IBM Research Report

A CMOS-compatible Process for Fabricating Electrical Through-vias in Silicon


IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

*IBM Systems and Technology Group
1000 River Street
Essex Junction, VT 05452
A CMOS-compatible Process for Fabricating Electrical Through-vias in Silicon

P.S. Andry, C. Tsang, E. Sprogis*, C. Patel, S.L. Wright, B.C. Webb,
L.P. Buchwalter, D. Manzer, R. Horton, R. Polastre and J. Knickerbocker
IBM T. J. Watson Research Center
1101 Kitchawan Road, Yorktown Heights, New York 10598
* IBM Systems and Technology Group
1000 River Street, Essex Junction, Vermont 05452
Email: andry@us.ibm.com, Tel: 914-945-3069

Abstract
In the past, traditional CMOS scaling has been one of the principal levers to achieve increased system-level performance. Today, scaling is becoming increasingly difficult and less effective, and a range of new two- and three-dimensional silicon integration technologies are needed to support next-generation systems. A silicon-carrier System-on-Package (SOP) is an advanced packaging solution, enabling interconnection between ICs and other devices at densities far beyond those of current first-level packaging. Silicon-carrier employs fine pitch Cu damascene wiring, high-density solder pads/joins and high-yielding electrical through-vias. A novel approach to fabricating robust through-vias in silicon is described. The key design feature enabling large-area, uniform arrays to be produced with high yield is the annular via shape. As compared to a standard cylindrical via shape, the annular via is easier to integrate into a standard CMOS copper back-end-of-the-line (BEOL) process flow. Two process flows are compared: the first having the conductor metal within the gap of the insulated annulus itself, the second having a conducting metal core enclosed within the inner wall of the annulus. For the first process flow, two annular conductors, plated copper and CVD tungsten, are compared in terms of ease of integration, yield and susceptibility to failure during thermal stressing. Large area (45 x 48 mm) silicon carrier modules containing more than 51,000 electrically measurable through-vias are used to compare overall yield and robustness of each process. Results on deep thermal cycling, current carrying capacity and thermomechanical modeling are discussed. Wafer-level via testing is used to statistically distinguish between via chain opens caused by bond and assembly issues versus failures in the vias or integrated wiring structures. Through-via resistances on the order of ~10 mΩ are typical, and through-via yields of 99.98% at module level have been demonstrated.

Motivation for Through-via Development
Development of manufacturable and robust silicon through-via technology has been widely pursued in recent years [1,2,3]. Myriad structures, processes and integration schemes are possible, as are the possible applications which include 3D wafer-to-wafer interconnects [4,5], chip stacks [6], rf & MEMS passive integration [7] and silicon-based System on Package (SOP) [8] among others.

In general, fabrication of through-vias comprises the steps of via etching, sidewall insulation, metallization, wafer thinning and connection to terminals or wiring on adjacent faces of the silicon wafer. Each of these steps comes with its own challenges depending on the geometry of the vias, the materials and/or processes used to insulate and metallize of the vias. Previously, two generic approaches to through-via fabrication sequences were described: one a “vias first” approach and the other a “vias last” approach [8]. The “vias first” scheme describes a sequence where the through-vias are etched, insulated and metallized before the fabrication of BEOL wiring levels; the “vias last” approach describes a sequence where the BEOL wiring levels are built first, and the through-vias are fabricated afterwards. In prior work, the via geometry was cylindrical. Described below are results of IBM Research activity on annular via structures and process flows which provide improved integration and enhanced yield.

Annular Through-Via Structure and Processing
The salient design feature allowing for high-yield through-via fabrication is the annular shape. The narrow annulus can be filled in a number of ways, and requires much less conductor cross-section than a standard cylindrical shape. Two distinct process flows are possible depending on whether the conductor is placed within the insulated annulus itself, or within the core region defined by the inner wall of the annulus. The annular conductor process flow is illustrated in Fig. 1. Following etch of the annular shape, thermally grown oxide is used to insulate the via, after which conductor is deposited within the insulated feature with the goal of complete fill within the gap. In this work we have chosen to compare two metals (electroplated Cu and CVD tungsten) but other metallization options are also possible. Wiring levels are built on top, and the wafer is attached to a handler, thinned and back side processed to expose the vias, after which UBM pads and solder C4s are added. The annular conductor via structure of Fig. 1 has several attractive features including a CTE-matched silicon core for improved mechanical reliability, and a reduction of process steps compared to the cylindrical hybrid plated Cu/paste approach [3] and a simplified overall BEOL integration scheme.

The second process enabled by the annular design via is the core conductor process illustrated in Fig. 2. Here, the annulus is thermally oxidized as before, but then filled with undoped poly-Si and polished. This process flow has the advantage that the wafer remains FEOL compatible after annular fill, meaning that FEOL devices or passives can be fabricated before the wiring levels are built. In this way, the process is comparable to the “vias last” approach described in earlier work [8]. Once the wiring levels are built, the wafer is
attached to a handler, thinned, and back side patterned. The silicon core is then etched away and replaced with a partial skin of electroplated copper. Complete filling of the entire via core was not attempted in this work. UBM pads and solder C4s are added in the final step as before.

Si carrier test vehicle: design and features

The primary test vehicle used in this investigation was designed specifically for rapid through-via yield evaluation of a very large array of evenly spaced vias. Shown in Fig. 3a, a 45 mm x 48 mm silicon carrier and a matching 49 mm x 50 mm silicon substrate were designed, with each having 2 micron thick alternating Cu wiring links to form complete chains at module assembly. The carrier contained the through-vias integrated to wiring links on the top side and evaporated UBM/C4s on the bottom side (Fig. 3b), whereas the substrate contained the matching array wiring links, and a large collection of 4 pt. probe taps and wiring traces (Fig. 3c) leading to vertically-aligned sets of 2x12 probe-card pads on left and right edges. The carrier employed Cr/CrCu/Cu/Au UBM pads and eutectic solder whereas the TSM of the silicon substrates were Cu finished with ~1 micron of electroless Ni and 100 nm of immersion Au.

Figure 1. Annular conductor process flow.

Figure 2. Annular via core conductor process flow.

The through-via array on the carrier is a simple 4 x 6 rectangular arrangement of sub-arrays, where each sub-array contains 60 rows x 36 columns of 4-on-8 mil spaced vias/pads. Thus, each sub-array contains 2,160 through-vias and the entire assembled module totals 51,840. Except for the sub-arrays at the very corners of the carriers, each sub-array in the completed module contains ten electrically measurable chains having 204 links each. Additional single via 4 pt. test sites are dispersed throughout the test vehicle, and two 144-link chains as well as a number of shorter chains ranging from 4 to 20 links (manually probed) are located in the very corners of the module for calibration and debugging. Automated 4 pt. step & repeat testing of the 420 probe-cad pads along each edge of the substrate allows for a total 232 of long chains (204 links), 96 single via/C4 sites and multiple substrate
wiring calibration structures to be probed in about 30 minutes, including set-up and alignment time.

![Figure 3](image)

Figure 3. a) 45 x 48 mm silicon carrier through-via array design (green) superimposed on custom silicon substrate (pink), b) silicon carrier wiring links and C4s, c) matching silicon substrate wiring links.

Module-level yield evaluation

Following dicing, silicon carriers, complete with their laminated glass backing were joined to the silicon substrates using a precision flip-chip bonder. The bonder has height control and allowed for any desired compression of the molten C4s during bonding. Certain modules were also joined on production equipment using standard flux, belt furnace reflow and post-reflow flux cleaning. Parts joined using standard tooling were found to have comparable via chain yield and resistance.

Via chain results for four different through-via types are shown in Table I. Chain resistance, standard deviation and yield are shown for two modules of each type. A total of 232 long chains (204 links) are probed for a total of 47,328 total vias per module. The first two entries in Table I are control modules having BEOL wiring and C4s (no deep silicon through-vias). For these, the deep via etch level was skipped during wafer processing, but the BEOL wiring link and contact via levels were fabricated as usual. Following glass lamination and wafer thinning, the remainder of silicon was completely removed by TMAH wet etching, exposing the BEOL contact vias. Pads and C4s were added directly on top of the contacts. Thus, these carriers are used to find a baseline for chain resistance and yield determined only by the wiring links, the C4s and the effectiveness of the joining process.

The following three pairs of entries show via chain results for two variations of the annular metal process as well as the copper core process described above. The first variation of the annular process used CVD tungsten as the metal filling process, and a via of core dimension 50 microns with an annular width of 4 microns. The second variation used standard damascene copper plating in vias of core dimension 50 microns with an annular width of 8 microns. The carriers fabricated using the copper core process had through-vias of core dimension 30 microns with an annular width of 2 microns. In this case, the small annular width makes them quite easy to fill using only a combination of thermal oxide and poly-Si fill. The resultant core via plated from the wafer backside is 28 microns in diameter, and the partial sidewall copper plating ranged from about 5 to 8 microns. All Si carriers were ~60 microns thick.

Table 1. Mean resistance, standard deviation and yield of 204-link via chains as a function via type for eight silicon carrier modules. Link yield is a maximum value assuming open chains contain a single open link.

<table>
<thead>
<tr>
<th>module</th>
<th>via type</th>
<th>$R_{chain}$ (Ω)</th>
<th>$\sigma_{chain}$ (Ω)</th>
<th>chain yield (%)</th>
<th>link yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>none: C4 only</td>
<td>51.6</td>
<td>1.9</td>
<td>93.5</td>
<td>99.97</td>
</tr>
<tr>
<td>2</td>
<td>none: C4 only</td>
<td>50.7</td>
<td>0.9</td>
<td>95.3</td>
<td>99.98</td>
</tr>
<tr>
<td>3</td>
<td>50/4 tungsten</td>
<td>53.9</td>
<td>2.0</td>
<td>94.8</td>
<td>99.97</td>
</tr>
<tr>
<td>4</td>
<td>50/4 tungsten</td>
<td>54.4</td>
<td>2.7</td>
<td>95.3</td>
<td>99.98</td>
</tr>
<tr>
<td>5</td>
<td>50/8 copper</td>
<td>48.4</td>
<td>1.8</td>
<td>95.3</td>
<td>99.98</td>
</tr>
<tr>
<td>6</td>
<td>50/8 copper</td>
<td>49.4</td>
<td>4.9</td>
<td>94.0</td>
<td>99.97</td>
</tr>
<tr>
<td>7</td>
<td>copper core</td>
<td>65.1</td>
<td>28.7</td>
<td>64.2</td>
<td>99.78</td>
</tr>
<tr>
<td>8</td>
<td>copper core</td>
<td>67.0</td>
<td>27.4</td>
<td>62.1</td>
<td>99.77</td>
</tr>
</tbody>
</table>

Note that the wiring links and C4s alone account for about 50 ohms of resistance, they have a fairly tight distribution around the mean (sigma ~2% to ~4%) and they produce modules having chain yield of ~95%. In the case of module 2, only 11 of the 232 chains were found to be open following bonding. Since every one of all 204 joins in a link must be good in order for the chain to function, this level of chain yield suggests that the upper limit on link yield (for uncorrelated defects) may be approximated by assuming chain yield = (link yield)$^{204}$. For example, for a random distribution of defects, a link yield of 99.9% (one bad C4 per
would give a chain yield of ~81.5% for this test vehicle. Admittedly, the link yield shown in Table I does not account for highly correlated defects such as lithography defects which may be localized to a very small region and may result in multiple closely spaced defects in a single chain. However, defects of this type may generally be well controlled in manufacturing environments.

All four annular via chain modules in Table I have chain yields which are equivalent to the two control modules, meaning that the addition of either tungsten or Cu through-vias does not detract from the overall yield for the as-bonded modules. The effect of the tungsten vias appears to be an net increase of ~4 ohms in average chain resistance. The fact that the annular Cu via modules have average chain resistances which are slightly less than the baseline can only mean that the damascene wiring links were slightly thicker than the 2 micron target on the source carrier wafers. Notably different from the rest are two the copper core modules. Chain yields starting in the low 60% range indicate a significant number open vias, statistically on the order of one out ~460. Furthermore, the average chain resistance is ~15 ohms higher than the baseline modules indicating that the vias themselves contribute significantly to the total resistance. Finally, the standard deviation is chain resistance is an order of magnitude larger than either the baseline modules, or the annular metal modules, indicating a large intrinsic variability in the vias themselves or perhaps in their integration scheme.

Table 2 shows a collection of single via data collected for the same set of 8 modules discussed in Table I. In this case, the sample of 4 pt. probe data was collected from 28 sites throughout the interior region of the module. Once again, the first two modules have no through-vias, so the resistance is due only to the C4 solder bump and the associated BLM and TSM Cu/Ni pads. An average value of ~ 5 milliohms is in line with values reported elsewhere [9]. The last column at the right of Table 2 entitled “chain adder” multiplies the mean single via value by 204 to predict the total resistive adder due to C4s and vias i.e. the predicted component of each long yield chain due to the vias and C4s.

<table>
<thead>
<tr>
<th>module</th>
<th>via type</th>
<th>$R_{\text{via}+\text{C4}}$ (mΩ)</th>
<th>$\sigma_{\text{via}+\text{C4}}$ (mΩ)</th>
<th>chain adder (Ω)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None: C4 only</td>
<td>4.7</td>
<td>4.6</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>None: C4 only</td>
<td>4.6</td>
<td>3.2</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>tungsten 50/4</td>
<td>21.6</td>
<td>4.9</td>
<td>4.4</td>
</tr>
<tr>
<td>4</td>
<td>tungsten 50/4</td>
<td>27.5</td>
<td>6.3</td>
<td>5.6</td>
</tr>
<tr>
<td>5</td>
<td>copper 50/8</td>
<td>5.7</td>
<td>3.3</td>
<td>1.2</td>
</tr>
<tr>
<td>6</td>
<td>copper 50/8</td>
<td>6.2</td>
<td>4.1</td>
<td>1.3</td>
</tr>
<tr>
<td>7</td>
<td>copper core</td>
<td>28.8</td>
<td>70.1</td>
<td>5.9</td>
</tr>
<tr>
<td>8</td>
<td>copper core</td>
<td>52.2</td>
<td>111.5</td>
<td>10.6</td>
</tr>
</tbody>
</table>

Table 2. Mean resistance and standard deviation of single via test sites as a function via type for eight silicon carrier modules as joined. Chain adder denotes the value obtained by multiplying the mean single via value by 204 i.e. the predicted component of each long yield chain due to the vias and C4s.

SEM cross-sections of typical through-vias corresponding to modules containing the three types of vias tested are shown in Fig. 4 a) through c). The scales are not exact, but are roughly equivalent. The copper core via is partially plated by design, since complete filling of the 28 micron diameter core would be impractical from a manufacturing and thermomechanical perspective. In this case, about 10 microns of copper was plated, however, even that amount required a very lengthy backside polish to complete the via; for damascene plating a thickness of ~5 microns might be considered a practical upper limit. Selective plating techniques could be employed to reduce polish time, but complete core fill at these dimensions would not be stable under thermal load due to large CTE-mismatch between copper and silicon [3].
chains having opens after 760 and 906 deep thermal cycles, respectively. In contrast, the copper core via chain module shows a higher level of chain opens with thermal cycling. Both copper core and annular tungsten modules show an increase in average chain resistance with thermal cycling while the annular copper and control modules exhibit virtually no change.

Figure 4. SEM cross-sections of a) 50/4 annular tungsten, b) 50/8 annular copper and c) copper core.

In order to verify the yield statistics extracted from the module data, and to better understand the behavior of the through-via structures under stress, but independently of the C4s and wiring BEOL links, a series of wafer level measurements were also performed.

**Wafer-level yield and reliability**

A special set of wafers were fabricated according to each process flow specifically for wafer-level evaluation during the standard test vehicle build. In the case of the annular conductor process, two wafers were fabricated, one having CVD tungsten fill, the other having electroplated copper fill. Via diameters were 50 microns, and annular width was 4 microns for both metallizations. Following the annular metal fill step, both wafers skipped the usual CMP overburden removal step, and were instead passivated using a thin PECVD SiN layer, laminated to handlers, thinned and backside processed up to the standard Cr/Cu/Au BLM pad deposition step (no C4s were deposited however). In this way, the wiring links were replaced with solid planes of metal connecting every via to all others, allowing a number of 2 pt. and 4 pt. electrical measurements to be performed.

Similarly, a third wafer was fabricated according to the copper core process, but having a shorting plane of electroplated Cu connecting all the through-vias together on one side. In this process, however, BEOL dielectric deposition and contact vias first had to be built on the top side of the wafer, since these contacts form the electrical connection once the core of the annulus is etched through from the thinned wafer back side. This is an important difference between the two via structures, and one which appears to have a strong bearing on performance of the via.

Following fabrication of the wafer-level test vehicles, two kinds of electrical measurements were performed. Four point probe measurements using combinations of 2, 3 and 4 through-vias allowed the extraction of true single through-via resistance free of any contributions due to C4s or wiring link structures. Wafer-level yield assessment, however, required a much greater number of via pair continuity measurements (i.e. tens of thousands) than a simple 2 pt. step-and-repeat system could provide in a reasonable amount of time. For these statistics, a high-speed opens/shorts (O/S) tester [10] used regularly in the IBM packaging facility was employed. The O/S tester, used to test ceramic substrate nets, employs two
rapidly actuated probes and a custom feedback and control system to achieve up to 60 measurements per second. To minimize the possibility of measuring correlated defects which could skew the statistics, the probes were flown a constant distance of 30 pads apart within each 36 x 60 sub-array of vias. A common location falling about midway between wafer edge and center was chosen, and the same 3 x 4 grid of sub-arrays was covered for each wafer. In this way, nearly 26,000 vias were measured per wafer.

Table 3. Single-via average resistance, standard deviation and yield for three types of via structures measured at wafer-level. Resistance was measured using 4 pt. probes, while yield was measured using high-speed flying probes in 2 pt. mode.

<table>
<thead>
<tr>
<th>via type</th>
<th>$R_{\text{via}}$ (mΩ)</th>
<th>$\sigma_{\text{via}}$ (mΩ)</th>
<th>via yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50/4 tungsten</td>
<td>16.3</td>
<td>0.8</td>
<td>100.00</td>
</tr>
<tr>
<td>50/4 copper</td>
<td>9.3</td>
<td>0.4</td>
<td>99.62</td>
</tr>
<tr>
<td>copper core</td>
<td>69.1</td>
<td>11.7</td>
<td>99.81</td>
</tr>
</tbody>
</table>

Figure 6. Strapped via pair IV curves for three via types identified in Table 3.

The results of the wafer-level evaluation are shown in Table 3. Note that the single via resistance values are in line with expectations, although the value measured for the annular copper process is slightly lower than expected based on the module data. In this case it was discovered that 63 out of a total of 98 opens were found to be tightly clustered within one of the 12 sub-arrays. Discarding this one sub-array would have resulted in a yield of 99.85%. Note also that not a single through-via open was detected on the wafer fabricated using the annular tungsten process.

The wafer-level test vehicles proved to be more useful than the assembled modules in assessing the current-carrying properties of the vias, primarily because the wiring links (10 microns wide and 2 microns deep) were susceptible to fuse-type failure at currents between 0.8 and 1.0 amps. Fig. 6 shows I-V curves for all three types of vias measured from one via to any nearest neighbor. In this way, the slopes of the curves represent the series resistance of two vias plus the bulk metal shorting “strap” between them. The latter contribution was generally much smaller than the via resistance for any given wafer. A full assessment of the upper current limits of these structures is the subject of future work.

Figure 7. Strapped via pair resistance distributions before (left) and after (right) deep thermal cycling for a) annular tungsten vias, b) annular copper vias, and c) copper core vias. Nearly 13,000 via pairs were probed in identical locations on each wafer.

As with the assembled modules, wafers were also subjected to deep thermal cycling to gauge the robustness of the via processes. Besides detecting fully open through-vias, the 2 pt. via pair resistance data before and after thermal cycling were binned to reveal the characteristic distributions shown in Fig. 7. Starting distributions are shown to the left
while results for > 1000 thermal cycles are shown to right. The annular tungsten vias (Fig. 7a) had zero opens after 1226 total cycles, and a very sharp distribution. The annular copper via resistance distribution (Fig. 7b) started out slightly wider than that of the tungsten vias, and broadened somewhat as a result of 1226 thermal cycles. A net increase of 8 electrical full opens was measured. For copper core vias (Fig. 7c), the initial distribution was broader than for the others, but after 1038 thermal cycles it was significantly transformed to a low, wide shape with a tail stretching out to 20 ohms. Further, a total of 165 initially continuous vias were converted to full electrical opens. Thus, there appears to be a clear correlation between module and wafer-level results for the copper core process.

The data thus far demonstrate that copper annular vias have a significant advantage in reliability during deep thermal cycling as compared to copper core vias. The results are supported by thermomechanical models of the two structures shown in Fig. 8. The models confirm that a copper annular via induces lower mechanical stress in the surrounding silicon than a copper core via when both are subjected to the same uniaxial stress.

Further, the results show that tungsten vias (see Fig. 9) can be fabricated with near perfect yield and reliability during deep thermal cycling. Research into the variability of the tungsten results at module level are currently underway.

Conclusions

A novel approach to fabricating robust CMOS-compatible though-vias in silicon has been described. The key enabling design feature is the annular via shape which is easier to insulate and metallize than comparable deep cylindrical vias, and in turn, simpler to integrate into a standard copper BEOL process flow. Two process flows were compared in detail: the first having the conductor metal within the gap of the insulated annulus itself, the second having a conducting metal core enclosed within the inner wall of the annulus. For the first process flow, two annular conductors, electroplated copper and CVD tungsten, were compared in terms of ease of integration, yield and susceptibility to failure during thermal stressing. Characterization at module level using long via chains, and at wafer level using high-speed via-to-via probing, indicate that through-via yields approaching 100% are achievable with both annular metals. For the dimensions chosen in this study, annular copper via resistance is typically less than 10 mΩ; tungsten vias are about a factor of 2 higher. Careful control of the metal fill process is required in order to achieve low resistance and stability during thermal cycling. Silicon through-via processing and integration remains an active area of study at IBM Research.

Acknowledgments

This work has been partially supported by DARPA under the PERCS program, Agreement NBCH30390004 and by Maryland Procurement Office (MPO) – Contract H98230-04-C-0920.

We would like to acknowledge the support of the MRL and CSS groups at IBM T.J. Watson Research Center, the development and support teams in Burlington, Vt. and East Fishkill, NY, and the IBM analytical group in East Fishkill, NY.

We would also like to acknowledge the support of F. Pompeo, T. Chainer, D. Seeger and T.C. Chen.

References


