RC24516 (W0806-114) June 26, 2008
Electrical Engineering

IBM Research Report

High Density 3D Integration

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 .

High Density 3D Integration

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

Abstract

This paper discusses the current and future needs in continued CMOS scaling, reviews the status of the transfer and joining (TJ) technology for MCM-D and wafer level 3DI integration, and explores the opportunities of the TJ technology in the realm of the "More than Moore" era.

Introduction

The continued CMOS device scaling in the past decades has enabled over a million fold increase in device content per chip and has exponentially increased device performance and reduced the device cost per bit by as much in the trend known as the "Moore's law" [1]. The recent inclusion and revision of "More Moore" and "More than Moore" directions in the ITRS road map for semiconductor industries [2] is a clear recognition that more content with better integration on the system level can be achieved through continued device scaling as well as through hybrid system integration for at least another decade. While transistor speed is gaining ground by switching to high–K gate dielectric and metal-gate coupled with SiGe gate stressor [3], the memory latency and wiring RC delay continue to challenge the 2D configuration in chip architecture. The current drive in device architecture toward the "More Moore" concept is largely the result of the shift from the frequency scaling toward the parallel scaling in order to alleviate power consumption and memory bottleneck while enabling the continued device size reduction [4]. The recent rise in device multi-cores approach has allowed better system performance architecture and a more balanced power management [5] while maintaining the trend in device density increase. The performance gain by multi-cores is well established by the "Amdahl's law" in the mainframe high performance computing [6]. It is recently that the multi-cores in device enabled the chip performance enhancement to continue while reducing the costly power consumption, and to some extent the memory bottleneck [7]. The eight cores Cell chip [8], IBM PowerPC architecture [9], and the Intel tera-scale 80-core processor [10], are some of the multi-cores examples. Multi-cores architecture helps to reduce some of the wiring delays while maintaining the device performance by parallelism [11]. Each generation device dimension scaling increases the proportion of RC delay between the CPU and the memory cache and reduces the efficiency of CPU access to the memory [12]. To place the memory closer to the CPU with a higher bandwidth access bus is critical to increase the chip overall performance. 3D architecture, placing the memory directly over the CPU using the through-Si-via (TSV) architecture offers the nearly perfect memory access (shortest distance and widest bus) for the CPU and the best chip overall performance [13]. In a 3D configuration, with a mere re-arrangement of the position of the cache to the logic from 2D to 3D, the device performance can be increased by as much as 30%, which is equivalent to complete generation of 2D device scaling (18 months) with everything else being equal [13].

Therefore, 3D integration will be a critical part of the device integration for CMOS scaling to continue in the future. With the recent rapid development in hybrid 3D integration technologies, the device performance at system level by innovative system schemes is expected to advance concurrently with or over the 2D scaling. Such cross-over transitions between the integrated (serial, homogeneous, device-centric) and the hybridized (parallel, heterogeneous, component-centric) systems had repeated many times since the very beginning of the IC revolution: Vacuum tubes (hybrids), transistors (integrated), discrete circuits (hybrid), integrated circuits (integrated), multi-chip modules (hybrid), system on chip (integrated), and now 3D system integration (hybrid). Technology maturity and the cost often dictate the prevalent technology options. The integrated and hybridized architectures both enable a much higher system density, and therefore system performance but in their unique ways: The integrated architecture allows a streamlined process, a high volume production, a better system scaling, and a diverse components specialization. The hybridized architecture offers new functionality, better system contents, specialized applications, and new market opportunities. We are in the middle of another such cross-over transition.

It must be noted that the underlying driving force to adopt of any new technology is its ability to concurrently enhance performance and reduce cost. CMOS 2D scaling in the past decades has been able to increase the chip content while reducing the cost per bit by a million times. Any successful 3D technology will also have to meet such a challenge in order to be a viable contender to maintain the continuation of the CMOS scaling as well as to enable and lead the "More than Moore" advancement.

This paper will discuss the 3D integration needs in both system hybrids and system stacks. The current status of chip-to-chip, chip-to-wafer, and wafer-to-wafer architectures will be reviewed. In particular, transfer and joining, a fine pitch system level 3D integration method will be presented. This transfer and joining method for 3D integration was initially developed for IBM high end system thin film multi-chip module (MCM-D). The technology was then further enhanced into a wafer level bonding, thinning, and fine pitch transfer joining connection technology with and/or without the aide of a glass carrier. We will demonstrate the transfer and joining technology in wafer level applications in fine pitch 2D CMOS integration, in CMOS wafer thinning, in CMOS to MEMS application, and in a proposed CMOS to opto-electronics application. We will review the results in 2D chip-to-wafer

(C2W) applications for chip-to-wafer placement, bonding, polishing, and wiring over the gaps between chips. We will also review the results of 3D fine pitch transfer-join contacts yield, reliability with varying contact via sizes. Finally, we will discuss whether fine pitch 3DI as applicable to packaging, as interconnect, or wafer level integration.

II. 3DI Overview

3DI introduces a new integration dimension in devices and various methods have been proposed to address the associated challenges [14]. The approaches can be grouped into package-centric (chip-chip, or C2C) and BEOL-centric (wafer-wafer, or W2W). In the package-centric approach the emphasis is on known good die (KGD), and on chip-chip or passive-chip integration where the wiring density improvement is secondary to functionality. Multi-stacking and lower cost are possible but with relatively higher difficulty [15]. In the BEOL-centric approach the focus is on wafer level connections and potentially with a high TSV density and lower cost [16]. The following is a brief overview of the commonly used 3DI approaches:

3DI with lateral connections (chip stacking):

Chip stacking is one of the earliest 3D integration concepts [17] in which memory chips are stacked together and the signal leads are taken out from the chip edges and are wire bonded to a logic chip. Such stacks are widely used in mobile devices today for low power and lower IO count devices to make the package with a higher density for information storage. The advantages of the chip stack methodology are its compactness and its requirement of limited power and IO density. Recently, through-Si-vias are incorporated in the chip stacks and allow memory-logic with either face-to-face (F2F) through C4 flip chip connection and form the basic building block in C2C and in W2W 3DI scheme [18]. More recently the system-in-a-cube (e-cube) concept is targeting high volumetric density for mobile and medical application [19].

3DI with through-connections:

For desktop and server application where IO band width and fast switching speed are required the 3DI technology favors through-Si-via (TSV) with a wider bandwidth. In such cases wafer-wafer (W2W), chip-chip (C2C) and chip-wafer (C2W) direct bonding and IO connections offer integration alternatives with potential for higher performance and cost reduction.

3DI covers several key components outlined below: 1) through-Si-via (TSV), 2) wafer bonding and thinning, and 3) interface-via-connection (IVC). Each 3D integration scheme addresses these aspects in its own unique ways:

1) **Through-Si-Via (TSV):**

TSV provides a direct, short and wide bus path for device-device communication in a 3DI stack in either C2C or W2W format [20]. Since it is more suitably formed at wafer level, TSV has brought the wafer level integration to the forefront. TSV process is strongly coupled with wafer bonding/thinning capability. For wafers with thickness 20-100 μm tungsten TSV is preferred. Tungsten has a closer thermal coefficient of expansion (TCE) to Si [15] than other metals, a high aspect ratio via fill by the CVD process, and is capable of a tight TSV pitch. However, tungsten's high resistivity, high stress and high temperature deposition requirements limit its wide use. 50:1 aspect ratio W via-fill has been demonstrated in Si as thick as 100 μm [15]. The W TSV pitch is normally limited by the wafer thickness and resistance requirements. Currently, W TSV is mostly considered applicable to Si thickness above 20 - 50 μm with the functional pitch about in about the same range.

When wafers can be thinned below 20 μm other TSV fill materials like Ni and Cu become alternatives to W. This is because the TSV size can be reduced and Ni and Cu TCE become less of an issue. Cu plating fill aspect ratio at 5:1 or less also restricts Cu to such a low thickness. Cu or Ni TSV filling is desirable since they conform to the back-of-the-line (BEOL) process and allow a greater TSV compatibility with the BEOL flow.

2) **Wafer and chip bonding**:

A reliable 3DI wafer bonding is vital to enable post-bonding Si thinning, TSV choice, interface via connection (IVC), post-bonding BEOL, and multi-layer device stacks. The primary options for wafer bonding are: A) metal bonding, B) dielectric bonding, and C) metal/dielectric hybrid. The type of bonding interface also determines the TSV build before the bonding (via-first) or after (via-last). There is also the stack direction. Wafers can be connected face to face (F2F) or face to back (F2B). This stacking direction is dictated by how the stacks are carried in the process, the registration and functionality required. Memory stacks tend to use F2B stacking while memory-logic prefers F2F stacking. F2F allows a better managed registration and performance while F2B enables a better managed design. In this aspect W2W and C2C share the requirements. The commonly used bonding methods and their salient features are summarized below:

A1) **Metal bonding with C4 or micro-C4:** C4 bonding uses solder for metal connection. It is widely used in chip stacking and TSV C2C with relatively thick 3DI chips with a more relaxed connection pitch. To reduce the TSV thickness (for smaller pitch) an underfill to secure the thin chip is normally required. For low connection density (> 50 μm pitch), this approach can be effective, especially if the wafer level yield is low and the known-good-die assembly KGD is necessary. This approach normally requires a via-first option. It is possible but can be quite complex for multiple chip stacks in this way [15].

A2) **Metal bonding with Cu-Cu thermal compression**: This type connection extends the connectivity ability of flip chip solder connection and can be used for both C2C and W2W. The metal contacts form interface connection as well as mechanical bonding between wafers. The metal interface is simple to form and has a relatively high strength. Normally this structure has a good vertical stress relief on the contacting surfaces during thermal cycles. They can have a tighter pitch down to a few microns and typically work well for thick wafers [21]. For tighter pitch 3DI with wafer thinning to less than 20 μm, the wafers tend to bend and crack along the bonding edges due to the lateral stress build-up. Since metal is

not as compliant as much as adhesive, it is relatively difficult to achieve a wafer level bonding with metal contacts alone. This approach is relatively successful for chip level bonding with a thick 3D chip. This approach also prefers a via-first option.

B1) **Dielectric interface with low temperature oxide-oxide bonding:** Oxide-oxide bonding stems from bulk SOI bonding [22]. In SOI the bonding oxide goes through over 1000C and forms a bonding interface. Low temperature oxide bonding takes advantage of surface activation to achieve the bonding strength [23]. Due to the nature of the bonds oxide-oxide bonding has a relatively low surface adhesion strength of several J/m^2, which is sufficient for wafer grinding and polish but might be marginal for continued BEOL and ability to sustain chip-package interaction (CPI). With a high content of through-via connections using via-last process the oxide-oxide bonding can be enhanced for continued BEOL and CPI. Oxide-oxide bonding normally starts from the wafer center and progresses toward the edge and provision is required for air to escape. Care must be given for alignment drift during such bonding.

B2) **Dielectric interface with adhesive bonding**: This alternative to oxide-oxide bonding, adhesive bonding originates from packaging technology. A layer of compliant (or B-staged) dielectric material or a dry adhesive is used to bond the wafers together. This approach frequently requires a vacuum bonder to ensure the interfaces are free of air pockets. Similar to oxide-oxide, adhesive bonding requires a via-last process for power and signal IO with slightly higher TSV aspect ratio due to the adhesive thickness of several μm. BCB and polyimide adhesives are commonly used as high temperature adhesives [24]. More recently, BEOL low-k dielectrics are being explored for enhanced thermal budgets [25]. After the bonded wafers are thinned the connections are added through the thinned top wafer in a BEOL-like process. For a fine pitch connection, this approach requires the top wafer to be thinned to less than 10 μm. Compared to oxide-oxide bonding, dielectric adhesive bonding has better topography conformity but has lesser registration accuracy, through-vias connection pitch, and BEOL thermal budget.

C1) **Hybrid with Metal/oxide**: In this case the metal contacts and the oxide dielectric are finely polished and are bonded together in one step. Due to planarity requirement this approach is suitable for C2C or C2W bonding. It is relatively difficult to achieve a wafer-wafer bonding with this method. Sometimes the oxide is replaced with BEOL inter-level dielectric (ILD) based on the bonding methods.

C2) **Hybrid with planar Metal/adhesive:** This method is similar to metal/oxide hybrid but replaces the oxide with an adhesive. The surface planarity in this approach is less stringent than that of oxide/metal hybrid due to the compliant nature of the adhesive [26].

C3) **Hybrid with interlocking metal/adhesive:** Another way to make the hybrid joining surface is to use an interlocking structure. Compared to the planar hybrid structure, the interlocking hybrid structure provides several additional features for wafer bonding. The first is the enabling of the direct metal-metal contact to improve connection yield. The second is the maintaining of registration without slippage during lamination. The third is the facilitation of air evacuation during lamination prior to pressurization. The fourth is the securing of the entire interface for the thinned wafer. In this paper we will focus in more detail on the hybrid option with interlocking adhesive/metal structure bonding. This metal/adhesive interlocking hybrid approach was developed during the earlier work for MCM-D structure made by thin film transfer and joining (TFTJ). In the case of wafer-wafer bonding, areas without bonding will have different stress than areas with bonding (either with metal or dielectric). This is particularly the issue for a 3D structure with a thinned top wafer for high density through-Si-connection. In the case of flip chip C4 connection the stress is maintained on the metal contacts due to the thickness of the chip and the stress concentration can be alleviated by increased C4 heights and the use of underfill. In the case of 3DI, the metal contacts alone can have only a maximum contact area of 25% which concentrates thermal stress and would lead to Si early fatigue fracture. With adhesive between the metal contacts this stress concentration can be distributed to the entire surface. The dielectric between the contacts has another (5th) function. It is to reduce the metal corrosion or migration due to the environment. The adhesive can prevent such issues and potentially increase reliability.

There are several ways to form this hybrid wafer level interconnection. This paper will review one of the methods we have practiced, transfer-join (TJ) method. We will review from its early application in MCM-D to recent results in wafer level 3D integration.

3) **Interface vias connection (IVC) (Imbedded and through-vias):** With oxide and adhesive bonding the via-last process is necessary to form the inter-wafer connections. This type via connection is also called through-via connections. Since the ability to form the TSV also depending on the top Si thickness, via-last is normally a good option for thin top Si (< 20 μm) and for power and ground connections to the bottom wafer with a low resistance and inductance. Theoretically a via-last through-via offers a better wiring pitch as the vias are lithographically defined. In reality, a high density through-via with via-last approach is density limited by the blocked wiring channels in the inter-level-dielectric (ILD) they pass-through and by the via fill aspect ratio. In addition, the wafer-wafer overlay accuracy in wafer bonding limits the pitch of the through-vias. Unless these practical issues are resolved, via-last offers less than the ideally projected through-via density advantages over the via-first approach. For via-first approach, the through-vias are formed as part of the BEOL layer build and are patterned with base wafer lithography and can allow better wiring channels. Since the wafer to wafer connections are made at the interface layer, via-last process is also called imbedded vias. Therefore, via-first is better suited for high density signals for the inter-wafer connections.

4) 3DI as BEOL (W2W), as packaging (C2C), or as a separate discipline:

Current C2C 3D favors adoption of packaging (C2C) approach while W2W 3DI tends to use BEOL-like flow. Both approaches have advantages and limitations. In the case of C2C 3DI, the materials and cost limit their use in board level

final assembly, therefore it is difficult to add additional structure beyond 2 or 3 levels. However, C2C has the known-good-die (KGD) advantage which is critical for early production. Wafer level 3DI in theory can lower the cost and build more stacks. However, the cumulative yield can be a major concern. In addition, to re-introduce the bonded wafers back to BEOL processing limits the material selection for an optimum wafer bonding. Alternatively, using C2W with wider choice of packaging materials in a dedicated line might be a better option for 3DI. This 3DI line would also enable "More than Moore" integration with functions beyond CMOS which will be discussed later.

III. Transfer and join (TJ) for MCM-D and for wafer level 3DI:

1) MCM/SCM Thin film Transfer and join (TFTJ) with face-up and face-down build:

In this section the recent results in transfer and join (TJ) technology from MCM-D and SCM to wafer level 3DI will be reviewed. The wafer level transfer and join (TJ) technology was based on the thin film transfer and joining (TFTJ) technology developed for an IBM high-end thin film multi-chip-module-deposited dielectric (MCM-D) [27]. In order to enhance the MCM-ceramic (MCM-C) wiring density, a low-K polyimide thin film high density and fine pitch wiring plane pair was added to the package [28]. This thin

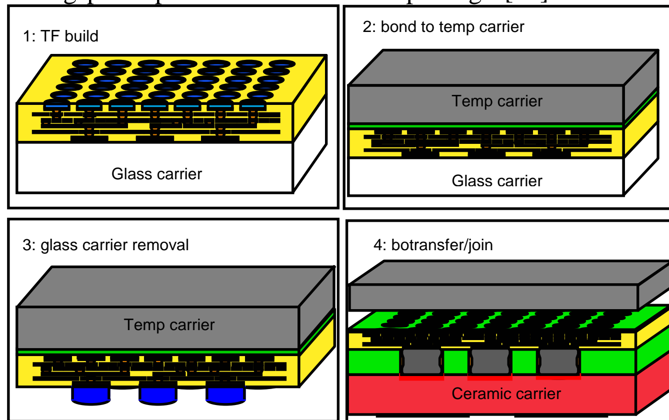

Figure 1: Top-surface up build [35]

film package was originally added to the MCM-C base through a serial build process. In order to simplify this process a parallel build for the thin film (TF) layers with glass plate as carrier was developed [29]. The TF was built on the glass carrier with either a face-up [30] or a face-down [31] build approach. The face-up approach is for complex systems with full level test and repair capability from both top and bottom surfaces of the TF modules. To better manage the film distortion during the build a two-step transfer method was developed. First the face-up TF (top surface up) was built, tested and flip transferred to a temporary glass carrier using a Teflon-PFA adhesion layer. The connecting IOs (bottom surface) are then prepared after the flip and tested for the bonding interface and the TF is then transferred from the temporary glass carrier to the MCM-C base through the bottom surface IO. The Teflon adhesion layer on the top surface also serves as a compliant layer during the lamination to the base MCM-C carrier. This "flip" approach ensures the TF always attaches to a rigid carrier and is the first successful film transfer method on a module level (Figure 1 from [35]). For the simpler single chip modules (SCM) the TF is preferred to build face-down (bottom surface up) and the TF is transferred directly from the glass carrier to the MCM-C base in one transfer step (Figure 2, also from [35]).

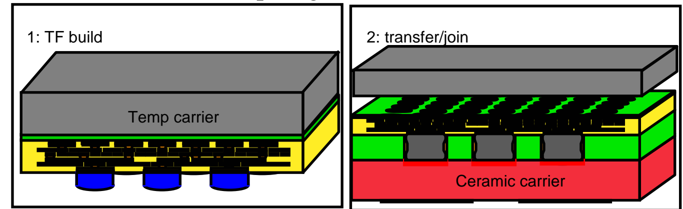

Figure 2: Bottom surface up build [35]

2) MCM/SCM Discrete and integrated interlocking bonding interface:

In the final transfer lamination of the TF to its MCM-C base, two bonding structures were developed. The first one is the discrete spacer structure in which a Kapton-EKJ 200 composite adhesion film (from DuPont) with punched through-vias was used as the bonding structure. Figure 3 shows the schematic of the discrete spacer structure. Figure 4

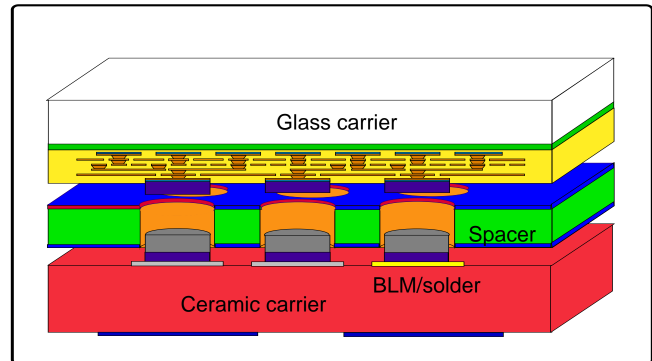

Figure 3: Discrete spacer (from [35])

Figure 4: Modules built with discrete spacer [35]

shows some of the MCM-D modules built with discrete spacer and Figure 5 is a typical SEM cross-section of the spacer structure. The Kapton-EKJ 200 film has a 37 μm kapton core and 6 μm EKJ adhesive on both sides of the core. The IO metal uses a normal C4 Ni ball limiting metal (BLM) and the solder is applied using 400 mesh grade 90Pb/10Sn paste with a glass molder transfer method [32]. The TF solder

balls are inserted into the punched through-vias in the spacer for both alignment and IO connection to the MCM-C module

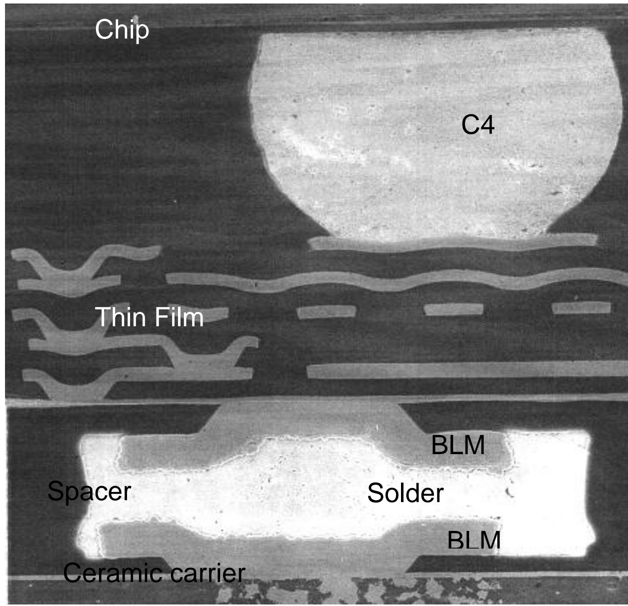

Figure 5: Discrete spacer x-section [35]

by a lamination. The spacer bonds TF to the base MCM-C and functions similar to C4 underfill structure. During the lamination the IO C4 BLM thickness is found to have a profound impact on the TF drift (or slippage) during the bonding lamination. When the BLM is thinner than the adhesive thickness (6 μm) the position of the TF IO BLM would slip out of the through-vias and the IO registration could not be maintained within the spacer through-vias. When the BLM thickness is increased to 12 μm (2x the adhesive thickness) the IO BLM and solder position is always contained within the spacer through-vias. The kapton core of

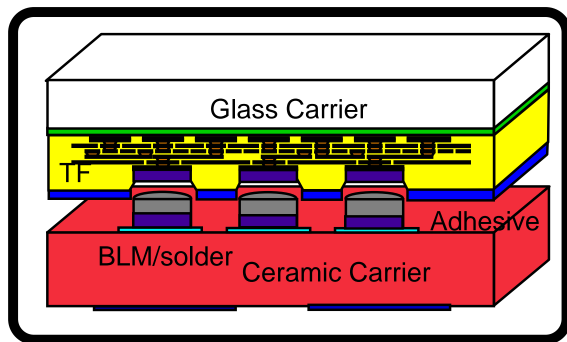

Figure 6: Integrated spacer [35]

the EKJ film and the BLM Ni metal form an interlocking structure when BLM is thicker than 6 μm and it is this interlock that provides C4 IO positioning stability during lamination. To control such a lamination slippage a lock and key interface structure was therefore always provided to confine the metal contacts in their allowed capture IO via area during lamination.

The original interlocking structure in discrete spacer was extended to an integrated spacer structure. Figure 6 show the schematic of the integrated spacer. Figure 7 shows some of the SCM modules built with integrated spacer structure and Figure 8 is a typical SEM cross-section for the integrated structure. In the integrated spacer a metal protrusion is

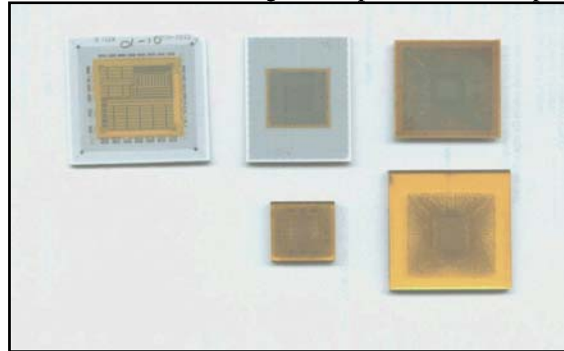

Figure 7: SCM modules built with integrated spacer [35]

provided on one interface. The other interface is a double

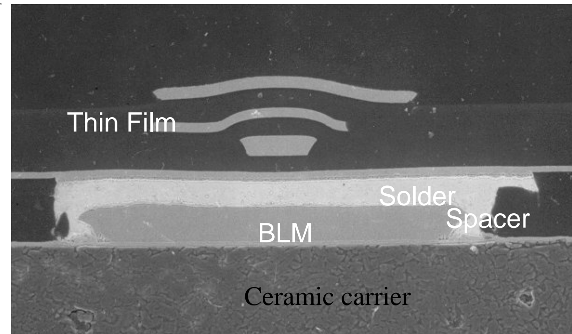

Figure 8: Integrated spacer x-section [35]

layer recess: a fully cured polyimide (such as PI5878G from HD-microsystems) and a thermoplastic polyimide adhesive (such as PI3003X1, also from HD-microsystems). The metal

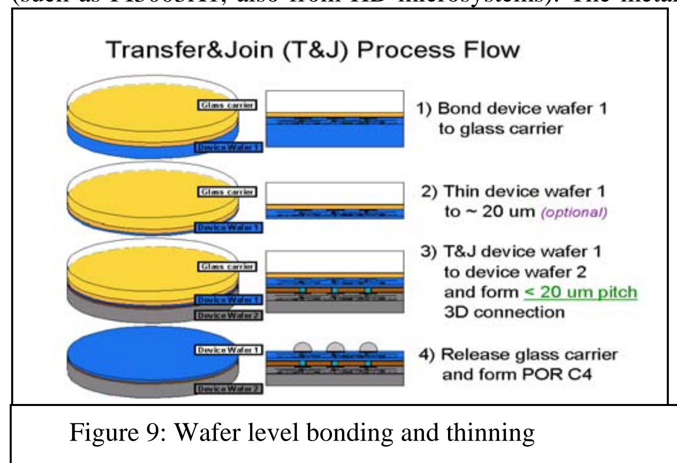

Figure 9: Wafer level bonding and thinning

provide a complete metal contact while allows film to outgas prior to pressurization for bonding. This integrated approach was used for single chip module (SCM) for both ceramic and laminate carriers.

3) 3DI Wafer-level bonding and thinning

The bonding and transferring of polyimide thin film module using glass carrier formed the basis for wafer level 3DI bonding, thinning, transferring, and joining. Figure 9 shows the schematic of extending TFTJ film handling

technology into wafer level bonding, thinning and transfer.

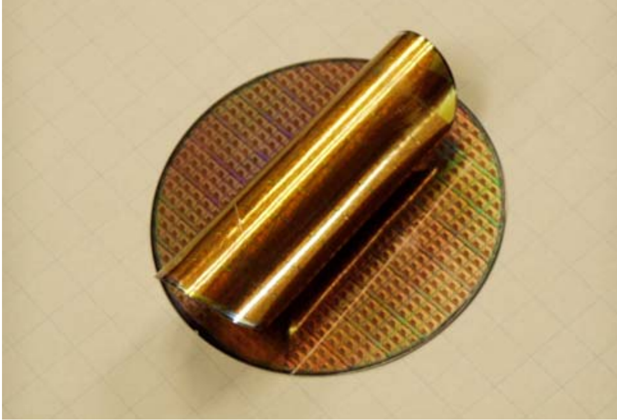

Figure 10: A 5 μm stand-alone wafer membrane [35] .

The process can be accomplished with or without a glass carrier. Figure 10 shows an 8" device wafer initially bonded to a carrier with the adhesive, was then thinned to 5 μm, and is released from the carrier as a stand-alone wafer membrane.

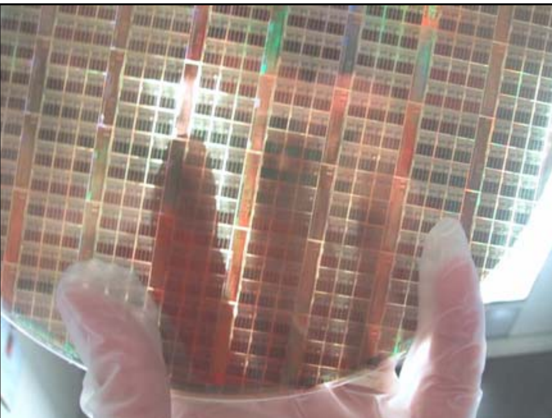

Figure 11: A bonded and thinned device wafer [22]

Figure 11 shows that the bonded wafer can be polished to remove all the bulk Si to its buried oxide (BOX) surface with devices only. Both Teflon-PFA and HD-3003X1 are evaluated for use in lamination at and beyond 300C [33, 35, 37a]. Kapton-KJ adhesive had shown to hold the structure through 370C chip reflow [35]. The works are under way to extend the wafer level bonding and thinning capability to 300 mm wafers.

4) Through-Si-via (TSV)

With the wafer bonding and thinning to sub-20 μm capability, TSV process with high density connection becomes possible. Several ways of deep Si via fill have been evaluated. Initial work showed that a 15 μm CD 45 μm pitch at 20 um deep TSV with electroless Ni is possible [35, 37]. Other TSV metal fills are also possible in this thickness regime [37c]. The through-Si via typically is formed before the wafer bonding (via-first). They can also be formed after the wafer thinning (via-last). Via-first normally allows for a wider material and processing choice as there is less constraint on the process temperature. Via-last, on the other hand, allows the TSV and interface via connection (IVC) form in one step, normally with a better production yield albeit at a lower via density. In most of our work, via-first is used with an imbedded IVC for inter-wafer connection.

5) W2W and C2W interface via connection (IVC):

One of the key 3DI aspects is the ability to interconnect stacked wafers. Thinning wafers to below 20 μm has enabled TSV via-last build which is typically better for power/ground through-via connections. For wiring critical applications the TSV via-first build is preferred for better wiring channel density. The integrated spacer concept developed for the MCM/SCM transfer bonding has been advanced to a finer pitch and was extended to wafer level transfer and joining (TJ) technology [33]. One of the initial via/stud interlocking interfaces demonstrating the alignment and connection

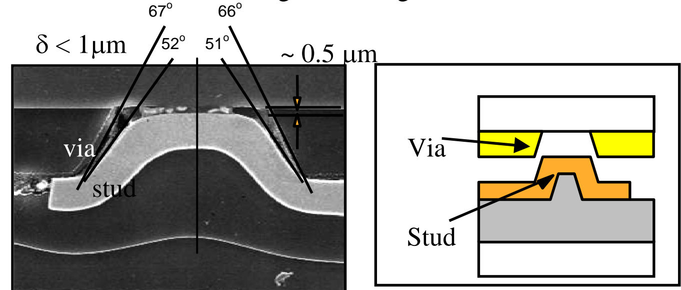

Figure 12: Via/stud (lock/key) alignment and accuracy [33]

capability is shown in Figure 12. One side of the interface is built with Cu protrusions (stud) and the other interface with recesses (via). The overlay accuracy by using the stud/via lock/key structure on a wafer level has been demonstrated to be within 1 μm as shown in the figure. The via-recess is normally a double layer structure with the lower layer as a distortion restrainer and the top layer an adhesive. The via/stud size has be demonstrated ranging from 1 μm to 20 μm with a typical 5 μm CD and a height of 3 to 5 μm in most cases. The distortion restrainer (lower layer of the double layer) can be any material with sufficient rigidity at processing temperature. We typically use a cross-linked

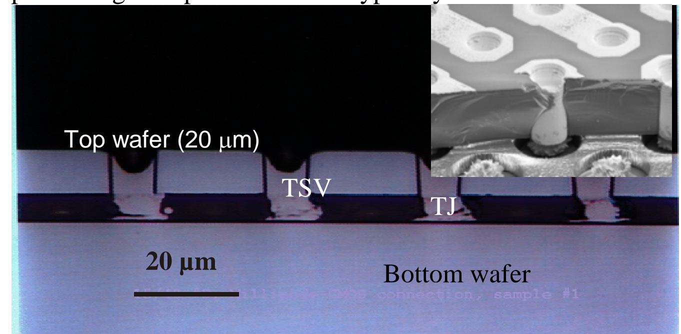

Figure 14: 20 μm thru-Si via-chain [35, 37a]

polyimide or a CVD oxide. The adhesive layer is a modified polyimide film with a Tg and Tm to allow flow during lamination without thermal decomposition [34]. The TJ assembly prior to full lamination has several features that address the major wafer bonding issues. The direct contact of the metal stud/via (lock/key) allows a metal bonding with a locally higher than applied pressure. The double layer recess controls the distortion and alignment during the heat ramp. A slight gap between the two mating surfaces in bonding enables outgas of the moisture trapped on the surfaces and in

the bulk. A less obvious feature of the TJ assembly is that it releases the stress between the interfaces. The small portion of the metal contact allows the two surfaces to remain stress free until bonding takes place. The adhesive absorbs any thermal and mechanical stress and protects the thinned wafer structure. This is very similar to C4 under fill used to spread the stress across the entire surface. Figure 14 shows a complete 20 µm CD 45 µm pitch Ni-TSV via-chain using TJ connects to a base wafer [33, 35, 37].

The bonding strength of the polyimide adhesive is very high. It typically exceeds the Si 4-point bend fracture strength. Figure 15 [from 21, 23, 34c]] compares the adhesion strength for several commonly used bonding surfaces in 3DI. Oxide-oxide is normally in the range of 1-3 J/m^2 while Cu-Cu can be as high as 20-30 J/m^2. However, polyimide adhesive strength is typically beyond 4-point bend measurement

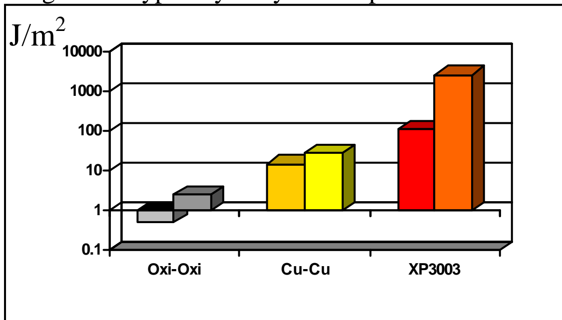

Figure 15: Bonding strength for various materials [21, 23, and 34c]

capability and Si beam break normally results. The adhesion strength is indirectly obtained by using peel test instead, which translates to over 100 J/m^2.

Because of the interfaces are interlocked TJ bonding preferably done in an iso-static vacuum assisted lamination process. Figure 16 shows a schematic of the lamination fixture. A conforming layer of air-tight materials is used as vacuum and pressure seal. Due to the self alignment nature of the lock and key structure, TJ structure has a high alignment accuracy and via connection yield, as mentioned earlier.

The TJ structure with a temporary bonding has been

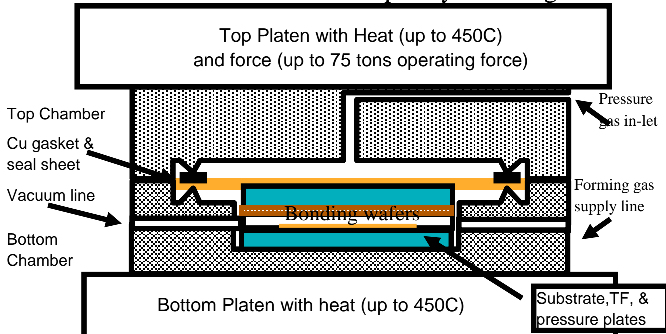

Figure 16: TJ lamination fixture schematic [35]

evaluated for chip-wafer placement for known-good-die (KGD) 3DI bonding at wafer level [35, 37b]. In this exercise we took two wafers with TJ mating structures. One of the wafers is diced into chips and the chips are re-aligned to the un-diced wafer using the TJ via/stud as alignment. The chips are tacked down to the surface wiring of the base wafer temporarily using a micro-tipped tack-down adhesive. After the chips are placed they are laminated to the base wafer and bonded to the base wafer using via/stud both as registration and as IVC via connection. The surface adhesive between the chips and the base wafer form the final permanent bond to secure the chips to the base wafer. In order to enable the polishing of the chip-wafer assembly after bonding, the edge chips, normally discarded after dicing, are put into the assembly as well to reconstruct the uniform surface required for lamination and for wafer thinning. After normal backside grinding step there is no damage to the chip corners if the chips are well bonded to the base wafer. There is no need for dicing channel back-fill and no concern about polish slurry and chemical getting to the joined surface as TJ joining seals the entire contact surfaces.

6) TJ Data and discussion:

Dozens of IBM MCM-D modules have been built by the TJ process and evaluated for construction and reliability, as well as manufacturability. A fully functional MCM module with 30 chips was assembled and that module was put in user condition for an extended period for field performance study. The reliability of the joining as well as the potential electrical impact in system level was evaluated. No performance disparity, due to the joint structure, was noted. The spacer thickness was found to have significant impact on the joints reliability. The thickness reduction from a 75 um spacer (EKJ320) to a 50 um spacer (EKJ200) significantly enhanced the thermal cycle life. TFTJ was also extended for use in SCM. 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. Based on the data from MCM modules SCM on alumina ceramic base carriers the spacer thickness below 20 µm was studied and no stress failure was reported in thermal cycles within the module expected lifetime. 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 add 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.

For wafer level 3DI application the TJ joining was refined based on TFTJ and was developed to enable a much finer pitch and registration accuracy. The typical applications are for join pitches below 50 µm. This tight pitch is designed for high density, high bandwidth 3DI applications. Figure 17 shows such a transfer of interconnects 10 µm lines on 20 µm pitch. There are two sets of wiring in the stitching pattern.

The x-wirings are formed on the base wafer. The ends of the

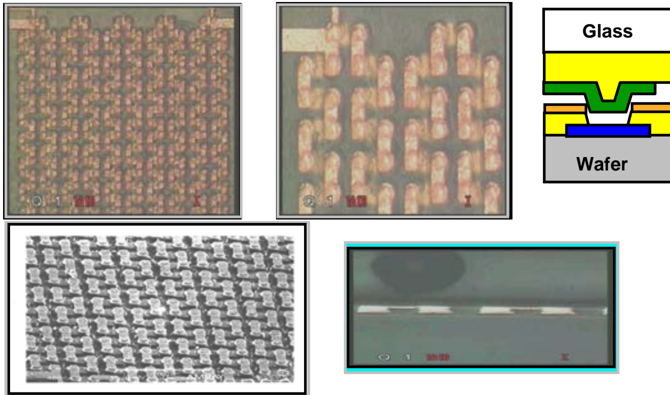

Figure 17: 10 on 20 μm TJ transferred via-chains [33,35]

x-wirings are protruding studs of 7 μm in height. The y-lines are formed on a transferring wafer or glass with the y-wiring

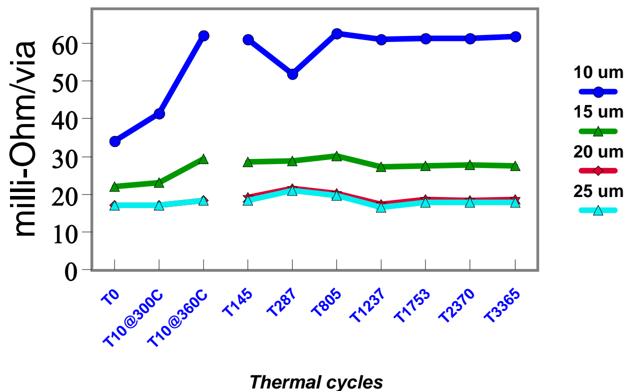

Figure 18: Chip join reflow and Thermal cycles results of the transferred TJ via-chains [33, 35]

ends with a recess of 5 μm. The x and y lines from both wafers are brought together with a lamination. The carrying wafer for the y-wirings is then removed. The y-wirings are then transferred to the base wafer and connect to the x-wirings and form via-chains. The transferred structures have the stitch sizes of 10, 15, 20, and 25 μm. After the transfer the joining yield was assessed to be 100% in 72 such transfers in each via size with 352 stitches in each chain. The transferred structures were then subjected to 13 times chip reflow cycles at 300C, then 350C to simulate module rework. The via-chain resistance is measured after the cycles are shown in Figure 18. The 10 μm stitch was found to have a resistance per via increase from 30 milli-ohms to 60 milli-ohms during the 350C reflow. For other via sizes the resistance is stable. The stitches were then put through 3650 cycles of 25 C to 150C thermal cycles for thermal fatigue test. The via-resistance values are plotted as a function of thermal stress cycles time. There is no appreciable change in the via-chain resistance during thermal cycles.

For chip-wafer 2D applications, the wiring resistance yield across the space between the chips dicing channels are measured before and after the lamination. No noticeable resistance change and yield loss was found for line width down to 5 μm.

The general approach described can in principle be extended to larger wafer sizes by using appropriate tooling option.

7) "More than Moore" applications

The ability of TJ technology to re-assemble chips into a pseudo-wafer with high registration accuracy enables the technology to be applied to much wider application spaces beyond CMOS scaling. ITRS roadmap anticipates that future CMOS advancement to follow "More Moore" and "More than Moore" parallel paths. TJ has been demonstrated a viable technology for "More Moore" CMOS 3DI applications. It is equally suitable for "More than Moore" integration. We will show some of the demonstrated applications with TJ:

A) **MEMS:** Using the same principle outlined in the previously discussions the micro electro-mechanical systems (MEMS) can be assembled accurately with CMOS devices.

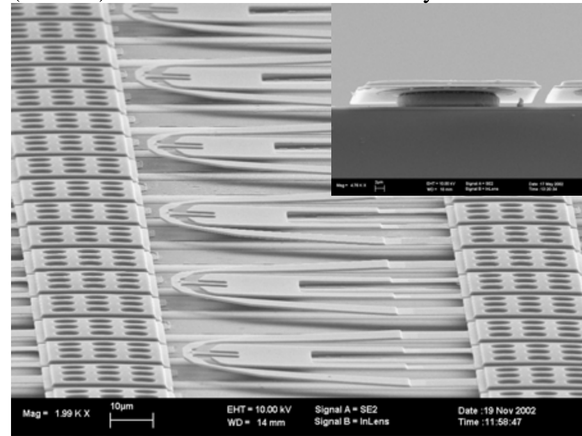

Figure 19: TJ for MEMS transfer [37a, 37b]

Millipede is an IBM project using nano-size dots as a recording device for high density low power mobile storage device. The central focus of the integration is to use a thermal-electrical logic device to drive micron scale mechanical levers for surface data read and write. TJ technology is used here to transfer the high density mechanical levers from one wafer to the surface of a CMOS wafer. Figure 19 shows the wafer level transfer of the levers with the levers sitting on the TJ studs. A high transfer and joining levers yield (>95%) and high planarity uniformity (20 nm / 100 μm) was demonstrated and the device successfully enters data system test [37a and 37b].

B) **Opto-electronics**: Optical interconnect plays a major role for box-box communication. With the advancement of integration it is conceivable that they might replace some of the on-chip wiring to increase the data bandwidth. One of the issues to incorporate optical channels into the electrical system is the placement accuracy and the fine size of the fiber channels with respect to the light emitters. TJ technology was demonstrated to be able to resolve both issues. Figure 20 shows an array of fine pitch optical channels formed by using polyimide conduits with metal reflector to form total reflection [38]. The channels are about 10 μm in size and 25 μm in pitch. Similar channels

using oxide on Si wafers can be made similar to the Millipede MEMS device. Figure 20 also shows a schematics of how a

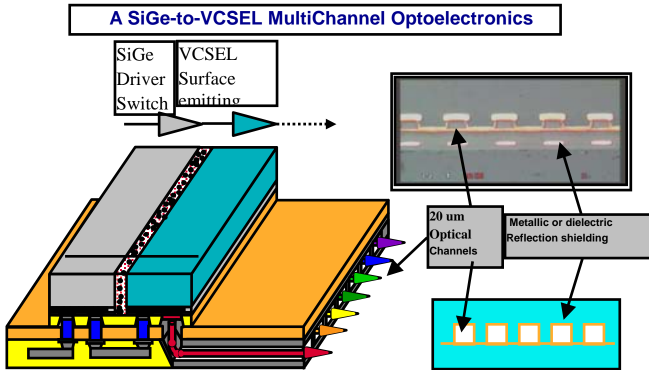

Figure 20: TJ for opto-electronics [33,35]

CMOS (SiGe) high speed logic and the laser drive (VCSEL) and the optical channels can be positioned accurately to send the laser signals into the multi-channels optical bus. The

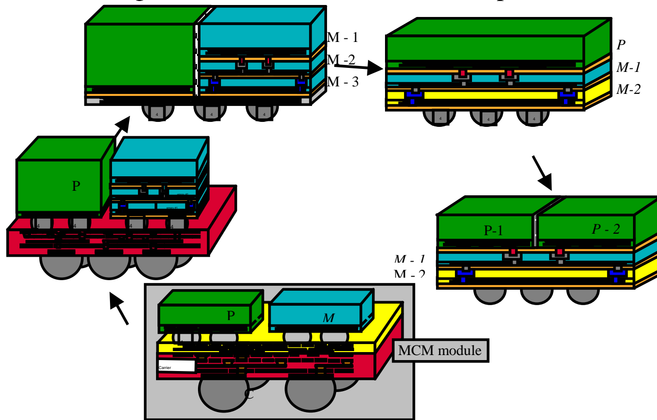

Figure 21: TJ enablement of hybrid systems [35]

receiving end of the optical channels will be a light detector device with the output to drive a logic CMOS, SiGe, or a III-V device. With the readily available TSV technology the optical channels can be used with vertical laser from wafer-wafer through 3DI connections.

C) **Hybridized device**: With the ability to integrating device both in 2D and 3D, it is possible now to construct a high density

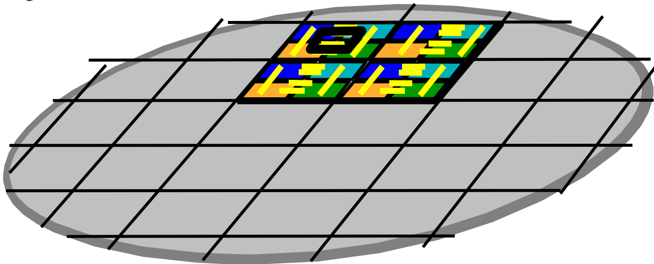

Figure 22: Wafer level Chip hybridization [33, 37b]

device in a wafer level with either chip-wafer or wafer-wafer to integrate SiGe, GaAs, III-V, logic/Memory in variety of ways. Figure 21 schematically shows some of the potential combinations of the 2D/3D to achieve the best circuit speed and or special functionality. Due to the material and substrate size compatibility issues the different materials normally can not be processed sequentially on the same wafer. With the T&J method the wafers can be separately processed and diced and then recombined on to a base carrier with proper connection as in discussed and illustrated schematically in Figure 22 (see [33, 37b, 39-42]).

IV. Summary:

The MCM-D thin film transfer and join (TFTJ) technology and the related wafer level transfer and joining (TJ) technology are reviewed. These integration technologies can enable high density fine pitch 2D and 3D integrations at wafer-wafer and chip-wafer level for “More Moore” scaling. The advances are also applicable for “more than Moore” applications in MEMS, opto-electrinics, hybrid devices. These advances can ultimately bring SOC and SIP together, achieving SOC performance with SIP versatility for the future CMOS.

Acknowledgments

The Author wishes to thank Dr. S. Purushothaman for his critical review and comments of this manuscript.

V. References

[1] G. Moore, Electronics, V38, pp114, 1965

[2] 2007 ITRS, http://www.ITRS.org/

[3] S. Joshi, et al, DRC2007 pp.53 – 54

[4] T.C. Chen, ICSICT2006, pp4-7.

[5] G.G. Shahidi, IEEE-CICC 2007, pp413-416.

[6] G. Amdahl, AFIPS Conference Proceedings, (30), pp.483-485, 1967.

[7] J.E. Fritts and E.D. Chamberlain, proc. 35th Simulations Symposium, 2002. Page(s):352 - 362

[8] S. Kapoor; HLVDT2007. Page(s):48 - 52

[9] B. stolt et al, V43No.1, Jan. 2008 pp21 – 28, JSSC.2007.

[10] S.R Vangal et al, V43No.1, 2008, pp29 – 41, JSSC.2007.

[11] L. Chai, et al, CCGRID2007. pp471-478

[12] R. Bergamaschi et al, ASPDAC2008. pp.708-713

[13] P. Jacob et al, IEEE Design&test, V22,No.6, 2005, pp540-547

[14] P. Leduc et al, Proc. IEEE (IITC), pp. 210-212, 2007.

[15] P.S. Andry, et al, Proc. of the 56th ECTC, pp. 831-837, 2006.

[16] P. R. Morrow et al, *IEEE Elect. Dev. Lett.* **27**, 335-337 (2006).

[17a] Val, C.; Lemoine, T.; IEEE Trans. on Components, Packaging, and Manufacturing Technology, Part A, B, C] V13, I4, Dec. 1990 Page(s):814 - 821

[17b] M. Karnezos; IEMT.2004. pp64-67

[18] K. Takahashi, et al; ECTC Vol.1,2004 Pp.601 - 609

[19] E. Beyne; VTSA.2006.pp1-9

[20] S. C. Jonson, *Semiconductor International*, pp. 40-45, June, 2007.

[21] K.-N. Chen, et al, *IEDM Tech. Digest*, 20-22 (2006).

[22] K. W. Guarini, et al, *IEDM Tech. Digest*, 943-946 (2003).

[23] J.A. Burns, et al, IEEE tran. Elec. Devices, V53, No.10, pp2507-2515, Oct. 2006.
[24] M. Wiemer et al, ESTC2006 pp.1401-1405
[25] D. Temple, et al, IEDM2006, pp:1 - 4
[26] R.S. Patti, Proc. Of IEEE V. 94, I6, June 2006 Pp1214 – 1224
[27] G.A. Katopis, et al, IBM J. Res.develop. Vol 43, No. 5/6, (1999), p621.
[28] K. Prasad and E. Perfecto, IEEE trans. Comp. Pkg. Mfg. Techno. Part B, Adv. Pkg., Vol 17, No.1 (1994) p38.
[29] C. Narayan, et al., IEEE trans. Comp. Pkg. Mfg. Tech., Part B, Vol 18, Issue 1, 1995, p42
[30]United States Patent 6,281,452, Prasad , et al. August 28, 2001
[31] E. Perfecto, et al, IEEE trans.Adv.Pkg., Vol 23, No.3 (2000), p1.
[32] United States Patent 6,099,935, Brearley , et al. August 8, 2000.
[33] H. B. Pogge et al., Proc. AMC'2001, pp.129-136,2001;
[34a] K. Kanakarajan and John A. Kreuz, US patent 5298331 (1994).
[34b] R. K. Kanakarajan, IPC national conf. On Flex Circuits, 1995, Minneapolis, MN
[34c] K. Patel et al, Proc. 10th Symposium on Polymers for Microelectronics, May 2002, Wilmington, DE, Session III-3.
[35] R. Yu, proc. VMIC2007, p223-230.
[36] F. Liu, et al, to be published.
[37a] 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
[37b] R. Guerre, et al, JMEMS, 2007, V17, No1, p157-165.
[37c] M.J.Wolf et al; ECTC2008.Page(s):563 – 570 via fill
[38] United States Patent 6,640,021, Pogge , et al. October 28, 2003
[39] United States Patent 6,444,560, Pogge, et al. September 3, 2002
[40] United States Patent 6,090,633, Yu , et al. July 18, 2000.
[41] United States Patent 6,835,589, Pogge , et al.
[42] United States Patent 6,856,025, Pogge , et al. February 15, 2005
[43]D. Gerty, (THERMINIC 2007), Budapest, Hungary, Sep. 17-19, 2007.