Chip Scale Review®

ChipScaleReview.com

The Future of Semiconductor Packaging

Volume 28, Number 1 January • February 2024

Heterogeneous chiplet integration to make megachips

- Making the golden connections in 3D-IC system-level design
- Backside power delivery: The new frontier for wafer bonding
- Exploring bond strength for an advanced chiplet with hybrid bonding
- Readying the supply chain for chiplets and heterogeneous IC packaging
- The great lithography debate: Copper clad laminate or glass substrates?
- Scaling up GaN and InP-based technologies for 5G and 6G wireless communication

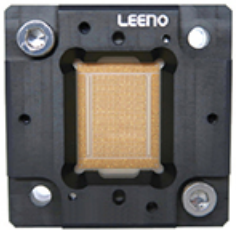
LEENO

www.leeno.com

10K Pins in 7 Days

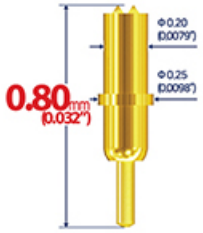
Fine Pitch Probe & Probe Head

Proven Mass Production Capability

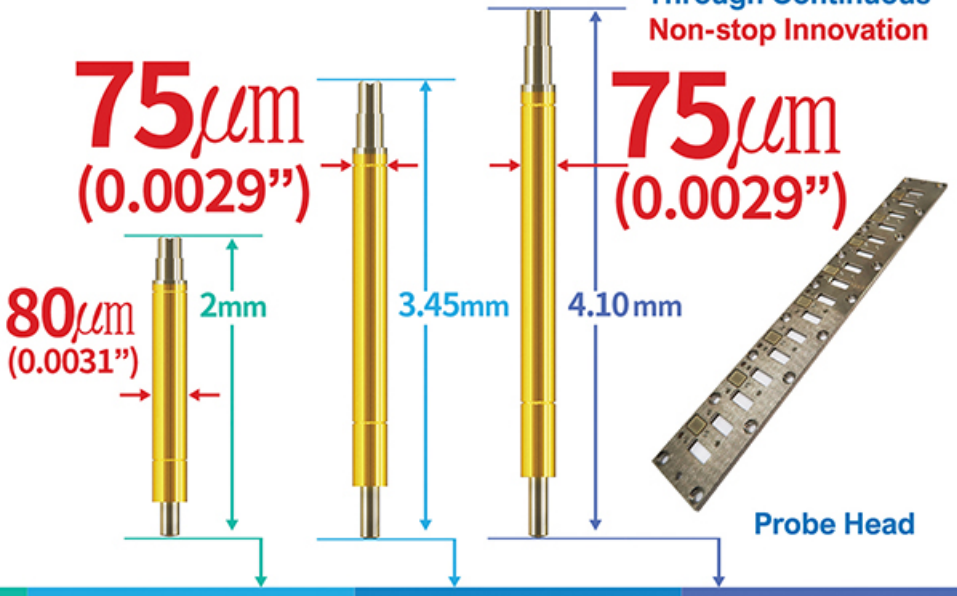


High Speed RF

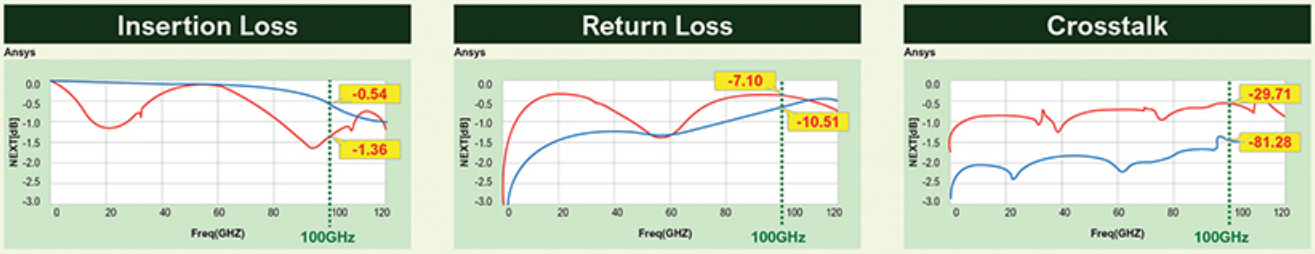
Short Signal Path 0.80mm with Spring Contact Probe!



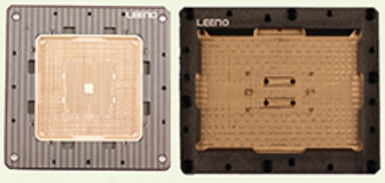
Through Continuous Non-stop Innovation



5G MP 100GHz Coaxial Socket



▪ Pitch: 0.35mm, Array



5 Sockets in 10 Days

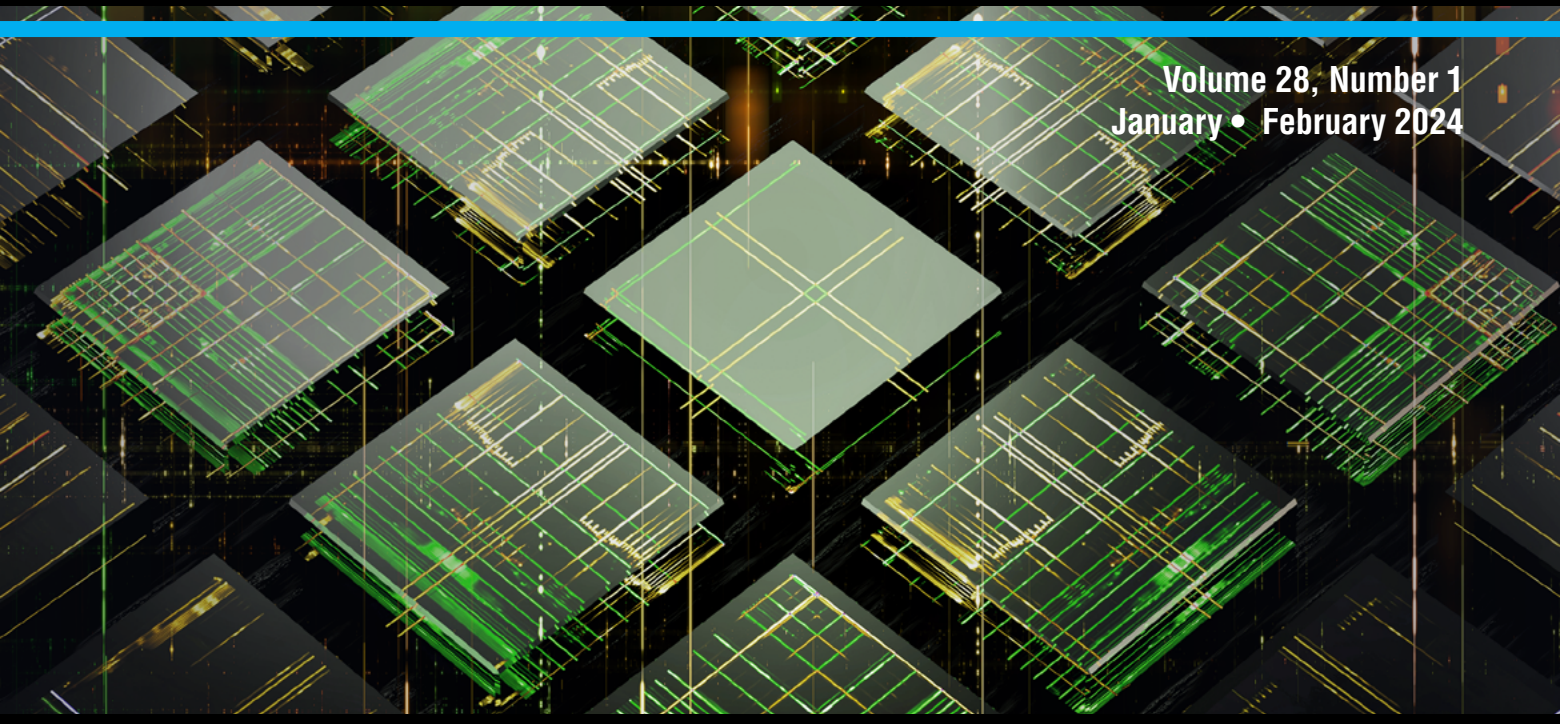
GLOBAL LEADER LEENO

HEAD OFFICE
10 105beon-gil MieumSandan-ro Gangseo-gu, Busan Metropolitan City, Korea

CONTACT
hskang@leeno.co.kr / +82 10 8737 6561
kimmj@leeno.co.kr / +82 10 5314 6560

CONTENTS

Volume 28, Number 1
January • February 2024



Heterogeneous chiplet integration to make megachips 13

The megachip approach helps to rearchitect heterogeneous chip tiling for developing highly complex systems having desired circuit density and performance. Recent work on large-area superconducting integrated circuits to join multiple individual die is highlighted in this article, with particular attention paid to the processing of the high-density electrical interconnects formed between the individual die.

Cover image courtesy of iStock/NiPilot

DEPARTMENTS

EXECUTIVE VIEWPOINT

7 Backside power delivery: The new frontier for wafer bonding
By Paul Lindner [EV Group]

TECHNOLOGY TRENDS

10 The great lithography debate: Copper clad laminate or glass substrates?
By Doug Brown [Onto Innovation]

FEATURE ARTICLES

13 Heterogeneous chiplet integration to make megachips
By Rabindra N. Das, Jason Plant, Alex Wynn, Matthew Ricci, Ryan Johnson, Matthew Stamplis, et al. [MIT Lincoln Laboratory]

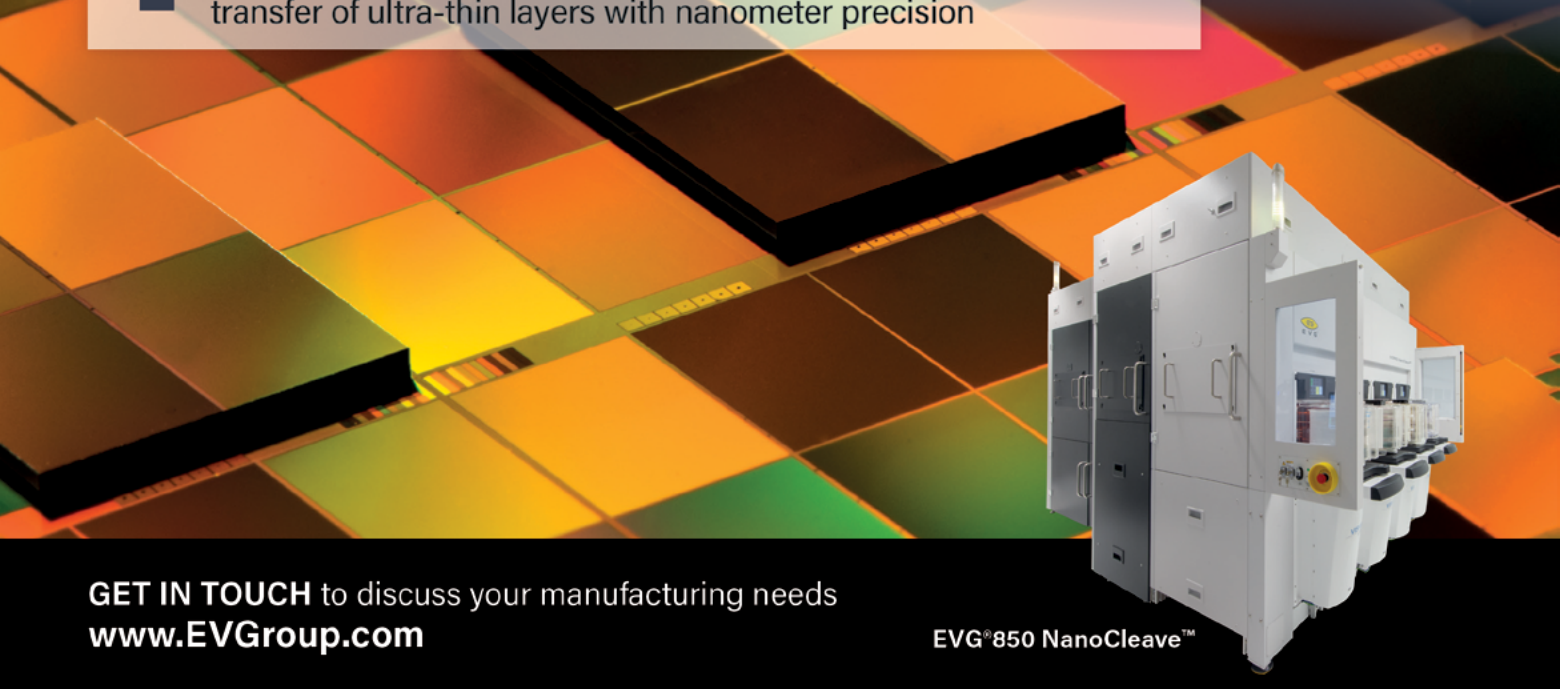
EVG®

invent
innovate
implement

www.EVGroup.com

FUSION AND HYBRID BONDING FOR HETEROGENEOUS INTEGRATION

- Enabling advanced 3D device stacking and chiplet integration for CMOS image sensors, memory and 3D system-on-chip (SoC)
- High-volume production equipment and process technology for Wafer-to-Wafer (W2W) and Die-to-Wafer (D2W) hybrid bonding
- GEMINI® FB Automated Production Wafer Bonding System delivers industry leading alignment accuracy and bonding performance
- Revolutionary NanoCleave™ technology enables release and transfer of ultra-thin layers with nanometer precision



GET IN TOUCH to discuss your manufacturing needs
www.EVGroup.com

EVG®850 NanoCleave™

CONTENTS

Chip Scale Review®
ChipScaleReview.com
The Future of Semiconductor Packaging

STAFF
Kim Newman
Publisher
knewman@chipscalereview.com

Lawrence Michaels
Managing Director
Editor-in-Chief
lmichaels@chipscalereview.com

Debra Vogler
Senior Technical Editor
dvogler@chipscalereview.com

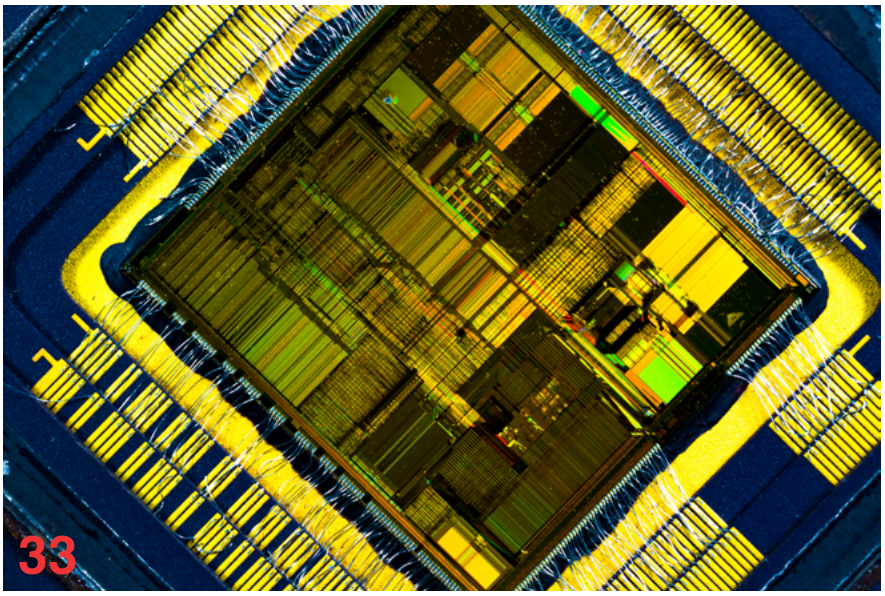
SUBSCRIPTION—INQUIRIES
Chip Scale Review
All subscription changes, additions, deletions to any and all subscriptions should be made by email only to
subs@chipscalereview.com

Advertising Production Inquiries:
Lawrence Michaels
lmichaels@chipscalereview.com

Subscriptions in the U.S. are available without charge to qualified individuals in the electronics industry.
Chip Scale Review is published six times a year with issues in January-February, March-April, May-June, July-August, September-October and November-December.

P.O. Box 2165
Morgan Hill, CA 95038
Tel: +1-408-846-8580
E-Mail: subs@chipscalereview.com

Printed in the United States



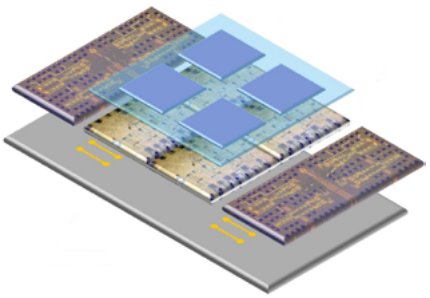
FEATURE ARTICLES (continued)

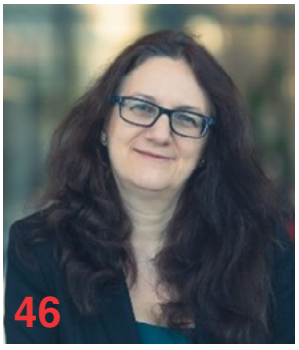
25 **Readying the supply chain for chiplets and heterogeneous IC packaging**
By Mike Kelly, Dave Hiner, Ruben Fuentes, Jonathan Micksch, Vineet Pancholi [Amkor Technology, Inc.]

33 **Making the golden connections in 3D-IC system level-design**
By Tarek Ramadan [Siemens EDA]

39 **Exploring bond strength for an advanced chiplet with hybrid bonding**
By Junya Fuse, Tomoya Iwata, Yuki Yoshihara, Marie Sano, Fumihiro Inoue [YOKOHAMA National University]

46 **Scaling up GaN- and InP-based technologies for 5G & 6G wireless communication**
By Nadine Collaert [imec]

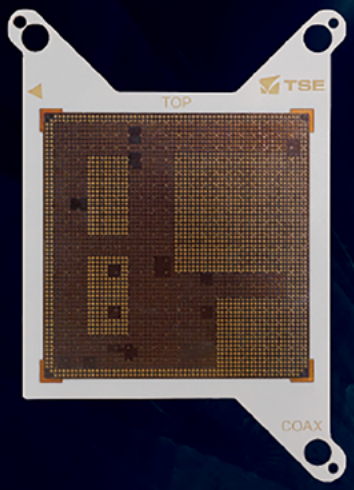




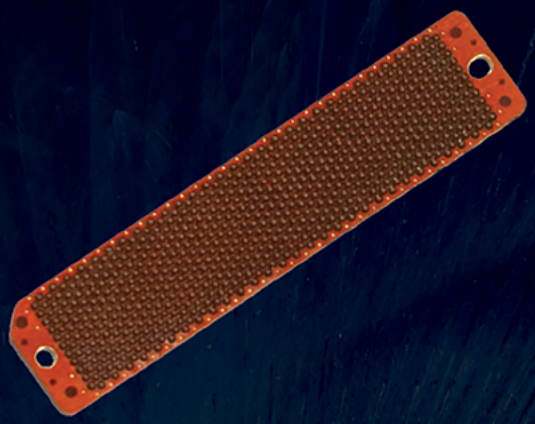
TSE

COAXIAL ELASTOMER SOCKET

for >64Gbps ATE/SLT Test
for Crosstalk free Board to Board Connector



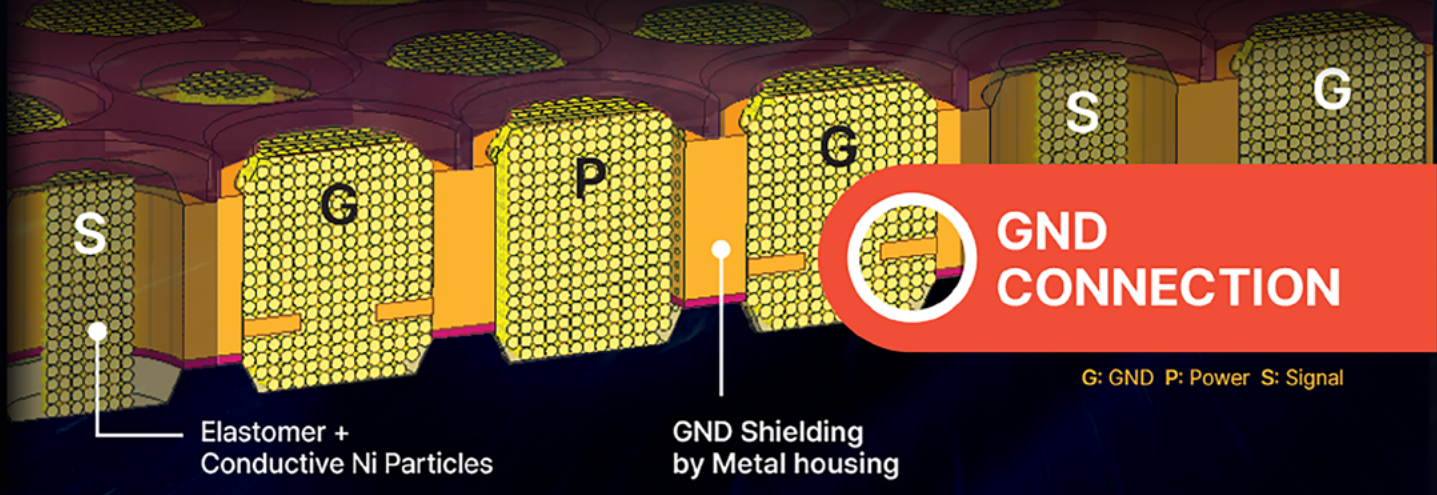
Semiconductor Package Test



Board to Board Interconnection

ELTUNE-coax™

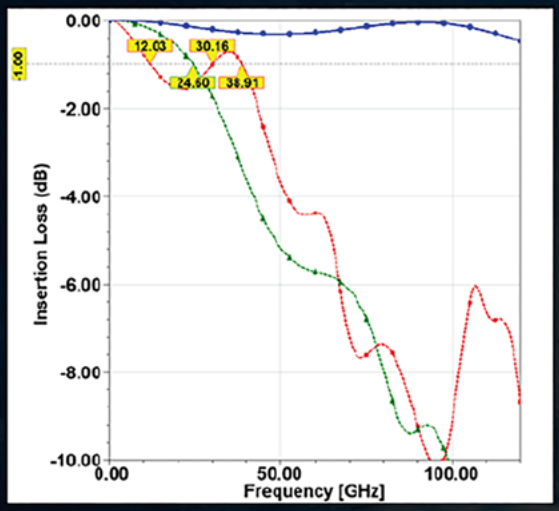
Metal GND Structure
Extremely Low Crosstalk >100Ghz@-20dB
Inductance <0.1nH
Min. pitch 0.6mm



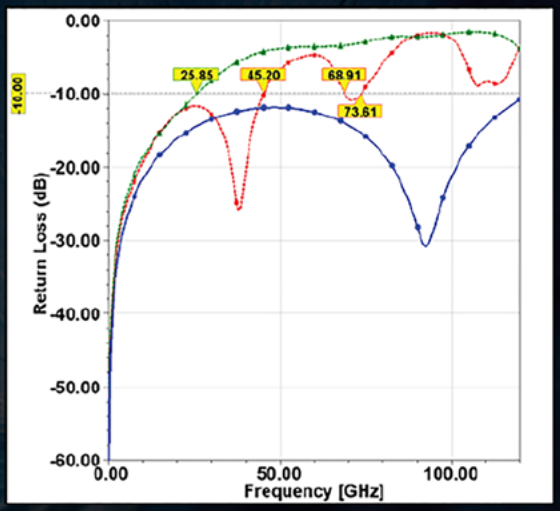
TSE

ELTUNE-coax™

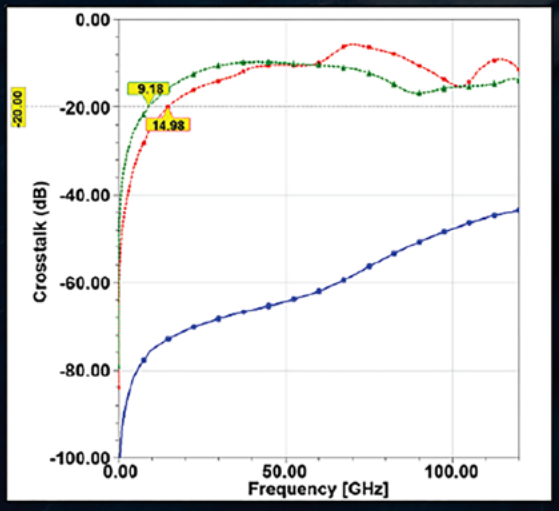
Electrical Specifications (unit: GHz)			
50Ω, 0.80mm pitch	Spring pin	Elastomer	ELTUNE-coax™
Electrical Length(mm)	3.05	0.60	0.60
Insertion Loss(S21) @-1dB	12.03	24.60	>100
Return Loss (S11) @-10dB	45.20	25.85	>100
Crosstalk (S31) @-20dB	14.98	9.18	>100



Insertion Loss



Return Loss



Crosstalk

TSE USA

Jong M. Lee jmlee@tse21.com
Kevin Kim kevin.kim@tse21.com

TSE HQ

Justin Yun justice@tse21.com
Jimmy Hwang jimmyhwang@tse21.com

ELTUNE-coax™
Elastomer
Spring Pin

www.tse21.com
189, Gunsu 1-gil, Jiksan-eup, Seobuk-gu,
Cheonan-si, Chungnam, 31032, Korea

Amkor Technology®

Enabling the Future

Powering Tomorrow with **Heterogeneous** & **Chiplet** Packaging Solutions

Amkor's heterogeneous packaging combines key technologies and increased chiplet integration to enable higher performance required by emerging technologies. Our innovative test solutions help deliver quality and reliability for all your heterogeneous packaging needs.

Backside power delivery: The new frontier for wafer bonding

By Paul Lindner [EV Group]

3D integration stands as a pivotal technology in advanced packaging and heterogeneous integration—it facilitates performance scaling at the system level. While the evolution of packaging has introduced 3D integration, progressing from systems in package to stacked integrated circuits (ICs) and 3D systems on chip, the industry is currently witnessing another significant turning point: backside power delivery networks (BSPDN).

In the traditional scaling approach, both signal and power distribution coexist on the frontside of the wafer. However, the growing demand for power, particularly in power distribution, is increasingly constraining the ability to achieve scalable solutions. Efficient transistor scaling, crucial for achieving higher transistor density, necessitates a proportional scaling of the power distribution network. This, however, encounters significant IR drop challenges, leading to detrimental transistor performance. Additionally, the interconnect design for signal and power is becoming highly interdependent, constituting a substantial portion (at least 20%) of the routing processes for power delivery. Moreover, power density experiences a rapid increase with scaling to the next node.

The industry consensus is to decouple signal and power by implementing BSPDN. This involves segregating the signal network on the frontside of the wafer and utilizing wafer-to-wafer bonding to efficiently access the transistor backside for power distribution and management. The key advantages include wider power lines with reduced IR drop, a more uniform voltage distribution, and, most significantly, more design space, thereby enabling further scaling of the standard cell height.

BSPDN eliminates the need to share interconnect resources between signal and power lines on the wafer frontside. As implied by its name, backside power delivery relocates power to the back of the wafer, leaving only signals to be transmitted through frontside interconnects. The consensus within the industry is that BSPDN will be implemented in nodes below 2nm, making it the industry standard in the upcoming advanced nodes. This transition is expected to commence in 2024, initially in client compute and server applications.

BSPDN manufacturing steps

The integration of BSPDN offers different approaches, each with unique advantages. In the first method, the logic cells have a buried power rail (BPR) to which the BSPDN is connected via a nano through-silicon via (TSV). The second approach leaves the power rail but instead uses a power via to transfer power directly from the BSPDN to the cell or transistor contact. Although this method is more complicated, it improves power efficiency and allows for better cell area scaling. The third approach involves a direct connection of the power supply from the BSPDN to the source and drain contacts of the individual transistors.

In the first BSPDN implementation, BPRs are created during the front-end fabrication of the device. BPRs are defined after a shallow trench isolation, which typically has a width of ~30nm and a pitch of ~100nm. The metal is then recessed and covered with a dielectric, followed by processing of the IC on the wafer frontside, with a copper metallization back-end-of-line (BEOL) signal network completing the front-side processing.

In the next step, the wafer with the devices and BPRs is flipped, with the active frontside bonded to a blanket carrier wafer. This bonding is achieved by dielectric fusion bonding at room temperature, followed by a post-bond anneal. The backside of the first wafer is then thinned in a sequential process of backside grinding, chemical mechanical polishing (CMP), and dry and wet etching steps.

After depositing a passivation layer on the backside, nano TSVs are patterned from the backside of the wafer using through-silicon alignment lithography. These nano TSVs, which are etched through several hundred nanometers of silicon, end at the tip of the BPRs. The process concludes with the implementation of one or more backside metal layers that electrically connect the backside of the wafer to the BPRs on the frontside via the nano TSVs.

Wafer bonding co-integration with lithography processes

Achieving low-distortion wafer-to-wafer bonding is critical to enabling BSPDN. Because nano TSV contacting of BPRs or later versions of BSPDN requires sub-10nm overlay accuracy to the buried frontside structures, there must be compensation by the lithography equipment for any incoming distortion from frontside processing and wafer bonding. This is accomplished through the co-optimization of wafer bonding and lithography.

It is noteworthy that for most applications, one of the wafers in the bonded stack undergoes thinning to a scale of several micrometers for subsequent processing. This thinning process induces a relaxation in the bonded stack, aligning it closely with the original shape of the thicker wafer. Several strategies can be proposed for lithography-based bonding overlay and distortion control. One such strategy involves the bonding of a patterned wafer to an unpatterned counterpart. In such instances, concerns regarding overlay between the wafers are mitigated because one of the wafers remains blank. However, the impact of bonding-induced distortion assumes prominence, especially considering that, post-bonding, the patterned wafer is typically thinned for continued processing with lithography from the backside of the device wafer. Consequently, any bonding-related distortion for which there is no compensation by a lithography scanner, significantly contributes to the overall overlay.

Two distinct approaches have been considered for controlling bonding fingerprints. The first relies exclusively on the actuation capability of the bonding tool, while the second requires concurrent actuation by both the bonding tool and the lithography scanner. The overlays resulting from these approaches are called "post-bond overlay" when exclusively achieved with the bonding equipment, and "post-litho overlay" when lithography correction is also needed.

Wafer bonding optimization

The intricate understanding of bond wave propagation and its influence on overlay underscores its strong dependence on various factors, encompassing the mechanical behavior and stress characteristics of the bonding wafers. Equally crucial is the direct bond interface, incorporating aspects such as via pitch, density, and nano topography variations.

Wafer parameters, including frontside and backside nano topography and the wafer bevel, are recognized as pivotal impact factors on process repeatability. Throughout the bonding process, accurate monitoring and control of these parameters, coupled with contributions from the bonding process itself, become imperative. Control over bond wave propagation and speed in distinct sections of the interface offers a means to manage the naturally occurring asymmetric and somewhat nonlinear bond wave behavior, thereby enabling compensation. These control parameters are equally instrumental in regulating local overlay performance.

The speed of the bond wave, contingent on angular-dependent acceleration, exerts a direct influence on post-bond scaling and residuals. Through the real-time monitoring and statistical evaluation of bond wave behavior, a correctional algorithm can be devised. A clear correlation emerges between bond wave speed and resulting residuals, with a slower bond wave correlating with lower residuals. The integration of these well-known process control parameters facilitates the establishment of an automated control system. Leveraging live data alongside statistical insights enables the regulation of both bond wave propagation speed and uniformity, thereby ensuring repeatable processes with reduced residual stress.

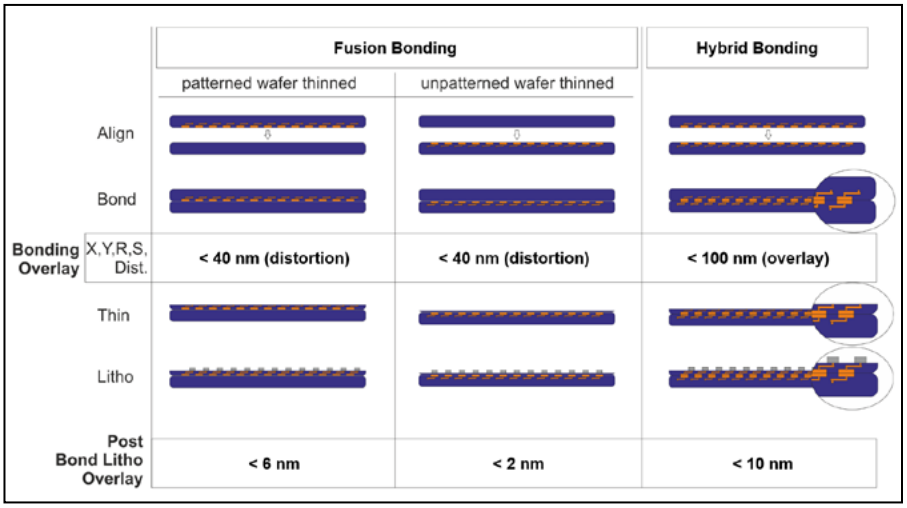


Figure 1: Process co-optimization of wafer bonding and subsequent lithography for different fusion and hybrid bonding integration flows. The figure also shows post-bond overlay at the bonding interface and post-bond litho overlay that describe contacting accuracy using the latest lithography compensation strategies. SOURCE: EV Group

Given that the resulting overlay, irrespective of bond overlay or post-bond litho overlay, is contingent on all processes involved (not just bonding), the importance of modeling potential overlay errors becomes evident. In-plane distortions (IPD) arising from silicon processing play a pivotal role in determining the maximum achievable overlay. Consequently, the systematic gathering and comparison of IPD data at various stages, considering potentially constant factors like wafer chucks, becomes essential. These factors can then be corrected systematically or through model-based approaches. The amalgamation of gathered IPD data with measured bond overlay facilitates post-bond litho overlay compensation through the application of appropriate compensation models on a scanner. This holistic approach ensures comprehensive control and optimization across all stages of the intricate semiconductor manufacturing process (Figure 1).

The future is bright for wafer bonding

Wafer bonding is expected to become increasingly significant to the advancement of the semiconductor industry in the coming years. Look no further than complementary field-effect transistor (CFET) architectures on the logic technology roadmap where wafer bonding emerges as a viable approach for uniting the n-type metal-oxide-semiconductor (n-MOS) with the p-type metal-oxide-semiconductor (p-MOS) through a direct bond. This facilitates the independent growth of n-MOS and p-MOS via epitaxy, exemplifying the potential of wafer bonding in enhancing semiconductor architectures. However, the benefits of wafer bonding extend beyond monolithic or quasi-monolithic integration. The bonding of chiplets, among other technologies, is expected to become increasingly supported by wafer-level processes, reflecting the evolving landscape of semiconductor manufacturing.

Biography

Paul Lindner is Executive Technology Director at EV Group, St. Florian am Inn, Austria. He heads the R&D, product and project management, quality management, business development and process technology departments. Customer orientation throughout all steps of product development, innovation and implementation in a production environment are among the main goals of the groups headed by Lindner. Since joining EVG in 1988, he has pioneered various semiconductor and MEMS processing systems, which have set industry standards. E-mail: P.Lindner@EVGroup.com

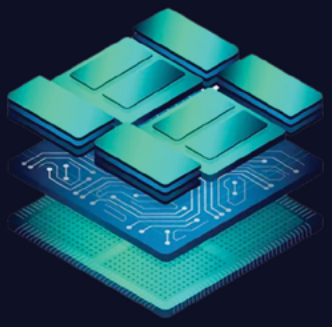
SIEMENS

Facing **semiconductor scaling** challenges?

These **5 key workflows for 3D IC** deliver the heterogeneous integration (HI) of chiplets necessary for 2.5D/3D design. Explore and deliver product differentiation faster using Siemens EDA's market-leading 3D IC technology solutions for 3D heterogeneous integration of node and performance-optimized chiplets.

1. **Architectural Planning & Analysis**
2. **Physical Design Planning & Analysis**
3. **Design Analysis**
4. **Reliability Analysis**
5. **Test Planning & Validation**

Your complete solution for 2.5/3D IC packaging design and verification



siemens.com/3DIC

The great lithography debate: Copper clad laminate or glass substrates?

By Doug Brown [Onto Innovation]

For many in the semiconductor industry, the future is clear. Glass substrates are destined to play an important role in advanced packaging. Compared to organic substrates, glass offers better thermal management, enhanced electrical properties, new form factor possibilities and improved conductor routing. All of which make glass substrates an innovative advanced packaging option for artificial intelligence (AI) and high-performance computing (HPC).

One thing is certain—the glass substrate future is poised to arrive later than originally expected. Previously, many in the industry thought that the transition from copper clad laminate (CCL) substrates to glass substrates would occur when redistribution layers (RDLs) shrank below 5/5µm line/space (l/s). But with today's equipment and processing solutions, the life of CCL in substrates may be extended beyond the 5/5µm l/s demarcation point and reach 2/2µm l/s.

Whether or not manufacturers should use substrates with CCL (**Figure 1**) or glass substrates is likely to be the subject of debate for years to come. In fact, the argument may only be resolved when one substrate, either those with CCL or made from glass, reaches the 2/2µm l/s finish line first—and does so while offering more reliability and lower cost. Until then, let's explore the pros and cons of CCL and glass substrates.

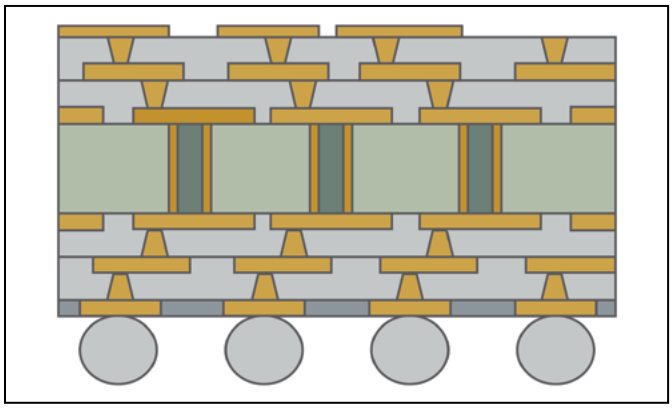


Figure 1: Copper clad laminate substrate.

The CCL argument

With most advanced IC substrates (AICS) currently using CCL and Ajinomoto build-up film (ABF), there is significant industry momentum to propel CCL technology to its full potential of 2/2µm l/s. After all, CCL has some significant advantages for AICS. For one, CCL's properties and limitations are well understood. Two, CCL is robust and nearly indestructible. But pushing AICS with CCL beyond the current RDL line/space requirements and overlay limits will require process innovation and additional lithography steps. For instance, laser-drilled vias in ABF will not support the less than 10µm via dimension requirements needed for state-of-the-art advanced packages with RDLs of 2/2µm l/s. The alternative process will be either photo-imageable dielectrics (PID) or ABF with a lithography patterned hardmask.

The benefits of using lithography for both RDL and via layers are significant. Not only will the overlay between via layers and RDL be improved, the via dimension could easily be reduced to less than 5µm. And by using lithography to pattern both the via and RDL structures, it will be possible to shrink the design rules of the via to the RDL landing pad, which currently limits package design rules, resulting in low interconnect density and additional RDL layers. At the moment, the via to RDL landing pad dimensions for 9/12µm RDL are more than 50µm, with a contingency for overlay errors between the via and the pad of more than 10µm. If overlay were improved by utilizing lithography for both RDL and vias, these dimensions could shrink significantly. However, this process adjustment would require a few additional steps. Regardless, the benefits would reduce overall costs, improve yield and, most importantly, extend the CCL roadmap to 2/2µm l/s.

Another lithography challenge for the 2/2µm l/s goal involves photoresist. Currently, the majority of AICS CCL manufacturers use dry-film negative-tone photoresist. This photoresist is laminated to the substrate and works well with larger RDL structures. However, at the 2/2µm l/s RDL node, a positive tone material would provide better resolution and process latitude. At this time, most positive photoresist is liquid and will require slit coating, or spray coating, unless a dry-film laminated version can provide the same imaging performance.

The AICS glass argument

When AICS CCL manufacturers identified the instability of CCL substrates and RDL design rule limits with their existing processes, they singled out glass as an attractive alternative (**Figure 2**). As it stands, glass has several selling points over CCL. At the top: glass provides a flat and distortion-free surface on which to build RDL and micro vias. The benefit here is that it enables even smaller features to be defined.

Glass, however, comes with its own set of challenges. Number one, it is fragile. This is especially true when it comes to the large panel sizes being employed today (510mm x 515mm and 600mm x 600mm) in advanced packages. Another drawback: glass substrates are also very thin. In some cases, less than 100µm. Given the fragility and thinness of glass substrates, sophisticated handling equipment will be required to process glass substrates through the various process steps to reduce the risk of breakage.

From a lithography perspective, a number of issues encountered with CCL can be immediately resolved by opting for glass

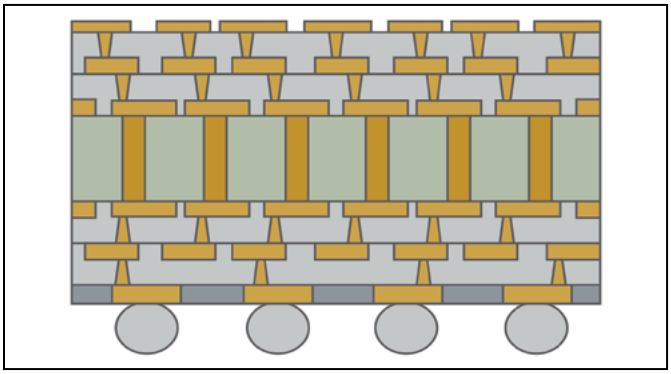


Figure 2: Glass substrate.

substrates instead. For instance, overlay of the via to RDL would be significantly improved as the glass would maintain its dimensions and not suffer from distortions by curing organic dielectric materials. The glass substrate would support higher resolution lithography. As such, the depth of focus budget would not be lost to substrate non-flatness. This would allow lithography tools to increase their numerical aperture (NA) to achieve the highly sought after 2/2µm l/s but sacrificing depth of focus (DOF) in the process, as described by Rayleigh's criterion. The reduced DOF, which continues to decline as resolution increases, would still be within a reasonable manufacturing DOF budget with glass substrates.

As with CCL, ABF, RDL and via processing would all need to be modified to meet 2/2µm l/s requirements. Some of the process steps described previously would be similar, but below 2/2µm l/s additional processes would need to be employed, especially for copper RDL plating seed removal. This particular process is isotropic and subjects the entire panel to a brief flash etch, which not only removes the seed material, but also the RDL metal, thereby reducing line width and impacting critical dimension (CD) uniformity. To resolve this issue, damascene processing has been proposed; this would require lithography to pattern RDL trenches in the ABF or PID and chemical mechanical planarization (CMP) to remove the excess copper to generate copper RDL interconnects.

Lithography system solutions

Currently, AICS CCL manufacturers are using extremely large field steppers (250mm x 250mm) with substrate formats of 510mm x 515mm and 600mm x 600mm. However, there is some discussion of 650mm x 650mm substrates, but these are not mainstream. These extremely large field steppers achieve high throughputs in excess of 110 panels per hour (PPH).

In addition to steppers, printed circuit board (PCB) manufacturers are experienced with laser direct imaging (LDI)

systems; as such, these lithography tools are also being used for RDL patterning involving AICS CCL. Of course, LDI has the advantage of not requiring reticles, but it is limited in resolution and throughput and not suitable for high-volume manufacturing (HVM). LDI is more of a research and development (R&D) tool to test out new designs and prototype larger geometry packages.

For AICS CCL processes, HVM steppers have low NA, which provide a large DOF; this allows steppers to easily accommodate the non-flatness of the substrate material. Currently, the RDL l/s resolution for extremely large field size HVM steppers is limited to 3µm. However, as we move closer to RDL of 2/2µm l/s, stepper solutions are available, albeit with smaller field sizes. The downside here is that such solutions limit package sizes to less than 60mm x 60mm—smaller than what most advanced packages will need—until the next generation of extremely large field steppers arrive with higher NA to support less than 2/2µm l/s.

At this time, R&D programs using CCL and/or glass substrates are racing toward the 2/2µm l/s node. HVM is still at 9/12µm l/s and is moving slowly towards 5/5µm l/s, with lithography requirements easily satisfied by the extremely large field size steppers. The HVM of glass substrates and 2/2µm l/s are not expected to occur until the end of the decade, so there is still time to gain a comprehensive understanding of the lithography requirements for 2/2µm l/s. Still, there are many lithography questions that need to be answered. For instance, what is the correct NA and DOF requirement and field size? Of course, customers are looking for these performance parameters to go beyond the laws of physics, so there needs to be more discussions and collaboration to determine what will be required.

In the past, the semiconductor industry referenced the International Technology Roadmap for Semiconductors (ITRS) to align original equipment manufacturers (OEMs) with material and substrate suppliers to deliver solutions with a clear indication of timing. Unfortunately, the Heterogeneous Integration Roadmap (HIR) does not have the level of detail required for lithography.

In the absence of a detailed industry-defined lithography roadmap, collaboration between OEMs and the material/substrate supply chain will be imperative. To help meet these challenges, Onto Innovation has established the Packaging Application Center of Excellence (PACE) in our Wilmington, Massachusetts, headquarters to address this issue. Collaborators are already engaged with the company in defining projects to help answer many of the most pressing lithography questions.

PACE will provide access to next-generation extremely large-field steppers, inspection, metrology and software capabilities that are currently in development. Furthermore, OEMs and supply chain partners will be able to develop next-generation materials using the center's infrastructure and its team's advanced packaging knowledge to provide customers with the solutions they need to accelerate their technology roadmaps, whether the future is in CCL or glass.

This collaborative opportunity posed by PACE may help determine the answer to the bigger question of which technology will win the race: copper clad laminate or glass. Until then, the debate continues.

Biography

Doug Brown is Senior Director of Product Management, Lithography, at Onto Innovation, Wilmington, MA. He is a graduate of the U. of Arizona, with a PhD in Electrical Engineering and a Master's in Electrical Engineering and another Master's in Physics. For the past 25 years, his focus has been on capital equipment development for the semiconductor industry, with tool introductions spanning plasma etch, deposition, ion implantation, liquid metrology and lithography. Email Doug.Brown@ontoinnovation.com

Heterogeneous chiplet integration to make megachips

By Rabindra N. Das, Jason Plant, Alex Wynn, Matthew Ricci, Ryan Johnson, Matthew Stamplis, et al. [MIT Lincoln Laboratory]

This paper describes a new, extremely large area integrated circuit (ELAIC) solution—we are calling a "megachip"—suitable for combining multiple chiplets of varying type (e.g., memory, application-specific integrated circuits [ASICs], central processing units [CPUs], graphics processing units [GPUs], power conditioning) into a single package on a common interconnect platform.

The megachip approach helps to rearchitect heterogeneous chip tiling for developing highly complex systems having desired circuit density and performance. Recent work on large-area superconducting integrated circuits to join multiple individual die is highlighted in this article, with particular attention paid to the processing of the high-density electrical interconnects formed between the individual die. A variety of megachip assemblies were fabricated and characterized using several techniques (i.e., scanning-electron microscopy (SEM), optical microscopy, confocal microscopy, X-ray) to investigate the integration quality, minimum feature size, silicon content, die-to-die spacing, and gap filling. Silicon dioxide, benzocyclobutene (BCB), epoxy, polyimide, and silicone-based dielectrics were used for gap fill, via formation and redistribution layers (RDLs).

For the megachip approach, the thermal stability is improved by reducing the die-to-die (D2D) gap and increasing the silicon content, allowing assemblers to mitigate the problem of mismatch in coefficient of thermal expansion (CTE) for different substrates/modules integration schemes, which is important for allowing the broad temperature range stability from reflow to operation at room or even cryogenic temperatures. Megachip technology facilitates more space-efficient designs and can accommodate most heterogeneous dies without compromising stability or introducing CTE mismatch or warpage. A variety of heterogeneous chips were used to fabricate megachip modules. The present process allows fabrication of megachip buildup layers having thicknesses in the range of 1-10µm, which allows packaging structures having both finer pitch and higher density. The processes and materials used to achieve smaller feature dimensions, satisfy stringent registration requirements, and achieve robust electrical interconnections are discussed.

Introduction

The increasing demand for digital computing, mobility, and connectivity is driving the microelectronics industry toward cost-driven, highly-integrated, miniaturized technology with increased performance and lower power consumption to bring next-generation devices into more and more applications [1-2]. Over the last decade, high-performance computing (HPC) has evolved to adapt smaller and more diverse technology nodes suitable for artificial intelligence (AI), machine learning, and embedded computing platforms—these applications consistently involve trade-offs between enabling more compute capability versus constraints in volume, weight, power, and thermal management.

Most of the power consumption for the above applications is due to moving data between chips in a system rather than the actual computing [3]. Furthermore, traditional Moore's Law scaling for developing next-generation devices faces various challenges including fabrication of larger chip sizes and associated yield improvement, development time, and cost scaling. This has forced the microelectronics industry to develop a number of alternative advanced packaging architectures and heterogeneous integration technologies [4]. A modern packaging architecture needs to integrate multiple processor and accelerator chips with minimum chip-to-chip spacing to minimize the interconnect length, on-chip memory, higher bandwidth connections, and management for greater heat densities, while being pushed into higher I/O counts, smaller pitches, and larger footprints [5-6]. This necessarily drives a requirement for improving the power efficiency of the chip-to-chip I/Os. In addition, new advanced packaging requires low-loss, mixed material, and versatile construction to accommodate the complexity associated with size, weight, and power (SWaP) optimization.

Conventionally, better wiring densities have been achieved by using filled dielectric to reduce via dimensions, lines, and spaces—thereby increasing the number of circuit layers—and utilizing microvias for interconnection. However, each of these methods has inherent limitations. For example, there are limitations related to laser drilling and electroplating of high aspect ratio blind- and through-vias, increased resistance of narrow (and) long circuit lines, and increased cost of fabrication related to additional wiring layers [7]. As a result, microelectronics packaging is moving toward alternative, innovative, low-cost approaches as solutions for miniaturization [8-10]. Fabrication, assembly, and heterogeneous integration are bridging the gap by enabling economic use of the third dimension (2.5D and 3D packaging). System-level integration is also emerging. These approaches include multi-die system-on-chip (SoC), system-in-package (SiP), stacked die, or package-stacking solutions.

In addition to the trend toward miniaturization, new materials and structures are required to keep pace with more demanding packaging performance requirements. Wafer-level packages (WLP), panel-level packages (PLP), silicon/organic interposers with redistribution

layers (RDLs), active interposers, and bridge die have become the preferred methods for lower-cost integration to meet the demands of higher functionality in ever-smaller packages, especially when coupled with the use of different technology node die [11-13]. The size of WLP increases with smaller technology nodes and causes more reliability and chip package mismatches. Today, various WLP technologies including WLPs with and without through-silicon vias (TSVs), WLPs with embedded-integrated passive devices, and use of low CTE, low-loss, high-glass transition temperature (Tg) material-based wafer-level substrates featuring fine traces and embedded/integrated passives, are used to reduce WLP chip package mismatch. Similarly, flip-chip integration with the bridge die embedded within the package substrate allows for shorter interconnect lengths for chip-to-chip communication. Active interposers with active-to-active bonding [14] are preferred for high-bandwidth, low-latency communication.

Although there are many packaging approaches available today for chiplet integration, the authors believe that there is room for further improvement. Tiling hundreds of known-good chips in proximity to one another and creating chip-like wiring and silicon content are highly desirable for creating next-generation chiplet-based computing architectures, but has yet to be demonstrated. Here, we present the implementation of such a chiplet-based tiling approach.

This paper discusses a heterogeneous chip tiling that enables the realization of extremely large-area integrated circuits (ELAICs)—or megachips—containing hundreds of closely spaced small chiplets that are interconnected using RDLs fabricated via a lithographic process. The ELAIC platform allows the tiling of known-good chiplets to make systems that perform as a single-chip monolithic device, despite being composed of many smaller heterogeneous chiplets. With this approach, one can fabricate a large-format single-chip-like SoC from advanced-node chiplets, screening for known-good die in order to increase the yield and performance of advanced-node technology beyond what is possible in a single-chip format.

A key focus of this paper is to address the scaling challenges faced when building large-scale processors. For example, the ELAIC solution is suitable for combining multiple types of chips (e.g., memory, ASICs, CPUs, GPUs, power conditioning) into a single system. This approach extends the number of chip tiles within a given space by enabling sufficiently high chip-to-chip connectivity to allow multiple chiplets to perform as a single-chip monolithic device. Connecting chiplets through our approach will enable a path to increase the format size of heterogeneous processors. The ELAIC structure, having 5-20µm chip-to-chip spacing, creates short (i.e., 50-500µm) electrical links for high-bandwidth, low-latency communication. An ELAIC has a chip-like silicon content (about 99%), allowing thermally-stable and inexpensive fabrication of a heterogeneous SoC with chip-like wiring densities. For HPC, power consumption comes primarily from moving data between chips in a system rather than from the on-chip computing operations. The ELAIC approach reduces data movement constraints by integrating multiple chiplets with minimum chip-to-chip spacing, thereby reducing the loading of these I/O paths by at least an order of magnitude. By integrating multiple chiplets into one large-area chip (2D array), the ELAIC approach can help solve many scalability challenges for high-end electronics.

The megachip approach

The following sections discuss the megachip (or ELAIC) approach with respect to tiling, physical characterization, and demonstration of the electrical interconnect.

ELAIC chip-tiling approach. The approach to ELAICs involves developing an integration process that can address the scaling challenges faced by many multi-chip systems. Integration of multiple chips that were produced using different (heterogeneous) fabrication technologies has been a persistent challenge. Typically, individually-packaged chips use a board-level assembly approach, and the associated "parasitic" electrical overhead and latency often become the limiting factors to a system's performance.

The ELAIC integration process will allow the tiling of known-good chips to make systems that perform as a single-chip monolithic device, despite being composed of several smaller heterogeneous chips. The primary goal of this effort is to develop a chip packaging interconnected with a RDL that is capable of integrating hundreds of chips in proximity to one another in a single system as shown in **Figure 1**. The RDL typically has multiple metal layers, each separated by a plasma-enhanced chemical vapor deposition

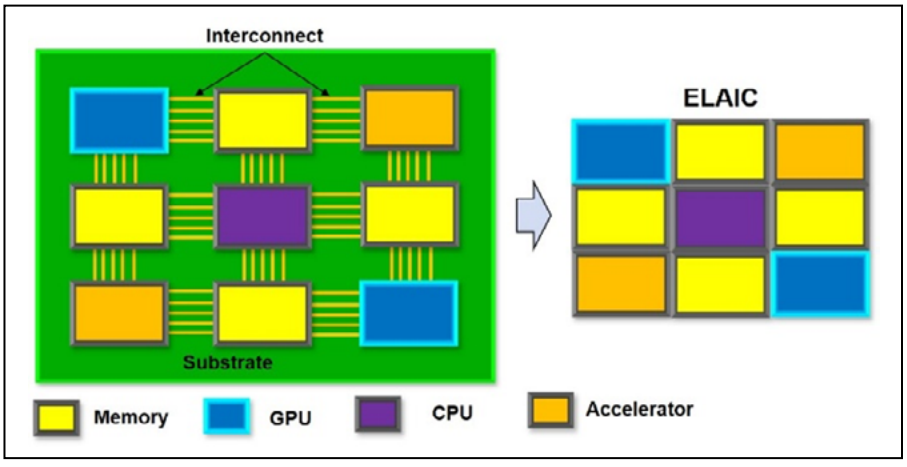


Figure 1: Extremely large-area integrated circuit (ELAIC), or megachip, concept: a) (left) Regular package where individual chips are attached to the substrate (organic or Si) and interconnected through the substrate; b) The megachip combines all the chips in a single plane where each individual chip will have at least two nearest neighbor chips for interconnection. The megachip enables chip-like wiring and eliminates the need for a substrate containing interconnects.

(PECVD) silicon dioxide layer, polyimide, or BCB dielectric, and uses micro-vias for interconnection. For our demonstration, the metal wiring layers were patterned using non-contact direct-write photolithography, which supports minimum wiring layer dimensions of 1µm and field sizes exceeding the largest relevant reticle size (50 x 50mm^2).

We evaluated various chip-like dielectric, wiring, and interconnection options. An ELAIC process flow is illustrated in **Figure 2**. The illustration displays the process flow for a double-layer RDL with a micro-bump layer on top of the RDL. The primary advantage of the ELAIC assembly is to produce a narrow (5-20µm) gap between the chips. This kind of gap is suitable for short (50 to 500µm) chip-to-chip interconnect lengths as shown in **Figure 2c**. Today, many high-performance electronic integrated circuits (e.g., field-programmable gate arrays [FPGAs]) use parallel interfaces for chip-to-chip communication. This approach requires small electrical length and more individual physical wires for data transmission. ELAIC enables narrow chip-to-chip spacing (10-20x smaller than the traditional approach) for smaller interconnect lengths and finer feature circuits, thereby enabling more physical wires for I/O connections with lower-latency, lower-power, higher-bandwidth chip-to-chip communication.

As a first step for chip tiling, we developed an assembly process for maintaining narrow gaps between the chips while maintaining top chip surface planarity in a larger scale ELAIC format. The top chip surface planarity enables thin dielectric deposition to make a finer pitch interconnection with chip-like RDL circuits. We assembled various ELAIC configurations using 5mm x 5mm to 20mm x 20mm chips in order to test chip surface planarity and chip-to-chip spacing/gap for the ELAIC structure. **Figure 3** shows various ELAIC

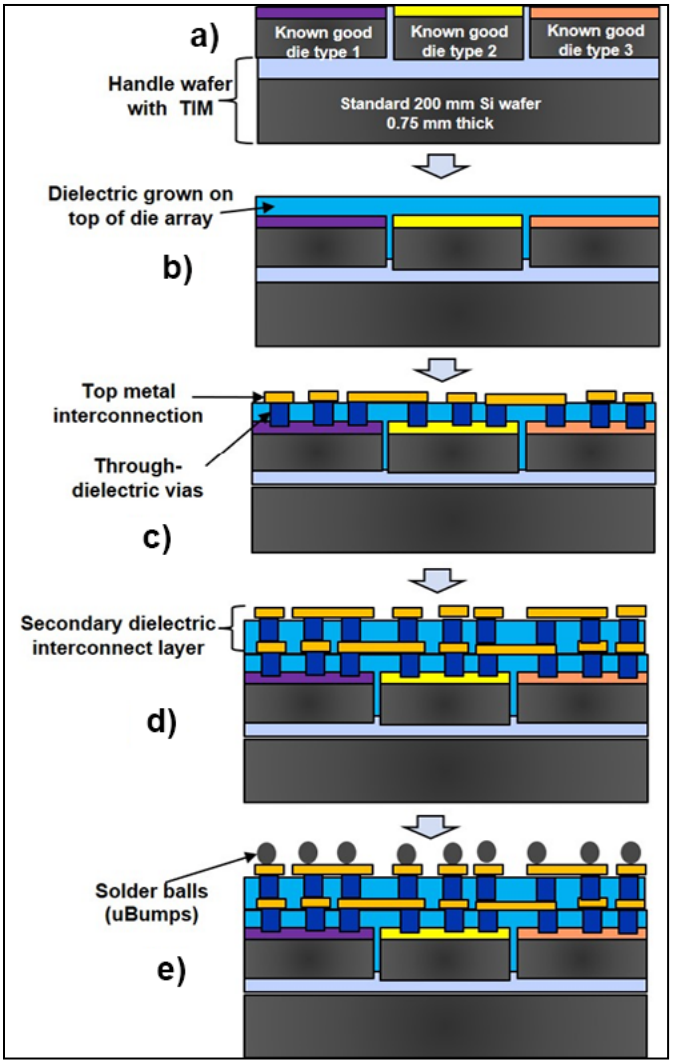


Figure 2: Process flow for an ELAIC construction. The chips are assembled on a handle wafer: a) Known-good die are placed face-to-face using a microscope. In general, the die use thermal interface materials (TIM) or related materials for die attachments; b) The dielectric layer is deposited; c) The first RDL is formed; vias are etched and top metal is deposited on the dielectric layer; d) The second RDL and additional dielectric layers for more complex interconnectivity are formed (target up to 4 RDLs); and e) Micro-bump fabrication—the bumps are deposited for flip-chip connection.

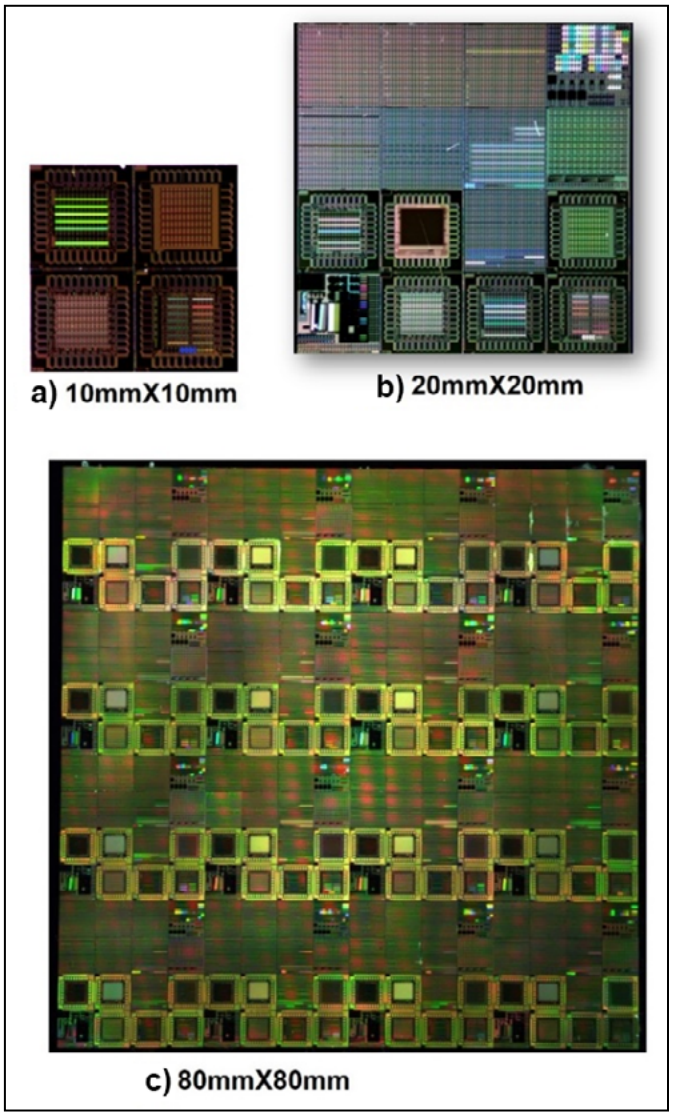


Figure 3: The ELAIC combines known-good die together to make systems that perform like an extremely large single chip. The scalability of the ELAIC fabrication process is shown—with assembly sizes ranging from 4 chips to 16 chips to 256 chips: a) Four 5mm x 5mm chip assembly; b) 16 5mm x 5mm chip assembly; and c) 256 5mm x 5mm chip assembly.

configurations using 5mm by 5mm chips ranging from 4 chips up to 256 chips as a representative example. The extremely large area (**Figure 3c**, 80mm by 80mm) circuits can be useful for advancing many systems, including those for HPC with diverse technology nodes for AI and deep learning, superconducting classical and quantum computing, large-format digital-pixel focal plane arrays [15-17] with minimum seam loss, photonic-chip tiling, millimeter-wave phased-array radar tiles, etc. The process involves the tiling of known-good chips to make systems that perform like a single-chip monolithic device, despite actually being composed of several smaller heterogeneous chips. Integration of multiple chips with different (heterogeneous) fabrication technologies has been a persistent challenge. The ELAIC (or megachip) platforms in many ways support chiplet-based system requirements by:

- Combining known-good chips together to make systems that perform like an extremely large single heterogeneous chip with very narrow inter-chip spacing for compact assembly. And for phased arrays, allowing the tightening of the lattice spacing (area is less) for better beam-steering.
- Providing aggressive interconnect pitch scaling for true process node interchangeability. And for RF, achieving lower interconnect parasitics that support more broadband connections.
- Enabling chip-like circuit content with good inter-chip planarity.
- Providing a built-in heatsink, thereby allowing for a better thermal solution for large chips.
- Supporting mixed-material construction with more Si/mm^3 (chip-like Si density), minimizing CTE mismatch, and supporting reliable operation ranging from room temperature to high (fabrication) and low (cryogenic) temperatures.
- Offering a path for introducing heterogeneous integration of non-silicon chips (not explored in this present work).
- Allowing active-to-active bonding (mix-and-match chiplets),

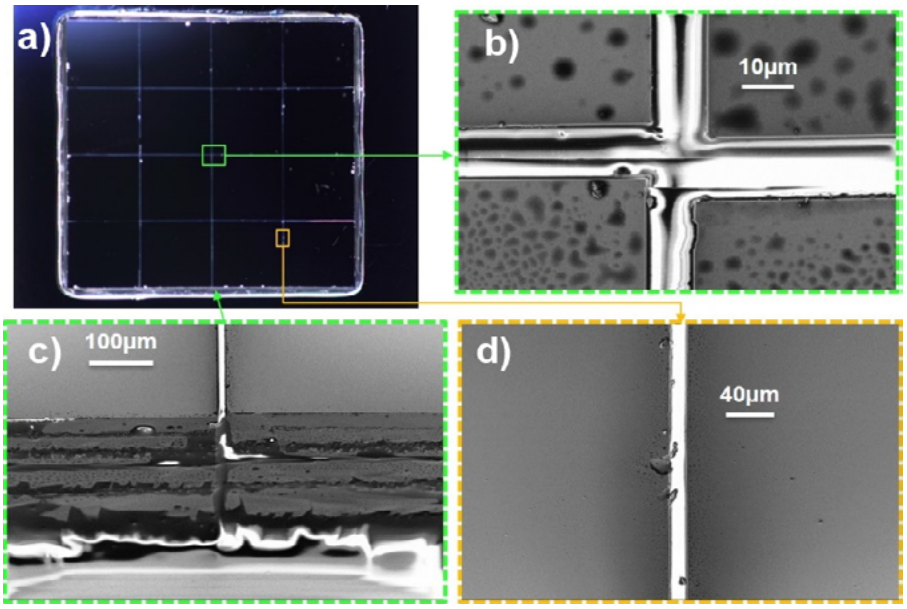


Figure 4: A 16-chip ELAIC assembly with very small 5-20µm chip-to-chip (C2C) spacing filled with dielectric. a) (top left): optical image of a 16-chip (each 5mm x 5mm) ELAIC assembly and b-d) (top right, bottom left, bottom right) corresponding enlarged SEM images that indicate a narrow C2C spacing filled (white area in SEM) with dielectric.

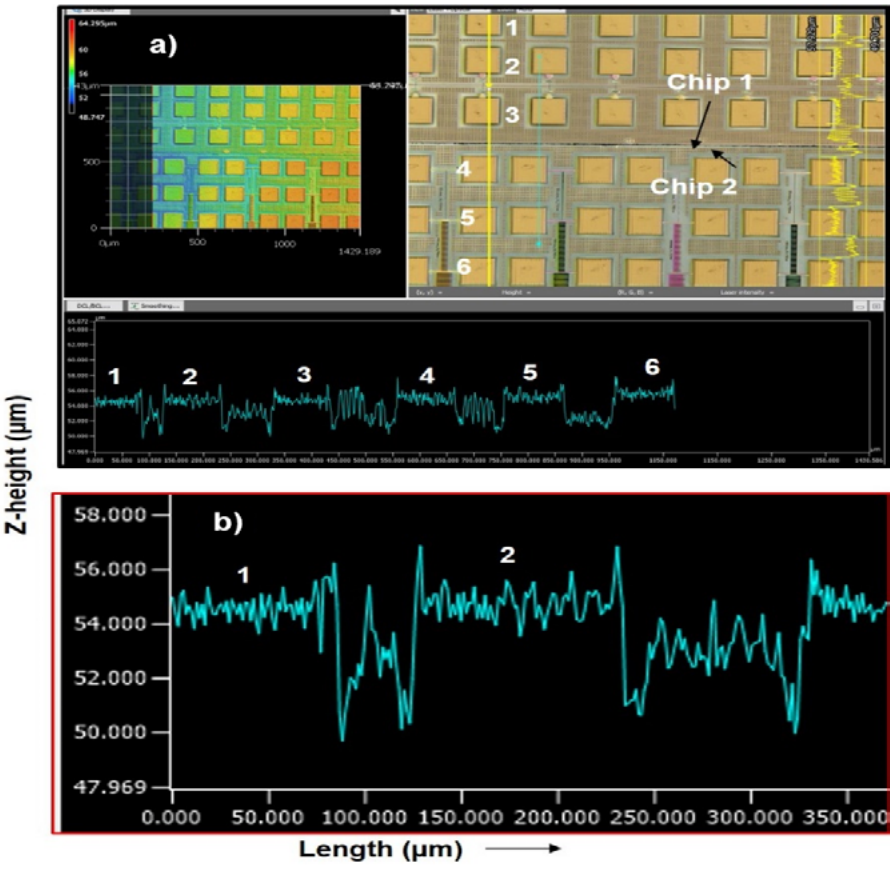


Figure 5: Selective-area confocal scan for a 4-chip ELAIC assembly. The figure shows confocal images and corresponding line scan between the chips to measure inter-chip planarity: a) Confocal micrograph and corresponding line scan. Confocal line scan from chip 1 metal pads (1,2,3) to chip 2 metal pads (4,5,6); and b) An enlarged confocal line scan; the confocal line scan shows metal pad height variation along the line as it scans from one edge to the other.

both side efficient metallic thermal interface materials (TIM), reduced die-die thermal resistance, and thermal cross-talk between neighboring die.
- Handling higher power density with a thermally-efficient Si floor plan.

ELAIC physical characterization. We used a variety of nondestructive analysis techniques for ELAIC physical device characterization. **Figures 4-5** show representative examples of ELAIC characterization. SEM, confocal scan, X-ray, and optical images are used to characterize key fabrication steps, which include chip-to-chip spacing, inter-chip planarity, dielectric deposition, via formation, feature size, and micro-bumping. **Figure 4** shows spacing between the stealth-diced chips in a 16-chip ELAIC assembly. The SEM data indicates that the ELAIC fabrication process maintains a narrow gap of 5-20µm between the chips and gap filling between the chips. Appropriate cleaning to remove dicing debris and give a smooth chip edge with minimal chipping is critical for minimizing chip-to-chip spacing.

Confocal microscopy was used to evaluate inter-chip planarity. **Figure 5** shows representative confocal images of a 4-chip ELAIC module measured using 100nm resolution in the z-axis. The confocal line scans show z-height variation along the line as it scans from one chip to the other. Metal pad height variation (pad 1-6) within and between the chip is negligible (less than 1µm). It is clear from confocal line scan data that the fabrication process maintains chip-like inter-chip planarity. We have developed a variety of ELAIC assembly approaches to optimize critical alignments between the chips. For example, the ELAIC devices used optical microscope for chip-to-chip alignment, and the post-process alignment accuracy was ±3µm. The gap fill and chip surface planarity allow us to select from a variety of dielectric material (PECVD oxide, benzocyclobutene [BCB], silicone, polyimide, etc.) to deposit on top of the ELAIC surface.

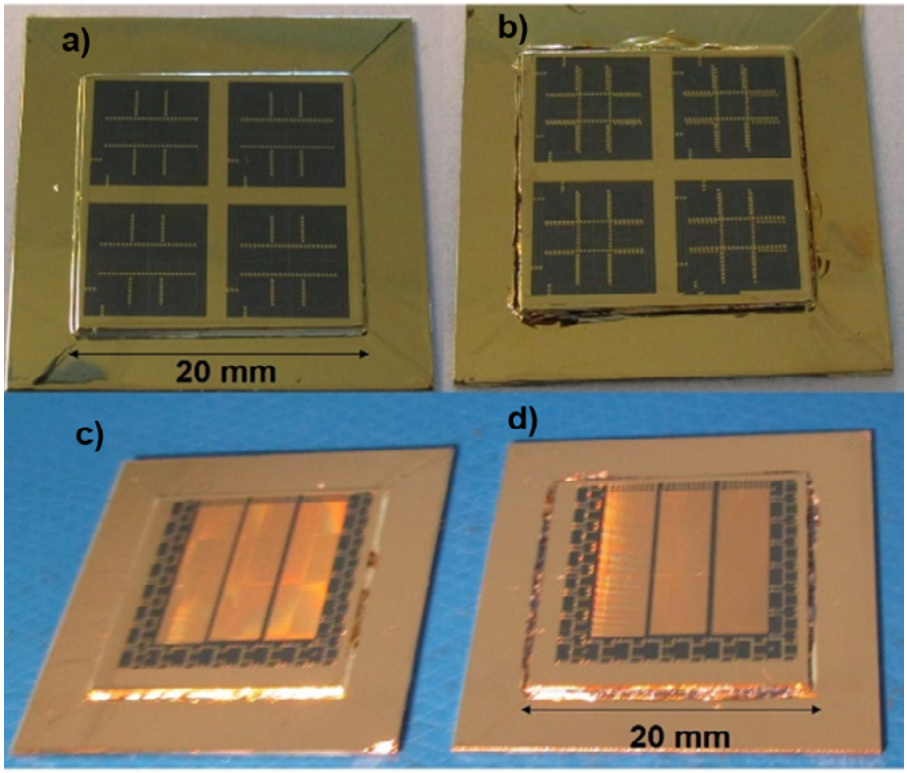


Figure 6: Passive circuit demonstration on top of a 16-chip ELAIC assembly: a) A single metal layer RDL; b) A double metal layer RDL deposited on BCB; c-d) A daisy chain circuit created on top of a 16-chip ELAIC assembly using multi-layer BCB dielectric.

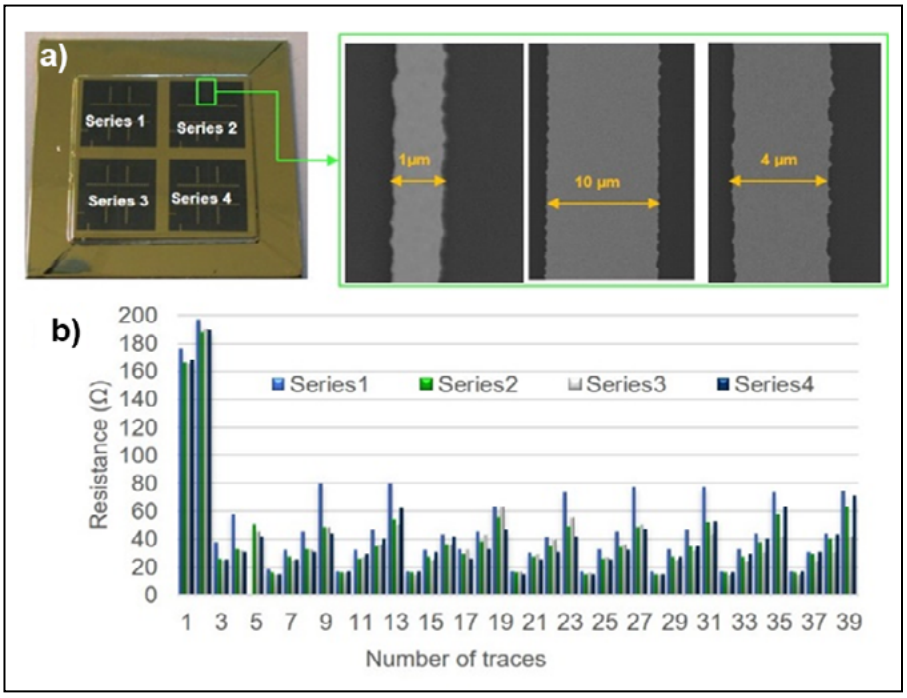


Figure 7: A passive interconnection circuit demonstration on top of a 16-chip assembly. The figure represents single-metal-layer passive circuits with four sections. Each 4-chip section has 1-10µm wide, 5-20mm long circuit traces going between the chips. a) Optical image of ELAIC and corresponding selective area SEM images of the circuit connecting the chips; b) Measured room temperature (RT) passive circuit resistance for four different sections. Resistance variation is due to the different widths and lengths of the electrical interconnect lines.

ELAIC electrical interconnect demonstration. As a next step for chip tiling, we selected PECVD oxide and BCB for RDLs. We have used single- and double-metal layers for implementing a passive electrical-interconnect demonstration. **Figure 6** shows a variety of passive interconnects deposited on a 16-chip (each 5mm x 5mm) ELAIC assembly. For example, it shows a variety of passive interconnects ranging from having 1-10µm wide and 5-20mm long circuit traces going between the chips. We also used a daisy chain circuit to access interconnections between metal layers. **Figure 7** shows a representative single-layer passive circuit example. **Figure 7a** shows an optical image and corresponding enlarged SEM images of a passive circuit lithographically-fabricated using BCB dielectric on a 16-chip ELAIC assembly. The SEM shows finer line circuits down to 1µm connecting multiple chips. These kind of fine-line circuits support chip-like wiring. **Figure 7b** shows room-temperature resistance of the passive circuits. It consists of four sections and each section has 1-10µm wide (trace width:1-10µm) and 5-20mm long traces going between the chips. Linewidth and linelength dictate the total resistance for individual passive circuits.

In addition to passive circuits, we also investigated interconnection between active superconducting chips containing tri-layer Josephson junctions (JJs) for larger system applications, such as quantum processors, readouts, control, and amplifier chips. Active chips can be connected together to create a multi-die SoC. These JJs and other active components may be on the same chip, or separate chips assembled into the ELAIC platform. In either case, a first step toward assessing the suitability of the ELAIC structure with Nb/Al-AlOx/Nb tri-layer junctions is to determine the impact of fabrication on the tri-layer junction performance. The addition of the RDL fabrication to the JJ chip may change the critical current, sub-gap voltage and other junction properties. In addition, multiple chip assembly, gap filling, and planarization may affect the stability and junction performance at 4K. To quantify the effects of fabrication on the tri-layer junction, we fabricated an ELAIC assembly where multiple superconducting chips with tri-layer junctions are attached to a single large ELAIC. This allowed us to determine the impact of ELAIC fabrication and to demonstrate basic desirable functionalities for multi-die SoC.

To assess the electrical performance of the chip assembly, multiple 4-chip ELAIC devices (**Figure 8**) were attached to a circuit card and wire bonded to measure I-V characteristics of tri-layer-based JJs at 4.2K. ELAIC-assembled superconducting chips had multiple sizes of junctions ranging in size from 700nm to 1000nm. Each measurement showed a typical

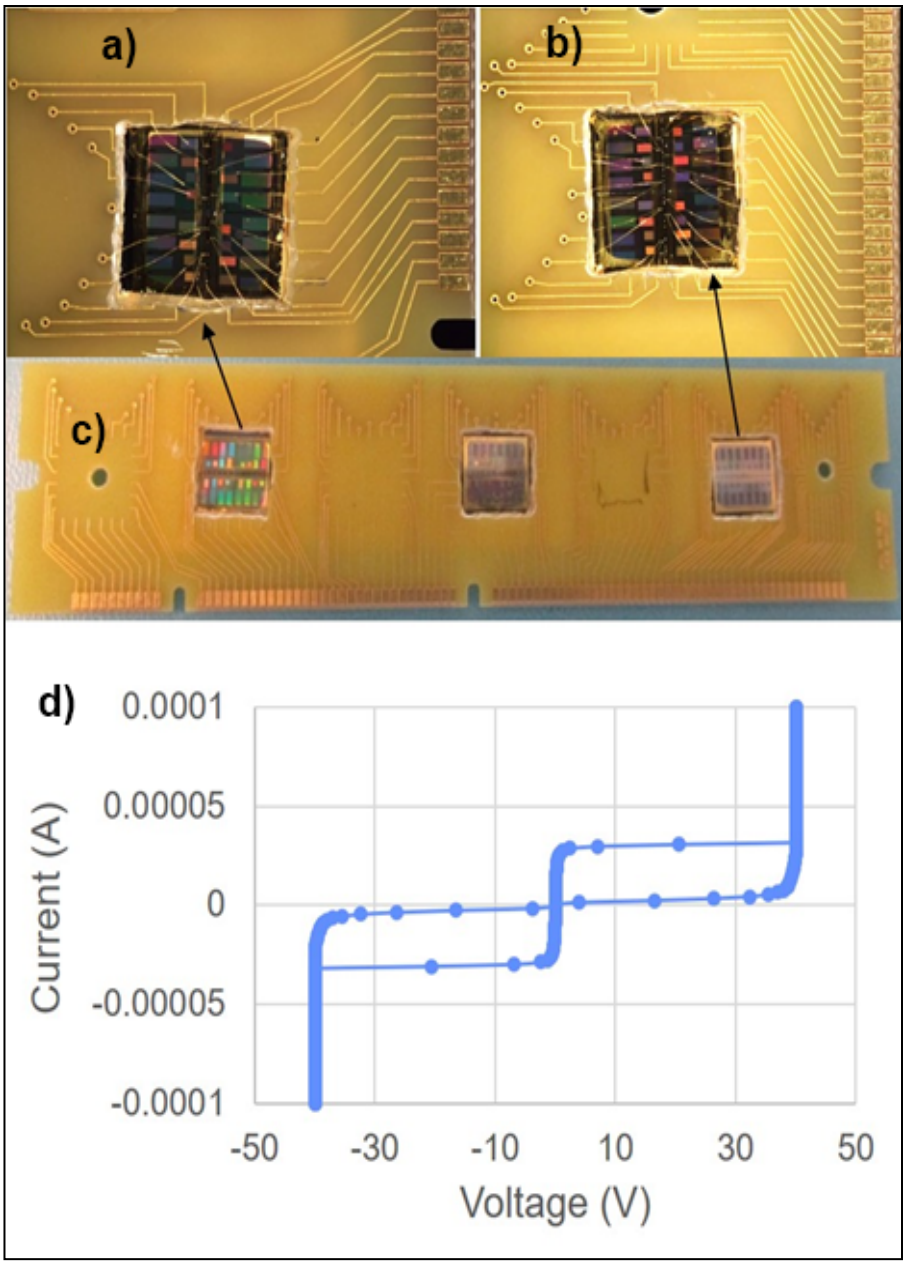


Figure 8: An active circuit demonstration. The figure shows three 10mm x 10mm ELAIC samples attached to a circuit card. Each ELAIC module consists of four active device chips containing superconducting junctions. These ELAIC devices used PECVD silicon dioxide as the dielectric and Ti-Au chip-to-chip interconnections. a-b) Optical image of 10mm x 10mm ELAIC attached to a circuit card for cold testing; c) Three 10mm x 10mm ELAIC samples attached to a circuit card for cold testing; d) The I-V characteristics of JJ series arrays connected between the chips through Nb and gold lines with a drawn JJ diameter of 0.7µm and 1µm.

I-V characteristic of Nb/Al-AlOx/Nb unshunted tunnel junctions (i.e., with respect to the Josephson critical current, sub-gap voltage, normal resistance at 4.2K). A variety of active superconducting chips with tri-layer junctions have been assembled to implement the ELAIC. The I–V characteristics and switching current of various tri-layer flip-chip JJ arrays were measured. We measured many ELAIC JJ arrays ranging from 40 to 20,000 JJs in series, with JJ drawn diameters ranging between 1.0µm and 0.7µm. I-V characteristics were measured for ELAIC JJ arrays connected across multiple chips through the RDL. **Figure 8d** shows a representative example of the typical I-V curve obtained from these multi-chip JJ series arrays connected through Nb and gold lines between the chips. This confirms connectivity between the chips through RDL and the preservation of JJ characteristics after RDL fabrication.

Single chip vs. the megachip concept

Figure 9 compares the ELAIC assembly with an equivalent single chip fabricated using a standard integrated circuits process. A 20mm x 20mm unsingulated chip was diced into 16 5mm x 5mm chips and reassembled to create a 20mm x 20mm ELAIC. We used X-ray imaging to inspect the single chip before dicing as well as the subsequent ELAIC. The X-ray image of the single chip and ELAIC looks similar. Dicing of the single chip removes Si from the dicing lane causing a smaller metal-to-metal gap between the chips as shown in **Figure 9d**. Overall, the ELAIC fabrication process produces a highly compact (>99% Si content) chip assembly with 5-20µm spacing between the chips, and maintains inter-chip planarity that is required for the finer line and smaller interconnect length needed to form parallel interfaces with wide I/O, high-bandwidth, and low latency for chip-to-chip communication. **Table 1** compares the single-chip SoC option with ELAIC multi-die SoC.

ELAIC can be used for flip-chip bonding to simplify fabrication and enhance connectivity and functionality in 3D for various applications. Flip-chip ELAIC (see **Figures 10-11**) offers a number of advantages over conventional monolithic SoC approaches:

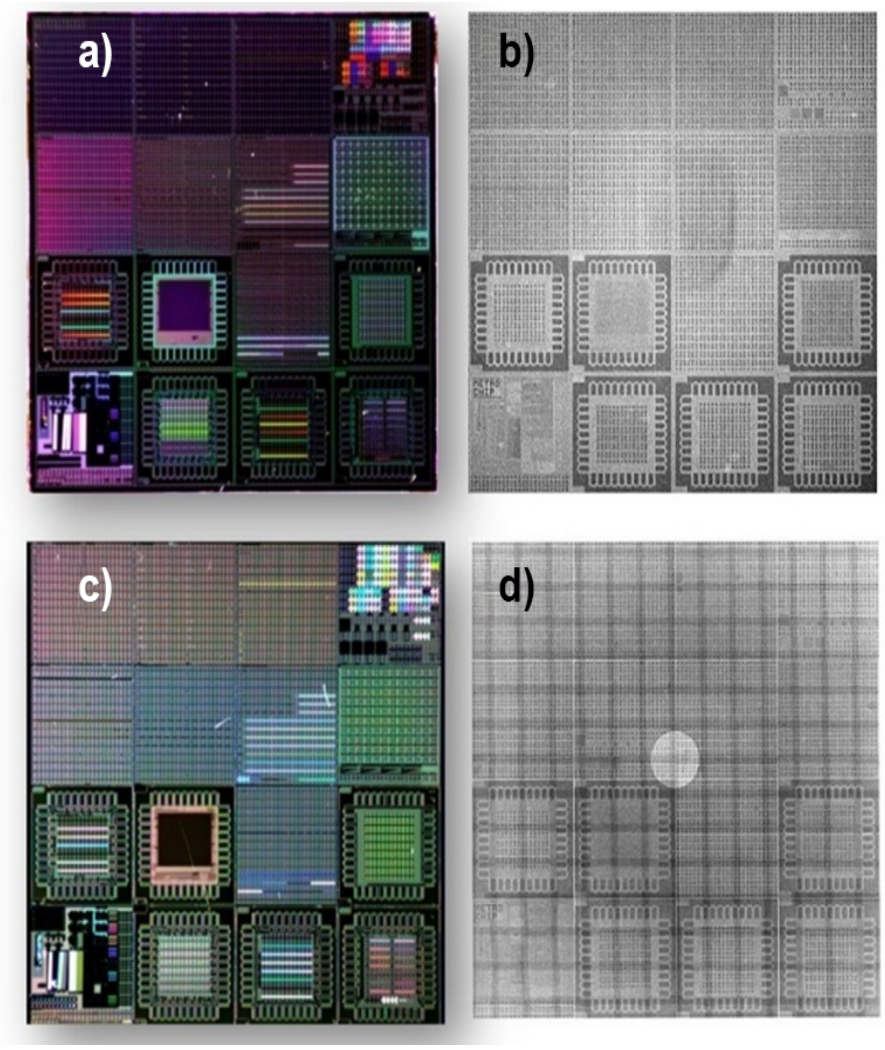


Figure 9: A single-chip vs. an ELAIC assembly: a) Optical image of a 20mm x 20mm unsingulated chip; and b) corresponding X-ray image of that unsingulated chip. This 20mm x 20mm single chip was subsequently diced into 16 5mm x 5mm chiplets. c) Optical image of a 20mm x 20mm ELAIC formed by recombining the 16 5mm x 5mm chiplets into single chip-like structures; and d) corresponding X-ray image of the ELAIC in c).

System	Single Chip/SoC	ELAIC
Si content	All (100%) Si	>99% Si
Interconnection	Through the wafer process	Through RDL
System-on-chip	Monolithic SoC	Multi-die SoC
Technology node	Single-node homogeneous system	Multi-node heterogeneous system

Table 1: A megachip vs. a single chip.

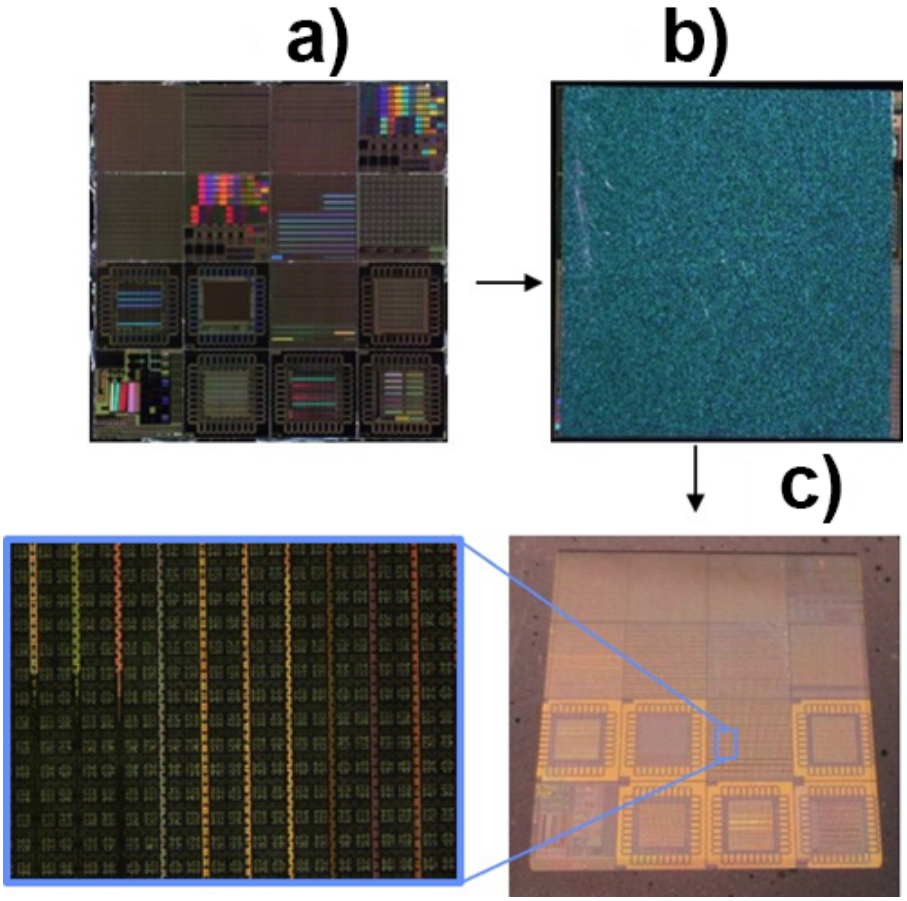


Figure 10: A flip-chip ELAIC: a) A 20mm x 20mm 16-chip ELAIC; b) An ELAIC flip-chip bonded to Si-die and underfilled to make a flip-chip ELAIC; and c) Si etched from flip-chip die and stopped at the oxide surface of the flip-chip die—this view shows a Si-less flip-chip ELAIC and the corresponding enlarged image looking through the oxide surface.

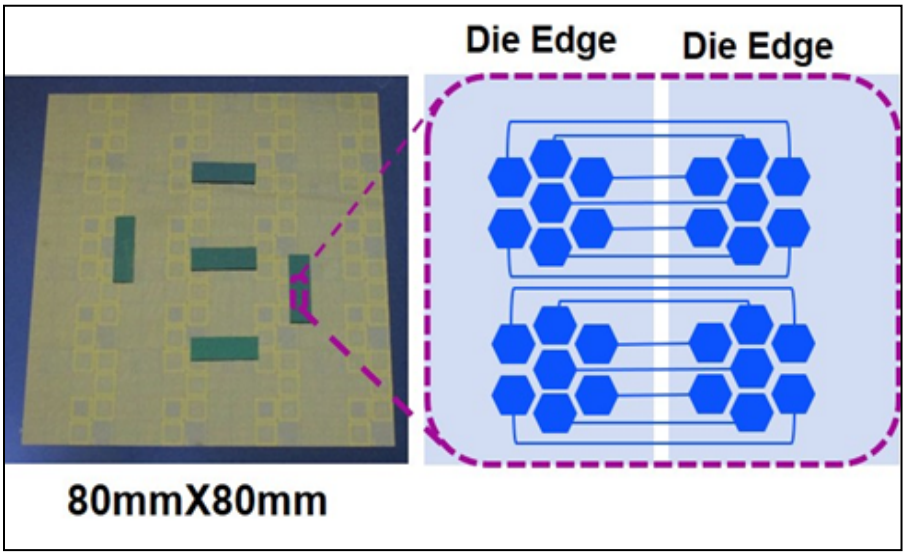


Figure 11: An 80mm x 80mm ELAIC (or megachip). It has sixteen 20mm x 20mm chips assembled to create the 80mm x 80mm megachip. The figure also shows a flip-chip bonded megachip and the corresponding enlarged schematic to show a chip-to-chip connection option through the flip-chip die. The next step is to remove Si from the flip-chip die (similar to **Figure 10c**) and stop at the oxide layer. Adding a Si-less interconnect layer adds 2-6µm thickness—necessary for creating the finer pitch megachip assembly.

- Provides various low-cost multi-chip read-out IC (ROIC) assembly for silicon avalanche photodiodes (Si-APDs) and other imagers [15-17].
- Introduces flip-chip Si-less active/passive bridge for chip-to-chip connection.
- Enables a thermally-optimized Si floorplan.
- Provides a cost benefit for yield and node optimization.

Summary

An integrated approach to develop ELAICs, or "megachips," using various heterogeneous die configurations has been demonstrated. This approach is suitable for high-end, expensive electronics where an SoC can be divided into chiplets with desired functionality and an ELAIC multi-die SoC can be created. The ELAIC can incorporate chips/chiplets from different foundry processes, and different technology nodes to improve mix and match capability, which further improves package performance. It also provides scalability to place a large number of chips onto the ELAIC platform, and enables a design that packages many different functionalities together, making it a viable approach to build larger systems.

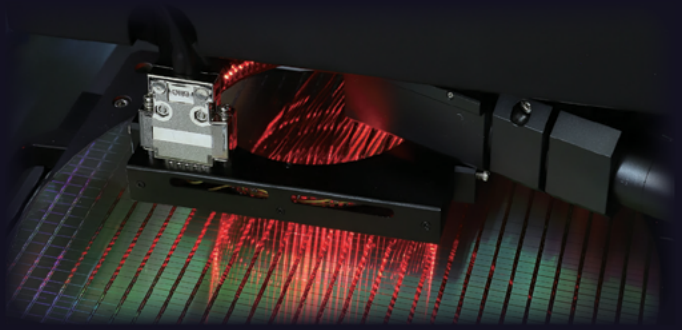
The ELAIC solution is suitable for making the right choices in terms of cost and partitioning–for each of the targeted applications, and to provide a heterogeneous path for large-scale fabrication. The ELAIC integration supports the capability to integrate hundreds of chips (also known as chiplets) in proximity to one another in a single system. This integration technology enables small (50-100µm) interconnects required for parallel interfaces for chip-to-chip communications. The extremely large area integrated circuit allows for connections between bare chips, and the wiring between chips to be as small as the wiring within a chip. The approach increases the circuit complexity that can be integrated within a given space by enabling sufficiently high chip-to-chip connectivity to allow multi-chip systems to perform as a single-chip monolithic device.

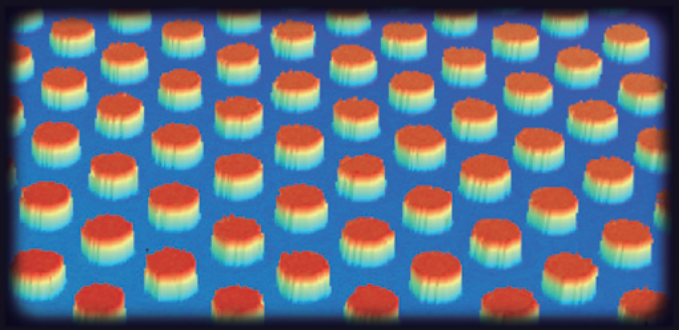
INTEK PLUS
Integrated Measurement System

KOSDAQ LISTED COMPANY

3D Solution for Wafer Bump Inspection

Bumps' 3D conditions such as Height and Coplanarity are among important quality factors in **wafer bumping process**.
INTEKPLUS has developed **WSI optics** used in precision measuring instrument to be applied for **mass-production 3D inspection** of **Fine-Pitch Bump** such as µBump and Cu Pillar.





Yield improvement

Analyze bumps condition to increase production yield.
3D Inspection of bumps before packaging to avoid Si Die loss due to NG bumps.

Fine-pitch inspection

Inspect finest pitch bumps in mass-production.
Shadow-free by WSI using coaxial illumination.

Accuracy and Precision

Bump height and coplanarity measured by high accuracy and precision, including transparent PI layer thickness measurement.

INTEKPLUS KR HQ
Sales person : YK Sung
sygs@intekplus.com

INTEKPLUS US Office
Sales person : Harry Yun
Harry.yun@intekplus.com

INTEKPLUS CN Office
Sales person : Kevin Lin
hisonic12@intekplus.com

INTEKPLUS TW Office
Sales person : Arthur Tung
arthur@intekplus.com

INTEKPLUS JP Office
Sales person : SS Kim
ssk@intekplus.com

INTEKPLUS CO.,LTD.

Visual Technology for Semiconductor Package / Wafer / EV Battery / Display
#263, Techno 2-ro, Yuseong-Gu, Daejeon, 34026 Korea Tel : +82-42-930-9900 Fax : +82-42-930-9999

Acknowledgments

Additional co-authors of this article are: Brian Tyrell, Kenneth Schultz, and Paul W. Juodawlkis, all from MIT Lincoln Laboratory. The authors also gratefully acknowledge Lenny Johnson, Karen Harmon, Peter Murphy, Chris Thoummaraj, John Liddell, Chris O'Connell, Terry Weir, Mike Miszka, Chad Stark, and Marcus Sherwin for useful discussions and valuable technical assistance. This research was funded by the Assistant Secretary of Defense for Research & Engineering under Air Force Contract No. FA8721-05-C-0002. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. government. This paper was originally presented at the 2023 IEEE 73rd Electronic Components and Technology Conference (ECTC).

References

1. R. N. Das, F. D. Egitto, J. Lauffer, B. Bonitz, B. Wilson, F. Marconi, et al., "3D-interconnect approach for high-end electronics," IEEE Elec. Comp. & Tech. Conf. Proc. (ECTC), 2012, 1333-1339.
2. R. N. Das, F. D. Egitto, S. G. Rosser, E. Kopp, B. Bonitz, R. Rai, "3D integration of system-in-package (SiP) using organic interposer: Toward SiP-interposer-SiP for high-end electronics," 46th Inter. Symp. on Microelectronics (IMAPS), 531-537(2013).
3. L. England, E. Tremble, I. Arsovski, W. Sauter, "Advanced packaging drivers/opportunities to support emerging artificial intelligence applications," 2019 Electron Devices Tech. and Manufacturing Conf. (EDTM), pp. 209-11.
4. R. Radojcic, "More-than-Moore 2.5D and 3D SiP integration," DOI 10.1007/978-3-319-52548-8_2.
5. R. N. Das, V. Calmidi, M. D. Poliks, V. R. Markovich, "Nanofluids, nanogels and nanopastes for electronic packaging," 2010 Proc. 60th ECTC, Las Vegas, NV, USA, 2010, pp. 1231-1238, doi: 10.1109/ECTC.2010.5490857.
6. R. N. Das, et al., "Fabrication and electrical characterization of embedded actives and passives for system-level analysis: Towards size, weight and power (SWaP) reduction," 2012 IEEE 62nd ECTC, San Diego, CA, USA, 2012, pp. 1593-1598, doi: 10.1109/ECTC.2012.6249049.
7. R. N. Das, F. D. Egitto, J. M. Lauffer, E. Chenelly, M. D. Polliks, "Versatile Z-axis interconnection-based core-less technology solutions for next-generation packaging," 2013 IEEE 63rd ECTC, Las Vegas, NV, USA, 2013, pp. 1728-1733, doi: 10.1109/ECTC.2013.6575808.
8. F. D. Egitto, R. N. Das, G. E. Thomas, S. Bagen, "Miniaturization of electronic substrates for medical device applications," 45th IMAPS, 186-191(2012).
9. Tom Coughlin, "Chiplets For All," Forbes, May 11, 2019.
10. C. Tseng, "InFO (wafer-level integrated fan-out) technology," 66th ECTC, May 2016, pp. 1–6.
11. A. Cardoso, L. Dias, E. Fernandes, et al., "Development of novel high-density system integration solutions in FOWLP-complex and thin wafer-level SiP and wafer-level 3D packages," 67th ECTC, May 2017, pp. 14–21.
12. *Fan-out Wafer-level-Packaging*, edited by John H. Lau, Springer ISBN 978-981-10-8883-4 ISBN 978-981-10-8884-1 (eBook) https://doi.org/10.1007/978-981-10-8884-1.
13. *Advances in Embedded and Fan-Out Wafer-Level Packaging Technologies*, edited by Beth Keser and Steffen Kroehnert, WILEY 2019, https://lccn.loc.gov/2018034374.
14. P. Coudrain, et al., "Active interposer technology for chiplet-based advanced 3D system architectures," 2019 IEEE 69th ECTC, Las Vegas, NV, USA, 2019, pp. 569-578, doi: 10.1109/ECTC.2019.00092.
15. K. K. Ryu, K. A. McIntosh, S. Rabe, J. Ciampi, R. D. Lambert, R. Das, et al., "Rapid prototyping of single-photon-sensitive backside-illuminated silicon avalanche photodiode arrays," Proc. SPIE 10980, Image Sensing Technologies: Materials, Devices, Systems, and Applications VI, 109800L (13 May 2019); https://doi.org/10.1117/12.2520798.
16. K. K. Ryu, B. F. Aull, M. Collins, N. Pestana, J. Ciampi, R. Das, et al., "Geiger-mode avalanche photodiode arrays fabricated on SOI engineered-substrates," Proc. SPIE 12089, Advanced Photon Counting Techniques XVI, 120890C (30 May 2022); https://doi.org/10.1117/12.2618610.
17. B. Tyrrell, K. Anderson, J. Baker, R. Berger, M. Brown, C. Colonero, et al., "Time delay integration and in-pixel spatiotemporal filtering using a nanoscale digital CMOS focal plane readout," IEEE Trans. on Electron Devices, vol. 56, no. 11, Nov. 2009, pp. 2516-23.

Biography

Rabindra N. Das is a Member of the Technical Staff in the Advanced Technology Division of MIT Lincoln Laboratory, Lexington, MA. Prior to MIT, he was a Principal Engineer at Endicott Interconnect Technologies (formerly IBM Endicott). Dr. Das has 20 years of experience in microelectronics packaging development for applications ranging from HPC to medical to quantum electronics. He holds 48 patents and more than 115 publications. Email Rabindra.Das@ll.MIT.edu

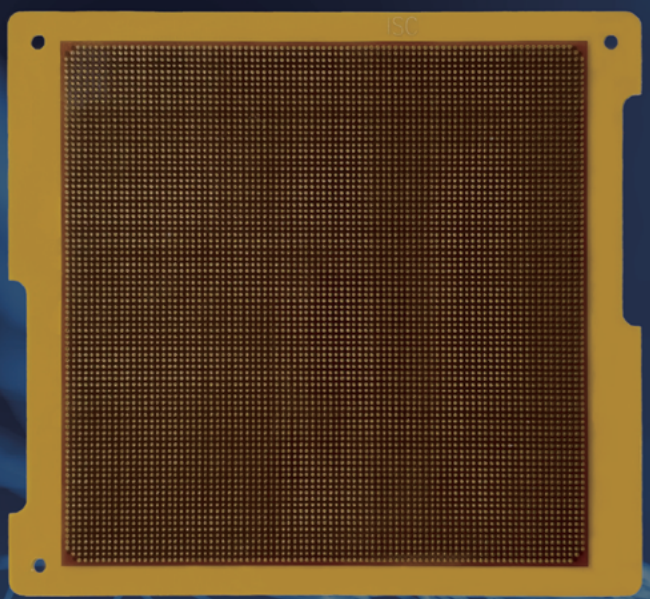
Jason Plant is a Member of the Associate Staff in the Advanced Technology Division of MIT Lincoln Laboratory, Lexington, MA. He is responsible for developing advanced compound-semiconductor optoelectronic components including the development and hybrid integration of high-power semiconductor lasers, optical amplifiers, photodetectors and optical modulators with silicon nitride based photonic integrated circuits (PICs).

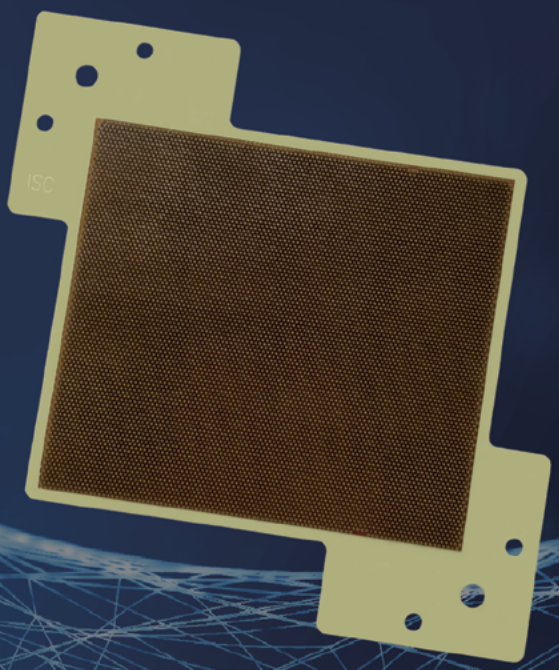
Global No.1
Total Test Solution Provider ISC

GLOBAL INNOVATION

System Package Test Solution
iSC-WiDER

- Up to 350um warpage can be covered
- Improved socket structure allowed 30% lower force
- Stroke increased by 50% to accommodate larger device size over 100x100





GLOBAL INNOVATION ISC

HEAD OFFICE 26, Geumto-ro 40beon-gil, Sujeong-gu, Seongnam-si, Gyeonggi-do, Republic of Korea

CONTACT sales@isc21.kr / +82 31 777 7675 / www.isc21.kr

LEENO

www.leeno.com

Spring Contact Probe

92㎛ Pitch~

200K Pins in 7 Days!

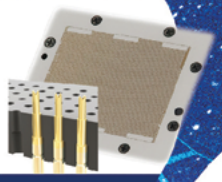
Probe Head

Probe Head Pin Counts 30K

10K Pins in 7 Days!

2 Probe Heads in 20 Days!

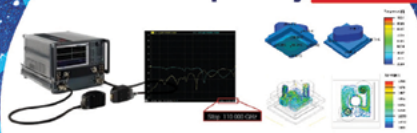
92㎛ Pitch~



RF/Thermal Simulation

110GHz Actual Measurement Capability

Report in 3 Days!



Actual Measurement

2-Port: 110GHz

4-Port: 43GHz

Logic & SLT Socket

10 Sockets in 7 Days!

Coaxial Probe 100GHz

Specification

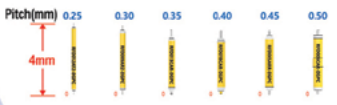
Frequency : 80GHz(BGA), 100GHz(QFN, LGA)

Pitch : 0.25mm~

Crosstalk : -60dB

Impedance : 50Ω±10%

1K Pins in 7 Days!



MP Socket

5 Sockets in 10 Days!

100GHz Coaxial Socket

Excellent RF & Electrical Performance!

Large Device Socket/Lid

Pin Counts 5K~

Device Power ~1,200W

10 Sockets in 10 Days!

LEENO LEENO INDUSTRIAL INC.

GLOBAL LEADER LEENO

HEAD OFFICE
10 105beon-gil MieumSandan-ro Gangseo-gu, Busan Metropolitan City, Korea

CONTACT
hskang@leeno.co.kr / +82 10 8737 6561
kimmj@leeno.co.kr / +82 10 5314 6560

Readying the supply chain for chiplets and heterogeneous IC packaging

By Mike Kelly, Dave Hiner, Ruben Fuentes, Jonathan Micksch, Vineet Pancholi [Amkor Technology, Inc.]

Products using heterogeneous integration (HI) and chiplets are here, they are in production, and the current trend is for more products and more customers to develop and qualify their products at outsourced semiconductor assembly and test suppliers (OSATs) and foundry providers.

Combinations of processors and memories have been around for years, culminating at the high end with processors and high-bandwidth memory (HBM) to address a fast-growing market for artificial intelligence (AI) algorithm training. Now, the functional de-partitioning of die functions into chiplets is enabling a broader and more potent impact on future designs. Chiplet approaches allow product performance increases to continue at a cost point that is still compelling. Total silicon costs can be lower due to better yields for smaller chiplets, and the opportunity to use a mixture of silicon process nodes to further optimize the cost of the silicon. Integrated circuit (IC) packaging for heterogeneous and chiplet approaches is more expensive, but this rise in package cost is offset by reduced total silicon expenses and favorable time-to-market advantages.

Moving to chiplets and HI implementations has required a new infrastructure to be established for IC and package design, IC and package fabrication, and electrical test. Design tools need to comprehend multiple ICs in 2D and 3D physical configuration, functional device electric test (E-Test) and higher power densities. IC packaging to support chiplets, and heterogeneous constructions, have been a primary focus for OSATs and foundries in recent years. Multi-die products must be integrated into one functional unit. Such integration is accomplished using higher-density integration approaches, namely wafer-scale (chip-on-wafer, CoW) high-density modules and high-density multi-chip modules (MCMs), or both. High-density modules are then attached to the IC package substrate in a production environment like the venerable system-on-chip (SoC) flip-chip ball grid array (FCBGA) packages, but with several key customizations.

Currently, the packaging methodologies employed in both production and development encompass the following: 1) 2.5D through-silicon via (TSV) modules relying on silicon interposers, i.e., 2.5D TSV; 2) modules utilizing high-density fan-out (HDFO) multi-layer redistribution layer (RDL) approaches; or 3) modules featuring bridges. These two-dimensional constructions can be used for discrete die combinations as well as combinations of discrete die and 3D die stacks. The sections below discuss this developing landscape, starting first with the 2.5D TSV.

2.5D TSV silicon interposers

2.5D TSV has been in high-volume manufacturing (HVM) at Amkor since 2017. The process flow begins with a full "TSV-reveal" capability, starting with full thickness interposer wafers from one of the foundries, thinning to reveal the Cu TSVs, followed with an inorganic passivation step, under bump metallization (UBM) and interposer backside bumping.

This product space is dominated by high-performance processors working in combination with high-bandwidth DRAM memory (HBM), including HBM2, HBM2e and HBM3. 2.5D TSV was one of the first modern heterogeneous integrations using a high-density module to permit integration of the processor and DRAM in the package itself. The silicon-based interposer uses a Cu backend foundry fabrication process, and this enables 1-2µm lines and spaces inside the IC package. This has been critical to enable a very wide parallel data base for HBM communication. In many ways, it was this process development to enable 2.5D TSV taking place in 2015-2018 that set the stage for a new class of high-density module-based products. These new approaches targeted the next wave of heterogeneous integrations as chiplets that were being designed and qualified in just the last few years. In addition to the latest processing know-how developed to support the TSV reveal process, a new class of ultra small Cu pillar bumps was required to support bump pitches in the 40-55µm range. This requires advanced plating tools and chemistries.

Many of the foundational technologies noted above were used as is or extended to intersect other high-density modules such as HDFO and bridge-based product developments (**Figure 1**).

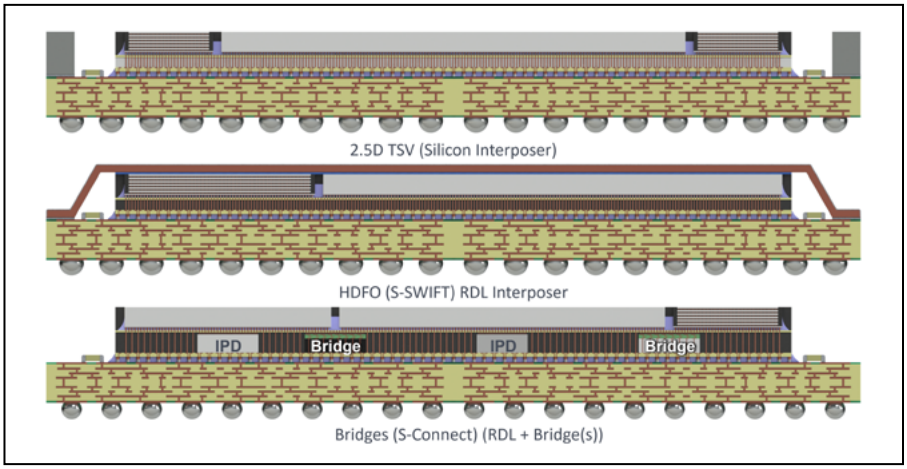


Figure 1: High-density module-based products.

HDFO

Modules based on HDFO interposers have been internally qualified and several of our customers' products are in qualification. Our internal terminology for HDFO is S-SWIFT (Substrate Silicon Wafer Integrated Fan-out Technology). HDFO technology is being applied to many markets and use applications, ranging from high-performance compute and AI, to automotive applications and beyond. Chiplet architectures are leading to a push for advanced packaging design rules that are enabled by HDFO and other module-based solutions. Our fabrication of this HDFO interposer is supported in both a chip-first and a chip-last construction. Each fabrication method has advantages and disadvantages, and in many cases the end customer may have a specific requirement for a given flow or construction.

Chip first, as the words imply, involves the placement of the active silicon chips at the beginning of the module fabrication. Chips are attached face up on a wafer carrier and the multi-layer RDL process is completed with direct metal connections to the active silicon. Chip last involves the fabrication of the RDL first followed by a traditional chip-on-wafer assembly process using solder joints and underfill. Once the modules are completed, they are assembled to a package substrate in a manner similar to non-module-based products. **Figure 2** shows a high-level comparison of these two approaches.

When considering a product intercept into HDFO, we encourage the use of a test vehicle (TV) that matches the design complexity of the final product. This TV phase can make use of the hierarchy of design rules within the design to test boundaries of HVM design rules and overall process capability. This process provides validation for the first product intercept and gives a first check on next-generation requirements. These TV designs utilize a series of daisy chains that can test historical areas of high stress, including die corners, die gaps, module corners, stacked vias and others. Having a well-defined daisy chain can give electrical data that can bolster the mechanical modeling work that is equally critical to the TV phase of development. Predictive modeling can provide key insights to the design and material choices at the start of a program, and it can be valuable at identifying challenges and solutions during development. This TV strategy for development has been used over many successful programs and is highly recommended.

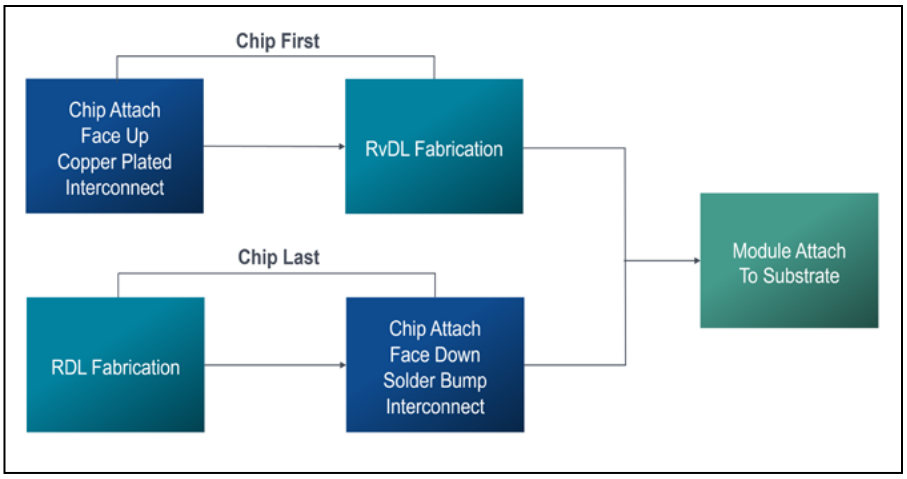


Figure 2: General chip-first vs. chip-last flow.

Our current SWIFT technology can support module designs down to 1.5-micron line and 1.5-micron space and with layer counts between 2 and 6 layers. Module sizes can be supported from smaller modules, to modules that are larger than reticle size using reticle stitching. **Figure 3** shows a typical 6-layer RDL module with the top die solder joint fabricated in a chip-last approach. **Figure 4** shows the subsequent module to substrate solder joint.

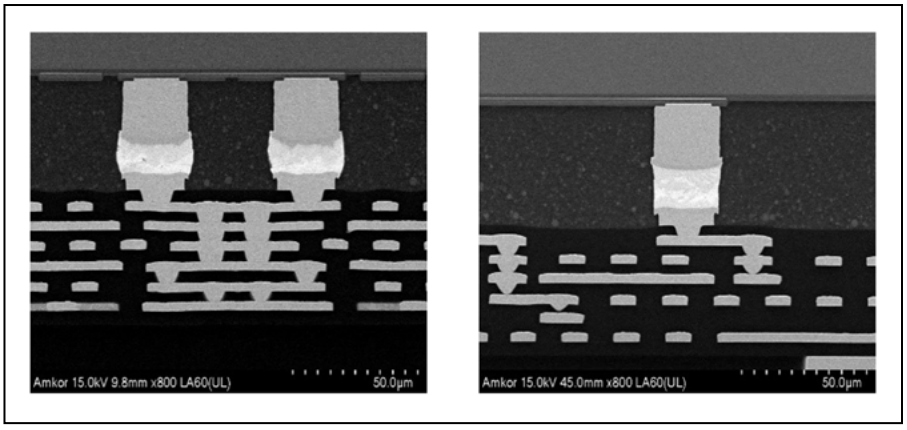


Figure 3: Test vehicle die connected to HDFO.

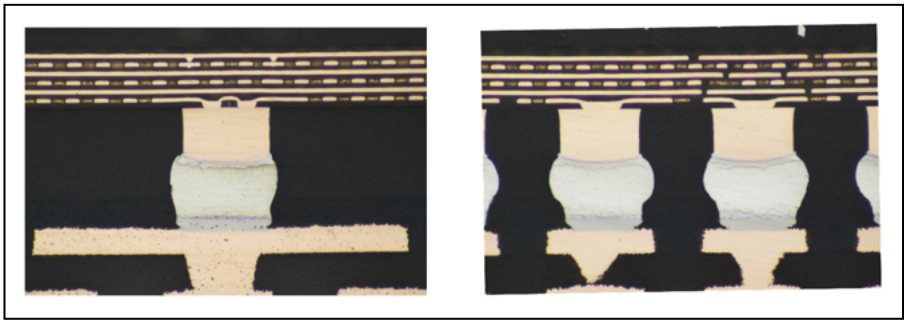


Figure 4: Test vehicle HDFO module connected to package substrate.

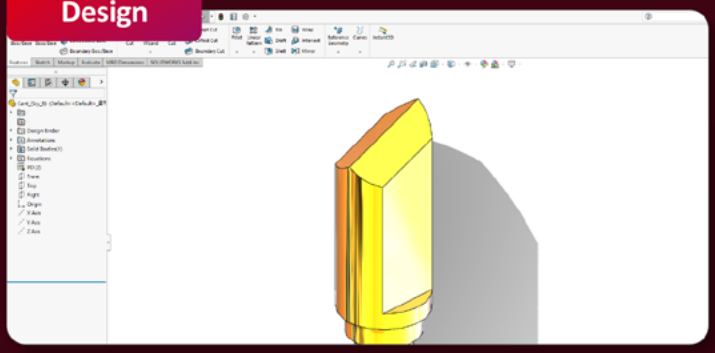
WinWay

Capable of producing probes with 75um outer diameter

Spring Probe

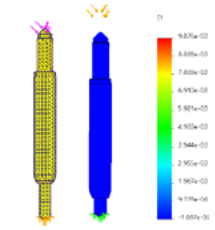
Customized design and manufacturing spring probe capability are available for different applications, we provide the most fitting materials, coatings, and probe structures for your ever changing IC packages and testing conditions.

Design

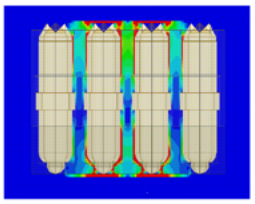


Customizable plunger

Simulation



Stress simulation



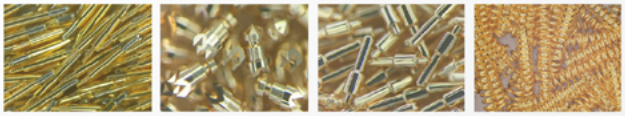
Electrical simulation

Ensuring probe performance and quality by simulation

Manufacturing

CNC

- ∨ Self-manufacturing parts
- ∨ Automated recipe configuration



Plating

- ∨ Various plating materials :

  Au/ AuCo/ AuCu/ Rh/ Ag/ PdNi/ NiP/ NiCo

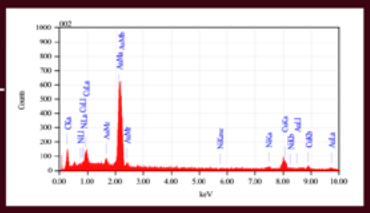
  Choose coating based on product attributes: varying in anti-adhesion, anti-oxidation, wear resistance, and corrosion resistance

Assembly

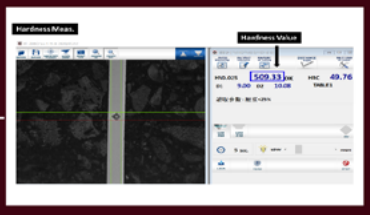
- ∨ 100% automation assembly and inspection (AOI)

Analysis

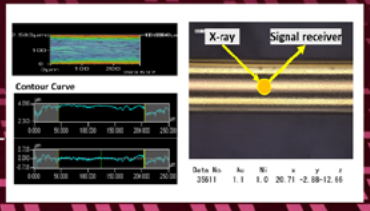
Material EDX



Mechanical property Hardness test



Plating Roughness Meas & Thickness Meas



Bridges

The next iteration of HDFO enables a 3D capability using bridge silicon and other embedded components under the active silicon. The fundamental building blocks developed in HDFO interposer fabrication are extended with the inclusion of embedded bridging components or other devices. These embedded components can have a basic ultra-high density routing function, such as between two chiplets, or they can be active or non-active components, such as an integrated passive device (IPD). The embedded silicon components may or may not contain TSVs allowing vertical connections through the embedded components.

To facilitate this new structure, several key process capabilities are necessary including accurate component placement, tall copper pillar plating, and warpage control. By using this approach, it is possible to utilize the high-density routing capability from the wafer fab to interconnect between chiplets, allowing a reduction in the RDL layer count in the HDFO interposer. The ability to add discrete component functionality is an added benefit for many customers. There exist three primary drivers for S-Connect in the market. First, silicon can offer sub-micron routing capability, which allows denser routing for system design. Second, the use of bridge die for the chip-to-chip connections can reduce the layer count requirement on the HDFO routing for the interposer and increase overall yields. Lastly, the process allows the placement of performance-enhancing non-bridge components such as silicon IPDs, providing closer proximity to key areas of the active silicon.

When engaging with our customers, we consider the tradeoff analysis between HDFO and S-Connect. Sourcing of the silicon bridge and potential IPDs is a key element of this assessment. A supply of these components is as critical to the decision as the active silicon. Process maturity is a key element because the industry has been in production with 2.5D TSV modules for many years. HDFO is, likewise, more mature than the S-Connect technology. These maturity levels can often influence the decision making of our customers.

Our version (S-Connect) of the embedded bridge HDFO module on substrate is shown in **Figure 5**. We have completed internal qualification of a chip-last non-TSV bridge TV, and we are working with several customers on plans to utilize these advanced capabilities.

Design support

Our SmartPackage Package Assembly Design Kits (PADK) (**Figure 6**) are ideal for preparing the design layout for a successful supply chain experience. The ability to build the various high-volume or advanced manufacturing and assembly design rule requirements into the design layout early in the design flow process can significantly facilitate a smooth supply chain support path. Notably, this solution is compatible with multiple electronic design automation (EDA) design tools, further securing the ability to align with many independent design workflows.

Two prevalent design workflows in contemporary practices are the Full-OSAT flow and the OEM/Fabless-focused project-based design processes. In the Full-OSAT design workflow, we have comprehensive design services and verification sign-off in accordance with the customer's instructions. In the OEM/Fabless design workflow, we collaborate with users who opt to design their package layout and necessitate the capability to finalize

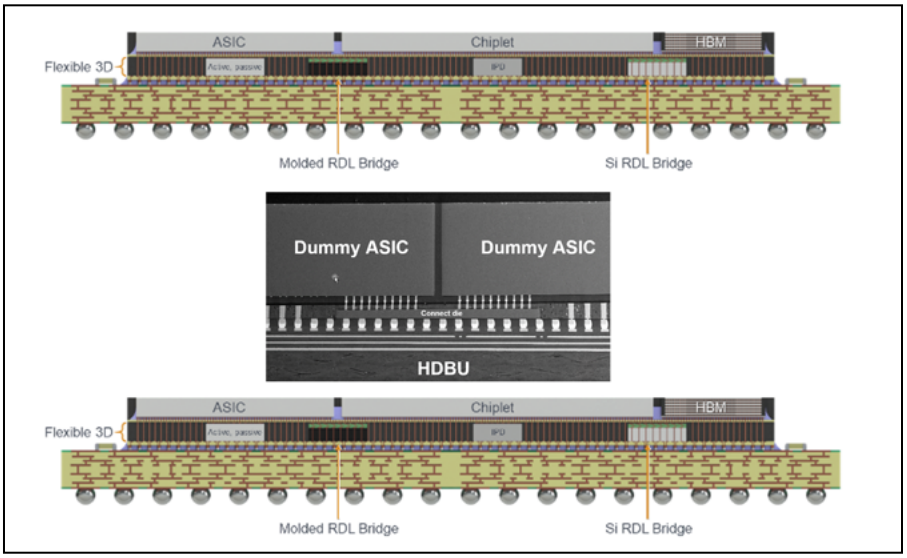


Figure 5: Amkor's S-Connect Technology.

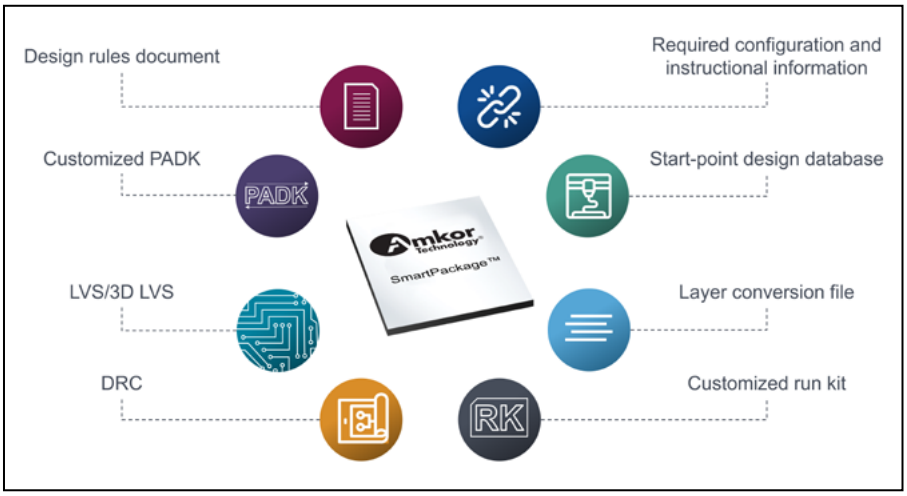


Figure 6: Amkor's SmartPackage PADK fine-tuned customized design rule requirements.

their verification signoffs before providing production data to us for manufacturing and assembly processes.

This ability for accurate implementation of design rules, manufacturing and assembly constraints into the EDA and computer-aided manufacturing (CAM) design tools, with the SmartPackage PADK enables users to quickly highlight design restrictions early in the design process limiting the number of product design cycles. To experience these benefits, some investment in software and hardware infrastructure may be necessary. Depending on the current design environment, a high-powered Linux server will need to be employed. This system will be tasked with running the extensive manufacturing and assembly design rule constraints on the manufacturing data that is destined for fabrication.

We pioneered the development of PADKs in 2016. Users receive support for PADKs to integrate them as a robust component of their device development approach. The OSAT can provide guidance for package design layout direction, provide specific application training, and provide ongoing design review support, using the latest software to incorporate their packaging knowledge and experience into SmartPackage PADK elements.

Which elements of a package assembly design kit exert the most significant influence on design readiness? The emphasis typically revolves around three key areas. The first is the EDA start-point database, the second is design rules constraint (DRC) sign-off verification, and the third is the ability to validate the production data with the connections list requirements.

A feature of SmartPackage PADK is the capability to fine-tune the specific customized design-rule requirements necessary for the device or design layout needs. The benefit of device-specific design rule decks versus a fixed-node rule deck system is that there is no need to provide careful and extensive waivers to receive a passing verification report.

Test

We have been providing test services for heterogeneously-integrated products since the inception of the 2.5D TSV development cycle. The systemic approach of designing and evaluating the test vehicles allows test engineers to develop tests for critical aspects of the design before the live product. There are a handful of test challenges that are common to all heterogeneous chiplet packages. Chiplet interconnect integrity is an important one.

Signal and power delivery to every chiplet within the package requires careful layout, design and test during the manufacturing process (**Figure 7**). Package material types used, and the package constructs described earlier in this article, impact the interconnect performance between the chiplets and the pins exposed at the package level. This includes both static connection quality with continuity, leakage and transient ac timing, impedance matching and signal crosstalk. Thermal performance of each of the chiplets also impacts production testing. Thermal gradients due to non-uniform chiplet temperatures are unavoidable. In a carefully designed overall product architecture, design for test (DFT) has access to all functional aspects of the product, which is an important consideration.

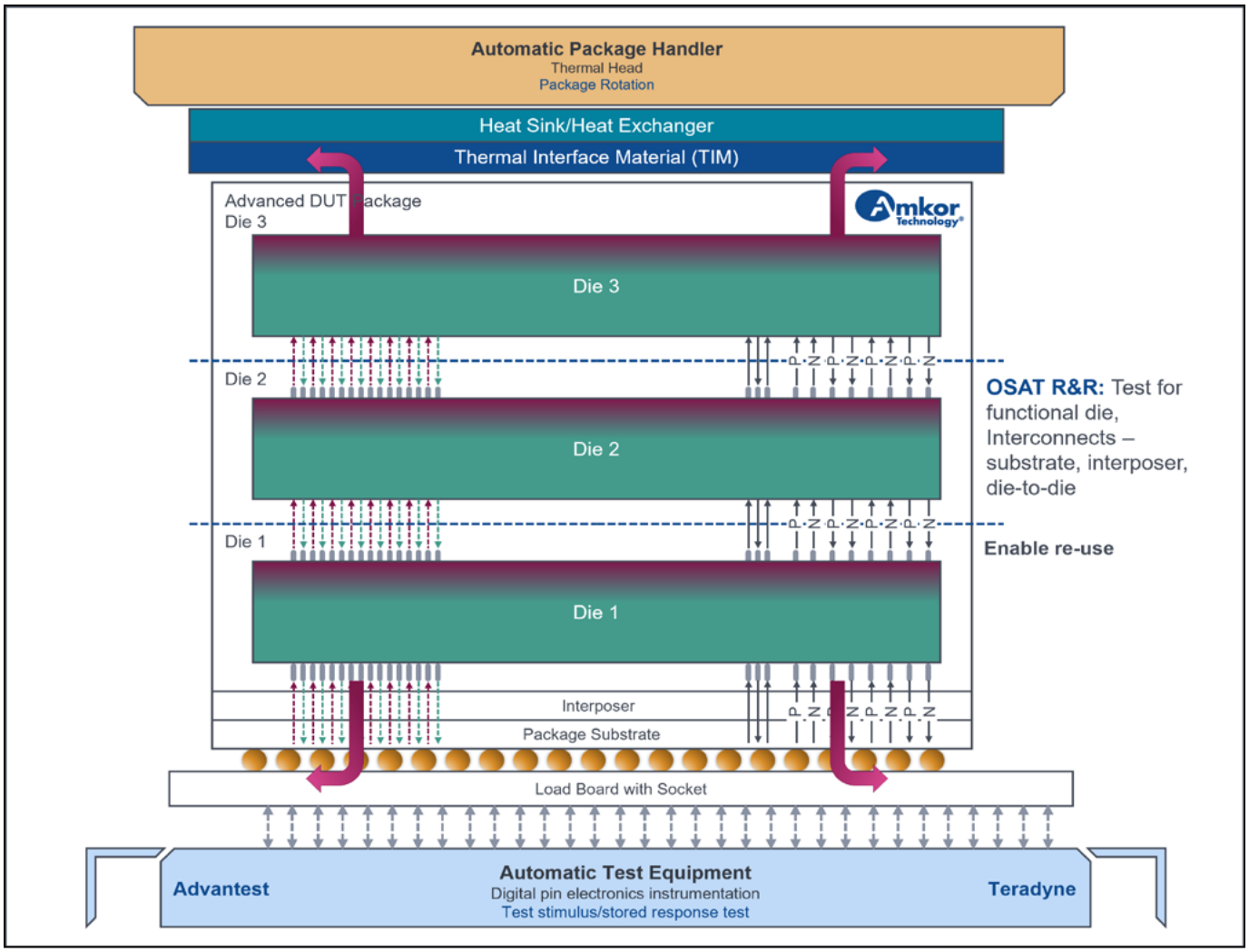


Figure 7: Production test for power, data I/O, bias and clocks in multi-die packages.

IEEE1838 [1] is one such standard that helps during the architecture phase of the product. Test access to each chiplet and all the functional blocks within is a "must have," to allow full production testing.

IEEE-1687 [2] describes the test methodology for accessing instrumentation embedded within a semiconductor device. Electronic data automation (EDA) vendors have defined intellectual property (IP) blocks to monitor environmental attributes including process, voltage and temperature (PVT) on-die. They have a similar concept of adding sensors within the logic design and have documented numerous benefits to the overall manufacturing workflow. PVT sensor placement in the vicinity of the thermal congestion is vital to analyze the severity and sensitivity of thermal densities

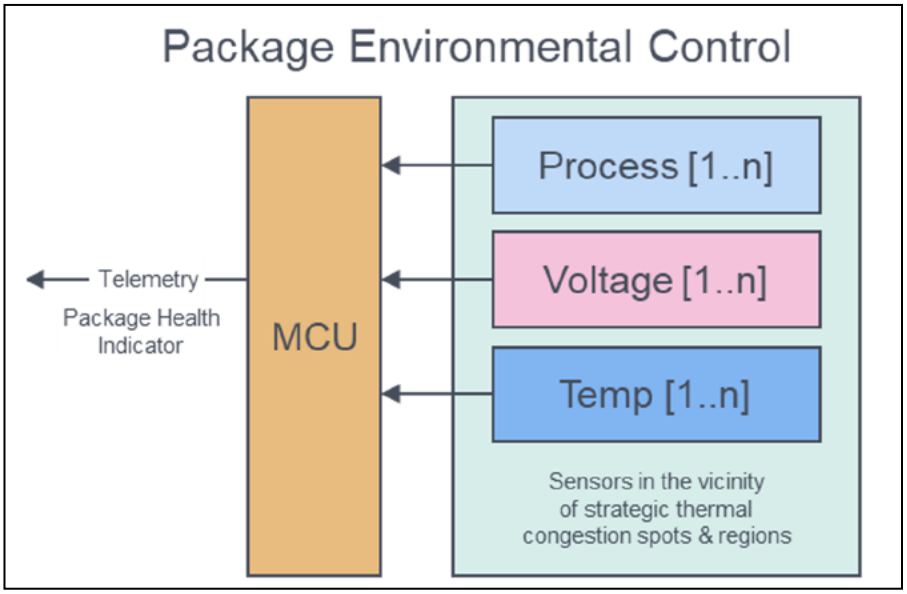


Figure 8: Block representation of Package Environmental Control with a variety of sensors that allow a telemetry stream to monitor package health during active operation, including the production test process.

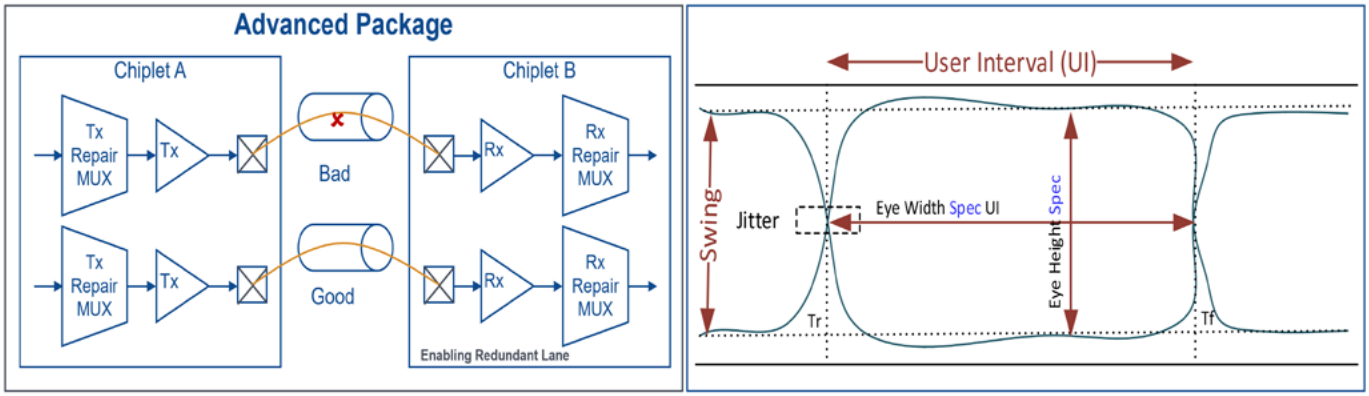


Figure 9: Standards are driving provisions for redundant lanes and on-chiplet mission mode eye characterization and margining.

within the package architecture and design implementation. It is simpler and more cost effective to sprinkle these sensors within the die rather than separately including them on-package. The telemetry stream under various corner cases is read back and analyzed to allow verification against the simulations. **Figure 8** shows a block representation of the Package Environmental Control for monitoring.

As an OSAT, production test workflow simplification is vital to our company. Industrywide test methodology standardization efforts are helping. For instance, the Universal Chiplet Interconnect Express (UCIe) standard includes constraining the shoreline on chiplets to be fixed. This allows for place and route simplification and interoperability.

The UCIe standard further includes design guidelines for redundancy repair and on-chiplet mission mode eye characterization and margining (**Figure 9**). Redundancy repair allows for yield recovery of packages that would have been a reject without this capability. Eye margining capabilities in production testing, allows the product architects and designers to monitor process variations and make systemic improvements, generation over generation.

A controlled, managed and repeatable production test environment ensures accurate feedback for future product design iterations and consistent yield. The test, package handling and optical inspection supply chain is continuing to refine the metrology to account for all the identified failure pinch points.

Thermal considerations

Power density continues to increase, and putting more functional silicon into a smaller volume requires close attention to the power dissipation path. We are developing optimized package-level solutions to assist in this effort. Polymer-based thermal interface materials (TIMs) continue to be a mainstay, but for the higher end power levels, metallurgical TIMs may be required, including for 2.5D TSV, HDFO and bridge modules. This is an active area of development.

Summary

The transition to heterogeneous chiplet-based integrations is well underway. The value proposition for the chiplet approach is strong, as evidenced by the recent successful market entries in the compute and AI market spaces. Heterogeneous and chiplet-based IC packaging plays a key role in this evolution, with 2.5D TSV, HDFO and bridge approaches providing a cost-effective path for these integrations.

Acknowledgments

SWIFT is a registered trademark of Amkor Technology, Inc. S-SWIFT and SmartPackage are trademarks of Amkor Technology, Inc. UCIe is a trademark Universal Chiplet Interconnect Express.

References

1. IEEE Standard for Test Access Architecture for Three-Dimensional Stacked Integrated Circuits," IEEE Std 1838-2019, vol., no., pp.1-73, 13 March 2020, DOI: 10.1109/IEEESTD.2020.9036129.
2. "IEEE Standard for Access and Control of Instrumentation Embedded within a Semiconductor Device," IEEE Std 1687-2014, vol., no., pp.1-283, 5 Dec. 2014, DOI: 10.1109/IEEESTD.2014.6974961.

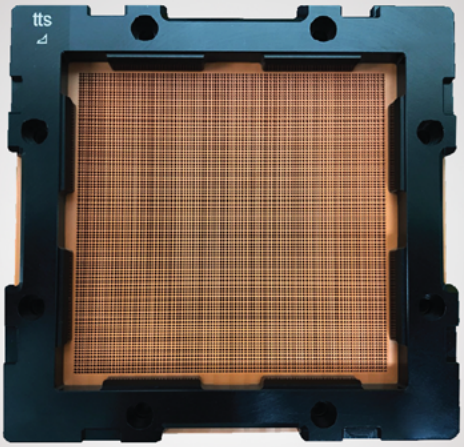
Biographies

Mike Kelly is VP, Chiplets/FCBGA Integration, at Amkor Technology, Inc., Tempe, AZ. Mike joined Amkor in 2005 and has led package developments for EMI shielding, thermally-enhanced packages, sensors and high-density MCM packages including 2.5D TSV and high-density fan-out (HDFO). He has worked in electronics and IC package design and manufacturing for 25 years. Mike has more than 40 patents in the field and holds master's degrees in Mechanical and Chemical Engineering. Email Mike.Kelly@amkor.com.

Dave Hiner is Sr. Director Package Development, at Amkor Technology, Inc., Tempe, AZ. Dave joined Amkor in 1997 and has held various management positions in quality, substrate technology and advanced interconnect and wafer-level package development. He has worked in the assembly and packaging industry for more than 30 years. Dave has 22 patents in semiconductor packaging and holds a bachelor's degree in Chemical Engineering from Arizona State U. and an MBA from U. of Phoenix.

tts

www.tts-grp.com

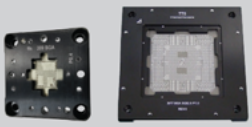


Epsilon Series

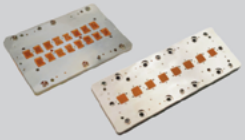
HIGH Pin Count

Stable test performance for large, high I/O count devices.

- *Pin coplanarity less than 0.35mm*
- *Pin compression up to 1.2mm*
- *Tri-temp testing between -40°C to +150°C*
- *Optimized for devices from 3000 to 10000 I/O's*
- *Pitch 0.80mm and greater*
- *Suitable for BGA and LGA devices*



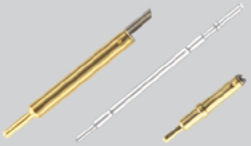
COAXIAL CONTACTOR



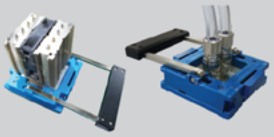
WLCSP PROBE HEAD

RF & mmWAVE CONTACTOR

STRIP CONTACTOR



PROBE PIN



HAND SOCKET LID

TestConX™ Expo

March 3-6 | Booth #28

BROCHURE

Making the golden connections in 3D-IC system-level design

By Tarek Ramadan [Siemens EDA]

3D integration is becoming a valuable alternative to the ongoing, yet challenging, transistor scaling known as Moore's Law. This is especially true for semiconductor products that are becoming limited by single-die designs (e.g., in terms of form factor, size, and technology node). Homogeneous 2.5D integrated circuit (IC) approaches have provided an acceptable yield and are currently helping bring about the era of chiplets—with its great emphasis on intellectual property (IP) reuse. Additionally, heterogeneous integration provides even greater value because different process nodes can be mixed in the same semiconductor product.

Unlike placing dies side-by-side in the 2.5D-IC approach, the 3D-IC approach allows the stacking of different dies vertically. This reduces undesired delays because the interconnects are shorter. True 3D stacking improves the form factor even more than is possible with 2.5D-IC and fan-out wafer-level packaging (FOWLP). Examples of both 2.5D-IC and 3D-IC integration are shown in Figures 1 and 2.

From a system-level design perspective, although each advanced packaging flavor (i.e., 2.5D-IC using silicon interposers, FOWLP, true 3D-IC) offers some unique challenges, some challenges apply to all of them. That is, the designer must ensure that the assembly is physically connected as expected compared to the "golden" design intent (captured as a system-level netlist). However, capturing a system-level netlist can be a challenge in the case of multiple substrates because each substrate usually requires a different design team, methodology, and/or format.

With current 3D-IC packaging technologies, because the system-level netlist (the 3D-IC design intent) drives system-level layout versus schematic (LVS) verification, designers need to ensure that the system-level netlist is "golden," i.e., it is the absolute reference of system connectivity. To demonstrate what golden means, assume that the designer is running LVS-type verification between the 3D-IC assembly layout and the 3D-IC assembly system-level netlist. When a connectivity error is reported in the LVS run, the designer should not have to stop and wonder whether the issue is coming from the system-level netlist or the layout physical routing. Although this undesirable situation might be inevitable in some cases (especially when a new design flow is being introduced and deployed for the first few projects), the goal of a system-level connectivity capture step is to raise the confidence in the netlist as much as possible to the point where the system-level netlist can be considered as a golden/frozen reference (when the design flow is mature enough). In the next section, the challenge of creating a golden system-level netlist is explored.

So, based on the above discussion, when running LVS-type verification between the 3D-IC assembly layout and the 3D-IC assembly system-level netlist, it is essential that any reported LVS error is related to the physical/logical routing—not the 3D-IC netlist. Therefore, a high level of confidence is expected in the connectivity

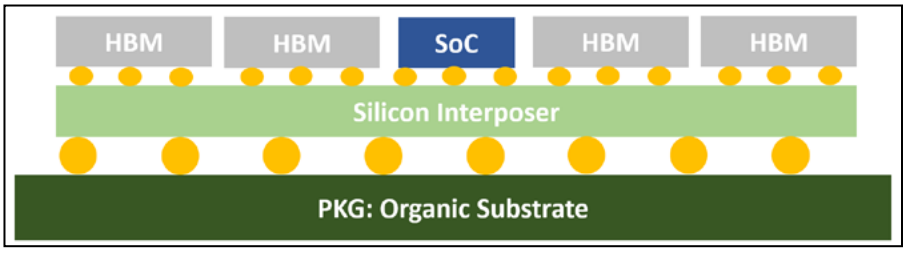


Figure 1: Heterogeneous integration using 2.5D-IC.

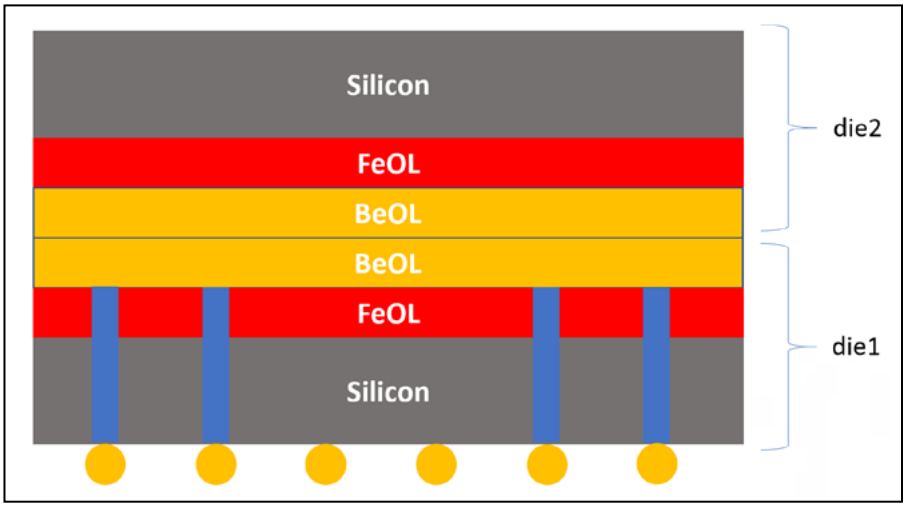


Figure 2: True 3D-IC achieved by stacking wafers.

information provided by the 3D-IC system-level netlist.

Another challenge when verifying the connectivity of a multi-substrate 3D-IC design is the project version-based connectivity exceptions, i.e., the need for creating shorts/opens while physically implementing the substrates. Designers need to treat these opens/shorts as "expected" errors and differentiate them from the "real" and "undesirable" errors.

Finally, there are other challenges associated with the "design dependent" physical verification nature of 3D-ICs. That is, an accurate assembly description is required so that LVS-type verification can be executed accurately.

Transitioning to advanced system-level connectivity flows

In the case of 2.5D-IC systems, designers usually build silicon interposers using similar approaches that are used to build digital system-on-chip (SoC) devices (using traditional, automatic place-and-route electronic design automation (EDA) tools), which means that the interposer connectivity is usually captured in a Verilog netlist format. As far as the designer can tell, this is the best system-level golden connectivity reference that he or she can generate using the traditional IC flows.

Now, however, when graduating to 3D-IC system-level design tools and approaches, how will designers know they built the connectivity correctly? They will need to employ equivalence checking, with which they can compare the output of the traditional connectivity capture approach (usually a single domain—CDL, Verilog, or CSV) vs. the output of a new, system-level connectivity capture approach (multi-domain). In other words, a quick, automated, and flexible "netlist-versus-netlist" comparison is required. Once the comparison results are clean, the designer will have more confidence in using the newly introduced connectivity capture flow.

To help designers make this transition, we developed an automated approach—using Siemens's Xpedition Substrate Integrator (xSI)—in which the designer can compare their traditional flow netlist (SPICE, Verilog, or CSV) versus the system-level assembly netlist generated from xSI (CSV). This automated netlist-versus-netlist comparison is done using Calibre tools accessed through an add-on feature integrated with xSI. Figure 3 shows the interactive graphical user interface (GUI; wizard) that is used for the netlist-versus-netlist comparison.

Users can see whether the xSI system-level connectivity matches their traditional flow connectivity by generating the netlist-versus-netlist "runset" from xSI, executing the comparison, and viewing the generated comparison report. If the report is clean, this means that xSI correctly built the system connectivity as intended.

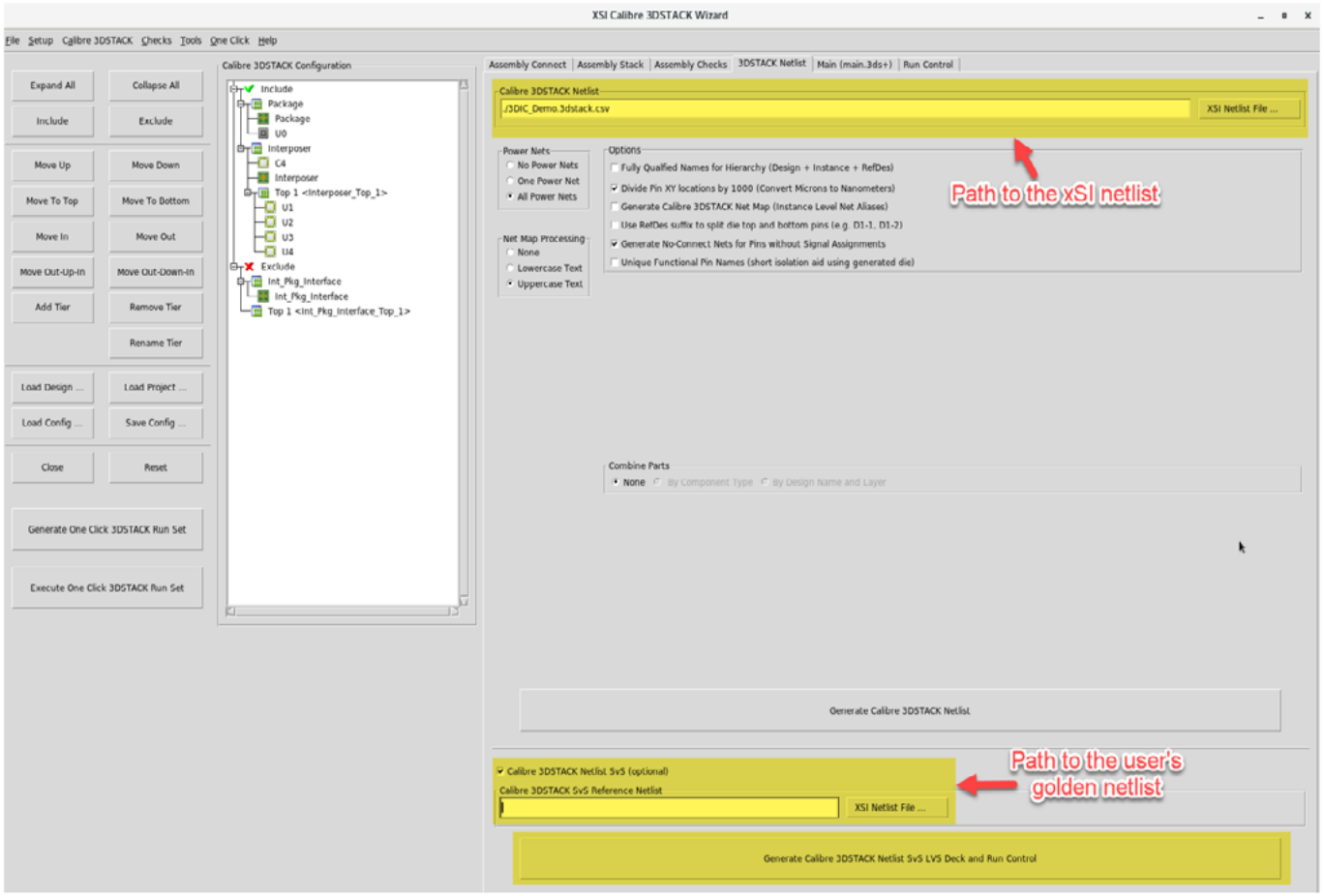


Figure 3: xSI wizard for netlist-versus-netlist feature.

System connectivity exceptions in 3D-IC multi-substrate assemblies

In some cases, a system source netlist is not enough for LVS verification of 3D-IC multiple substrate systems. In these cases, designers need to "short" some signals together in the substrate layout. These shorts create a mismatch between the layout and the source netlist. Although these differences are reported in the LVS comparison results, they are intended by the user (usually temporarily).

For the user to differentiate between the intended LVS issues versus the real LVS issues, the intended LVS issues need to be somehow waived from the LVS results. One obvious solution is to modify the system source netlist to match the physical layout implementation. However, this is not desirable because the system source netlist is required to be golden and frozen across different physical implementations and iterations. Therefore, designers need a quick, automated, and flexible way to handle connectivity exceptions in a 3D-IC design that enables them to waive intended shorts in the LVS comparison results. Our Calibre 3DSTACK supports a "net mapping" feature that can be utilized to account for the intended shorts. The shorts list can be considered a net map file, and it can be automatically included in the runset using a wizard that is integrated with xSI.

Another challenge can arise when an assembly includes both an interposer and a package substrate for the same connection; the IC design team (building the silicon interposer) may use a different net naming methodology than the packaging team (building the organic substrate). The system-level designer then ends up in a situation where the same port name can be assigned to two different net names. For example, the same die-to-ball grid array (BGA) connection can be assigned to two different net names: C4_PKG (packaging team naming) and C4_INT (interposer team naming).

Our xSI can recognize a connection across substrates, even when two different net names are assigned to the connection. This is achieved by applying the xSI interface part feature. The interface part function is used to connect two different substrates in xSI (called "floorplans" or designs). An example of an interface part is highlighted (on the left) in Figure 4.

As seen in Figure 4, there are two substrates: "Interposer" and "Package." For the "TEST_CLK" signal, there are two different net names: "TEST_CLK" in the interposer domain and "pkg_TEST_CLK" in the package domain. Although the net names are different, the "TEST_CLK" signal is correctly tracked across the system due to the existence of the "C4P" part (the interface part).

Another thing to consider is tool support for known shorts. In multi-substrate 3D-IC assemblies it is a common practice for designers to short some unneeded signals to the ground plane or short different power planes

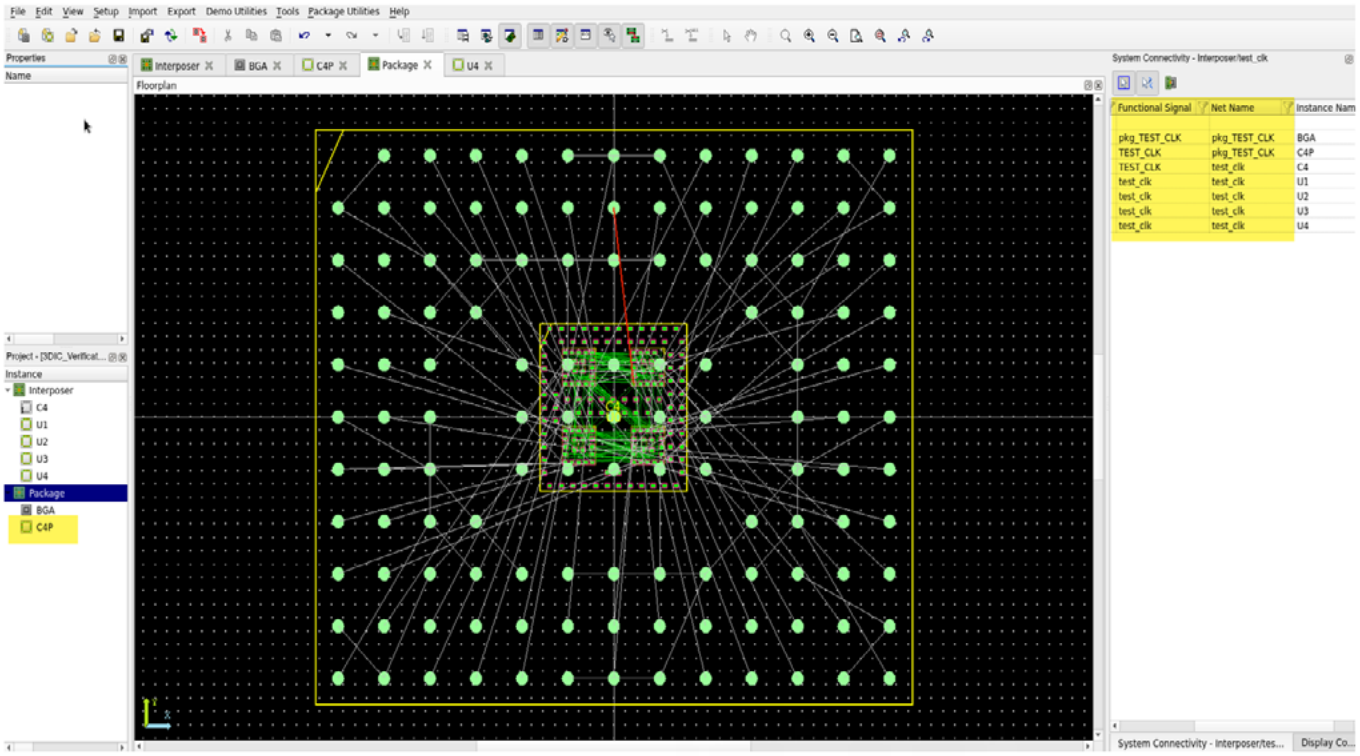


Figure 4: An interface part in xSI.

together (for example: short VDD_SENSE and AVDD on the interposer to VDD on the package) while these power nets are disconnected in the system source netlist. Users often include this information in a list (e.g., text file, CSV). Without including this list in the xSI/Calibre 3DSTACK flow for system-level LVS, many LVS errors are reported. In this case, it is difficult for users to distinguish between the intended errors and the real errors. To resolve this challenge, the shorts list should be considered while generating the source netlist for Calibre 3DSTACK so that the nets present in the shorts list are merged.

Calibre 3DSTACK allows a net mapping feature, as presented in the previous section. This feature can be utilized to account for the shorts list. The shorts list can be considered as a net map file, and it can be automatically included in the Calibre runset using a wizard that is integrated with xSI, as shown in Figure 5.

Assembly verification challenges

Let's consider a typical silicon interposer case in which one foundry is manufacturing both the interposer and the dies included in the 2.5D-

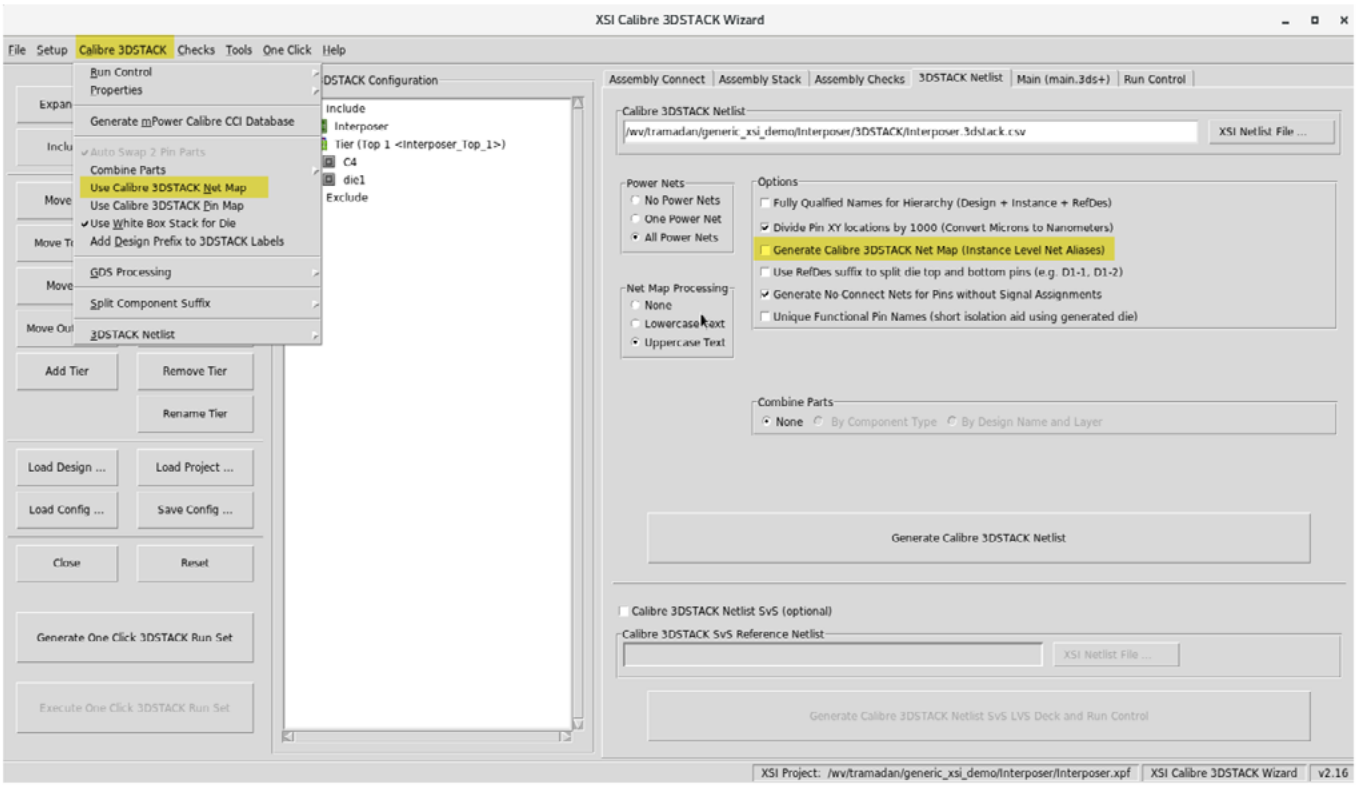


Figure 5: Inserting a net map file in Calibre 3DSTACK runset.

IC system. Although the physical verification of the individual dies is a well-established process, the physical verification of the full assembly requires two more steps: 1) Interposer design rule checks (DRC) and LVS; and 2) Die alignment and connectivity to the interposer.

The foundry can provide the interposer DRC and LVS decks as part of the standard physical design kit (PDK). This is because the interposer formats and design methodologies are somehow similar to standard die(s). However, die alignment and connectivity to the interposer imposes a challenge because the die's location and orientation can change from one design house to another and even from one project to another in the same design house. Figure 6 shows an example of different assemblies that include the same interposer. Consequently, inter-die DRC and inter-die LVS through the interposer might be completely left to the system-level designer to own and execute.

Multiple manufacturers

The challenge of full assembly verification exponentially increases when verification encompasses multiple substrates from multiple manufacturers in an assembly. In this case, the assembly verification considers a silicon interposer (from a foundry) stacked on top of an organic substrate (from an outsourced semiconductor assembly and test [OSAT] supplier). No single manufacturer can provide a PDK for the full system. The best-case scenario is when the foundry provides a standard interposer PDK and the OSAT provides a simple way of checking the substrate connectivity. However, the assembly verification of the whole system (die to interposer

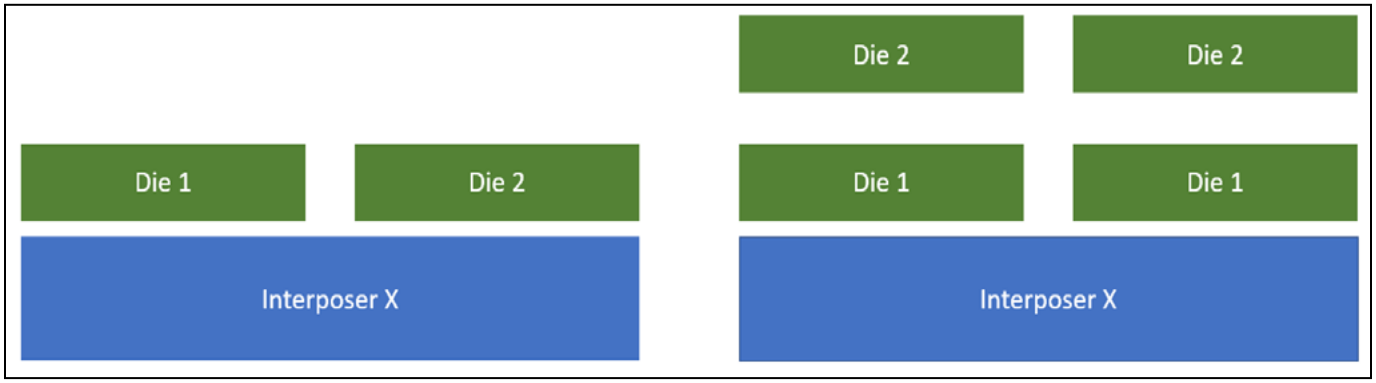


Figure 6: Different assemblies that include the same interposer.

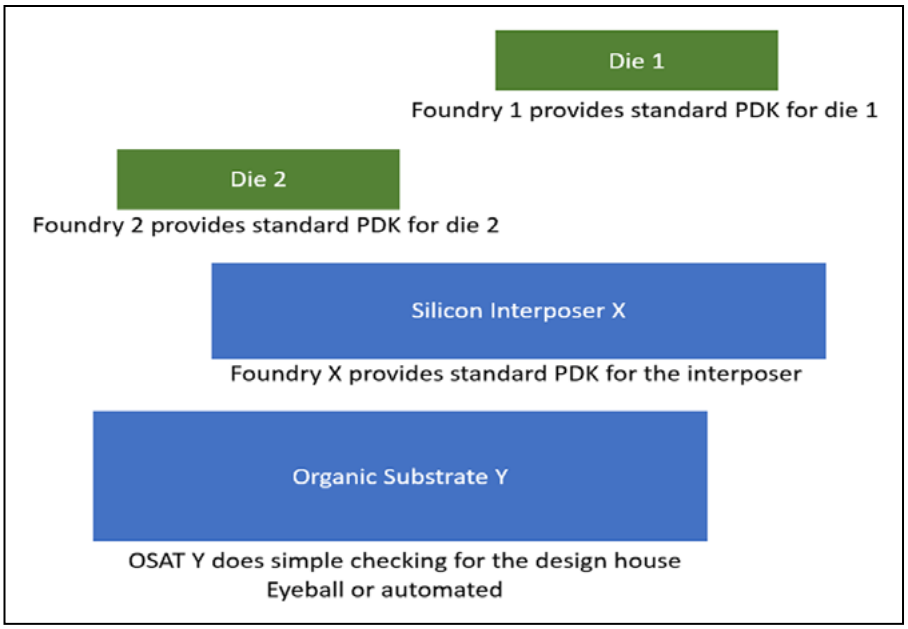


Figure 7: "Standalone" checking for every component.

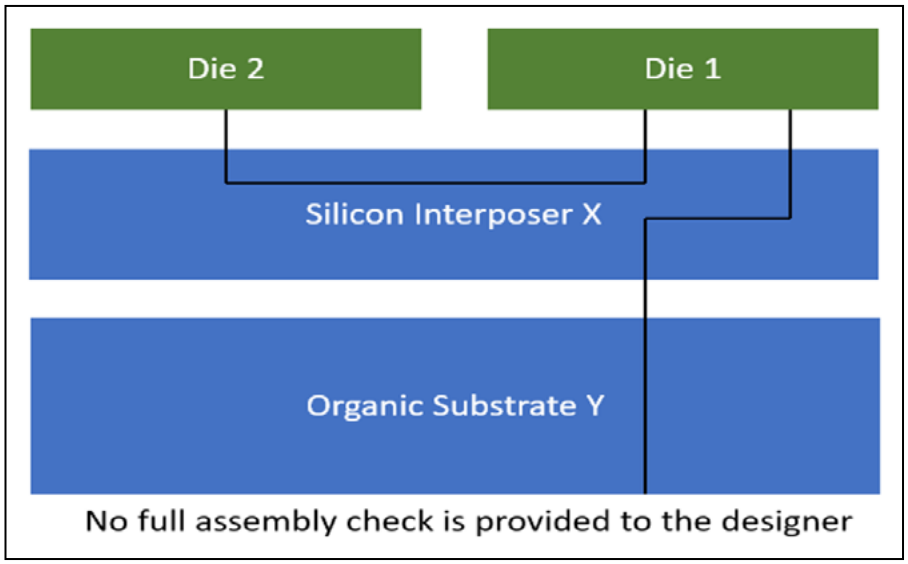


Figure 8: "Assembly" checking for the full assembly.

to organic substrate) is left to the system-level designer as illustrated in Figures 7 and 8.

System-level designers, therefore, need a reliable and mature approach of generating the setup files needed for assembly verification. Ideally, this approach should account for the full assembly and should be completely automated so that it fits right in the system-level design's environment and setup.

To resolve the above challenge, the xSI and Calibre 3DSTACK currently include a plug-in that allows the system-level designer to utilize the information stored in the xSI database to automatically generate a complete Calibre 3DSTACK runset (full assembly description and comprehensive assembly checks) along with a system source netlist. This designer-centric approach is agnostic to the different die technology nodes, the different substrates involved in the assembly, and the different manufacturing vendors (for example, a silicon interposer from foundry X and an organic substrate from OSAT Y).

Summary

For 3D-IC assemblies, the designer must ensure that the system-level netlist is golden. This can be a challenge in the case of a newly-introduced design flow. The xSI and Calibre 3DSTACK offer a fast, automated, and flexible netlist-versus-netlist approach so users can be confident that they built the system-level connectivity correctly.

Another challenge in the 3D-IC design flow are connectivity exceptions, in which different design versions can include intended shorts that need to be waived for more user-friendly, system-level LVS debugging. Again, we have shown how our solutions allow the support of known shorts and the tracking of connectivity between the IC design and package design domains can be handled seamlessly.

Finally, the assembly physical verification of 3D-ICs requires an accurate assembly description that is "design dependent." We overcame this challenge by providing a plug in on top of xSI that supports the full automation of the Calibre 3DSTACK deck.

Biography

Tarek Ramadan is the Applications Engineering Manager for the 3D-IC Technical Solutions Sales (TSS) organization at Siemens EDA, Cairo, Egypt. He drives EDA solutions for 2.5D-IC, 3D-IC, and wafer-level packaging applications. Prior to that, Tarek was a Technical Product Manager in the Siemens Calibre design solutions organization. Ramadan holds BS and MS degrees in Electrical Engineering from Ain Shams U., Cairo, Egypt.

Exploring bond strength for an advanced chiplet with hybrid bonding

By Junya Fuse, Tomoya Iwata, Yuki Yoshihara, Marie Sano, Fumihiro Inoue [YOKOHAMA National University]

Because of the high demand for high-density vertical interconnections using advanced chiplet technology, hybrid bonding is now being comprehensively investigated as an alternative to micro-bumps. However, there are many obstacles and challenges in the R&D phase that remain to be overcome with respect to die-level hybrid bonding. For example, a bonding strength measurement method for die-level hybrid bonding has not yet been well established. Therefore, the bonding strength measurement method at the die level was studied using the Cube Corner indentation method. The validity of the nanoindentation test was demonstrated by comparing the bonding strength calculated by using double-cantilever-beam and nanoindentation tests. Then, by performing nanoindentation tests on the die-to-wafer (D2W) sample, we compared it to the wafer-to-wafer (W2W) sample, which has a similar bonding film. The comparison showed the difference in bonding strength due to the different bonding films.

Introduction

Pitch scaling of vertical interconnections is a crucial development aspect for advanced chiplet integration. So far, the use of micro-bumps has fulfilled the demand for vertical interconnections in 3D integration, however, it is facing the limitation of scaling below 10µm pitch due to reliability concerns and the alignment limitation with thermal compression bonding. Cu-Cu hybrid bonding can be an alternative to accomplishing vertical interconnections, particularly where the technology is getting more mature for W2W integration [1-2]. However, several challenges remain in D2W hybrid bonding in order for it to be implemented in a wide variety of chiplet integration schemes. The focus of a novel integration is to develop collective and reconstructed D2W hybrid bonding that will enable mass production [3-5]. In addition, there are many fundamental issues to be overcome for D2W integration. For example, the bonding strength measurement for D2W bonding is not well defined.

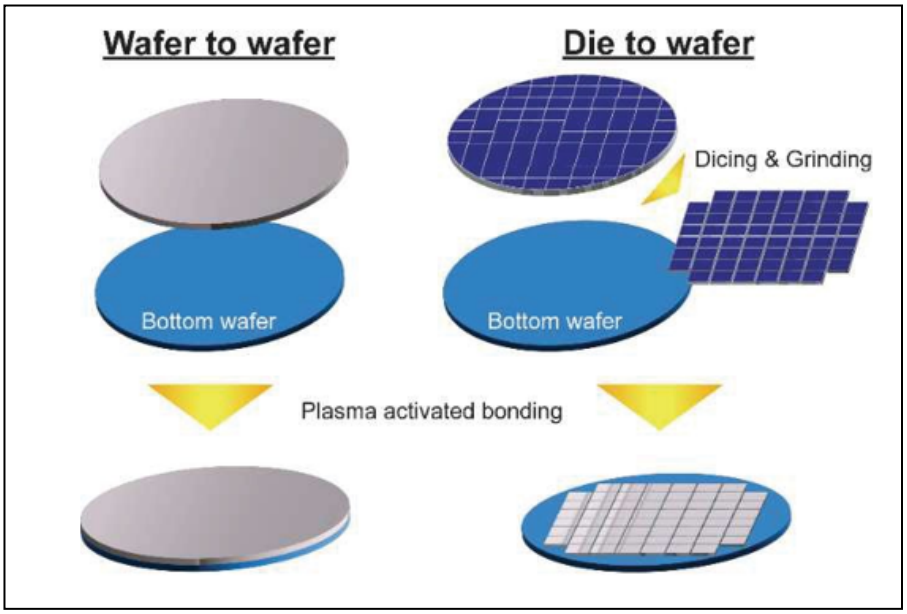


Figure 1: Schematic illustrations of W2W vs. D2W process flow.

Figure 1 shows the simplified schematic image of the process flow of W2W and D2W hybrid bonding. For wafer-level direct (fusion) bonding, the double-cantilever-beam (DCB) method is used to determine the bond strength (Figure 2a-b) [7-11]. However, it is not applicable to die-level bonding because there is no initiation at the bonding interface to insert the blade (Figure 2c). Furthermore, because the measurement by DCB is only available at the wafer edge, the inner adhesion strength is not precisely measured. Although several measurement methods are proposed [12], the comparison between precisely measured W2W and D2W examples has not yet been made. Therefore, we investigated an alternative approach to evaluate bond strength for actual die-level direct-bonded samples in this study. We used nanoindentation (NI) as the bond strength measurement method. Furthermore, exactly the same sample (wafer) is used to compare the bonding method between DCB and NI. Finally, we compared the adhesion strength of W2W and D2W bonded samples by using the NI method.

Experimental methods

The sections below discuss sample preparation and the bond strength measurement methods that were used.

Sample preparation. All the samples were fabricated from 300mm Si wafers (775µm) as the original substrate. A thermal silicon dioxide (Th-SiO_2) film with a thickness of 100nm was used as the reference bonding interface. The reference W2W bonding was performed using a full-auto wafer bonding system. In addition, a low-temperature silicon dioxide (LT-SiO_2) film 100nm thick (deposited using plasma-enhanced chemical

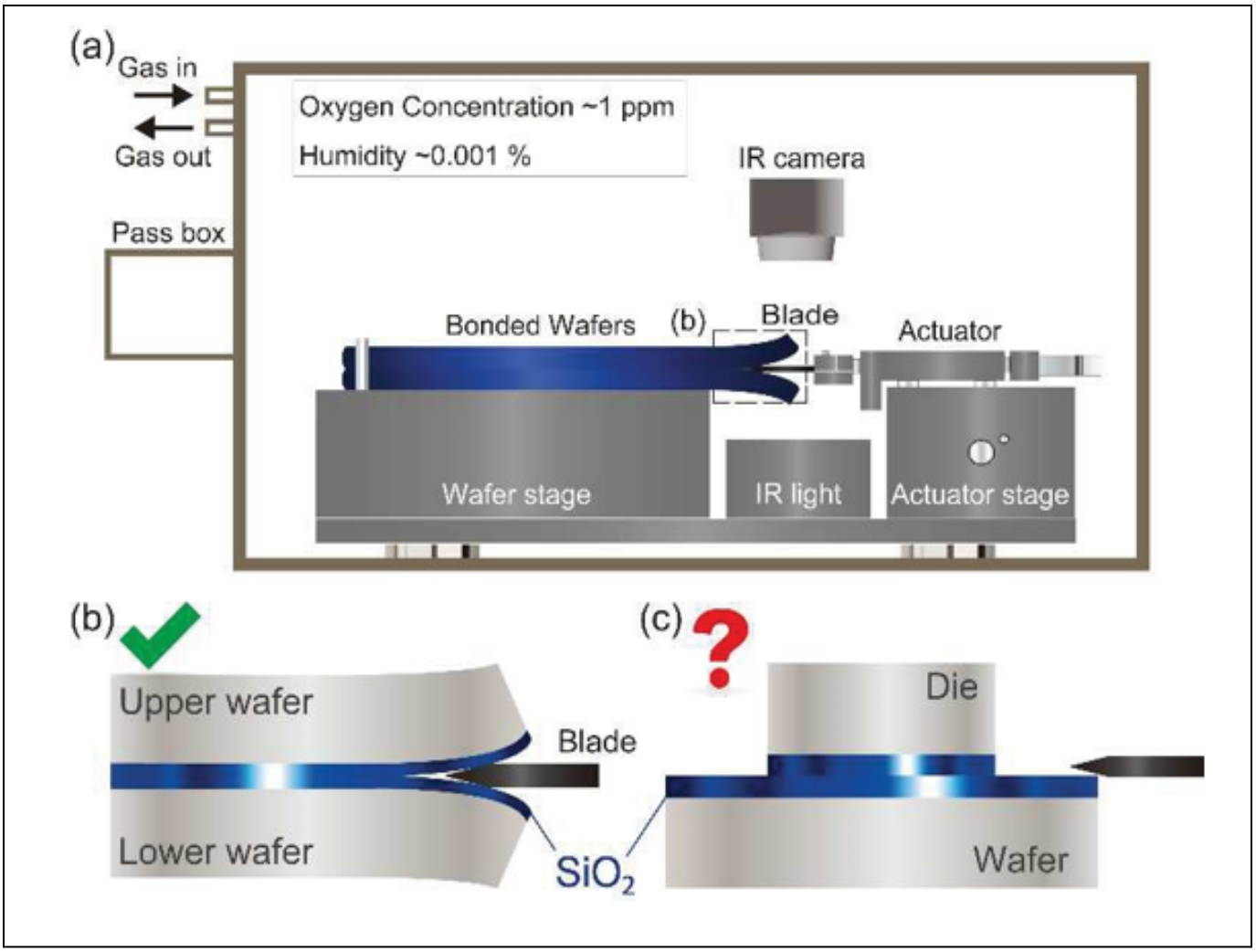


Figure 2: Schematic illustrations of: a) an in-house designed DCB tool; the W2W strength measurement with b) DCB, and c) for D2W.

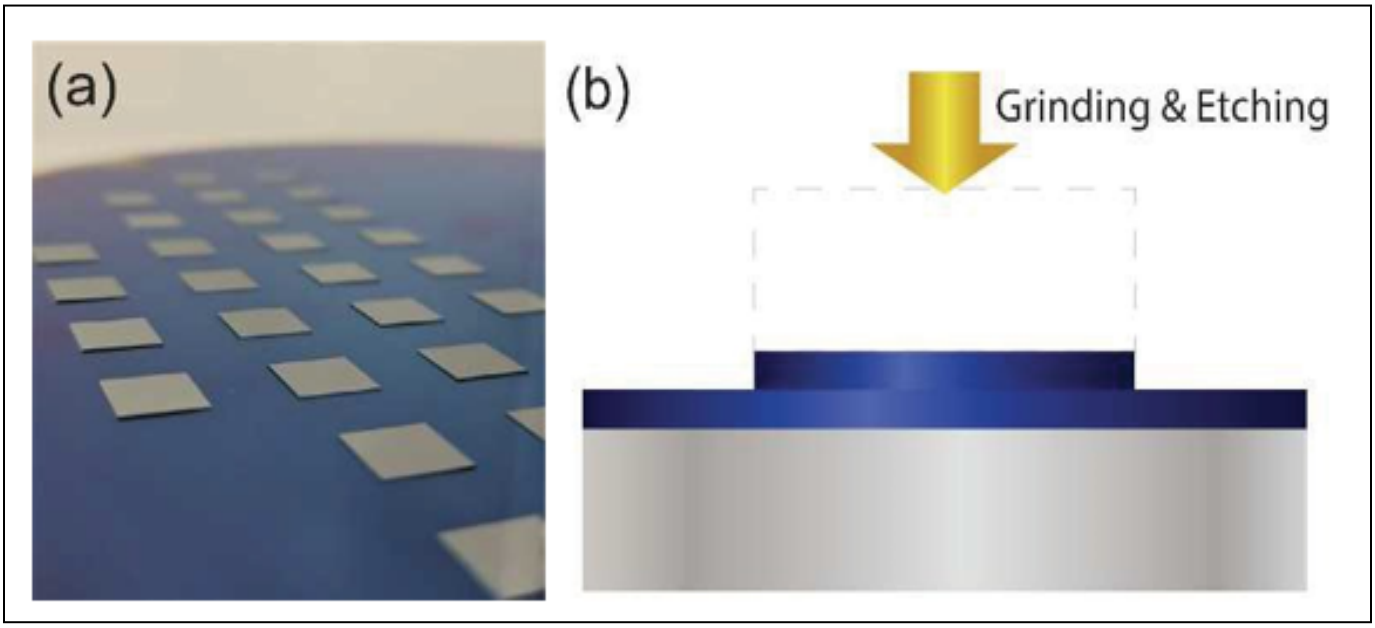


Figure 3: a) Photo of an actual D2W bonded sample; and b) A schematic illustration of sample preparation for the NI test.

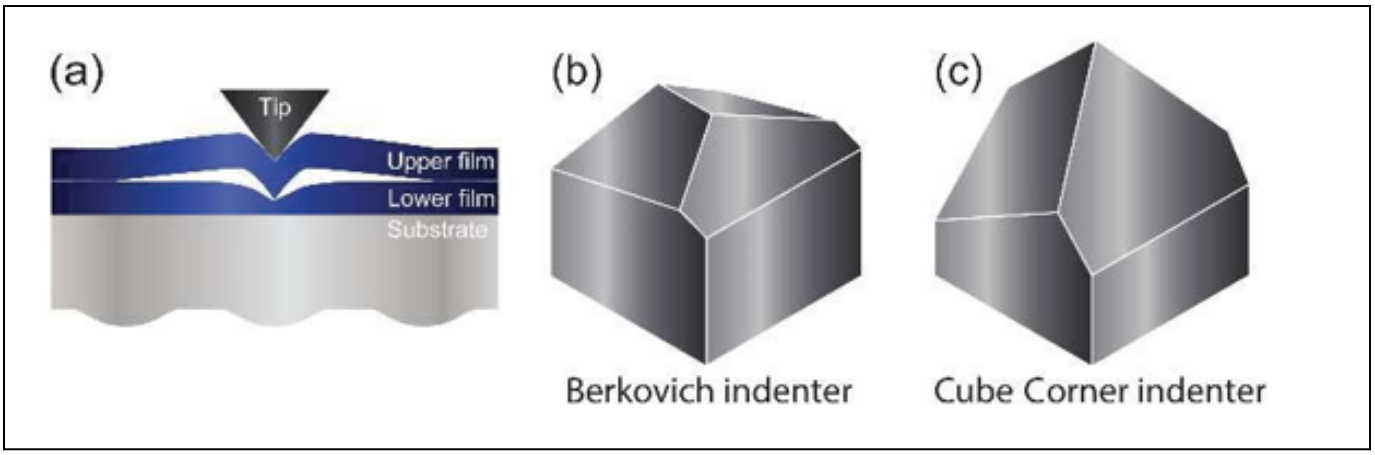


Figure 4: Schematic illustrations of: a) the bond strength measurement and delamination caused by nanoindentation; b) a Berkovich tip; and c) a Cube Corner tip.

vapor deposition [PECVD]) was used as the bonding interface. After planarization of the LT-SiO_2, the top wafers were thinned down and singulated into a die. The die and wafer surface are plasma activated by a standalone plasma activation system that is compatible with the wafer and dicing frame. The die stacking was then carried out using a full-auto die bonder compatible with D2W hybrid bonding. After bonding, the wafers were annealed in an N_2 atmosphere at various temperatures for one hour. **Figure 3a** shows the picture of the test sample right after bonding. Both W2W and D2W samples were observed for voids at the interfaces using a scanning acoustic microscope (SAM) after bonding and post-annealing.

Bond strength measurement methods. The bond strength of W2W bonded pairs is measured using the DCB method in a glovebox (GB) with an anhydrous environment and at ambient temperature as a reference. More details of the measurement condition can be found in reference [13].

For the NI test, the bonding interface needs to be revealed. For this purpose, the bulk silicon of the top dies (or wafers) was removed by back-grinding and wet etching (**Figure 3b**). Instead of using time-consuming techniques such as GHz SAM or transmission electron microscopy (TEM), we employed scanning probe microscopy (SPM) images from the top surface to calculate the delamination. The obtained SPM image was analyzed, and the delamination area was defined from the image.

The schematic drawing of NI equipment is shown in **Figure 4a**. The Cube Corner indenter tip is used to measure the interfacial bond strength. Generally, a Berkovich indenter is used to measure mechanical characteristics with NI. Nevertheless, Cube Corner indenter tips were used in this study because the bonding dielectrics are relatively thin. The geometry of these indenters is described by Morris, et al. [14]. The ideal Cube Corner has an axisymmetric equivalent angle of 35.26° and that of the Berkovich indenter's is 65.27°. The Cube Corner tip also has a sharper shape than the Berkovich indenter (**Figure 4b-c**).

It has been reported that the threshold indentation failure load (i.e., the load at which the material does not crack) for bulk materials decreases significantly as the sharpness of the Cube Corner's tip increases, and the threshold indentation load for radical failure decreases significantly with hardness [15].

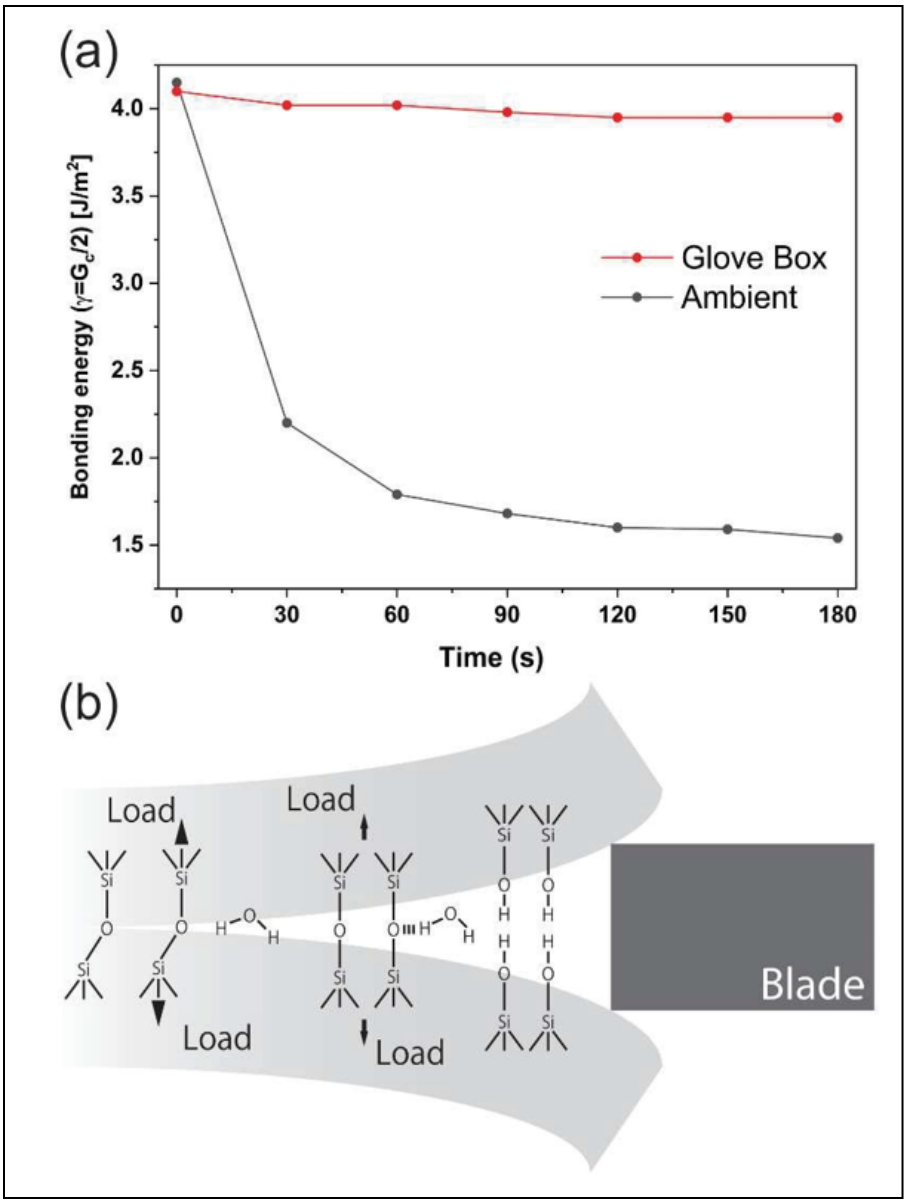


Figure 5: a) The bond strength of a W2W bonded pair measured by DCB; and b) A schematic image of water stress corrosion.

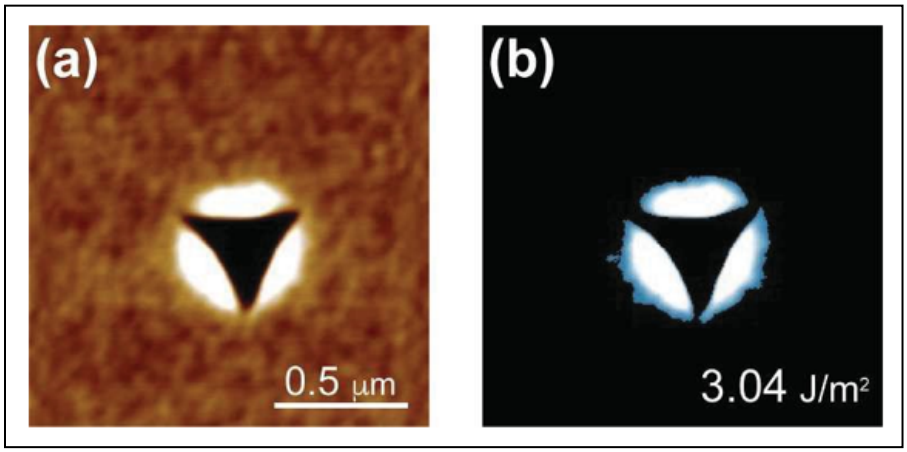


Figure 6: a) An SPM image of indentation; and b) An analysis image from the interface (thermal SiO_2).

Results and discussion

The sections below discuss the precise bond energy measurement done by DCB and by NI.

Precise bond energy measurement by DCB. We investigated the delamination length using DCB by taking the measurement at ambient temperature and in an anhydrous atmosphere (i.e., in the glovebox). The bonding energy that was calculated varied depending on the measurement time. The time transient with respect to the bonding strength after blade insertion is shown in **Figure 5a**. There was a decrease in apparent bonding energy (e.g., propagation of delamination) due to water stress corrosion under ambient temperature. The cause was the hydrolysis reaction of siloxane bonds at the bond interface as shown in **Figure 5b** [10-12]. The measurement result can be considered to be robust and repeatable from our previous report [13]. Therefore, a comparison is made for the DCB result and the NI method.

Bond strength measurement by NI (W2W bonding). The bond strength was measured using NI on exactly the same sample but at a different location from the wafer pair measured using DCB. **Figure 6a** shows the SPM image of the indentation holes formed during the indentation. A nano-sized delamination area observed in **Figure 6a** was converted into a visually recognizable area for further analysis and estimation of bond strength (**Figure 6b**). The calculation of bond strength by NI has some variation compared to the DCB method.

The bonding strength γ (= $G_C/2$) (J/m^2) is expressed as follows [16]:

$$G_C/2 = \gamma = (E_f h V_O^2) / 4Vc^2 \ (J/m^2) \quad \textbf{Eq. 1}$$

where E_f (GPa) is Young's modulus of the bonded film, h (m) is the thickness of the SiO_2 film on one side, V_O (m) is the contact area of the Cube Corner, and Vc (m) is the delamination area taken from **Figure 6b**. In the DCB test, the calculations could be performed without considering the thin film because the test was performed on a large scale. Therefore, the Young's modulus used in the calculations was that of silicon, and it was different for each crystal orientation. On the other hand, Young's modulus used in the estimation is that of the thin film being delaminated by the NI test (i.e., the amorphous Th-SiO_2 film). In this study, 70GPa was used as Young's modulus of the SiO_2 film [17].

Figure 7 shows the bonding energy measured by NI at different locations in a wafer (two edges and the center) for two sets of bonded wafers, whose bonding energy was measured by DCB. Edge 1 is the edge of the silicon wafer in the <100> crystal orientation, and Edge 2 is the edge in the <110> crystal orientation.

The first sample set has weaker plasma before bonding (bond strength 3.08J/m^2) and the second sample has stronger plasma before bonding (bond strength 4.1J/m^2). In total, 30 points were measured per each location. Then, some obvious outliers and/or failed measurements were removed for analysis We did this because we considered it necessary to allow for a certain range of measurement error—considering that the NI test is a destructive test—to obtain some certain range of measurement.

For the case of low plasma, the bonding energy in the DCB method was 3.08J/m^2. For the case using measurement by NI, the average bonding energy of Edge 1 was 3.23J/m^2, the average bonding energy of Edge 2 was 3.51J/m^2, and the average bonding energy of the Center was 3.14J/m^2. The bonding energy in the DCB test fell within the quartile range for all samples. Furthermore, the average bonding energy at Edge 1 and the Center differed from the DCB test by 4.8%, and 2.0%, respectively.

The bonding energy in the DCB method was 4.1J/m^2 for the case using high plasma before bonding. When the measurement was done by NI, the average bonding energy of Edge 1 was 3.96J/m^2, the average bonding energy of Edge 2 was 4.3J/m^2, and the average bonding energy of the Center was 3.98J/m^2. The bonding energy in the DCB test also fell within the quartile range for all samples. Furthermore, the average bonding energy at Edge 1, Edge 2, and the Center differed from the DCB test by 3.3%, 4.9%, and 2.0%, respectively.

As expected, the measurement error of the NI test is large, therefore it is difficult to measure a valid bonding strength by measuring only a single point, so it is necessary to statistically estimate the bonding strength from a large number of measurement points. This method, however, can be considered as effective as a wafer bonding strength measurement method because the statistical measurement results lead to a similar range as the bonding energy obtained by the DCB test.

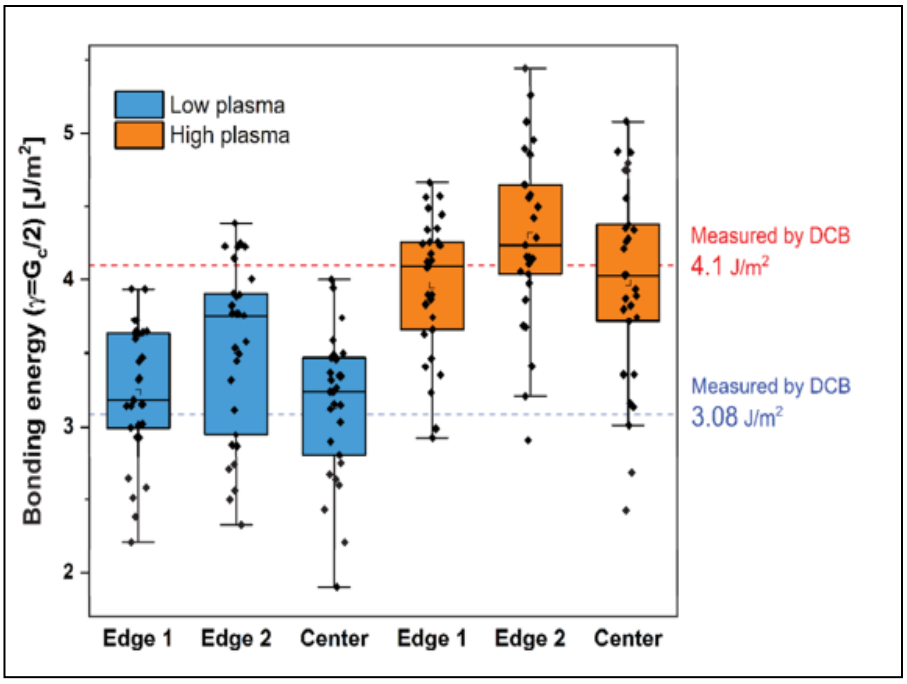


Figure 7: Bonding energy for each wafer using NI.

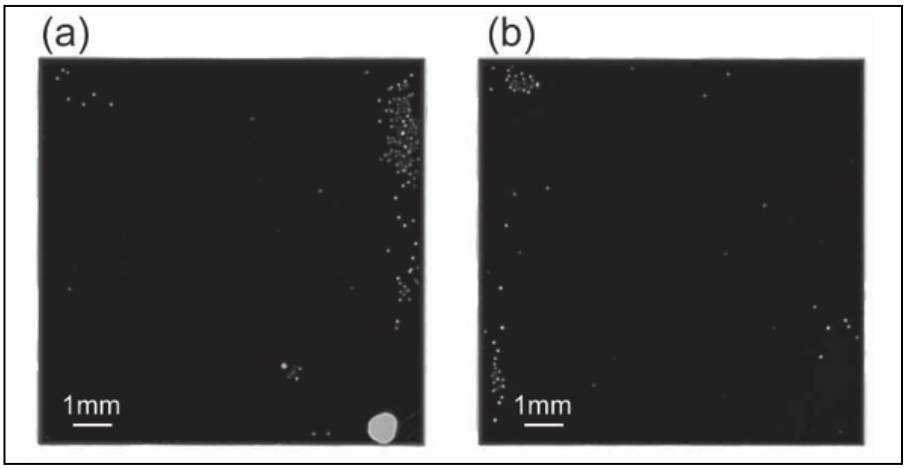


Figure 8: SAM image of indentation: a) with a thermal SiO_2 film; and b) with an LT-SiO_2 film.

It should be noted that in this study, the bonding energy did not change between the edge and the center in W2W. This indicates that the plasma radiation and annealing during W2W bonding are uniform. Additionally, it is worth noting that the bond strength is equivalent for the case with the obtained value in "GB." This indicates that the bonding interface is not exposed to air by the indentation. Therefore, the impact of water stress corrosion in this case can be negligible, which enables precise bond strength measurement by NI.

Bond strength measurement by NI (D2W bonding). Based on the result obtained in W2W, we utilized the bond strength measurement method for D2W bonded samples. Figure 8 shows SAM images with D2W. The interface dielectric material for bonding is thermal SiO_2 film (thickness=100nm) and LT-SiO_2 film (thickness=100nm).

In Figure 8, one can see the SAM image of D2W bonded samples after annealing at 250°C. There are some minor voids at the die edge. The root cause of the void formation is currently being studied. However, we are quite sure that it is not due to the dicing or bonding processes because the bonding voids obviously appeared after annealing. Therefore, it might be due to a non-optimized film condition that can outgas water after annealing. However, we used this sample for the measurement because the

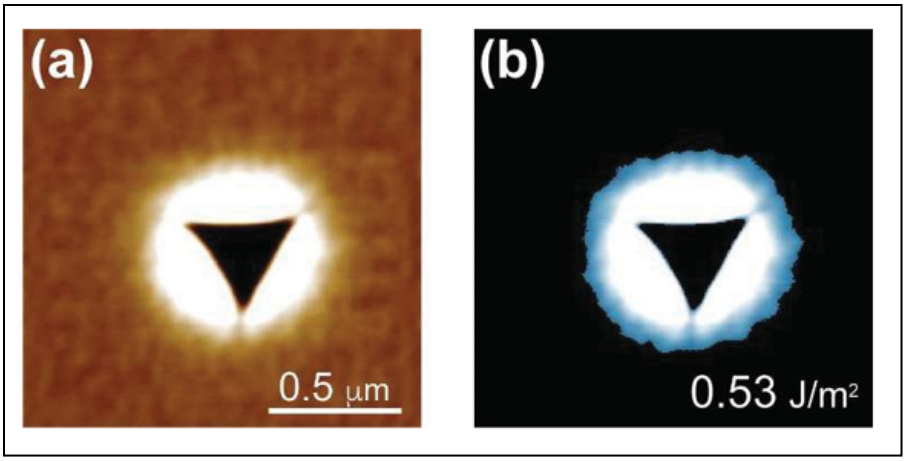


Figure 9: a) An SPM image of indentation; and b) An analysis image from the interface (thermal SiO_2).

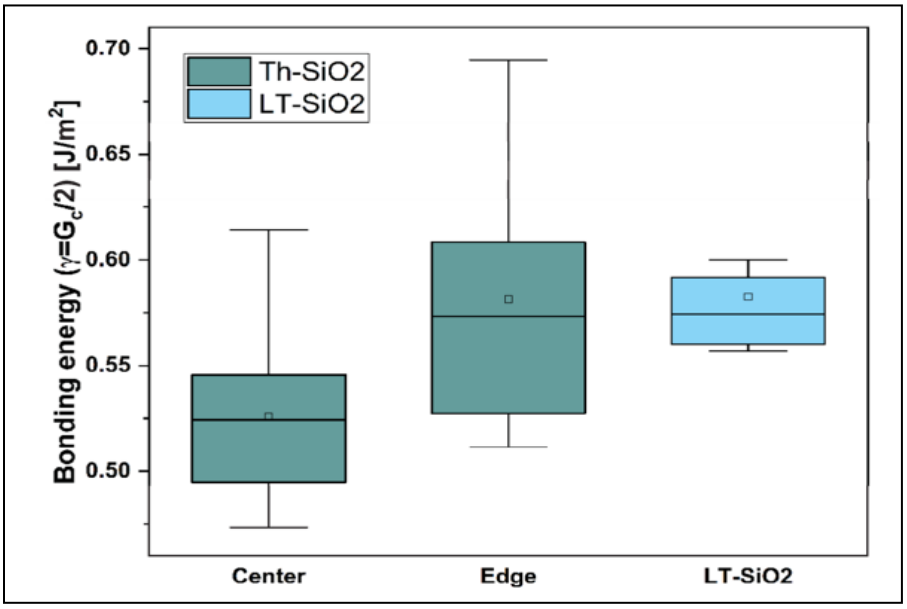


Figure 10: The bonding energy of D2W samples for different bonding dielectric interfaces and locations.

existence of voids and the impact of bonding voids on bond strength is another area of interest for our investigations.

For the case of D2W bond strength measurement, three different samples were prepared. The bonding is homogeneous (i.e., the same dielectric layer is used for both top die and the bottom wafer). **Figure 9** shows the SPM image of D2W bonded samples with Th-SiO_2 as the bonding dielectric layer. The delamination area of D2W is clearly larger than that of W2W, even though the interface of W2W and D2W is the same thermal SiO_2 film.

In the case of Th-SiO_2, the average bonding energy of the Center (in this case, the center of the die) was 0.53J/m^2, the average bonding energy of the Edge was 0.58J/m^2 (**Figure 10**). Comparing the Center and Edge quartiles, there is a variation within the die as they do not completely overlap. In the case of LT-SiO_2, the average bonding energy was 0.58J/m^2. This indicated that the process for D2W direct bonding that we performed is not well optimized compared to the one used for W2W.

Continuing the above discussion for our case in particular, the plasma activation and bonding had some time interval. As we observed for the case of W2W bonding, the plasma condition makes a large difference with respect to the bond strength. The activated site, e.g., the OH group, or even a dangling bond, might have been deactivated due to the queue time for the D2W samples. The further optimization of the D2W direct bonding process is being evaluated based on the obtained bond strength values.

Summary

The investigation of an alternative approach to evaluate bond strength for actual die-level direct bonded samples was executed. The estimation of bond strength at the die level is crucial in assessing the reliability of advanced chiplets processed using D2W hybrid bonding. The use of the nanoindentation test made it possible to measure the in-plane uniformity of W2W bonding. Furthermore, it is now possible to measure the bonding strength of D2W samples, which could not be measured in the DCB test. Although the accuracy of the measurement is still a challenge, the measurement can accelerate the development of D2W hybrid bonding with a precise bond strength value.

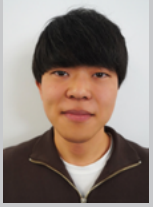
Acknowledgments

This article was presented at the 25th Electronics Packaging Technology Conference (EPTC 2023) and was edited for publication in *Chip Scale Review*. This work was supported by the "Intensive Support for Young Promising Researchers" project from the New Energy and Industrial Technology Development Organization (NEDO) and Strategic Research Seed Cultivation Project from KISTEC.

References

1. F. Inoue, et al., "Influence of composition of SiCN as interfacial layer on plasma activated direct bonding," ECS J. of Solid State Science and Tech., vol. 8, pp. 346-350, 2019.
2. J. H. Lau, "Recent advances and trends in advanced packaging," IEEE Trans. on Components, Packaging and Manufacturing Tech., vol. 12, no. 2, pp. 228-252, 2022.
3. A. Elsherbini, et al., "Enabling next generation 3D heterogeneous integration architectures on Intel process," 2022 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, pp. 27.3.1-27.3.4.
4. T. Uhrmann, "Accelerating 3D and heterogeneous integration with high-volume D2W hybrid bonding," *Chip Scale Review* Mar/Apr 2023.
5. F. Inoue, et al., "Inorganic temporary direct bonding for collective die to wafer hybrid bonding," 2023 IEEE 73rd Electronic Components and Tech. Conf. (ECTC), Orlando, FL, USA, pp. 556-563.
6. I. Nathan, N. Nejadsadeghi, N. Kohama, H. Tanoue, K. Motoda, "Distortion simulation for direct wafer-to-wafer bonding process," 2023 IEEE 73rd ECTC, Orlando, FL, USA, pp. 694-698.

7. W. P. Maszara, G. Goetz, A. Caviglia, J. B. McKitterick, "Bonding of silicon wafers for silicon-on-insulator," J. Appl. Phys. vol. 64, pp. 4943-4950, 1988.
8. F. Fournel, L. Continni, C. Morales, J. Da Fonseca, H. Moriceau, F. Rieutord, et al., "Measurement of bonding energy in an anhydrous nitrogen atmosphere and its application to silicon direct bonding technology," J. Appl. Phys. vol. 111, p. 104907, 2012.
9. K. Takeuchi, T. Suga, "Quantification of wafer bond strength under controlled atmospheres," Jpn. J. Appl. Phys. vol. 61, SF1010, 2022.
10. T. A. Michalske, B. C. Bunker, "Slow fracture model based on strained silicate structures," J. Appl. Phys. vol. 56, pp. 2686-2693, 1984.
11. V. Masteika, J. Kowal, N. St. J. Braithwaite, T. Rogers, "The effect of atmospheric moisture on crack propagation in the interface between directly bonded silicon wafers," Microsystem Technologies, vol. 19, pp. 705-712, 2013.
12. K. Baek, J. Kim, M. -s. Han, K. Lim, D. M. Rhee, "Chip level evaluation of wafer-to-wafer direct bonding strength with bending test," IEEE 73rd ECTC, Orlando, FL, USA, 2023, pp. 310-317.
13. J. Fuse, T. Iwata, S. Ebiko, F. Inoue, "Robust measurement of bonding strength for wafer-to-wafer 3D integration," International Conf. on Electronics Packaging (ICEP), Kumamoto, Japan, 2023, pp. 105-106.
14. D. J. Morris, S. B. Myers, R. F. Cook, "Sharp probes of varying acuity: Instrumented indentation and fracture behavior," J. Mater. Res. 19, 165 (2004).
15. D. Morris, R. Cook, "Indentation fracture of low-dielectric constant films: Part I. Experiments and observations," J. of Materials Research, 23(9), 2429- 2442 (2008).
16. A. A. Volinsky, N. R. Moody, W. W. Gerberich, "Interfacial toughness measurements for thin films on substrates," Acta Materialia 50 (2002): 441-466.
17. R. J. Jaccodine, W. A. Schlegel, "Measurement of strains at Si-SiO_2 Interface," J. of Applied Physics 1 May 1966; 37 (6): 2429-2434.



Biographies

Junya Fuse received a BS degree in Mechanical Engineering from YOKOHAMA National University, Japan. He is currently pursuing an MS degree with a focus on chiplets at the same university.

Contact author: Fumihiro Inoue is an Associate Professor at YOKOHAMA National University, Japan. He worked at imec, specializing in unit processes for 3D integration until 2021. In recognition of his outstanding contributions to the field, he was awarded the IEEE EPS Outstanding Young Engineer Award for 2022. Email inoue-fumihiro-ty@ynu.ac.jp

Scaling up GaN- and InP-based technologies for 5G and 6G wireless communication

By Nadine Collaert [imec]

Year after year, more and more data are being transmitted wirelessly by an ever-growing group of users. To keep up with this trend and make data transfer faster and more efficiently, the fifth generation (5G) of mobile communication is being rolled out, and the industry is already looking at what lies beyond. While 5G enables peak data rates of 10Gbit/s, 6G is projected to operate at 100Gbit/s from 2030 onwards. In addition to coping with more data and connections, researchers are investigating how the next generation of wireless communication can support new use cases such as autonomous driving and holographic presence, among others. To enable these exceedingly high data rates, the telecom industry has been pushing up the frequencies of wireless signals. While 5G initially uses sub-6GHz frequency bands, products targeting 28/39GHz are already showcased. Additionally, there is a growing interest in using FR3 (6-20GHz) frequency bands for 5G networks because of their ability to balance coverage and capacity. For 6G, frequencies above 100GHz are being discussed.

Moving toward higher frequencies has several advantages: new frequency bands can be used—thereby solving the spectrum scarcity issue within existing bands. Also, the higher the operating frequency, the easier it is to obtain wider bandwidths. Frequencies above 100GHz and bandwidths up to 30GHz allow telecom operators, in principle, to use lower-order modulation schemes within the wireless data links, which reduces power consumption. Higher frequencies are also associated with smaller wavelengths (λ). As the antenna array size scales with λ^2, antenna arrays can be packed more densely. This contributes to better beamforming, a technique that ensures that a larger fraction of transmitted energy reaches the intended receiver.

The advent of higher frequencies, however, comes at a price. Today, complementary metal-oxide-semiconductor (CMOS) is the preferred technology for building the critical components of transmitters and receivers. These include the power amplifiers within the front-end modules that send the radio frequency signals to and from the antennas. The higher the operating frequency, the more the CMOS-based power amplifiers struggle to deliver the required output power with sufficiently high efficiency. And that's where technologies such as GaN and InP come into play. Because of their outstanding material properties, these III/V semiconductors are more likely to provide the required output power and efficiency at high operating frequencies. GaN, for example, has a high current density, high electron mobility, and large breakdown voltage. The high power density also allows for a small form factor and, therefore, a reduction in overall system size at the same performance.

GaN and InP outclass CMOS at higher operating frequencies

In a modeling exercise, researchers at imec compared the performance of three different power amplifier implementations at 140GHz operating frequency: a full CMOS implementation, a CMOS beamformer with SiGe heterojunction bipolar transistor (HBT), and an InP HBT (**Figure 1**) [1]. InP was the clear winner in terms

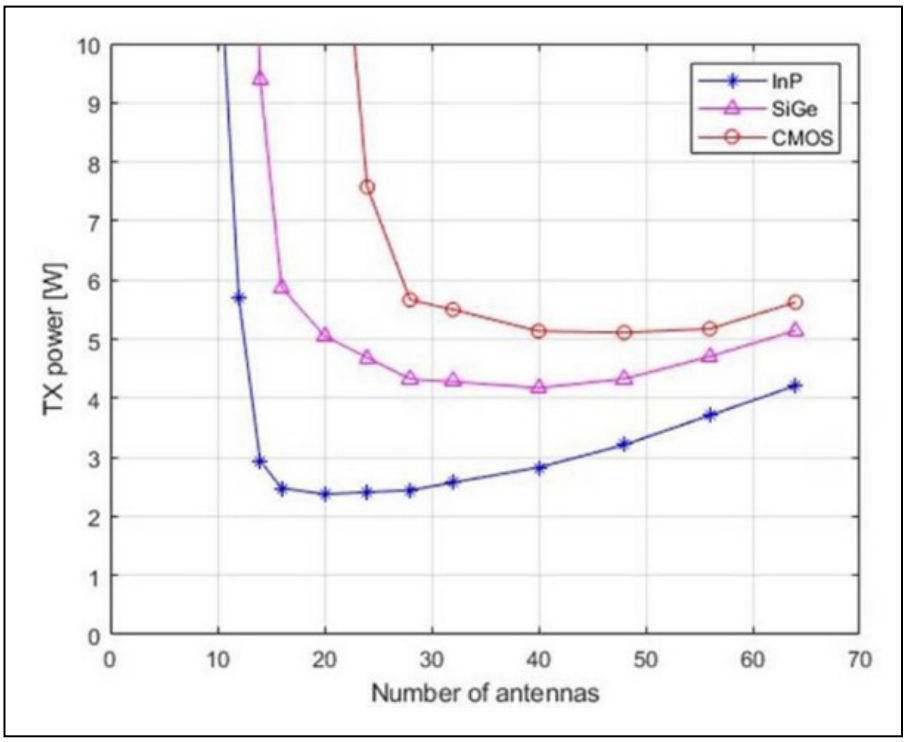


Figure 1: Comparing the power consumption of CMOS, SiGe, and InP devices in transmitter architectures as a function of the number of antennas [1].

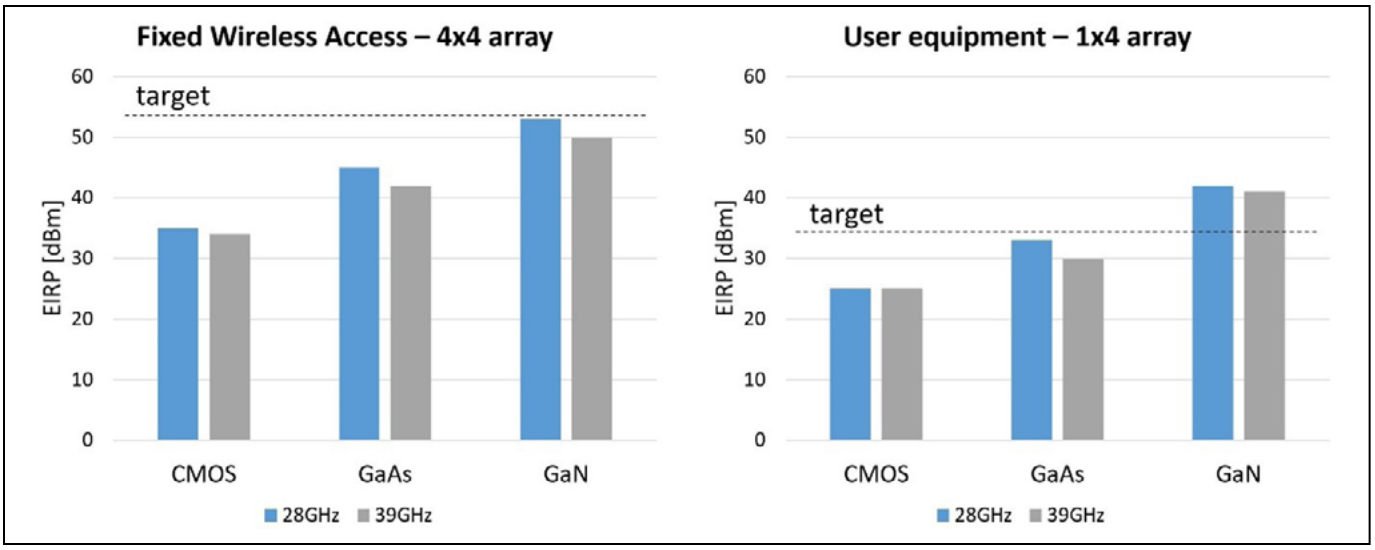


Figure 2: Output power for 28GHz and 39GHz operating frequencies in: a) (left) FWA, and b) (right) user equipment: a comparison of three different technologies [1].

of output power (over 20dBm) and energy efficiency (20-30%). Modeling results also indicate that for InP, the optimal point for energy efficiency is obtained with a relatively low number of antennas. This is especially interesting for footprint-restricted use cases like user equipment (e.g., mobile devices).

At lower mm-wave frequencies, however, GaN shows excellent performance. For both 28 and 39GHz, high-electron mobility transistors (HEMTs) made of GaN-on-silicon-carbide (GaN-on-SiC) are observed to outclass CMOS-based devices and GaAs HEMTs in terms of output power and energy efficiency. Two different use cases were considered, i.e., fixed wireless access (FWA, with 16 antennas) and user equipment (with four antennas) (**Figure 2**).

Opportunities and challenges for upscaling

If we consider cost and ease of integration, GaN and InP device technologies cannot yet fully compete with CMOS-based technologies. The III/V devices are typically made on small and costly non-Si substrates, relying on less suitable processes for high-volume manufacturing. Integrating these devices on 200 or 300mm Si wafers is an interesting approach to achieving overall optimization while maintaining superior RF performance. Not only are Si substrates cheaper, but the CMOS-compatible process also enables large-scale manufacturability.

Integrating GaN and InP on a Si platform requires a combination of new transistor and circuit design approaches, materials, and manufacturing techniques. One of the main challenges relates to the large lattice mismatch: 8% for InP, and 17% for GaN. This is known to create many defects in the layers, which ultimately degrade device performance. In addition, we will have to co-integrate the GaN-on-Si- and InP-on-Si-based components with CMOS-based components into a complete system. GaN and InP technologies will

initially be used to realize the power amplifiers within the front-end modules. Also, low-noise amplifiers and switches could potentially benefit from the unique properties of these compound semiconductors. But in the end, CMOS will still be needed for calibration, control, and beamforming.

Within its Advanced RF Program, imec, along with its industry partners, explores various approaches to integrate GaN and InP devices on large-size Si wafers, and how to enable their heterogeneous integration with CMOS components. Pros and cons are being assessed for different use cases—infrastructure (such as FWA) as well as user equipment.

Improving GaN-on-Si technology for RF performance

Depending on the starting substrate, there are several flavors of GaN technology: GaN bulk substrates, GaN-on-SiC, and GaN-on-Si. Today, GaN-on-SiC is widely explored and already used for infrastructure applications, including 5G base stations. GaN-on-SiC is more cost-efficient than bulk GaN technology, and SiC is an excellent thermal conductor that helps to dissipate the generated heat in high-power infrastructure applications. However, the cost and limited size of the substrate make it less suitable for mass production.

GaN-on-Si, on the contrary, has the potential to be upscaled to 200mm and even 300mm wafers. Thanks to years of innovation for power electronics applications, the integration of GaN on large-size Si substrates has made tremendous progress. But further improvements are needed to ready GaN-on-Si technology for optimal RF performance. The main challenges lie in achieving comparable large signal and reliability performance to GaN-on-SiC and raising the operating frequency. This requires continued innovations in the material stack design and choice of materials, reduction of the gate length of the HEMTs, suppression of parasitics, and keeping the RF dispersion as low as possible.

Imec's GaN-on-Si process flow for RF starts with the growth (by metal-organic chemical vapor deposition (MOCVD)) of an epitaxial structure on 200mm Si wafers. This structure is comprised of a proprietary GaN/AlGaN buffer structure, a GaN channel, an AlN spacer, and an AlGaN barrier. GaN HEMT devices with TiN Schottky metal gates are subsequently integrated with a (low-temperature) 3-level Cu back-end-of-line process.

Recently, competitive results have been obtained on imec's GaN-on-Si platform, bringing the output power and power added efficiency (PAE) for the first time closer to those of the GaN-on-SiC technology. The PAE is a commonly used metric to rate the efficiency of a power amplifier, which takes into account the effect of the amplifier's gain on its overall efficiency (**Figure 3**) [2-9].

Complementing the technology development with modeling activities will ultimately help achieve even better performance and reliability. For example, at IEDM 2022, imec introduced a simulation framework to better predict thermal transport in RF devices. In a case study with GaN-on-Si HEMTs, the simulations revealed peak temperature rises up to three times larger than previously predicted. Modeling work such as this provides further guidance in optimizing RF devices and their layouts early in the development phase [10].

Exploring InP-on-Si for 6G sub-THz frequencies

For the longer term, InP HBTs are being explored for 6G applications. As previously demonstrated, of all technology implementations, InP HBTs offer the best output power/efficiency trade-off at the operating frequency of 140GHz. Researchers also know how to design InP HBTs for optimal RF performance. But the fabrication usually starts from small (InP) substrate wafers (<150mm), using lab-like processes that are not CMOS-compatible. But what happens to the performance when we integrate InP on Si? Depositing InP on Si is known to introduce many defects, mainly threading dislocations and planar defects. These defects induce leakage currents that can dramatically deteriorate device performance or cause reliability issues.

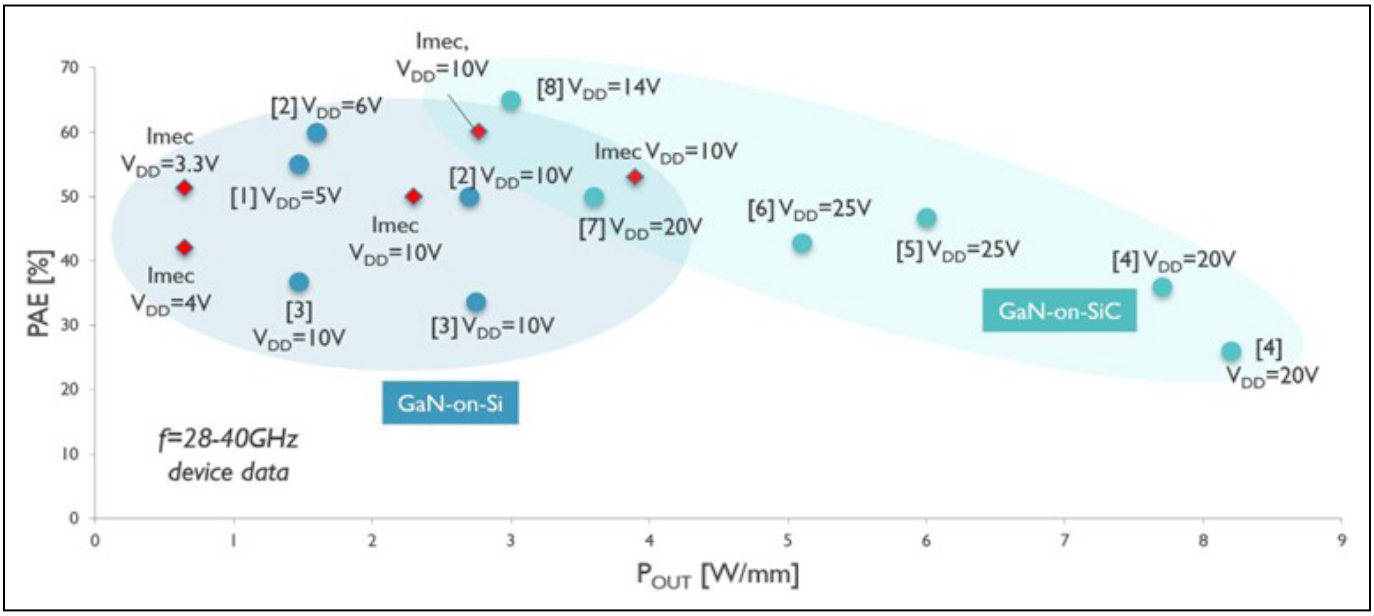


Figure 3: GaN-on-Si benchmarking data. The imec data in red is among the best reported for GaN-on-Si devices and comparable to GaN-on-SiC substrates [1-9].

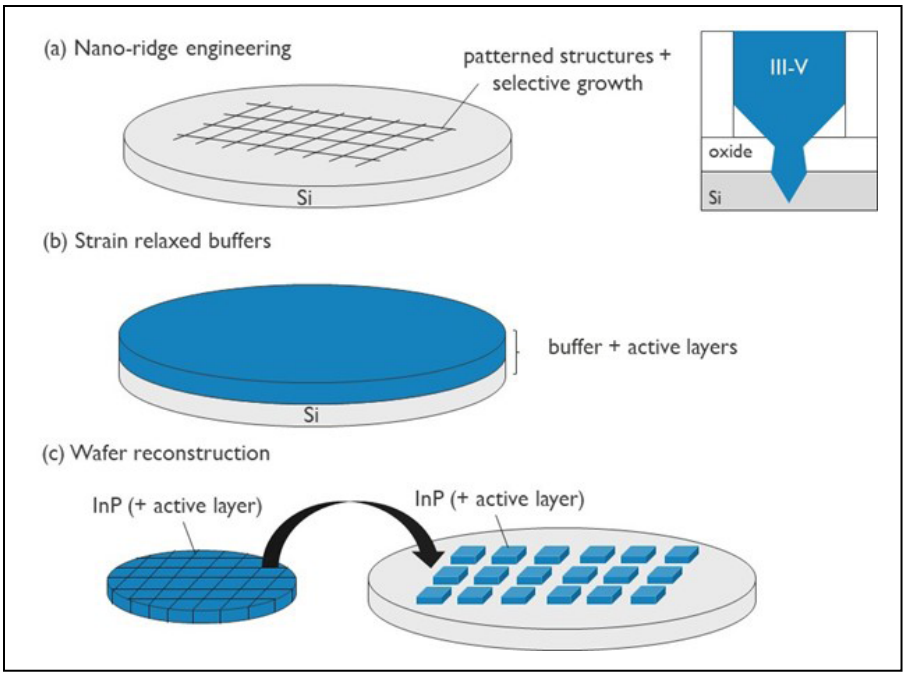


Figure 4: Schematic representation of the different InP-on-Si growth approaches: a) nano-ridge engineering; b) blanket growth with strain relaxed buffers, and c) wafer reconstruction.

Three approaches are being considered for upscaling (**Figure 4**). Two of them rely on the direct growth of InP on Si, and another on wafer reconstruction. All three approaches are envisioned to offer more cost-effective solutions than current technologies that use small InP substrates. But they all have pros and cons regarding performance, cost, and heterogeneous integration potential. Imec has taken on the role of assessing benefits and challenges for the various use cases: infrastructure as well as mobile devices.

A first approach (**Figure 4b**) to making InP-on-Si HBTs uses strain-relaxed buffer layers deposited directly on top of Si to compensate for the 8% lattice mismatch between Si and InP. Next, InP is grown directly on top of this buffer layer. The ability to use larger wafer sizes, especially in cases where some of the Si could be reused, provides a significant cost advantage. However, optimizations are needed to reduce the number of defects further.

Departing from this "blanket" growth approach, imec proposes nano-ridge engineering (NRE) as an alternative technology to cope with defects more efficiently (**Figure 4a**). NRE relies on

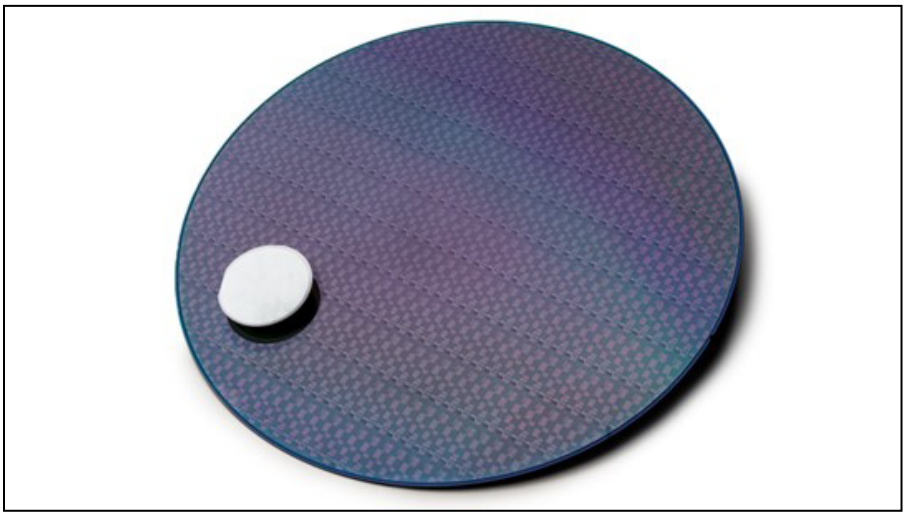


Figure 5: A two-inch InP wafer, and a 300mm Si wafer with a InP NRE HBT stack.

selectively growing the III/V material in pre-patterned trenches in Si. These high-aspect-ratio trenches are very effective in trapping the defects in the narrow bottom part and allowing the growth of high-quality, low-defectivity material out of the trench. Overgrowing the nano-ridge widens it towards the top, forming a solid base for a device stack. The first insights obtained from a GaAs/InGaP case study will guide the optimization of the target InGaAs/InP NRE HBT devices (**Figures 5-6**).

Beyond direct growth, InP can also be placed on Si using a wafer reconstruction technology (**Figure 4c**). In this case, high-quality InP substrates—with or without the active layers—are diced into tiles during wafer constitution. The tiles are subsequently attached to a Si wafer using a die-to-wafer bonding technique. The key challenges lie in the efficient transfer of the materials and the removal of the InP substrate, for which several techniques are being considered.

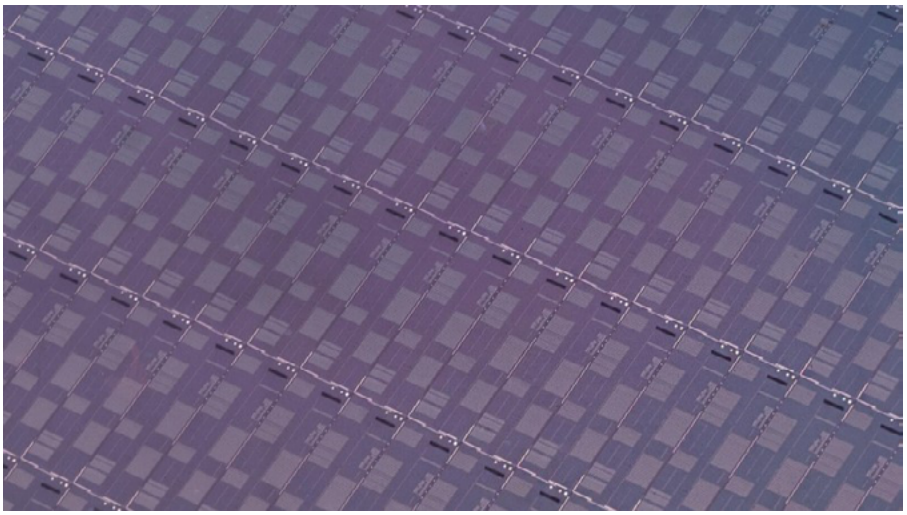


Figure 6: Zoom-in of the 300mm Si wafer showing the dies with the InP NRE HBT structures.

Towards heterogeneous integration

Ultimately, the III/V-on-Si power amplifiers must be combined with CMOS-based components that take care of, e.g., calibration and control. Imec is looking into various heterogeneous integration options and weighing their pros and cons for various use cases.

Advanced laminate substrate technology is the most common way to integrate different RF components in a system-in-package, and optimizations to make it adaptable to higher frequencies are ongoing. Additionally, imec is exploring more advanced heterogeneous integration options, including 2.5D interposer and 3D integration technologies.

Especially for frequencies above 100GHz, it is important to note that the antenna module starts to define the area available for the transceiver. Indeed, when going to higher frequencies, the wavelength decreases, and the area of the antenna array scales accordingly. Above 100GHz, the antenna size becomes smaller than the front-end module size, which hardly scales in size with increasing frequency. An interesting option for large antenna array configurations is to move the RF front-end module under the antenna array. And this is where 3D integration technologies (either die-to-wafer or wafer-to-wafer) come into play, enabling short and well-defined connections between the front-end module and the antenna modules. However, thermal management remains a great concern for 3D integration, and being able to provide effective heatsinks will be crucial. Today, at imec, we are performing a comprehensive system-technology-co-optimization (STCO) analysis to evaluate different technologies for 3D integration and to guide the technology choices from a system-level perspective.

For handheld devices, where a reduced number of antennas can relax the constraints, 2.5D interposer technology is considered an interesting approach. This

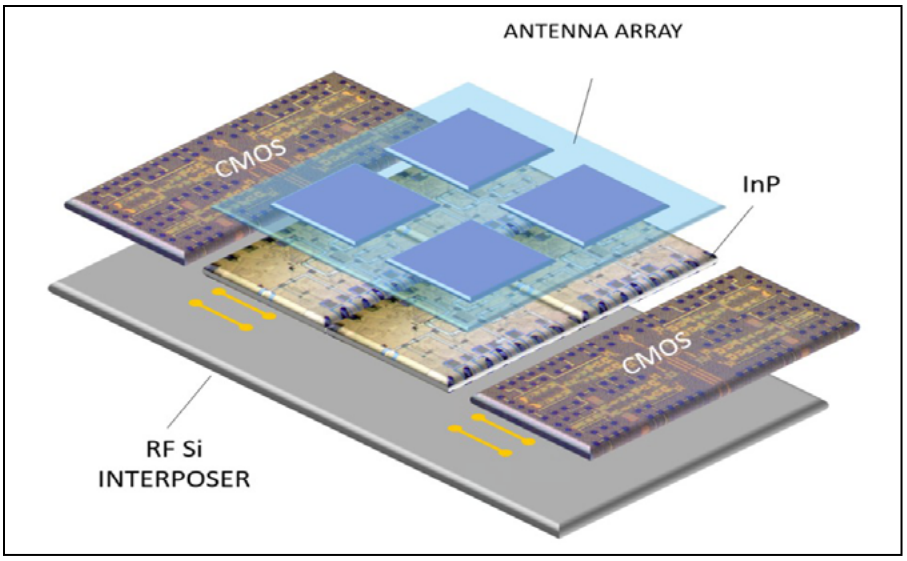


Figure 7: Schematic representation of an RF Si interposer with integrated InP and CMOS devices and antenna array in a package.

heterogeneous integration option uses a layer stack with lithography-defined connections and even through-Si vias to communicate between III/V- and CMOS-based components. In this case, the III/V devices sit next to the CMOS chip, enabling better thermal management because both chips can be in direct contact with a heat sink. Such architecture, however, only allows for 1D beam steering. We are currently evaluating hardware implementations of 2.5D interposer technology, looking into the most optimal combinations of substrates, dielectrics, and redistribution layers to minimize losses. For example, we have shown a first version of an RF-tailored Si interposer technology using a standard Si substrate, copper semi-additive interconnect, and thick spin-on low-k dielectrics that exhibit very low interconnect loss, even above 100GHz (**Figure 7**).

Summary

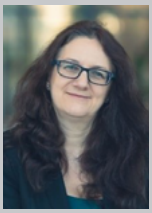
Recent upscaling and integration efforts show that GaN-on-Si and InP-on-Si can become viable technologies for next-generation high-capacity wireless communication applications.

Acknowledgment

This paper was originally published in *Compound Semiconductor* (Vol. 29/ Issue 5/pp. 46-51). It has been edited for republication in *Chip Scale Review*.

References

1. N. Collaert, et al., "III-V/ III-N technologies for next-generation high-capacity wireless communication," 2022 International Electron Devices Meeting, DOI: 10.1109/IEDM45625.2022.10019555
2. H. W. Then, et al., "Advances in research on 300mm gallium nitride-on-Si(111) NMOS transistor and silicon CMOS integration," 2020 IEEE International Electron Devices Meeting (IEDM), pp. 27.3.1-27.3.4, doi:10.1109/ IEDM13553.2020.9371977.
3. H. W. Then, et al., "Advanced scaling of enhancement mode high-K gallium nitride-on-300mm-Si(111) transistor and 3D layer transfer GaN-silicon Finfet CMOS integration," 2021 IEEE IEDM, pp. 11.1.1-11.1.4, doi:10.1109/ IEDM19574.2021.9720710.
4. W. Wang, et al., "Improvement of power performance of GaN HEMT by using quaternary InAlGaN barrier," IEEE Journal of the Electron Devices Society (Vol: 6), 2018, pp. 360 - 364.
5. Y. C. Lin, et al., "Gallium nitride (GaN) high-electron-mobility transistors with thick copper metallization featuring a power density of 8.2 W/mm for Ka-Band applications," *Micromachines* 2020, 11(2), 222; https://doi.org/10.3390/ mi11020222
6. M. Mi, et al., "Millimeter-Wave Power AlGaN/GaN HEMT Using Surface Plasma Treatment of Access Region," IEEE Trans on Electron Devices, vol. 64, no. 12, pp. 4875-4881, Dec. 2017, doi: 10.1109TED.2017.2761766.
7. Y. Zhang et al., "720-V/0.35-mΩ. cm^2 Fully Vertical GaN-on-Si Power Diodes by Selective Removal of Si Substrates and Buffer Layers," IEEE Electron Device Letters, vol. 39, Issue 5, pp. 715-718.
8. K. Harrouche, et al.,"High performance and highly robust AlN/GaN HEMTs for millimeter-wave operation," IEEE Jour. of the Electron Devices Society, 2019, 7, pp. 1145-1150; https://hal.science/hal-02357492
9. J.-S. Moon, et al., "Novel high-speed linear GaN technology with high efficiency," 2019 IEEE MTT-S International Microwave Symp. (IMS), pp. 1130-1132, doi: 10.1109/ MWSYM.2019.8700832.
10. B. Vermeersch, et al., "Thermal modeling of GaN & InP RF devices with intrinsic account for nanoscale transport effects," 2022 International Electron Devices Meeting (IEDM), DOI: 10.1109/ IEDM45625.2022.10019370



Biography

Nadine Collaert is a Fellow and a Program Director at imec, Leuven, Belgium. She's currently responsible for the advanced RF program looking at the heterogeneous integration of III-V/III-N devices with advanced CMOS to tackle the challenges of next-generation mobile communication. Previously, she was a Program Director of the logic beyond Si program and has also been involved in the theory, design, and technology of FinFET devices, emerging memories, and more. She has a PhD in Electrical Engineering from the KU Leuven, (co-) authored more than 400 publications, and holds more than ten patents in device design and process technology. Email Nadine.Collaert@imec.be

ADVERTISER INDEX

Amkor Technology	6
Brewer Science	29
CyberOptics	OBC
DL Technology	45
E-tec Interconnect	47
ECTC	52
EV Group	2
IMAPS	IBC
Intekplus Co., Ltd	21
Ironwood Electronics	35
ISC Co., Ltd.	23
Kulicke & Soffa	12
Leeno Industrial	IFC, 24
MicroCircuit Laboratories	11
Onto Innovation	49
Plasma Etch	52
Siemens EDA	9
Sonix	41
Technic	36
TestConX	50
Test Tooling Solutions Group	32
TSE Co. Ltd	4,5
WinWay Technology	27

January February 2024
For Advertising Inquiries: ads@chipscalereview.com

20TH INTERNATIONAL CONFERENCE & EXHIBITION ON DEVICE PACKAGING

FOUNTAIN HILLS, AZ • WWW.DEVICEPACKAGING.ORG

Keynotes

Pooya Tadayon - Intel
Hemanth Dhavaleswarapu - AMD
Rajendra Pendse - Meta
Sitaraman (Sita) Balasubramanian - Aptiv
Pradeep Lall - Auburn University

PLATINUM SPONSOR

ASE

GOLD SPONSORS

Amkor Technology
DECA
IBM
SUSS MicroTec

SILVER SPONSORS

adeia
cadence
KLA
NAMICS
PacTech
QP TECHNOLOGIES
SIEMENS
XYZTEC

CORPORATE SPONSORS

ASMPT enabling the digital world
DELO INDUSTRIAL ADHESIVES
evatec
IMAT
JSR
MacDermid Alpha ELECTRONICS SOLUTIONS
SETNA

ADDITIONAL SPONSORS

MRSI MYCRONIC
NxQ

WWW.DEVICEPACKAGING.ORG

March 18-21, 2024

We-Ko-Pa Resort & Conference Center Fountain Hills, Arizona

NEW! ***DPC 2024*** *will be co-located with the Workshop on Advanced Packaging for Medical Microelectronics (March 21-22)*

12 Confirmed Professional Development Courses (PDCs) on Fan-out Wafer Level Packaging, Chiplets, System-in-Package, Flip Chip, Sensor Packaging, Wire Bonding, and more.

Featuring 3 Technical Tracks and an Interactive Poster Session across 2.5 days with a record number of abstracts submitted this year!

Heterogeneous 2D & 3D Integration

Fan-Out, Wafer Level Packaging & Flip Chip

Next Gen Applications (Automotive, 5G/6G, Photonics, HiSpeed RF)

SOLD-OUT EXHIBIT HALL!

Exhibits sold-out earlier than ever before, with a waitlist of 10+ companies.

organized by the
International Microelectronics Assembly and Packaging Society (IMAPS)

Seriously Fast.

WX3000™ Metrology and Inspection Systems for Wafer-Level and Advanced Packaging


Copper Pillar


Micro Bump


Flip Chip (c4)


BGA Ball

Powered by Multi-Reflection Suppression® (MRS®) Technology

2-3X Faster with High Resolution and High Accuracy

WX3000 3D and 2D metrology and inspection system provides the ultimate combination of high speed, high resolution and high accuracy for wafer-level and advanced packaging applications to improve yields and processes.

Powered by Multi-Reflection Suppression (MRS) Sensor Technology

The 3-micron NanoResolution (X/Y resolution of 3 micron, Z resolution of 50 nanometer) MRS sensor enables metrology grade accuracy with superior 100% 3D and 2D measurement performance for features as small as 25-micron.

Ideal for wafer bump metrology, solder on wafer metrology and for measuring photoresist and plating uniformity to improve the plating process.

100% 3D and 2D metrology and inspection can be completed simultaneously at high speed (25 300mm wafers/hour and 55 200mm wafers/hour) as compared to a slow method that requires two separate scans for 2D and 3D, and only a sampling process.

nordson.com/testinspect | +1 760.918.8471

Nordson
TEST & INSPECTION