22]. A second is the cost of the registration-and-punching operation, which could prove substantial when substrates are small, layer count high, and layer alignment critical. Third is the creation of debris and planarity defects that could adversely affect yield or reliability. Ideally, we'd like an approach that makes alignment hole punching unnecessary.

The third challenge relates to the poor stability of unreinforced polyimide and other organic films. Manufacturers' data shows that these materials have a low-modulus, a high thermal coefficient of expansion (TCE) and a varying amount of shrinkage with elevated temperature exposure. They are also hygroscopic and have a high coefficient of humidity expansion (Table 1). By comparison, a silicon wafer has a 100x higher modulus, 1/10th the TCE, is not hygroscopic and has no shrinkage with elevated temperature exposure.

Property range for all Kapton/Upilex/Apical polyimide films	
Expansion (TCE), ppm/C	11-32
Humidity, ppm/% RH	7-22
Initial shrinkage, ppm*	100-1700

* Shrinkage range includes TYP & MAX values for 30min-2hr @ 150C & 200C

Table 1 Stability parameters for polyimide films.

We choose to address this instability by restricting our edge interconnects to the two long edges of the module. While this does reduce our inter-layer routing resources by a factor of two in the worse case (e.g. for a square module), it enables three advantages: (1) single-axis processing, (2) using web tension for dimensional adjustment during lay-up, and (3) using the edge-exiting traces themselves as fiducials during continuous alignment and layer tacking. Single-axis processing is broadly beneficial, offering the potential to simplify most material handling and processing operations including inspection, coating, drying, alignment and tacking, lamination and thermal processing, and edge sawing. The other two advantages apply to assembling the stack.

Polyimide is transparent at wavelengths from 0.6-5 µm, that is, from green through infrared, and it may be possible to illuminate the partially completed stack from below and use the edge-exiting traces at the top of the stack as fiducials for aligning the subsequent layer. If bottom

illumination proves insufficient, light can be coupled into the edge of the stack at or near the topmost layer. By curving either the partially completed stack or the layer being added over a large-radius, transparent drum[2], it may be possible to perform alignment as a continuous roll-based process. If interfacial forces are insufficient to stabilize layer alignment within the stack through lamination and curing, the joint may require discrete or continuous tacking of the B-staged adhesive using a thermode, a laser, or some other means.

After lamination, the stack is sawn to expose the exit traces on the long edges of the individual modules. This sawing will likely be done with equipment similar to that used for singulating chip-scale packages or other high-density single-chip packages. The saw must stop short of cutting through the stack to facilitate subsequent handling in un-singulated format. If this edge sawing is done in the cross-web direction, it may allow handling of the sawn substrates in a segmented or "lumpy" roll format for subsequent interconnect processing and encapsulation. This comes at the cost of having to locally halt the web motion for each crosscut. If the edge is sawn parallel to web, the stack may prove too stiff for subsequent roll processing. However this orientation facilitates using web tension during lay-up to reduce the dimensional mismatch between the partially completed stack and the next layer being added. While humidity control and high- spatial resolution active temperature control also offer potential solutions, utilizing web tension may allow shorter control-system time constants and higher production speeds.

The finish and shape of this sawn groove are likely to be extremely critical to subsequent interconnect formation. While straight sidewalls and generous corner radii simplify saw procurement and maintenance, sloped sidewalls and sharp corners may be required to optimize subsequent processing, handling and inspection. Custom saws are available with up to a 30° sloped sidewall, but they cost more than plain saws, have a substantial delivery lead-time, and cannot be re-dressed. The sawn groove is cleaned and plasma de-scummed to expose the exiting traces. This operation may remove only enough material to clean the metal surface, or a longer polymer etch may be used to leave the traces projecting a few microns from the polymer surface. The surface is now ready for D2D wiring.

[2] While registration requirements are quite different, our proposed lay-up process bears a superficial resemblance to that used to wind stacked-film capacitors on large diameter core wheels.

Geometric Analysis of D2D Interconnection Processes

The primary goal for our geometric analysis of D2D is to create models that accurately predict the minimum wiring pitch attainable for any given set of edge interconnect process parameters. While in its most general form, D2D is applicable to a wide range of geometries, we'll begin with circular wiring elements built-up from dimensionless points on a plane. We'll briefly show how the process might be extended to multi-layer geometries with crossovers. The most important aspects of 3-D buildup will be explored, followed by a series of models using D2D to connect finite-sized pads on a sawn module edge and including the most important expected process variations. Our general approach will be to develop a set of inequalities defining bounds on acceptable geometry, and to show where particular bounds apply.

Basic 2-D models:

We begin with a simple two-dimensional model, and consider a collection of dimensionless points scattered about on a plane (Fig. 1). Each point represents a discrete terminal node in some external electrical network. The points are collectively *activated*, and each point grows or builds-up an electrically conductive circle of radius b. Wherever two points are separated by a distance of less than 2b, their respective circles will join and their respective terminals will be electrically connected. Where separation is greater than 2b, no circle-circle connection is made. Long wires connecting a series of terminals are created wherever a series of points are separated by <2b. Wires branch wherever a circle contacts more than two neighboring circles. A planar surface is not required; this model could be extended to arbitrary surface topologies.

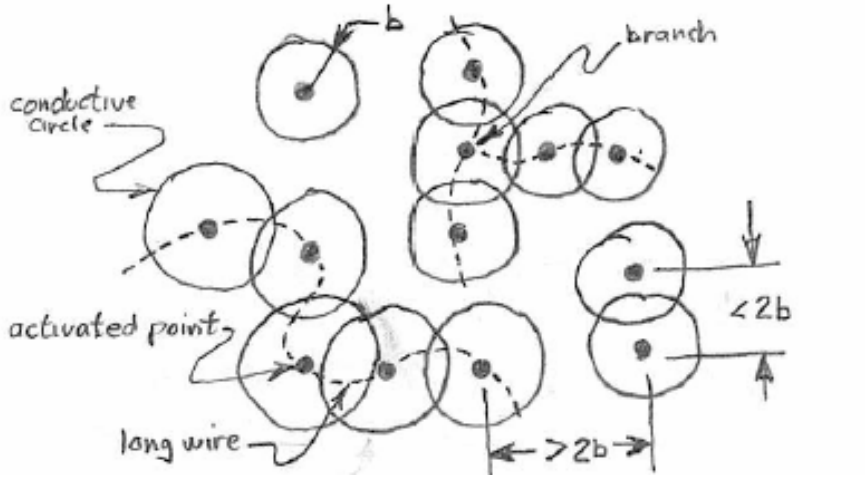


Fig. 1 D2D connections from dimensionless points.

These dimensionless point assumptions can also be applied to create far-field models for finite-sized activated regions, and will provide an accurate connectivity approximation when the activated regions are much smaller than their separations. We'll first perform a dimensionless point analysis to create a series of far-field models specific to our laminated stack geometry. These models will initially be expressed as a series of inequalities defining bounds on minimum buildup b_{MIN} and maximum buildup b_{MAX}. Input parameters will include the minimum build-up required to make a connection, the maximum allowed to avoid shorts between non-connected terminals, and the effect of inter-layer registration accuracy. These models pave the way for subsequent near-field models that more accurately account for finite-sized activated regions with dimensions comparable to their separation.

We begin with a rectangular grid of dimensionless activated points that represent a perfectly aligned layer stack with terminals having dimensions much smaller than their distance from each other. Terminals within a layer are characterized by their minimum horizontal spacing s_1. Layer thickness is specified by the nominal vertical spacing s_2. As shown in Fig. 2, a vertical column connection will always be made when:

$$b_{MIN,1} > s_2 / 2$$

And a horizontal short to an adjacent column will not occur as long as:

$$b_{MAX,1} < s_1 / 2$$

As a limiting case, with perfect control of build-up ($b_{MIN} = b_{MAX}$) it is sufficient to require that $s_2 < s_1$ to assure correct interconnection.

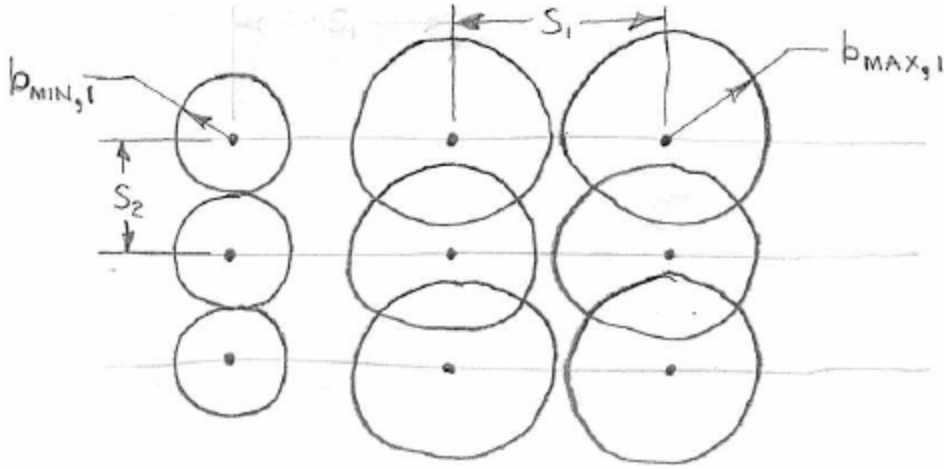


Fig. 2 Perfectly aligned stack with dimensionless terminals.

But life is not quite so simple. Inter-layer misalignment or *registration error* is expected to be an important factor limiting edge interconnect density. Adding registration error r to our model, we apply the Pythagorean theorem to calculate a second set of bounds on buildup. Fig. 3 shows that to assure a connection within the column now requires:

$$b_{MIN,2} > [s_2^2 + r^2]^{½} / 2$$

And to avoid shorting to adjacent columns:

$$b_{MAX,2} < [s_2^2 + (s_1\text{-}r)^2]^{½} / 2$$

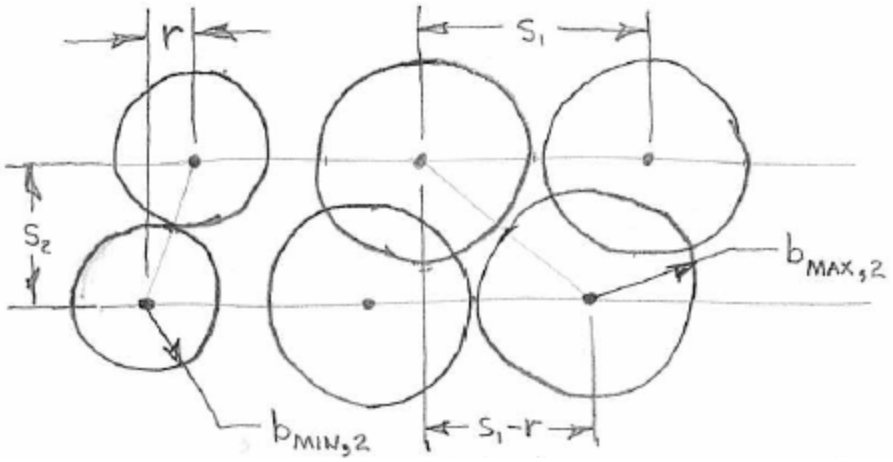


Fig. 3 Misaligned layer in stack with dimensionless terminals.

Combining the two inequalities confirms our intuition that when $r > s_1 / 2$, $b_{MIN,2} > b_{MAX,2}$ and connectivity conditions cannot be met; short circuits between columns would occur before the connections are made within a column. This bounding inequality will prove useful for interpreting parametric studies. We'll call it the *global base case* and restate it as:

$$s_1 = 2\,r$$

The inequality $b_{MIN,1} > s_1 / 2$ is no longer required, as the expression for $b_{MIN,2}$ correctly covers the case where $r = 0$. But the inequalities for both $b_{MAX,1}$ and $b_{MAX,2}$ are required to bound b_{MAX}. When $r \leq$ some threshold value r_0, the closest shorting opportunity is within a horizontal layer and the $b_{MAX,1} < s_1 / 2$ constraint is binding. For $r_{MAX} > r_0$ however, the closest shorting opportunity is to an adjacent layer, and the $b_{MAX,2}$ constraint will be binding. Setting $b_{MAX,1} = b_{MAX,2}$, applying the quadratic equation, and selecting the correct root by observing that for $s_1 >> s_2$, $r_0 > 0$, we find the threshold value for registration error:

$$r_0 = s_1 - (s_1^2 - s_2^2)^{½}$$

This threshold registration error r_0 is significant in that there is no benefit to improving the accuracy of an alignment process beyond this value.

Thus far we have required only that adjacent circles touch for an electrical connection to be established, and that circles separated only by an infinitesimally small distance will not short. In more realistic scenarios, obtaining adequate joint strength or conductance will require some minimum joint overlap width w (Fig. 4). Similarly, the need to optically inspect completed interconnects and to meet leakage or breakdown specifications will require some minimum gap v between adjacent columns. We must extend our model to incorporate w and v.

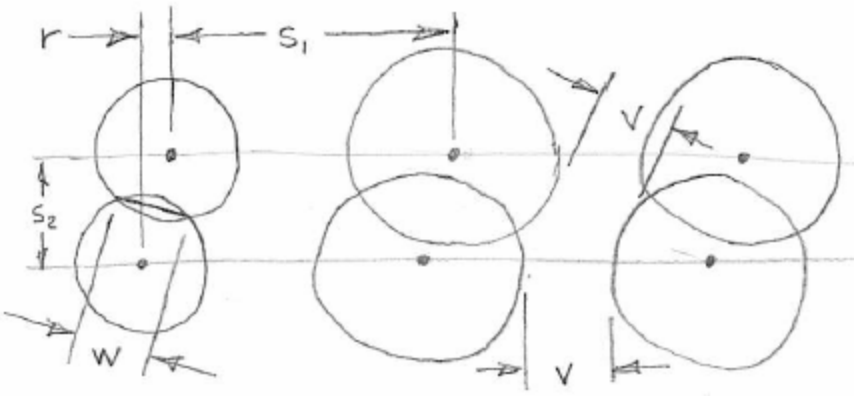


Fig. 4 Minimum joint width w and minimum visual gap v.

Applying the Pythagorean theorem twice we find the minimum buildup to meet joint width requirements:

$$b_{MIN,3} = [s_2^2 + r^2 + w^2]^{½} / 2$$

This new inequality encompasses cases where $r = 0$ and $w = 0$ so it supercedes both previously derived bounds on b_{MIN}. For bounds on $b_{MAX,3}$, we subtract $v / 2$ from the previous two inequalities to find:

$$b_{MAX,3a,\, r < r0} \leq (s_1 - v) / 2$$

$$b_{MAX,3b,\, r = r0} \leq \{ [s_2^2 + (s_1 - r)^2]^{½} - v \} / 2$$

While our analysis has so far treated b_{MIN} and b_{MAX} as independent variables, their difference $b_{MAX} - b_{MIN}$ can be considered the build-up tolerance range, which is a function of process variability. We can express the tolerance range by a linear model:

$$b_{MAX} - b_{MIN} = b_0 + c\, b_{MIN}$$

b_0 is the tolerance range at $b_{MIN} = 0$, and could represent start time uncertainty. c represents uncertainty in the spatial and temporal uniformity of the build-up rate. Depending on our requirements, we can choose b_0 and c to represent either the entire range possible with a given process, or to more accurately model a limited build-up range. For a sufficiently limited range, one might dispense altogether with the rate term c.

We can now apply our far-field models to predict the minimum attainable wiring pitch P, which is equal to the horizontal spacing s_1 when the terminals are dimensionless. We rearrange the tolerance model to express b_{MAX} in terms of b_{MIN}:

$$b_{MAX} = b_0 + (1+c)\, b_{MIN}$$

And replace b_{MIN} with the bounding value of $b_{MIN,3}$ to find the minimum value of b_{MAX}

$$b_{MAX} = b_0 + (1+c)\, [s_2^2 + r^2 + w^2]^{½} / 2$$

Substituting this result into the expressions for $b_{MAX,3a,\, r < r_0}$ and $b_{MAX,3b,r=r_0}$ and solving for s_1, we obtain two lower bounds on pitch P_{MIN}. As before, which constraint is binding depends on the value of r_0:

$$P_{MIN:\, r<r_0} = s_1 = v + 2b_0 + (1+c)\, [s_2^2 + r^2 + w^2]^{½}$$

$$P_{MIN:\, r=r_0} = s_1 = r + \{[v + 2b_0 + (1+c)(s_2^2 + r^2 + w^2)^{½}]^2 - s_2^2\}^{½}$$

Multilayer D2D:

Before extending our 2-d models to encompass 3-d buildup and finite terminal dimensions, we'll illustrate how a planar D2D process could be used to define additional layers, creating crossovers and arbitrarily complex multi-layer circuits. Conceptually, this is done by repeatedly:

1) Building up a set of circles of radius b to create a layer of D2D wiring.
2) Depositing an infinitesimally thin electrically insulating layer over the entire surface.
3) Patterning a new set of points electrically connected to the layer below.
4) Activating those points.

Note that while steps 1, 2, and 4 can be bulk processes, step 3 is cannot. Either mask-based or sequential patterning is required, as with conventional via opening techniques. Therefore, D2D will not be attractive in many multilayer applications, particularly those where the number of layers or vias is large. For applications where a single layer of wiring plus a small number of crossovers are sufficient, a D2D multi-layer process could prove economical.

As an illustration of such process, consider the following highly idealized multi-layer geometry where only two wiring layers are required to create a crossover and layers can be added ad infinitum. Design rules constrain activated points on all layers to a square grid of dimension a (Fig. 5a). If build-up radius b is greater than a, points spaced at distance 2a will be connected. However, points spaced a distance $5^{½}$a apart (e.g. those separated two grid-points in one axis, and one grid-point in the other) will not be connected as long as $b < 1.25^{0.5}$ ($b \approx 1.118a$). Setting the nominal value of b to the geometric mean of these two bounds results in $b = 1.25^{0.25}$ ($b \approx 1.057a$). In this system, a crossover can be created between two points originally a distance of 4a apart by adding a second layer with just two activated points on the grid a distance of 2a apart, as shown in Fig. 5b.

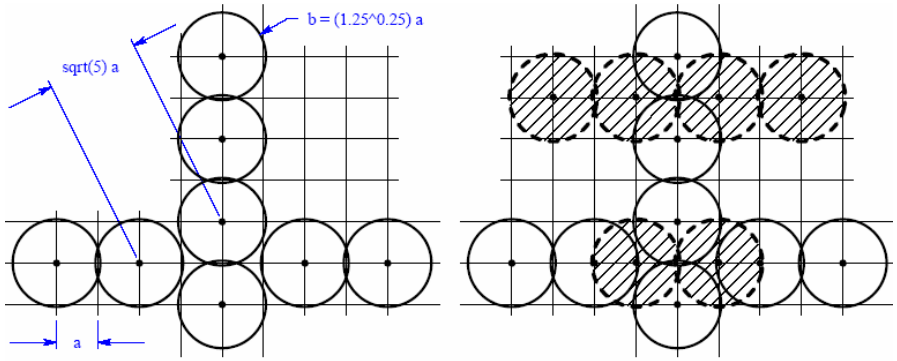


Fig Za. First layer on grid with vertical wire. Fig Zb. Adding a second layer creates crossovers.

Fig. 5 Multilayer D2D: a) First layer on grid. b) Adding second layer creates crossovers

Note that where only a single layer of wiring plus crossovers is required, or where the wiring pitch can change with layer count, design rules for subsequent layers can differs from those on the first. In the single layer plus crossovers case, we could increase the spacing between the two activated points used to define the crossover, and use a larger buildup to bridge the gap.

Hybrid approaches combining mask-based or imprint lithography, sequential processing, and D2D are also possible. For example, crossovers could be added to a single-layer fine-pitch flex circuit by coating the wiring with a dielectric, opening vias on both sides of each crossover with a laser, and using D2D to bridge the gap.

Three-dimensional build-up:

While our 2-dimensional, far-field models illustrate some of the fundamentals of laminated-stack D2D geometry, further analysis is needed to accurately predict wiring density in the *near field*, where the terminal size and spacing are comparable. The technologies used to pattern the layer-exit traces are generally specified with rules for minimum line width l_1 and spacing s_1. To maximize stack-edge wiring density, we'd ideally like our stack-edge interconnect pitch $P=l_1+s_1$ to approach what is possible within the individual layers. We'd like to extend our models to provide accurate predictions in this near-field regime. It is useful to begin this analysis by first considering the three-dimensional geometry common to most real-world D2D build-up processes.

We extend our original 2-d, dimensionless-terminal, far-field model to three dimensions by considering a set of activated points scattered about a 3-d space. Around each point there is a 3-d conductor buildup forming a sphere of radius b (Fig. 6).

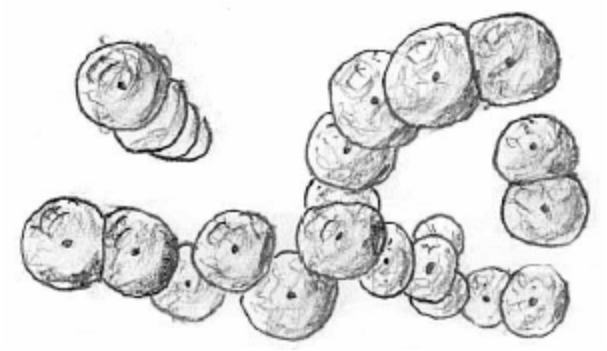


Fig. 6 Spheres surrounding points in space form 3-D D2D.

As before, points separated less than 2b are connected and those with a greater separation are not. While it's challenging to imagine a real physical process that instantiates this most general case, 3-d spherical buildup from points on a 2-d surface is a useful model for our laminated stack edge interconnect and many others. In D2D processes currently under consideration, buildup will indeed be three-dimensional, initiating from a planar array of finite-sized activated terminal regions called *pads.* To visualize this, first consider the idealized buildup from a dimensionless point on a planar surface, forming a conductive hemisphere as shown in Fig. 7a. When two points are separated less than 2b, their respective hemispheres overlap to form a *knit plane*, the semicircular region shown in Fig. 7b. For a joint width w, this overlapping region has area $A_w = \pi w^2 / 8$

Fig. 7 Buildup on plane surface from: a) One dimensionless point. b) 2 points, intersecting at a knit plane.

Extending this model, a line on a surface can be considered a one-dimensional array of points with inter-point spacing approaching zero, a knit plane semicircle approaching radius $b_{0,}$ and cross-sectional area:

$$A_b = \pi\, b_0^2 / 2$$

The complete build-up volume around a line forms a linear conductor, which can be thought of as the superposition or union of the set of associated hemispheres.

Real pads have finite extent in two dimensions. The circular pad in Fig. 8a can be thought of as an infinite number of points, the union of their associated hemispheres defining the shape of the conductor buildup. Note that the top of this conductor has a flat region the same size and shape as the underlying pad. The same is true for Fig. 8b, which depicts buildup on a square pad.

Fig. 8 Buildup from finite pads: a) Circular pad. b) Square pad.

A comparison of this idealized square pad an actual EL-Ni bump in Fig. 9 suggests that this union-of-hemispheres model is a reasonable representation for the geometry of our targeted process.

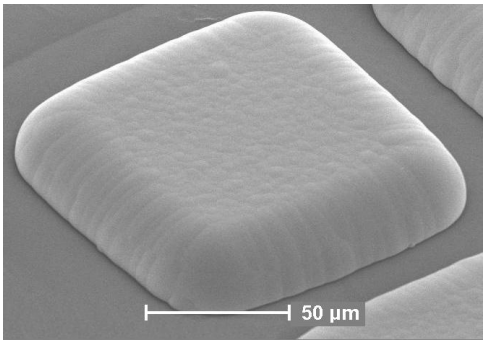


Fig. 9 35 µm high electroless-Ni bump plated on 72 µm square copper pad.

Where buildups from arbitrarily shaped pads meet and continue to grow, the resulting knit region may no longer be a plane, and complex 3-D models may be needed to characterize the joint in detail. Instead, we'll rely on the simplified intersecting hemisphere joint models for the remainder of our analysis. This approach will provide conservative approximations of the actual knit contact area and minimum attainable wiring pitch for the edge interconnect geometries of interest.

Finite activated regions:

We now return to our dimensionless point models, and apply them to interconnections at the edge of our laminated stack of flex circuits. Our assumption is that we want an electrical connection between each flush pad (the sawn end of a flex circuit trace) within a vertical column, but with no shorts between adjacent columns. It's worth noting that if this full-column connection architecture is adhered to for an entire stack, and no traces are deleted to create intentional breaks in the vertical conductor columns, then the order in which the layers are stacked does not affect the DC connectivity of the circuits.

The metal on a real circuit layer has thickness l_2, and the patterned traces have linewidth l_1 and spacing s_1. Where these traces intersect the sawn stack edge, approximately rectangular pads of height l_2 and width l_1 will be formed (Fig. 10). We note that the trace conductor thickness l_2 has no apparent role in determining if buildup will result in connections or shorts, and that our dimensionless point analysis can be applied without modification to evaluating shorts between columns. As before, there are two bounding inequalities for b_{MAX}, one to cover shorts within a layer ($r < r_0$) and another for column shorts between adjacent layers ($r > r_0$).

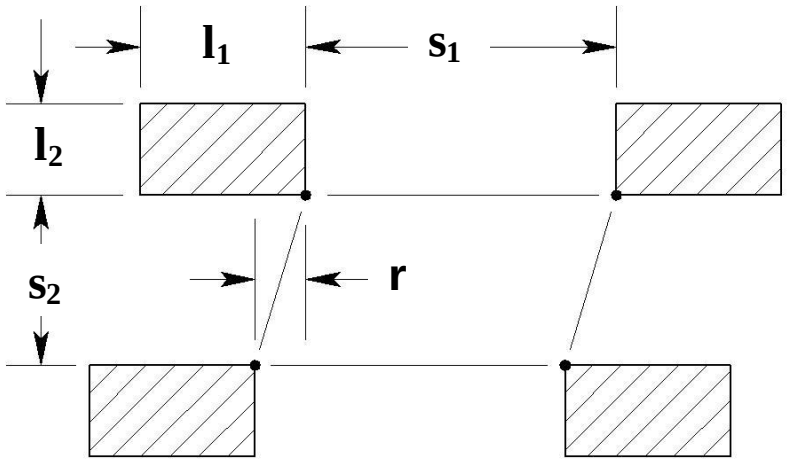


Fig. 10 Finite activated regions: pads formed by layer wiring traces at sawn edge of laminated stack.

Our prior analysis needs some modification to correctly describe connectivity conditions within a column. Fig. 10 shows that as long as the registration error r is less than the trace width l_1, the minimum buildup b_{MIN} required to initiate a column connection is constant at $s_2/2$. In this $r < l_1$ regime, we could try to specify a required minimum knit surface area, or a minimum knit line width w, but we find that the geometry has become more complex. The knit surface is no longer planar, and it becomes important to ask more detailed questions about joint requirements. If we expect the knit surface to be weaker or more electrically resistive than the rest of the material, then the actual area of the knit surface will be important. If however we expect the knit surface to be indistinguishable from the bulk, a full 3-D analysis may be required to assess the strength or resistance of the column and its connections to the traces. It's also possible the local joint properties may change continuously during buildup as the knit area forms.

To simplify our finite-pad analysis, we'll assume the column joint is formed by buildup initiated at only two points where adjacent column pads are closest to each other. By considering only two points, our prior dimensionless point analysis can be directly applied, providing a conservative bounding model. Any finite pad pair should have a more robust column joint than one initiated at just the two closest points.

We now have two bounding inequalities for b_{MIN}, one for $r < l_1$ and another for $r = l_1$:

$$b_{MIN,r<l_1} = [s_2^2 + w^2]^{½} / 2$$
$$b_{MIN,r=l_1} = [s_2^2 + (r-l_1)^2 + w^2]^{½} / 2$$

The bounding inequalities on b_{MAX} are unchanged, and are restated here with consistent nomenclature to clarify their new finite-pad role:

$$b_{MAX,r<r_0} \leq (s_1 - v) / 2$$
$$b_{MAX,r=r_0} \leq \{ [s_2^2 + (s_1-r)^2]^{½} - v \} / 2$$

As with our dimensionless point analysis, we can derive expressions for the minimum attainable space s_1, but in this instance we have four equations. Which of the four applies depends on the relationship of the maximum registration error r to both its threshold value r_0 and the trace or pad width l_1.

$$s_{1,MIN:\, r<l_1, r<r_0} = v + 2b_0 + (1+c)\,[s_2^2 + w^2]^{½}$$
$$s_{1,MIN:\, r<l_1, r=r_0} = r + \{[v + 2b_0 + (1+c)(s_2^2 + w^2)^{½}]^2 - s_2^2\}^{½}$$
$$s_{1,MIN:\, r=l_1, r<r_0} = v + 2b_0 + (1+c)\,[s_2^2 + (r-l_1)^2 + w^2]^{½}$$
$$s_{1,MIN:\, r=l_1, r=r_0} = r + \{[v + 2b_0 + (1+c)(s_2^2 + (r-l_1)^2 + w^2)^{½}]^2 - s_2^2\}^{½}$$
$$\text{where: } r_0 = s_1 - (s_1^2 - s_2^2)^{½}$$

Since $r_0 = f(s_{1,MIN})$, more than one equation must be evaluated and have its corresponding r_0 calculated in order to determine which equation applies. Any solution that specifies $s_{1,MIN} < s_2$ violates our most fundamental connectivity requirement and must be rejected. Note that setting $l_1 = 0$ corresponds to our far-field case, and the equations above reduce to those we derived earlier.

Any real trace patterning process will specify some minimum linewidth $l_{1,SPEC}$ and space $s_{1,SPEC}$ and we must also verify that these specified minimums have been satisfied. While the value of l_1 can be checked during input variable selection, $s_{1,MIN}$ cannot be checked until its value is calculated. If $s_{1,MIN}$ is found to be out of range, it must be set to its specified minimum value:

$$\text{if: } s_{1,MIN} < s_{1,SPEC}, \text{ then:}$$
$$s_{1,MIN} = s_{1,SPEC}$$

We now have everything we need to determine the minimum attainable pitch:

$$P_{MIN} = l_1 + s_1$$

A final comment: We could eliminate one input parameter by non-dimensionalizing our model, such as dividing all our linear dimensions by the substrate thickness s_2. We'll resist this temptation because it would tend to obscure the overwhelming importance of reducing the registration error r, as will be shown in the following section.

Parametric Studies:

The goal of our parametric studies is to predict the minimum wiring pitch attainable with practical layer stacking and D2D processes. We'll first explain our choice of parameters believed representative of proposed processes, then use our models to evaluate a series of graphs to clarify the key issues in creating D2D connections, and finish by predicting the maximum edge wiring density likely to be attained with our PIRM module and related stacked, flex-based heterogeneous systems.

The substrate thickness s_2 imposes a fundamental bound on D2D wiring density. Polyimide films are available in a variety of grades from DuPont (Kapton), Ube Industries (Upilex), and Kaneka (Apical). All brands and types are available in 25 µm and 50 µm thicknesses, with wide availability from 12.5-75 µm, and more limited availability from 7.5-125 µm. Our development is currently based on 50 µm film and much thinner films expected in production. To span our entire range of interest, we'll select three thicknesses: s_2 = 6.25, 25, and 100 µm. The 6.25 µm value provides a reasonable lower bound on future film availability and R2R web handling capability. The 100 µm value will require thick (>50 µm) plating, but is typical for thin layers in conventional multi-layer PCB fabrication. It also approximately represents the state-of-the-art for embedded thinned-die packaging.

The maximum layer-to-layer registration error r will be seen to have profound effect on attainable wiring density. As mentioned earlier in this paper, a 250 µm maximum registration error specification is typical for pin-aligned commodity multilayer PCBs. Considerably less than 75 µm is expected to be possible for a single-axis, closed-loop, continuous alignment, tacking and lamination process.

For both minimum trace width l_1 and spacing s_1, we'll evaluate values of 6.25, 25, and 100 µm. While much denser wiring has been demonstrated with imprint lithography and with additive processes, l_1=s_1= 25 µm is approximately state-of-the-art for high volume tape-automated-bonding and flex circuit vendors [26] and l_1=s_1= 100 µm is typical for conventional subtractive process PCBs.

The EL-Ni build-up process we are evaluating was originally developed by researchers at Fraunhofer IZM and Technical University Berlin as a means of fabricating low-cost bumps and under-bump-metallurgy on silicon wafers having aluminum wire bonding pads. Wafer bumping services and equipment – including EL-Ni bumps on copper pads – are now offered commercially [18],[19]. Typical plating thicknesses vary from a few microns when used as under-bump metallurgy for solder balls, to 15-20 µm for bumps for adhesive-based flip-chip attach, but thicker plating is possible. Vendor design rules typically require a minimum pad spacing of two times the nominal bump height plus 10 µm. Taken by itself, this pad spacing rule suggests that we might evaluate our models using b_0 = 5 µm and c = 0 as process parameters. However, vendors also routinely specify that bump heights of up to 20 µm can be controlled to better than ± 2 (or + 4 / – 0 µm). This suggests that a parameter set with b_0 = 4 µm and v = 2 µm might be more appropriate for our purposes. To allow for additional plating rate uncertainty at > 20 µm build-ups, we'll set c = 5%. Our studies will treat these parameters collectively, and apply the label "Std_EL-Ni" when b_0 = 4 µm, v = 2 µm, and c = 5%. Otherwise, all three parameters will be set to zero.

While prior to our experiments, we had no hard information to support an informed choice of minimum overlap w, we expected that the mechanical properties of the columns would be important and that stresses in the plated deposits might be an issue [20]. So for our parametric studies, when w is set to a non-zero value, that value will be within a factor of four of the substrate thickness. This range was chosen to clarify how much significance the overlap specification could have on minimum attainable pitch.

While we've greatly reduced our parameter space by limiting the number of values each may assume, a full combinatorial analysis would still result in a large amount of data to interpret. Instead, we'll present a more limited series of graphs with a goal of illustrating the effect of variations in the most important parameters and predicting the potential wiring density for various levels of D2D technology development. Substrate thickness s_2 imposes a fundamental bound on wiring density, and each graph will depict a single substrate thickness with the registration error r as the independent variable and the minimum pitch P_{MIN} as the dependent variable. Each graph will include a global base case plot where every parameter is set to zero and a second local base case plot with all parameters except substrate thickness s_2 are set to zero. Other plots will be added, setting other parameters to non-zero values, and typically finishing with all parameters set to non-zero values believed feasible in practice.

We begin with side-by-side graph pairs having parameter sets that are identical except for substrate thickness s_2. The first two graph pairs have plots for two non-zero linewidths l_1, and two non-zero buildup overlaps w, which match the two linewidths. The third graph pair is similar to the second, but fixes the linewidth and introduces the effect of the plating process tolerance "Std_EL-Ni". The seventh and final graph depicts our proposed initial development target, with substrate thickness s_2 = 25 µm, linewidth l_1 = 25 µm, and standard plating process tolerances. This final graph varies the overlap w over a wide range to emphasize the importance of future work to quantify overlap requirements.

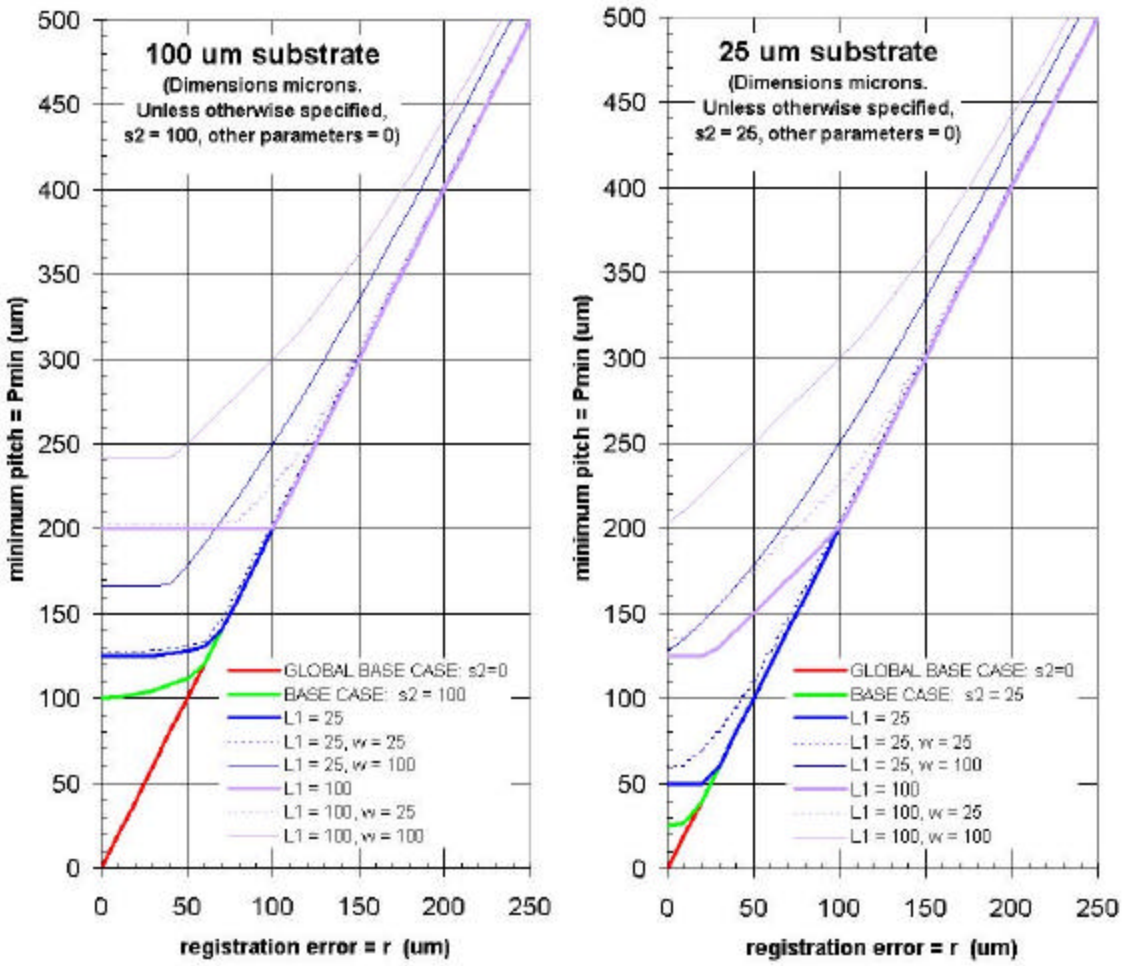


Fig. 11 100 and 25 µm substrates: 25 and 100 µm linewidths and overlaps, 0-250 µm registration error.

Fig. 11 depicts 100 and 25 µm substrates with 25 and 100 µm linewidths and overlaps, and 0-250 µm registration error r. The global base case, with all parameters set to zero, is a straight line with slope of two since the minimum wiring pitch P_{MIN} can never be less than twice the registration error r. The base case plot (everything except s_2 set to zero) shows that $P_{MIN} = s_2$ when r = 0, and that P_{MIN} approaches the global base case as the registration error increases. For both 25 and 100 µm values of linewidth l_1, at r = 0 we find that P_{MIN} increases beyond the base case by an amount equal to the linewidth. Both graphs show that when the linewidth is equal to the substrate thickness, the minimum pitch is constant with increasing registration error until the base case lines are met and tracked. When the substrate is thicker than the linewidth, as when s_2 = 100 µm and l_1 = 25 µm, this flat portion of the plot extends only to r = l_1, where the column conductor spacing begins to increase. When the substrate is thinner than the linewidth, as when s_2 = 25 µm and l_1 = 100 µm, the flat portion extends only to where the registration error approaches the substrate thickness; then inter-layer shorts become the binding constraint. When the required overlap W becomes non-zero, the minimum pitch increases. This increase is modest for small registration errors since as the buildup circles touch and then merge, their intersecting region increases quickly in size. At larger overlaps, the effect can become quite significant. Finally, these graphs illustrate the fundamental importance of reducing the inter-layer registration error, and the impetus for developing accurate single-axis alignment. With conventional pin-alignment having r = 250 µm, all the parameter sets shown approach P_{MIN} = 500 µm, but there is no way to get below it.

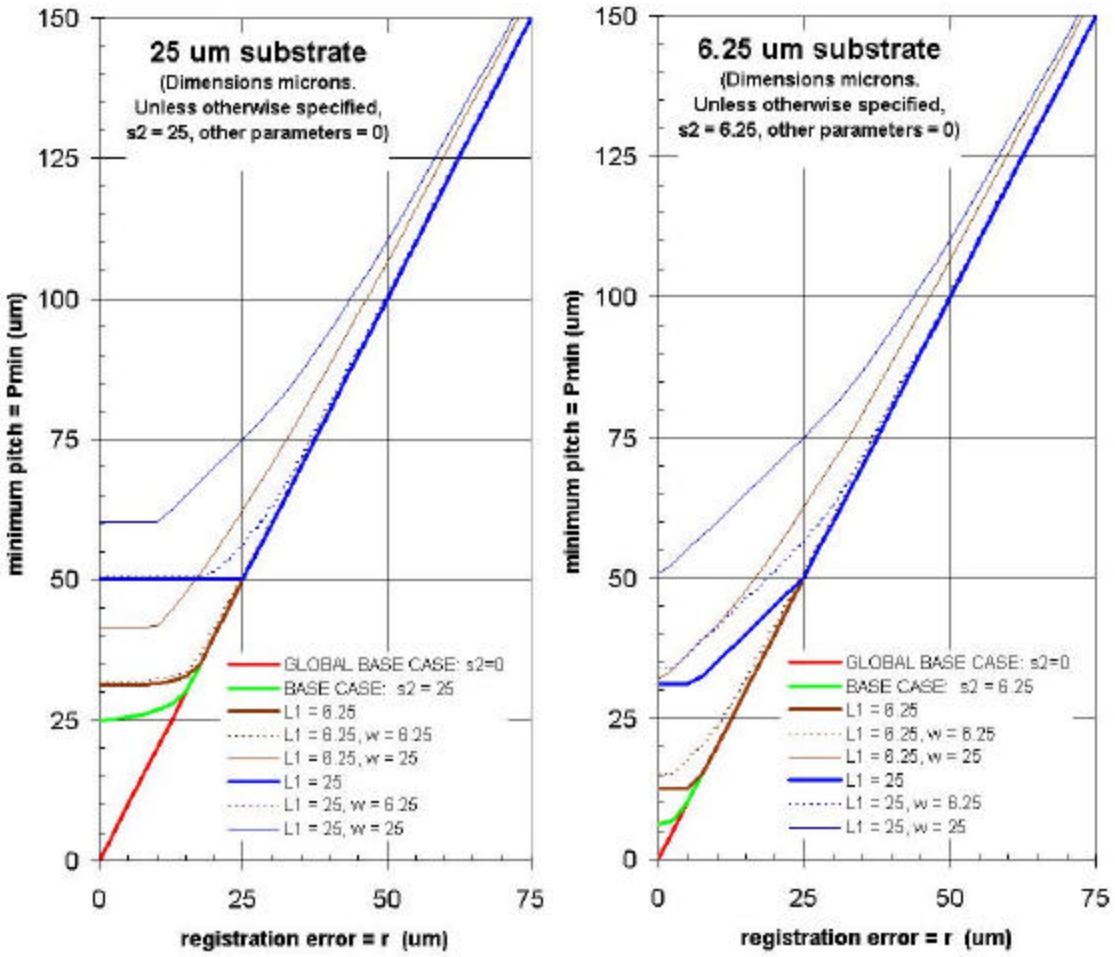


Fig. 12 25 and 6.25 µm substrates: 6.25 and 25 µm linewidths and overlaps, 0-75 µm registration error.

The pair of graphs in Fig. 12 looks virtually identical to the previous pair, and indeed they are. They depict a 4x shrink of all dimensions, with 25 and 6.25 µm substrates and 6.25 and 25 µm linewidths and overlaps. These substrate thicknesses cover the range from our initial target implementation down to the minimum believed practical. The registration error r ranges from 0 to 75 µm, again illustrating the importance of improving the 250 µm typical for pin-aligned multi-layer circuit lay-up. In the 6.25 µm substrate graph, the l_1 = 25 µm plots all exhibit a large range where the increase in minimum pitch appears directly proportional to the increase in registration error. The governing equation for $r < l_1$, $r \geq r_0$, shows that this is indeed the case. Finally, it is worth noting that the two graphs are identical for $r \geq r_0, l_1$, where the governing equation with these parameter sets is the same for both.

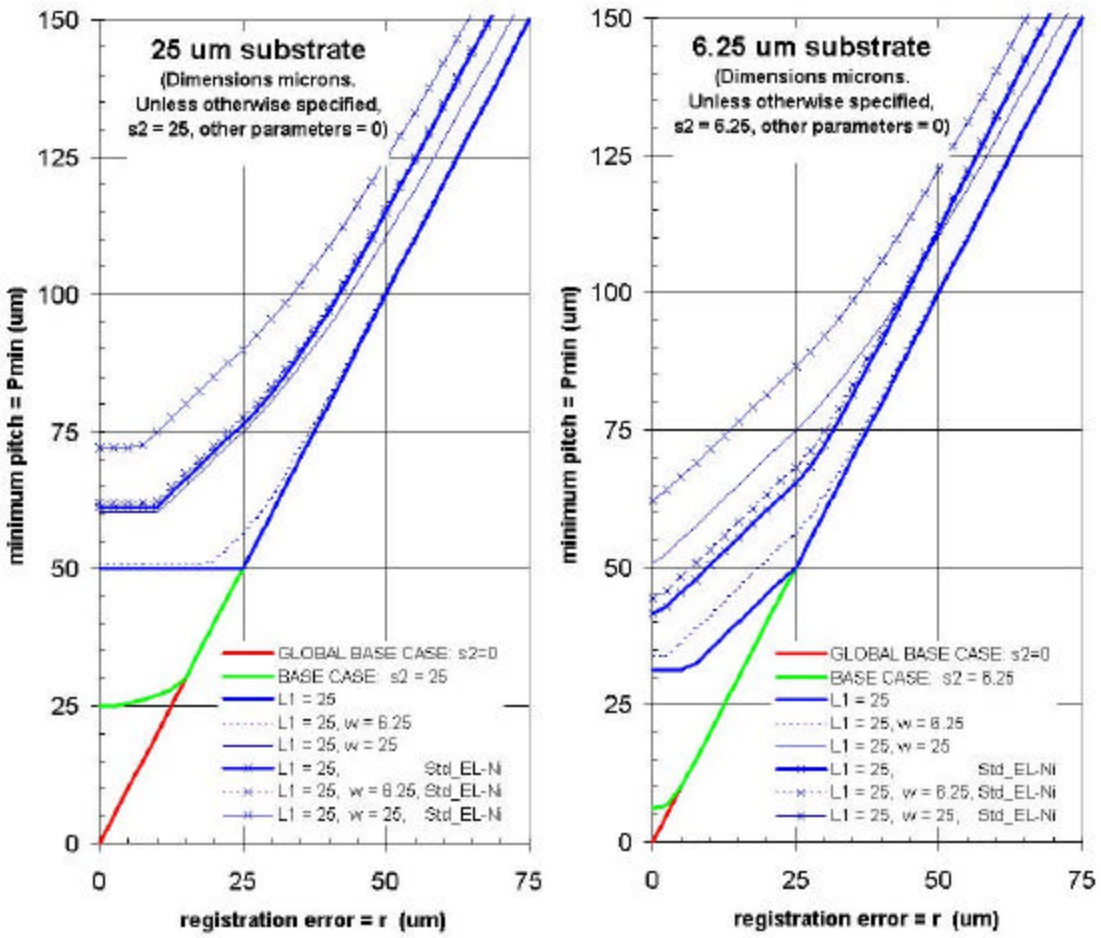


Fig. 13 25 and 6.25 µm substrates with 25 µm linewidth and Std_EL-Ni plating tolerances.

Fig. 13 is similar the previous pair but is limited to a 25 µm linewidth and includes the effect of the plating tolerance parameter set "Std_EL-Ni" (b_0 = 4 µm, v = 2 µm, and c = 5%), characteristic of today's EL-Ni wafer bumping services. Unlike the previous graph pairs, we no longer see a value of r beyond which the two graphs are identical; with the introduction of the plating tolerance, the substrate thickness assumes greater importance.

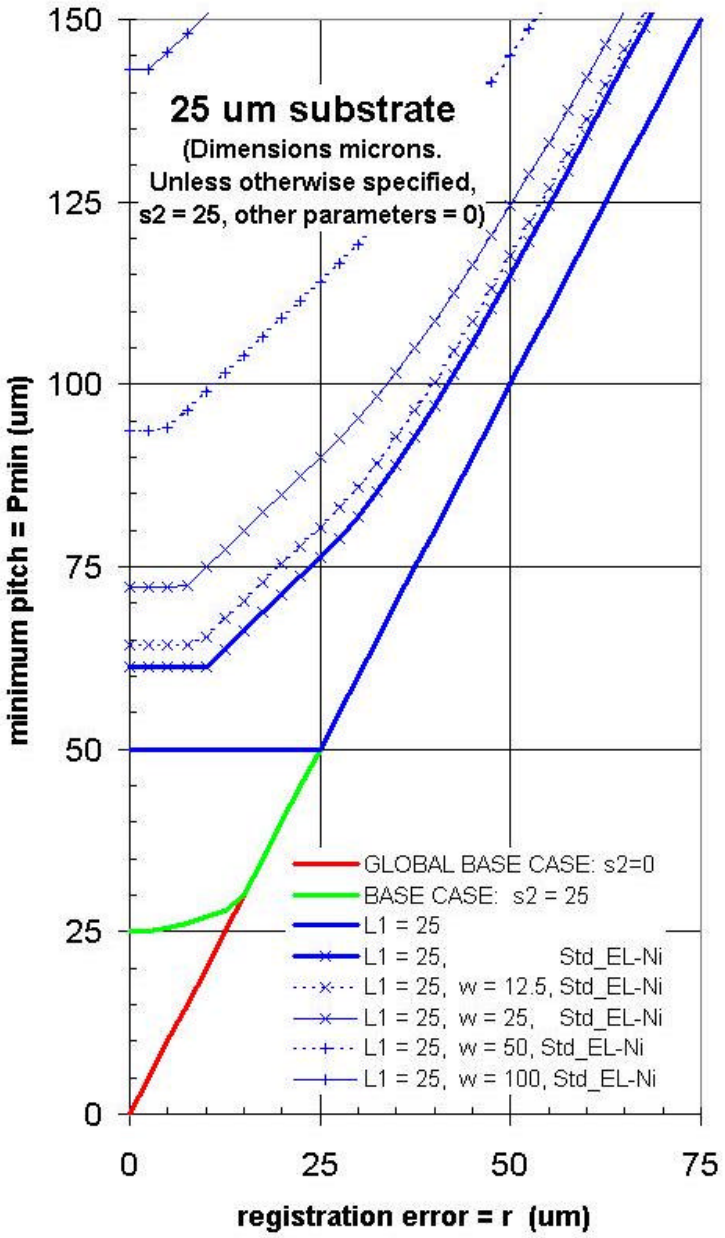


Fig. 14 Development target: 25 µm substrate with Std_EL-Ni plating tolerances and 12.5-100 µm overlap.

Fig. 14 depicts our proposed initial development target, with substrate thickness s_2 = 25 µm, linewidth l_1 = 25 µm, and with standard plating process tolerances. The overlap w is varied over a wide range to emphasize the importance of future work to evaluate the mechanical and electrical properties of D2D joints so that overlap requirements can be correctly specified. The graph does suggest that if adequate registration error can be attained, and modest overlaps prove sufficient, sub- 200 µm pitches should be within easy reach.

Experimental Results:

The goal of initial experiments was to answer three questions: 1) What is required to create an individual EL-Ni D2D connection? 2) What pitch is needed to assure isolation between adjacent columns? 3) How difficult will it be to attain high yields? Early results support the viability of EL-Ni D2D in all three areas. The plating buildup behaved much as predicted in the preceding geometric analysis, with D2D connections readily established and column isolation easily maintained. Numerous isolated test patterns with r = 0 demonstrated wiring pitches below 50 µm, with increasingly robust conductors and higher yields at larger pitches.

The experiments used a copper-Kapton flex circuit as a proxy for the module sidewall. Our flex circuit lacks some important features of an actual sawn sidewall, such as registration errors and the potentially activation-inducing presence of saw damage, tramp metal and adhesive interfaces. However the simple material system and smooth polymer surface present a best-case scenario for an economical set of initial experiments. Our substrates were fabricated on 50 µm thick Kapton having a thin Cr tie-coat and 2um of Cu. Using a 5" glass mask and spun-on negative photoresist, the Cu was etched to form an array of test patterns (Fig. 15) with square and rectangular pads having mask dimensions of 20, 40 and 80 µm. The spaces between pads assumed one of nine values from 20-80 µm in approximately $2^{0.25}$ (~20%) increments. The 20x20 µm pads rounded during etching, becoming ~12 µm diameter and easily delaminating at the Cr-polyimide interface during etching and subsequent processing. The larger pads remained firmly attached, and had length and width typically 8 µm less than the mask artwork. To help us assess the importance of surface preparation, the test substrates were subjected to five different plasma-cleaning processes (none, Ar, Ar-O_2, O_2, and O_2-CF_4), each in a horizontal stripe across the width of the substrate.

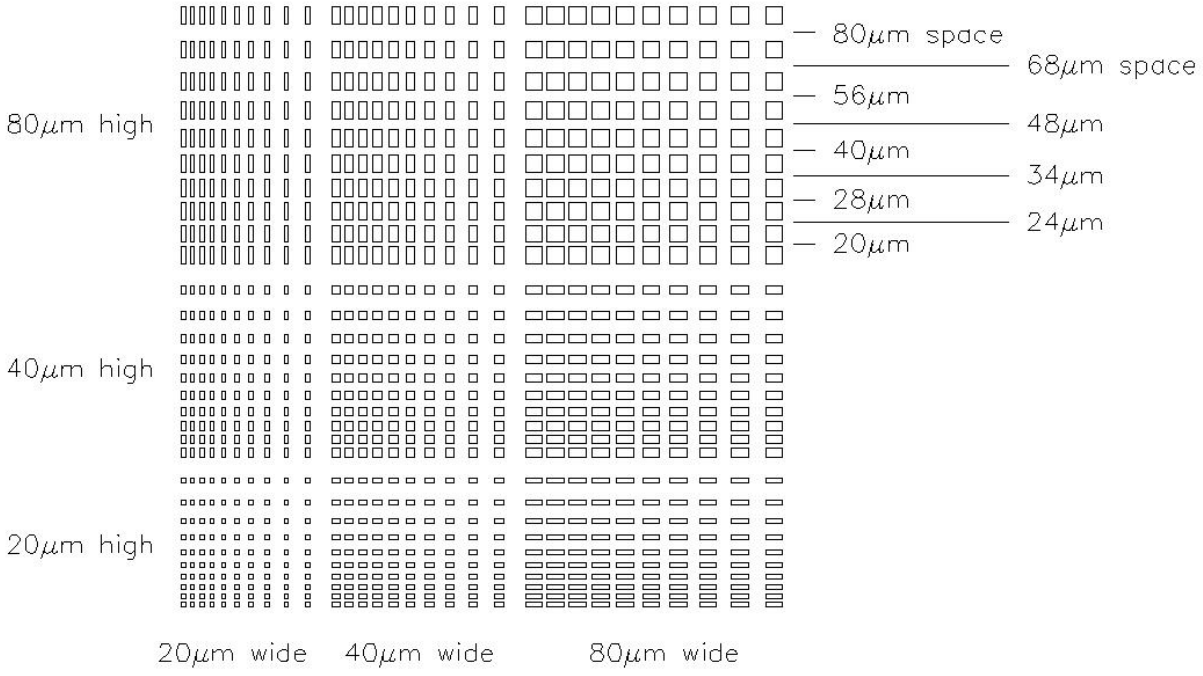


Fig. 15 Mask dimensions of test pattern pads and spaces.

We provided the cleaned substrates to two vendors with extensive EL-Ni bump plating experience (Vendor A and Vendor B). The vendors cut the flex into ten vertical ~8mm wide strips, each strip containing the five zones that had been subjected to different cleaning processes. The ends of each strip were affixed with tape to a handling wafer and plated to differing nominal buildups, ranging from approximately 5-50 µm in 5 µm increments. As part of the standard wafer metallization process, each strip received a 50 nm layer of immersion gold as an oxidation barrier after nickel plating.

Our first goal was to assess the quality of individual EL-Ni D2D connections. Samples from both vendors were hand probed with a pair of tungsten needles and demonstrated good connections between dots. Wherever optical microscopy showed an apparent connection, the resistance between dots was well under 1Ω, and wherever there appeared to be a space, the resistance was over 1 MΩ. Where bump overlap was sufficient, groups of connected bumps could be separated from the substrate with a hobby knife, separating at the Cr-polyimide interface and leaving the D2D joints intact. While not quantitative, this suggests that mechanically robust joints can be created. On the backside of the pried-off material, immersion gold had covered the nickel on the underside of the bump edges to within a few microns of the pads (Fig. 16a). It will be important to understand this phenomenon better to assure that corrosive solutions are not trapped in this region. At 500x magnification under white light, no knit line between bumps could be seen on the backside of undamaged samples. We were also able to image the backside of the bumps through the polyimide, allowing in-plane plating build-up to be measured (Fig. 16b). Though not as clear as directly viewing pried-off samples, viewing through the polyimide is non-destructive.

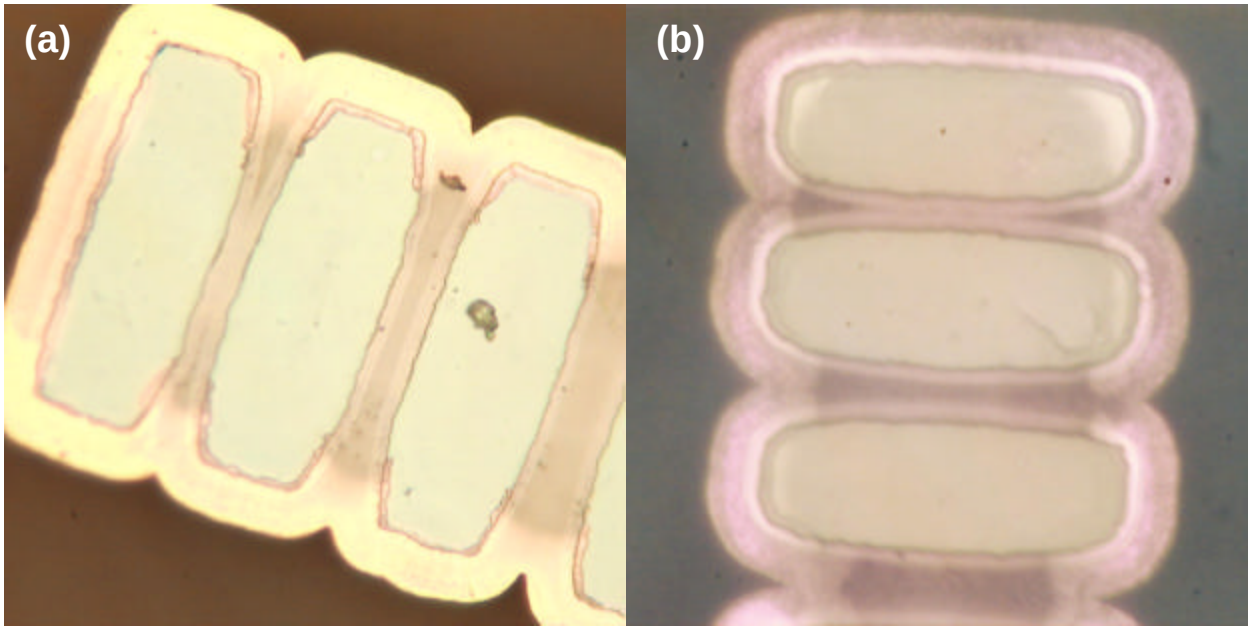


Fig. 16 Back of 20x80 µm NOM detached bumps: a) Direct view. b) Through polyimide.

Scanning electron microscopy (SEM) of intact plating revealed a phenomenon that has important implications on required D2D overlap. When the deposits on adjacent pads first make contact, the joint appears not to have truly knitted (Fig. 17a). We hypothesize that even with vigorous agitation, mass transport into and out of the forming joint is limited and the plating solution can rapidly become depleted.[3] This situation seems to persist even as the bumps grow, as shown by the joint on the left-hand side of Fig. 17b. But when the included angle between the meeting deposits reaches a sufficiently large value, the deposits appear to begin to knit as shown by the right-hand joint in Fig. 17b. Indeed one can imagine that for a given plating process there is some angle beyond which fresh plating solution can reach the knit line almost as easily as the more exposed surfaces. At this point, uniform deposition across the joint initiates and begins to form a strong bond at the knit line. From Fig. 17b and Fig. 17c, it appears that for an included angle of greater than ~90°, the deposition across the joint is sound. Fig. 17d shows a large angle joint at very high magnification. While the location of the knit line is still clear, the surface texture appears to be continuous across the joint. This evidence suggests that local joint integrity may be found to be a strong function of the included angle. Sectioning studies are a clear next step toward understanding the details of joint formation, which will be vital to successful practical application of EL-Ni D2D.

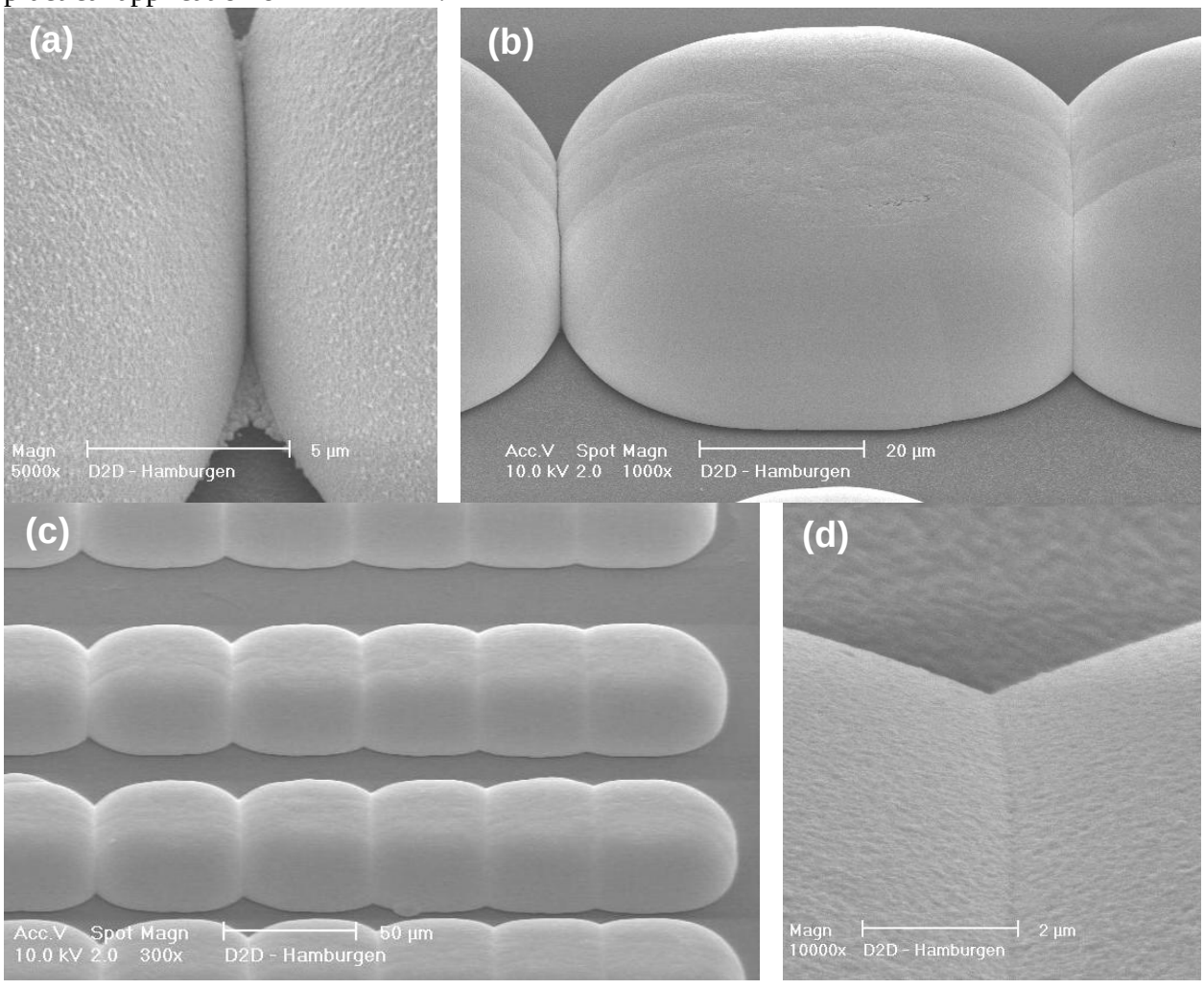


Fig. 17 EL-Ni D2D joint details, showing development of knit line.

[3] The deposits of the sort seen at the edges of the joint in Fig. 17a were also seen at other locations having a shallow included angle at the joint. They form only adjacent to the substrate, where we might expect contaminants to come to rest due to capillary action following liquid-based cleaning and drying.

The second goal of our experiments was to demonstrate the pitch required to assure isolation between adjacent columns. SEM photos of the best test patterns plated by vendor B show the minimum pitch possible without shorting (Fig. 18). The leftmost photo has a nominal 10 µm bump height, and no shorts at a 48 µm pitch. The center photo shows 15 µm high bumps and a slight gap between bumps on a 54 µm pitch. The right-hand photo shows much thicker 35 µm bumps, with the minimum pitch increasing proportionately to 76 µm.

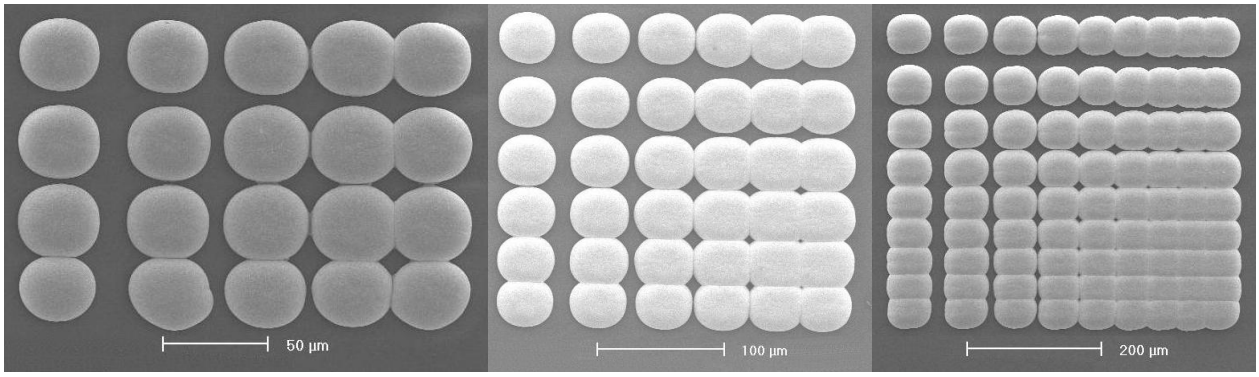


Fig. 18 Plan view of best vendor B test patterns: 10, 15 and 35 µm nominal bump heights.

For each of the ten plating thicknesses, we visually inspected vendor A's best patterns to determine the maximum pitch at which adjacent pads would short-circuit (e.g. establish a D2D connection) and the minimum pitch for which an open circuit would be assured. Fig. 19 depicts these results plotted against the bump height, as measured by the vendor. The graph includes a line showing where the boundary between the two regions would be if the pads were exactly 12 µm diameter and the plating buildup was perfectly uniform in all directions. While the slope is as expected, the data suggest that plating in this process proceeds somewhat faster in plane than normal to the plane. The graph shows that with perfect registration, all but the thickest plating could be used to create 100 µm pitch interconnections on the edge of stacks assembled from 25 µm substrates.

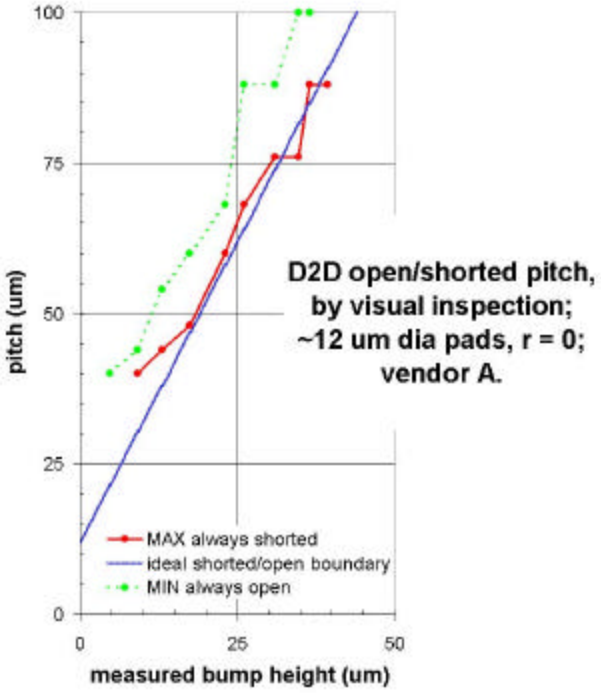


Fig. 19 D2D open/shorted pitch by visual inspection: ~12 µm diameter pads, r=0, vendor A.

The third goal for our experiments was to assess yield-related issues, and here the two vendors delivered markedly different results. The Vendor A parts showed no evidence of pads failing to activate and plate, but there were a multitude of spurious activation sites on the polyimide surface, the incidence increasing markedly with plating thickness. This suggests that something on the polyimide surface was changing during plating. The degree of spurious activation was also correlated with the gas used in plasma cleaning. Pure argon was by far the worst actor, and with the longest plating times there was wild growth that covered everything with nickel. Ar-O_2 was somewhat better than pure argon. The O_2 and O_2-CF_4 plasmas appeared to give the best results, and "no plasma treatment" fared only slightly worse.

The samples from Vendor B looked much better. Fig. 20 and Fig. 21 compare optical microscopy of the very best 20x20 µm patterns from both vendors at approximately 10, 25, and 40 µm bump heights. Note that with the thickest plating, the Vendor A sample has mis-shaped deposits and spurious activation sites and even the ~25 µm sample has several nodules that could have caused shorts had they occurred elsewhere. By contrast, the Vendor B samples show relatively uniform deposition over the entire range. The difference between vendors is amplified when the full sample set is considered rather than comparing just the best patterns. The material processed by Vendor B was largely free of defects over the entire range of plasma pre-treatments and plating thicknesses, while it was quite difficult to find any defect-free sites on Vendor A's material.

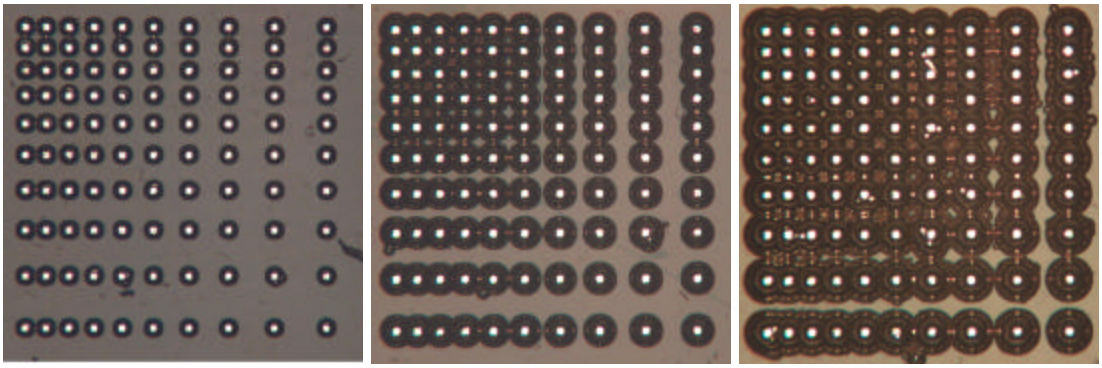


Fig. 20 Vendor A: b ˜ 10, 25, 40 µm.

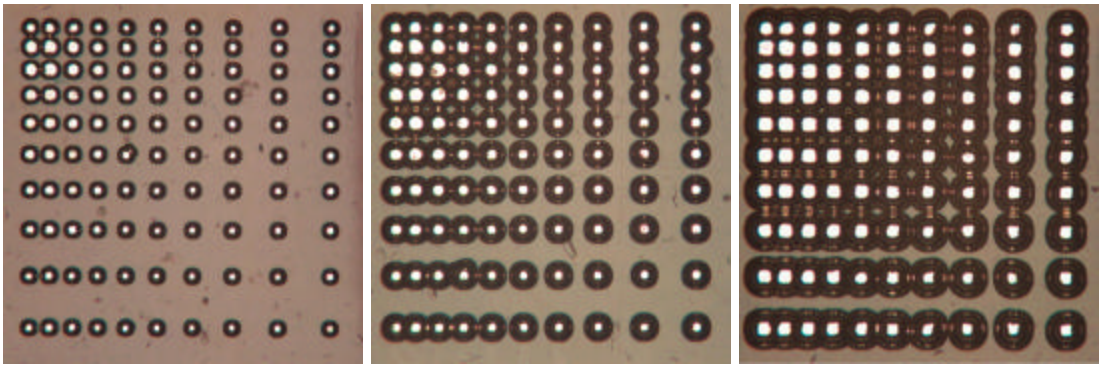


Fig. 21 Vendor B: b ˜ 10, 25, 40 µm.

Development Issues:

There are at least three materials initially present at the sawn module edge: trace metal, polyimide substrate, and the interlayer adhesive, and to this we add our EL-Ni plating. While with any wet processing there are concerns about the potential reliability impact of liquid and contaminant entrapment, particularly at interfaces between dissimilar materials, our greatest concerns relate to features specific to the EL-Ni plating process. We've already mentioned the need to understand how bumps knit together within a column when they touch and join. We must also consider the thermo-mechanical aspects of the material system and their effect upon reliability. Nickel, for example, has a different coefficient of thermal expansion than most candidate substrate materials and lamination adhesives, and polymers exhibit significant dimensional change with variations in humidity while metal does not.

The most significant long-term issues are expected to center on process yield. The EL-Ni on copper plating process uses a palladium-based chemistry to selectively activate only the metal trace ends. Discussions with suppliers and the results of our initial experiments have led to a concern regarding the density of undesired activation sites due to polymer surface roughness, inclusions or metal contamination, and perhaps at the interfaces between materials. Considerable effort may be required to develop processes that adequately clean the metal and polymer surfaces without creating spurious activated sites. Managing surface activation will be an unending task. To maintain the highest possible yields, activation issues will require substantial attention during process development and will have to be closely monitored during volume production.

Much could yet be learned from our initial batch of samples. Further tests might include detailed dimensional measurements to validate and improve our geometric and process capabilities models. Connections should be sectioned to learn how joint integrity evolves as adjacent bumps knit and form a connection. 4-wire electrical testing could also be performed on the joints to accurately assess their resistance. Automated optical inspection will be an important part of any eventual production process, and it might worthwhile to send samples to inspection equipment vendors to assess their capabilities.

There are several possible ways to proceed with D2D development. While a flex circuit is a reasonable proxy for a sawn module edge, many problems will not manifest themselves until laminated and sawn edge patterning is attempted. Major attributes of such laminated test modules will need to be selected: 1) substrate material and thickness, 2) substrate metal system, and 3) lamination adhesive system. While the most obvious substrate material is some flavor of unreinforced polyimide, D2D has stacked-chip applications that suggest considering reinforced resins such as bismaliamide-triazine and multifunctional FR-4. As the reinforcing material can be expected to affect the creation of spurious nucleation sites, a variety of both woven and non-woven materials should be considered. Metal systems should include copper as well as aluminum, chromium and perhaps even silver. Lamination adhesives could include the thin layers of the ultra-violet- cured acrylated-urethanes under consideration for their compatibility with the PIRM layer process, as well as the acrylics and epoxies generally used in multilayer rigid-flex circuits. The sawing and cleaning processes will require considerable attention. While diamond sawing is the most obvious approach, water jet and liquid nitrogen jet cutting processes might provide a surface that requires less pre-plating preparation.

Costs for EL-Ni D2D Process:

While in production we hope to be able to perform plating in a roll-to-roll process optimized to our exact requirements, vendors' wafer bumping pricing provides us with an upper bound on production costs of roughly $0.50 per memory card. A lower bound on EL-Ni plating cost can be devised by considering just the price of the chemicals. According to one supplier, in high volume the chemicals for commodity El-Ni plating cost under $1.50 per "mil- square-foot" or $0.63 per cc [17], equivalent to about $0.0005 to plate interconnect on the long edges of a memory card. A published paper quoted chemical higher costs, but demonstrated that at least one commodity-priced plating bath is suitable for wafer under-bump metallurgy and bumps [16].

While our lower bound ignores the substantial non-chemical costs of plating, it is 3 orders-of-magnitude below our service bureau sales price estimate. Assuming that in very high volumes the plating could be done at $1/10^{th}$ the service bureau price and 100x the cost of the chemicals, D2D interconnect could be plated for around $.05 per memory card. This compares favorably with proposed laser patterning approaches.

Scaling issues are also important. The time spent laser patterning is proportional to the distance the beam traverses, thus cost increases with wiring density. For D2D the plating time is reduced as wiring density increases, leading to lower costs. While other issues such as surface preparation requirement and yield obviously need to be considered, this trend toward lower cost with increasing density is a major attraction of D2D processes.

Alternate Maskless, Bulk Interconnection Process Implementations:

EL-Ni plating is not the only maskless, bulk process for creating D2D interconnects. There are numerous other alternatives that use pad geometry to define placement of wiring. These include hybrid variants of D2D and other maskless techniques that avoid the unfavorable scaling attributes of sequential (e.g. laser-based) processing. We'll consider both approaches that preserve the full two-dimensional generality of "pure D2D", as well as schemes that exploit process anisotropies to form wires preferentially in one direction. These alternatives are considered to have less broad applicability, less potential for scaling to smaller dimensions, or higher complexity, cost and risk than our baseline EL-Ni D2D proposal.

While wet processing is likely least expensive at larger dimensions, dry processing may also be possible. Reactive conductor deposition from gas phase may allow D2D to be used in wafer fabrication, particularly where it reduces the mask count in very high layer count memory devices or perhaps even stacked wafer structures.

One disadvantage of the conventional EL-Ni bump process is that it is slow, typically proceeding at only 15-20 µm/hour. It may be possible to reduce costs by doing only a partial buildup, and completing the column connections with a solder dip and air knife process similar to conventional wave soldering. The air knife would be directional, and introduce anisotropy by preferentially removing solder bridges between the columns. It is also possible to completely eliminate the need for any EL-Ni plating by using a very long plasma etch on the module sidewalls, converting the pads into posts projecting a distance h above their surroundings. When h, l_1, and s_1 are all much greater that s_2, one would expect stable fillets to form preferentially

between posts on adjacent layers, forming column connections. If the required plasma processing proved less expensive than plating, this might provide a low cost solution, albeit with a significantly larger wire pitch than D2D with EL-Ni.

In both of the above scenarios, it may be possible to replace the solder with a conductive adhesive via a process akin to the dip-transfer technique that is used for attach of flip-chip parts to circuit boards [9]. For package attach, the adhesive is uniformly coated on a platen and the package vertically dipped to transfer a controlled amount to the bumps. In our case however, the module would be dragged laterally on the platen, introducing a preferred direction for column connections to occur.

Others have demonstrated using electroplating to make connections between pairs of adjacent terminals [10], [11]. In both cases cited, one terminal of each pair was bussed together with all other such terminals. To facilitate removal after plating, the bussing structures were located remotely from where the connections were to be established. While these techniques do not work where only a single terminal for each circuit net is accessible, a somewhat different approach may make it possible to use electroplating to create D2D connections. The key is to choose a workpiece and plating system geometry that will allow fine-grained control of electric field strength and plating bath depletion. After a long sidewall etch to convert the pads to tall posts, a thin metal seed layer would be sputtered to bus together all the posts and form the cathode. The anode design and bath flow field would be designed to cause plating to preferentially build up on the posts where the fields and bath concentrations are highest, eventually forming thick bridges and defining columns between adjacent layers. A chemical bath would then be used to uniformly back-etch, removing material until the seed layer between adjacent columns disappears, isolating the columns electrically. Assuming the preferential build-up during electroplating is sufficiently pronounced, the column connections would remain intact.

While less broadly applicable than the preceding methods, it may be possible to introduce anisotropy to preferentially create column connections. Others have assembled nanowires after growth into parallel arrays by fluid flow [24]. At larger scale-lengths, either an electroless or electroplating plating solution could be directed across the surface, increasing the plating rate within column connections and reducing it between columns, creating a self-aligning effect. A similar effect could perhaps be attained in the back-etch portion of the previous electro-plating example, causing a higher removal rate in the channels between the columns than in the comparatively stagnant regions within the columns.

The operation used to expose the module edges offers still another possibility. The conventional way to expose the traces at a module edge is by sawing. Since the saw penetration is only a small fraction of its diameter, considerable care must be exercised to obtain a smooth finish, one free of metal smearing that could short adjacent columns. If however we either shear the module edge, or finish it with a secondary grinding operation oriented parallel to the columns, any metal smearing will occur in a direction that will tend to create desired connections within a column rather than causing undesired column-to-column shorts. It may be possible to exploit this phenomenon to improve any of the previously proposed connection processes.

A completely different alternative is to use resist-based processes we'll refer to collectively as *picket fence photolithography*. They typically begin with a module edge that has been sawn with a sloping sidewall of angle a. The sidewall is plasma etched until the trace pads become posts of height $h > s_2 \tan a$. A thin metal seed layer is then coated over the entire surface. This may prove challenging, as the coverage must include the hard-to-reach region below the base of each

post. A thick layer of negative photo resist is deposited and baked; electrophoretic resists may be useful to obtain uniform coverage. The resist is exposed from above with a collimated light source, fully dosing the spaces between the columns. However, since $h > s_2 \tan a$, each post shadows most of the post below, as well as the spaces between the posts within a given column. When the resist is developed and undeveloped material removed, the seed metal layer is left exposed between the posts in a column. It can then be electroplated up to the thickness of the resist layer, the resist stripped and the seed layer removed. Many variations of picket fence photolithography are possible, including those using positive photoresists or mass CO_2 laser ablation of polymer between columns.

Others have investigated the self-assembly of superparamagnetic colloidal particles onto a patterned ferromagnetic film, and shown how an external magnetic field can be used to connect adjacent pads in a preferred direction within an array [25]. While not directly applicable to our module edge interconnect requirements, it illustrates yet another physical phenomenon which might be useful for patterning D2D interconnects.

D2D connection geometries may be relevant to molecular-scale electronics, where a variety of related approaches have been proposed [23]. With considerable arm waving, we propose the following illustrative scenario. Assume a surface containing molecular-scale electronic elements, with the terminals of the elements consisting of individual dangling bonds or clusters that comprise activated sites. These sites might be the result of the molecular-scale self-assembly process that defined the underlying electronics, or perhaps they were created by surface interaction with a probe tip. A buildup of a conductive complex occurs at each active site, to a depth controlled either by the reaction duration, or perhaps by a self-terminating reaction. For a multi-layer structure, the surface would be passivated, sites activated as required, and the build-up process repeated.

Conclusions:

We have described the D2D process, and shown how it might be used to create inexpensive single and perhaps multi-layer interconnects over a wide range of scale lengths and potential applications. Our analysis showed that in our targeted substrate stacking application, substrate thickness sets a fundamental limit on wiring density, and layer-to-layer registration is likely to provide another important limit. Our demonstration of 50 µm pitch columnar interconnect on a planar substrate having perfect inter-row alignment suggests that a <200 µm pitch will be possible with improved single-axis alignment. Analytical models predict that 600 µm pitch on all sides of a module is likely possible with the tool set used for fabricating standard multi-layer PCBs. We hope that this work marks the starting point for development of a D2D interconnect technology that will greatly reduce the cost of producing of our very smallest electronic systems.

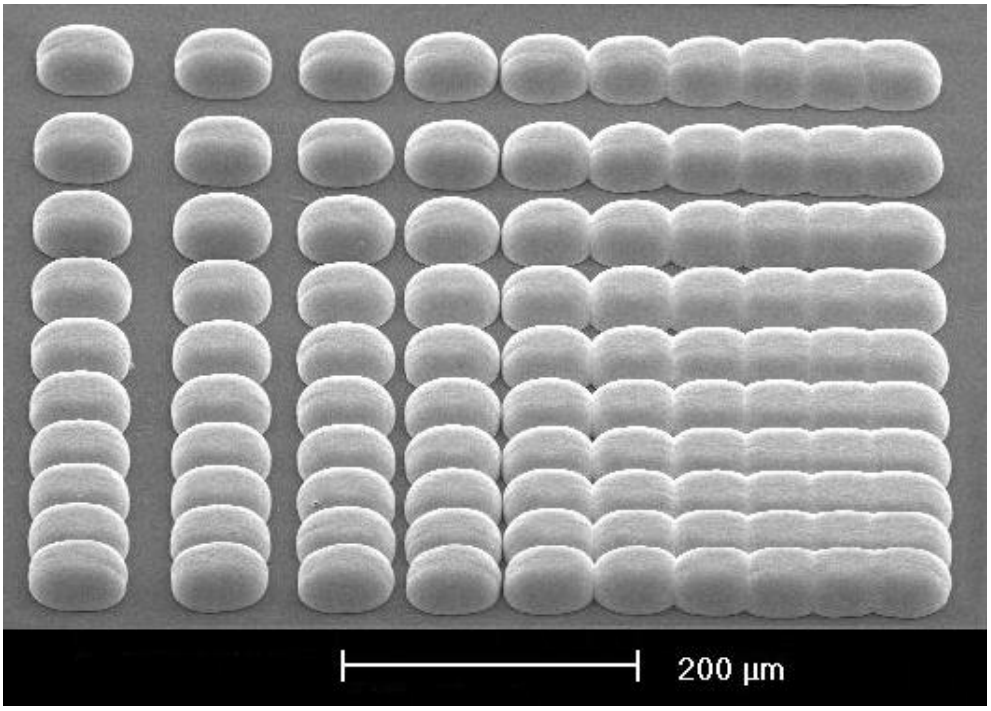


Fig. 22 Oblique view of 35 µm thick EL-Ni plating on ~12 µm pads forming D2D interconnect.

Acknowledgements: I'd like to thank my manager Carl Taussig for his patience and support as I groped my way along, Albert Jeans for his SEM artistry, and all the other members of the PIRM team for the many ideas and suggestions they contributed. Thanks also to Thorsten Teusch of Pac Tech USA, Andrew Strandjord of IC Interconnects and Bob Strolle of Microconnex for generously sharing their time and expertise.

References:

[1] S. F. Al-sarawi, D. Abbott, P. D. Franzon, "A Review of 3-D Packaging Technology," *IEEE Trans. CPMT – Part B*, vol. 21, no. 1, Feb 1998.

[2] N. Tanaka, Y. Yamaji, T. Sato, K. Takahashi, "Guidelines for Structural and Material-System Design of a Highly Reliable 3-D Die-stacked Module with Copper Through-vias," in *Proc. 53rd IEEE ECTC*, New Orleans, May 27-30, 2003, pp. 597-602.

[3] http://www.toshiba.co.jp/about/press/2004_01/pr2101.htm

[4] M. Sunohara, K. Murayama, M. Higashi, M. Shimizu, "Development of Interconnect Technologies for Embedded Organic Packages," in *Proc. 53rd IEEE ECTC*, New Orleans, May 27-30, 2003, pp. 1484-1489.

[5] H. Kwon, T. Cho, K. Choi, J. Shim, S. Kim, Y. Lim, D. Shin, U. Shin, S. Oh, "SIP Solution for High-End Multimedia Cellular Phone," in *Proc. 36th IMAPS Intl Symp on Microelectronics*, Boston, Nov 18-20, 2003, pp 223-228.

[6] C. Val, "New Approaches to 3-D Interconnection Systems in Package Applications," in *Proc. 36th IMAPS Intl Symp on Microelectronics*, Boston, Nov 18-20, 2003, pp 229-234.

[7] A. Matsumura, K. Uwada, Y. Hotta, "Development of Low-K Encapsulating Film for Stacked Packages," in *Proc. 36th IMAPS Intl Symp on Microelectronics*, Boston, Nov 18-20, 2003, pp 189-193.

[8] R. Fillion, R. Wojnarowski, C. Kapusta, R. Saia, K. Kwiatkowski, J. Lyke, "Three-D Vertically Stacked Electronic Structures for Array Sensors," in *Proc. 36th IMAPS Intl Symp on Microelectronics*, Boston, Nov 18-20, 2003, pp 218-222.
[9] J. Clayton, "Very High Pin Count Flip Chip Assembly Using Conductive Polymer Adhesives," in *Proc. 36th IMAPS Intl Symp on Microelectronics*, Boston, Nov 18-20, 2003, pp 275-281.
[10] C. Lin, S. Chiang, T. Yang, "Bumpless Flip Chip Packages for Cost/Performance Driven Devices," in *Proc. 53rd IEEE ECTC, New Orleans*, May 27-30, 2003, pp. 554-559.
[11] Y. Joung, M. Allen, "High Density Electroplating Bonding Interconnect Technology: Chip Packaging and High Aspect Ratio Passive Elements," in *Proc. 53rd IEEE ECTC*, New Orleans, May 27-30, 2003, pp. 640-646.
[12] T. Sugiyama, Y. Yano, S. Ishihara, T. Maruyama, H. Juso, T. Kimura, M. Kada, "Board Level Reliability of Three-Dimensional Systems in Package (SIPs)," in *Proc. 53rd IEEE ECTC*, New Orleans, May 27-30, 2003, pp. 1106-1111.
[13] J. Balde, ed, *Foldable Flex and Thinned Silicon Chips for Multichip Packaging*, Kluwer Dec 2002, ISBN 0-7923-7676-5.
[14] "Move over, silicon," *The Economist Technology Quarterly*, Dec 14, 2002, p 21.
[15] S. Möller, C. Perlov, W. Jackson, C. Taussig, S. Forrest, "A Polymer/Semiconductor Write-Once Read-Many-Times Memory," *Nature*, vol. 426, Nov 13, 2003, pp. 166-169.
[16] H. Lu, C. Kang, S. Wong, H. Gong, "Evaluation of Commercial Electroless Nickel Chemicals for a Low Cost Wafer Bumping Process," Semiconductor Science & Tech., vol. 17, no. 9, Sept 2002, pp. 911-917.
[17] M. Feldman, Surface Technology, Inc (STI), private communication, Feb 2004.
[18] T. Teutsch, T. Oppert, E. Zakel, D. Tovar, "A Low Cost Bumping Process for 300 mm Wafers," in *Proc. 32nd IMAPS Intl Symp on Microelectronics*, Chicago, Oct 26-28, 1999.
[19] A. Strandjord, S. Popelar, C. Erickson, "Electroless Nickel/Gold Plating on Copper Based Semiconductors," in *Proc. 33rd IMAPS Intl Symp on Microelectronics*, Boston, Sept 20-22, 2000, pp 196-201.
[20] J. Song, J. Yu, "Phase Transformation and Residual Stress Evolution in Electroless Ni-P UBM in Low Cost Flip Chip Technology," in *Proc. 35th IMAPS Intl Symp on Microelectronics*, Denver, Sept 4-6, 2002, p 261-266.
[21] C. Taussig, "Thin Film Electronics on Polymer Substrate Fabricated Using Roll-to-Roll Plasma Processing and Self Aligned Imprint Lithography," *USDC 3rd Annual Flexible Displays and Microelectronics Conf.*, Phoenix, February 10-12, 2004.
[22] M. Roesch, Hewlett-Packard, private communication, Feb 2004.
[23] M. Islam, S. Sharma, T. Kamins, R. Williams, "Ultrahigh-Density Silicon Nanobridges Formed Between Two Vertical Silicon Surfaces," *Nanotechnology*, vol.15, 2004, pp. L5-L8.
[24] Y. Huang, X. Duan, Q. Wei, C. Lieber, "Directed Assembly of One Dimensional Nanostructures into Functional Networks," *Science*, vol. 291, 2001, pp. 630-633.
[25] B. Yellen, G. Friedman, "Printing Superparamagnitic Colloidal Particle Arrays on Patterned Magnetic Film," *J Applied Physics*, vol. 93, no. 10, May 15, 2003, pp. 7331-7333.
[26] D. Numakura, "Manufacturing of the Ultra High Density Flex Circuits," *IPC 2nd Intl Conf on Flexible Circuits, Chip Scale and Bare Die Packaging*, Tempe, Arizona, Jan 6-8, 2004.
[27] H. Goldstein, "Packages Go Vertical," *IEEE Spectrum*, Aug 2001, pp. 46-51.
[28] J. Minahan, A. Pepe, R. Some, M. Suer, "The 3-D Stack in Short Form," in *Proc. 42nd IEEE ECTC*, San Diego, May 18-20,1992, pp. 340-344.
[29] W. Hamburgen, "Dot-to-Dot Maskless Bulk Interconnect Process for Stacked Chip Packaging," in *Proc. 37th IMAPS Intl Symp on Microelectronics*, Long Beach, CA, Nov 14-18, 2004.