RC24337 (W0708-088) August 29, 2007
Electrical Engineering

IBM Research Report

3D Integration - Past, Present and Future

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

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home .

3D Integration – Past, Present and Future

Roy Yu
(IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598)
yur@us.ibm.com

Abstract

This paper presents an overview of the fine pitch transfer and joining (T&J) method for 3D device integration. T&J fine pitch transfer and joining technology was originally developed for IBM high-end G5 multi-chip-module (MCM-D) to enable the enhanced wiring density needs. The method enables processing and handling of an ultra thin functional layer such as a dense wiring stack or a MEMS device on a temporary carrier, joining to connect and secure the ultra thin media to a permanent base substrate and the releasing of the temporary carrier. This method has been demonstrated in a wide range of applications, including multi-chip and single-chip modules (MCM/SCM), in chip-to-wafer and wafer-to-wafer 3DI thru-via integration (3DI) for micro-electro-mechanical systems (MEMS) and in micro-opto-electro system (MOES) for hybrid devices. The salient unit process elements of the T&J technology and the process commonality that underpins the wide range of applications made possible with T&J will be reviewed. Potential extendibility of T&J for 3D integration of microelectronic device layers will also be discussed.

Introduction

The semiconductor industry is approaching an interesting cross road at this time. On one hand, the ITRS road map continues to project the scaling trend from CMOS-32 nm node device toward CMOS 15 nm node for the near and long term future [1]. On the other hand, with the reduced returns in performance and the significant cost increase for each new generation of scaling, it is arguable as to whether a scaling only path would make performance and financial sense for the future. In the mean time, 3D integration (3DI) technology activities have made substantial incremental advances in recent years and key elements such as chip and wafer stacking, and thru-Si connections, have begun to show promise and a broader consideration by the microelectronic community. 3DI technology enables shortened signal path, and a potentially bigger data transfer bandwidth among the functional elements of a microelectronic system, leading to potentially higher performance even if one uses an existing, mature base CMOS technology.
It is also plausible to argue that the circuit density could be maintained comparable to scaling if the system is suitably designed to take advantage of the 3DI architecture and the volumetric scaling enabled by stacking layers. Thus one might say that the question is not whether 3D integration should occur but when-- in concert with or after some sort of end point in scaling. This largely depends on how 3DI process technology and the 3DI targeted design mature on the one hand and to what extent novel material and processing solutions continue to extend the cost effective scaling of conventional 2D integration.

The current stacking vs. scaling situation is reminiscent of the one a few decades ago when CMOS technology started to replace bi-polar technology in computing systems. Complex multi-chip modules (MCM) were developed during the bi-polar age to provide the extensive chip to chip interconnects required between the numerous chips that formed the system. CMOS devices and scaling enabled to shrink the much needed circuit content into fewer chips and reduced the reliance on the packaging interconnects and the MCM. The system-on-a-chip (SOC) option became dominant and became the manufacturing method of choice.

As CMOS scaling is pushed to finer and finer ground rules, the fundamental device physics begins to be influenced more and more by the device size and pattern size variations which can significantly lessen the performance gain one would anticipate from continued scaling. This slow down in performance gain from area scaling offers the opportunity for 3DI to take the integration to the third dimension which could enable continued system performance gain and cost reduction.

Beyond the scaling related drive, another important driver for 3DI is the anticipated need to have CMOS devices along with other devices such as opto-electronic devices, hybrid SiGe or InP devices, micro-electro-mechanical or micro fluidic devices to be part of the total system for the overall function, value and performance. In these areas, system-in-a-package (SIP) currently has the edge due to the versatile ability for integration of disparate chips. As in the case of the bipolar to CMOS transition mentioned before, as the 3DI process matures and is able to integrate these components into a system on a chip, there will be a leap forward in system integration and performance gain for these hybrid systems through the use of 3D integration.

It is therefore not surprising that there is increased R&D activity in 3DI and the enabling unit processes and tooling over the last few years. A key new aspect of 3DI is the need to be able to reliably transfer and join multiple functional layers on to a common substrate. Thin film transfer and joining is a collection of process methodologies we have developed to address this need.

In this paper, T&J technology is described beginning with the MCM-D application where many of the fundamental thin film transfer and joining (TFTJ) methods were developed. Further refinement of TFTJ for fine pitch wafer level 3D integration is also discussed which includes wafer level bonding, thinning and thru-silicon via formation. As thin as 1 μm Si films and flexible MEMS circuits were made with the aid of a temporary carrier. This thinned device layer (typically less than 20 um) was then transfer joined to a permanent carrier through a fine pitch via/stud (or lock/key, pin/socket) joining interface. The metal contact in the interface is typically of Cu-Cu (but can be other metals as needed to enable a more reliable bond) with a high temperature adhesive surrounding the contacts for bonding strength. We have demonstrated both chips-to-wafer and wafer-to-wafer fine pitch connections with this transfer and joining method. We also evaluated the joined structures for thermal cycling reliability and the ability to provide wiring across chip to chip gaps in chips to wafer applications.

Looking beyond the present, future system integration will likely involve 3D integration of multiple device layers which may be Si CMOS layers, hybrids containing Si CMOS and MEMS, MOES, SiGe, III-V or micro fluidic device systems. Some examples of early work in this area will also be discussed.

1) Thin film transfer and join for MCM-D

1.1) Thin film fabrication on temporary glass carrier

In Thin film transfer and join (TFTJ) technology a properly selected glass plate is used as a temporary carrier. This type glass has a thermal coefficient of expansion (TCE) of ~ 3 ppm/C and is compatible with TF processing tools and subsequent transfer and joining processes [2]. With a temporary glass carrier the format of thin film panel size is no longer limited to that of the base carrier. This enables thin films to be produced in large panels and multi-up for various application needs, similar to format used in LCD flat panel display production or surface laminated circuitry (SLC) board lines [3]. Large panel size not only allows increased production volume and ability to fabricate multiple part numbers on the same panel but also reduces per carrier recurring cost significantly, making it a cost competitive carrier technology.

With glass panel as carrier, thin film (TF) can be built either top-surface (connecting to chip side at C4 pitch) up or top-surface down depending on the complexity of the package to which the TF is being attached [4, 5]. For simple single chip modules (SCM) it is preferable to build the thin film top-surface (C4) down so that its package I/O is accessible and can be joined to the permanent base carrier surface directly (Figure 3) prior to release of the temporary glass carrier.

In the case of a complex MCM package, top-surface up build of the TF on glass would be preferred as interconnection wires between chips require testing to ensure goodness prior to joining to the permanent substrate. For top-surface up build and after testing, a flip of thin film by bonding its top surface to a second glass carrier is needed. The first glass carrier is then removed by laser ablation to expose the package I/O surface (Figure 1). This flip step allows preparation of the package I/O surface for joining to base carrier. Once the package I/O surface is conditioned for joining, the entire glass plate is diced into single module size for transfer joining of thin film to its base carrier [2, 3, 4, 5].

1.2) Package I/O joining using solder

In TFTJ the package I/O from the TF is actively and directly connected to that of base carrier. Therefore before transfer the base carrier I/O is preferably pre-tinned with solder. The maximum amount of solder is determined by the volume of joining surface adhesive cavities wherein the solder will be contained. A minimum solder volume is required to ensure that there is residual solder available after pretinning the base carrier I/O so as to enable a reliable metallurgical bond between the package I/O pads or studs to the base carrier I/O. For a ceramic base carrier, the joining surface is normally finished with Ni/Au [6]. There are several ways to apply solder to such a base carrier I/O. A simple and cost effective way is to use solder paste with mold transfer technique [7] which is illustrated in Figure 4 (top). In this method a glass plate with matching TCE to the base carrier is etched with cavities of certain size and depth which determine the solder volume. The patterns in the cavities are the mirror image of that of the base carrier I/O. The cavities of the mold are then filled with solder paste. For a 350C reflow process a 90/10 PbSn solder paste with a 400 mesh grade can be used. This mold transfer technique can be used with any type of solder paste, including lead fee solder, offering process versatility. In the case of ceramic carrier with Ni/Au surface, the 90/10 PbSn solder paste can be transferred at a temperature between 310C to 330C [8]. The solder can also be transferred through standard solder screening methods for coarser pitch I/O arrays. Since the surface of the base carrier is not always perfectly flat, mold transfer method typically performs better with a high transfer yield. Figure 4 (below), shows a solder mold filled with solder and an IBM G5 base carrier after receiving the transferred solder.

1.3) Adhesive and TF tack attach lamination

A key aspect in TFTJ is the use of a thermoplastic polyimide adhesive, such as Kapton-KJ, from Dupont [9, 10]. Some of its properties are listed in figure 5 (top). Kapton-KJ has a Tg of ~220C. Above this Tg Kapton-KJ undergoes a second phase plateau. It remains thermally stable with a relatively strong modulus in this plateau until 400C. In this 220C-400C window Kapton-KJ can form a strong adhesive bond to nearly any type of material layers of interest and requires only relatively low (~200 PSI) bonding pressure (Figure 5, below).

Kapton-KJ is available either as a dry composite film or as spin-on solution (3003X1 from HD Microelectronics). If a dry film is used, the film is preferred to be an adhesive-core-adhesive composite. The adhesive film we have studied was kapton-EKJ200 which has a 38 μm core and 12 μm KJ adhesive on both sides. EKJ dry film can be easily punched to form I/O via receptacles for solder connection. Figure 2 illustrates lamination of thin film with dry film (discrete) adhesive. When spin-on solution is used, Kapton-KJ is coated with another base layer of polyimide on the I/O surface to form a double layer composite adhesive structure. Laser ablation or lithography and RIE is then used to open the joining surface for solder connection. Figure 3 shows the lamination assembly in the case where spin-on KJ is used. With solder and adhesive in place, the thin film is tack attached to the base carrier through a 330C reflow. During reflow a moderate iso-static pressure is applied to hold the glass with TF in place (Figure 6). For dry film composite adhesive somewhat higher pressure is required to prevent the composite adhesive from thermal movement. The pressure not only helps form the preliminary adhesive bond but also forces solder to make a metallurgical connection between the TF contacts and base carrier I/O. The stability of the adhesive core (or the base layer of the spin-on double layer) maintains the solder in its designated join surface space. At this stage the adhesive is not totally bonded between TF and base carrier surfaces but has sufficient strength to secure the thin film to the base carrier surface. Solder joints add additional security that the glass plate can be removed by laser ablation, as described in the next step, without compromising thin film structural integrity or its alignment to the base carrier.

1.4) Glass carrier laser release and post release lamination

After tack reflow to secure TF to the base carrier, the glass plate is released using UV excimer laser ablation [2,4]. The basic mechanism of the ablation process is the absorption of 308 nm laser energy in polyimide leading to the cleaving of bonds and volatilization of a very small and controllable thickness of the polyimide (Figure 7, left). Laser release of TF from glass uses the same principle. The glass plate

selected as temporary carrier has adequate transmission to the 308 nm laser wavelength and allows thin films to separate from glass within 100 nm of the polyimide/glass interface. The ablation onset fluence at the interface is determined to be around 70-100 mJ/cm^2 (Figure 7, right). This laser fluence range is selected to be just above the ablation threshold of the polyimide release layer and has been shown to impart minimal dynamic stress on the TF during release [2].

After glass plate release, the thin film and the base carrier undergo a vacuum assisted iso-static lamination which is performed to bond the TF permanently to base ceramic carrier surface through the KJ adhesive while maintaining I/O connection with the solder. During lamination the substrate is placed in a vacuum system which maintains the base pressure below 100 mTorr [2, 9,10]. Figure 6 (top) shows the schematic of the laminator (above) and the vacuum assisted iso-static fixture (figure 6, below). After the system reaches 350C lamination temperature the isostatic gas lamination pressure is applied through a flexible foil to a pressure block [2]. The foil also acts to isolate the evacuated portion of the fixture where the parts sit from the pressurized side. There is a compliant conforming layer placed between the pressure block and the thin film. The pressure block transfers the gas pressure through the compliant layer to the thin film, achieving a uniform pressure across the thin film surface regardless of the base carrier surface roughness and topography where the thin film is bonded to the base carrier.

After the thin film is laminated to the base carrier, the C4 contacts are opened using either laser ablation or a RIE process depending on how the C4 metal is structured [11,12]. Ni/Au is plated on the C4 metal for chip to join to.

2) Transfer and joining for wafer level 3DI

2.1 Ultra thin wafer on a glass carrier

For the 3DI wafer thinning application, the thin film on glass with the flip build (Figure 1) is extended to enable the transfer of a thinned wafer on glass [5, 13, 14a, 14b]. To ensure the wafer release in a later step from the glass carrier, the wafer is protected with a layer of polyimide (PI), such as that used for wafer passivation. A free-standing 12 μm Teflon-PFA film is used as the adhesive between the flip glass and wafer [5, 13, 14a, 14b]. The glass-Teflon-PI-wafer stack is then laminated together, after a proper surface treatment. The lamination is typically at 380C for 30 min under an iso-static pressure of 50-100 PSI. Once bonded, the wafer can be thinned to almost any thickness with the glass as a carrier. As thin as 1 μm Si films can be obtained and processed in this way. Ultra thin Si wafer is actually very flexible. Figure 10 (top) shows a wafer level transfer process flow. Figure 10 (middle) shows a 5 μm Si film with protective coating after release and shows that it can be bent in a complete circle. Figure 10 (below) shows a wafer which was bonded and thinned to remove much of the Si bulk [14b]. To minimize damage to the devices, the wafer is normally polished to ~ 100 μm and then RIE trimmed to the desired final thickness of 5-20 μm [14a].

2.2) Thru-Si Via formation

The thru-Si vias can be formed either before the wafer bonding to glass (via-first) or after the wafer thinning (via-last) [14a, 17]. Via first normally allows for a wider material and processing choice as there is less constraint on the process temperature due to the Teflon adhesive. Teflon-PFA retains its form to 280C. Beyond that it can blister or can cause the thin Si to distort due to its TCE. The disadvantage of via-first approach is that via depth uniformity can not be always guaranteed and care must be taken in opening of the via bottom with CMP which could lead to Si cracking. This is one reason a final RIE open of the vias is preferred for a "soft" landing. Figure 11 (below) [14a] shows a cross section of a thru-via chain with Ni fill for a 20 μm via depth in a thru-Si via-chain test vehicle.

2.3 Wafer-wafer and chip-wafer T&J joining

The T&J interlocking interfaces are the most important structure for the joining step [4, 5, 13, 14a, 15, 16]. A set of stud/via (lock/key, or pin socket) are formed on the bottom of the thinned wafer and on the top of the receiving wafer, respectively (Figure 11, top). This T&J interlocking interface serves two

important functions. One function is to provide metal to metal direct electrical contact. Since the studs are slightly taller than the vias are deep, the electrical contact is guaranteed. The other function is to provide improved registration between the aligned wafers. Better than 1 μm overlay across an entire wafer can routinely be achieved after both pre and post lamination bonding [13, 14a].

With the wafers aligned, the lamination bonding is typically performed at 330C-380C with 100 -150 PSI iso-static pressure, depending on the types of the adhesives used. Figure 12b shows such a final stack structure with 3DI via chain with via-first Ni electro-less filled thru-vias for the top wafer. The studs are on the top of the bottom wafer and are provided with plated Au [14a]. The via-recess is on the bottom of the top-wafer with Ni as contact metal. The via-recess is a double-layer polyimide. The via-chain thus formed has an almost 100% contacts yield. The defects are frequently associated with the via-fill rather than the bonding operation. Due to this fill issue an electro-plating fill with Cu, Ni, or Au via-fills is some times preferred for vias with high aspect ratios.

The improved alignment enabled by this T&J interlocking assembly structure is particularly important for placing chips on a wafer for mixed chip architecture as each individual chip registers to its own via/stud for alignment regardless of where the chips come from or the type of the registration patterns. Figure 12 (below) shows the concept of chip-wafer attachment and Figure 12 (top) shows a corner of 4 chip set aligned with via/stud scheme described above, viewed through the glass carrier. With the T&J interface, chips aligned to a base wafer can be lamination joined and thinned just like a bonded wafer, allowing chip-wafer assembly [5, 13, 14a, 17].

3) Hybrid systems integration

Further integration beyond the scaling/3DI is anticipated from hybrid systems in which SiGe, III-V devices, MEMS, opto-electro systems are integrated into a single chip. T&J technology is applicable for these as well. These will only be described briefly. The detailed process for the construction is similar to that used for MCM and wafer transfer. The interested readers are referred to the original publications [14a, 16, 18]

3.1 MEMS application

Figure 13 shows a MEMS high density storage media transferred onto a CMOS driver. The details of the module level assembly are described in reference [14a].

3.2 Opto-electro systems

Figure 15 shows an array of microns sized optical channels encased in a metal reflector. Integration of opto-electro with CMOS and the process flow also follows that of T&J [18].

3.3 SiGe, III-V hybrid with CMOS layers

Si CMOS has the mature infrastructure, large wafer capability and high circuit density while SiGe, Ge (or GaAs) devices can achieve higher circuit speed and or special functionality. If these two components can form logic-memory combination then the processor speed and memory size can be greatly enhanced. Due to the material and substrate size compatibility issues the different materials normally can not be processed on the same wafer. With T&J, the wafers can be separately processed and diced and then recombined on to a base carrier with proper connection as in discussed in [18] and illustrated schematically in Figure 16.

Figure 15 shows such a concept in which chips from various sources (CMOS, SiGe, GaAs, VCSEL, MEMS, stacked memory, among others) are brought together and fine pitch (~ 20 μm) connected. This can enable SOC performance for processor to memory, CMOS to SiGe, CMOS to MEMS, SiGe to VCSEL, and other chip mixes with an SIP packaging flexibility and versatility.

4) Test Data and discussion:

4.1 MCM/SCM

To-date, over 100 IBM G5 modules have been built and evaluated for construction and reliability, as well as manufacturability. A fully functional G5 module was put in user condition for extended period for field performance study. We also generated extensive electrical data between conventionally fabricated and TFTJ built modules to assess electrical impact due to the introduction of the joining surfaces in the final structure. No performance disparity, due to the joint structure, was noted. TFTJ build meets all IBM high end MCM performance requirements. Figure 8 shows a G5 module (above, right) and several other MCM modules (above, left) after the TF transfer and the chip attachment. Figure 8 (below) shows the typical cross section of the TFTJ connections in such MCMs. TFTJ is also under evaluation for SCM on alumina ceramic base carrier. This is a thin film patch concept which applies thin film only to a localized area of a chip carrier or a card where high density wiring is required. Figure 9 (above) shows several such products being evaluated. Figure 9 (below) shows an SEM cross section of the joint. This local patch arrangement by TFTJ combines the advantages of low cost base carrier with nominally low complexity and wiring density provided with high wiring density of thin film wiring only in a small local area. This enables optimized chip carrier performance at system level while reducing over all module cost. This is possible because with TFTJ, the base carrier can remain at a relatively coarse ground rule with a better electrical resistivity, high production yield, and low cost. With the small area thin film patch, more patches can be made on each temporary panel reducing cost per patch especially when compared with serially additive processing of the same local pattern on the base carrier one at a time. Most significantly, TFTJ allows thin film to be processed on a large panel line such as flat panel display line which would further reduce processing cost. One interesting application space is to apply a TFTJ patch directly to a low density micro card. In this way any low density card can be transformed into higher density by a local patch.

4.2 Wafer level 3DI

Another significant advancement made in TFTJ is in wafer level fine pitch SOC/SIP packaging. We have recently demonstrated TFTJ transfer and join density down to 10 μm line and 20 μm pitch. Part of this work was reported in [13]. Figure 14 shows such a transfer of interconnect for 10 μm line at 20 μm pitch stitch pattern and 5 μm lines on a wafer level. This work significantly increases workable joining and wiring density for interconnects. Figure 17 shows the reliability test results for these types of via chains. The structure with joining via sizes of 10, 15, 20, 25 μm were subjected to 13 chip reflow cycles to 350C to simulate chip joining and module rework and then put through 3000 cycles of 25 C to 150C thermal cycles for thermal stress. The via resistance values are plotted as a function of thermal stress time. For 10 μm via-chains, there is about a 2x resistance increase during chip join reflow and the resistance remains unchanged for the thermal cycles. For via size 15 μm and larger, there is no appreciable change in the via-chain resistance.

Summary:

A very versatile thin film transfer and join (TFTJ) technology has been developed. This technology enables the use of high density thin film wiring locally to where it is needed on a nominally low density base carrier. The technology has been evaluated for ceramic high end MCM and low end SCM applications. Significant advances in TFTJ have also been made at finer pitch (20-50 μm) for wafer level 3DI and mixed chip integration applications. These advances can ultimately help bridge the gap between and leverage SOC and SIP, achieving SOC performance with SIP versatility.

Acknowledgement

I would like to thank Dr. S. Purushothaman for his critical review of this manuscript and suggestions.

[1] ITRS 2006 updates, http://www.ITRS.org/
[2] C. Narayan, et al., IEEE trans. Comp. Pkg. Mfg. Tech., Part B, Vol 18, Issue 1, 1995, p42
[3] G. White, et al, IEEE Trans. Comp. Pkg. Mfg. Tech. Part B, Vol. 18, Issue 1, 1995, p37
[4] United States Patent 6,281,452

Prasad , et al. August 28, 2001
[5] United States Patent 6,444,560
Pogge , et al. September 3, 2002
[6] G.A.Katopis, et al, IBM J. Res.develop. Vol 43, No. 5/6, (1999), p621.
[7] United States Patent 6,099,935
Brearley , et al. August 8, 2000, Solder
[8] S.J.Jeng et al, IEEE electron device letters, Vol 22, No. 11, 2001, p542.
[9] R. K. Kanakarajan, IPC national conf. On Flex Circuits, 1995, Minneapolis, MN
[10]K. Kanakarajan and John A. Kreuz, US patent 5298331 (1994).
[11] K. Prasad and E. Perfecto, IEEE trans. Comp. Pkg. Mfg. Techno. Part B, Adv. Pkg., Vol 17, No.1 (1994) p38.
[12] E. Perfecto, et al, IEEE trans.Adv.Pkg., Vol 23, No.3 (2000), p1.
[13] H. B. Pogge et al., Proc. AMC'2001, pp.129-136,2001;
[14a] M. Despont, et al; TRANSDUCERS, Solid-State Sensors, Actuators and Microsystems, 12th International Conference on, 2003,Volume 2, 8-12 June 2003 Page(s):1907 - 1910 vol.2
[14b] Guarini, K.W., et al; IEMD2002 p943 – 945.
[15] United States Patent 6,090,633
Yu , et al. July 18, 2000 ML-TFTJ
[16] United States Patent 6,835,589
Pogge , et al.
[17] United States Patent 6,856,025
Pogge , et al. February 15, 2005
[18] United States Patent 6,640,021
Pogge , et al. October 28, 2003

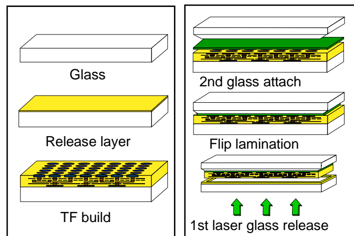


Figure 1: top-side up build

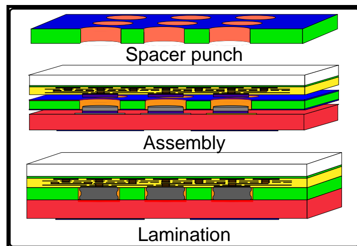


Figure 2: discrete build

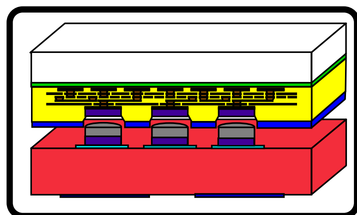


Figure 3: integrated build

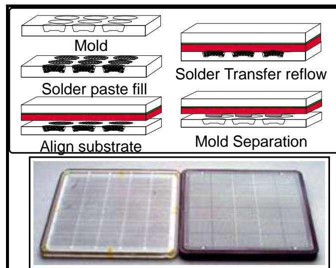
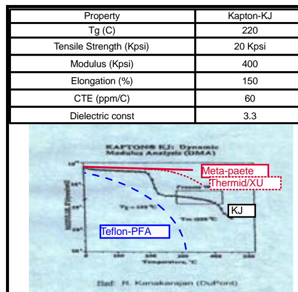


Figure 4: solder transfer



Property	Kapton-KJ
Tg (C)	220
Tensile Strength (Kpsi)	20 Kpsi
Modulus (Kpsi)	400
Elongation (%)	150
CTE (ppm/C)	60
Dielectric const	3.3

Figure 5: Adhesive

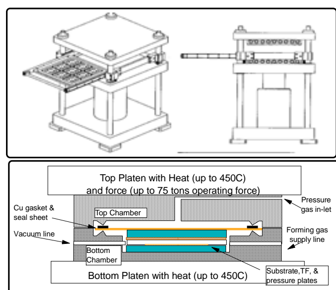


Figure 6: Lamination

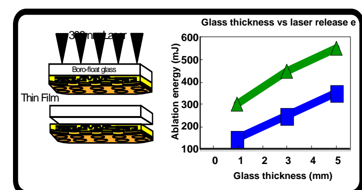


Figure 7: Laser release

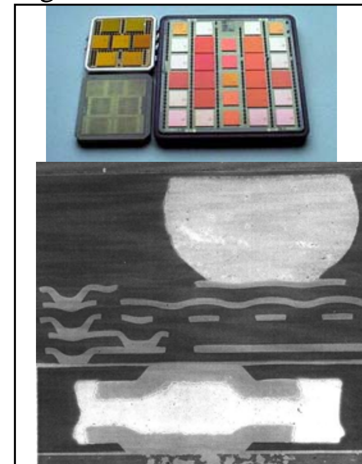
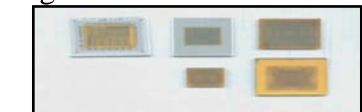


Figure 8: MCM



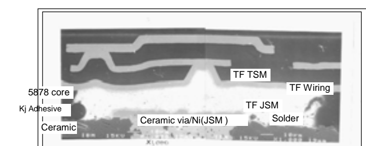
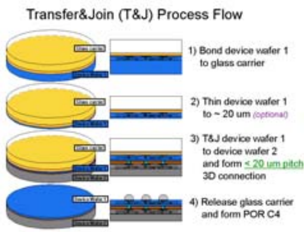


Figure 9: SCM



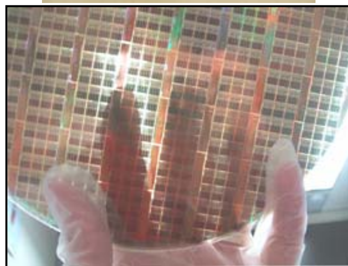


Figure 10: Wafer thinning

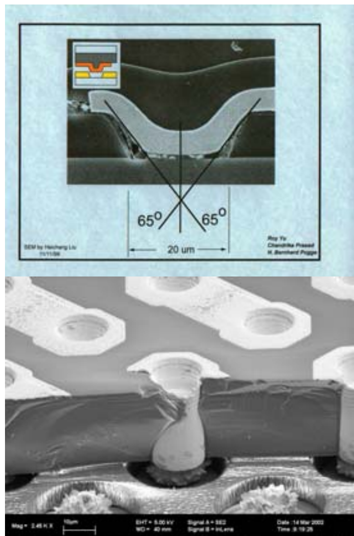
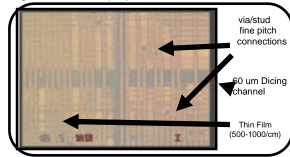


Figure 11: Thru-via TJ join



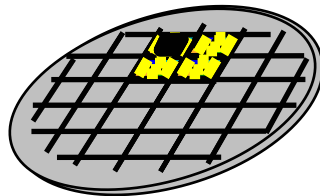


Figure 12: Chips-wafer assembly

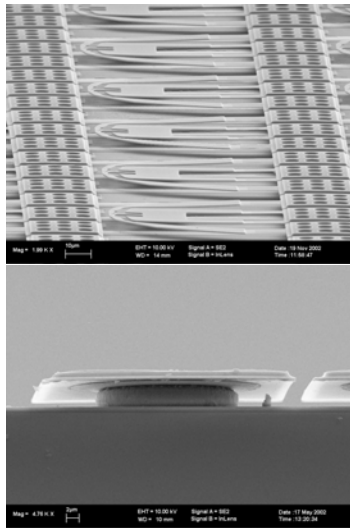


Figure 13: MEMS

Figure 14: Fine pitch T&J

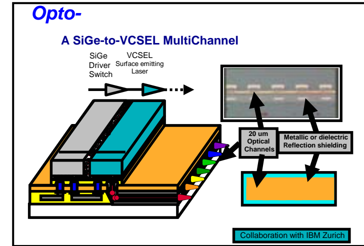
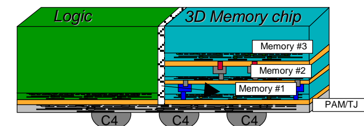
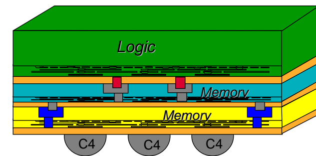


Figure 15: Hybrid application





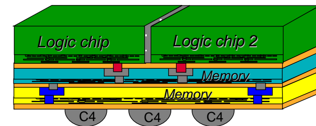


Figure 16: wafer/chip 3DI

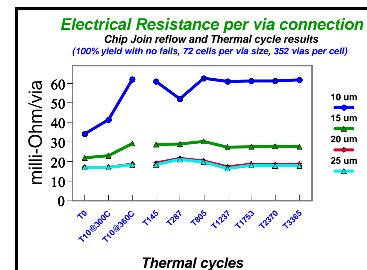


Figure 17: Fine pitch TJ reliability