

Chip Scale Review®

ChipScaleReview.com

The Future of Semiconductor Packaging

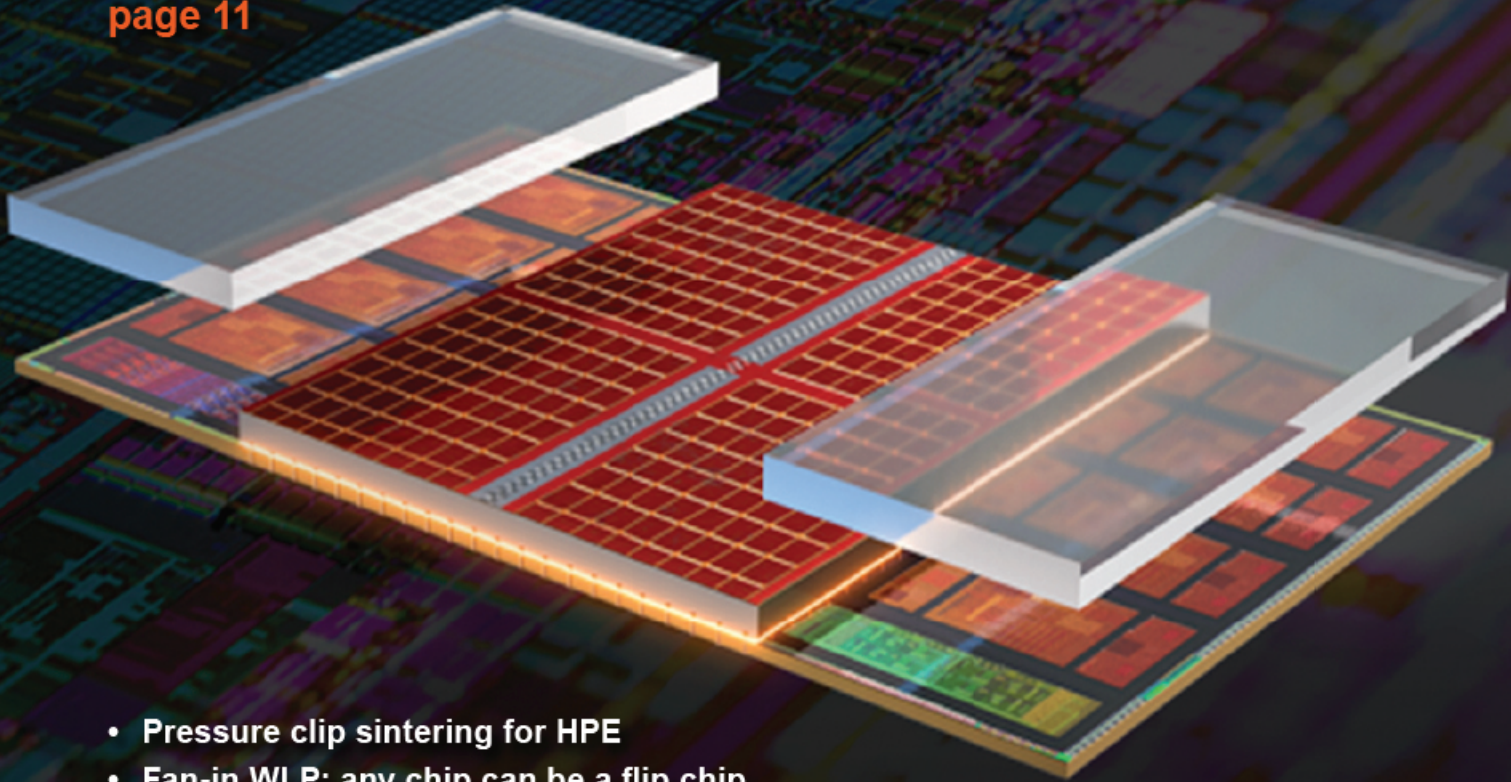
Volume 26, Number 3

May • June 2022



ADVANCED PACKAGING

Enabling Moore's Law using heterogeneous integration
page 11



- Pressure clip sintering for HPE
- Fan-in WLP: any chip can be a flip chip
- Scalable silicon photonics using optical bump NIL
- Semiconductor test: staying ahead of nanodevices
- Status and outlook for FO wafer/panel-level packaging

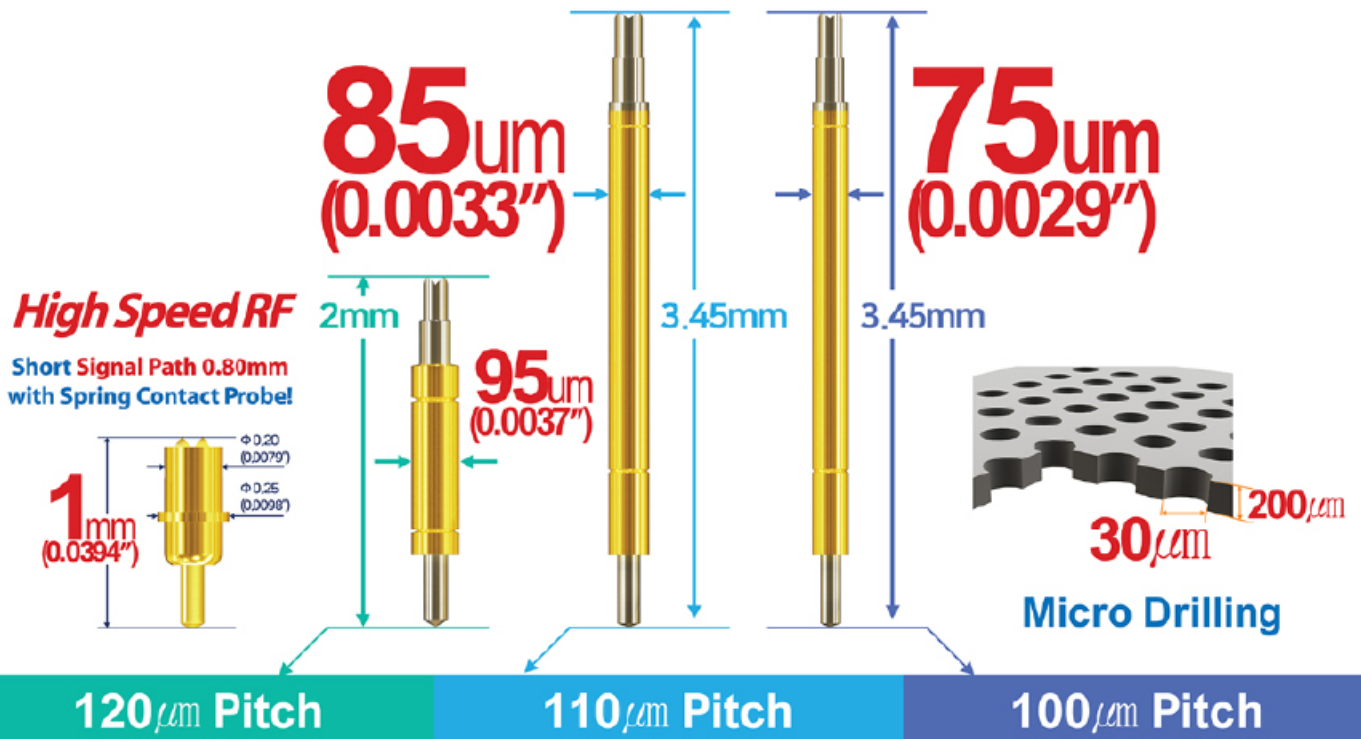
Subscribe



Fine Pitch Probe & Probe Head

Continuous Non-stop Innovation!

Proven Mass Production Capability

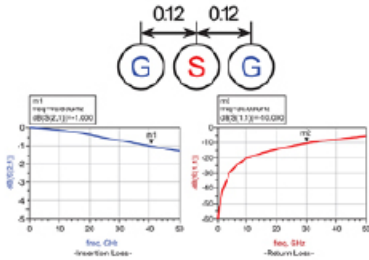


Mechanical Spec.

- Spring Force : 0.281oz (8.0g) @ .0098 (0.25mm)
- Recommended Travel : .0098 (0.25mm)
- Full Travel : .0118 (0.30mm)
- Material : Terminal – Pd Alloy / No plated
Plunger – Pd Alloy / No plated
Barrel – Ni-Au Alloy / Au plated
Spring – Music Wire / Au plated

Electrical Spec. (Simulation Data)

- Current Rating : 1.0A
- Propagation Delay : 20.80ps
- Capacitance : 0.21pF
- Inductance : 0.38nH
- Insertion Loss : 40.83GHz @ -1.000dB
- Return Loss : 30.03GHz @ -10.000dB (Dielectric material : CERAMIC)

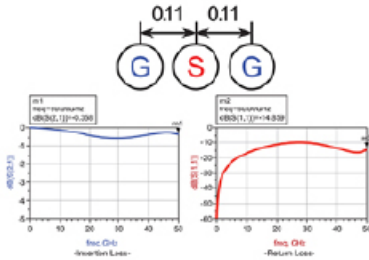


Mechanical Spec.

- Spring Force : 0.212oz (6.0g) @ .0118 (0.30mm)
- Recommended Travel : .0118 (0.30mm)
- Full Travel : .0138 (0.35mm)
- Material : Terminal – Pd Alloy / No plated
Plunger – Pd Alloy / No plated
Barrel – Ni-Au Alloy / Au plated
Spring – Music Wire / Au plated

Electrical Spec. (Simulation Data)

- Current Rating : 0.9A
- Propagation Delay : 38.25ps
- Capacitance : 0.47pF
- Inductance : 0.63nH
- Insertion Loss : > 50.00GHz @ -1.000dB
- Return Loss : > 50.00GHz @ -10.000dB (Dielectric material : CERAMIC)

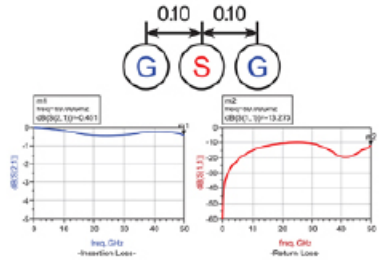


Mechanical Spec.

- Spring Force : 0.247oz (7.0g) @ .0118 (0.30mm)
- Recommended Travel : .0118 (0.30mm)
- Full Travel : .0138 (0.35mm)
- Material : Terminal – Pd Alloy / No plated
Plunger – Pd Alloy / No plated
Barrel – Ni-Au Alloy / Au plated
Spring – Music Wire / Au plated

Electrical Spec. (Simulation Data)

- Current Rating : 0.8A
- Propagation Delay : 35.55ps
- Capacitance : 0.44pF
- Inductance : 0.66nH
- Insertion Loss : > 50.00GHz @ -1.000dB
- Return Loss : > 50.00GHz @ -10.000dB (Dielectric material : CERAMIC)





Extending chiplet integration to 3D enables the placement of dies on top of each other, thereby providing added capacity without the added lateral distance. This keeps the latency low, and the dynamic power low. By freeing up valuable space inside the package, you can also fit more cores, and more transistors within a given package size.

Cover image courtesy of AMD

DEPARTMENTS

TECHNOLOGY TRENDS

6 Scalable silicon photonics packaging using optical bump nanoimprint lithography

By Hesham Taha

[Teramount Ltd]

Martin Eibelhuber

[EV Group]

49 Pressure clip sintering for high-power electronics

By Eric Kuah

[ASM Pacific Technology Ltd]

FEATURE ARTICLES

11 The next frontier: Enabling Moore's Law using heterogeneous integration

By Raja Swaminathan

[AMD]

We put **Heterogeneous Packaging** to the test.

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.

Enabling the Future |

© 2022 Amkor Technology, Inc. All Rights Reserved.



www.EVGroup.com

ACCELERATING HETEROGENEOUS INTEGRATION

- EV Group's Heterogeneous Integration Competence Center™ accelerates new product development fueled by heterogeneous integration and advanced packaging
- Wafer-to-Wafer (W2W) and Die-to-Wafer (D2W) hybrid bonding processes ready for sample test, product development and qualification
- Open access innovation incubator for EVG customers and partners across the microelectronics supply chain, guaranteeing the highest IP protection standards
- Combining EVG's world-class wafer bonding, thin-wafer handling and maskless, optical and nanoimprint lithography products and expertise, as well as pilot-line production facilities and services

VISIT US AT BOOTH #935

SEMICON[®]
WEST

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

STAFF

Kim Newman
Publisher
knewman@chipscale.com

Lawrence Michaels
Managing Director/Editor
lmichaels@chipscale.com

Debra Vogler
Senior Technical Editor
dvogler@chipscale.com

SUBSCRIPTION—INQUIRIES

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

Advertising Production Inquiries:
Lawrence Michaels
lmichaels@chipscale.com

Copyright © 2022 Haley Publishing Inc.
Chip Scale Review (ISSN 1526-1344) is a registered trademark of
Haley Publishing Inc. All rights reserved.

Subscriptions in the U.S. are available without charge to qualified
individuals in the electronics industry.
Chip Scale Review, (ISSN 1526-1344), is published six times a
year with issues in January-February, March-April, May-June, July-
August, September-October and November-December. Periodical
postage paid at Gilroy, Calif., and additional offices.

POSTMASTER: Send address changes to Chip Scale Review magazine
P.O. Box 2165
Morgan Hill, CA 95038
Tel: +1-408-846-8580
E-Mail: subs@chipscale.com

Printed in the United States

FEATURE ARTICLES

25 Semiconductor test: staying ahead of nanodevices

By Tucker Davis, Brian Brecht
[Teradyne]

32 Status and outlook for fan-out wafer/panel-level packaging

By John H. Lau
[Unimicron Technology Corporation]

43 Fan-in wafer-level packaging: any chip can be a flip chip!

By Ray Fillion
[Fillion Consulting]

Call for Technical Articles and Technology Trends Columns

Original, well-written articles are the lifeblood of *Chip Scale Review*. We welcome your contributions on advanced semiconductor packaging technologies, processes and materials. CSR features columns on current technology trends, market updates and insightful guest editorials are typically included in every issue. Our spectrum of content coverage is specifically focused on these *BEOL & MEOL* topics:

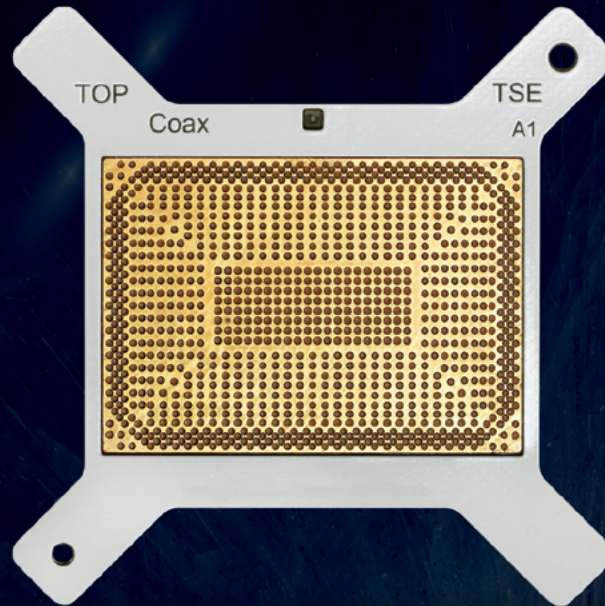
- Wafer-level Packaging (WLP)
- Panel-level Packaging (PLP)
- Heterogeneous Integration (HI)
- AI and Quantum Technologies
- Through-Silicon Vias (TSVs)
- 3D Stacking & RDL Interconnects
- System-in-Package (SiP)
- Package-on-Package (PoP)
- System-on-Package (SoP)
- Chip-on-Wafer (CoW)
- Multi-Chip Modules (MCM)
- Chiplets
- Lithography
- Organic and Inorganic Substrates and Interposers, Glass Based Solutions
- Advanced IC Assembly/Packaging
- Bonding/Debonding
- Integrating ICs with Nanotechnology
- MEMS, RF/Wireless
- Optoelectronic/Photonic Devices
- Inspection/ Metrology
- Wafer/Device/System Test & Burn-in

Send abstracts to: editor@chipscale.com



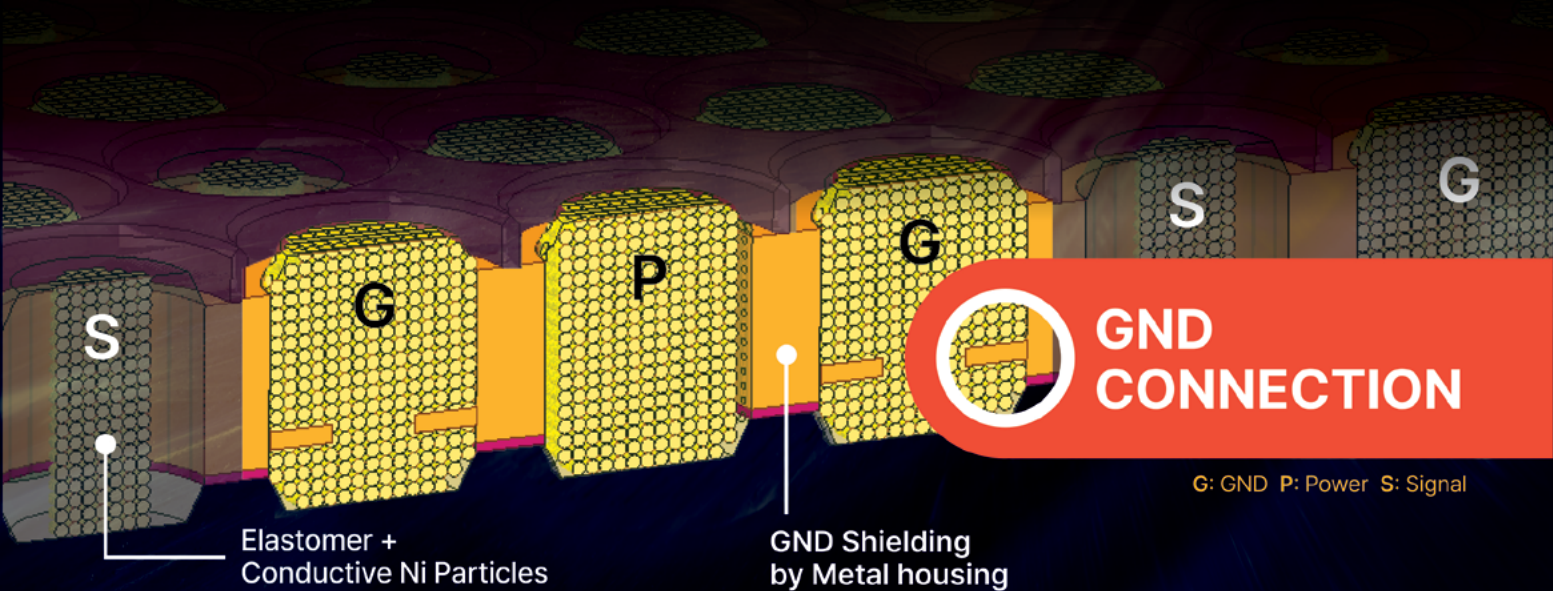
COAXIAL ELASTOMER SOCKET

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



ELTUNE-coax™

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

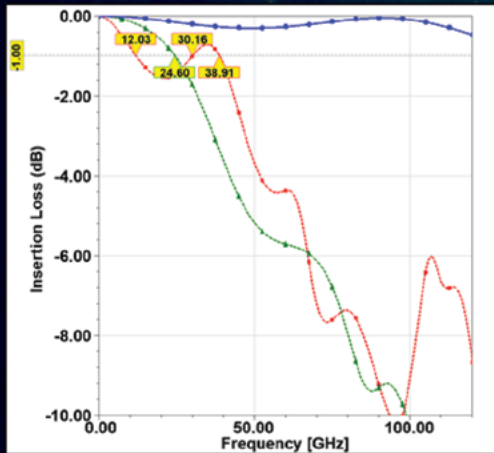


COAXIAL ELASTOMER SOCKET

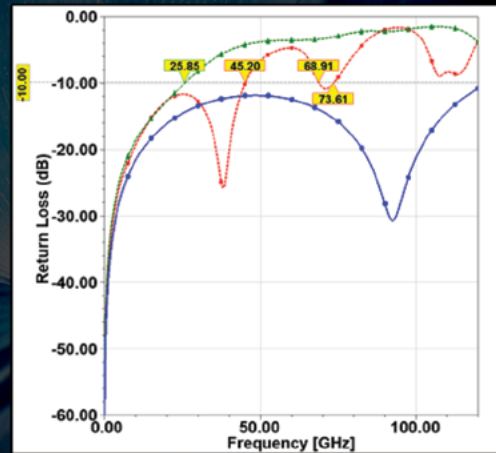
for >64Gbps ATE/SLT Test
for Crosstalk free Board to Board Connector

ELTUNE-coax™

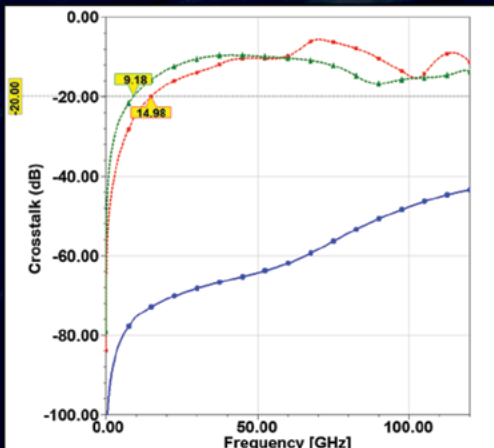
Electrical Specifications (unit: GHz)			
50Ω, 0.80mm pitch	Spring pin	Elastomer	ELTUNE-coax™
Electrical Length(mm)	3.05	0.60	0.60
Insertion Loss(S21) @-1dB	12.03	24.60	>100
Return Loss (S11) @-10dB	45.20	25.85	>100
Crosstalk (S31) @-20dB	14.98	9.18	>100



Insertion Loss



Return Loss



Crosstalk

TSE USA

Jong M. Lee jmlee@tse21.com
Matt Hansen matt@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



Scalable silicon photonics packaging using optical bump nanoimprint lithography

By Hesham Taha [Teramount Ltd] and Martin Eibelhuber [EV Group]

Silicon photonics has emerged as a promising platform for supporting the ever-growing demand for high-speed data transfer, low-power consumption and low latency, which are required for the next generations of data centers, advanced computing, and 5G/6G networks and sensors. The silicon photonics market has expanded significantly in the last few years and is expected to grow at a 26.8% compound annual growth rate (CAGR) over the next five years [1]. While wafer manufacturing capabilities for silicon photonics are well advanced through the use of standard semiconductor mass-production processes and existing infrastructure, silicon photonics packaging and testing are still behind and lack production scalability, which limits wider deployment of silicon photonics. Photonic Bump technology, a new wafer-level implementation of optical elements for scalable packaging and testing capabilities, is presented in this article. The Photonic Bump is an equivalent of electrical solder bumps and has the potential to align silicon photonics with standard semiconductor wafer manufacturing and packaging lines, thereby bridging the gap in silicon photonics toward high-volume manufacturing.

Fiber-to-chip assembly is the main limiting factor in existing silicon photonics packaging solutions, which use direct fiber bonding on a photonic chip with adhesives through active alignment or specialized high-precision alignment equipment. These are limited in their volume manufacturability, scalability to large numbers of fibers, compatibility with packaging processes such as reflow requirements, and integration with electronics packaging. The essence of the problem is related to geometrical constraints of tight assembly tolerances when packaging single-mode fibers with silicon or nitride waveguide channels on a photonic chip, as well as related to the complex side-coupling geometry. These impose critical obstacles for silicon photonics to be able to be applied to wider applications such as co-packaged optics in ethernet switches, advanced computing and future chip-to-chip optical connectivity.

Teramount and EV Group have collaborated to adopt wafer-level optics technologies in order to enhance silicon photonics packaging processes. Under this collaboration, nanoimprint lithography (NIL) has been used for wafer-level implementation of Photonic Bumps on silicon photonics wafers.

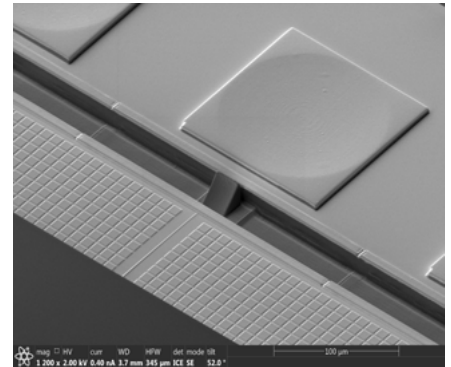


Figure 1: Photonic Bump wafer-level imprint on a silicon photonics wafer at accurate placement relative to waveguide channel.

Photonic Bump elements provide unique optical coupling functionalities that include: 1) a vertical beam deflection to enable wide-band surface coupling as a replacement for the complicated side-coupling geometry; and 2) a spot size conversion for mode matching between single-mode fiber and the chip's waveguide (see **Figure 1**). In addition, Photonic Bumps are used to enable the “self-aligning optics” scheme when connected with the Teramount PhotonicPlug fiber connector [2], which enables fiber-chip assembly tolerances of larger than $\pm 20\mu\text{m}/1\text{dB}$ (see **Figure 2**).

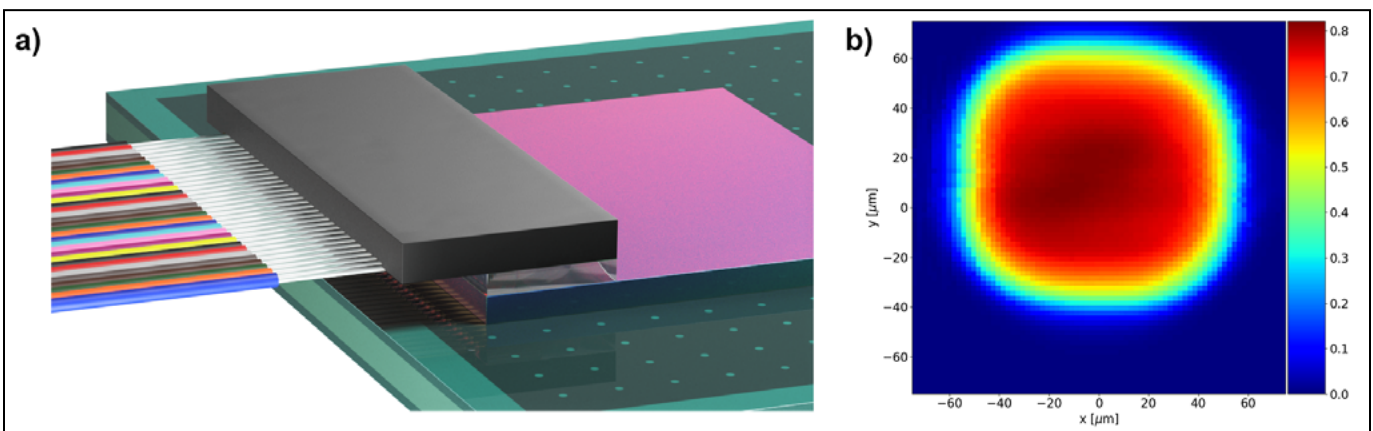
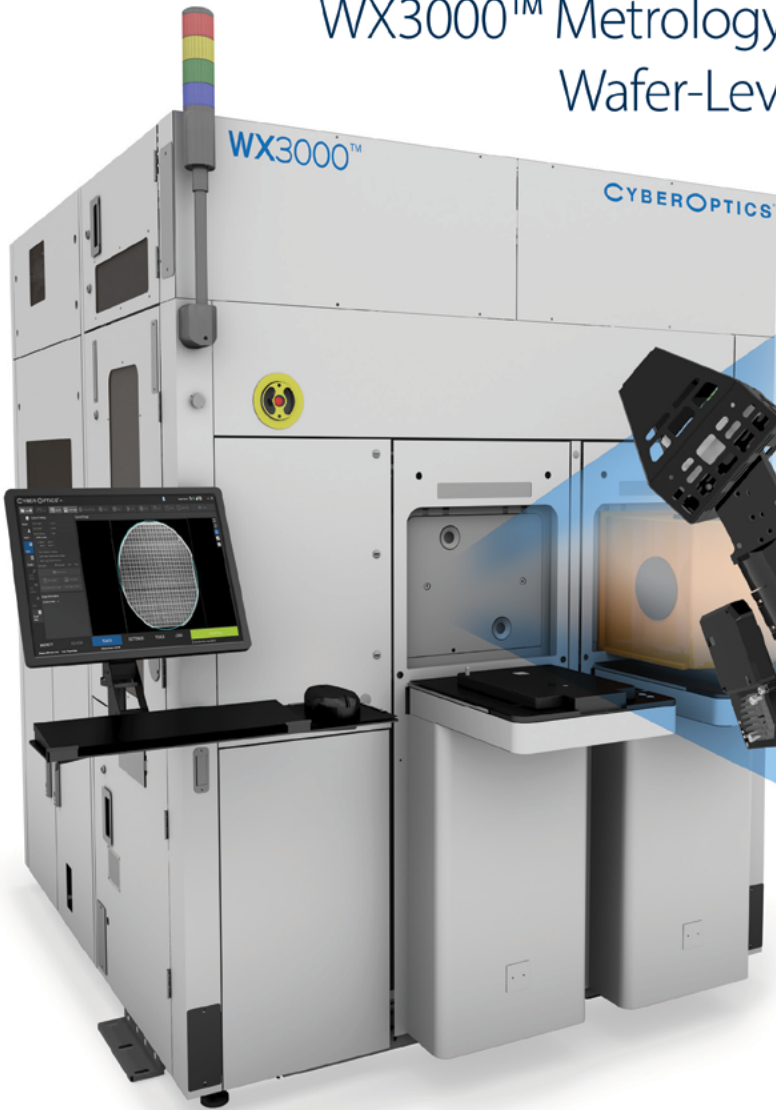


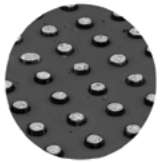
Figure 2: a) (left): PhotonicPlug assembled on a photonic “bumped” silicon photonic chip. b) (right): Measured XY fiber-chip tolerance providing larger than 100 times of fiber-chip assembly tolerance compared with existing technologies.

Seriously Fast.

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



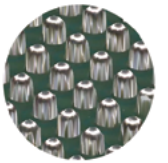
Copper Pillar



Micro Bump



Flip Chip (c4)



BGA Ball



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.

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.

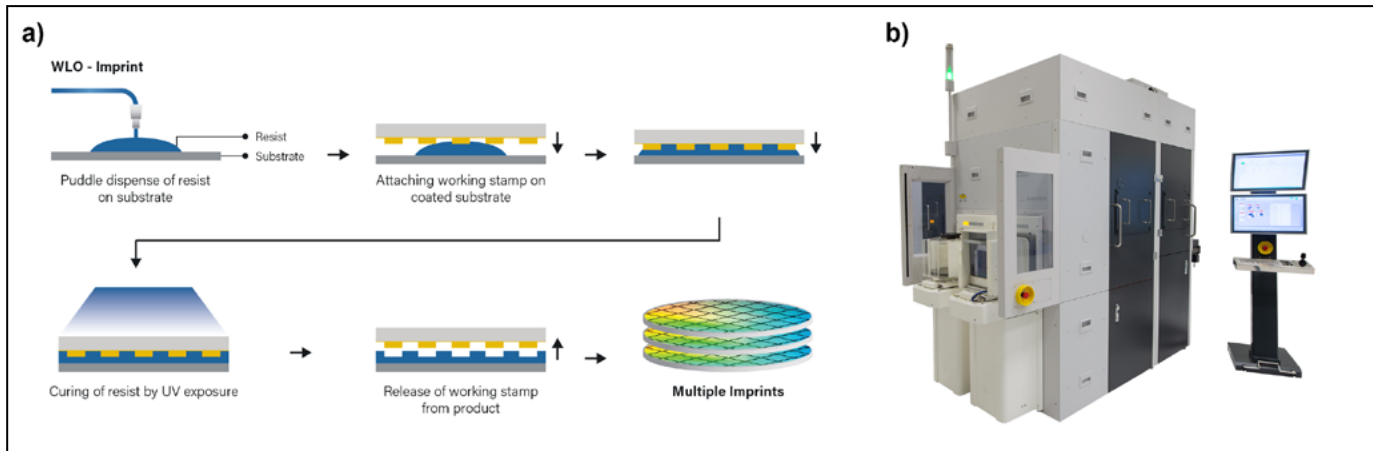


Figure 3: a) (left) Typical process flow for imprinting wafer-level optics, which can be accomplished by b) (right) an EVG7300 UV nanoimprint lithography system. The EVG7300 UV-NIL system can support multiple processes, including SmartNIL, wafer-level optics and device stacking.

The combination of large assembly tolerances and wide-band surface coupling enables the transition of silicon photonics packaging from specialized equipment to standard passive alignment assembly protocols and tools. This transition supports high-yield and high-volume packaging. In addition, it allows unique packaging protocols such as detachable and post-reflow fiber connectivity, which are optimized for assemblies with large numbers of fibers and co-packaged optics applications. Moreover, the PhotonicPlug and the Photonic Bump, fueled with their surface coupling and large tolerances, create an effective wafer-level testing capability prior to wafer dicing, thereby enhancing silicon photonics wafer manufacturing yields.

Application of NIL to silicon photonics

NIL has proven to be the most effective method of replicating complex structures, such as 2.5D features, grayscale patterns and freeform optics, because it is not limited to the constraints of optical lithography. Standard optical lithography is optimized to build up structures layer by layer. While this layer by layer approach makes it ideally suited to the needs of the electronics industry, it is not sufficient for manufacturing photonic structures. In contrast, NIL enables the patterning of 3D structures in a single-step process, which is ideally suited for the photonics industry where light-matter

interaction relies largely on shape and geometry [3]. For example, NIL allows the imprint of complex geometries such as sharp edges of deflector mirrors, curved surfaces, high and low aspect ratio structures as well as imprinting in deep cavities. Wafer-level optics (WLO) processes have long proven their high repeatability in high-volume production for optical sensors and are now being leveraged for photonic packaging.

The NIL process offers significant yield and cost advantages for the above-mentioned structures compared to conventional manufacturing methods, such as diamond drilling, laser direct writing and electron-beam writing, which have very low throughput and are therefore difficult to scale up to larger substrates and volume-production environments. Incorporating the NIL process enables the use of best-performing dies and the ability to efficiently bring these high-quality patterns into production lines.

Teramount worked with EVG to establish suitable manufacturing process solutions for Teramount's Photonic Bump. In the development work, a wafer-scale master stamp with the Photonic Bump structures was produced from a single-die "hard master" using EVG's Step and Repeat (S&R) NIL process. This scaling enables wafer-level mass-production processes, and is typically based on two steps. First, the S&R master is used to replicate multiple working stamps. Next, the working stamps are used to imprint

the functional photonic structures on the target substrates (see **Figure 3**). While multiple replications are needed to support the scaling and to avoid wear out of the single die master, the final imprints of the fully-functional optical structures demonstrated high pattern fidelity, precise alignment and precise control of desired layer thicknesses. Scanning electron microscope (SEM) inspection showed residual layer thickness at <1% of the structure height and high alignment accuracy to within less than 500nm. In particular, the precise alignment to the optical structures underneath the photonic chip is crucial for the excellent coupling performance described above.

Working in conjunction with the Photonic Bump packaging technology, NIL is now making wafer-scale packaging possible in the photonics industry, which could have a profound impact on lowering packaging and overall product costs. Whereas packaging is still a relatively small (but growing) share of overall complementary metal-oxide semiconductor (CMOS) production costs, it represents the majority of overall cost in photonics manufacturing, which still relies on single-device packaging schemes. Wafer-level integrated photonics, enabled by NIL and Photonic Bump packaging, has the potential to flip this equation.

The ability of NIL to provide accurate placement of optical elements on silicon photonics wafers plays a critical role in shifting the typical fiber packaging

complexity from the assembly domain to the wafer manufacturing domain. It provides an ideal platform for post-processing of silicon photonic wafers for the photonic “bumping” process to be performed either at semiconductor foundries, or at outsourced semiconductor assembly and test (OSAT) facilities. As part of the joint collaboration, EVG provided NIL process development and prototyping services through its NILPhotonics Competence Center, as well as expertise in both CMOS and photonics manufacturing, to assist Teramount in accelerating the development and production of its PhotonicPlug technology.

Summary

The Photonic Bump is a transformational solution for establishing a scalable silicon photonics packaging platform that generates, for the first time, an effective “through-chip optical via” for seamless photonics and electronics integration through 3D packaging and interposer geometries. It holds the promise to align silicon photonics with the standard semiconductor manufacturing ecosystem and to leverage silicon photonics to volume manufacturing for a variety of emerging applications.

References

1. MARKETSANDMARKETS Report, “Silicon Photonics Market with COVID-19 Impact Analysis by Product (Transceivers, Switches), Application (Data Center & High-performance Computing, Telecom.), Waveguide, Component, and Geography - Global Forecast to 2027,” Nov. 2021.
2. A. Israel, et al., “Photonic plug for scalable silicon photonics packaging,” Proceedings Volume 11286, Optical Interconnects XX; 1128607 (2020).
3. M. Eibelhuber, et al., “Nanoimprint Lithography Enables Cost-effective Production of Photonics,” *Photonics Spectra*, Feb. 2015.

Biographies

Hesham Taha is CEO at Teramount, Jerusalem, Israel. He has a PhD in

Applied Physics from the Hebrew U. of Jerusalem with a focus on nano-lithography. Formerly, he worked as R&D scientist, Sales & Marketing manager for Nanonics Imaging. Email: hesham.taha@teramount.com.

Martin Eibelhuber is Product Manager at EV Group, Florian am Inn, Austria,

where he focuses on NIL-related equipment and technology. He has a doctorate in technical physics from the Johannes Kepler U. Linz, specializing in nanoscience and semiconductor physics.



Global Services for Wafer Level Packaging

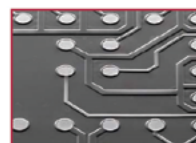
Electroless Plating

- NiAu for Low Cost Bumping
- High Reliability NiPdAu
- Cu & Au Wire Bonding



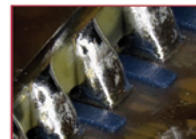
Electroplating

- Wafer Level Redistribution
- Cu Pillars
- NiFe for MEMS



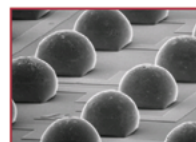
Solder Ball Bumping

- 4"-12" Wafers
- BGA-like devices
- 3D-Applications



Solder Rework & Reballing

- for CSP, BGA, LGA, CLCC, PCB, MEMS etc.
- No Tooling Required



Wafer Backend Processing

- Backside Metallization
- Thinning & Dicing
- Tape & Reel



www.pactech.com





KOSDAQ
LISTED COMPANY

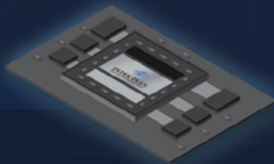


SEMICONDUCTOR PACKAGE INSPECTION

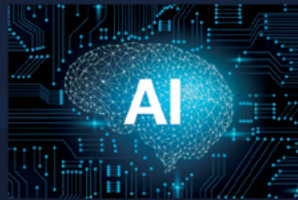


INTEKPLUS specializes in 2D and 3D Surface Inspection and Metrology for Semiconductor Package Products. We provide the World's Fastest Inspection Capability to achieve High Productivity.

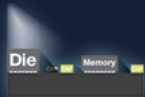
Supreme Vision Solution



Heterogeneous Integration



AI Deep Learning



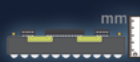
Shadow Free



Micro Crack (Die Mold)



SFF to LFF Size Coverage



Total Height

Product Line

Out-Tray Platform

- World Fast Pick & Place Inspection Solution
- Advanced Solution for Heterogeneous Package



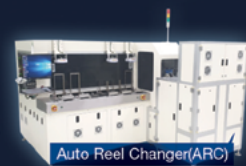
In-Tray Platform

- Full PVI Function for 6 Side Inspection
- World Fastest Productivity



Inspection + Tape & Reel

- Integrated with Vision System and TR
- Full Automation for Post Reel Process



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

For more information

www.intekplus.com / sales1@intekplus.com



The next frontier: Enabling Moore's Law using heterogeneous integration

By Raja Swaminathan [AMD]

The explosion of connected devices over the last 40 years in our industry has driven an explosion of semiconductor content riding the back of Moore's Law. Starting with the personal computer cycle then continuing with smartphones and Internet of Things (IoT) devices, silicon has permeated every aspect of our lives. This explosion of semiconductor content in everything from devices in our pockets to integrated into our clothing has led to the birth and growth of a new era of high-performance computing (HPC), as all the data being generated is processed into useful information to improve our lives.

From the cloud to the edge, and from artificial intelligence (AI) to 5G communications, the insatiable demand for HPC has become a driving force within the microelectronics industry and it will shape the next several generations of technology and design innovation. The demand for compute is accelerating rapidly with the doubling of HPC system performance every 1.2 years. This trend is much faster than Moore's Law, which currently has slowed to doubling of transistor density every 2-3 years; so, the compute capability is clearly driven

by innovations outside of the raw silicon. The demand for HPC is not simply bragging rights of being in the top 500 supercomputers. These devices are solving problems that are pressing for humanity including drug discovery, climate models, new energy exploration and many more. Today's best compute platforms only whet our appetite for more as the possibilities for solution finding become more compelling.

Demand for computation is outpacing Moore's Law

The next bit of sobering data regards the much-discussed cracks in Moore's Law. As we know, silicon technology node introductions have been slowing down, and simultaneously delivering less benefit, while at the same time, the costs per yielded mm² of silicon are going up. This is particularly challenging because the semiconductor industry has thrived on delivering more performance and features in each generation by adding transistors. With these trends, the cost per transistor will stop scaling in the next few years, which creates notable economic headwinds to meeting the demand. These

costs are not just a result of inflationary pressure but based on the underlying physics and complexity of these new nodes.

The next aspect of the slowdown in node introductions is that scaling factors are diverging between different intellectual property (IP) types, with static random-access memory (SRAM) and especially analog circuits lagging well behind the scale factors of logic. This leads to the chip-level view of area scaling where, with a mix of logic, SRAM and analog content, we will not be able to shrink chip designs appreciably toward the end of this decade. This illustrates that the irresistible force of compute demand is colliding with the immovable object of device physics, creating an environment where new architecture approaches and non-device innovations are critical for our ecosystem.

It is now recognized that conventional computing is approaching fundamental limits in energy efficiency. Historical trends show that general purpose CPU energy efficiency worsens with higher performance, so new approaches are required (Figure 1). We are also finding new approaches to reduce energy for compute. Modular design, chiplets and 3D

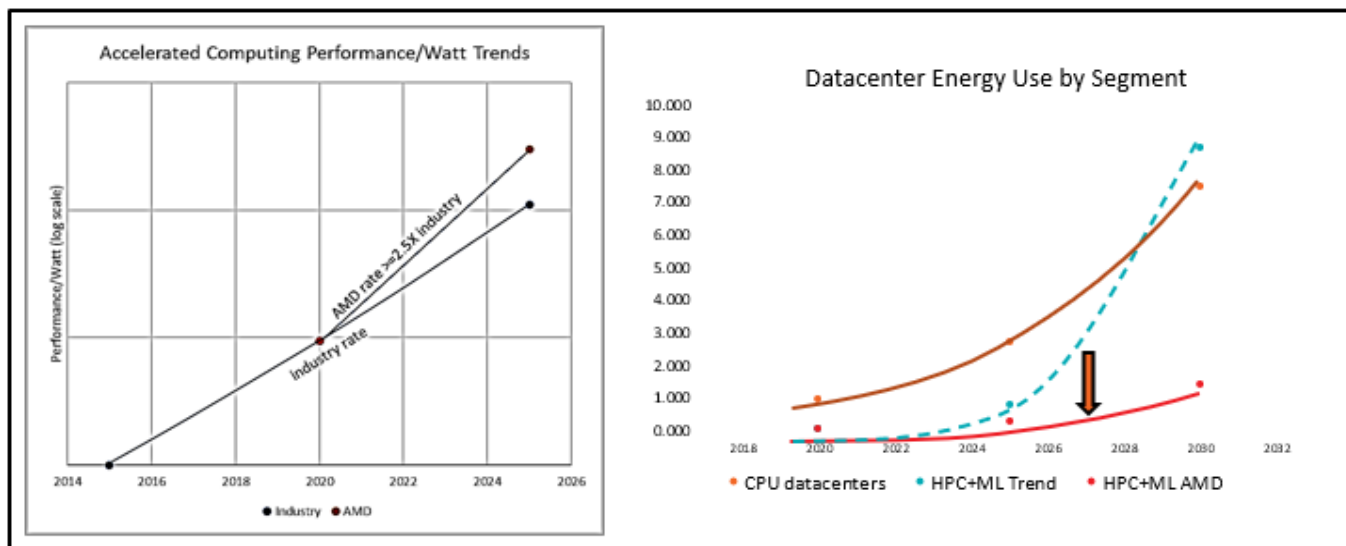


Figure 1: AMD's efficiency goal for HPC/AI applications.

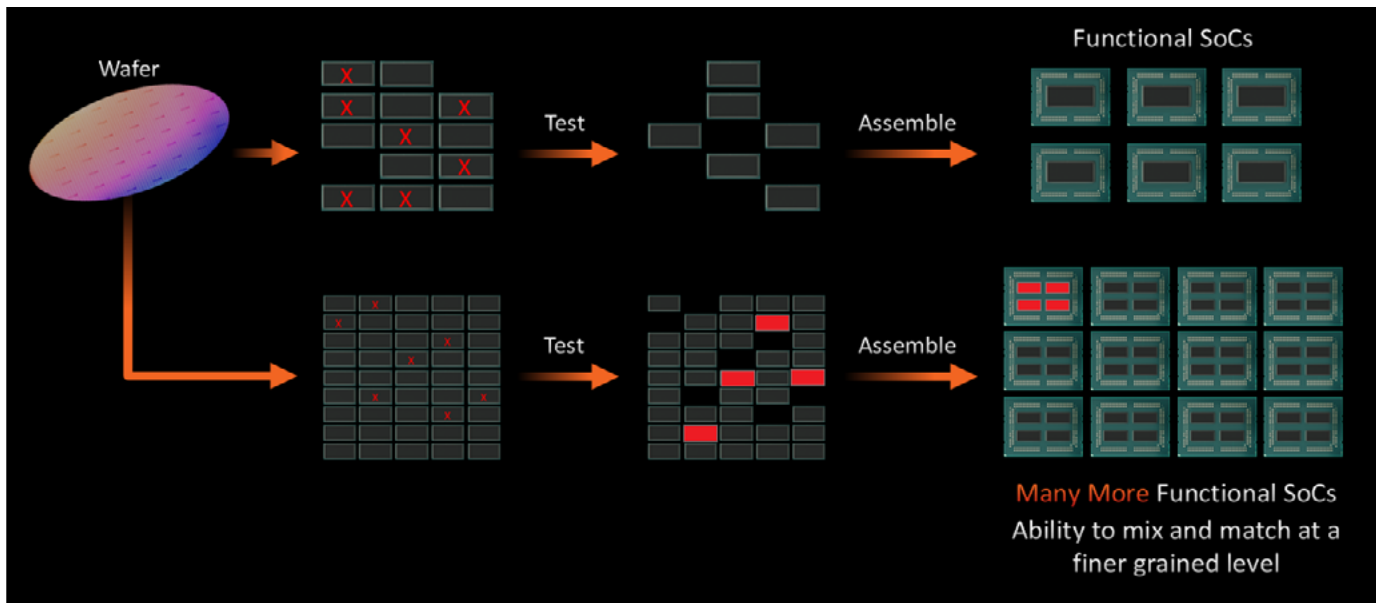


Figure 2: High-level approach to chiplets.

stacking are the next frontier for efficiency gains. Application-specific optimization provides better performance-per-Watt. The last five years show an industry efficiency improvement rate of 12X for HPC and AI nodes. The AMD goal is to dramatically accelerate this improvement rate to 30x by 2025.

So, we have the bright future of exploding compute demand and simultaneously the dark cloud of technology headwinds. The trillion-dollar question is how to architect, design, and build future systems that solve these challenges. The answer is increasingly clear that modular, multi-chip design is a fundamental enabler. Systems must be more specialized for the task they are running. General purpose is no longer generally applicable. We need efficient accelerators, and we need economically viable ways to continue to deliver this performance in the face of the formidable cost trends. Let us take a closer look at what modular design can do, and what the enabling technology requirements are.

Let us start off with the magic of chiplet-based design, which is becoming much more pervasive. AMD led the way in this approach with our heterogeneous technology server and desktop products back in 2019. An initial motivation for chiplets was economics. Back in the day, Moore’s Law enabled a doubling of transistors and capability in each generation, and all was good. Lately, this has not worked out as well. With shrink factors slowing down while compute demand has not, die sizes have been growing at an unsustainable rate.

With chiplets, we can split a formerly monolithic system-on-chip (SoC) into two components to improve performance. However, this results in a non-trivial overhead associated with “chipletizing” the design. Each die needs test capability, power management, and an interface so it can talk to the other chiplets. These interfaces will not be as small, low latency, or power efficient as on-die wires; therefore, the architecture needs to accommodate new boundaries and complexity.

To illustrate benefits of the chiplet approach, let us consider the yield dynamics. With a single large die and a fixed number of defects on a wafer, we yield a small set of functional SoCs for a wafer’s worth of chips (**Figure 2**). As soon as we split that big SoC up into, say, four chiplets, the yield dynamics start to work in our favor. The same number of defects now just take out a small chiplet, and we can use our wafer sort capabilities to select the good ones and build more functional SoCs from the same silicon. This is one factor that has helped AMD to meet market demand better when wafer supplies are so constrained.

We also gain the flexibility of building chiplet SoCs with varying numbers of chiplets to address different markets. Perhaps a less obvious benefit is that we can cherry pick faster chiplets from the wafer and assemble them into higher-performance and higher-priced SoCs for customers who want, and will pay for, the greatest performance possible.

The benefits described above are substantial, though modular design is bigger than just decomposing an SoC into chiplets.

We want to build tailored products for specific markets by mixing and matching chiplet types. Some chiplets can be general purpose CPUs, others can be more specialized. We can now specialize a domain-specific chiplet and include more or fewer of them for a given product. However, the success of this approach is heavily dependent on the package technologies used to assemble these dice and enable them to communicate with each other.

Package architectures

Many package architectures exist in the industry to enable die-to-die interconnections across various product segments (e.g., mobile, PC, server, and desktops) (**Figure 3**).

Examples include:

- Multi-chip module (MCM) architectures from AMD and other industry players.
- Other 2D architectures based on redistribution layer (RDL)-like interconnects (or 2D-organic) like integrated fan-out with redistribution layer (INFO-R), and fan-out chip-on-substrate (FoCoS).
- 2D silicon-based architectures like embedded multi-die interconnect bridge (EMIB), AMD’s elevated fan-out bridge (EFB), Integrated fanout with integration of an LSI (INFO-L), and Si interposer where the die-to-die interconnect is achieved using passive Si; as well as
- 3D architectures — defined as active-on-active Si stacking, such as the AMD 3D V-Cache™, Foveros/



LBSemicon

Look Beyond **Best Solution Provider**

WLP / Bumping

Test / COF

Back-end / RDL

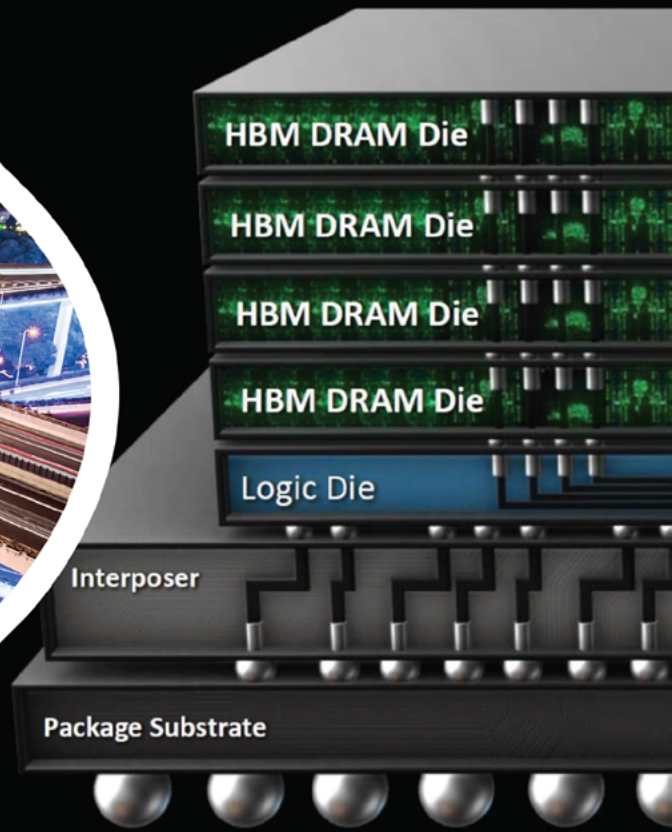
Since semiconductor chip manufacturers seek to produce increasingly lighter and thinner products, wafer bumping technology is expanding its application scope and creating more value. In response to this trend, LB Semicon is deeply committed to continuous R&D to support your business and provides advanced bumping solutions to top IC manufacturers both at home and abroad. As the best solution provider, LB Semicon has contributed to its even more advanced technology of today.



HQ: Korea

www.lbsemicon.com

V I S I T U S



When Moore's Law no longer gets your ICs where they need to go.

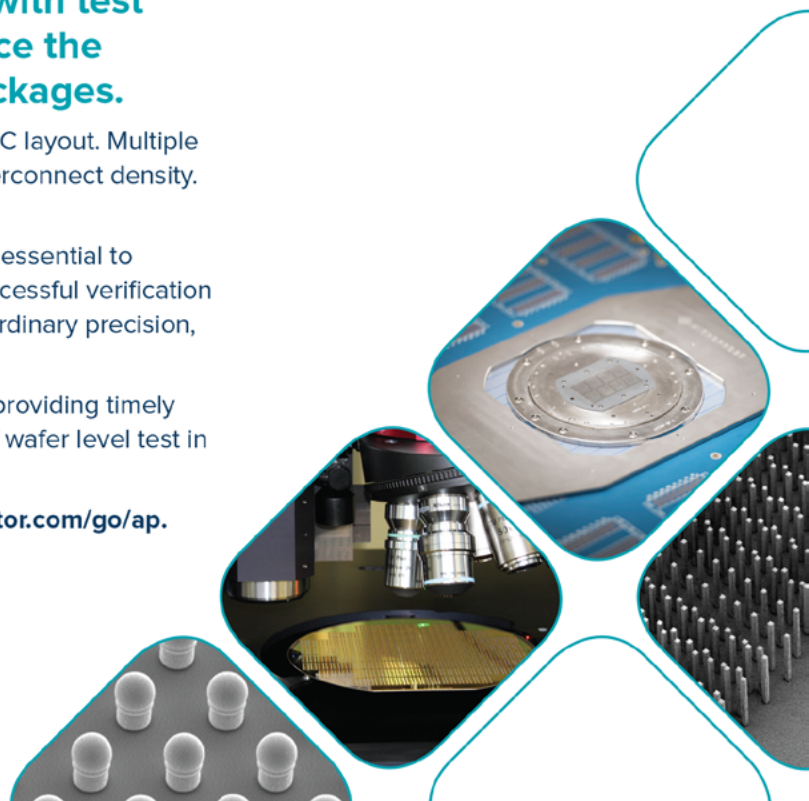
FormFactor takes on the challenge with test and measurement solutions to reduce the manufacturing cost of advanced packages.

Advanced packaging adds a new vertical dimension to IC layout. Multiple dies merge into single systems with unprecedented interconnect density. Performance goes up. Power consumption goes down.

And wafer-level test and measurement becomes nearly essential to guarantee cost-effective fabrication and packaging. Successful verification at this level requires probing and measuring with extraordinary precision, optically, electrically and mechanically.

FormFactor not only understands the problem, it's also providing timely solutions to advance yield knowledge at every phase of wafer level test in advanced IC packaging.

Let us help you rise to the occasion. Visit www.formfactor.com/go/ap.



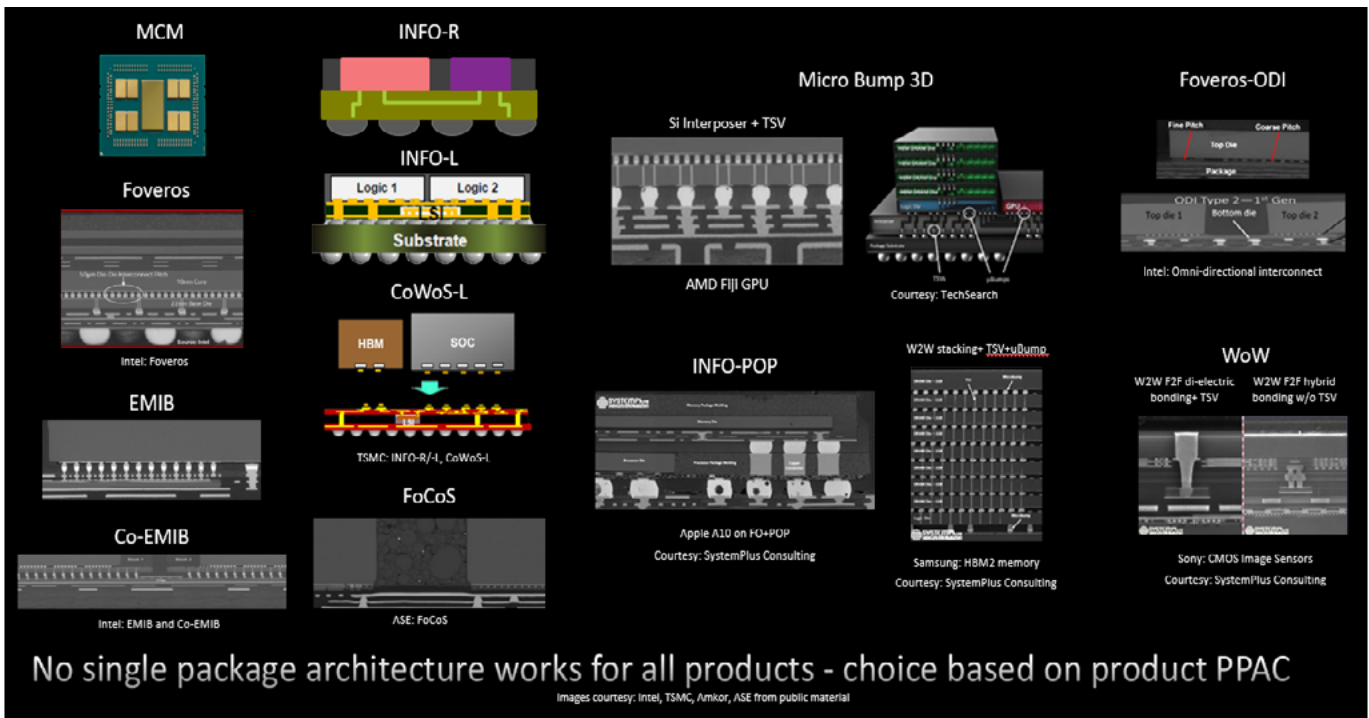


Figure 3: Sample package architecture options for die-to-die chiplet interconnects.

omni-directional interconnect (ODI), wafer-on-wafer (WoW) architecture found in image sensors and the memory markets.

Chiplet package architecture choice is not a one size fits all approach, rather it is made based on specific power, performance, area, and cost (PPAC) requirements per product. A critical dimension of making this all work is driving the overhead of those interfaces down. One way to quantify this is to tabulate the linear interconnect density and the areal interconnect density of packaging approaches (Figure 4).

MCMs are great, low-complexity designs, but the low connection density of this technology limits its applications to specific boundaries for the chiplets. For instance, the AMD EPYC™ and Ryzen™ lines chose to put the CPU cores on one chiplet and the IO and memory interfaces on another one. This works with MCM because the CPU bandwidth requirements are relatively modest and can be supplied across highspeed SERDES routes.

To accomplish more exotic SoC chiplet configurations, higher bandwidths are required. In the middle of Figure 4 is an example Radeon Instinct™ design, which requires high-bandwidth memory to feed the compute engines. To supply over a terabyte per second of bandwidth to memory, a

higher density interconnect is required. We chose passive silicon interposers for the first instance, and most recently, the elevated fan-out bridge approach.

The holy grail of chiplet architecture is of course 3D stacking. The 3D hybrid bond approach that we have recently introduced with AMD 3D V-Cache™ provides dramatically higher bandwidth density, which has enabled us to connect a 64MB cache chiplet directly on top of the 32MB of existing cache, which required thousands of signals—so the package technology choice is very specific to the architecture. The choice can be visualized in a simplified way. The higher density package technologies are more

expensive because they require more precise patterning and many more processing steps; however, with that density comes the benefits of a reduction in interface area, and of course, lower energy for data movement. Chipletizing comes with the overheads including IO area, additional design effort and complexity, additional assembly and testing steps. Getting to the right architecture requires that we must ensure that the value of our newly modular solution with its configuration flexibility and yield has benefits that more than outweigh the costs. Getting this right is a highly multi-disciplinary endeavor, requiring engineers from different domains to rapidly iterate and provide solutions in new ways.

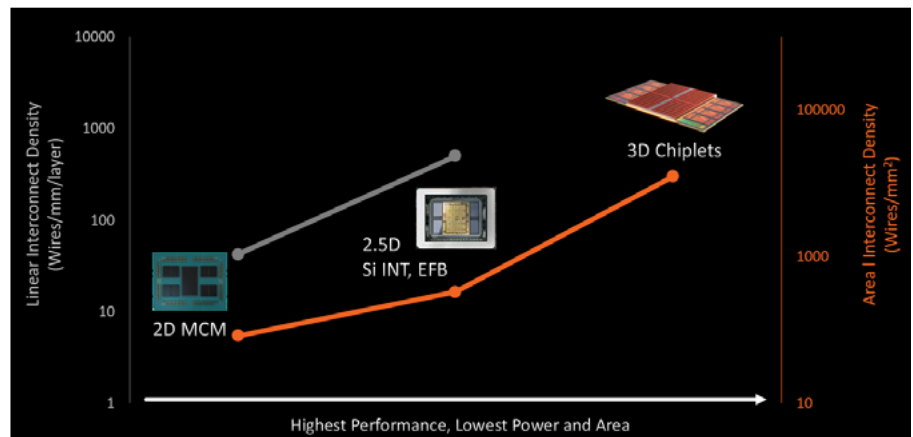


Figure 4: Improving key parameters that drive high-performance computing forward.

AMD elevated fan-out bridge architecture

To illustrate the result of one of these optimization challenges, let us focus on the implementation of a new package architecture called elevated fan-out bridge (Figure 5) that we recently announced for the MI200 GPU compute product. As noted earlier, these products require terabytes of memory bandwidth and therefore need denser connections than organic packages provide. One industry approach is shown on the left in Figure 5 that embeds the silicon bridge die, containing the interconnect wires, into a cavity carved out of the organic package. This has better electrical behavior than legacy 2.5D silicon interposer approaches because it does not require through-silicon vias (TSVs) though does come with challenges associated with the substrate embedding approach.

We decided to develop a cleaner approach that elevates that silicon bridge to live in the shadow of copper pillar bumps. We can thin these silicon bridges down so that there is no significant height impact to the compute die. We now avoid having to carve out a cavity in the substrate and can also lithographically define this module as a unit without dealing with micro bumps on the substrate. Getting better placement accuracy with this method provides an example of the evolution of package technologies and the innovation going on in this space. Chiplet designs can get quite complex with eight high-bandwidth memory (HBM) stacks, two compute die chiplets, and the elevated fan-out bridges (EFBs) to connect them. By choosing a technology that is robust and manufacturable, we have been able to deploy the tens of thousands of these required for the Frontier supercomputer.

AMD 3DVCache™

As computer architects know well, large on-die L3 caches can provide instructions per clock (IPC) uplifts for CPU performance, which is especially important in today’s world of ever-increasing appetites for compute and for large data sets. Not surprisingly, as we survey products across the industry over the past decades, there has been a steady increase in on-die cache sizes. So that begs the question, can this trend continue indefinitely? In fact, why is it that the on-die cache integration is starting to slow?

The answers to these questions lie in the barriers to large on-die caches. As noted earlier, Moore’s Law slowdown impacts different silicon functions differently. Analog circuits have not scaled much into the advanced nodes, and SRAMs, upon which on-die caches are largely based, are also not scaling as well as logic.

Increasing the on-die cache capacity, which also increases the die size and lowers the yield, is becoming increasingly cost prohibitive and also becomes a challenge for product flexibility. The performance afforded by large caches is important for some markets, though it can be overkill for other market segments to bear the added cost. Finally, larger area also means longer data path distances, which increases cache access latency power and can offset the performance gains.

Up to this point, chiplet integration had mostly meant 2.5D integration. For example, in a hypothetical CPU with a large cache, one can separate part of the cache into a separate die, or chiplet, and place them side by side. The smaller die sizes can improve yield, and therefore the cost, and it provides

the flexibility to have the CPU die with a smaller cache as a standalone product to address different markets.

As valuable as these benefits are, extending chiplet integration to 3D can break even more barriers. By placing the dies on top of each other, you can have the added capacity without the added lateral distance, so you can keep the latency low, and the dynamic power low by freeing up valuable space inside the package. You can also fit more cores and more transistors within a given package size. All these incentives led to the creation of the AMD 3D V-Cache™—the industry’s first high-performance processor product with 3D integration based on hybrid bond technology.

The AMD V-Cache™ consists of three main components. The first one is the “Zen 3” CPU core complex die (CCD). It is manufactured using TSMC 7nm FinFET technology. Each CCD contains eight cores in a core complex (CCX) and the eight cores share a 32MB L3 cache. It was able to achieve a 19% average IPC uplift over the previous “Zen 2” design, and it has a die size of 81mm² (Figure 6). What is important to point out here is that the AMD 3D V-Cache™ support, both architecturally and physically, was planned for and integrated into the CCD, from the beginning of “Zen 3” design.

The second component of the AMD 3D V-Cache™ is the extended L3 Die (L3D). Like the CCD, it was also built using TSMC 7nm FinFET technology. It has a die size of 41mm², which is roughly half of the CCD die size. The sizing was intentional to allow the L3D to fit over the CCD’s L2 and L3 cache area. The relatively low power density of the caches allowed the thermal impact

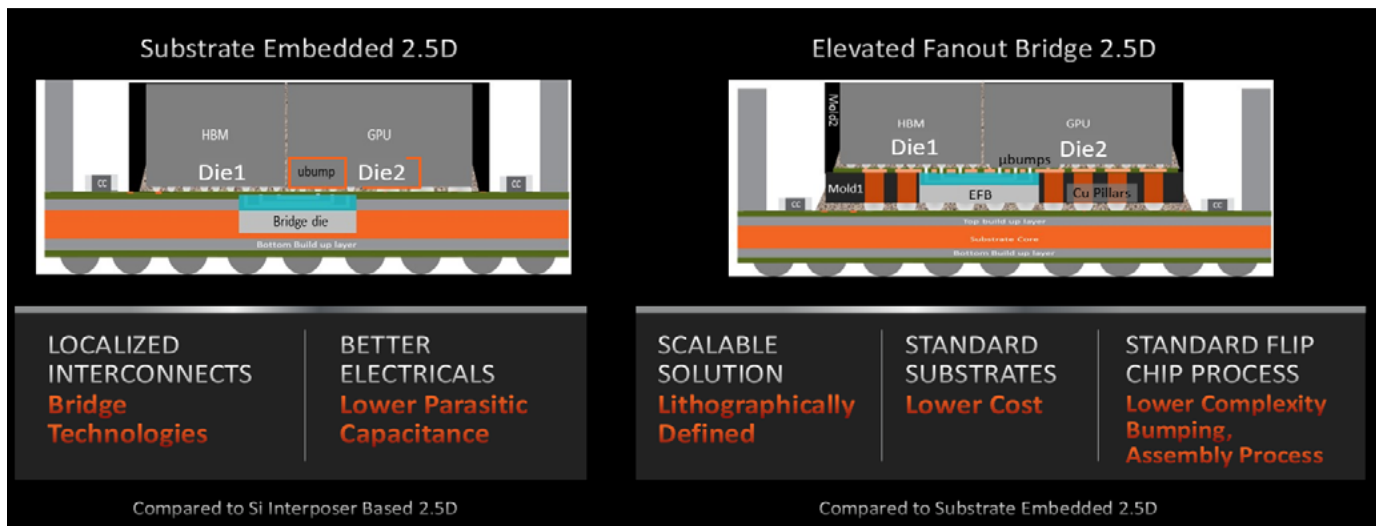


Figure 5: 2.5D “bridge” architecture landscape.

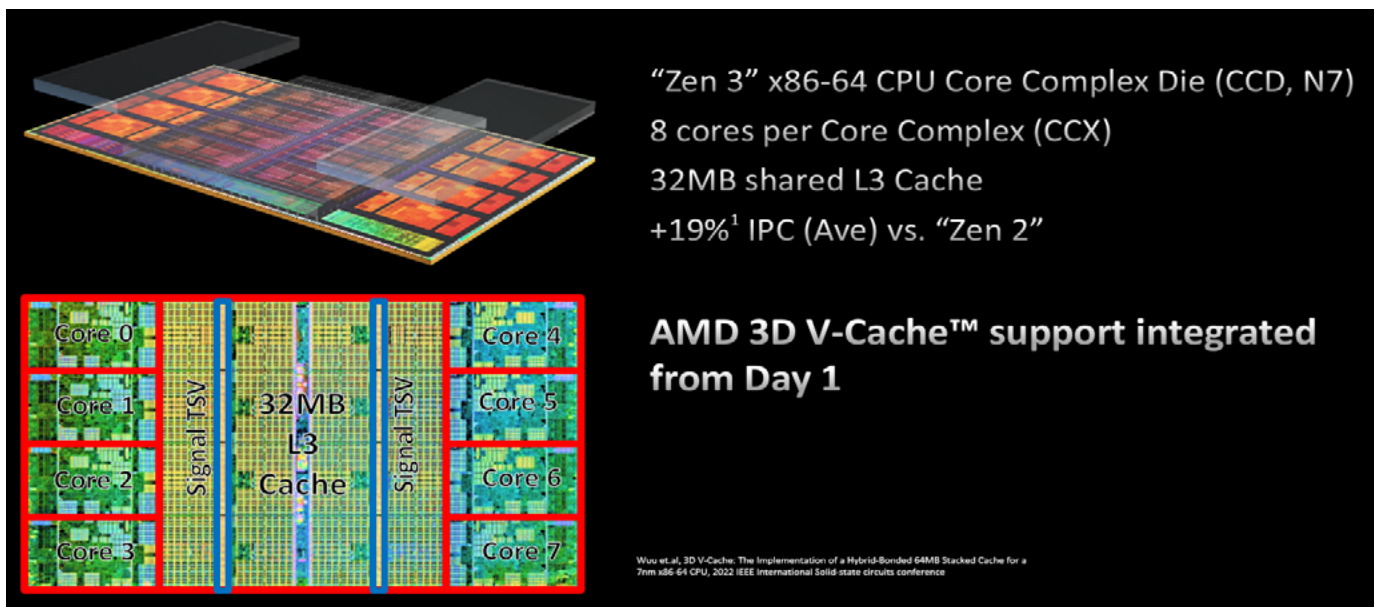


Figure 6: AMD 3D V-Cache™ components: CCD.

because of overlapping the two dies from becoming a limiter.

The final component of the AMD 3D V-Cache™ is the structural die. Two structural dies, which are dummy silicon dies, are placed over the CCD area not covered by the L3D (**Figure 7**). The

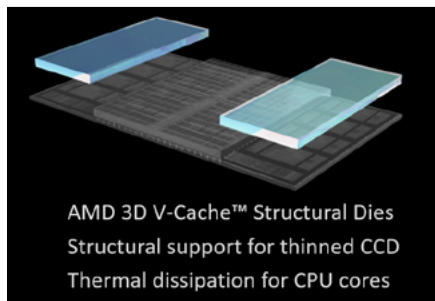


Figure 7: AMD 3D V-Cache™ components: structural die.

structural dies serve two purposes: 1) as the name implies, they provide structural support for the thinned down CCD die; and 2) because silicon is a good thermal conductor, the structural dies are also used for thermal dissipation from the high-frequency, high-power density CPU cores to the heat sinks.

A closer look at the AMD 3D V-Cache™ hybrid bond technology is shown in **Figure 8**. It uses the TSMC SoIC™ process. The image shows the backside of the face-down bottom die, and the face-down top die hybrid-bonded onto the bottom die. The Cu interface between the dies is called bond pad metal (BPM), which connects to the TSV from the bottom die. On the other side of the BPM is the bond pad via (BPV), which is used to connect the

BPM to the Cu metal 13. It is through these TSV, BPM, and BPV structures that power delivery and signals are exchanged between the top and bottom dies. The technology supports a 9µm minimum TSV pitch.

Physically, the CCD is placed face down with C4 interfaces to the substrate. The backside of the CCD is thinned down to reveal the TSVs, which serve as the interconnects to the L3D. The L3D is then also placed face-down and hybrid-bonded to the back of the CCD. Finally, the structural dies are placed on the two sides of the CCD and oxide-bonded to the CCD. Please note, this hybrid bond technology differs from the common 3D approach of connecting the dies through micro-bumps.

Now, we compare the AMD Cu-based 3D architecture versus the current best in class solder-based micro-bump 3D architecture (**Figure 9**). Solder-based micro-

bump technology with tall TSVs is based on traditional solder-based packaging technologies and can scale from 50µm to ~36µm and is acceptable for low-bandwidth applications. AMD 3D chiplet architecture, as shown to scale relative to micro-bump technology, by contrast, uses silicon fab-like manufacturing methods with back-end design rule-based TSVs with Cu-only interconnects without the presence of solder. This is a transformational point in the industry's advanced packaging journey, where interconnect technologies are now being enabled using silicon fab-based techniques to enable extreme bandwidth architectures. As a result of the extreme scaling, we are also able to achieve >3x higher interconnect energy efficiency, >15x higher interconnect density, as well as better signal and power performance compared to micro-bump 3D architectures.

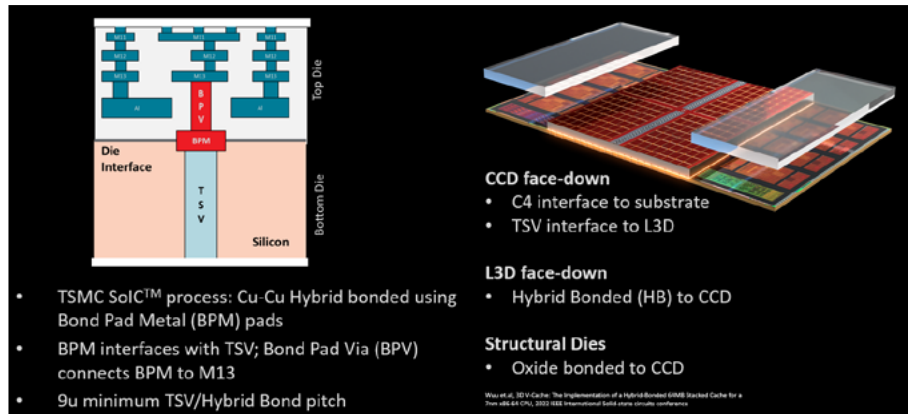


Figure 8: 3D V-Cache™: bringing it together.

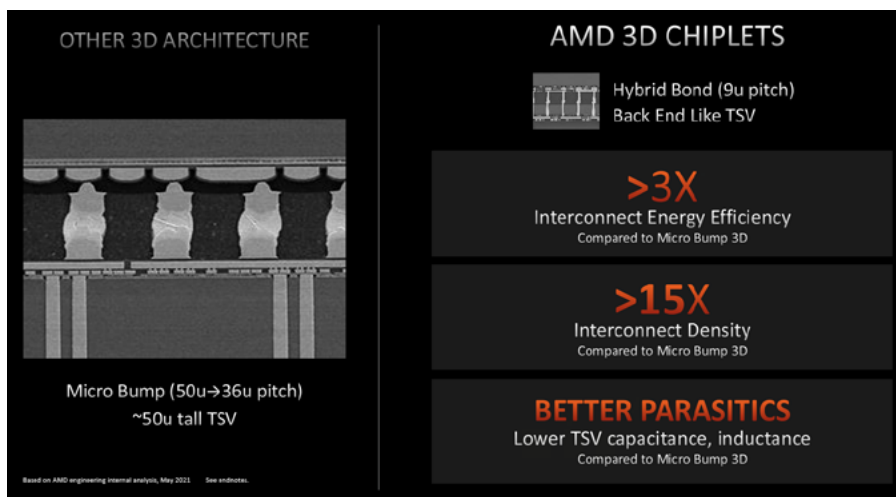


Figure 9: AMD hybrid-bonded 3D chiplet architecture comparison to solder-based 3D architectures.

Regarding “Zen 3” cache hierarchy, each core has a 32KB I-cache and a 32KB D-cache, along with a private 512KB L2 cache. There are eight cores per CCD, and all eight cores share a 32MB L3 cache. The L3 cache is 16-way set associative, with a 32B/cycle interface to each core. DECTED ECC, which can correct double bit errors and detect triple bit errors, is included for enhanced data reliability. When the L3D is bonded on top of the CCD, it expands the 32MB shared L3 cache to 96MB. The 96MB cache continues to be shared between the eight cores, and it continues to be 16-way set associative. It also maintains the L3’s 32B/cycle interface to each core, which provides more than 2TB of total

L3 bandwidth per second. Despite tripling the L3 size, AMD 3D V-Cache™ only adds four cycles of additional latency, which can only be achieved through 3D stacking.

Power delivery was a key architecture focus when we architected AMD V-Cache™. The CCD has three primary power supplies (Figure 11) – there is RVDD in orange, which is the raw, unregulated supply upon which the L3 cache logic runs. Then there is VDD, which each core regulates independently from RVDD. Finally, there is VDDM, which is the supply for the L2 and L3 SRAM bit cells. Of course, there is also VSS, which is shown in grey in the diagram (Figure 11). When the L3D is stacked onto the CCD,

both RVDD and VDDM are delivered to the L3D through power TSVs. To better convey the power delivery RDL, the construction in Figure 11 is flipped upside down with the top L3D die on the bottom. RVDD supplies the logic portion of the L3D die, while VDDM powers the SRAM bit cells. The power TSVs are primarily placed in the channels between the SRAM macros in the CCD.

The SRAM arrays on the L3D die consist of 512 128KB data macros, and 1088 6KB tag and the (LRU) macros located near the signal TSV columns. It is a dual-rail design using VDDM for the SRAM bitcells and RVDD for the peripheral circuits. As added power can negatively impact performance in a power constrained environment, the L3D arrays are optimized not only for high density, but for low power as well. To that end, the SRAM arrays on the L3D uses extensive power reduction features.

3D interface signals are extremely simple flop-to-flop signals that can be enabled only with the use of a hybrid-bonded architecture with its low parasitics. On the transmission side, the signal after leaving the flop is buffered and sent through the TSV to the other die. On the receiving side, the signal first goes through a minimal electrostatic discharge (ESD) circuit to protect against ESD events that can occur during the 3D assembly process. The signal then goes through an isolation circuit, which properly isolates the interface signal that would be floating when the other die is not attached.

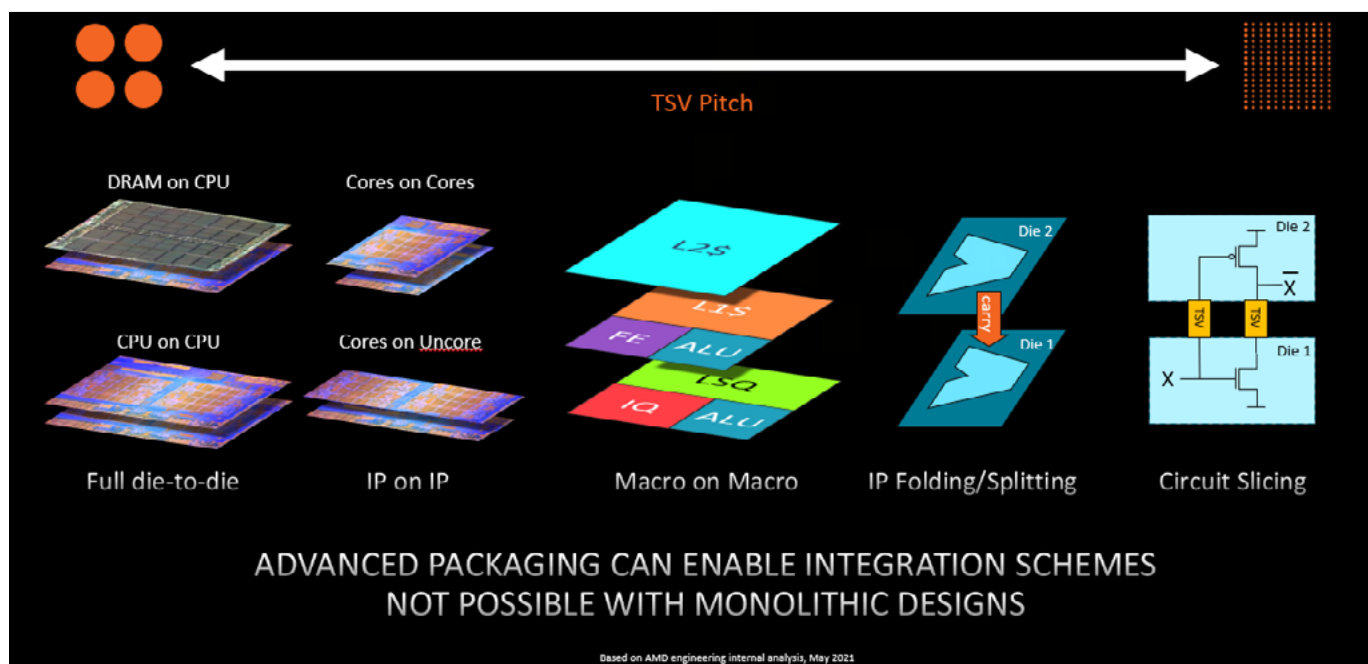


Figure 10: Vision for enabling new architectures with future 3D stacking innovations.

MAXIMIZE YIELDS

for Heterogeneous Integration

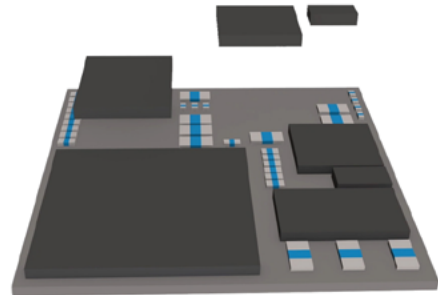
Heterogeneous Integration is enabling higher-bandwidth, lower power consumption, and increased functionality in virtually all of the newest high-tech products – all within a smaller form factor.

However, building the HI modules that power these devices brings a host of new challenges, demanding a comprehensive solution that breaks traditional boundaries for efficient multi-die assembly.

THIN DIE HANDLING

One such challenge is precision die processing for multi-die packages requiring die stacking with a silicon interposer and chiplets.

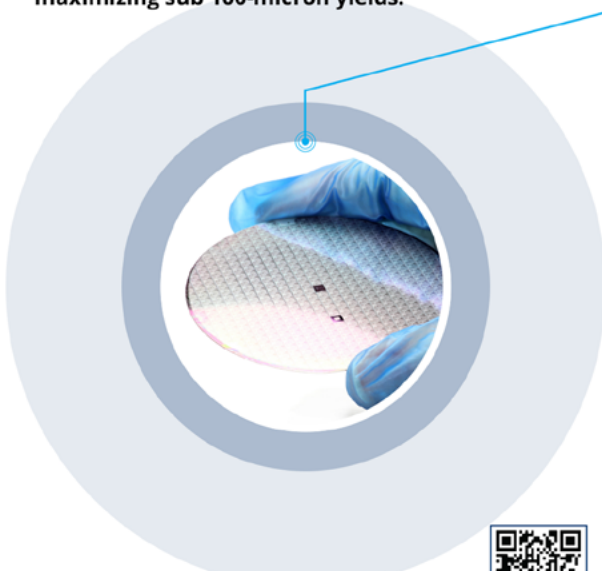
Universal's FuzionSC + HSWF solution has the thin die handling, precision ejection capability and system accuracy to transform this challenge into a significant competitive advantage by **maximizing sub 100-micron yields**.



FUZIONSC & HIGH-SPEED WAFER FEEDER

The High-Speed Wafer Feeder (HSWF) is the world's fastest rapid-exchange multi-die feeder. Combined with Universal's FuzionSC™ Platform, it is the **ultimate multi-die solution for heterogeneous integration**.

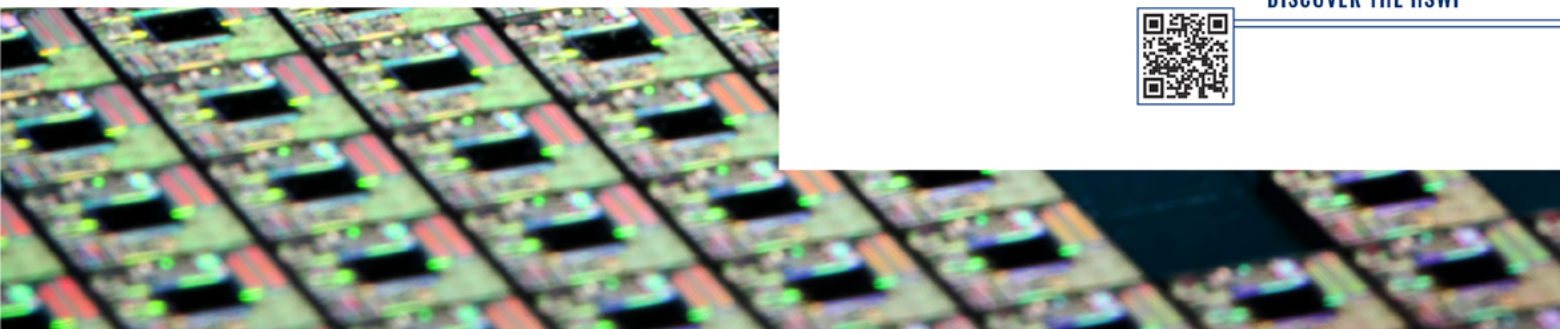
- **WAFER CAPACITY** – to minimize replenishment rate
- **SPEED** – to meet volume requirements
- **MULTIPLE DIE TYPES** – to maximize utilization
- **LARGE SUBSTRATE** – to reduce manufacturing costs
- **THIN DIE** – to maximize sub 100-micron yields



Scan the QR code to be contacted by a Universal Representative and learn more:



DISCOVER THE HSWF

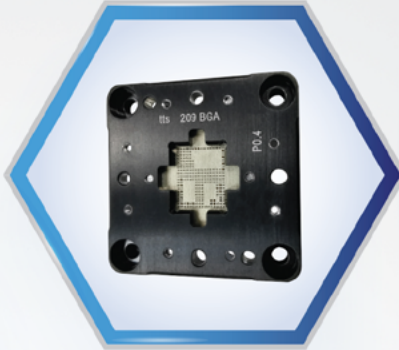
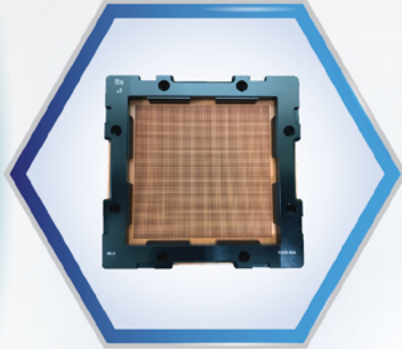




www.tts-grp.com

High Pin Count

Pitch $\geq 0.80\text{mm}$
Pin Count $\geq 10,000$
Coplanarity $< 0.35\text{mm}$

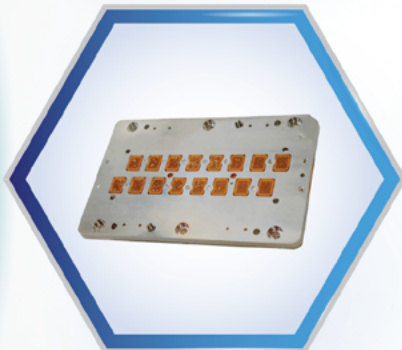


Coaxial

Pitch $\geq 0.40\text{mm}$
Insertion Loss $> 40\text{GHz}@-1\text{dB}$
Crosstalk $> 35\text{GHz}@-52\text{dB}$

WLCSP Probe Head

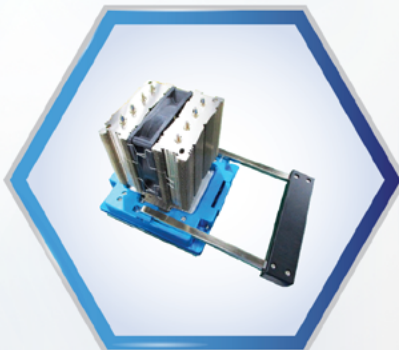
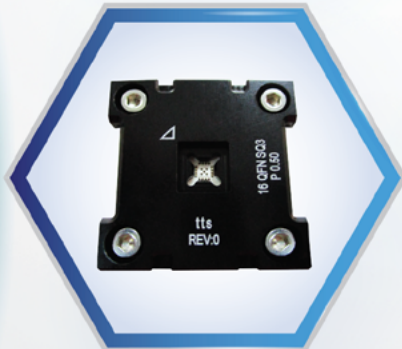
Pitch $\geq 0.15\text{mm}$
Pin Count ≤ 6000
Longevity $> 1000\text{K}$



Probe Pin

Pitch $\geq 0.12\text{mm}$
Power $\leq 6.5\text{A}$
Frequency $> 140\text{GHz}$

RF
Pitch $\geq 0.35\text{mm}$
Insertion Loss $> 60\text{GHz}@-1\text{dB}$
Return Loss $> 30\text{GHz}@-20\text{dB}$



Hand Socket Lid

Heatsink $< 100\text{W}$
Heatpipe $100\text{W} - 1000\text{W}$
Liquid Cooling $300\text{W} - 1500\text{W}$



BROCHURE



Finally, the incoming signal is captured by a flop. What is interesting here is the simplicity and the compactness of the fully-digital IO circuitry, which contributes to the power efficiency and low latency of this hybrid-bonded 3D interface. So how does this translate to performance? In a desktop gaming system, AMD 3D V-Cache™ delivered on average 15% faster gaming performance when compared with its non stacked Ryzen™ counterpart. This 15% is truly a generational leap in performance, which in the past has been enabled only by silicon node transitions.

Milan-X server implementation of the AMD V-Cache™ architecture enables three times the L3 cache compared to standard Milan processors. This additional L3 cache relieves memory bandwidth pressure and reduces latency –that in turn dramatically speeds up application performance.

3D stacking: future and challenges

3D cache stacking over CPU cores is just the beginning of the 3D journey (Figure 10).

The future of 3D stacking is a function of TSV pitch and can spawn many architectural innovations including IP-on-IP stacking, macro-on-macro stacking, IP folding/splitting, as well as circuit-level slicing 3D stacking technology progression. These innovations, along with other advanced packaging techniques, will enable beyond-Moore's-Law scaling this decade and enable complex heterogeneous integration schemes not possible even with monolithic designs.

There are multiple challenges to enable 3D chiplet architectures. All these chiplets need to be tested thoroughly before assembly or we throw away the entire expensive module. Stacking encounters challenges with higher

power densities. This predicament comes along at the same time as Moore's Law is doing less and less for power. Managing and mitigating thermal issues is going to be an interesting and exciting area for innovation, along with power delivery solutions and high current densities across multiple dice means a 3D power grid, among other things. All our tricks of integrated regulators and power gating will need to be deployed to support the power demands of all the layers in the design. Silicon and package are merging with this architecture. Enabling the right design tools that can seamlessly move from system to package to C4 to 3D interface, to truly deliver the best-in-class DTCO, is critical.

Finally, as mentioned at the outset, performance is delivered at the system level, and these heterogeneous modular SoCs will need to be connected with the right software to deliver system-level performance where an increasing amount of differentiation can be delivered.

Summary

We are truly at a new era of computing. Design and innovation must take a step up. The new paradigms will combine traditional CPU compute engines heterogeneously with accelerators, using continuously evolving and improving package technology to enable levels of integration that today are at the board level. In the future, they will be at the integrated modular silicon level. System architectures, previously only in massive supercomputers, are now coming to the masses. It will be an incredibly exciting next era of computing innovation driven by advanced packaging and I look forward to the opportunities ahead!

Endnotes

1. AMD 3D Chiplet Technology -Competition 3D architecture picture from SystemPlus. Intel Core i5-L16G7: the first utilization of Intel's Foveros Technology with Package-on-Package configuration in a consumer product.. <https://www.systemplus.fr/reverse-costing-reports/intel-foveros-3d-packaging-technology/>
2. MLNX-001R: EDA RTL Simulation comparison based on AMD internal testing completed on 9/20/2021 measuring the average time to complete a test case simulation. Comparing: 1x 16C 3rd Gen EPYC CPU with AMD 3D V-Cache Technology versus 1x 16C AMD EPYC™ 73F3 on the same AMD "Daytona" reference platform. Results may vary based on factors including silicon version, hardware and software configuration and driver versions.
3. MLNX-021R: AMD internal testing as of 09/27/2021 on 2x 64C 3rd Gen EPYC with AMD 3D V-Cache (Milan-X) compared to 2x 64C AMD 3rd Gen EPYC 7763 CPUs using cumulative average of each of the following benchmark's maximum test result score: ANSYS® Fluent® 2021.1, ANSYS® CFX® 2021.R2, and Altair Radioss 2021. Results may vary.
4. MLN-075A: Altair™ Radioss™ comparison based on AMD internal testing as of 09/27/2021 measuring the time to run the neon, t10m, and venbatt test case simulations using a server with 2x AMD EPYC 75F3

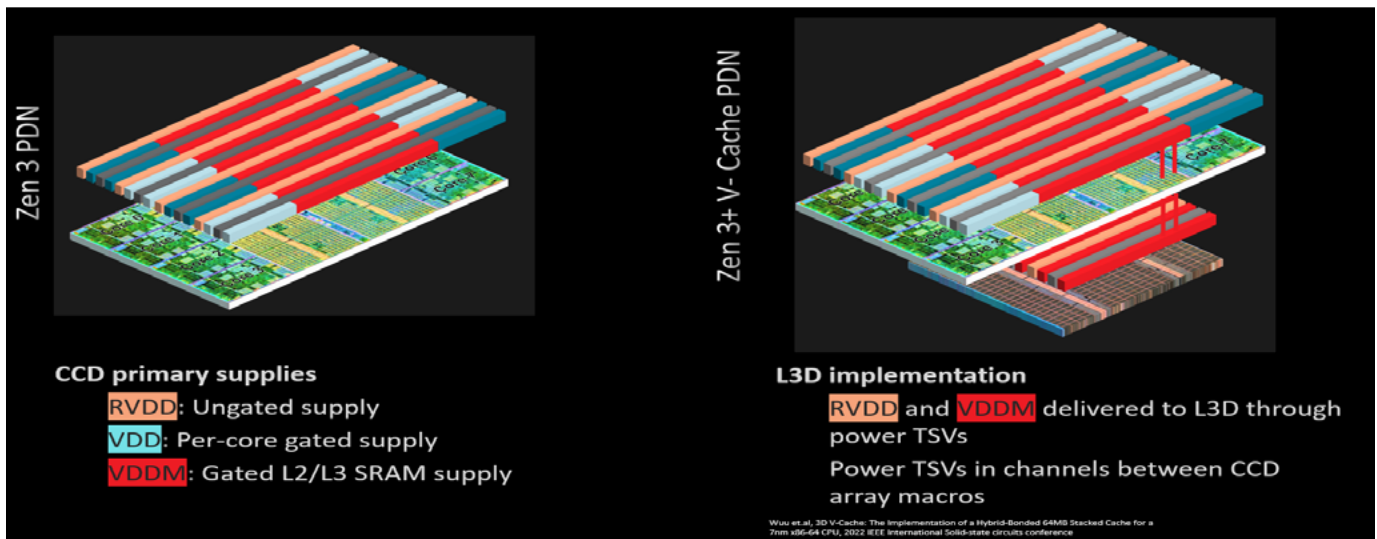


Figure 11: Schematic of 3D AMD V-Cache™ power delivery.

versus 2x Intel Xeon Platinum 8362. Neon crash impact is the max result test case. Results may vary.

5. MLN-080B: ANSYS® CFX® 2021.1 comparison based on AMD internal testing as of 09/27/2021 measuring the average time to run the Release 14.0 test case simulations (converted to jobs/day - higher is better) using a server with 2x AMD EPYC 75F3 utilizing 1TB (16x 64 GB DDR4-3200) versus 2x Intel Xeon Platinum 8380 utilizing 1TB (16x 64 GB DDR4-3200). Results may vary.
6. MLN-130A: ANSYS® Mechanical® 2021 R2 comparison based on AMD internal testing as of 09/27/2021 measuring the average of all Release 2019 R2 test case simulations using a server with 2x AMD EPYC 75F3 versus 2x Intel Xeon Platinum 8380. Steady state thermal analysis of a power supply module 5.3M (cg1) is max result. Results may vary.
7. MI200-01 - World's fastest data center GPU is the AMD Instinct™ MI250X. Calculations conducted by AMD Performance Labs as of Sep 15, 2021, for the AMD Instinct™ MI250X (128GB HBM2e OAM module) accelerator at 1,700 MHz peak boost engine clock resulted in 95.7 TFLOPS peak theoretical double precision (FP64 Matrix), 47.9 TFLOPS peak theoretical double precision (FP64), 95.7 TFLOPS peak theoretical single precision matrix (FP32 Matrix), 47.9 TFLOPS peak theoretical single precision (FP32), 383.0 TFLOPS peak theoretical half precision (FP16), and 383.0 TFLOPS peak theoretical Bfloat16 format precision (BF16) floating-point performance. Calculations conducted by AMD Performance Labs as of Sep 18, 2020 for the AMD Instinct™ MI100 (32GB HBM2 PCIe® card) accelerator at 1,502 MHz peak boost engine clock resulted in 11.54 TFLOPS peak theoretical double precision (FP64), 46.1 TFLOPS peak theoretical single precision

matrix (FP32), 23.1 TFLOPS peak theoretical single precision (FP32), 184.6 TFLOPS peak theoretical half precision (FP16) floating-point performance. Published results on the NVidia Ampere A100 (80GB) GPU accelerator, boost engine clock of 1410 MHz, resulted in 19.5 TFLOPS peak double precision tensor cores (FP64 Tensor Core), 9.7 TFLOPS peak double precision (FP64), 19.5 TFLOPS peak single precision (FP32), 78 TFLOPS peak half precision (FP16), 312 TFLOPS peak half precision (FP16 Tensor Flow), 39 TFLOPS peak Bfloat16 (BF16), 312 TFLOPS peak Bfloat16 format precision (BF16 Tensor Flow), theoretical floating-point performance. The TF32 data format is not IEEE compliant and not included in this comparison. <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/nvidia-ampere-architecture-whitepaper.pdf>, page 15, Table 1.

8. MI200-02 - Calculations conducted by AMD Performance Labs as of Sep 15, 2021, for the AMD Instinct™ MI250X accelerator (128GB HBM2e OAM module) at 1,700 MHz peak boost engine clock resulted in 95.7 TFLOPS peak double precision matrix (FP64 Matrix) theoretical, floating-point performance. Published results on the NVidia Ampere A100 (80GB) GPU accelerator resulted in 19.5 TFLOPS peak double precision (FP64 Tensor Core) theoretical, floating-point performance. Results found at: <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/nvidia-ampere-architecture-whitepaper.pdf>, page 15, Table 1.
9. MI200-07 - Calculations conducted by AMD Performance Labs as of Sep 21, 2021, for the AMD Instinct™ MI250X and MI250 (128GB HBM2e) OAM accelerators designed with AMD CDNA™ 2 6nm FinFet process technology at 1,600 MHz peak memory clock resulted in 3.2768 TFLOPS peak

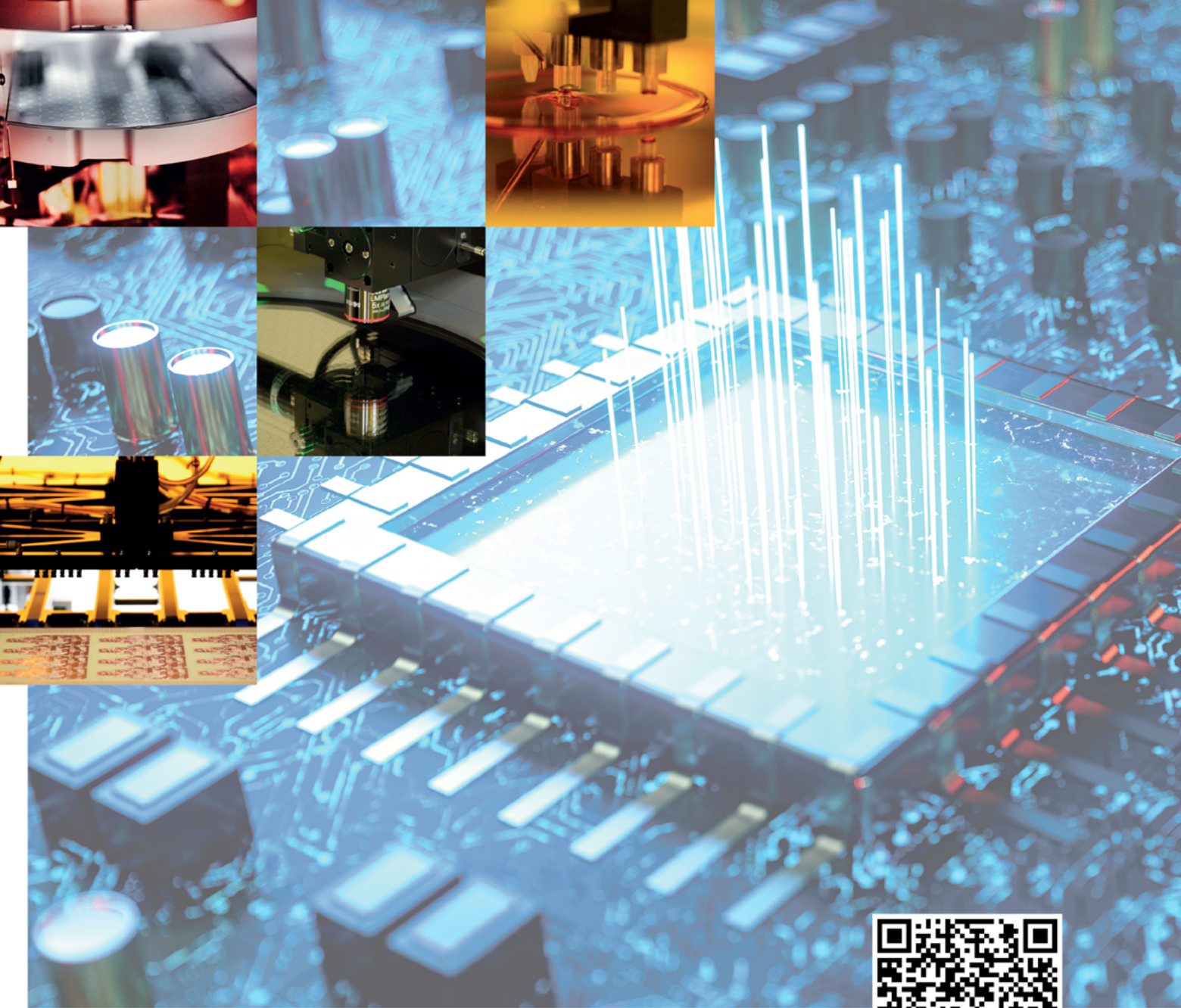
theoretical memory bandwidth performance. MI250/MI250X memory bus interface is 4,096 bits times 2 die and memory data rate is 3.20 Gbps for total memory bandwidth of 3.2768 TB/s ((3.20 Gbps*(4,096 bits*2))/8). The highest published results on the NVidia Ampere A100 (80GB) SXM GPU accelerator resulted in 2.039 TB/s GPU memory bandwidth performance. <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/nvidia-a100-datasheet-us-nvidia-1758950-r4-web.pdf>

10. MI200-15A - Testing Conducted by AMD performance lab as of 10/7/2021, on a single socket Optimized AMD EPYC™ CPU server, with 4x AMD Instinct™ MI250X OAM (128 GB HBM2e) 560W GPUs with AMD Infinity Fabric™ technology, using LAMMPS ReaxFF/C, patch_2Jul2021 plus AMD optimizations to LAMMPS and Kokkos that are not yet available upstream resulted in a median score of 4x MI250X = 19,482,180.48 ATOM-Time Steps/s Vs. Dual AMD EPYC 7742@2.25GHz CPUs with 4x NVIDIA A100 SXM 80GB (400W) using LAMMPS classical molecular dynamics package ReaxFF/C, patch_10Feb2021 resulted in a published score of 8,850,000 (8.85E+06) ATOM-Time Steps/s. <https://developer.nvidia.com/hpc-application-performance> 19,482,180.48/8,850,000=2.20x (220%) the/1.2x (120%) faster. Container details found at: <https://ngc.nvidia.com/catalog/containers/hpc:lammips> Information on LAMMPS: <https://www.lammips.org/index.html> Server manufacturers may vary configurations, yielding different results. Performance may vary based on use of latest drivers and optimizations.



Biography

Raja Swaminathan is a Senior Fellow & Advanced Packaging Leader at AMD, Austin, Texas, USA. Prior to AMD, he was at Apple, architecting and developing the packaging technologies for the M1x series of processors and Principal Engineer, Silicon Package Architecture at Intel. He holds 35 patents on semiconductor packaging technologies. He received his BS in metallurgy from Indian Institute of Technology, Madras, India, and a PhD in Materials Science from Carnegie Mellon U. Email: raja.swaminathan@amd.com



SUSS EQUIPMENT SOLUTIONS ENABLING LEADING-EDGE PACKAGING TECHNOLOGIES

SUSS MicroTec is a leading supplier of equipment and process solutions for the semiconductor industry enabling state-of-the-art packaging technologies from R&D to high volume production.

With next generation lithography solutions for FOWLP and leading edge wafer bonding technologies for 2.5D and 3D stacking, we contribute to the advancement of innovations in the key areas of packaging and system integration.

Always be one step ahead — with solutions from SUSS MicroTec.
Contact us for more information!

SUSS MicroTec
info@suss.com
www.suss.com

SUSS MicroTec
+

5G

Beyond Connectivity Galileo Test Socket

- Extremely short electrical paths to deliver excellent signal integrity
- Low inductance, high bandwidth >40GHz
- High Current Carrying Capacity
- Very high thermal conductivity
- Field-replaceable contact set with minimal tooling and technical expertise required
- 3D printing technology for quick-turn socket body manufacture
- Pre-stocked elastomer sheets enabling very short lead-times
- Solder-down performance for applications where signal integrity loss is not an option



Galileo is an innovative, low-profile test socket engineered to support today's high performance Digital and RF applications. It leverages proven interposer elastomer technology and advanced 3D printing manufacturing to provide a high-performance solution for BGA, LGA, QFP, SOIC, or QFN packaged devices with extremely short lead times.



COMMUNICATION

COMPUTER

CONSUMER

AUTOMOTIVE

Semiconductor test: staying ahead of nanodevices

By Tucker Davis, Brian Brecht [Teradyne]

In the semiconductor fabrication process, engineers continue to innovate, enabling smaller transistors and higher density circuits. The transition to FinFETs allowed 7nm and 5nm processes to realize circuits of amazing density, and the progress of 3nm and gate-all-around (GAA) transistors provides a clear path for future advancement of digital circuit cost reduction and performance improvement.

As higher transistor counts lead to devices that are larger and more complex, there is increased pressure to achieve better test throughput and yields to maintain manufacturing cost efficiency. Here, we will explore how innovations to the signal and power delivery architectures will allow semiconductor manufacturers to achieve better throughput and yield to manage overall costs.

Processing demand: no sign of slowing

The past two years have witnessed unprecedented growth in the semiconductor industry driven by advances in artificial intelligence (AI), natural language processing, automated vehicles, and augmented/virtual reality. All of these applications require enormous computational processing and communications bandwidth to make sense of the proliferation of sensing and real-world interfaces, which depend heavily on advancements in semiconductors. We've witnessed the results as trends in processors move to the forefront of semiconductor processing, adding new AI cores and increasing quality standards at very high volume while controlling costs.

Figure 1 [1] illustrates the historical growth of transistors per microprocessor, demonstrating that the pace of processing demand, as expressed by device complexity, has continued on an exponential growth path for 50 years. We see no signs of the demand for more processing power slowing. With transistor counts reaching the 100s of billions per die, we must

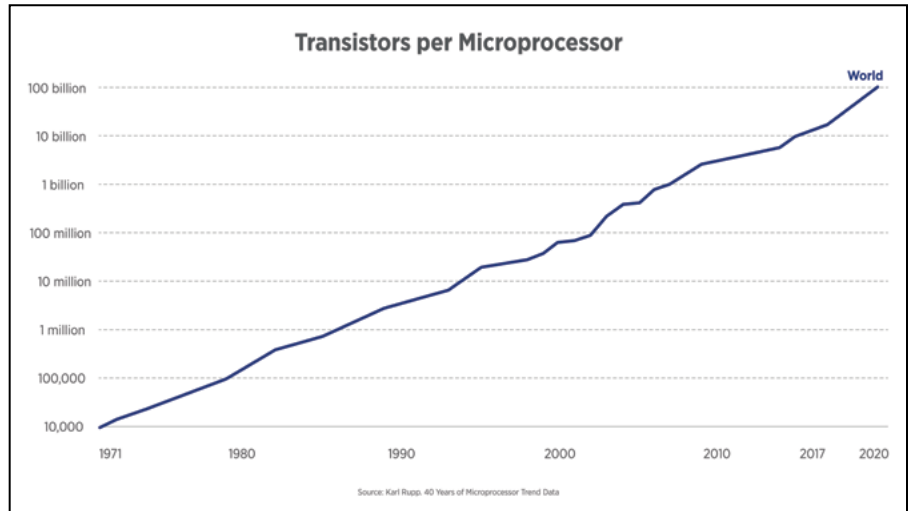


Figure 1: 42 years of microprocessor trend data. SOURCE: [1]

turn our focus to how these devices will be designed and tested to keep up with this growth in complexity.

From FinFET to GAA

Looking forward, GAA is shaping up to be the enabler for transistor count to continue increasing, ensuring complexity continues its exponential growth and with it, an expansion of test requirements. The new nanowire or nanosheet structures forming GAA allow more transistors in each device, but also bring new defects, in addition to complexity. At a high level, GAA boosts transistor density, which

increases test vectors and test times, and creates a higher need for repair and trim. This drives new test challenges in terms of signal delivery to the device under test (DUT), which are summarized in Table 1. This new set of challenges will be disruptive to many traditional test strategies, but a new generation of automatic test equipment (ATE) systems are available to meet these requirements.

Modern automatic test equipment has extremely dense instruments to measure and provide thousands of high-performance signals and power supplies to a wide variety of devices

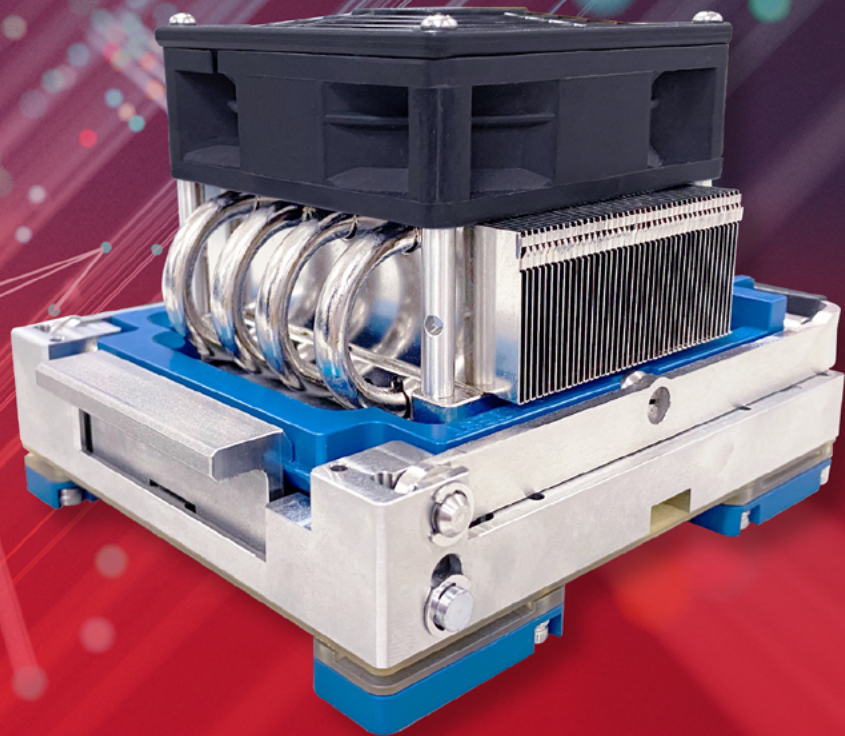
Device Characteristics	Signal Delivery Challenges
More transistors per mm ²	Higher signal count, more power delivered to the device
New defect mechanisms with lower initial yields	Defect-free interfaces that don't "pollute" the device yield signature
Smaller geometries with lower power supply and gate voltages	Low noise and accurate power delivery
More transistor switching during device operation	Power delivery capable of handling high inrush currents
Smaller leakage currents	Lower leakage signal delivery

Table 1: Summary of test challenges for GAA devices.



WinWay Technology

Your trusted partner in IC testing



Burn-in Socket

- Innovative lid design.
- Individual temperature control design.
- Supporting high power chips of burn-in testing.

Performance	Heating rate	Precision of temperature
850W (<125°C)	< 150 sec (25°C to 125°C)	±1°C

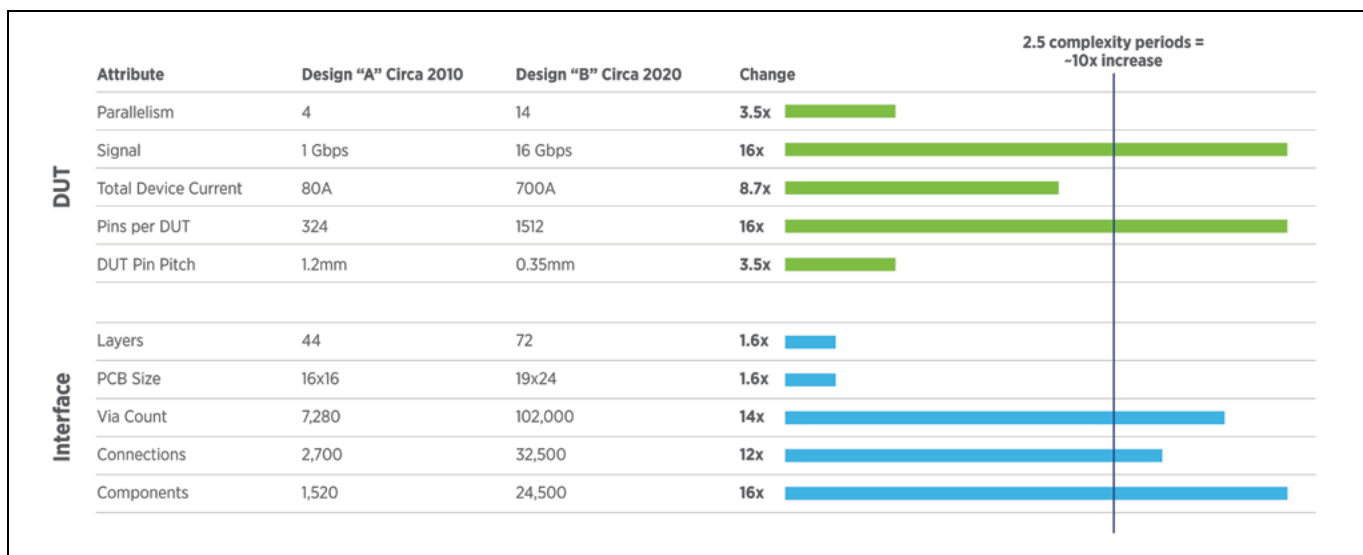


Figure 2: Complexity trend for two examples of 2X4 scaling. SOURCE: Teradyne

under test. A typical, multiple-site, digital ATE configuration has thousands of digital channels, thousands of amps of power, and tens of thousands of interconnections. To house this many signals, measurement units, sources, and supporting electronics, along with cooling to maintain constant temperature on integrated circuits (ICs) after calibration, the tester volume is on the order of a cubic meter. The DUT size, packaged or wafer, is on the order of square millimeters. There is a significant architectural challenge to “funnel” the thousands of signals and power supplies from an area in meters to millimeters, while maintaining full performance, and making the DUT interface board or probe card producible.

By looking back through the past dozen years, interface board complexity typically doubles every four years. The doubling of complexity is seen across many different attributes: site count increases, via count, pin pitch reduction, application circuitry, signal speeds, power distribution network impedances, and more. The modern-day interface board is as, or more, complex than the instrumentation in the ATE. The ATE architecture needs to be able to keep up with these increasing DUT demands, while providing a path to fast-turn, acceptable and predictable yield, highly-reliable, interface board builds.

The complexity trend has been based on two times the performance and two times the pins, every 4 years, or what we call “2x4 scaling,” as seen in **Figure 2**,

which compares two design examples from 2010 and 2020. A ten-year span of time is 2.5 complexity periods and therefore, the expectation is the attributes will be 10x or more difficult. Some attributes are significantly higher than 10x and others below, but taken in the aggregate, the design differences demonstrate this “2x4 scaling” trend. As we project out in time, the next 10 years will likely exhibit a similar increase in attribute difficulty. The key question is: how do we recognize and interpret device trends to engineer ATE and interface solutions that meet those increasingly difficult requirements?

Translating transistor technology to interface requirements

Transistor scaling has resulted in a consistent decrease in power rail voltage from 5V at the 0.35µm planar transistor node (circa 1995) down to 0.8V or lower with the recent N7 (7nm) FinFET transistor node. This translates to an 8% reduction per year in power rail voltage. The projections are the power rail voltage will continue to decrease with the GAA transistor architecture and by 2030, typical power rails in high-performance digital applications will be 0.5V or lower.

The reduction in power rail voltage influences the power integrity of the device, which is the combination of the ATE power supply and interface board power impedance. Even though the power rail voltage is decreasing, the transistor density is increasing, which results in the total power rail current staying flat or

slightly increasing for a given application. This means power rail impedance must decrease at the same rate or more than the decrease in power rail voltage to avoid device “brown out” events under heavy parallel scan loads, or even worse, power rail voltage spikes that could damage the device because of transistor over-stress.

Take a basic case of a 0.8V power rail with a total rail current of 50A peak during the parallel scan execution. To achieve a target of 10% droop and kick response requires a power impedance of no greater than 1.6mΩ across the operating frequency range. To maintain the same 10% droop and kick response at 0.5V power rail voltage with the same 50A peak current requires a power impedance of 1mΩ—a 38% reduction. If the current increases to 80A as more transistors are packed in the device, consistent with transistor scaling trend to date, the power impedance must be 0.63mΩ—a 61% reduction.

If the impedance reduction is not met, the consequences could be significant. In one scenario, the way increased impedance will manifest itself is the scan test results will become unstable from device to device or lot to lot, thereby impacting yield. One strategy to compensate and achieve stable scan results and yield is to reduce the scan clock rate or reduce the number of parallel scan execution blocks to reduce the peak current load. The effect of this strategy is longer test times, however, given the high cost of lost yield, this is preferable to throwing away good die.

Supporting DUT trends with ATE architecture

The ATE and device interface board (DIB) architectures are critical to the ability to sufficiently test current and future devices. The following are examples of architectures that enable the DUT trends discussed in this article.

Cleaner paths from instruments to DUTs and to application circuitry.

A traditional interface board has instruments connecting to extremely dense “clusters” of function types (power, digital, analog, utility signals, etc.), as illustrated in **Figure 3a**. As site and channel counts increase, a large number of signals must cross over each other. The crossovers increase layer counts and signal losses, make site to site test correlation difficult, and add time to route the board. By rotating the orientation of the instruments 90 degrees relative to the DUT, the signals are organized in “strips,” instead of “clusters,” as seen in **Figure 3b**. Consequently, the routing from instrument to the DUT area is significantly improved into clean routing channels. This results in improved route utilization per layer and the ability to implement a clean “site copy” approach to DUT layout, which then aids in site-to-site matching and device correlation.

Higher-performance instruments and delivery paths to the DIBs.

Another DUT requirement the ATE must address is the need for extremely fast responding voltage sources. These voltage sources have loop responses reaching into the Megahertz frequency range. By reaching this speed, there is no need to use space on the DIB for bulk capacitance in the multiple farads using tantalum or aluminum electrolytic capacitors to maintain acceptable droop, kick, and settling time at the DUT. This space savings can be used for more higher-frequency ceramic caps, or more

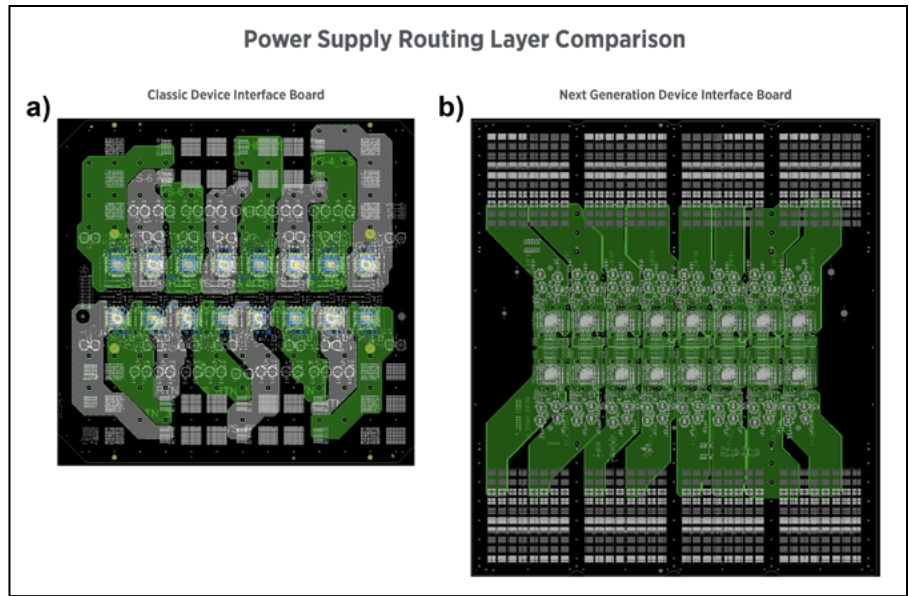


Figure 3: Power supply routing layer comparison for: a) (left) a classic DIB; and b) (right) a next-generation DIB. SOURCE: Teradyne

DUTs and application circuitry. Another benefit of the reduced total capacitance is less damage to probes because of the possibility of the high energy storage of the capacitor being released rapidly, causing needle or socket pin burn. **Table 2** compares traditional solutions with better solutions for satisfying stable DUT voltages across a wide frequency range of transient currents.

The elimination of the large-value bulk capacitor packages also reduces the chance of resonances in the frequency response of the power network when combined with the high-frequency capacitors. Resonance can happen if there is a significant gap in the self-resonant frequencies of different capacitor banks on the board.

Capacitors and their connections to the board can be simplified into a series model of capacitance (C), equivalent series resistance (R), and equivalent series inductance (L). Impedance (Z) of an ideal capacitor is $Z_c = 1/(j\omega C)$, where ω is $2\pi \times$ frequency. Adding in

the series equivalent R and L, adds the terms, $Z_l = j\omega L$ and $Z_r = R$. Z total is these three terms in series. $Z = R + 1/(j\omega C) + (j\omega L)$. Solving for magnitude, $|Z| = \text{square root } \{R^2 + [\omega L - 1/(\omega C)]^2\}$. At the frequency point where $(\omega L) = 1/(\omega C)$, the reactive terms cancel out, and all that remains is the R term. Typically, R is small, which sets a low impedance resonant point at that frequency.

Having a very large value capacitor in parallel with small value capacitors can set multiple resonant points as shown in **Figure 4**. Operating near these resonant points can have adverse or inconsistent results on power performance and test results. Reducing to fewer capacitor types can eliminate resonant frequency points. Another strategy is to add many different capacitor values, with varying self-resonant frequencies, to “fill in” the gap between the two resonance points. This method often requires power integrity simulation to properly select capacitor values and quantities.

Frequency	Traditional Solution	Better Solution
DC to 100KHz	Typical ATE power supply	Faster and more accurate power supply
10KHz to 1MHz	Bulk capacitance on DIB	Faster and more accurate power supply
500KHz to 10MHz	More ceramic caps on DIB	Ceramic caps to solve ESR/ESL spike
1MHz to 30MHz	More caps in DUT package or probe hardware	Get lowest ESL caps available and use as many as will fit

Table 2: Comparison of solutions for satisfying stable DUT voltages across a wide frequency range of transient currents.



Enhancing Surfaces and Materials for the World's Technology Leaders

Production-Proven HVM Systems for Wafers, Panels, and Advanced Substrates

Thermal

- Low-temp curing
- High-vacuum curing
- Degas and dehydration
- Fluxless solder reflow
- UV and thermal curing
- Hybrid bonding

Coating

- Monolayer coating
- Monolayer coating with in-situ plasma cleaning
- LPCVD coating

Wet

- Plating (electrolytic and e-less)
- Clean and surface prep
- Desmear
- EBR

VertaCure™ XP

High-volume automated system for polyimide vacuum cure



Yield Engineering Systems, Inc.

Call: **1-510-954-6889** (worldwide) or **1-888-YES-3637** (US toll free)

www.yieldengineering.com

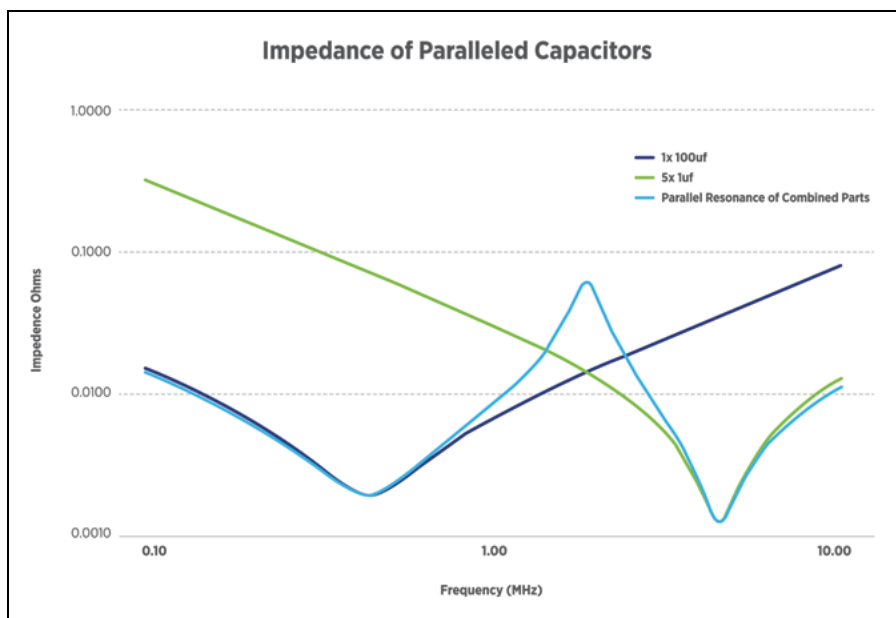


Figure 4: Impedance of capacitors in parallel vs. transient current frequency. SOURCE: Teradyne

Another trend in DUT power supplies is smaller resolution of amps per channels, with the ability to merge a higher number of channels together to reach high peak currents. This gives flexibility for test strategies to test different sections of the DUT independently, or merge into larger combined rails. The ability to measure and control power in subsections of a DUT provides added test control and data collection.

Typically, ATE power supplies are merged using a power plane on the device interface. An alternative to this technique is to split the power rail into more, narrower paths compared to a fewer, wider paths, but with several significant tradeoffs and consequences. The narrow paths require more space for power rail to power rail separation, as well as added room for more sense line routings. It is more difficult to connect bypass capacitors into multiple small planes compared to fewer larger planes. The narrow paths typically result in higher inductance, which leads to increased power rail impedance, and worse overall droop and recovery performance. Each power rail could utilize individual wide traces, then connect together inside the DUT along with their sense lines. Unfortunately, the added complexity of implementing this method often leads to more probe damage, not less.

Merging supplies on the device interface is accomplished by connecting voltage sources that are typically set up to equally share current between all the channels in a group. If the supplies were to be routed to the DUT individually and one of the channels has high-resistance probes, it will still force the same current as the other channels that have low-resistance probes. The power dissipated in the high-resistance probes will eventually fail. In the case of a combined plane with the power rail, the current balances between the low- and high-resistance probes thereby preventing large power dissipation on an individual pin.

As an example of the discussion at the end of the previous paragraph, assume 10 probe pins, each nominally at 50mΩ, and the DUT requires 1V at 5A, for 0.5A per probe. In this case, the power dissipated per probe is $I^2 * R = .05^2 * .5 = 12.5mW$. If one of the ten pins changes state from 50mΩ to 1Ω (e.g., because of dirty needle tips), then in the individual wide trace per VS case, the current per pin is still the same due to current sharing, and that pin goes to $I^2 * R = .05^2 * 1 = 250mW$, compared to the 12.5mW design target. This is a 1900% increase in power dissipation, which will likely damage pins. It is impractical to derate the nominal pin power rating to handle this increase.

Whereas, in the common plane for the VS rail, the voltage on the interface and the DUT pad remains the same for each pin. The current is redirected to the 9 pins at 50mΩ resulting in a current increase from 0.5A/pin to 0.55A/pin, which is 15mW compared to the original design target of 12.5mW. A standard design derating should allow for pins to safely handle the 20% increase in power dissipation, or more, given careful planning and analysis. The best practice is to derate the power dissipation per pin to not exceed 50% of the design target. In this example, it takes more than 3 of the 10 pins with high resistances to exceed the derating target.

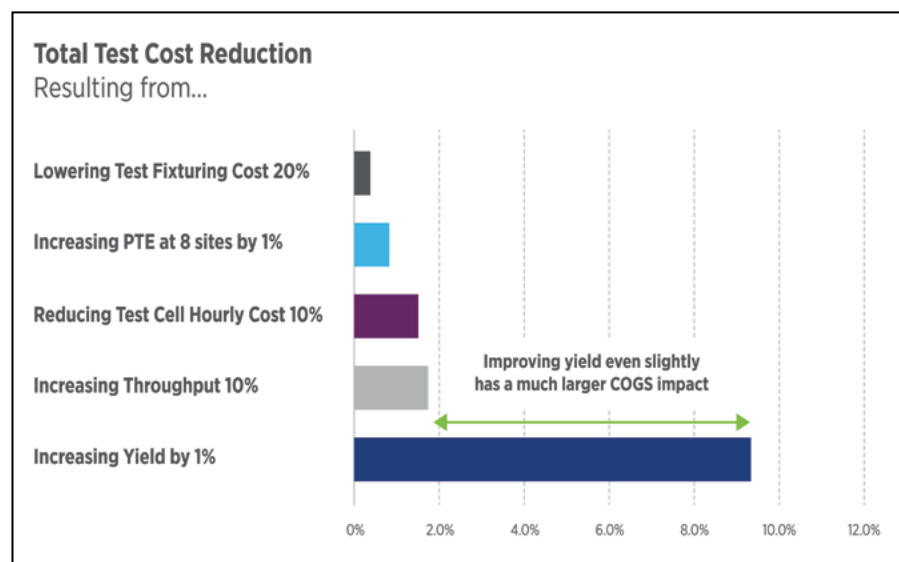


Figure 5: Total test cost reduction summary. SOURCE: Teradyne

Better performance/improved architecture: best yield

The goal of the signal delivery improvements described above is optimization of the overall cost function. Traditionally, semiconductor test has been viewed as an added “cost of quality.” Many hours are devoted to highly optimizing the test process to reduce the cost of test, either by increasing throughput or lowering overall capital costs for the test cell and fixtures. The key is focusing test on optimizing yield through a balanced strategy of data analytics and yield learning, higher performing test solutions with improved accuracy and wider guard bands, or using test for trim and repair.

As seen in **Figure 5**, a breakout of a typical device, even on a relatively high-yielding device, a minor improvement in yield (in this case 1%), produces almost 10x the impact of the typical test cost optimizations, like reducing the hourly test cell cost, improving the overall test throughput by shortening the test time, or increasing parallel test efficiency.

As the semiconductor industry marches forward on the path of technological advances, the economics of the manufacturing process must adapt to the higher costs of the silicon used to implement it. The path to achieve the best possible yields is enabled by the best possible ATE architectures integrated with the best possible interface performance.

Reference

1. K. Rupp, “40 Years of Microprocessor Trend Data,” Retrieved at: Our World in Data; <https://ourworldindata.org/grapher/transistors-per-microprocessor>

DDR GHz Sockets

Industry's Smallest Footprint

- Up to 500,000 insertions
- Bandwidth to 94 GHz
- 2.5mm per side larger than IC
- Ball Count over 5500, Body Size 2 - 100mm
- <25 mΩ Contact Resistance throughout life
- Five different contactor options
- Optional heatsinking to 800W
- Six different Lid Options



Ironwood
ELECTRONICS 1-800-404-0204
www.ironwoodelectronics.com



Biographies

Tucker Davis is the Product Manager for UltraFLEX^{plus}, at Teradyne, North Reading, MA. He is focused on complex digital devices, such as mobile application processors. Prior to joining Teradyne, he was an Applications Engineer and Product Manager for National Instruments. Tucker holds a Master's degree in Electrical Engineering from the University of Oklahoma, where he graduated summa cum laude. Email: tucker.davis@teradyne.com

Brian Brecht is an Engineering Manager at Teradyne, Agoura Hills, CA. With more than 25 years of experience in the design of automated test equipment instruments and platforms. He holds a number of patents and has led design teams around the world in the delivery of high-performance device interface boards. In his current role, he is focused on advanced interface structures for testing next-generation devices.

Status and outlook for fan-out wafer/panel-level packaging

By John H. Lau [Unimicron Technology Corporation]

The biggest difference between fan-out technology and flip-chip technology is that fan out needs to fabricate the redistribution layers (RDLs), but flip chip uses the substrate with RDLs. There are at least three different formations of fan-out RDLs, namely: a) chip first with die face down; b) chip first with die face up; and c) chip last (or RDL first). In this brief article, recent advances in fan-out are presented, such as: 1) RDL formations; 2) heterogeneous integration of the baseband chip and antenna-in-package (AiP); and 3) heterogeneous integration of photonic integrated circuits (PIC) and electronic integrated circuits (EIC). Some recommendations are also provided.

Chip first with die face down

Figure 1 shows an example of heterogeneous integration of four chips and four capacitors using chip-first with die face-down fan-out packaging [1]. The package size is 10mm x 10mm, which consists of one 5mm x 5mm chip, three 3mm x 3mm chips, and four 0402-capacitors. The process flow is very simple. First, the chips are picked up and then placed face down on a temporary carrier with a double-sided thermal release tape. Then, the carrier and the chips are molded with epoxy molding compound (EMC) using the compression method and then post-mold cured (PMC) before removing the carrier and the double-sided tape. Next comes building the RDLs from the Al or Cu pads on the chips. Finally, solder balls are mounted and the whole reconstituted carrier (with chips, EMC, RDLs, and solder balls) is diced into individual packages as shown Figure 1.

There are two RDLs in each package. Each RDL consists of the photosensitive polyimide dielectric layer and the Cu conductor layer. Because an under bump metal (UBM)-

less pad has been used for the solder ball, the Cu conductor layer of RDL2 is thicker than that of RDL1. This is because of the Cu consumption due to solder reflow and during operation. For detailed information on the design, materials, process, fabrication, and reliability of the PCB assembly of the heterogeneous integration package, please see [1,2].

Figure 2 shows an example of heterogeneous integration of mini-light-emitting diodes (LEDs) for an RGB display using chip-first with die face-down fan-out packaging [3]. The mini-

LEDs are red (R) ($125 \times 250 \times 100\mu\text{m}$), green (G) ($130 \times 270 \times 100\mu\text{m}$), and blue (B) ($130 \times 270 \times 100\mu\text{m}$). The spacing among the RGB mini-LEDs is $80\mu\text{m}$, the pixel-to-pixel spacing is also $\sim 80\mu\text{m}$, and the pixel pitch is $625\mu\text{m}$. There are two RDLs in each package. A printed circuit board (PCB) ($132\text{mm} \times 77\text{mm}$) is designed and fabricated for the drop testing that is done on the mini-LED package. Thermal cycling of the mini-LED surface mount device (SMD) PCB assembly is also performed by a nonlinear temperature- and time-dependent finite-element simulation [3].

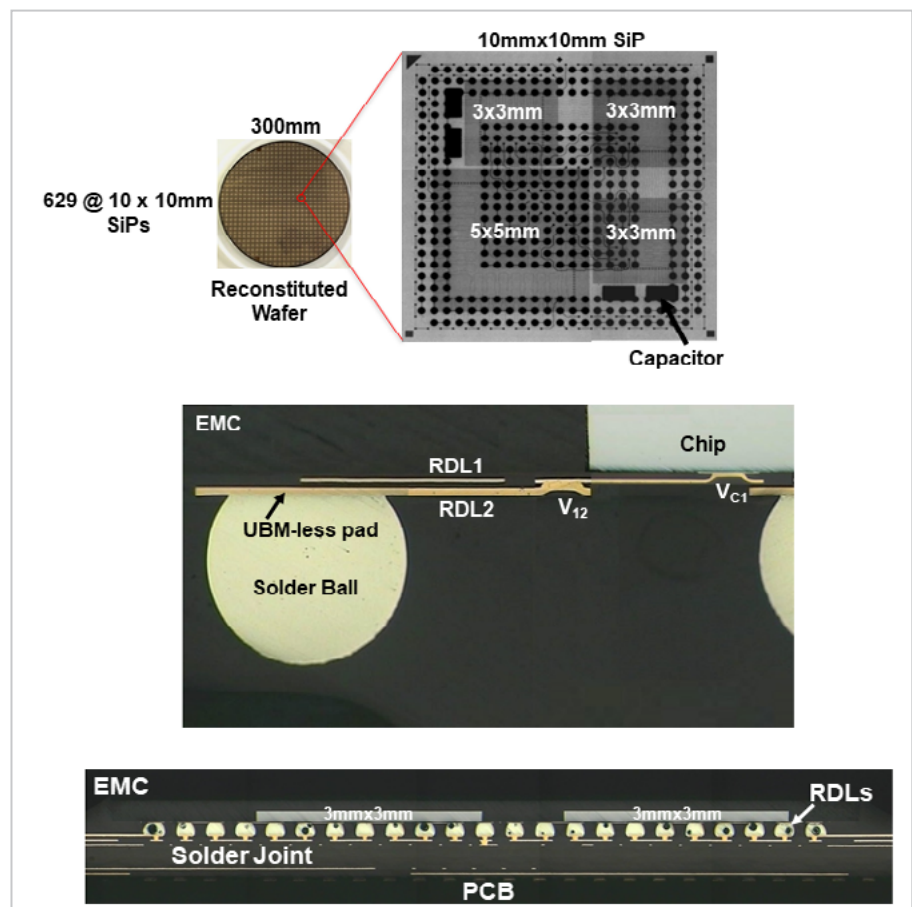


Figure 1: Heterogeneous integration of four chips and four capacitors (chip first die face down with a temporary wafer process).

XPERI IP BUSINESS IS NOW ADEIA

[AHH-dee - uh]

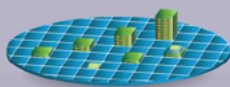
Adeia turns ideas into innovations

Our name may be new, but our roots run deep with decades of continued innovation. We invent, develop and license innovations that advance how we live, work and play.

Adeia invented and pioneered Direct and Hybrid Bonding

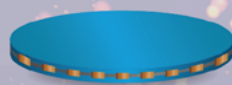
DBI® Ultra

Die-to-Wafer Hybrid Bonding



DBI®

Wafer-to-Wafer Hybrid Bonding



ZiBond®

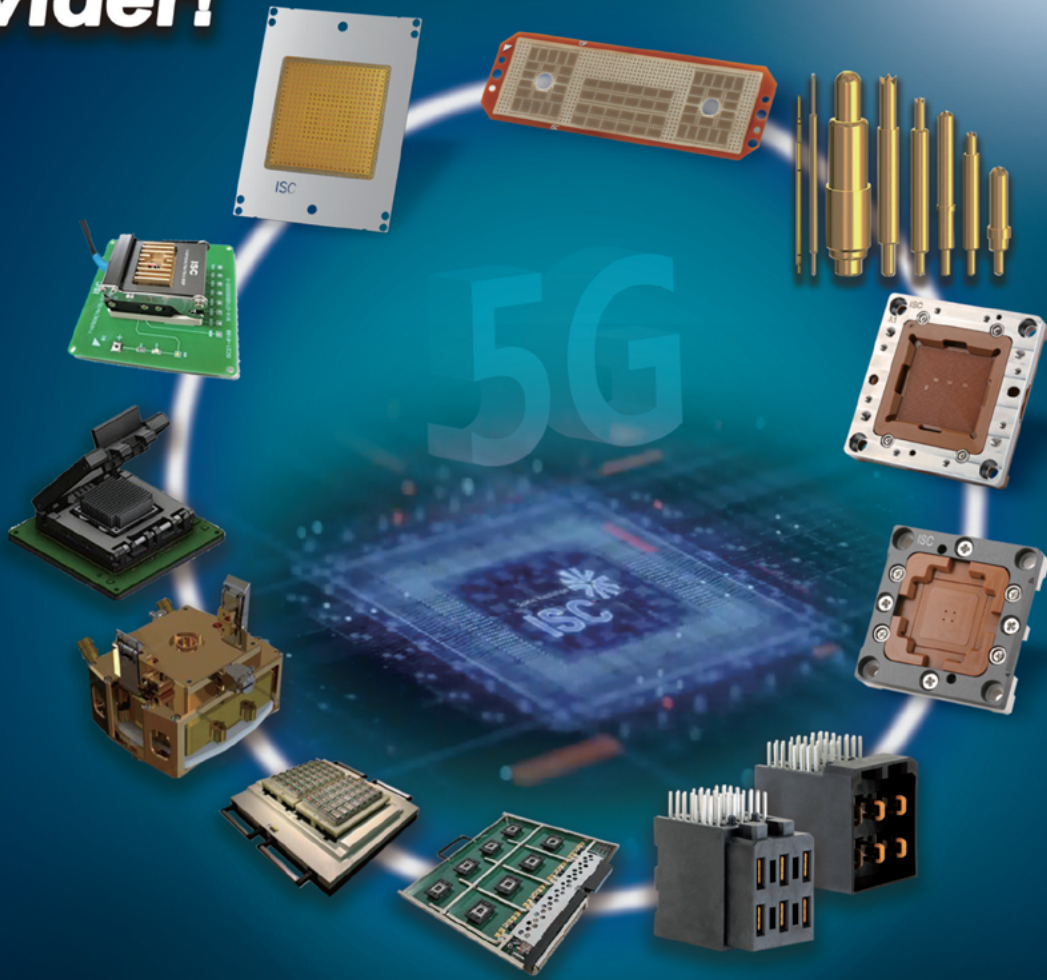
Wafer-to-Wafer Direct Bonding



Global No.1! Total Test Solution Provider!



www.isc21.kr
ISC
ISP
ISB
ISTIU
ISCON
ISTCU



ISC BURN-IN SOCKET SOLUTIONS

- No ball damage
- Direct inserting on Burn In Board without soldering
- No damage on Burn In Board land
- Simple structure without sub PCB & easy maintenance

ISC POGO SOCKET SOLUTIONS

- Various design available
- Excellent gap control & long lifespan
- High bandwidth & low contact resistance
- Adjustable Impedance
- Fully Metal Shielding

ISC Hi-fix & Burn In Board

- High performance and competitive price
- Test fine pitch, high speed device at hot & cold temperature
- Customized design to meet individual requirements

ISC ELASTOMER SOCKETS & INTERPOSERS

- High performance and competitive prices
- High speed & RF device capability
- No load board pad damage & no contact trace on the ball
- Customized design to meet challenging budget constraints
- Full thermal and electrical simulation

ISC Connector

- ISC Connector Solutions solve many problems across a wide range of circuit sizes, configurations, pitch, and PCB-attach terminations.
- Designed to have the strong resistance against torque forces on mating area to achieve the high reliability in the natural fall shock.

ISC THERMAL CONTROL UNIT

- Extreme active temperature control
- Customized design to meet challenging requirements
- Price competitiveness through self designing and fabrication
- Safety auto shut down temperature monitoring of the device & thermal control unit
- Full FEA analysis for strength, deflection, air flow and any other critical requirements



CONTACT
ISC CO., LTD

ISC HQ
Seong-nam, Korea

ISC International
Siliconvalley, CA

Tel: +82 31 777 7675 / Fax : +82 31 777 7699
Email:sales@isc21.kr / Website: www.isc21.kr

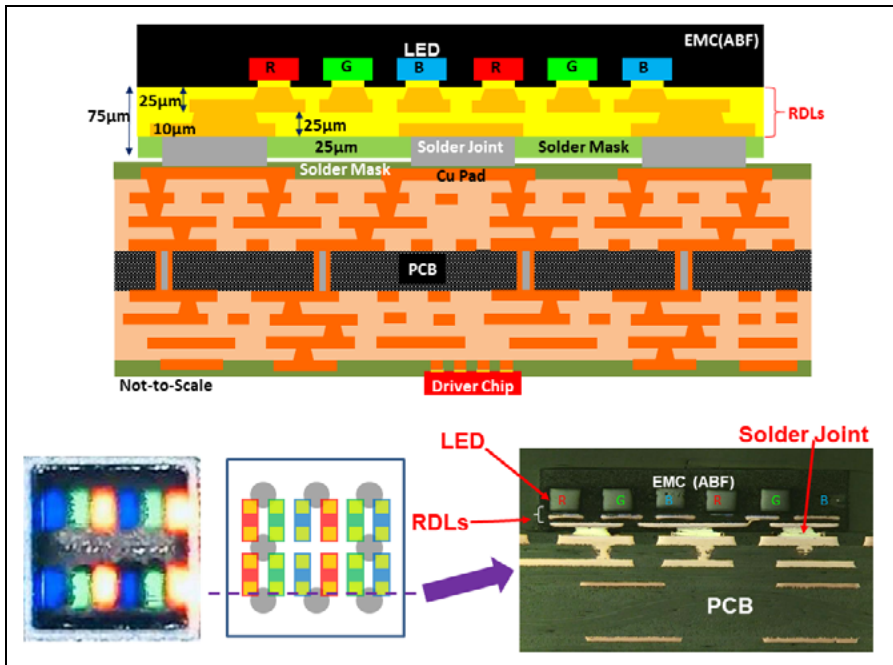


Figure 2: Heterogeneous integration of mini-LEDs for an RGB display (chip first die face down with a temporary panel process).

Chip first with die face up

Figure 3 shows an example of chip-first with die face-up fan-out packaging [4]. The chip size is 10mm x 10mm and the package size is 13.42mm x 13.42mm. The process steps of chip first with die face up is a little more complicated than that of chip first with die face down.

On the device wafer, one is to fabricate a Cu stud (about 15µm) on the original (Al or Cu) contact pads and the other is to laminate a die-attach film (DAF) on the bottom side of the device wafer. The function of the DAF is to attach (adhere) the die solidly onto the temporary carrier to avoid die shift caused by compression molding of the EMC; and the function of the Cu stud is to protect the original contact pads during backgrinding of the EMC, which is done to expose the Cu stud.

On the temporary glass carrier, a light-to-heat conversion (LTHC) layer (about 1µm) is spin coated onto the temporary glass wafer carrier. The chips are picked and placed face up on the LTHC carrier. In order to cure the DAF, a bonder with temperature and pressure should be used. The DAF process is carried out at 120°C (both bond head and bond stage) with a bond force of 2kg for 2s for each chip. The temporary carrier, therefore, will expand during the pick and place process. However,

during patterning/photolithography of the RDLs, the reconstituted carrier (temporary carrier + chips + EMC) is at room temperature. Therefore, pitch compensation caused by the DAF heating is needed [4]. After EMC

dispensing, compression molding, and then PMC are done. Then, the following are done: 1) backgrinding of the EMC to expose the Cu stud; 2) fabricating the RDLs; 3) and mounting the solder balls. Those processes are then followed by scanning a laser through the temporary glass carrier to the LTHC layer—the LTHC layer becomes powder, and the temporary glass carrier is then very easy to remove. Finally, the reconstituted wafer (with chips, EMC, RDLs, and solder balls) is diced into individual packages. There are three RDLs in each package and the minimum metal line width (L) and spacing (S) are 5µm. For detailed information on the design, materials, process, fabrication, and reliability of the chip-first with die face-up fan-out packaging, please see [4]. TSMC’s integrated fan-out (InFO) [5] used for Apple’s application processor is one of the chip-first with die face-up fan-out processes.

Die shift issues

In [4], we determined the die shift caused by compression molding by measuring the position of each chip before and after molding. (The die size is 10mm x 10mm and the minimum metal L/S are 5µm.) **Figure 4** shows the

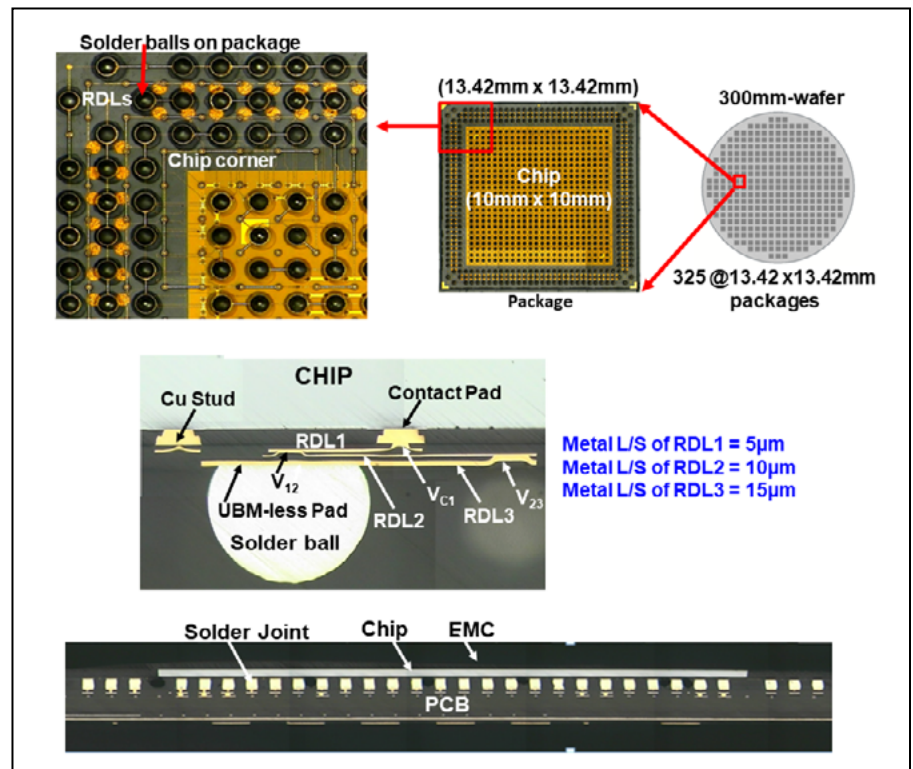


Figure 3: A chip-first with die face-up packaging process.

Because of the die-attach film (DAF), the die-shift is very small!

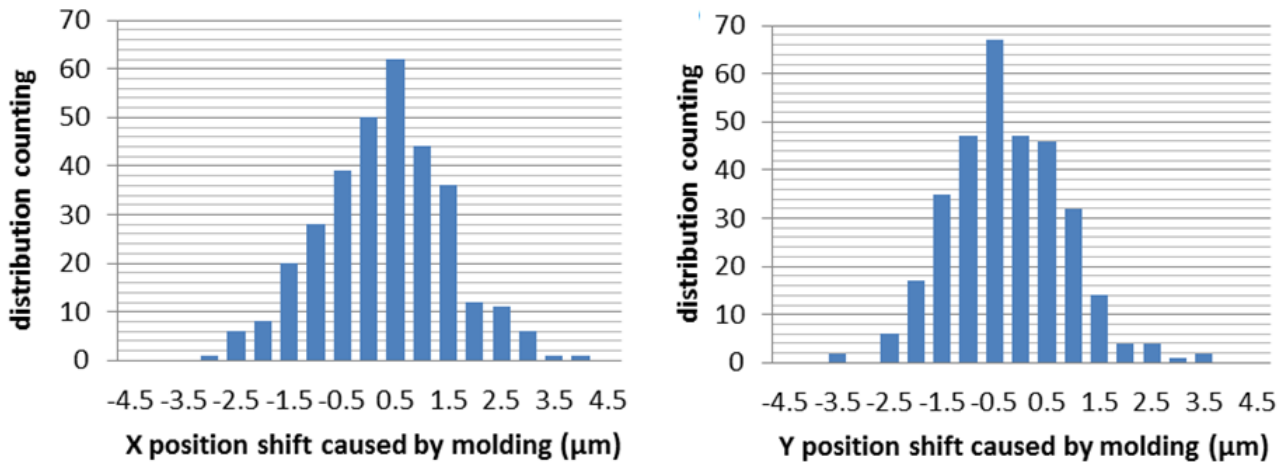


Figure 4: Die shift measurement of a chip-first with die face-up packaging process.

statistical plots of the x-position die shift and y-position die shift caused by the compression molding. It can be seen that because of the DAF (which solidly holds the chip to the carrier), the die shift (can be controlled within $\pm 3\mu\text{m}$) is too small to be an issue when making the RDLs.

In general, in order to avoid the die shift issues, the chip-first with die face-down process is used mostly for smaller die ($\leq 5\text{mm} \times 5\text{mm}$) and larger metal L/S RDLs ($\geq 10\mu\text{m}$), and chip-first with die face-up processing is used for larger die ($\leq 12\text{mm} \times 12\text{mm}$) and smaller metal L/S RDLs ($\geq 5\mu\text{m}$).

Warpage issues

Another critical issue for chip-first fan-out packaging is warpage [6,7]. There are at least two kinds of warpage about which we should be concerned: 1) the warpage of the reconstituted carrier should not be too large to affect the downstream fan-out process flow such that the reconstituted carrier cannot be placed/operated on the RDL equipment; and 2) the warpage of the individual fan-out package should not be too large so that it affects the quality and reliability of the surface mount technology (SMT) assembly, such as causing a stretched solder joint, for example. For detailed discussion and the allowable warpage for chip-first fan-out packaging, please see [6,7].

For the chip-first with die face-up process, it is interesting to note that the warpage of the temporary carrier + chips + EMC right after PMC has been found to be in the shape of a smiling face [7].

The average maximum warpage is equal to $609\mu\text{m}$ (Figure 5a). The shadow Moiré measurement result has been found to be in excellent agreement with the simulation result (Figure 5b). The warpage of the temporary carrier + chips + EMC right after backgrounding of the EMC to expose the Cu stud has been found by the shadow Moiré method to have changed from a smiling face to a crying face (Figure 5a). A similar trend has been found by the simulation method (Figure 5b) [7].

Chip last (RDL first)

The very first paper on chip-last (or RDL-first) technology was published by NEC Electronics Corporation (now Renesas Electronics Corporation) at IEEE/ECTC 2011 [8]. In the past few years, many companies such as Amkor, IME, ASE, SPIL, TSMC, Samsung, Shinko, and Unimicron, have also published papers on this topic. The process steps of the chip-last approach

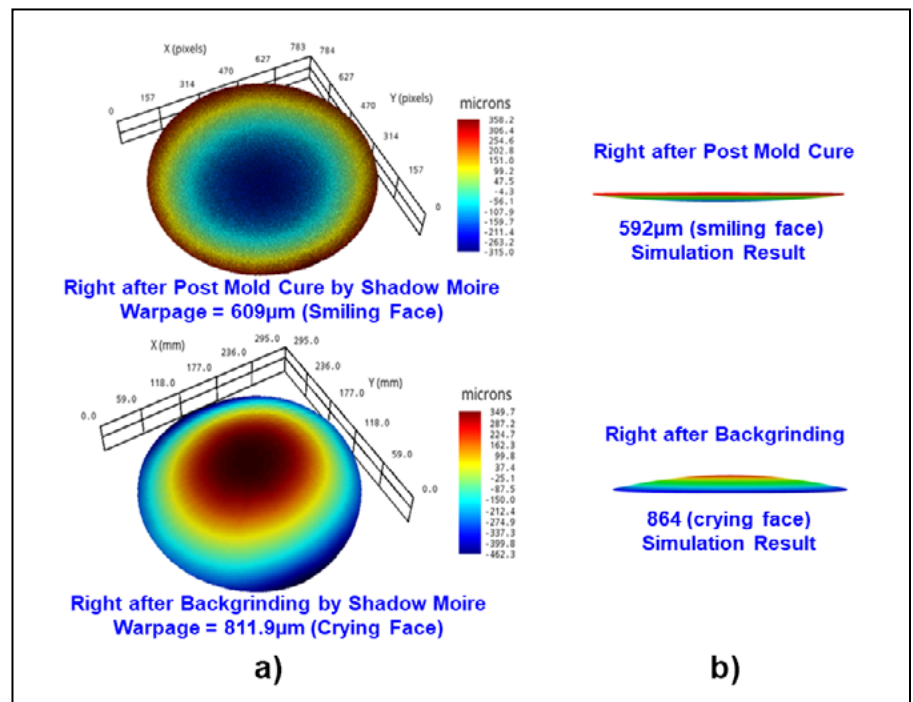


Figure 5: Warpage measurement and simulation of a reconstituted wafer fabricated by the chip-first with die face-up packaging process.

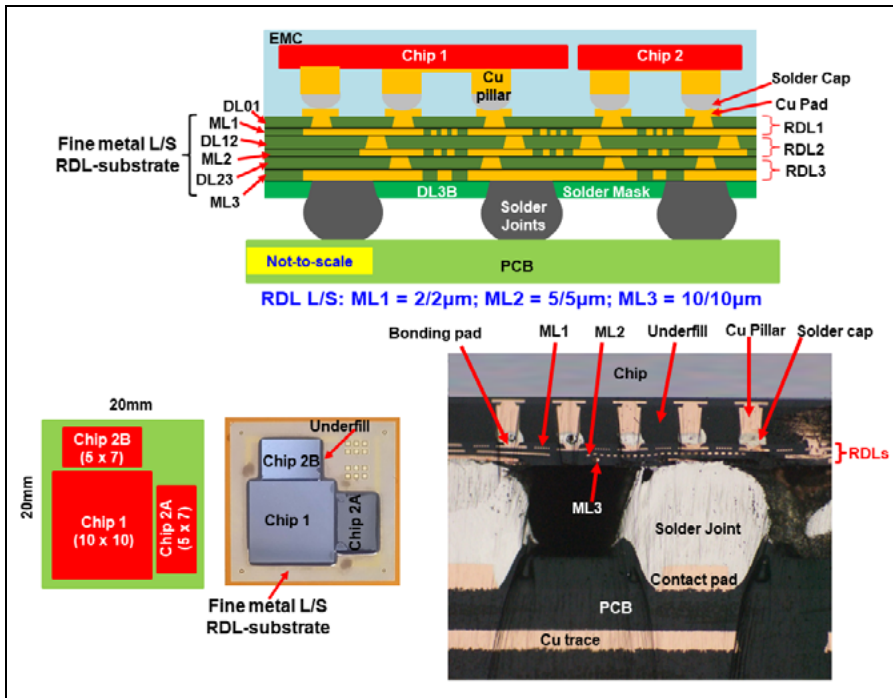


Figure 6: Heterogeneous integration of three chips on a fine-metal L/S (2 μ m-minimum) RDL substrate using a chip-last fan-out panel process.

are much more complicated than those of chip-first with face-up and face-down processes. The chip-last process is meant for high-density and high-performance (and therefore, higher cost) applications.

Figure 6 shows an example of heterogeneous integration of three chips on a fine-metal L/S RDL-substrate [9,10]. The size of the large chip is 10mm x 10mm, and that of the smaller chip is 5mm x 7mm. There are three layers of the RDL-first substrate, and the minimum metal L/S is equal to 2 μ m. One practical application of heterogeneous integration is for the application processor chipset, i.e., the large chip could be an application processor and the small chips could be memories.

The process steps for fabricating the RDL-first substrate are as follows. First, a LTHC film (1 μ m) is slit coated on a temporary rectangular glass carrier (515mm x 510mm) and that step is followed by slit coating a photo-imageable dielectric (PID) for the solder mask (or passivation layer) dielectric layer (DL) DL3B, as shown in **Figure 6**. Then, a Ti/Cu seed layer is formed by physical vapor deposition (PVD). That step is followed by applying photoresist, then using laser direct imaging (LDI), followed by photoresist development. Then,

electrochemical deposition (ECD) of Cu is done following stripping off the photoresist and etching off the Ti/Cu to obtain the metal layer (ML) ML3 of RDL3. Those steps are followed by slit coating a PID and then using LDI to obtain the DL (DL23) of RDL3. The next steps are: sputtering the Ti/Cu seed layer, slit coating

the photoresist, using LDI and then developing the photoresist, and then using ECD to deposit the Cu. These steps are followed by stripping off the photoresist and etching off the TiCu seed layer to get the ML (ML2) of RDL2. Next comes slit coating a PID and LDI to get the DL (DL12) of RDL2. The same process steps are repeated to obtain the ML (ML1) and DL (DL01) of RDL1. Next comes sputtering the Ti/Cu, slit coating the photoresist, LDI and develop, and using ECD to deposit the Cu. Those steps are followed by stripping off the photoresist and etching off the TiCu to get the bonding pad (lead) for the chips. The last step in the fabrication of the RDL substrate immediately before the chips-to-substrate bonding is the surface finishing of the Cu bonding pads. Electroless palladium and immersion gold (EPIG) surface finishing is used. The fabrication of the fine-metal L/S RDL substrate is thereby completed.

In parallel with the fabrication of the RDL-first substrate, the wafer bumping of the large and small chips with the standard PVD and ECD Cu and solder process is performed. The next step is dicing the wafers into individual chips. For all the chips, the bump consists of the Cu pillar, Ni barrier, and SnAg cap.

Now, we are ready to do the chips-to-RDL substrate bonding. It should be noted that, because of the support

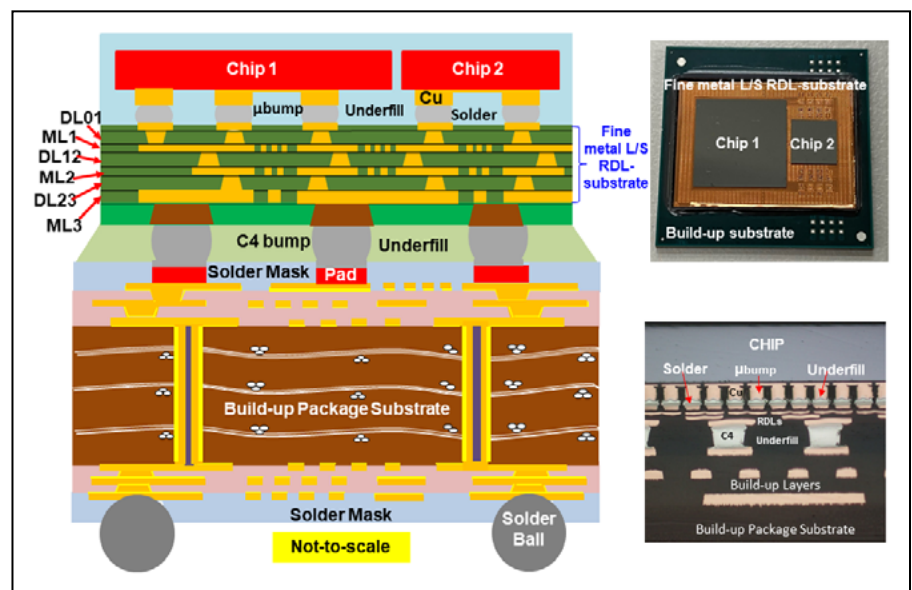
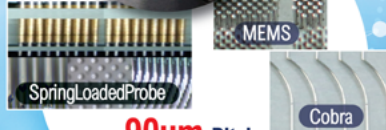
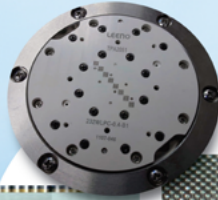


Figure 7: Heterogeneous integration of two chips on a hybrid substrate (a combination of the fine-metal L/S RDL substrate using a chip-last fan-out panel process, and the build-up package substrate with the C4 solder bump and underfill).

90um Pitch~

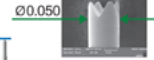


Spring Contact Probe



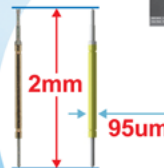
90um Pitch~
Probe Head

Tip Type

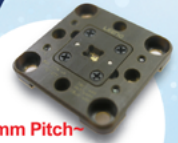


Specification

Pitch : Min.120um
Spring Force : 8.0g@ 250um
Current Rating : 1.0A
Inductance : 0.38nH



RF Probe for Fine Pitch Probe Head



0.18mm Pitch~

RF Coaxial Spring Probe & Impedance Controlled



Logic Test Socket

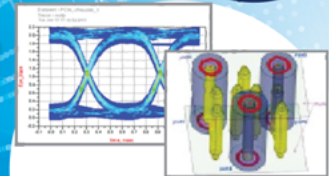
High Speed

Frequency : >20GHz
VSWR : < 1.2

Automatic Coaxial Probe



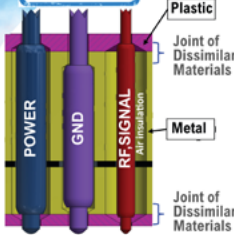
5G



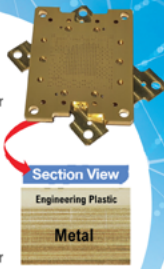
Electrical Analysis

CCC Test, HFSS, TDR
Eye Diagram
4 Port VNA Test

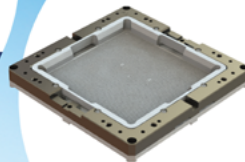
PATENT



MP Socket



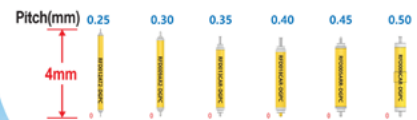
120mm x 120mm
> 10k Probe Count



Large Device Socket

Specification

Frequency : 80GHz(BGA),
100GHz(QFN, LGA)
Pitch : 0.25mm~
Crosstalk : -60dB
Impedance : 50Ω±10%



Coaxial Probe
100GHz

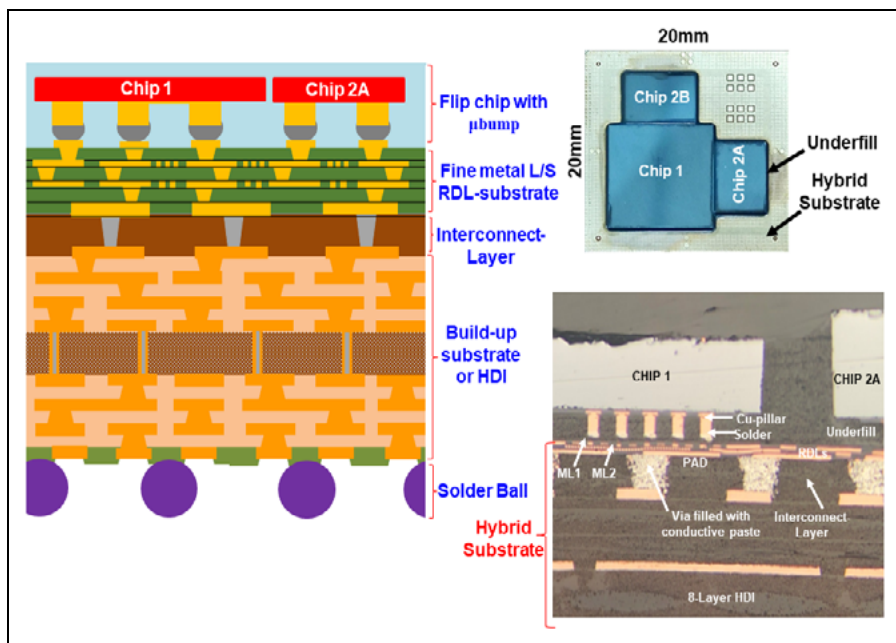


Figure 8: Heterogeneous integration of three chips on a hybrid substrate (a combination of the fine-metal L/S RDL-substrate by chip-last fan-out panel process and the build-up package substrate with the interconnect layer).

of the temporary glass carrier, the substrate is very stiff and flat prior to bonding. After the chips-to-RDL substrate bonding is complete, the next step is underfilling. The temporary glass carrier is removed by a laser so that we can make the solder resist opening and perform surface finishing on the Cu contact pads. Those steps are followed by solder ball mounting and dicing into individual packages. Finally, the individual package on the PCB is surface mounted. For more information on the design, materials, process, fabrication, and reliability of the PCB assembly of the heterogeneous integration package described above, please see [9,10].

Recently, 2.3D IC integration – where the fine-metal L/S RDL substrate and the build-up package substrate or high-density interconnect (HDI) are interconnected (combined) into a hybrid substrate through the controlled collapse chip connection (C4) solder joints that are enhanced with underfill – has been gaining traction thanks to companies such as STATS ChipPAC, Cisco, Amkor, ASE, MediaTek, SPIL, Samsung, TSMC, Shinko, and Unimicron. **Figure 7** shows an example of the hybrid substrate supporting two chips with microbumps. The large chip could be a system-on-chip (SoC) and the smaller

chip could be memory or a memory cube. For more information on the design, materials, process, fabrication, and reliability of the heterogeneous integration of two chips with 50 μ m pitch on a hybrid substrate by a fan-out RDL-first panel-level package, please see [11].

The fine-metal L/S substrate and the build-up package substrate, or

HDI substrate, can also be combined through an interconnect layer [12] into a hybrid substrate. This is very similar to [11] except the C4 solder joint and underfill are replaced by an interconnect layer as shown in **Figure 8**. For more information on the design, materials, process, fabrication, and reliability of the heterogeneous integration of three chips on a hybrid substrate with an interconnect layer by a fan-out RDL-first panel-level package, please read [12]. Again, Chip1 could be a SoC and Chip2A and Chip2B could be memories or memory cubes.

Heterogeneous integration of EIC and PIC devices

Figure 9 shows a conceptual layout of a 2.3D IC integration of a switch, PIC and EIC devices with a chip-last fan-out process to achieve lower power, higher speed, smaller form factor, and lower cost needed to achieve a higher data bandwidth for data center applications. It can be seen that the package substrate is supporting the fine-metal L/S RDL substrate, which is supporting the ASIC/switch, EIC and PIC with μ bumps. This structure is believed to be lower cost than the 2.5D IC integration of a switch, PIC and EIC devices with a through-silicon via (TSV) interposer as shown in Figure 25 of [13].

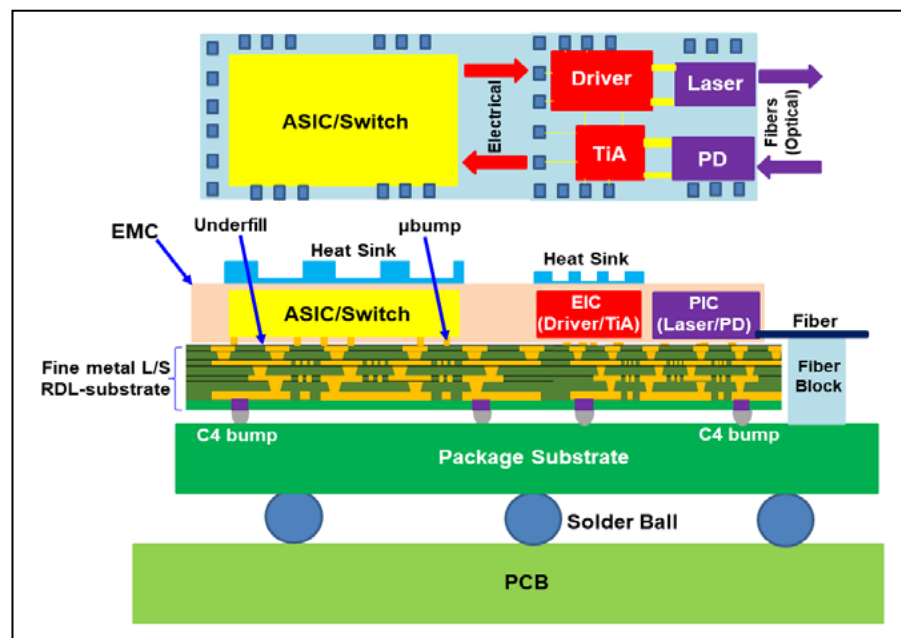


Figure 9: Heterogeneous integration of switch, EIC, and PIC on a fine-metal L/S RDL substrate for a data center application.

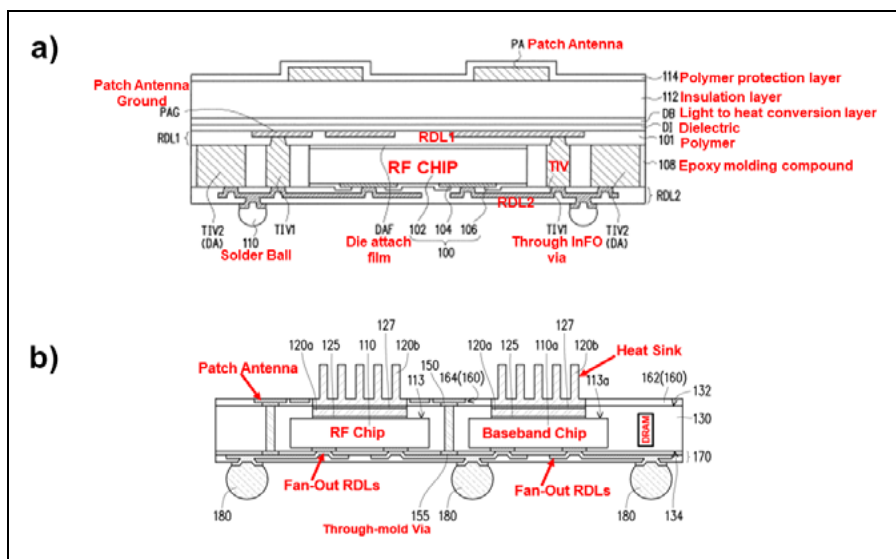


Figure 10: a) TSMC's AiP patent: US 10,312,112, June 4, 2019; b) Unimicon's heterogeneous integration of baseband and AiP patent: TW 1,209,218, November 1, 2020.

Antenna-in-package

In [14], TSMC demonstrated that the InFO_AiP for high-performance and compact 5G millimeter-wave system integration is superior than that of solder-bumped flip-chip AiP on substrate: 1) in the 28GHz frequency range, InFO RDLs transmission loss (0.175dB/mm) is 65% less than that on a flip-chip substrate trace (0.288dB/mm), and 2) in the 38GHz frequency range, the transmission loss for InFO RDLs (0.225dB/mm) is 53% less than that (0.377dB/mm) on a flip-chip substrate trace. TSMC's patent on InFO_AiP is shown in **Figure 10a**—it is a chip-first with die face-up fan-out process.

Figure 10b shows the Unimicon patent of the heterogeneous integration of AiP and a baseband chipset using a chip-first with die face-down fan-out process. It can be seen that the radio frequency (RF) chip and the baseband chipset (modem application processor and the dynamic random access memory [DRAM]) are placed side-by-side with RDLs and coupled with the antenna patches. A heat spreader/sink is also proposed, which is almost impossible using a chip-first with die face-up fan-out process.

Summary

Some important results and recommendations given the information presented are summarized as follows:

- The most important task in fan-out packaging is to fabricate the RDLs. There are at least three RDL

formations, namely chip first with die face down, chip first with die face up, and chip last or RDL first.

- The chip-first with die face-down package is for smaller chip sizes ($\leq 5\text{mm} \times 5\text{mm}$), larger metal L/S ($\geq 10\mu\text{m}$) RDLs, and smaller package sizes ($\leq 10\text{mm} \times 10\text{mm}$). Because of the small chip sizes and large metal L/S RDLs, the impact of die shift on the manufacturing yield is small. A couple of examples of the heterogeneous integration of chips and mini-LEDs have been provided.

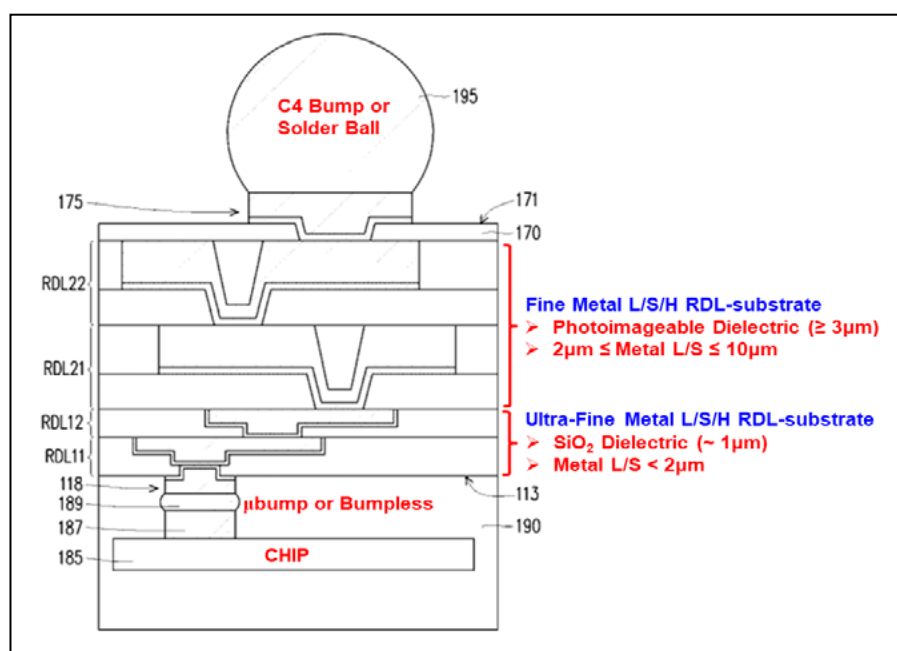


Figure 11: Unimicon's future ultra-fine metal L/S RDL-substrate. US Patent filed on April 19, 2021.

- A heterogeneous integration of the baseband chipset and AiP with heat spreader/sink fabricated using the chip-first with die face-down fan-out process has been proposed.
- The chip-first with die face-up process is for larger chip sizes ($\leq 12\text{mm} \times 12\text{mm}$), smaller metal L/S ($\geq 5\mu\text{m}$) RDL, and larger package sizes ($\leq 25\text{mm} \times 25\text{mm}$). Because of the DAF process, there is no die shift issue, which has been demonstrated with an example.
- The chip-last process can be used for very large chip sizes ($\leq 20\text{mm} \times 20\text{mm}$), very large package sizes ($\leq 55\text{mm} \times 55\text{mm}$), and very small metal L/S ($\geq 2\mu\text{m}$) RDLs (the so-called fine-metal L/S RDL substrate). A few examples of the heterogeneous integration of multi-chip on fine-metal L/S RDL substrate have been provided.
- With respect to chip-last with ultra-fine metal L/S ($< 2\mu\text{m}$) RDL substrates, we can say this: before the fine-metal L/S RDL substrate, one should first fabricate the ultra-fine metal L/S RDL substrate as shown in the US patent application (**Figure 11**) by Unimicon.
- A heterogeneous integration of the application-specific IC (ASIC)/switch, EIC, and PIC on a fine-metal L/S RDL substrate fabricated by the chip-last fan-out process has been proposed.



AT&S

HIGH-TECH PROVIDER FOR BEST SOLUTIONS WORLDWIDE

AT&S Austria Technologie & Systemtechnik Aktiengesellschaft

smc@ats.net www.ats.net

Acknowledgments

The author would like to thank his co-authors of the papers [1-4,6,7,9-13] cited in this brief article. Their useful and constructive contributions are greatly appreciated.

References

1. J. H. Lau, M. Li, M. Li, T. Chen, I. Xu, X. Qing, et al., "Fan-out wafer-level packaging for heterogeneous integration," *IEEE Trans. on CPMT*, Vol. 8, Issue 9, Sept. 2018, pp. 1544-1560.
2. C. Ko, H. Yang, J. H. Lau, M. Li, M. Li, C. Lin, et al., "Chip-first fan-out panel-level packaging for heterogeneous integration," *IEEE*

- Trans. on CPMT, 2018, Vol. 8, Issue 9, Sept. 2018, pp. 1561-1572.
3. J. H. Lau, C. Ko, C. Lin, T. Tseng, K. Yang, T. Xia, et al., "Fan-out panel-level packaging of Mini-LED RGB display," *IEEE Trans. on CPMT*, Vol. 11, No. 5, May 2021, pp. 739-747.
4. J. H. Lau, M. Li, Q. Li, I. Xu, T. Chen, Z. Li, et al., "Design, materials, process, and fabrication of fan-out wafer-level packaging," *IEEE Trans. on CPMT*, Vol. 8, Issue 6, June 2018, pp. 991-1002.
5. C. Tseng, C. Liu, C. Wu, D. Yu, "InFO (wafer-level integrated fan-out) technology," *IEEE/ECTC Proc.*, May 2016, pp. 1-6.
6. J. H. Lau, M. Li, D. Tian, N. Fan, E.

- Kuah, K. Wu, et al., "Warpage and thermal characterization of fan-out wafer-level packaging," *IEEE Trans. on CPMT*, Vol. 7, Issue 10, Oct. 2017, pp. 1729-1738.
7. J. H. Lau, M. Li, Y. Li, M. Li, I. Au, T. Chen, et al., "Warpage measurements and characterizations of FOWLP with large chips and multiple RDLs," *IEEE Trans. on CPMT*, Vol. 8, Issue 10, Oct. 2018, pp. 1729-1737.
8. N. Motohashi, T. Kimura, K. Mineo, Y. Yamada, T. Nishiyama, K. Shibuya, "System in a wafer-level package technology with RDL-first process," *IEEE/ECTC Proc.*, May 2011, pp. 59-64.
9. J. H. Lau, C. Ko, K. Yang, C. Peng, T. Xia, P. Lin, et al., "Panel-level fan-out RDL-first packaging for heterogeneous integration," *IEEE Trans. on CPMT*, Vol. 10, No. 7, July 2020, pp. 1125-1137.
10. J. H. Lau, C. Ko, T. Peng, K. Yang, T. Xia, P. Lin, et al., "Chip-last (RDL-first) fan-out panel-level packaging (FOPLP) for heterogeneous integration," *IMAPS Trans., Jour. of Microelectronics and Electronic Packaging*, Vol. 17, No. 3, Oct. 2020, pp. 89-98.
11. J. H. Lau, G. Chen, J. Huang, C. Yang, N. Liu, T. Tseng, "Hybrid substrate by fan-out RDL-first panel-level packaging," *IEEE Trans. on CPMT*, Vol. 11, No. 8, Aug. 2021, pp. 1301-1309.
12. C. Peng, J. H. Lau, C. Ko, P. Lee, E. Lin, K. Yang, et al., "High-density hybrid substrate for heterogeneous integration," *IEEE Trans. on CPMT*, Vol. 12, No. 3, Mar. 2022, pp. 469-478.
13. J. H. Lau, "Recent advanced and trends in advanced packaging," *IEEE Trans. on CPMT*, Vol. 12, No. 2, Feb. 2022, pp. 228-252.
14. C. Wang, T. Tang, C. Lin, C. Hsu, J. Hsieh, C. Tsai, et al., "InFO_AiP technology for high-performance and compact 5G millimeter wave system integration," *Proc. of IEEE/ECTC*, May 2018, pp. 202-207.



Semcon[®]

Semiconductor Equipment for a full range of applications

Technic supplies a full range of semiconductor electroplating tools including laboratory and small-scale production plating technologies and fully automated production equipment for electroless processes such as ENIG and ENEPIG. From single-cell countertop testing devices to fully enclosed automatic rack transport systems with multi-programmable functions, Technic has the right Semcon tool for your operation.

TECHNIC
www.technic.com



Biography

John H. Lau is a senior special project assistant at Unimicron Technology Corporation, Taoyuan City, Taiwan (ROC). He has more than 40 years of R&D and manufacturing experience in semiconductor packaging, 510 peer-reviewed papers, 40 issued and pending US patents, and 22 textbooks. He is an ASME Fellow, IEEE Fellow, and IMAPS Fellow. He earned a PhD degree from the U. of Illinois at Urbana-Champaign. Email John_Lau@unimicron.com

Fan-in wafer-level packaging: any chip can be a flip chip!

By Ray Fillion [Fillion Consulting]

In the November/December issue of *Chip Scale Review*, seven advanced packaging technologies were described [1]. The article covered embedded chip packaging (ECP), fan-in wafer-level packaging (FIWLP), fan-out wafer-level and panel-level packaging (FOWLP, FOPLP), 3D chip stacking, package-on-package (PoP) stacking and system-in-package (SiP) and compared how well each meets the basic functions of microelectronics packaging. In this article, we will go into a more in-depth look at FIWLP device structures and how they are fabricated. We will also look at the trends in semiconductors and microelectronics and discuss the developments needed for FIWLP technologies to support these trends in terms of I/O density, power dissipation, input supply voltage levels and chip operating frequencies.

Fan-in technologies

FIWLP technologies were developed to enable the flip attachment of chips that were designed for wire bonding, using area solder bumps. More than 90% of semiconductor chips produced have perimeter I/O pads designed for wire bonding onto a chip carrier package. A packaged wire-bonded chip has a footprint 4 to 10 times larger and 2 to 4 times thicker than a chip, whereas a flip-chip package is chip size. Manufacturers of hand-held electronics, particularly smartphones and smart watches, demanded

smaller footprint, thinner devices for thinner, lighter products with increased functionality. Without fan-in technologies, personal electronic products could not be as thin, as light and as small as they are, and still have the functionality and performance that they have.

As indicated by their name, fan-in devices reconfigure I/O pads from the chip perimeter to an area over the surface of the chip. FIWLP devices are processed in one of two formats, on-wafer or on a molded-wafer formed on a 300mm diameter footprint. **Table 1** compares flip-chip, fan-in and fan-out technologies for four key features: footprint, I/O capacity, costs and maturity. Flip-chip and fan-in devices are chip size, while fan-out devices are 2X to 5X larger. As for I/O capacity, flip-chip devices can have 1000s, while fan-out devices can have in the high 100s, and fan-in devices can have in the low 100s. Flip-chip devices have the lowest cost as there are no post-wafer processing steps required, while panel-level processing has lower costs than wafer-level processes, and fan-in has lower costs than fan-out. Flip-chip processing has the highest maturity, while wafer-level processes have higher maturity than the panel-level processes.

FIWLP structures

There are two basic FIWLP approaches: 1) on-wafer redistribution layer (RDL) processing, which is

directly on a semiconductor wafer, and 2) molded-wafer RDL processing, which is processing on a reconstituted, molded wafer. Both of these FIWLP approaches utilize 300mm semiconductor wafer processing equipment.

On-wafer FIWLP devices. A typical on-wafer FIWLP device is depicted in perspective view (top) and cross-sectional view (bottom) in **Figure 1**. The perspective view shows a chip with 24 wire bond pads on the perimeter area of the two narrow ends of the chip, in single rows of 12 pads each. A 5 by 6 area array of 30 solder bumps are located over the center of the chip. The cross-sectional view shows a first organic dielectric layer covering the chip with microvias formed through the dielectric directly to the chip pads. A patterned RDL metallization layer is formed on the top of the first dielectric layer connecting through the microvias to the pads and routing out to array pads. A second dielectric layer or passivation layer covers the RDL lines and has openings to the array pads. Optionally a second metallization layer is used to form the array pads. Solder bumps are attached to each array pad.

Molded-wafer FIWLP devices. A typical molded-wafer FIWLP device is depicted in **Figure 2** with perspective view (top) and cross-sectional view (bottom). The RDL structures over the chip are identical to the on-wafer FIWLP device in **Figure 1**. As seen in the cross-sectional view, the chip's

Technology	Footprint	I/O Capacity	Costs	Maturity
Flip Chip	Chip Size	Very High	Low	Very High
FI-WLP	Chip Size	Low	Moderate	High
FI-PLP	Chip Size	Low	Low	Low
FO-WLP	>> Chip Size	Moderate	High	High
FO-PLP	>> Chip Size	Moderate	Moderate	Low

Table 1: Comparisons of flip-chip to fan-in and fan-out devices for size, I/O capability, costs and maturity.

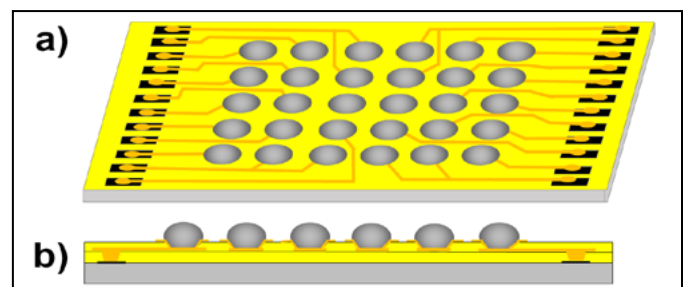
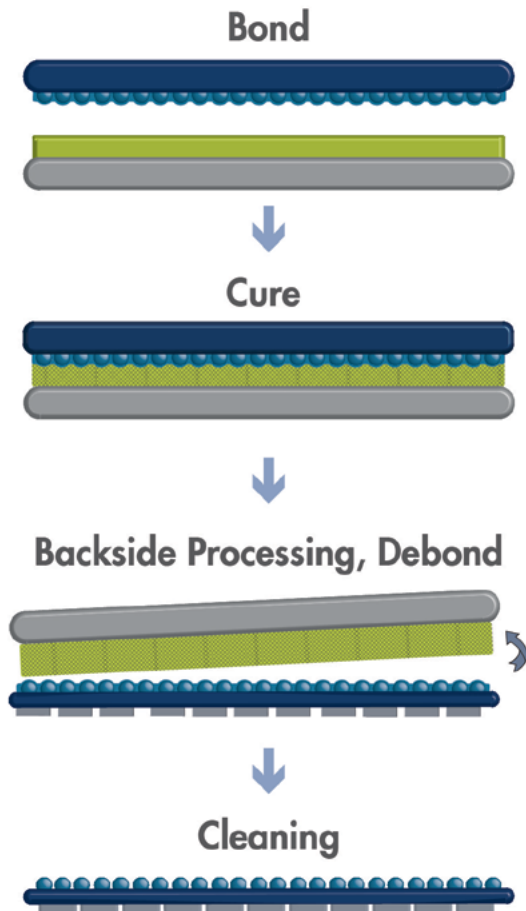


Figure 1: Typical on-wafer FIWLP device in a) (top) perspective view; and b) (bottom) cross-sectional view.

Dual-Layer Solution

A path-changing direction for temporary bonding



Our unique dual-layer solution for high-temperature & high-stress applications found within the semiconductor industry

www.brewerscience.com

©2020 Brewer Science, Inc.

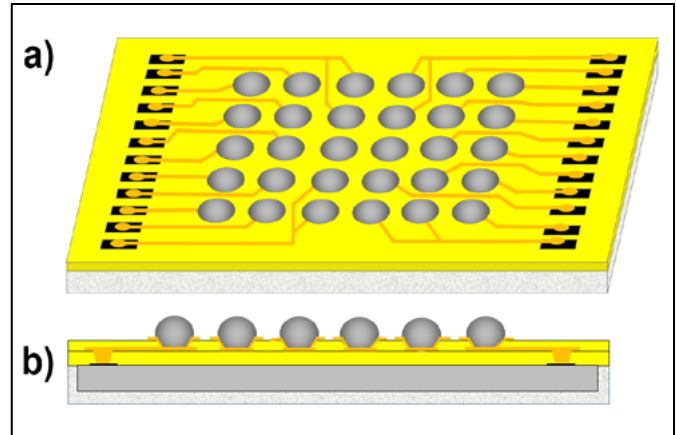


Figure 2: Typical molded-wafer FIWLP device in a) (top) perspective view, and b) (bottom) cross-sectional view showing molding material covering the chip's side edges and backside surface.

sides and its back surface are covered with the organic molding material. This molding material, in addition to forming the molded-wafer structure, provides protection to the chip from mechanical damage, from moisture and from processing fluids (i.e., fluxes, etch materials, cleaners).

Fan-in pad and bump structure

Figure 3 depicts an enlarged view of a typical on-wafer or molded-wafer FIWLP device structure. It shows a microvia to a chip pad, RDL routing from the chip pad to the array solder pad and a solder bump on the pad. The microvia typically has a diameter of 25 to 50 μ m and is formed through the lower dielectric layer (typically 5 to 10 μ m-thick) to the chip perimeter I/O pad. A barrier metal separates the chip pad (typically aluminum), from the RDL metallization (typically copper). The RDL metallization connects through the microvia onto the dielectric, and routes to the center area and the array pad. A passivation layer or a second dielectric layer covers the first RDL layer. A second patterned metallization layer connects through a via in the second dielectric to the array pad. A solder bump is attached on the array pads.

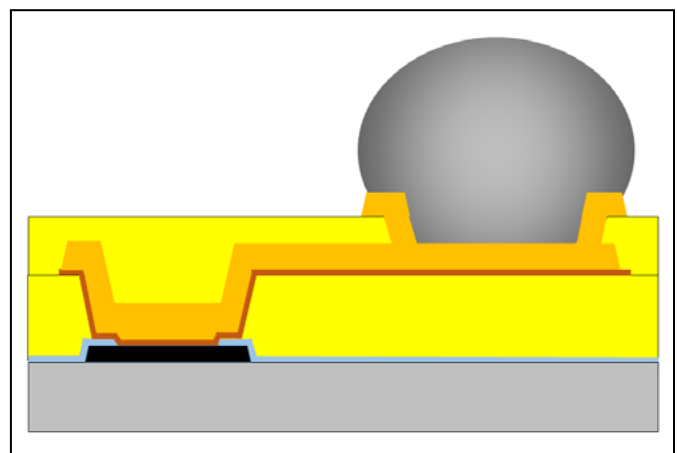


Figure 3: Expanded cross-sectional view of a FIWLP device showing an RDL microvia to the chip perimeter pad, RDL routing, and a solder bump on an RDL solder pad located on the dielectric layer.

FIWLP processes

The processing steps used to convert a chip designed for wire bonding into a FIWLP are nearly identical for both on-wafer processing and molded-wafer processing. The exceptions are the processing steps used to fabricate the molded wafer. Once the molded wafer is complete, exactly the same processing steps, processing equipment and processing materials are used by both.

On-wafer FIWLP processing steps.

Figure 4 depicts the typical processing steps for an on-wafer FIWLP. The on-wafer FIWLP approach starts with a completed wafer that is fully tested and ready for wafer dicing. The RDL processing is done using back-end-of-line (BEOL) wafer fabrication equipment. In step a), a thin dielectric layer (BCB, polyimide) is applied to the wafer surface by spin coating. In step b), microvias are formed through the dielectric layer to the chip perimeter pads. In step c), a thin metal layer (5 to 10 μ m) is deposited on the surface of the dielectric and into the microvias and is patterned to form the RDL layer. Complex devices with more I/Os or with less chip area may require one or more additional RDL layers. In step d), a passivation layer or a second dielectric layer is applied over the first RDL layer. In step e), openings are formed through that layer to the first RDL metal layer. In step f), the pad metallization layer is applied to the top surface and patterned forming the array pads. In step g), solder paste is applied and reflowed forming the solder bumps. In step h), the wafer is diced forming multiple FIWLP devices with their perimeter I/O pads reconfigured into an array of solder bumps and in effect, forming pseudo flip-chip devices.

Molded-wafer FIWLP processing steps. **Figure 5** depicts typical processing steps for a molded-wafer FIWLP process. In step a), thermal release tape is laminated to the top of a processing carrier, typically in a 300mm diameter wafer format. In step b), multiple bare chips are mounted face down onto the tape and held in place. In step c), molding material is applied by compression molding to embed the chips and form a molded-wafer with an array of bare chips. In step d), the molded-wafer is removed from the release tape and the processing carrier. In step e), the molded-wafer is back ground to thin the structure and provide a planar surface. In steps f) through h), the RDL processing steps are identical to processing steps a) through g)

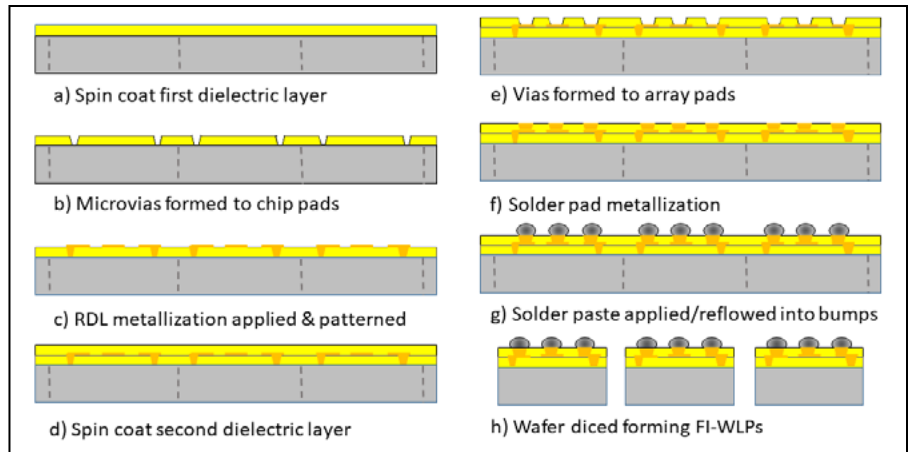


Figure 4: Typical on-wafer FIWLP device processing steps forming RDL structures on a wafer.

of the on-wafer processing steps in **Figure 4**. In step i), the molded-wafer is diced to form FIWLP devices with molding material covering the sides and the back surface of each device.

FIWLP development needs

FIWLP devices are rapidly increasing in both number of devices shipped and total market value. Yole reported in 2020 that

FIWLP had revenues of over \$2.5B in 2019 and were estimated to rise to \$3.5B in 2025 [4]. Semiconductor trends are continuing with more gates per chip, which in turn is driving higher I/Os per chip, higher power dissipation per chip, faster switching frequencies and lower supply voltages. All of these trends affect the FIWLP requirements in the coming years. In order to extend FIWLP technologies to meet the needs

QUALITY IS EVERYTHING

Your yield. Your profitability. Your reputation depends on 100% quality assurance for every wafer, device and package.

Sonix is the leader in ultrasonic technology and expertise for inspecting wafer bonds, device interconnects and package integrity.

Find smaller defects faster, at any layer. Learn more and request a free sample analysis at [Sonix.com](https://www.sonix.com).

sonix[™]

© 2016 Sonix, Inc. All rights reserved.

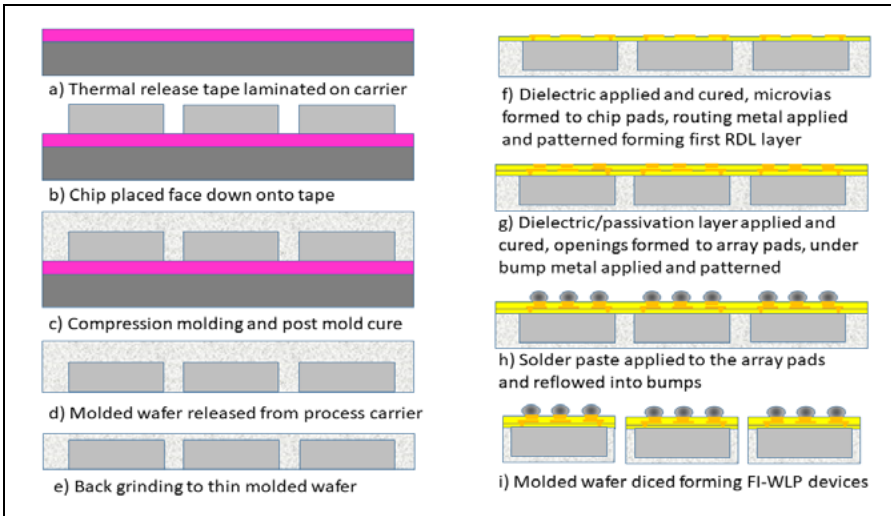


Figure 5: Typical molded-wafer FIWLP device processing steps, from forming a molded-wafer and applying RDL structures.

of these next-generation semiconductor devices while maintaining high yields and lowering costs, FIWLP fabricators need to develop the capability for: 1) smaller solder bump pitches; 2) tighter RDL line widths and spacing; 3) higher thermal conductivity; 4) reduced RDL interconnect parasitics; 5)

reduced chip placement tolerances and chip movement; and 6) adaption of via locations and RDL routing.

Smaller solder bump pitches.

Increasing chip I/O counts requires shrinking array pad pitches of FIWLP devices. Although smaller solder bump

itches (55µm and below) have been demonstrated for flip-chip devices, they are only done where a flip-chip device is mounted onto a silicon substrate or onto a substrate having a similar low coefficient of thermal expansion (CTE) and are inherently planar. Tighter array pad pitches will require smaller diameter solder bumps with much lower solder height. On-wafer FIWLP devices can be fabricated with bump pitches down to 50µm and below when used on a silicon or other low coefficient of thermal expansion (CTE) substrate. Molded-wafer FIWLP devices are not dimensionally-stable enough to support these low solder bump pitches. Incorporation of Cu pillars on array pads would allow FIWLP devices to reduce solder bump height and, therefore, array pitch. Although most FIWLP devices are mounted on fine-line circuit boards, a growing number will be assembled onto a SiP along with flip-chip devices and 3D chip stacks. These will utilize a silicon or glass substrate permitting small pitch solder bump attach.

Reducing RDL line widths and spacing.

Increasing the I/O count on a FIWLP will also require RDL line widths and line spacing to decrease proportionally. Metallization and metal patterning techniques particularly for molded-wafer processing, need to shift from subtractive metal patterning (standard PC panel processing) to semi-additive metallization techniques (standard semiconductor processing) providing rectangular line cross-sections, finer line width control and lower interconnect resistance. Molded wafers have a mix of low-CTE silicon chips and high-CTE molding material and RDL dielectrics. This combination can cause warpage, poor planarity and variable and nonuniform molded-wafer shrinkage. Molded-wafer FIWLP has chip location issues related to chip placement tolerances and chip movement after placement. All of these make going to finer features on molded-wafer FIWLP devices problematic.

Higher power dissipation. Higher power dissipation chips can cause over heating of the chips and softening of organic materials unless a low thermal resistance cooling path is provided. Higher power dissipation can also cause device hot spots that can exacerbate solder fatigue in smaller solder bumps. Since all fan-in devices have at least one organic layer overlying the chip surface, cooling a higher power dissipation chip through its top surface or through its solder bumps is less efficient than cooling

Visit us on our website
www.e-tec.com

NEW
in the Test Socket World:

E-tec Interconnect is pleased to present its new Clamshell Open Top Socket

- **High reliability** up to 500k insertion cycles
- **High frequency** up to 27 GHz in pogo pin (up to 40GHz validated, with Elastomeric contact technology)
- Thru-hole technology, SMT, Solderless Type
Also pluggable into adapter MiniGridSocket series (see E-tec catalog TS01)
- All kinds of packages, even your latest special custom packages

Flat Pack PGA LGA Ceramic spatial LCC BGA QFN / Power QFN Custom

E-Tec Interconnect AG, Mr. Pablo Rodriguez, Lengnau Switzerland
Phone : +41 32 654 15 50, E-mail: p.rodriguez@e-tec.com

EP Patents 0897655, 1385011, 0829188, US Patents 6249440, 6190181, 6390826 and Patents in other countries

a flip-chip device. One option to improve topside cooling of a FIWLP device is to form a thermal pad under each array solder pad. This could be done by opening vias in the first dielectric layer directly under each array pad to the chip passivation layer. During the first RDL metal processing step, the thermal vias could be metallized, forming thermally-conductive posts that provide each solder bump with a direct thermal path to the chip.

Higher device switching frequencies.

Higher device switching frequencies can create additional RDL line cross talk that can affect switching margins. Faster switching may require isolating ground planes within the RDL layers to either provide a controlled impedance transmission line, or to minimize cross talk between adjacent lines. Lower supply voltages generate higher power and ground currents that will increase resistive line losses and a higher device sensitivity to power rail droop and surges, thereby lowering device operating margins.

Chip misplacement and movement.

A unique problem for all molded-wafer FIWLP devices, are lateral and rotational chip placement tolerances and chip movement after placement. Whereas chip locations for on-wafer FIWLP processes are precise down to fractions of microns, all molded carrier processes use pick and place machines that inherently have placement tolerances more than an order of magnitude higher, i.e., multiple microns. Depending on the processes and the organic materials used to bond the chip to the carrier before molding, there can be chip movement after placement included during adhesive curing, molding and molding material curing. The molded carrier shrinks during the molding process, thereby adding a varying global offset. Finally, the molded-wafer has a higher composite CTE than a wafer does resulting in a varying chip position caused by temperature changes from the elevated temperature during adhesive and molding material curing and the lower temperature during photolithography steps. These chip position tolerances will limit the ability of molded-wafer FIWLP processes to shrink RDL line widths and solder bump pitches.

RDL adaption. One approach to overcoming chip misplacement and chip movement issues is adapting the locations of microvias and RDL routing for each device based on its precise position. This was first done using the GE high-density interconnect (HDI) multi-chip interconnect

technology [5]. This embedded chip technology mounted chips face up in cavities using liquid-dispensed chip attach adhesive. The combination of chip placement tolerances and excessive chip movement resulted in die location errors of 25 to 100 μ m. This technique measured the exact position of each chip corner pad using an automated vision system and adjusted microvia locations and RDL routing. This is only applicable to FIWLP

using laser-based photolithography. A similar technique has been implemented by Deca on its FOPLP processing line. It uses an automated imaging system to measure each chip's exact location based upon its corner I/O pads. It then recomputes the laser drill data base and RDL laser patterning data and forms the microvias in the correct locations and correctly routes the RDL lines and pads [6]. This technique would be needed to extend molded-wafer



LEADERS IN MICRO DISPENSING TECHNOLOGY

SMALL REPEATABLE VOLUMES ARE A CHALLENGE, BUT NOT IMPOSSIBLE IF YOU HAVE BEEN CREATING THEM AS LONG AS WE HAVE.

TO DO IT WELL, WE PROVIDE THREE THINGS:

Dispensing Expertise in a variety of microelectronic packaging applications.

Feasibility Testing & Process Verification based on years of product engineering, material flow testing and software control.

Product Development for patented valves, dispensing cartridges, needles, and accessories.

Our Micro Dispensing product line is proven and trusted by manufacturers in semiconductor, electronics assembly, medical device and electro-mechanical assembly the world over.

www.dltechnology.com.

216 River Street, Haverhill, MA 01832 • P: 978.374.6451 • F: 978.372.4889 • info@dltechnology.com



THE TEST...WITH BURN-IN COMPANY

Micro Control Company offers solutions to meet the challenges created by the burn-in requirements of high-power logic devices. By providing individual temperature control to each device under test guarantees that the correct thermal stress is applied to each device during the burn-in cycle. Lower power device applications not requiring individual temperature control can go through burn-in with test cycles in greater quantities.

Features:

- 150 watts per DUT power dissipation with air cooling
- 128 I/O channels with 64 M of vector memory behind all independent channels
- Individual pattern and voltage zones per slot
- Multiple temperature zones
- 1080 amps of current per burn-in board
- Future product development capabilities
- Low cost per DUT



Have Other Requirements?

Micro Control Company offers burn-in carts, prescreen stations, burn-in boards, and test program development.

MICRO CONTROL
COMPANY SINCE 1972

CELEBRATING 50 YEARS

7956 Main Street NE | Minneapolis, MN 55432 | 763-786-8750
www.microcontrol.com

FIWLP devices to the higher densities needed to meet the next-generation microelectronic devices. FIWLP devices fabricated on-wafer do not need microvias and RDL adaption as there is no chip placement step and no possibility of chip movement.

Summary

The long running advances in semiconductor processing capabilities are continuing with higher I/O counts forecast for the foreseeable future. One might assume that many of these higher I/O count devices would require FIWLP processing, but the fact is that high I/O count chips, such as in high-performance microprocessors and graphic processors that go into personal computers, mainframes and servers, as well as application and graphic processors that go into mobile devices, will be designed as flip-chip devices. They do not need additional package-level processing, such as FIWLP processing, as they will be area-array devices. FIWLP will continue to target chips designed for wire bonding with I/O counts in the low (~10 to ~50) to medium (~50 to a couple 100) range. FIWLP technologies still need to go to finer lines and tighter pad pitches as low and medium I/O count chips go through die shrinks that will require finer line RDL and tighter pitch solder bumps.

References

1. R. Fillion, "Advanced microelectronics packaging technologies and their performances," *Chip Scale Review*, Nov/Dec 2021 pp.10-17.
2. G. Ridly, "Introduction to flip chip: what, where, why, how," *Flipchips.com*, Oct. 2000.
3. R. Fillion, "Embedded chip build-up using fine-line interconnect," *ECTC 2007*.
4. "Advanced packaging quarterly market monitor," *Yole Développement*, March 2020.
5. "High-density copper/polyimide overlay interconnections," R. O. Carlson, et al., *IEPC 1988*.
6. "New commercialization of Deca's fan-out technology," *i-micronews.com*, March 2022.



Biography

Ray Fillion is Managing Director at Fillion Consulting, Schenectady, NY. He retired after 40 years from the GE Global Research Center where he worked in various engineering, management, business development and licensing positions in embedded chip, MCMs, 3D modules and power electronics. He has over 100 publications, has 45 issued U.S. patents and was the lead inventor on the GE Embedded Chip Build-Up and the GE Power Overlay technologies. Email fillion.consulting@gmail.com



Pressure clip sintering for high-power electronics

By Eric Kuah [ASM Pacific Technology Ltd]

The focus on vehicle electrifications and clean emissions has garnered much attention and press coverage. In light of this attention, major vehicle companies in the US, Europe, Japan, China, and South Korea have announced their intentions of building modules for battery electric vehicles (BEV) or plug-in hybrid electric vehicles (PHEV). This global demand further drives the market valuation – previously valued at \$163.01 billion in 2020 – it is projected to hit \$823.75 billion by 2030—a cumulative average growth rate (CAGR) of 18.2% from 2021 to 2030 [1].

As BEV and/or PHEV products continue to grow in demand, one thing that remains constant is their requirement for high-power electronics (HPE). This class of HPE package will then require pressure sintering,

a method whereby the semiconductor chip is attached to a substrate using a silver or copper paste as the main elemental component. For the chip to connect to the other connecting point to generate high voltage and deliver high current, a clip is an alternative interconnect to heavy wedge aluminum wire bonding. We will discuss the motivations of using a clip interconnect scheme for HPE and its part in the assembly journey while exploring the positives of pressure clip sintering.

Motivation for using clip interconnection in HPE

A clip interconnection used for HPE usually has copper as the base material because of its high thermal conductivity. Copper clips are used to replace heavy wedge aluminum interconnect because

they improve the thermal conductivity pathway, thereby averting overheating in the package by avoiding the formation of hot spots. Copper clips, therefore, improve the electrical performance by reducing parasitic inductance. Parasitic inductance is an unwanted inductance effect that is present in electronic modules that prevents electrical current flowing through electrical circuitry. It should be noted that inductance is only welcome when it is deliberately created using an inductor, along with a function in mind. The reduction of parasitic induction, therefore, will translate into improvement for the HPE package reliability when a clip interconnect is employed.

Another consideration for clip sintering in HPE is to account for built-up stress between the interface of the semiconductor die and the metallic clip. The resulting stress tensor

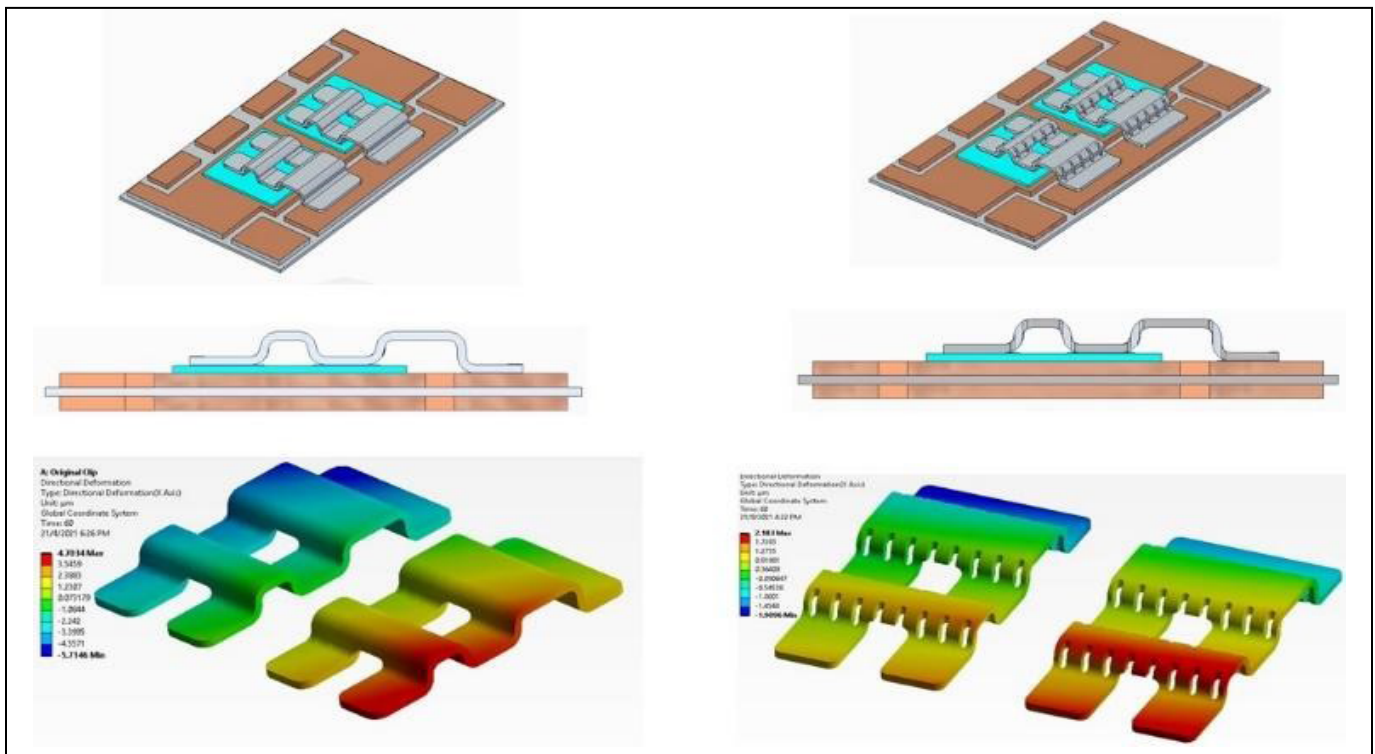


Figure 1: Clip designs in HPE for stress relief.

SEMI AMERICAS EVENTS

Technology Leadership Series of the Americas Events



The **Technology Leadership Series** of the Americas is a collection of eight major events representing the various technology communities in our industry. These programs connect 2,400+ member companies and over 1.3 million global professionals to advance the technology and business of electronics manufacturing and the design-supply chain.



There are numerous ways for you to participate in these events to connect with the industry and grow your business.

ATTEND | PRESENT | SPONSOR | EXHIBIT



SAVE THE DATES FOR THESE 2022 EVENTS

Explore these upcoming programs and fill out the Interest Form on each website to get connected for the latest event updates.



www.semi.org/semi-americas



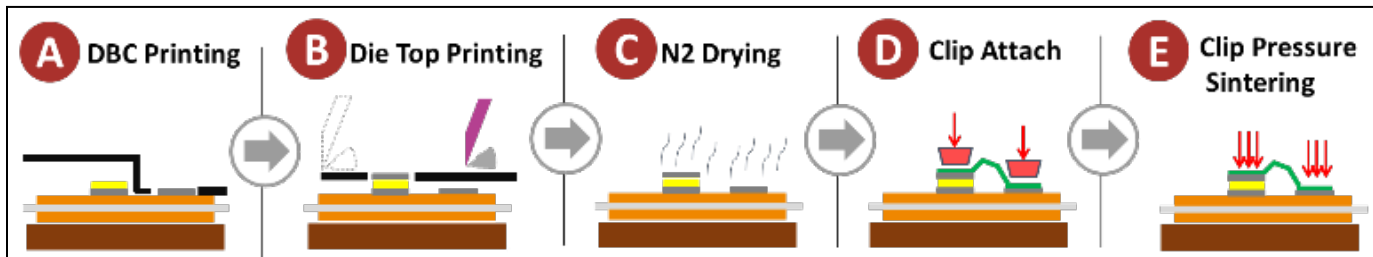


Figure 2: Clip sintering assembly process.

is because of the coefficient of thermal expansion (CTE) mismatch. To understand this importance, we can look at the CTE between the semiconductor die and metal clip. For example, if the semiconductor die is made from silicon carbide with a CTE range of 3-4ppm per °C, and the CTE for clip copper with a few microns of plating on its surface ranges from 16-18ppm per °C, the order of thermal expansion would be about 4 to 5 times the stress differential. If this stress differential for the clip sintering is not designed with sufficient stress relief features, such stresses during thermal cycle testing may lead to breakage of the HPE. Therefore, designs of stress relief features play a crucial role in reducing the interfacial stress between the copper clip and semiconductor chip metallization. **Figure 1** shows some potential ways to reduce the stress because of CTE mismatch.

Clip interconnect pressure sintering assembly processes

The clip assembly for HPE begins with printing of the sintered material onto the substrate. The example shown in **Figure 1** is a direct-bonded copper (DBC) substrate. It requires the application of pressure sintering paste to form a bond with the metal clip. The sintering paste comes in many formats,

with the most commonly used being paste and film—based on our observation with HPE package developer end users.

Figure 2 shows a typical clip pressure sintering assembly process that starts with the application of paste onto a substrate via a custom-designed and fabricated stencil. The challenges faced during printing are to avoid overprinting and smearing, as well as having to deal with the need for a two-step printing process. Printing is performed in two separate steps because the clip attached has a two-ended location – namely the DBC, where the source pad area is located and the top of the semiconductor chip – to complete the flow of electrical current. The printed area where the clip will be attached is a fraction of the area of the total die size. There exists a height difference between the source pad area and the substrate, therefore, merging a two-step printing into a single step would lead to insufficient printing pressure on at least one of the connecting points. Furthermore, fabricating a 3D stencil would be a challenge for a single-print process. The first printing is performed on the DBC, which we term here as the first connection on the source pad area (A), and the second printing (B) is performed on top of the semiconductor top as shown in **Figure 2**.

After printing, the next assembly step is nitrogen drying (see **Figure 2**, (C)).

The purpose of drying is to remove the solvent within the printed paste. A properly dried printed paste increases the paste evenness and firmness to avoid paste spluttering or roll-up during clip placement. The last step of the clip sintering assembly process is pressure sintering, where the two connection points – namely the source pad area and the top of the semiconductor chip – are pressure sintered with metal stamps. Optimal stamp design is critical within the sintering tool to ensure sintering pressure is evenly applied onto the interconnection of the clip and its contact point, i.e., the source pad and the top of the die. After the pressure sintering, the quality of the sintered clip bond is characterized by automated shearing or peel testing. The testing method employed is dependent on the design and size of the clip. If the thickness of the clip is sizable for the shear tool to contact during shearing, it will be used to evaluate the clip bond force, otherwise peeling of the clip would be the alternative—which is most commonly used. Scanning acoustic microscopy will be employed to check for voiding, delamination, and the uniformity of the pressure of the completed sintered bond.

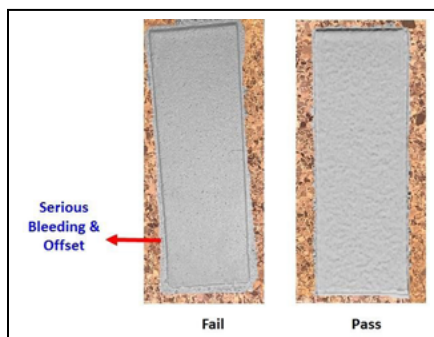


Figure 3: Difference between poor and good printing.

Exploring optimally-printed results

To obtain optimally-printed sintering paste, the following factors must be optimized: 1) paste viscosity at room temperature, 2) paste mixing and rollability, 3) printing speed, 4) printing force, 5) printing direction, 6) stencil frame spring force, and 7) squeegee release speed and distance. **Figure 3** demonstrates the difference between a poor printing outcome versus a good printing outcome for clip sintering, if printing parameters are not optimized.

Using local facilities for a compatible clip sintering process

To cope with the rapid growth of HPE applications, the manufacturing requirements of the whole production line should be considered. The clip sintering process for HPE needs to be compatible with all the upstream and downstream processes. Therefore, many manufacturers today are looking for local facilities to speed up the development processes and time to market.

Many semiconductor packaging solution providers will set up different local laboratories to update their customers on the initial development and characterization of the important processes for the complete HPE modules in manufacturing. It also provides a means to work with collaborative critical material (e.g., nano silver paste and ceramic substrate) suppliers for characterization and even qualification purposes. This action will help accelerate the industry's technology development and reduce the initial risk for setting up the production line.

Summary

Clip sintering is an alternative interconnect to heavy aluminum wire bonding when a HPE device or system is required to transmit high voltage or to allow high current to flow. In order to ensure a good pressure sintering bond for the clip, it is crucial to select a sintering paste that flows easily when applied. In addition to a

printable sintering paste, the design of an optimal printing area within the HPE will avoid issues such as smearing, uneven paste printing and offsetting of printed paste. Applying a uniform force and pressure during clip placement followed by pressure clip sintering is also a critical process parameter. The results of using optimal processes for each step of a clip-bonded HPE is a high-quality sintered clip bond that is robust enough to handle the influx of stress when subjected to high thermal load during actual operation of the HPE.

Acknowledgments

The author thanks Nelson Fan, Ding Jia Pei, Deivasigamani Mouleeswaran, Yuan Bin, Liao Jian, Tim Lu Fei, Wilson Kwok and Eugene Wee of ASM Pacific Technology Ltd.

Reference

1. A. Jadhav, S. Mutreja, "Electric vehicle market share, growth, size,

analysis 2022-2030," Allied Market Research, Jan. 2022. Retrieved Feb. 23, 2022, from <https://www.alliedmarketresearch.com/electric-vehicle-market#:~:text=The%20global%20electric%20vehicle%20market,electric%20vehicle%20industry%20as%20well>.

Biography

Eric Kuah, DBA, is VP of Technology APET 1-2 at ASM Technology Singapore Pte Ltd., Singapore. He has been with ASM since 1993 and his current responsibilities in the APET group include development of new technology for encapsulation packaging in the IC and LED areas, silver sintering, ToF, and active alignment for cameras, and managing a team of engineers in the various engineering disciplines from mechanical to materials engineering. He is a holder of more than 25 US patents. Email eric.kuah@asmpt.com

PLASMA ETCH
PROGRESS THROUGH INNOVATION

PLASMA IMPROVES BONDING!

Our fully automated plasma cleaners offer:

- Improved Markability
- Enhanced Adhesion
- Better Bonds
- Easier Assembly
- Surface Modification

100% removal of organic contaminants with a low environmental impact!

STARTING AT ONLY \$ 5,900 USD

TO LEARN MORE, VISIT WWW.PLASMAETCH.COM
OR CALL 775-883-1336

ADVERTISER INDEX			
Adeia	33	LB Semicon	13
Advantest	OBC	Leeno Industrial	IFC, 38
Amkor Technology	1	Micro Control Company	48
AT&S	41	PacTech	9
Brewer Science	44	Plasma Etch	52
Chip Scale Review	3	SEMI	50
CyberOptics	7	Smiths Interconnect	24
DL Technology	47	Sonix	45
E-tec Interconnect	46	SÜSS MicroTec	23
EV Group	2	Technic	42
FormFactor	14	Test Tooling Solutions Group	20
INTEKPLUS CO., LTD.	10	TSE Co. Ltd	4,5
Ironwood Electronics	31	Universal Instruments	19
ISC CO., LTD.	34	WinWay Technology	26
Johnstech International	IBC	Yield Engineering Systems	29

July August 2022 • Space Close July 5th • Ad Materials Close July 8th
For Advertising Inquiries • ads@chipscale.com



THE FUTURE STARTED WITH
JohnstechTM
SINCE 1991

FROM DC TO 100GHz+

Johnstech was the first to create a high-performance test contactor for production testing of IC devices.

- 30 YEARS OF DEVELOPING IP AHEAD OF MARKET NEEDS
- HIGH PERFORMANCE TEST SOLUTIONS
- RAPID LEAD TIMES WITH OUR RELIABLE SUPPLY CHAIN
- WORLD CLASS FIELD SERVICE AND SUPPORT WHERE YOU NEED US

FOR MORE INFO, VISIT JOHNSTECH.COM/FUTURE

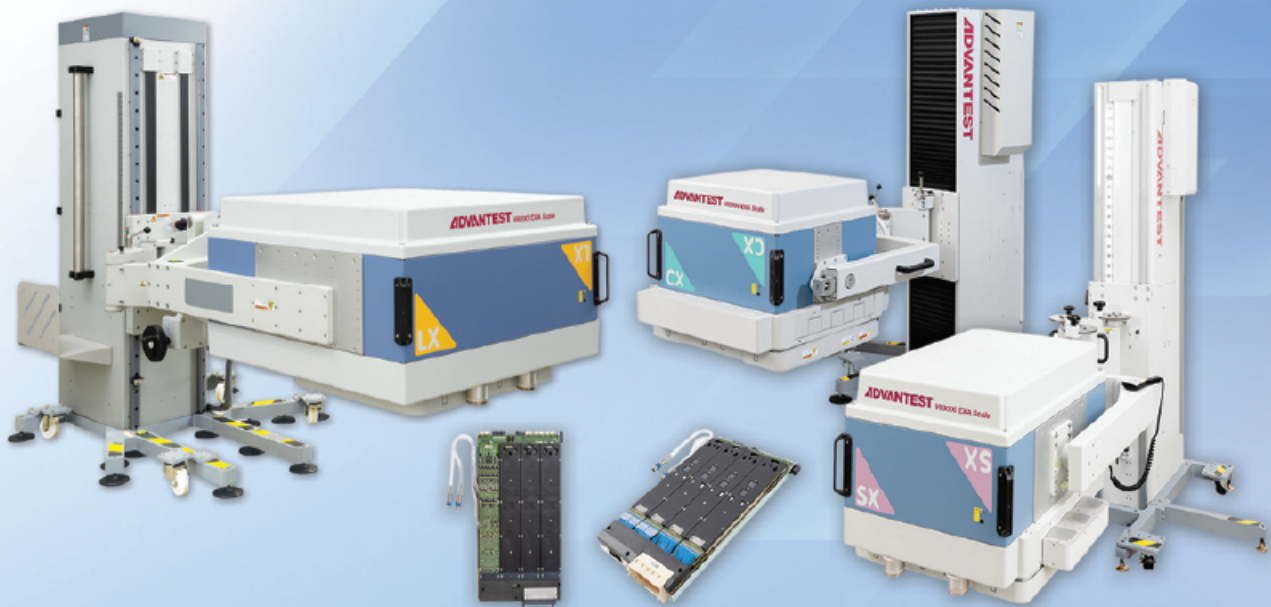
© 2022 Johnstech International Corporation



Advantest. Enabling the Age of Convergence and Exascale Computing.

V93000 EXA Scale

Power of Innovation. Strength of Scale.



A powerful synergy is taking place as high-performance computing intersects with artificial intelligence, causing a major shift in the evolution of semiconductor design. As the amount of data being processed grows exponentially and scalability creates new testing challenges, Advantest responds with the V93000 EXA Scale™ SoC Test Systems offering solutions targeted at advanced digital ICs up to the exascale performance class.

As technologies continue to converge, Advantest is enabling its customers to address Big Data and Smart Manufacturing challenges with innovative test solutions that ensure superior performance of their most advanced device designs, and is helping them to quickly bring those products to market with the greatest cost efficiency.