

Chapter 2

3D Process Technology Considerations

Albert M. Young and Steven J. Koester

Abstract Both form-factor and performance-scaling trends are driving the need for 3D integration, which is now seeing rapid commercialization. While overall process integration schemes are not yet standardized across the industry, it is now important for 3D circuit designers to understand the process trends and tradeoffs that underlie 3D technology. In this chapter, we outline the basic process considerations that designers need to be aware of: strata orientation, inter-strata alignment, bonding-interface design, TSV dimensions, and integration with CMOS processing. These considerations all have direct implications on design and will be important in both the selection of 3D processes and the optimization of circuits within a given 3D process.

2.1 Introduction

Both form-factor and performance-scaling trends are driving the need for 3D integration, which is now seeing rapid commercialization. While overall process integration schemes are not yet standardized across the industry, nearly all processes feature key elements such as vertical through-silicon interconnect, aligned bonding, and wafer thinning with backside processing. In this chapter we hope to give designers a better feel of the process trends that are affecting the evolution of 3D integration and their impact on design.

The last few decades have seen an astonishing increase in the functionality of computational systems. This capability has been driven by the scaling of semiconductor devices, from fractions of millimeters in the 1960s to the tens of nanometers in present-day technologies. This scaling has enabled the number of transistors on a single chip to correspondingly grow at a geometric rate, doubling roughly every

A.M. Young (✉)

IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA
e-mail: yalbert@us.ibm.com

18 months, a trend originally predicted by Gordon Moore and now referred to as Moore's law [1]. The impact of this trend cannot be underestimated, and the resulting increase in computational capacity has greatly influenced nearly every facet of society.

The tremendous success of Moore's law, and in particular, the scaling of Si MOSFETs [2], drives the ongoing efforts to continue this trend into the future. However, several serious roadblocks exist. The first is the difficulty and expense of continued lithographic scaling, which could make it economically impractical to scale devices beyond a certain pitch. The second roadblock is that, even if lithographic scaling can continue, the power dissipated by the transistors will bring clock frequency scaling to a halt. In fact, it could be argued that clock frequency scaling has already stopped as microprocessor designs have increasingly relied upon new architectures to improve performance. These arguments suggest that, in the near future, it will no longer be possible to improve system performance through scaling alone, and that additional methods to achieve the desired enhancement will be needed. Three-dimensional (3D) integration technology offers the promise of being a new way of increasing system performance even in the absence of scaling. This promise is due to a number of characteristic features of 3D integration, including (a) decreased total wiring length, and thus reduced interconnect delay times; (b) dramatically increased number of interconnects between chips; and (c) the ability to allow dissimilar materials, process technologies, and functions to be integrated.

Overall, 3D technology can be broadly defined as any technology that stacks semiconductor elements on top of each other and utilizes vertical, as opposed to peripheral, interconnects between the wafers. Under this definition, 3D technology casts a wide net and could include simple chip stacks, silicon chip carriers and interposers, chip-to-wafer stacks, and full wafer-level integration. Each of these individual technologies has benefits for specific applications, and the technology appropriate for a particular application is driven in large part by the required interconnect density. For instance, for wireless communications, only a few through-silicon vias (TSVs) may be needed per chip in order to make low-inductance contacts to a backside ground plane. On the other hand, high-performance servers and stacked memories could require extremely high densities (10^5 – 10^6 pins/cm²) of vertical interconnects. Applications such as 3D chips for supply voltage stabilization and regulation reside somewhere in the middle, and a myriad of applications exist which require the full range of interconnect densities possible.

A wide range of 3D integration approaches are possible and have been reviewed extensively elsewhere [3]. These schemes have various advantages and trade-offs and ultimately a variety of optimized process flows may be needed to meet the needs of the various applications targeted. However, nearly all 3D ICs have three main process components: (a) a vertical interconnect, (b) aligned bonding, and (c) wafer thinning with backside processing. The order of these steps depends on the integration approach chosen, which can depend strongly on the end application. The process choices which impact overall design points will be discussed later on in this chapter. However, to more fully appreciate where 3D IC technology is heading, it

is helpful to first have an understanding of work that has driven early commercial adoption of 3D integration today.

2.2 Background: Early Steps in the Emergence of 3D Integration

Early commercialization efforts leading to 3D integration have been fueled by mobile-device applications, which tended to be primarily driven by form-factor considerations. One key product area has been the CMOS image-sensor market (camera modules used in cellular handsets), which has driven the development of wafer-level chip-scale packaging (WL-CSP) solutions. Shellcase (later bought by Tessera) was one company that had strong efforts in this area. Many of these solutions can be contrasted with 3D integration in that they (1) do not actually feature circuit stacking and (2) often use wiring routed around the edge of the die to make electrical connections from the front to the back side of the wafer. However, these WL-CSP products did help drive significant advances in technologies, such as silicon-to-glass wafer bonding and subsequent wafer thinning, that are used in many 3D integration process flows today. In a separate market segment, multichip packages (MCPs) that integrated large amounts of memory in multiple silicon layers have also been heavily developed. This product area has also contributed to the development of reliable wafer-thinning technology. In addition, it has helped to drive die-stacking technology as well, both with and without spacer layers between thinned die. Many of these packages have made heavy use of wirebonding between layers, which is a cost-effective solution for form-factor-driven components. However, as mobile devices continue to evolve, new solutions will be needed.

Portable devices are taking on additional functionality, and product requirements are extending beyond simple form-factor reductions to deliver increased performance per volume. More aggressive applications are demanding faster speeds than wire bonding can support, and the need for more overall bandwidth is forcing a transition from peripheral interconnect (such as die-edge or wirebond connection) to distributed area-array interconnect. The two technology elements that are required to deliver this performance across stacked circuits are TSVs and area-array chip-to-chip connections. TSV adoption in these product areas has been limited to date, as cost sensitivities for these types of parts have slowed TSV introductions. So far, the use of TSVs has been dominated by fairly low I/O-count applications, where large TSVs are placed at the periphery of the die, which tends to limit their inherent advantages. However, as higher levels of performance are being demanded, a transition from die-periphery TSVs to TSVs which are more tightly integrated in the product area is likely to take hold. The advent of deep reactive-ion etching for the micro-electro-mechanical systems (MEMS) market will help enable improved TSVs with reduced footprints. Chip-on-chip area-array interconnect such as that used by Sony in their PlayStation Portable (PSP) can improve bandwidth between memory and processor at reasonable cost. One version of their microbump-based

technology offers high-bandwidth (30- μm solder bumps on 60- μm pitch) connections between two dies assembled face-to-face in a flip-chip configuration. This type of solution can deliver high-bandwidth between two dies; however, without TSVs it is not immediately extendable to support communication in stacks with more than two dies. Combining fine-pitch TSVs with area-array inter-tier connectivity is clearly the direction for the next generation of 3D IC technologies.

Wafer bonding to glass, wafer thinning, and TSVs-at-die-periphery are all technologies used in the manufacturing of large volumes of product today. However, the next generation of 3D IC technologies will continue to build on these developments in many different ways, and the impact of process choices on product applications needs to be looked at in detail. In the next section, we identify important process factors that need consideration.

2.3 Process Factors That Impact State-of-the-Art 3D Design

The interrelation of 3D design and process technology is important to understand, since the many process integration schemes available today each have their own factors which impact 3D design in different ways. We try to provide here a general guide to some of the critical process factors which impact 3D design. These include strata orientation, alignment specifications, and bonding-interface design, as well as TSV design point and process integration.

2.3.1 Strata Orientation: Face-to-Back vs. Face-to-Face

The orientation of the die in the 3D stack has important implications for design. The choice impacts the distances between the transistors in different strata and has electronic design automation (EDA) impact related to design mirroring. While multiple die stacks can have different combinations of strata orientations within the stack, the face-to-back vs. face-to-face implications found in a two-die stack can serve to illustrate the important issues. These options are illustrated schematically in Fig. 2.1.

2.3.1.1 Face-to-Back

The “face-to-back” method is based on bonding the front side of the bottom die with the back side (usually thinned) of the top die. Similar approaches were originally developed at IBM for multi-chip modules (MCMs) used in IBM G5 systems [4], and later this same approach was demonstrated on wafer level for both CMOS and MEMS applications. Figure 2.1a schematically depicts a two-layer stack assembled in the face-to-back configuration. The height of the structure and therefore the height of the interconnecting via depend on the thickness of the thinned top wafer. If the aspect ratio of the via is limited by processing considerations, the substrate

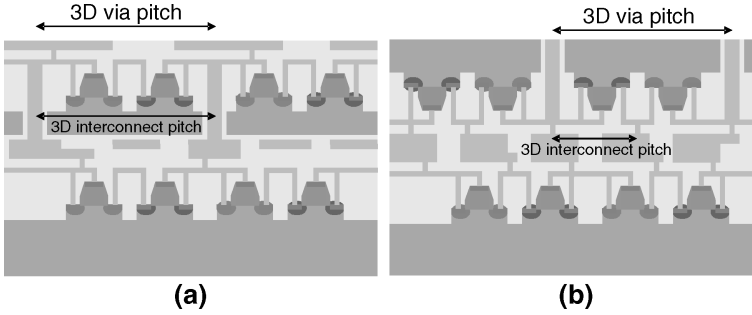


Fig. 2.1 Schematic illustrating face-to-back (a) and face-to-face (b) orientations for a two-strata 3D stack. Three-dimensional via pitch is contrasted with 3D interconnect pitch between strata

thickness can then directly impact the number of possible interconnects between the two wafers. It is also important to note that in this scheme, the total number of interconnects between the wafers cannot be larger than the number of TSVs. In order to construct such a stack, a handle wafer must be utilized, and it is well known that handle wafers can induce distortions in the top wafer that can make it difficult to achieve tight alignment tolerances. In previous work we have found that distortions as large as 50 ppm can be induced by handle wafers, though strategies such as temperature-compensated bonding can be used to reduce them. For advanced face-to-back schemes, typical thicknesses of the top wafer are on the order of 25–50 μm , which limit the via and interconnect pitch to values on the order of 10–20 μm . Advances in both wafer thinning and via-fill technologies could reduce these numbers in the future.

2.3.1.2 Face-to-Face

Figure 2.1b shows the “face-to-face” approach which focuses on joining the front sides of two wafers. This method was originally utilized [5] at IBM to create MCMs with sub-20- μm interconnect pitch with reduced process complexity compared to the face-to-back scheme. A key potential advantage of face-to-face assembly is the ability to decouple the number of TSVs from the total number of interconnections between the layers. Therefore, it could be possible to achieve much higher interconnect densities than allowed by face-to-back assembly. In this case, the interconnect pitch is only limited by the alignment tolerance of the bonding process (plus the normal overlay error induced by the standard CMOS lithography steps). For typical tolerances of 1–2 μm for state-of-the-art aligned-bonding systems, it is therefore conceivable that interconnect pitches of 10 μm or even smaller can be achieved with face-to-face assembly. However, the improved inter-level connectivity achievable can only be exploited for two-layer stacks; for multi-layer stacks, the TSVs will still limit the total 3D interconnect density.

2.3.2 Inter-strata Alignment: Tolerances for Inter-layer Connections

As can be seen from some of the above discussions, alignment tolerances can have a direct impact on the density of connections achievable in the 3D stack, and thus the overall performance. Tolerances can vary widely depending on the tooling and process flow selected, so it is important to be aware of process capabilities. For example, die-to-die assembly processes can have alignment tolerances varying from the 1- μm range to the 20- μm range, depending on the speed of assembly required. Fine-pitch capability in these systems is possible, but transition to manufacturing can be challenging because of the time required to align each individual die. It is certainly plausible that these alignment-throughput challenges will be solved in coming years; when married with further scaling of chip-on-chip area connections to fine pitch, this could enable high-performance die-stack solutions. Today, wafer-to-wafer alignment offers an alternative, where a wafer fully populated with die (well registered to one another) can be aligned in one step to another wafer. This allows more time and care to be used in achieving precise alignment at all chip sites. Advanced wafer-to-wafer align-and-bond systems today can typically achieve tolerances in the 1- to 2- μm range. Although the set of issues for different processes is complex, a deeper discussion of aligned wafer bonding below can help highlight some of the issues found in both wafer- and die-stacking approaches.

Aligned wafer bonding for 3D integration is fundamentally different than blanket wafer bonding processes that are used. For instance, in silicon on insulator (SOI) substrate manufacturing, the differences are several-fold. First of all, alignment is required since patterns on each wafer need to be in registration, in order to allow the interconnect densities required to take advantage of true 3D system capabilities. Second, wafers for 3D integration typically have significant topography on them, and these surface irregularities can make high-quality bonding significantly more difficult than for blanket wafers, particularly for oxide-fusion bonding. Finally, due to the fact that CMOS circuits (usually with BEOL metallization) already exist on the wafers, the thermal budget restriction on the bonding process can be quite severe, and typically the bonding process needs to be performed at temperatures less than 400°C.

Wafer-level alignment is fundamentally different from stepper-based lithographic alignment typically used today in CMOS fabrication. This is because the alignment must be performed over the entire wafer, as opposed to on a die-by-die basis. This requirement makes overlay control much more difficult than in die-level schemes. Non-idealities, such as wafer bow, lithographic skew or run-out, and thermal expansion can all lead to overlay tolerance errors. In addition, the transparency or opacity of the substrate can also affect the wafer alignment. Tool manufacturers have developed alignment tools for both full 200-mm and 300-mm wafers, with alignment accuracy in the $\sim 1\text{--}2\ \mu\text{m}$ range.

Due to the temperature excursions and potential distortions associated with the bonding process itself, it is standard procedure in the industry to first use aligner

tools (which have high throughput) and then move wafers to specialized bonding tools for good control of temperature and pressure across the wafers and through the 3D stack. The key to good process control is the ability to separate the alignment and pre-bonding steps from the actual bonding process. Such integration design allows for a better understanding of the final alignment error contributions. That said, the actual bonding process and the technique used can affect overall alignment overlay, and understanding of this issue is critical to the proper choice of bonding process.

For example, an alignment issue arises for Cu–Cu bonding when the surrounding dielectric materials from both wafers are recessed. In this scenario, if there is a large misalignment prior to bonding, the wafers can still be clamped for bonding. In addition, this structure cannot inhibit the thermal misalignment created during thermo-compression and is not resistant to additional alignment slip due to shear forces induced during the thermo-compression process. One way to prevent such slip is to use lock-and-key structures across the interface to limit the amount of misalignment by keeping the aligned wafers registered to one another during the steps following initial align and placement. This could be a significant factor in maintaining the ability to extend the 3D process to tighter interconnect pitches, since the allowable pitch is often ultimately limited by alignment and bonding tolerances.

Wafer-thinning processes which use handle wafers and lamination can also add distortion to the thinned silicon layer. This distortion can be caused by both differences in coefficients of thermal expansion between materials and by the use of polymer lamination materials with low elastic modulus. As one example, if left uncontrolled, the use of glass-handle wafers can introduce alignment errors in the range of 5 μm at the edge of a 200-mm wafer, a value which is significantly larger than the errors achievable in more direct silicon-to-silicon alignments. So in any process using handle wafers, control and correction of these errors is an important consideration. In practice, these distortions can often be modeled well as global magnification errors. This allows the potential to correct most of the wafer-level distortion using methods based on control of temperature, handle-wafer materials, and lamination polymers.

Alignment considerations are somewhat unique in SOI-based oxide-fusion bonding. This process is often practiced where the SOI wafer is laminated to a glass-handle wafer and the underlying silicon substrate is removed, leaving a thin SOI layer attached to the glass prior to alignment. Unlike other cases, where either separate optical paths are used to image the surfaces to be aligned or where IR illumination is required to image through the wafer stack, one can see through this type of sample at visible wavelengths. This allows very accurate direct optical alignment to an underlying silicon wafer, in a manner similar to wafer-scale contact aligners. Wafer contact and a preliminary oxide-fusion bond must be initiated in the alignment tool itself, but once this is achieved, there is minimal alignment distortion introduced by downstream processing [6, 7].

2.3.3 Bonding-Interface Design

Good design of the bonding interface between the stacked strata involves careful analysis of mechanical, electrical, and thermal considerations. In the next section, we briefly describe three particular technologies for aligned 3D wafer bonding that have been investigated at IBM: (i) Cu–Cu compression bonding, (ii) transfer-join bonding (hybrid Cu and adhesive bonding), and (iii) oxide-fusion bonding. Similar attention to inter-strata interface design is required for die-scale stacking technologies, which often feature solder and underfill materials between silicon die.

2.3.3.1 Copper-to-Copper Compression Bonding

Attachment of two wafers is possible using a thermo-compression bond created by applying pressure to two wafers with Cu metallized surfaces at elevated temperatures. For 3D integration, the Cu–Cu join can serve the additional function of providing electrical connection between the two layers. Optimization of the quality of this bonding process is a key issue being addressed and includes provision of various surface preparation techniques, post-bonding straightening, thermal annealing cycles, as well as use of optimized pattern geometry [8–10].

Copper thermo-compression bonding occurs when, under elevated temperatures and pressures, the microscopic contacts between two Cu regions start to deform, further increase their contact area, and finally diffuse into each other to complete the bonding process. Key parameters of Cu bonding include bonding temperature, pressure, duration, and Cu surface cleanliness. Optimization of all of these parameters is needed to achieve a high-quality bond. The degree of surface cleanliness is related to not only the pre-bonding surface clean but also the vacuum condition during bonding [8]. In addition, despite the fact that bonding temperature is the most significant parameter in determining bond quality, the temperature has to be compatible with BEOL process temperatures in order to not affect device performance.

The quality of patterned Cu bonding at wafer-level scales has been investigated for real device applications [9, 10]. The design of the Cu-bonding pattern influences not only circuit placement but also bond quality since it is related to the available area that can be bonded in a local region or across the entire wafer. Cu bond pad size (interconnect size), pad pattern density (total bond area), and seal design have also been studied. Studies based on varying Cu-bonding pattern densities have shown that higher bond densities result in better bond quality and can reach a level where they rarely fail during dicing tests. In addition, a seal design which has extra Cu bond area around electrical interconnects, chip edge, and wafer edge could prevent corrosion and provide extra mechanical support [9].

2.3.3.2 Hybrid Cu/Adhesive Bonding (Transfer-Join)

A variation on the Cu–Cu compression-bonding process can be accomplished by utilizing a lock-and-key structure along with an intermediate adhesive layer to improve

bond strength. This technology was originally developed for MCM thin film modules and underwent extensive reliability testing during the build and qualification [11, 12]. However, as noted previously, this scheme is equally suitable for wafer-level 3D integration and could have significant advantages over direct Cu–Cu-based schemes.

In the transfer-join assembly scheme, the mating surfaces of the two device wafers that are to be joined together are provided with a set of protrusions (keys) on one side that are matched to receptacles (locks) on the other, as shown in Fig. 2.2a. A protrusion, also referred to as a stud, can be the extension of a TSV or a specially fabricated BEOL Cu stud. The receptacle is provided at the bottom with a Cu pad which will be bonded to the Cu stud later. At least one of the mating surfaces (in Fig. 2.2a, the lower one) is provided with an adhesive atop the last passivation dielectric layer. Both substrates can be silicon substrates or optionally one of them could be a thinned wafer attached to a handle substrate. The studs and pads are optionally connected to circuits within each wafer by means of 2D wiring and/or TSVs as appropriate.

These substrates can be aligned in the same way as in the direct Cu–Cu technique and then bonded together by the application of a uniform and moderate pressure at a temperature in the 350–400°C range. The height of the stud and thickness of the adhesive/insulator layer are typically adjusted such that the Cu stud to Cu pad contact is established first during the bonding process. Under continued bonding pressure, the stud height is compressed and the adhesive is brought into contact and bonded with the opposing insulator surface. The adhesive material is chosen

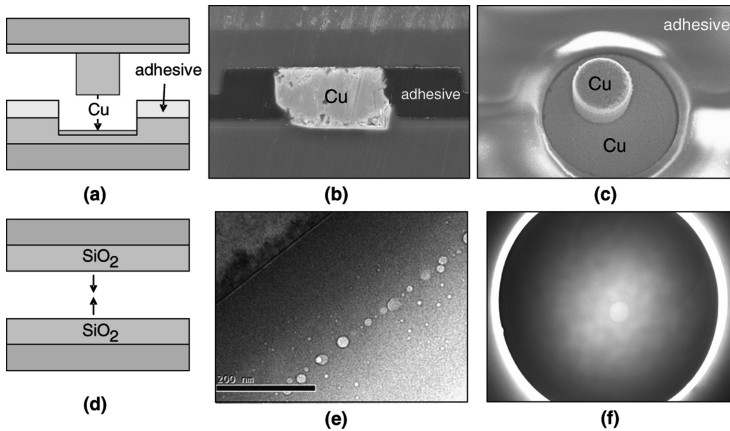


Fig. 2.2 Bonding schemes: (a) cross section of the transfer-join bonding scheme; (b) polished cross section of a completed transfer-join bond; (c) a top-down scanning-electron micrograph (SEM) view of a transfer-join bond after delayering, showing lock-and-key structure and surrounding adhesive layer; (d) oxide-fusion bonding scheme; (e) cross-sectional transmission-electron micrograph (TEM) of an oxide-fusion bond; (f) whole-wafer infrared image of two wafers joined using oxide-fusion bonding

to have the appropriate rheology required to enable flow and bonds the two wafers together by filling any gaps between features. Additionally, the adhesive is tailored to be thermally stable at the bonding temperature and during any subsequent process excursions required (additional layer attachment, final BEOL wiring on 3D stack, etc.). Depending upon the wafers bonded together, either a handle wafer removal or back side wafer thinning is performed next. The process can be repeated as needed if additional wafer layer attachment is contemplated.

A completed Cu–Cu transfer-join bond with polymer adhesive interlayer is shown in Fig. 2.2b. Figure 2.2c additionally shows the alignment of a stud to a pad in a bonded structure after the upper substrate has been delayered for the purpose of constructional analysis. This lock-and-key transfer join approach can be combined with any of the 3D integration schemes described earlier. The fact that the adhesive increases the mechanical integrity of the joined features in the 3D stack means that the copper-density requirements that were needed to ensure integrity in direct Cu–Cu bonding can be relaxed or eliminated.

2.3.3.3 Oxide-Fusion Bonding

Oxide-fusion bonding can be used to attach two fully processed wafers together. At IBM we have published extensively on the use of this basic process capability to join SOI wafers in a face-to-back orientation [13], and other schemes for using oxide-fusion bonding in 3D integration have been implemented by others [14]. General requirements include low-temperature bonding-oxide deposition and anneal for compatibility with integrated circuits, extreme planarization of the two surfaces to be joined, and surface activation of these surfaces to provide the proper chemistry to allow robust bonding to take place. A schematic diagram of the oxide-bonding process is shown in Fig. 2.2d, along with a cross-sectional transmission-electron micrograph (TEM) of the bonding interface (Fig. 2.2e) and a whole-wafer IR image of a typical bonded pair (Fig. 2.2f). The transmission-electron micrograph shows a distributed microvoiding pattern, while the plan-view IR image shows that, after post-bonding anneals of 150 and 280°C, excellent bond quality is maintained, though occasional macroscopic voids are observed.

The use of multiple levels of back-end wiring typically leads to significant surface topography. This creates challenges for oxide-fusion bonding, which requires extremely planar surfaces. While it is possible to reduce non-planarity by aggressively controlling metal-pattern densities in mask design, we have found that process-based planarization methods are also required. As described in [15], typical wafers with back-end metallization have significant pattern-induced topography. We have shown that advanced planarization schemes incorporating the deposition of thick SiO₂ layers followed by a highly optimized chemical-mechanical polishing (CMP) protocol can dramatically reduce pattern-dependent variations, which is needed to achieve good bonding results. The development of this type of advanced planarization technology will be critical to the commercialization of oxide-bonding schemes, where pattern-dependent topographic variations will be encountered on a routine basis. While bringing this type of technology into manufacturing poses many

challenges, joining SOI wafers using oxide-fusion bonding can help enable very small distances between the device strata and likewise can lead to very high-density interconnect.

2.3.4 TSV Dimensions: Design Point Selection

Perhaps the most important technology element for 3D integration is the vertical interconnect, i.e., the TSV. Early TSVs have been introduced into the production environment by companies such as IBM, Toshiba, and ST Microelectronics, using a variety of materials for metallization including tungsten and copper. A high-performance vertical interconnect is necessary for 3D integration to truly take advantage of 3D for system-level performance, since interconnects limited to the periphery of the chip do not provide densities significantly greater than in conventional planar technology. Methods to achieve through-silicon interconnections within the product area of the chip often have similarities to back-end-of-the-line (BEOL) semiconductor processes, with one difference being that a much deeper hole typically has to be created vertically through the silicon material using a special etch process.

The dimensions of the TSV are key to 3D circuit designers since they directly impact exclusion zones where designers cannot place transistors, and in some cases, back-end-of-the-line wiring as well. However, the dimensions of the TSV are very dependent on the 3D process technology used to fabricate them and more specifically are a function of silicon thickness, aspect ratio and sidewall taper, and other process considerations. These dimensions are also heavily dependent on the metallization used to fill the vias. Here we will take a look at the impact of these process choices when dealing with two of the most important metallization alternatives, tungsten (W) and copper (Cu).

2.3.4.1 Design Considerations for Tungsten and Copper TSVs

Impact of Wafer Thinning

Wafer thinning is a necessary component of 3D integration as it allows the interlayer distance to be reduced, therefore allowing a high density of vertical interconnects. The greatest challenge in wafer thinning is that the wafer must be thinned to ~5–10% of its original thickness, with a typical required uniformity of <1–2 μm . In bulk Si, this thinning is especially challenging since there is no natural etch stop. The final thickness depends on thinning process control capabilities and is limited by the thickness uniformity specifications of the silicon removal process (that being mechanical grinding and polishing, wet or dry etching). Successful thinning to a uniform Si thickness of a few microns has been demonstrated, but typically thicknesses of greater than 20 μm are necessary for a robust process. Standard process steps for Si thinning are as follows: First a coarse grinding step is performed in order to thin the wafer from its original thickness (~700–800 μm) to a thickness of

125–150 μm . This type of process is usually performed using a 400 mesh grinding surface. This step is followed by a fine grinding step to thin to ~ 100 μm using an 1800–2000 mesh surface. Next a mechanical polishing step can be performed down to the desired thickness of 30–60 μm . For most processes it is desirable for these grinding steps to be performed only on uniform regions of the silicon, due to the fact that stresses associated with mechanical grind/polish steps can damage fine features in the silicon. If it is necessary to expose the TSV from the backside, it is typically desirable to complete the thinning process using a plasma-based etching process (such as reactive-ion etching) to expose the TSVs. Limitations on the uniformity of the backside thinning sets the practical limit on the TSV depth, and therefore, the via pitch that can be achieved using this type of blind-thinning process. After vias are exposed from the back side of the wafer, a redistribution layer is often fabricated using process technology similar to that used in wafer-level packaging. In some cases more than one level of wiring can be fabricated.

Different wafer thicknesses can lead to the use of different via geometries, as seen in Fig. 2.3. For the aggressively thinned case where the silicon thickness is < 30 μm , fully filled vias made of a single-conductor metallized with either tungsten or copper can be realized in a fairly straightforward manner. However, in the case where wafer thickness is larger, say > 100 μm , the via possibilities take on different forms. For copper vias, the silicon thickness has a fairly direct impact on the size of the TSV footprint at the wafer surface; thicker silicon substrates force larger copper-TSV footprints, which will, at some point, become economically infeasible. Also, copper-based vias need to deal with the low aspect ratios that are reliably possible with copper plating. As wafer thicknesses get larger, there tends to be a transition from full-plating to partial-plating methods in order to maintain manufacturable copper-plating thicknesses. The footprint of tungsten vias tends to be less sensitive to silicon thickness. Although the dimensions of tungsten-filled vias can be limited by deposition-film thickness, the aspect ratios achievable using tungsten deposition can be quite impressive. A wide range of silicon thicknesses are generally possible within a small TSV footprint; at IBM we routinely fabricate tungsten TSVs spanning silicon thicknesses in excess of 100 μm .

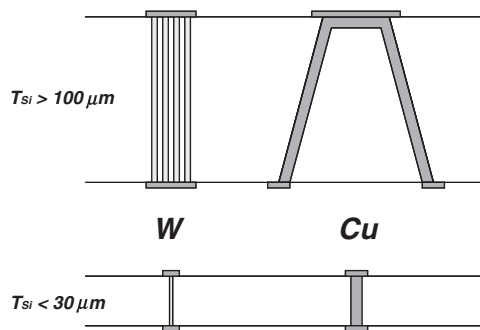


Fig. 2.3 Schematic illustration of the effect of different silicon thicknesses on via geometry for tungsten (W) and copper (Cu) via cases

Impact on Via Resistance and Capacitance

The choice of conductor metallization used in the TSV has a direct impact on important via parameters such as resistance and capacitance. Not only do different metals have different intrinsic values of resistivity, but their respective processing limitations also dictate the range of via geometries that are possible. Since the choice of metallization directly impacts via aspect ratio (i.e., the ratio of via depth to via width), it also has a direct impact on via resistance. Tungsten vias are typically deposited as thin films with very high aspect ratio ($>20:1$) and hence are narrow and tend to have relatively high resistance. To mitigate this effect, multiple tungsten conductors can be strapped together in parallel to provide an inter-strata connection of suitably low overall resistance, at the cost of increased area, as shown schematically at the top left of Fig. 2.3. Copper has a better intrinsic resistivity than tungsten. The allowable geometries of plated-copper vias also have low aspect ratios (typically limited from 6:1 to 10:1) and therefore lend themselves well to low-resistance connections.

Via capacitance can be impacted strongly by the degree of sidewall taper introduced into the via design. Tungsten vias typically have nearly vertical sidewalls and no significant taper; copper vias may more readily benefit from sloped sidewalls. Although sidewall taper enlarges the via footprint at the wafer surface which is undesirable, the introduction of sidewall taper can help to improve copper plating quality and increase the deposition rate of via-isolation dielectrics. These deposition rates are strongly impacted by via geometry, and methods which help to increase final via-isolation thickness will enable lower via capacitance levels and thus improve inter-strata communication performance.

2.3.4.2 Ultra-High Density Vias Using SOI-Based 3D Integration

It is also possible to utilize the buried oxide of an SOI wafer as a way of enhancing the 3D integration process, and this scheme is shown in Fig. 2.4. This so-called SOI scheme has been described extensively in previous publications [13, 16] and is only briefly summarized here. Unlike other more conventional 3D processes, in our SOI-based 3D integration scheme, the buried oxide can act as an etch stop for the final wafer thinning process. This allows the substrate to be completely removed before the two wafers are combined. A purely wet chemical etching process can be used; for instance, TMAH (Tetramethylammonium hydroxide) can remove $0.5\ \mu\text{m}/\text{min}$ of silicon with excellent selectivity to SiO_2 . In our process, we typically remove $\sim 600\ \mu\text{m}$ of the silicon wafer by mechanical techniques and then employ 25% TMAH at 80°C to etch ($40\ \mu\text{m}/\text{h}$ etch rate) the last $100\ \mu\text{m}$ of silicon down to the buried oxide layer. The buried oxide has a better than 300:1 etch selectivity relative to silicon and therefore acts as a very efficient etch stop layer. Overwhelming advantages of such an approach are that all of the Si can be uniformly removed, leaving a very smooth ($<10\ \text{nm}$) surface for the application of bonding oxide films, and the fact that the layer of transferred circuits is automatically a minimal thickness across the wafer, facilitating the fabrication of high-density inter-strata connections later in

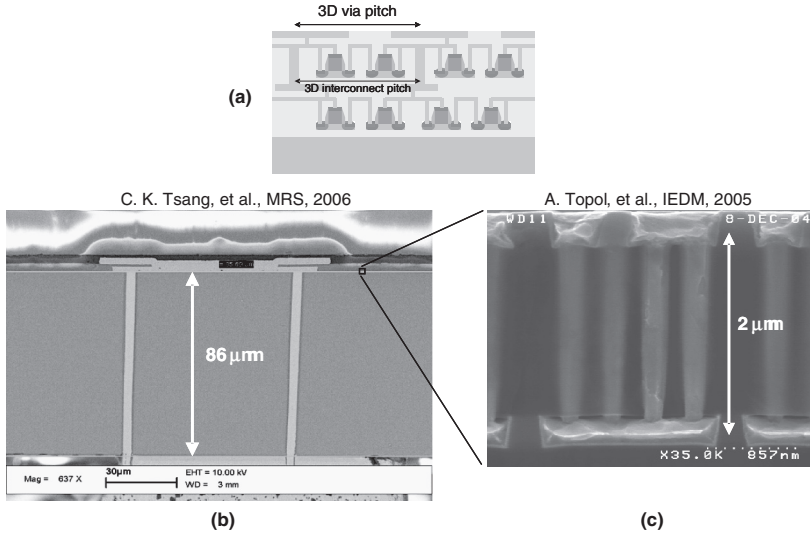


Fig. 2.4 (a) Schematic illustrating face-to-back SOI-3D assembly; (b) annular TSV with a fairly conventional silicon thickness, shown for size comparison with (c) interlayer 3D-via in SOI-3D technology

the process flow. Because the distance between the layers is much smaller than in conventional TSV schemes, the via pitch and size can be dramatically reduced. The minimum height of an interconnecting via could be as low as 1–2 μm, allowing via dimensions as small as 0.2–0.25 μm [13]. If extremely tight wafer-level alignment can also be achieved, then via pitches on the order of 1–2 μm are conceivable, opening the door to numerous new system-level opportunities not achievable with looser interconnect pitches. Figure 2.4b, c shows a comparison of more conventional TSVs (fabricated with tungsten deposited in an annular-shaped via region) with Cu-filled vias fabricated using our SOI-based integration scheme. The size comparison shows the potential advantages of SOI-based 3D integration.

2.3.5 Via-Process Integration and a Reclassification of Via Types

The specific process flow used to fabricate the TSV is of importance to the circuit designer since the method of via-process integration drives specific design rules that must be adhered to. As one example, certain via processes inherently require exclusion rules in BEOL wiring to allow the TSV to pass through, whereas others do not. Another example follows from the fact that alignment tolerances differ for front-side and back-side via processing, which drives via-process-specific design rules. In much of the early literature, the use of the terms “vias-first” and “vias-last” has been common, but such terms have led to significant confusion in the industry. At times these designations have been used to denote whether vias are processed

before or after the base integrated-circuit wafers have been completed (e.g., for vias processed from the front side of the wafer, these terms have been used to distinguish between vias processed before and after back-end-of-line wiring), and in other cases these terms have been used to refer to whether via formation on completed base wafers occurs before or after wafer thinning (e.g., from the front side or the back side of the thinned wafer).

An alternate classification scheme that is based on the timing of the TSV via etch in the process flow can be used to improve clarity. This redefined classification is based on the two most important practical considerations to the designer: (1) Does the process use frontside or backside via etch? (2) If it uses frontside via etch, is this etch done before or after back-end-of-line (BEOL) wiring? This classification leads us to three primary cases of interest:

- (i) Pre-backend frontside via (type-F1)
- (ii) Post-backend frontside via (type-F2), and
- (iii) Backside via etch (after wafer thinning, type-B).

Since backside via etches generally land on the lowest available BEOL metal layer, we ignore here the case where the backside via is etched deeper into the BEOL. Properties of the three primary classes of TSVs are outlined in Table 2.1 and are discussed in further detail below.

Table 2.1 Characteristics of various TSV types

Via type	F1	F2	B
Via etch	Frontside	Frontside	Backside
Via formation	Via before BEOL	Via after BEOL	Via after thinning
High-temperature materials compatibility	+	—	—
Reduced via dimensions	+	—	+
Low circuit blockage	+	—	+
Ease of process-integration	—	+	+

2.3.5.1 Pre-backend Frontside Via (Type F1)

One approach to TSVs that has been demonstrated at IBM and elsewhere is a type of “vias-first” process flow. In this approach, F1 vias are formed before BEOL processing, and thus the approach has the advantage of allowing higher-temperature dielectric and metal fill processes and allows high-aspect-ratio vias to be formed. The F1 via tends to be more challenging to integrate with conventional CMOS, but has the advantage of low-wiring blockage in the BEOL. One particular approach that has been described is an annular-via geometry for large-area contacts [17]. This structure utilizes an etched ring-shaped via geometry, where the ring width is narrow

enough to allow complete fill of the structure utilizing a variety of materials, including doped polysilicon, electroplated copper, or CVD tungsten. For higher-density 3D integration applications, annuli with small central cores (the region defined by the inner diameter of the annulus) can be used, where the annular region is filled with isolation dielectrics and the central core subsequently etched and metallized. The term “pre-backend frontside via” can be extended to all cases where via integration occurs immediately before any particular level of metal (e.g., pre-M1 frontside via, pre-M2 frontside via). It is distinguished from the F2 via in that F2 vias are processed after CMOS fabrication is complete.

2.3.5.2 Post-backend Frontside Via (Type F2)

TSVs can be etched from the frontside of the wafer, taking advantage of frontside alignment capabilities, and yet still be fabricated after completion of BEOL wiring. This F2 via will be preferred in many cases since it is easier to integrate with CMOS process technology and will have the yield advantages that come from not interfering with the standard CMOS process flow. It does have the downside of having slightly larger vias due to the requirement of extending the vias through the BEOL and, more importantly, has the significant disadvantage of blocking wiring channels through the entire BEOL stack. For this reason, unless via dimensions or the total number of TSVs required are small, this type of via can become difficult to implement, especially for designs of high complexity.

2.3.5.3 Backside Via (After Wafer Thinning, Type B)

A variety of vertical through-silicon-interconnect technologies have been developed by IBM [17–19]. Often the TSV is formed only after wafer thinning by utilizing a backside deep reactive ion etch. Like the F2 via, this type B via also has the significant advantage that the CMOS technologies used in the base wafers do not need to be modified. After backside deep reactive ion etching, insulation can subsequently be applied to the interior of the via and selectively removed from the bottom to allow electrical contact to the front side of the wafer. Metallization of large vias prepared in this manner has been demonstrated by various means. For example, the use of an initial partial fill with plated copper followed by evaporation of Cr/Cu BLM and Pb/Sn solder has been demonstrated. Backside vias typically have aspect ratios limited by process capabilities for dielectric and metal fill. They will also be significantly hampered by the alignment capabilities of backside lithography on thinned wafers, and thus may not be well suited for high-density 3D integration.

2.4 Conclusion

In order to be an effective 3D circuit designer, it is important to understand the process considerations that underlie 3D technology. In this chapter, we have tried to outline the basic process considerations that 3D circuit designers need to be aware

of: strata orientation, inter-strata alignment, bonding-interface design, TSV dimensions, and integration with CMOS processing. These considerations all have direct implications on design and will be important in both the selection of 3D processes and the optimization of circuits within a given 3D process.

Technology developments in the areas of CMOS image-sensors, wafer-level chip-scale packages, multichip packages for memory applications, TSVs, and chip-on-chip area-array interconnect are all well on their way and are pushing us rapidly toward the development of full 3D circuit integration. We hope that circuit designers reading this volume will be inspired to learn how to best take advantage of the unique aspects of 3D circuit integration.

Acknowledgments The authors wish to thank the following people for their contributions to this work: Roy R. Yu, Sampath Purushothaman, Kuan-neng Chen, Douglas C. La Tulipe, Narender Rana, Leathen Shi, Matthew R. Wordeman, Edmund J. Sprogis, Fei Liu, Steven Steen, Cornelia Tsang, Paul Andry, David Frank, Jyotica Patel, James Vichiconti, Deborah Neumayer, Robert Trzcinski, Latha Ramakrishnan, James Tornello, Michael Lofaro, Gil Singco, John Ott, David DiMilia, William Price, and Jesus Acevedo. The authors also acknowledge the support of the IBM Microelectronics Research Laboratory, and Central Scientific Services, as well as the staff at EV Group and Suss MicroTec. This project was funded in part by DARPA under SPAWAR contract numbers N66001-00-C-8003 and N66001-04-C-8032.

References

1. G. Moore, Cramming more components onto integrated circuits, *Electronics* 38, 114–117 (1965).
2. R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, Design of ion-implanted MOSFETs with very small physical dimensions, *IEEE Journal of Solid-State Circuits*, SC-9, 256–268 (1974).
3. A. W. Topol, D. C. La Tulipe, L. Shi, D. J. Frank, K. Bernstein, S. E. Steen, A. Kumar, G. U. Singco, A. M. Young, K. W. Guarini, and M. Jeong, Three-dimensional integrated circuits, *IBM Journal of Research & Development*, 504: 91–506 (2006).
4. C. Narayan, S. Purushothaman, F. Doany, and A. Deutsch, Thin film transfer process for low cost MCM-D fabrication, *IEEE Transaction Transactions on Components, Packaging, and Manufacturing Technology*, B18: 42–46 (1995).
5. E. D. Perfecto, R. R. Shields, A. K. Malhotra, M. P. Jeanneret, D. C. McHerron, and G. A. Katopis, MCM-D/C packaging solution for IBM latest S/390 servers, *IEEE Transaction on Advance Packaging*, 23: 515–520 (2000).
6. S. E. Steen, D. C. La Tulipe, A. Topol, D. J. Frank, K. Belote, and D. Posillico, Wafer scale 3D integration: overlay as the key to drive potential, *Microelectronic Engineering*, 84: 1412–1415 (2007).
7. A. W. Topol, D. C. La Tulipe, L. Shi, S. M. Alam, A. M. Young, D. J. Frank, S. E. Steen, J. Vichiconti, D. Posillico, D. M. Canaperi, S. Medd, R. A. Conti, S. Goma, D. Dimilia, C. Wang, L. Deligianni, M. A. Cobb, K. Jenkins, A. Kumar, K. T. Kwietniak, M. Robson, G. W. Gibson, C. D’Emic, E. Nowak, R. Joshi, K. W. Guarini, and M. Jeong, Assembly technology for three dimensional integrated circuits, *Proceedings of 22nd International VLSI Multilevel Interconnection Conference (VMIC)*, pp. 83–88, Fremont CA, October 4–6, (2005).
8. K.-N. Chen, C. S. Tan, A. Fan, and R. Reif, Morphology and bond strength of copper wafer bonding, *Electrochemical and Solid-State Letters*, 7: G14–G16 (2004).

9. K.-N. Chen, C. K. Tsang, A. W. Topol, S. H. Lee, B. K. Furman, D. L. Rath, J.-Q. Lu, A. M. Young, S. Purushothaman, and W. Haensch, Improved manufacturability of Cu bond pads and implementation of seal design in 3D integrated circuits and packages, *23rd International VLSI Multilevel Interconnection (VMIC) Conference*, Fremont CA, September 25–28, 2006, VMIC Catalog No 06 IMIC-050, pp. 195–202 (2006).
10. K.-N. Chen, S. H. Lee, P. S. Andry, C. K. Tsang, A. W. Topol, Y. M. Lin, J. Q. Lu, A. M. Young, M. Jeong, and W. Haensch, Structure design and process control for Cu bonded interconnects in 3D integrated circuits, *IEDM Technical Digest*, 20–22 (2006).
11. H. B. Pogge, et al., Bridging the chip/package process divide, *Proceedings of AMC*, 129–136 (2001).
12. R. Yu, Wafer level 3D integration, *International VLSI Multilevel Interconnection (VMIC) Conference*, Fremont CA, September 24–27, (2007).
13. A. W. Topol, D. C. La Tulipe, L. Shi, S. M. Alam, D. J. Frank, S. E. Steen, J. Vichiconti, D. Posillico, M. Cobb, S. Medd, J. Patel, S. Goma, D. DiMilia, T. M. Robson, E. Duch, M. Farinelli, C. Wang, R. A. Conti, D. M. Canaperi, L. Deligianni, A. Kumar, K. T. Kwietniak, C. D’Emic, J. Ott, A. M. Young, K. W. Guarini, and M. Jeong, Enabling SOI based assembly technology for three-dimensional (3D) integrated circuits (ICs), *IEDM Technical Digest*, 363–366 (2005).
14. P. Leduc, F. de Crécy, M. Fayolle, B. Charlet, T. Enot, M. Zussy, B. Jones, J.-C. Barbe, N. Kernevez, N. Sillon, S. Maitrejean, D. Louis, and G. Passemard, Challenges for 3D IC integration: bonding quality and thermal management, *Proceedings of IEEE International Interconnect Technology Conference (IITC)*, pp. 210–212 (2007).
15. A. W. Topol, S. J. Koester, D. C. La Tulipe, and A. M. Young, 3D fabrication options for high performance CMOS technology, *Wafer Level 3D ICs Process Technology*, C. S. Tan, R. J. Gutmann, and L. R. Reif, Eds., Springer, New York; ISBN 978-0-387-76532-7 (2008).
16. K. W. Guarini, et al., Process technologies for three dimensional integration, *Proceedings of the 6th Annual International Conference on Microelectronics and Interfaces*, American Vacuum Society, pp. 212–214 (2005).
17. C. K. Tsang, P. S. Andry, E. J. Sprogis, C. S. Patel, B. C. Webb, D. G. Manzer, and J. U. Knickerbocker, CMOS-compatible through silicon vias for 3D process integration, *Proceedings of Material Research Society*, 970: 145–153 (2006).
18. P. S. Andry, C. Tsang, E. J. Sprogis, C. Patel, S. L. Wright, and B. C. Webb, A CMOS-compatible process for fabricating electrical through-vias in silicon, *Proceedings of the 56th Electronic Components and Technology Conference*, San Diego, CA, pp. 831–837 (2006).
19. C. S. Patel, C. K. Tsang, C. Schuster, F. E. Doany, H. Nyikal, C. W. Baks, R. Budd, L. P. Buchwalter, P. S. Andry, D. F. Canaperi, D. C. Edelstein, R. Horton, J. U. Knickerbocker, T. Krywaczyk, Y. H. Kwark, K. T. Kwietniak, J. H. Magerlein, J. Rosner, and E. J. Sprogis, Silicon carrier with deep through vias, fine pitch wiring and through cavity for parallel optical transceiver, *Proceedings of the 55th Electronic Components and Technology Conference*, pp. 1318–1324 (2005).

Three-Dimensional Integrated Circuit Design

EDA, Design and Microarchitectures

Xie, Y.; Cong, J.J.; Sapatnekar, S. (Eds.)

2010, XII, 284 p., Hardcover

ISBN: 978-1-4419-0783-7