Chiplets

Designing, manufacturing and testing.

A Semiconductor Engineering Research Report: 2023

SEMICONDUCTOR ENGINEERING
DEEP INSIGHTS FOR THE TECH INDUSTRY

Table of Contents

Introduction 3

Why chiplets 5

What is a chiplet 7

How chiplets are packaged 9

SIPs, MCMs 11

2D, 2.5D, 3D-IC 12

2D 12

2.5D 13

3D-IC 14

Connecting chiplets 16

Design issues 21

Teams become larger 21

Physical issues — stress, materials, thermal, aging 24

Inter-die connectivity 25

IP and IP reuse 28

Security 32

Assembly and test 38

Assembly, interconnect 38

Test 39

Standards 43

What needs to be standardized 43

Who is working on what standards 44

Who is working on ecosystems — mini-consortia 45

Government gets in the act 47

Commercialization of chiplets 50

Startup funding 53

Chiplets 53

Packaging 57

Conclusion 60

Appendix 61

Terminology 61

Sources, references, further reading 61

Relevant staff-written articles 61

Blogs 63

Videos 63

Technical papers 64

White papers 65

Introduction

"It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected. The availability of large functions, combined with functional design and construction, should allow the manufacturer of large systems to design and construct a considerable variety of equipment both rapidly and economically."

— Gordon Moore, "Cramming more components onto integrated circuits," Electronics, Volume 38, Number 8, April 19, 1965.

As Moore's Law and Dennard scaling reach their limits, new semiconductor architectures are being used and developed to increase performance, reduce time-to-market, save power, and ultimately reduce costs.

This shift has been in the works for at least the past decade, and it has been discussed in one form or another since the introduction of Moore's Law in 1965. But the cost and predictability of device scaling were so attractive that the chip industry didn't take advanced packaging and multi-chip architectures seriously until the past decade, when the laws of physics made it economically unfeasible to continue shrinking anything but the most advanced digital logic.

Those laws are immutable. The only way to achieve big improvements in performance, power, and area/cost going forward is with heterogeneous, domain-specific architectures, not by shrinking features. This has been proven repeatedly since the 28nm node, and the concept has been gaining steam in applications ranging from data centers to smart phones.

While some functions work better using older nodes — particularly analog and those operating under extreme temperatures or stress — others require more digital transistors to achieve sufficient performance. Many devices must process and move huge amounts of data faster with very low

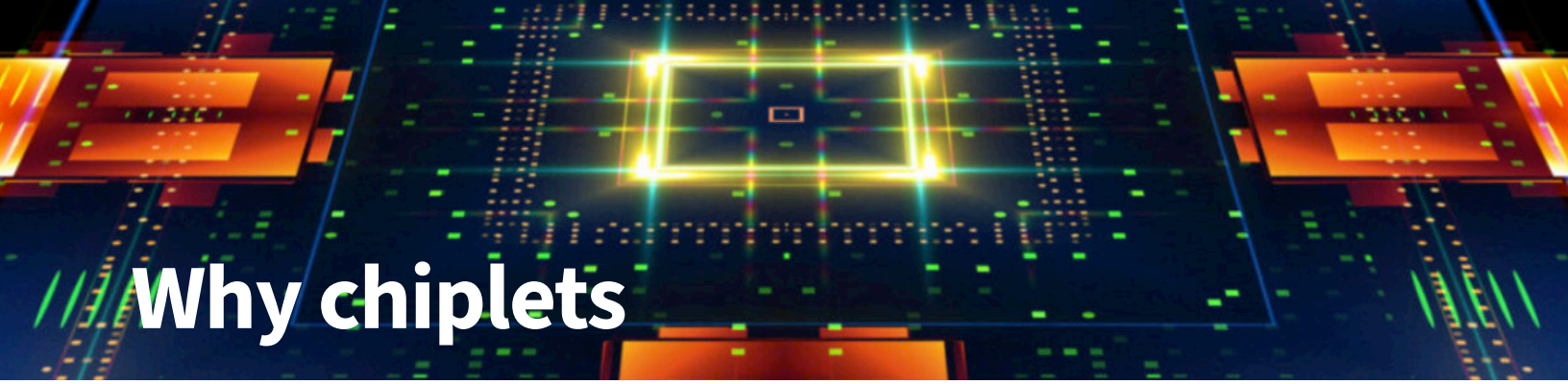
latency. Continuing to improve performance in light of physical limitations no longer can be done using a single planar die.

Chiplets promise the best of both worlds. A single advanced package can include 5nm or 3nm chiplets (sometimes called tiles or dielets, depending on a vendor's marketing lingo) for tasks that need faster processing, or a mix of standard or custom accelerators for specific data types. They also can include chiplets developed at mature nodes for analog or less important functions. Put simply, chiplets offer customization, optimization, and scalability, and they can be assembled more quickly using a standard interconnect scheme such as Bunch of Wires (BoW), or Universal Chiplet Interconnect Express (UCIe), which is an extension of the Peripheral Component Interconnect Express (PCIe) standard that was developed and introduced by Intel in 1992.

But chiplets also add new layers of complexity, such as how to design, test, inspect, measure, and package what is essentially hardened IP from multiple manufacturers, and there needs to be a predictable and repeatable way to ensure interoperability with other components in a package. This has limited their use to a few large industry players that have developed their own chiplets.

Turning this into a commercial opportunity will be more difficult. It will require cooperation and information sharing at many levels across the industry.

This ebook examines what chiplets are, what they are being used for today, and what they will be used for in the future. And it looks at the hurdles they face, the technologies involved, and how standards and ecosystems are forming in these early days of the technology.



Why chiplets

When the integrated circuit (IC) was first developed, it was a single, monolithic device. Every couple of years the various components packed into a fixed area — logic, SRAM, I/Os — were shrunk down for a new manufacturing process or process node. But cramming more transistors, memories, and other components into the same physical space on a planar substrate has been limited by physics for the last decade, as evidenced by the reduced PPAC benefits at each new node. The rising costs are a reflection of the challenges in manufacturing 3D transistors, whether those are finFETs, gate-all-around FETs starting at 3nm, or complementary FETs at some future node. Issues involving dynamic power density, current leakage, higher operating temperatures, and a host of physical effects are piling up at each new node.

While process technology scaling will continue — current roadmaps by TSMC, Intel Foundry Services, and Samsung Foundry — extend well into the 1 nanometer/10 angstrom range. But it no longer makes sense to put everything on a single die. For one thing, there are so many features being incorporated into devices these days that they literally will not fit on a single SoC because it exceeds the size of the reticle. Rather than "stitching" together large SoCs, it makes more sense to disaggregate a system down into smaller chips, or chiplets, each with a very specific function, and manufactured at whatever process node makes sense. This is particularly useful for analog components, which do not benefit from scaling. And it means that not everything has to be developed from scratch for every application, because it can be re-used regardless of which process node is used for the digital logic.

This is why there is so much attention being focused on chiplets. Using a consistent interface scheme, IP can be hardened and assembled into packages in LEGO-like fashion, and it can be manufactured at whatever process node makes sense for that particular IP. If it works as planned, the chiplet approach can speed time to market, reduce costs, and allow much more customization of designs for specific domains. In effect, it provides a way to mass-customize portions of a design, and this approach has been shown to work quite well by Marvell, AMD, Intel, and IBM using internally sourced chiplets.

That gives design teams flexibility to make tradeoffs about which parts they want to shrink, which ones they don't, and it gives them more time to innovate on the critical features and where there is real added value.

Why chiplets

“What we are talking about now is an economy of design,” said William Chen, ASE fellow and senior technical advisor. “Previously we talked about an economy of scale. Now we are thinking about economy of design meeting the needs of economy of scale.”

With chiplets, the architect or design team can explore multiple options based on cost, power, performance, and market segment. Those chiplets can be highly targeted, or more general, such as a standard I/O. So while the final device is not a general-purpose chip, it may be comprised of some generic components.

Customization using proven silicon is a key benefit to the chiplet concept, along with keeping the ability to choose various footprints, voltages, clock speeds, redundant circuits, etc.

Companies that successfully use chiplets today have created what is essentially a chassis with interchangeable parts. Every architecture has its own blueprint for saving power, dissipating heat, and reducing latency, which can include everything from external and embedded bridges to vertical stacking to reduce distances that critical signals must travel.

What is a chiplet

A chiplet is a die/dielet/tile that is designed to be integrated with other chiplets in some type of advanced package. That package could be a system-in-package (SiP), 2.5D (planar or vertical), or 3D-IC.

Nearly all chiplets in use today are being developed by big chip companies, which design the chiplets with some sort of standard interconnect or protocol and re-use them across multiple designs. That is likely to change as more standards are developed by the chip industry, and as more mini-consortia begin working through all of the possible problems for their particular market segment.

Logic, memory, sensors, MEMS, mixed-signal, or any one of a variety of different functions can be developed as chiplets. And a 28nm chiplet may be placed next to, and work together with, a 5nm chiplet. The concept is basically to shrink a printed circuit board and put all the pieces into a much smaller space. It also allows more features to be added to a design, which in the past required different dies to be connected. That increasingly exceeds the size of a reticle, which 858mm^2 with EUV steppers, and 429mm^2 with high-NA EUV steppers.

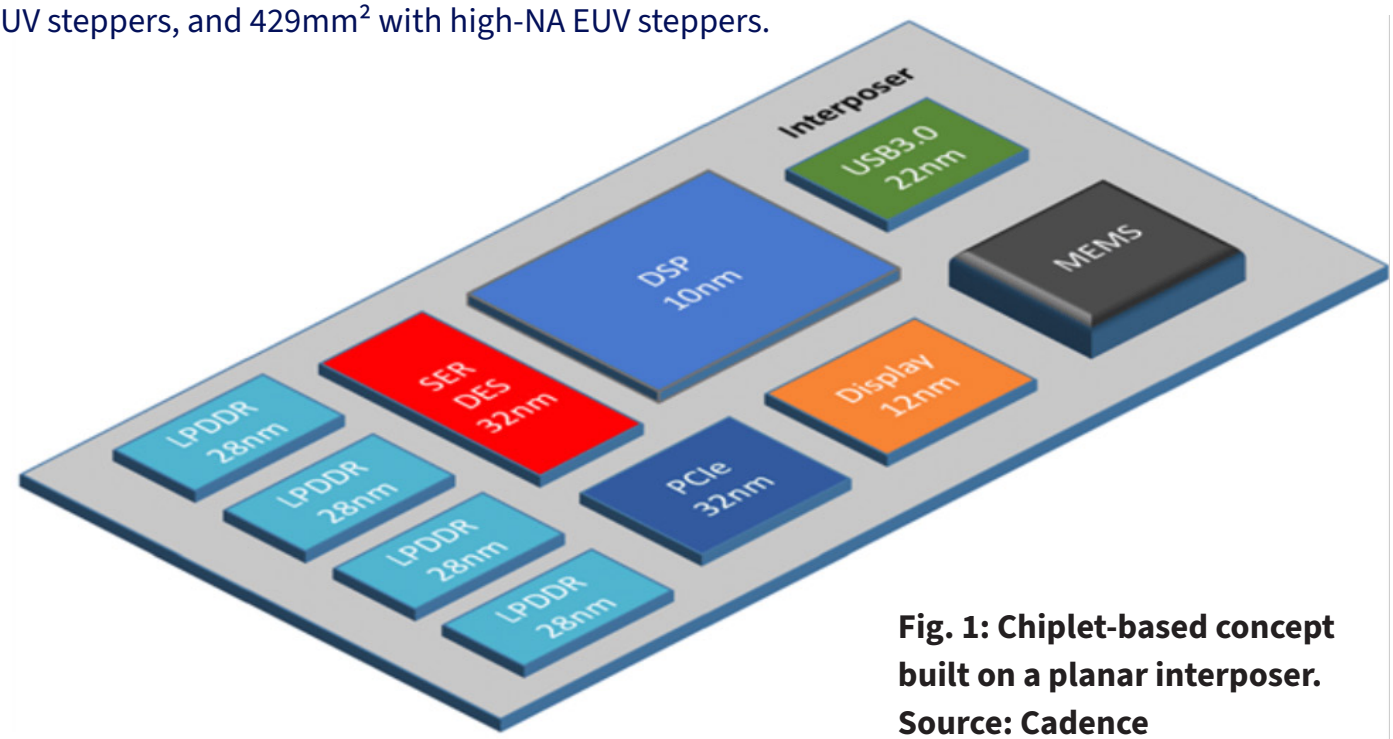


Fig. 1: Chiplet-based concept built on a planar interposer. Source: Cadence

Using chiplets gives the IC design team or OEM the option to take the best — or most appropriate — version of each function from the most expert designer/manufacturer, and to integrate it into some type of package. It's often called a mix-and-match, or LEGO-block type of concept, and the chiplets can be selected from a menu of options.

"The biggest problem is to establish an ecosystem," said Andy Heinig, head of department for efficient electronics in Fraunhofer IIS' Engineering of Adaptive Systems Division. "Packaging, chip design, or IP are not the biggest issue. It's more about getting everything in place. This is the real challenge today. Then we can discuss performance, because any chiplet solution will have to compete against existing solutions. It's a bit of a chicken-and-egg problem, though, because you need performance — especially because the chips are more expensive than an SoC solution — but you can't focus on performance because you have to get it running first.

The hard part is understanding how different chiplets may interact with each other, and how they will behave under different use cases, which requires very detailed characterization. Even with the best simulation tools, however, there frequently is insufficient data to predict how a chiplet will behave in the context of a system. It can vary greatly by application, by use case, and over time, which is why there is so much attention paid today around thermal mapping and management, various types of noise and stress, and inconsistent aging, all of which can affect manufacturing yield and long-term reliability in the field.

On the other hand, chiplets have been proven to work extremely well in controlled situations. AMD and ASE have been collaborating on chiplets for graphics subsystems for the better part of a decade. Marvell has used chiplets commercially since 2016. And Intel is customizing systems for its data center customers based on chiplets.

The challenge now is to develop chiplets and chiplet-friendly architectures that can be sold commercially for any vendors' devices.

How chiplets are packaged

Packaging is usually the last step in IC manufacturing, but with chiplets it needs to be considered in depth up front in the design cycle.

The package encases and protects the die or dies inside the package and connects them to the outside world. Usually the packaging is done by an OSAT (outsourced semiconductor assembly and test) company after the fab process and the interconnect process, also known as back-end-of-the-line (BEOL).

But packages also must be developed specifically to handle the characteristics of the chiplets within them, so packaging has become a much bigger decision for chipmakers than in the past. Consequently, the OSAT may take on more responsibility and be involved with the design team much earlier in the process. Also the fab may take on more of the BEOL interconnect work. The assembly, test, and packaging (ATP) may be done by the foundry, the OSAT, or some combination of both.

TSMC also has proposed etching chiplets into a die at the front end of the line, a process it calls System on Integrated Chips (SoIC). That approach relies on hybrid bonding, which is an interconnect approach initially used for image processors. Hybrid bonding is a permanent bond that combines a dielectric bond (SiOx) with embedded metal (Cu) to form interconnections.

Today, chipmakers can leverage the chiplet model and integrate dies in existing advanced package types, such as fan-outs and 2.5D. There also are options to stack logic-on-logic, or memory-on-logic, in a 3D-IC.

But the optimal solution for one application may be very different for another, which is the whole point of chiplets.

“The question is what are the devices we are targeting,” said C.P. Hung, vice president of R&D at ASE, during a panel discussion at the 2021 IEEE Electronic Components and Technology Conference (ECTC). “With multiple chips, you have to look at I/O density. We can handle that in flip-chip. If that’s not enough, we can consider fan-out. If we have multiple memories that need to be integrated, we may need to use 2.5D.”

How chiplets are packaged

Synopsys explains on its website that when designing multi-die SoCs, system architects are faced with multiple packaging options. Among them:

- 2D packaging, where dies are assembled on an organic substrate and laminate;
- 2.5D packaging, where an interposer using silicon or redistribution layer (RDL) fanout is used to route the signal between dies in the SoC;
- 3D packaging, where dies are stacked vertically using hybrid bonding techniques, or
- Some combination of the above.

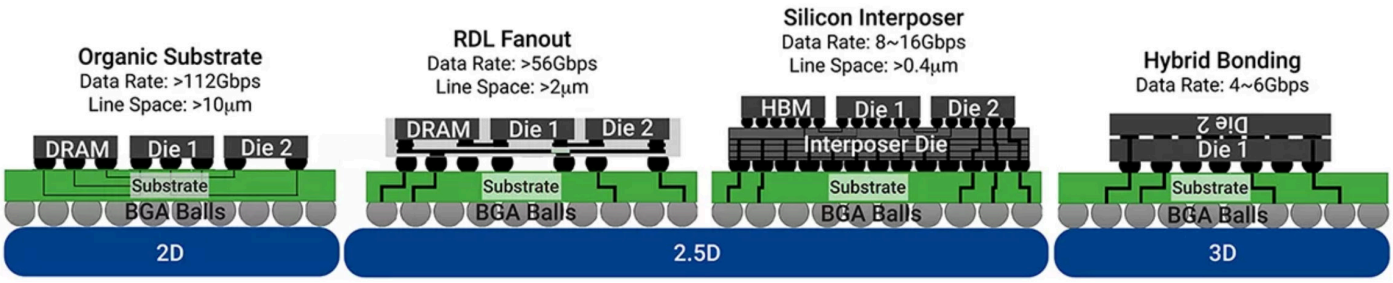


Fig. 2: Packaging options available to designers. Source: Synopsys

Choices of package can depend on a variety of physical factors, such as speed, mechanical stability, voltage, heat, bonding and interconnect options, and cost.

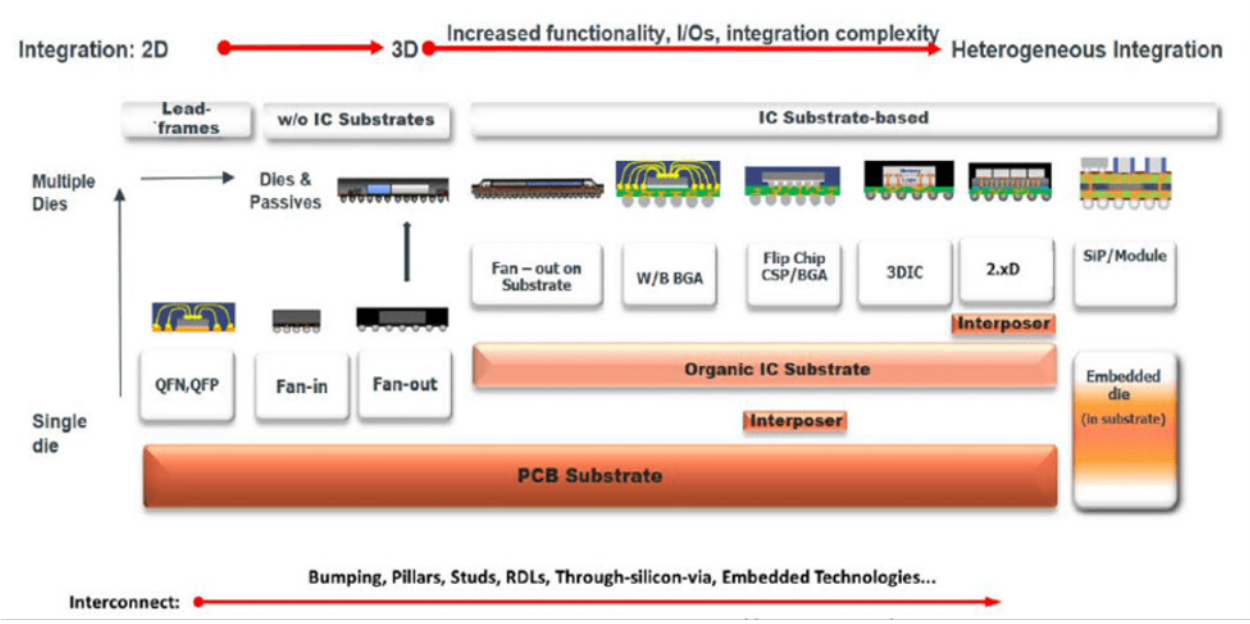


Fig. 3: Different packaging technologies. Source: IEEE/hir/Fraunhofer IIS EAS

SIPs, MCMs

A system-in-package (SiP) is an umbrella term for an integrated circuit (IC) package with multiple dies, as opposed to a system-on-chip (SoC) that has all the system blocks on one die. The SiP disaggregates all the blocks into separate dies/chiplets on the same substrate, but they are connected to function as one IC. In general, a SiP can be thought of as using hard IP, while an SoC uses soft IP.

But a SiP also is a particular configuration. It is meant to be a lower cost, smaller-form-factor package with multiple chiplets in it.

The precursor to SiPs were multi-chip modules (MCMs), which date back to the 1980s. They were basically two or more chips in a basic package. The terms are often used interchangeably, which has created some confusion in the market.

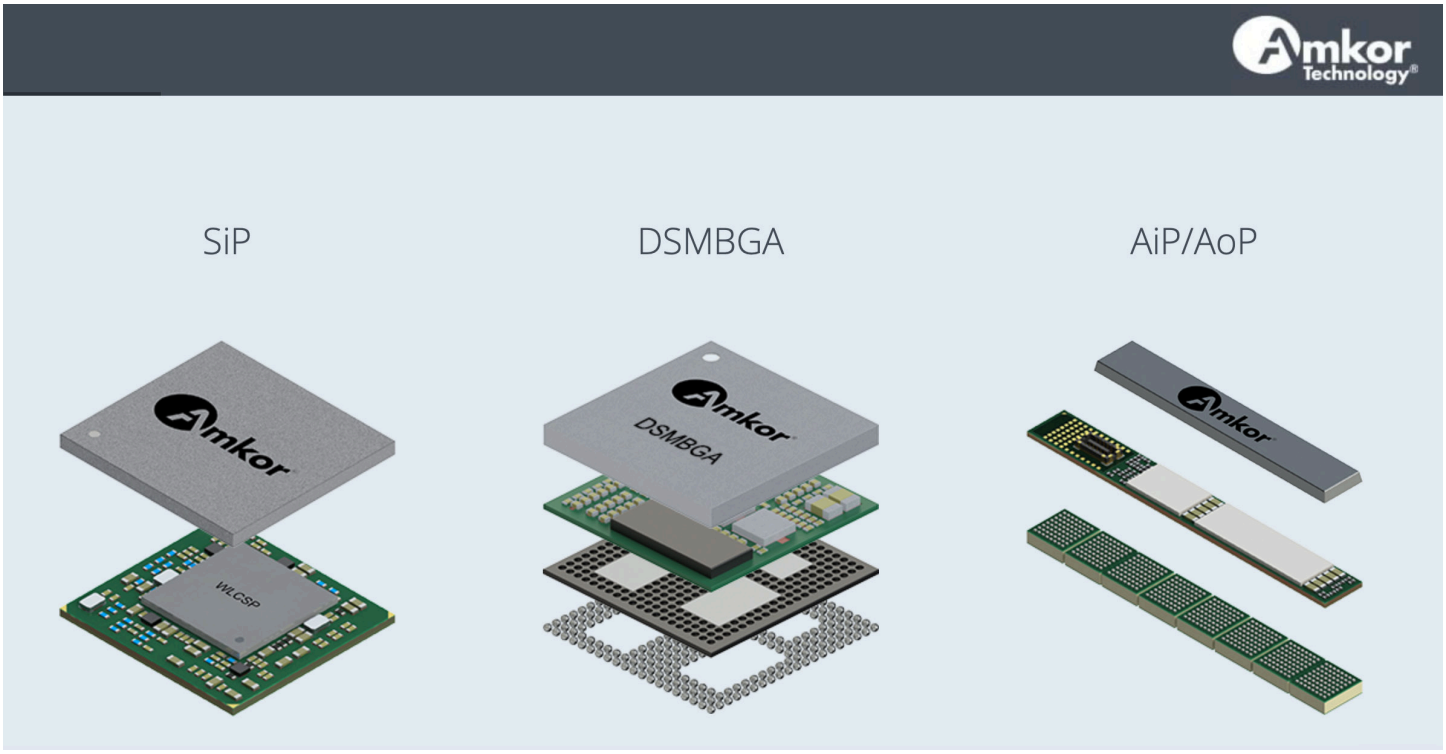


Fig. 4: Multiple packaging options, including SiP, double-sided molded ball-grid array, and antenna-in-package and antenna-on-package. Source: Amkor

2D, 2.5D, 3D-IC

Advanced packaging has made the chiplet concept possible, but each type of package has it benefits and issues.

2D

In 2011, Xilinx broke apart its Virtex 7 chip into four parts connected to an interposer on a planar substrate. The company said at the time that its goal was higher yield, because it's easier to manufacture a smaller chip with high yield than a large chip. So by breaking apart a large FPGA into smaller piece it would be able to reduce the cost, and with a silicon interposer it would actually increase performance.

For the next decade, chipmakers and packaging houses began experimenting with ways to improve performance and simplify packaging, essentially settling on some version of a fan-out packaging approach. The idea is that given the limited number of pins on an SoC, the various components could be spread out to simplify layout and packaging.

To make fan-out packages, dies are placed in a wafer-like structure using an epoxy mold compound. The RDLs are formed. The individual dies are cut, forming a package. But there are some challenges here, so it's not as simple as it sounds. First, dies or chiplets can shift when they are placed into the RDL. And second, fan-outs are typically slower than other approaches, which is why there has been so much focus on high-density fan-outs, which reduce lines and spaces to 2nm or less.

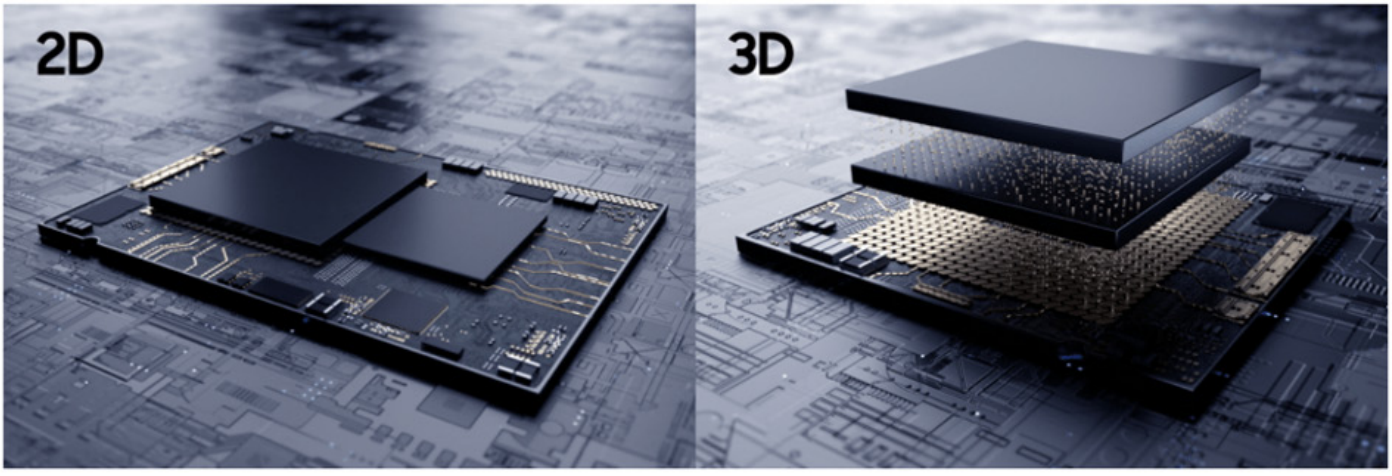


Fig. 5: 2D/3D-IC packaging approaches. Source: Samsung/Fraunhofer IIS EAS

2.5D

With 2.5D integration, different functions are placed side-by-side next to an interposer, or stacked vertically on top of that interposer.

The approach typically has been widely used in data centers, where performance and low power are critical and where there is some cost resiliency. Communication between chips is accomplished using either a silicon or organic interposer, typically a chip or layer with through-silicon vias for communication.

While communication between chips is slower than between different components on a single die, the distances that a signal travels actually may be shorter, and there are more conduits for signals. So collectively, data moves faster and requires less energy to move. On a single chip, skinny wires also slow data movement, and increased resistance adds heat.

2.5D architectures have been paired with stacked memory modules, particularly high-bandwidth memory, to further improve performance. The downside of this approach is that the interposers are expensive, and IP needs to be developed specifically for this architecture.

Fig. 6: A model showing a system in package consisting of chiplets from differing nodes on an interposer. One chiplet shown in orange is 3D with at least two chiplets stacked vertically. The two chiplets shown in red are identical and represent using multiple identical chiplets to scale up their function. Source: Semiconductor Engineering, 3D-printed model by Michael Posner/Synopsys

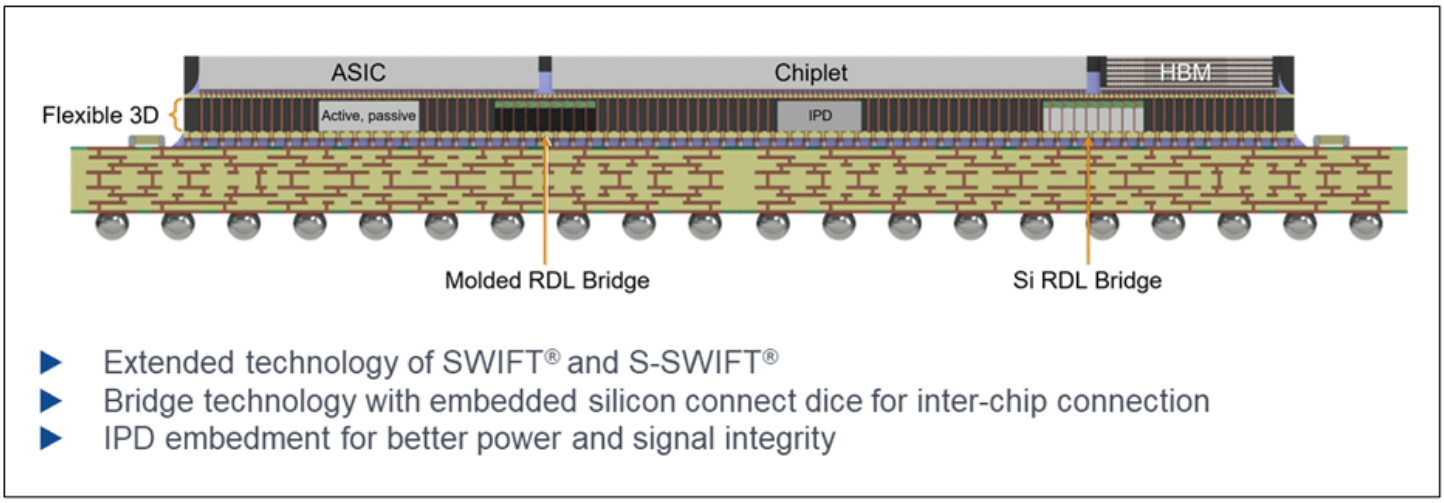


Fig. 7: Amkor's S-Connect technology. Source: Amkor

3D-IC

With Moore's Law, the migration from one technology node to another invariably had new effects and new restrictions or limitations to ensure an easy path to sign-off. With that came new opportunities that could be exploited. The migration from 2.5D to 3D makes those previous migrations look simple. 3D-IC is disrupting every aspect of the tools, models, flows, and even organizations.

3D-ICs have logic and sometimes memory stacked on top of each other, and sometimes logic-on-logic. 3D configurations helps somewhat by reducing dynamic power over the 2.5D because the connections are shorter on 3D. Nevertheless, while placing logic-on-logic may sound like a small step, several problems must be overcome to make it a reality.

True 3D involves chiplets stacked on top of each other in a highly integrated manner. The biggest concern has been thermal dissipation, but that seems to be increasingly under control because less energy is required to push signals shorter distances than with other types of packaging or on planar die. That, in turn, improves overall efficiency and performance, but it makes it even more critical to understand potential interactions between chiplets.

"There will be some test chips happening toward the latter part of this year that involve full logic-on-logic," said Vinay Patwardhan, product management group director in the Digital & Signoff Group at Cadence. "By the middle of next year, we can expect some realistic logic-on-logic chips, particularly something with multiple AI

cores. Those companies are running out of area on a single die. Many of these designs are close to the reticle size limit, something over 600 or 700 millimeters squared. They are desperately trying to go full-3D stack for the next-generation design because it does not require too much architectural change. But cutting them and stacking them is a physical change."

Other targets for 3D may be further out, such as the ability to stack heterogeneous dies. "That would require a true 3D placer and router that was working on a heterogeneous stack," said Rob Aitken, a distinguished architect at Synopsys. "It would have to know that to build any logic path that crosses the dies, you'd need two separate libraries. They could quite well be two different technology nodes, but existing tools and flows make the assumption about consistent libraries, and those assumptions are baked quite deep. It's not inconceivable that the tools can be modified to deal with that. But some base assumptions of the tools would need to change."

Connecting chiplets

There are proven ways of connecting chiplets that work, such as UCIe, Bunch of Wires (BoW), silicon interposers, bridges, and even hybrid bonding. In the future, it's likely that more than one of these approaches will be used in complex designs, opening the door for much more innovation.

Chiplets are connected to each other, and at least one of them (if not all) connects to an interposer or substrate.

Moving from separate chips that are packaged and placed on a PCB, to a package that integrates multiple dies, dramatically changes the interconnect schemes.

"A traditional ASIC has large I/O drivers necessary to drive signals through the package, board, and external interfaces," said Tony Mastroianni, advanced packaging solutions director for Siemens EDA. "This could range from tens of millimeters to several meters. 2.5D die-to-die interfaces deploy smaller I/O drivers that are only required to drive horizontal connections to adjacent die through the interposer, which may be on the order of tens to hundreds of microns. 3D die-to-die interfaces deploy even smaller I/O drivers, which are only required to

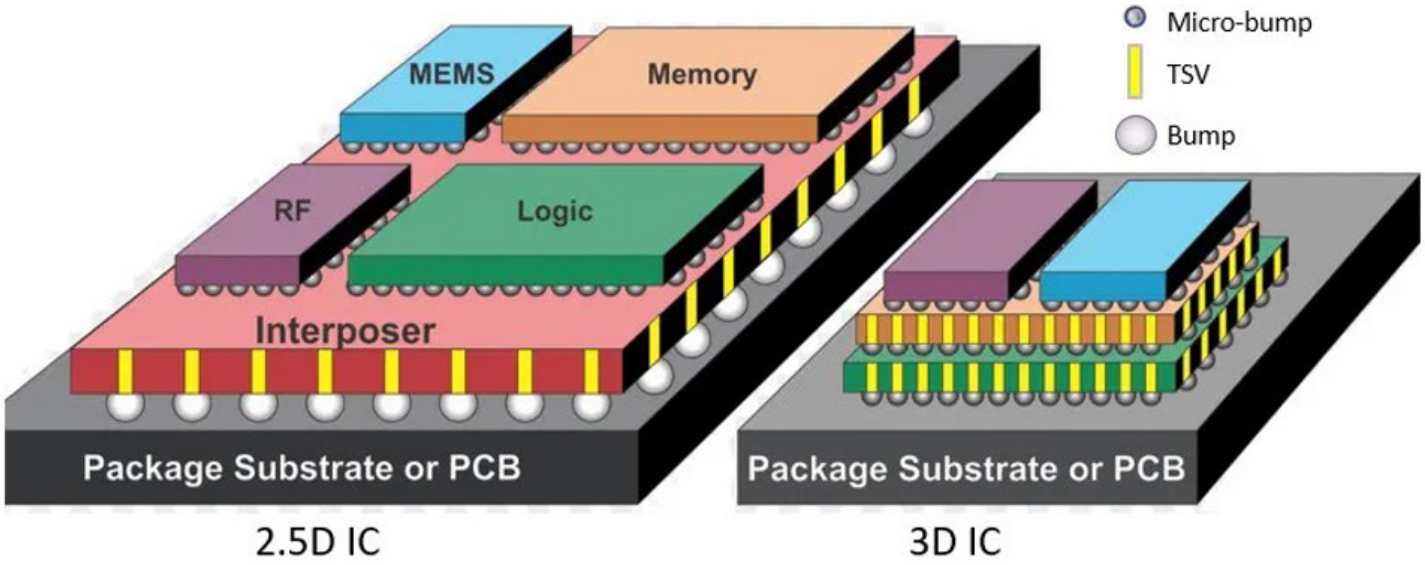
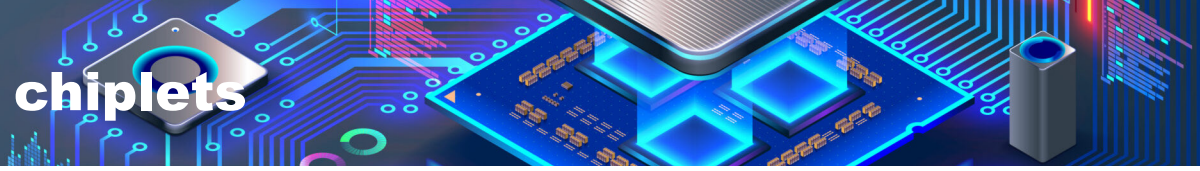


Fig. 8: 2.5D versus 3D-IC; Chiplets connect to interposers through bumps. Source: Siemens EDA



drive vertical connections directly to the die stacked above or below. Those may be on the order of a few to hundreds of nanometers. The reduced drive strength and shorter trace lengths inherent in the 2.5 and 3D approaches enable dramatic reductions of power and increased I/O bandwidth, which offers orders of magnitude of improved energy efficiency (pJ/bit)."

There are several options available at this point. "One method of chiplet integration avoids the use of fine-geometry interconnect altogether," said Brian Holden, vice president of standards for Kandou. "With this method, the interconnect between the chiplets is simply the organic package substrate. This avoids complex manufacturing processes, and the extra cost and yield loss associated with silicon interposers. Low-power ultra-short reach (USR) SerDes are used to enable high-speed interconnect between the chiplets."

With packaging, as with chiplets, there are numerous considerations and tradeoffs.

"When you disaggregate your die into multiple dies, you can put it either on a substrate or put it on an interposer," said Ketan Mehta, senior director of SoC IP product marketing at OpenFive. "That creates a big distinction. With an interposer you can do thousands of signals, whereas with the substrate you can only do a few hundred at most. If the customer is implementing a large die that also has HBM, for example, they have no choice but to implement it on an interposer. That leads you toward a parallel interface because the interposer will accommodate thousands of signals."

Intel has developed its own chiplets strategy using multiple elements such as its Embedded Multi-die Interconnect Bridge (EMIB). Instead of using a large silicon interposer typically found in 2.5D approaches, EMIB uses a very small bridge with multiple routing layers. This bridge is embedded as part of the substrate fabrication process.

There is an ongoing debate in the standards community between which is better, parallel or serial interfaces. Each has pros and cons.

"What customers really care about are lowest possible latency, lowest possible power, bandwidth for beachfront, performance in terms of reach, and then cost, which is basically the yield," explains Manmeet Walia, senior product manager for high-speed SerDes at Synopsys. "Serial connections use very lightweight SerDes. They have minimalistic PHYs, and you don't need any decision feedback equalization — simply DLL clock-based forwarded approach."

The serial standards are being driven by the Optical Internetworking Forum (OIF), and are referred to as 112G USR, or extra short reach (XSR) links.

On the parallel side, there are several standards efforts. One is Open High Bandwidth Interconnect (OpenHBI), which is being driven by the Open Compute Project's (OCP) Open Domain-Specific Architecture (ODSA) sub-project. The OpenHBI parallel die-to-die standard may enable a way to drop the SerDes.

"Parallel die-to-die interfaces fundamentally assemble a large number (thousands) of I/O pins that drive single-ended signals across dies," wrote Manuel Mota, senior staff product marketing manager at Synopsys. "Because the data rate per pin is only a few gigabits per second (Gbps) — from 8 to 16 Gbps — and the distance between dies is only a few millimeters, from 3 to 5mm, the drivers and receivers can be simplified while achieving system Bit Error Rates (BER) well below the 1e-22 to 1e-24. The system BER can be met without requiring additional error correction mechanisms, such as forward error correction (FEC) and retry, which add link complexity and latency. By simplifying the I/O, eliminating the serialization and deserialization (SerDes) steps, and avoiding very-high-speed signaling, parallel die-to-die interfaces can achieve remarkably high power-efficiency and low-latency while supporting very high throughput across the link. For these reasons, parallel die-to-die interfaces are very appealing to SoCs for high-performance computing applications that are not tightly constrained by packaging cost and assembly."

Intel, meanwhile, has developed the Advanced Interface Bus (AIB). "The specification for AIB 2.0 is already in the CHIPS Alliance GitHub," said Jose Alvarez, senior director in the CTO Office for the Programmable Solutions Group at Intel. "It is work in progress, and very close to being released. Our goal is 4 gigabits per second per wire, a total of about 7.6 terabits per second of bandwidth per interface. But it's not just about bandwidth itself. It's about energy efficiency. Today we are at 0.85 picojoules per bit of energy utilization. We went to 0.5 picojoules per bit, and the DARPA PIPES program wants to push this to a 0.1 pJ/bit. That's a much longer horizon, but we are leading toward that."

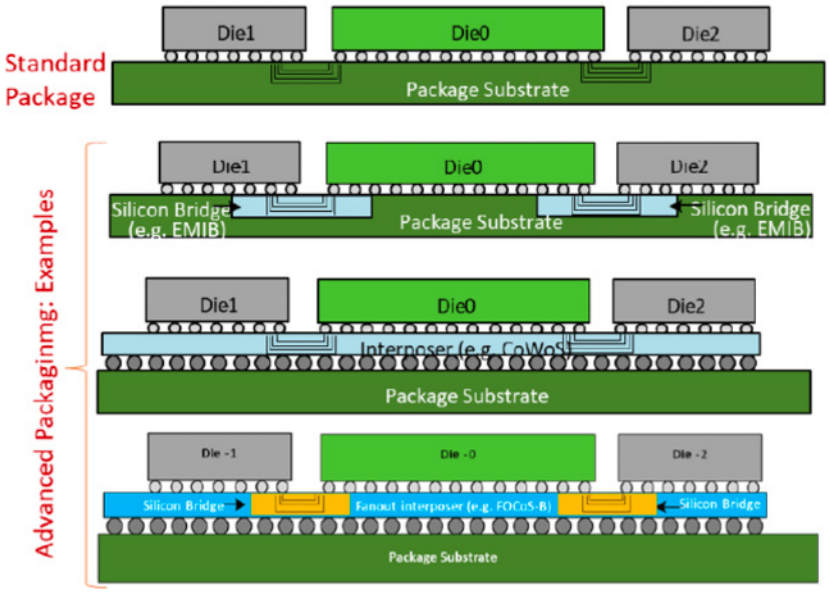


Fig. 9: UCIe open chiplet ecosystem.
Source: Fraunhofer IIS EAS/UCIe/Chiplet Summit

Some companies have deployed an approach called Bunch Of Wires (BOW). A November 2020 press release from GUC showed some performance numbers for that interface and demonstrates some of the performance tradeoffs. It quotes error-free communication between dies with full duplex 0.7 Tbps traffic

per 1mm of beachfront, consuming 0.25 pJ/bit. GUC believes the next generation will support 1.3 Tbps error-free full duplex traffic per 1mm of beachfront, with the same 0.25 pJ/bit power consumption using TSMC 5nm process.

GUC said that power consumption for a parallel connection is 6 to 10X lower than alternative solutions using ultra-short reach SerDes-based communication through package substrate.

For example, Eliyan, a startup that develops chiplet interconnects, has focused on eliminating the interposer from UCIe-compliant designs by building a physical interconnect layer (PHY) on both sides of a chiplet. "That gets rid of any complexity for manufacturing, thermal management, and allows us to bank all the things we've learned from the old MCM (multi-chip module) days," said Patrick Soheili, co-founder and head of business and corporate development at Eliyan. "We have plans to build a bunch of chiplet stuff, using our technology as a basis. So we would have our technology on one end, and some other stuff on the other side, and connect two, three, or four things together. Maybe they're HBM devices, maybe they're other I/O controllers."

The number of possible schemes for connecting devices together is growing rapidly. In late 2022, TSMC introduced its 3D Fabric Alliance to connect different layers and devices in a 3D package.

"We have EDA, IP, and design services, and we're also adding memory partners, OSATs that help us assemble these devices, and the substrates, which become extremely important in 3D, as well," said Dan Kochpatcharin, director in TSMC's Design Infrastructure Management Division. "These devices can be 10 centimeters high, and the substrates can be 20 or more layers. So, we need to make sure we align their roadmaps with our roadmaps so we can interface with them, and maybe there will be different materials that work together. And then you have to think about

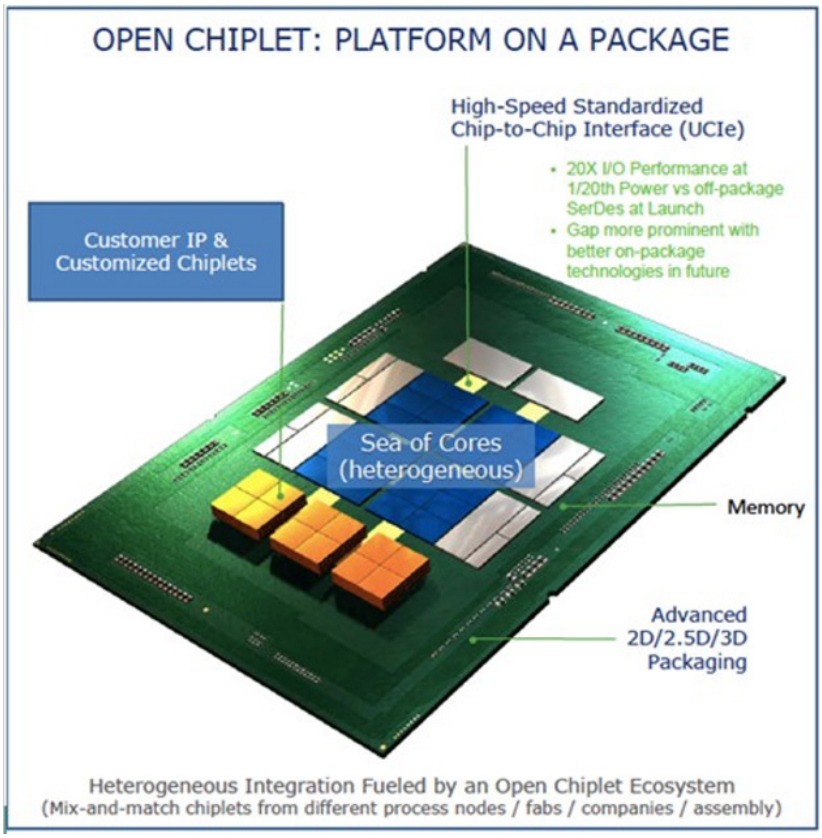


Fig. 10: UCIe heterogeneous integration with chiplets. Source: UCIe

testing the whole system, which is not easy. So we're working with Advantest and Teradyne, and also the EDA vendors. And IP is important in tests because we need to design for reliability."

And that's just the beginning of some of the generalized integration schemes emerging. There will be many more before this marketplace gets sorted out, and there will be an increasing number of proof points about what works and what doesn't, as well as some new issues that have never been considered. For example, uneven aging in chiplets can cause all sorts of reliability problems that have never been tackled before, particularly in markets where devices are expected to remain functional for years. Cost savings in one area may be offset by cost increases in another as the economics of chiplets and packaging evolve, and costs that customers pay today may become less attractive as the chiplet model evolves.

"Today's protocols do not scale well to multi-die interconnectivity," said Synopsys' Posner. "Up to this point there has been many proprietary standards — also, proprietary interfaces used to connect those die. To move this to mainstream, you need a more public open standard. There's been a few — XSR, openHBI, and the latest being UCIe. "

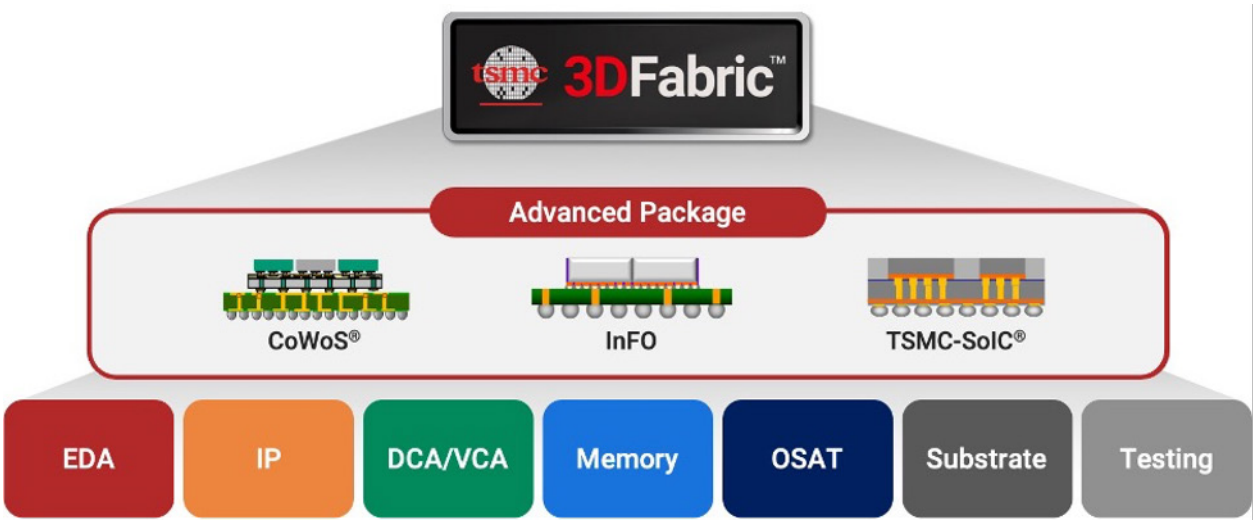
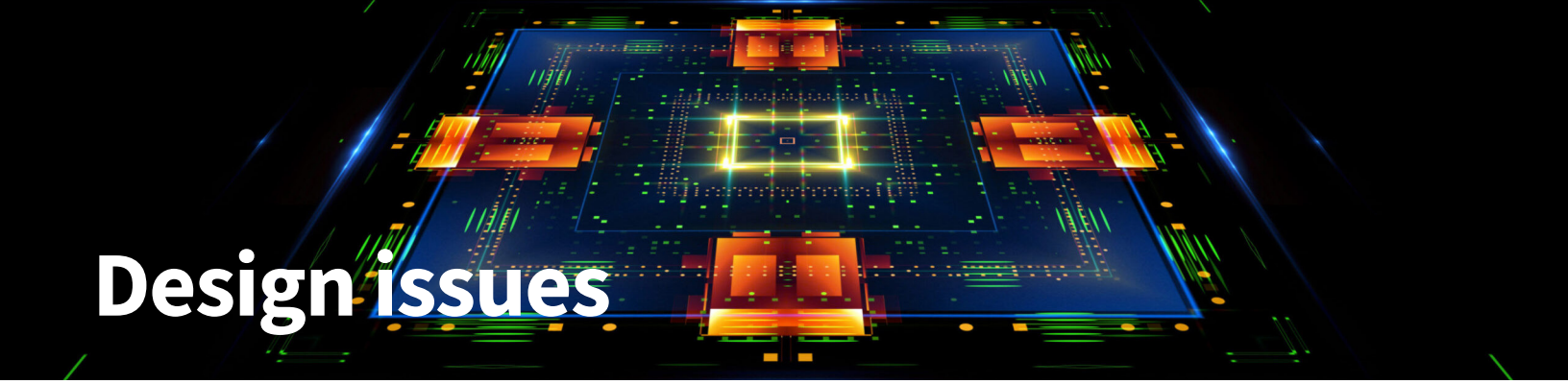


Fig. 11: TSMC's 3DFabric ecosystem model. Source: TSMC



Design issues

The complexity of chiplet systems, 2D, 2.5D, 3D, or some mix of those, is making the design both different and difficult during these early days. The entire ecosystem — design, test, manufacturing, packaging, and in-field monitoring — will require modifications to tools and methodologies, if not outright replacement.

At a starting point, designers will need to look at a multitude of issues. At ECTC in 2021, Bryan Black, a senior fellow at AMD, outlined the design considerations and challenges with chiplets:

- How to partition the dies in a system
- Design reuse
- Managing parametric variation
- Power delivery
- Interconnect speed
- Partitioning overhead
- Global clocking
- Die security
- Thermal management

Teams become larger

IP reuse via chiplets appears to be an effective and feasible solution, and a potentially low-cost alternative to shrinking everything to the latest process node. But that reuse — and just the chiplet concept, in general — means more people are involved in the design now. Multiple parties drive a chiplet project, not just the system or packaging or SoC architect. It's now all of the above, and more.

"It used to be the system designers would give a spec to the RTL guys, then the RTL guys would write the code and maybe work with the ASIC guys to pick the IP, then the packaging guys would get involved," said Tony Mastroianni, advanced packaging solutions director at Siemens Digital Industries Software. "It was very decoupled. Now, everything is coupled. The architects are working with the RTL architects and the package designers. When you're designing these chiplets, there may be six or

seven designs being done in parallel. You can think of them as blocks, but they're not being connected in the top level of the chip. They're being connected in the package, so the package guy is in the middle of that. That means they need to look at floor-planning that system-in-package. They're making sure all the I/Os are aligned on those high-speed internal interfaces. And they have to worry about thermal coupling between the chips. That's something that you don't normally do in ASIC design. You just look at the package."

The challenges are significant for chiplet design today. Companies working in this space typically have a lot of experience designing advanced ASICs.

"Since there is no ecosystem of chiplets, they need to be doing multiple chip designs concurrently, because all these chips need to be designed to work together to build a whole system," Mastroianni said. "That means the architects really have to work with the package architects, and a little bit of ASIC, but it's really deciding how to decompose into chiplets. This is a huge exploration space. There are a lot more options to split it up. You also can build much larger systems. You are limited, even in the advanced nodes, to a certain number of gates and die size, reticle size, which is a physical limitation. But when you get into these advanced packages, that's no longer the case."

EDA tools need upgrades

All the chiplet integrations are pushing current EDA tools to their limits, but 3D-IC will require some significant upgrades to existing EDA tools and flows.

"We call them the three M's," said Marc Swinnen, director of product marketing at Ansys. "It is multi-physics, multi-scale, and multi-organizational challenges:

- The multi-physics challenge includes thermal, electromechanical, mechanical, and electromagnetic. The chip designer traditionally didn't worry much about these issues, except for the RF guys.
- The multi-scale challenges occur when you go from nanometers on a chip through millimeters on a package through centimeters on a 3D-IC interposer. That is six orders of magnitude you're encompassing. Traditionally, those have been handled by three different sets of tools. For 3D-IC, these all need to be consolidated into one.
- It has become a multi-organizational issue. The skills for this do exist in industry, but they're separated across different teams, different companies sometimes. For 3D-IC companies, they will have to re-architect their organizations to bring together a single team that encapsulates all the expertise you need to solve this. You can't throw it over the wall to some remote

team, and then to another team across the world, and back to the design team."

Many problems are extensions to ones being used today and may be introduced in phases. "The first phase will probably support homogeneous dies," said Mastroianni. "All the dies will be in the same technology. That makes it a little bit easier, but ultimately to really leverage this technology, you want to be able to leverage different process technologies, different nodes. That will require common data models to be able to perform timing closure. In addition, when all devices are not on the same die, you can't assume that they're all fast-fast, or slow-slow. You have to deal with that. On-chip variation is a statistical technique where you make some assumptions about how much your timing can vary. It's basically overhead margin you're building into your design. But if you have different chips that are manufactured in separate runs, you can't assume any correlation. They're totally uncorrelated. So you have to do a more extreme corner optimization analysis."

Place-and-route tools have to be re-architected for the Z dimension, and they must become a lot more thermally aware. "We already have 2D design tools that are activity-aware," said Cadence's Patwardhan. "The placer and other tools have the ability to take a VCD file, representing worst-case activity. You can aggregate that from a simulation, and then the placement of cells happens in such a way that the hotspots are spread out. It is an iterative flow where we do the first placement, do the cropping, and after clock-tree synthesis, we can refine the placement using some of the activity data. That is a power density-based flow. This can be expanded to 3D. We are working on this and have some early prototypes, where we can take activity information and then use early thermal analysis, either based on your static currents or full dynamic activity, and based on that decide 3D placement. We have enhanced our 2D placement engine now. We have to expand it to take the Z dimension, and it's a multi-objective placer. Thermal effects can be directly modeled as an objective to the placer."

Another simplification likely to be seen in early tools is a restriction on where a Z-dimension partition will be made. If macro cells or IP blocks are kept on a single die, they can be signed-off within a die and not have to wait until the entire stack is logically assembled.

"There are people talking about removing that restriction," said Cadence's Park. "They call it macro-folding. In the analog world they call it circuit folding. If you have a really small form factor in a planar sense, but they have some space to go vertically, there are people talking about folding macros on top of each other. I don't know of any design where that's actually in production, but there are certainly some of our customers talking about that capability. With folding you can make it half the size in a planar sense, and just a little bit thicker in a vertical sense."

Physical issues — stress, materials, thermal, aging

Physical issues are a big challenge in three dimensions. The entire chiplet team has to worry about mechanical stress particularly with large silicon interposers.

"They must be concerned about the pitch of the micro-bumps and look at reliability types of issues," said Tony Mastroianni, advanced packaging solutions director at Siemens EDA. "Test is a huge issue, as well. The test folks have to work to make sure that for all the chips you have a test strategy, so they get involved very early to figure out how they're going to test the whole system-in-package, and then the chips need to talk to each other. Die-to-die interfaces are used for the high-speed functional layout, but for test, all that test I/O needs to get connected in the package. So again, the package guys are in the middle of that."

One of the key learnings from 2.5D integration is that significant mechanical issues exist at the junction of two dies. "Whenever two dies are integrated together, you introduce stress," said Thomas Uhrmann, director of business development at EV Group. "If you look at problems with interposers, most of the fracture points are at the connections, and that creates reliability issues. You shouldn't underestimate the complexities that come from dealing with a mix of materials. In the middle of the die, you may have underfill. When you cure that, it shrinks. That creates stress, even though it stabilizes the connections. With 3D integration, that problem shifts to another dimension."

Still, those problems are reasonably well understood at this point, even if they are not always easy to correct.

"More interesting heterogeneous stack questions come up when you start mixing different materials," said Synopsys' Aitken. "When you've got CMOS stacked on CMOS, even if it's a different node, it's probably going to behave mechanically in a way that makes sense. If you decided to stack a gallium nitride device on silicon, or put a layer of some other objects, there's a bunch of cool things you could do. But you start to get some interesting mechanical questions that require a lot of thought."

Thermal dissipation adds another challenge, and it affects different components differently and often unevenly. "If you're moving the HBM away from the processor, it's going to be stacked, so then you've still got thermal problems on each of these pieces," said Frank Ferro, senior director product management at Rambus. "If they're not all on the same SoC, it's even worse, because you've got so many watts on the HBM and so many watts on the processor, and that's in a very tiny footprint."

Thermal issues can crop up anywhere in a package, caused by heavy computation due to high transistor density or resistance related to pushing signals through skinny wires. "Thermal is probably the biggest challenge today," said Mastroianni. "While HBM is doing stacks of 12

die, that is a very different problem because it is memory, and you're only going to enable one of those stacks at a time. They're not all firing at the same time. They do not have to worry about thermal management. The practical limitation today is probably three die, and even that's going to be a challenge."

But it is not all bad news. "The 3D configuration helps a little bit by reducing dynamic power," said Cadence's Patwardhan. "With 2.5D, a signal has to travel across a large die and then travel on the interposer to another die, resulting in long wire lengths. When you have a stacked die, you can route in the Z direction, reducing the wire length. So the dynamic power, the switching power, is reduced in a 3D configuration. If the stacking happens correctly, where the switching elements on both dies are not switching at the same time, you can effectively use a 3D stack to reduce the power or thermal footprint. If there is too much switching happening on both tiers at the same time, thermal effects, the chimney effect, comes into play."

That is one application for 3D technology. "If you believe in the concept of dark silicon, that not all of the device needs to be on all the time, then you could conceptually architect a 3D stack in a way that you were able to manage the thermal so that power and the heat were not a problem," said Synopsys Aitken. "You could take something that previously was implemented as a large 2D object and implement it as a smaller 3D object instead."

In any case, it will require early analysis. "You must do some thermal characterization very early on, before place-and-route," said Cadence's Park. "You need the ability to put in parametric data describing the power dissipation of each chiplet, the molding compound to be used, parametrically describing what you expect a potential heatsink to look like, the size of the package it is going on, because that's a natural part of the system that helps distribute the heat, and even the size of the PCB, which further helps. In the prototyping stage, you are starting to look at what things can be stacked, and even in a 2D world, how close they can be to one another, what types of chips or chiplets are best suited for stacking based on early knowledge about the design."

Also, chiplets may age at different rates. As mentioned earlier, uneven aging in chiplets can cause all sorts of reliability problems that have never been tackled before, particularly in markets where devices are expected to remain functional for years.

Inter-die connectivity

Ensuring high-speed communication between chiplets requires dealing with complex physical layers (PHYs), SerDes, and communications protocols to ensure the reliable transmission of data. But from a tools perspective, this is particularly difficult with chiplets because of the number of possible configurations. That design flexibility is the whole point of chiplets, but creating tools that are specific is key.

"You need PHYs for high-speed interfaces in 2.5D because you're driving up to two millimeters," said Siemens' Mastroianni. "You have to worry about timing and synchronization and dealing with signal integrity issues. But with true 3D, since the logic is nanometers or microns away, you can just use regular gates, the regular standard cells. They do have special cells that have a little bit of ESD built in, but essentially you don't need those PHYs. Instead, you just have those logical interfaces talking through regular logic. You have to do some synchronization for clock but that's normal STA logic type stuff and timing optimization."

That creates different problems. "You have the opportunity for many more interconnects between the stacks, and putting tens of thousands of PHYs is a non-starter," said Aitken. "But you do have to be concerned about the testing of these things, the sign-off of these things. What are you actually going to drive? Will you have an inverter that drives up to a piece of metal and connects to a matching buffer on the other side? Or are you going to put in some kind of a MUX so that you can do some testing? Or are you going to try to contact them for wafer probe, or are you going to forget about the whole thing and not test it until you build it?"

The design community is attempting to answer those questions. "There is normally pre-bonding and post-bonding tests that OSATs perform," said Patwardhan. "Direct probing of these micro-bumps, which are less than 10 microns, may or may not be possible with today's testing techniques. A lot of the testing happens through test paths that are defined across the two dies. They insert programmable e-fuses that can run open circuit tests. We have to make sure that whatever test insertion we do, from an EDA point of view, we follow the emerging IEEE 1838 standard and make sure all those checks are available through the full EDA flow. Testing will evolve as these hybrid bonds become more mainstream."

UCIe is somewhat limited right now. "The current version of the UCIe standard is designed to have one processor in the chiplet, the capabilities of which are extended by additional accelerators on other circuits of the chiplet," writes Andy Heinig, group leader for advanced systems integration and department head for efficient electronics at Fraunhofer Institute of Integrated Circuits, Division of Engineering and Adaptive Systems. "However, system architectures in heterogeneous systems (e.g. for autonomous driving) will be designed in a substantially different way, namely with distributed multi-processor systems, each on different circuits of the chiplet. This necessitates the development of new system concepts, which must themselves be standardized to facilitate interoperability. Since these concepts are yet to be developed, however, this poses a significant challenge. It is therefore too early to develop standards in this area."

While AMD, Intel, Samsung, and Marvell have successfully deployed their own chiplets, a commercial chiplet market is more in the

planning stages than reality. HBM is the closest thing to a standard chiplet today, and it's expensive due to the cost of the 2.5D interposer, so its use has been limited.

"The supply chain problems didn't help over the last two years," said Rambus' Ferro. "That's eased a little bit, but it did highlight some of the problems when doing these complex 2.5D systems because you've got to get a lot of components and substrates. If any one of those pieces is not available, that disrupts the whole thing or results in long lead times. That's given rise to methods to build an HBM chiplet or an HBM system connected through a traditional PCB or other means. If I've reduced the size of my processor — at least in the sense that I've taken some of the I/O away, but also because now I don't have to put that big processor on a 2.5D silicon interposer — I can build that in a traditional package and then talk to the 2.5D. If you just have that I/O device with either HBM or even GDDR, or with something like PCIe, you can have different I/Os on those chiplets. We are seeing a really strong push now for that technology. And while still early, the rubber is now starting to hit the road where real solutions are starting to be proposed. UCIe is taking the flag for the moment."

Connecting everything in a consistent way is one of the big challenges. "A PHY alone will not make the interconnect fully compatible with other chiplets with the same PHY," said Javier DeLaCruz, senior director of system integration at Arm. "The control logic and protocol have to align, as well. The choice of protocol is likely more religious than the PHY. Steps can be taken to make die-to-die interconnects more efficient from a PPA standpoint, particularly with latency, but these tradeoffs reduce the ability to extend the use of these chiplets to other products."

For example, two UCIe interfaces may use the same PHY, but if one uses a PCIe protocol and the other uses a streaming protocol with a direct to CHI interface, there is a mismatch that needs to be addressed.

"There is hope in the chiplet reuse model," DeLaCruz said. "In most use cases, there will be one chiplet with the product's secret sauce that would be custom-designed. This custom chiplet can match the PHY from the pre-existing chiplet and/or employ protocol translation."

Others say chiplets will not be so straightforward. "From a technology point of view, re-using IP in the form of pre-made chiplets is far more difficult and risky," noted John Ferguson, director of product management at Siemens Digital Industries Software. "In an SoC, you have a single process and a corresponding PDK. Things like thermal impacts and mechanical stresses are fairly mitigated as everything is on a single silicon substrate. With heterogeneous chiplets combined in a package, you now have a vertical degree of freedom, corresponding to new stresses from stacked dies or other materials, and greater challenges of heat dissipation through multiple layers of different materials. These affect electrical behavior and ultimately can impact yield or

reliability. Bottom line, just because a chiplet works in one package configuration, does not guarantee it will behave the same in another. 3D-IC verification will require continuous levels of analysis at different levels of the design viability to ensure it will behave as desired."

This is much more complex than what is needed at the SoC level, and it also translates into a business issue in terms of investing in new tools, training users, the time required to do all these iterations, and the risk of failed designs, Ferguson said.

While 2.5D is the predominant stacking technology today, full 3D-IC is coming, bringing a whole new set of problems. "The technology is there, where you can actually stack die and make them work, but it's primarily done manually," Mastroianni said. "What I call true 3D is where the tools will actually do all that detailed chip decomposition, making sure power is delivered through all the different layers and closing timing through multiple dies. That's probably two to three years away before it becomes mainstream."

A large ASIC, in contrast, is broken into hierarchical blocks, and there's reusable IP primarily for analog and high-speed I/O. "With chiplets, now you're decomposing your 'super ASIC' into smaller chiplets, but you're not necessarily stuck with having to use the same process, and you can use that to your advantage," he said. "If you have a large processor, you can use a 5nm or 3nm. If you have an analog/mixed signal, you can use a cheaper process that works better for that. And there may be specialized IP that's only available in a very expensive node. If you only need that for one interface, why not just build that into a chiplet?"

Still, depending on the type of the IP, as well as the application, different types of adaptations are necessary, noted Fraunhofer's Heinig. "The most important part is the connection of the IP to the die-2-die bridge. If you have a single core/accelerator or an FPGA, a direct connection via an AXI interface is possible. But here, too, it must be clear how the accelerator gets the data. If it is a row streaming interface, it is easy to transfer the data. But if a shared memory is used, the access to the memory must be defined. If there is a multi-core system, new system concepts must be established with local memory on the accelerator card to avoid unnecessary data transfers. It's not only a problem of the IP itself, but maybe also for the whole chiplet. It must be clear who is responsible if something isn't working in a complex and expensive chiplet system."

IP and IP reuse

At the same time, chiplets raise a whole host of interesting business impacts for IP vendors, much more so than technical impacts.

"On the technical side, for pure digital logic IP (such as all processor IP), chiplets won't have any impact to the design itself — i.e., processor

core logic connects to on-chip interconnect — and whether that interconnect bridges to other on-chip elements or off-chip through I/Os of some flavor makes no difference to what the processor vendor delivers," noted Steve Roddy, vice president of marketing at Quadric.

On the other hand, the business side will be quite interesting, Roddy said. "The predominant pricing model in processor IP is license fee plus running royalty, where the royalty is most often calculated as a percentage of the average selling price (ASP) of the semiconductor product. Back in the day when 95% of the business for IP companies was selling to semiconductor vendors who sold single packaged die to system OEMs, the percentage ASP model was well understood and well accepted. But what if what previously was a $20 packaged, tested, final-piece price with a single die suddenly becomes a $20 multi-die module on a substrate? Where does that percentage of ASP get calculated? Does the calculation happen on the digital chiplet price or the final assembled multi-chiplet package part? What about situations where the licensee sells only the chiplet and someone else integrates multiple chiplets into the final assembly?"

A further wrinkle is that the IP business worked, historically, because the IP block is one of many such pieces of IP assembled into a larger chip, Roddy said. "Processor Vendor X and Silicon Company Y were able to agree on, for instance, a 3% ASP royalty largely because the IP core occupied only 3% or 5% or 10% of the die area of the final chip. If chiplets suddenly change the equation to where a chiplet consists solely of a single IP, and for instance a chip assembler builds a module from one or more CPU chiplets plus a DSP chiplet plus a GPNPU chiplet plus a wireless baseband chiplet, then the entire value of a specific chiplet is mostly the realization of that single IP block into silicon. IP vendors won't sit idle and accept 3% of the revenue from a chiplet that contains only that one vendor's IP. The IP vendor will have tremendous incentive to transition into a chiplet supplier directly, not an IP supplier. The design content and engineering effort are almost the same, but the delivery vehicle is changing. The more that chiplet interconnect and assembly technology standardizes across the industry, the faster that transition could occur."

This also means companies that used to get away with just delivering code now have to actually deliver silicon, observed Rich Goldman, director at Ansys. "They've got to understand all the tape-out issues, and that it's actually got to work. A second consideration is they no longer have to rely on royalties and extracting royalties from the users and from the foundries. They now can sell actual silicon. Royalties are an age-old problem, as old as IP itself, and one that has never really been solved except by Arm — but only because Arm had such a central position. From previous experience, royalties were something that we tried to stay away from because it sounds like a great thing, and you can negotiate royalties, but as soon as you're

very successful in quantities, your customer will renegotiate those royalties down so the payoff is never really there, with the exception of Arm. With actual silicon, you get to participate in the volume part of the sales."

Siemens' Mastroianni argues that for chiplets to become mainstream, there will need to be reusable chiplets. "It's coming, but it's a very different business model. The likely candidates who are going to do this are IP providers. It could be the PHY guys, because they can build high-speed SerDes chiplets. Or it could be processor-type companies. In any case, it's a pretty big transition because they're licensing IP right now for integration within a package. The difference here is, there's a couple of models that they can take. One is they can do a test chip. Several IP companies have to do test chips anyhow, so they can design a general-purpose chiplet with their IP. They'll do a test chip, and they could just license that. But then the customer would have to go off and basically do the fab, do the testing, etc. Another model would be to get into the silicon business. In that case, they would assume all that operations type stuff, from package and testing to manufacturing, and distribution of known good die. If they take that route, that would be more of a traditional piece-part type, and there could be a license on top of that as well. The business model still remains to be seen."

Mixel is one of those IP companies that uses test chips. Ashraf Takla, CEO of Mixel, looks at the chiplet opportunity as either an IP kind of approach, or one in which its test chips are converted into chiplets. "For the C-PHY/D-PHY combo, many companies acquire hundreds or thousands of these. Equipment companies, for example, need to test C-PHY but they don't have an FPGA solution or some way to implement that on the tester side. So, they buy these chips for that purpose. We're already selling our test chips at low-volume, and at the right time it makes sense to turn some of these into chiplets."

Takla noted it may not make sense to have many different chiplets similar to each other. So if it's a transmitter CSI, that could be one chiplet. "Maybe you would have a transmitter CD-PHY as a chiplet, but you also can add programmability into it, such as how many lanes are needed. Is it CSI (camera serial interface) or DSI (display serial interface)? So just like it's programmed into the IP, maybe you can build even more programmability on top of that. On one level, chiplets aren't that different from the IP business. Everyone is trying to re-use IP and do things in a way that they don't have to customize it as much, and for chiplets that will be even more important."

In the short term, chiplets need enough volume to be able to address the complexity they bring, whether that involves tools or interfaces such as UCIe/XSR (extra short reach). "Everyone wants to do chiplets," said Mick Posner, product line senior group director for IP at Synopsys. "Can you remember the last inflection point of an introduction of a significant new protocol that's applicable across all market segments? No, and

that's why you see 100,000 different customers trying to get into this. Long term, there are going to be some challenges because it's the ultimate of reuse, along with the business model. Typical IP sales are based on a per-project basis. With chiplets, what do you now define as a project? If that die is used across 10 different chips, is that 10 uses? Is that one use? And customers know this very well. They see it as, 'I've just produced a piece of silicon. You never put any kind of restrictions on my silicon before, outside of royalty models.' But when it comes down to the license usage, that potentially could have an impact on the whole industry."

Posner believes "chiplet" is a loaded term. "It has a connotation that it's this tiny little chip. But I can guarantee, from the 100+ opportunities that we're tracking, a chiplet is a 50 x 50mm^2 die. It's just that there are now two or four of them in a package. Chiplets are fantastic for mixing and matching your nodes. It's the whole notion that you can mix and match different die."

It is possible to combine chips from different foundries into a chiplet, but it's not as simple as it sounds. "You have to worry about standards and making sure you get the all the right voltages," said Siemens' Mastroianni. "Even if it's from the same foundry you have to worry, because the chips will be coming from different lots. By definition, they're two different chips, so you have different corners. If they're coming from different processes, that even makes it a little more challenging, but it still needs to be dealt with. A lot of that is handled through the die-to-die interfaces, almost like a SerDes interface, which kind of decouples it. Those interfaces are designed to be like that. It's more of an issue on those other signals. For instance, the low-speed I/O may need some connection, so you have to worry about it. But typically, those interfaces are not as critical. And the high-speed I/O are covered pretty much through those standard protocols."

Flexible partitioned functionality can help make chiplet architectures more modular and scalable. "At the same time, chiplets make it possible to optimize some design processes," said Sue Hung Fung, product line marketing manager for UCIe at Cadence. "Choices of process nodes, foundries, a wide range of package types, and IP types, etc., are made possible through chiplet flexibility. For portable and modular chiplet systems, die can be re-used and can also interoperate across the various process nodes in a die-to-die fashion. Digital logic circuits are far better at scaling down in geometry than analog circuits, RF, and memory. A cost-effective way to achieve a reduced die size and higher yield is by dividing the analog/digital domain into separate areas, re-using the analog, and scaling down the digital."

Then, when there is a need for variation in the common denominator function, IP reusability is relevant.

"The port speeds, for instance, could be 100G, 200G, or 400G. However, scaling of the same chiplet would re-use the 100G IP and scale it to 800G, which wasn't accessible in the earlier

chiplet design, when a future-looking need for 800G comes into play," she said. "The reuse of IP is an important factor in chiplet design here. For example, there can be a case that it is necessary to redesign the chiplet to handle an 800G bandwidth. The area, power, bump pitch, etc. may all have a possible impact to incorporate the new bandwidth updates needed. However, the basic functionality of the IP and the circuitry may be re-used to achieve a higher bandwidth."

Security

Security is an issue for any chip design, but disaggregating SoCs into chiplets significantly alters the cybersecurity threat landscape.

Unlike a monolithic multi-function chip, which typically is manufactured using the same process technology, chiplets can be developed anywhere and at any process node. In fact, among the key reasons for developing heterogeneous chiplets is that not every function benefits from the latest process technology, and not all of them can be crammed onto a single die. But that also raises the threat level, and the industry is struggling to deal with security issues in a repeatable and cost-effective way.

"Some of these are built in a way that they can be affordably reverse-engineered or re-manufactured, with some malicious functionality added to the chiplet because it's a much smaller die," said Scott Best, senior director, silicon security products at Rambus. "It's a million transistors and it does a very specific function. There are state-funded actors around the world that could clone that functionality, put it into a compatible process node, and add a malicious capability to it. And now the risk is they somehow insert that malicious component into the supply chain, and it gets inadvertently integrated. It's a little bit scary, especially in its do-ability. I'm not sure a lot of people fully appreciate how easy it is to clone small chips built in trailing-edge technology nodes. Trailing edge starts to mean more and more things as time goes by. Even 22nm, which was once state-of-the-art, is three or four generations behind us now. They become very appetizing targets, especially if coming up with that chip gives an adversary access to a much larger end market."

A key challenge is to develop a supply chain with robust traceability. "Let's say a customer is concerned about whether security is built into a chip," said Siemens' Mastroianni. "A lot of the die-to-die interfaces have security built in. But there are other considerations for traceability, which is a different issue. We want to make sure that for the whole design process you're not compromised, and that requires covering all the way from RTL design to manufacturing delivering the parts. It's a whole different beast."

Integrating multiple chiplets into a heterogeneous package opens the door for security breaches and potential risks associated with malicious modifications or attacks on the individual chiplets during design, assembly, or test. "Also, because chiplets are often designed and manufactured by different vendors, there is

a risk that a malicious actor could compromise one of the vendors and use that access to compromise the entire chiplet-based system," Mastroianni said.

Some of these concerns began with SoCs, but chiplets are raising them to a whole new level. "Security issues still exist, but with chiplets they can be much harder to prevent," said Guillaume Boillet, senior director of product management at Arteris IP. "Now, coordination is needed across who's doing the chiplet since some countermeasures require the chiplets and the software to work hand-in-hand. Beyond that, with chiplets, there is a bigger surface of attack, and out of this two categories of problems emerge. First is hardware Trojans. In heterogeneous systems where the chiplets could come from different companies, the idea that some companies will reverse-engineer some of the chips that comprise the chiplet and replace it with a malicious one — that sounds like science fiction. But this threat exists, especially when it comes to IoT and all those things is probably more achievable than it is for a big system."

The second problem is a wider attack surface, in part because the connections between the chiplets are a lot easier to track. That creates the possibility of a man-in-the-middle attack.

Some of this depends on the complexity of the supply chain for chiplets. Companies like Intel, AMD, and Marvell develop their own chiplets, which minimizes supply chain threats. But companies assembling and integrating commercially developed chiplets will have a much tougher time.

"Chiplet authentication is not such a problem if you're a single company," said Synopsys' Posner. "You've disaggregated your SoC, you own both sides, you are familiar with your supply chain, so the potential for a malicious chiplet getting into your supply chain is probably pretty low. You've created a package that has your die and some third-party die, and they need to communicate with each other. One of them has to take the lead and say, 'I'm the manager, you're the sub. Please declare you are an authentic chiplet.' Today, there is no standard around that. It is mentioned in the UCIe spec as something to be addressed in the future."

It's a different story with a chiplet marketplace. "At the moment, because there's not really mixing and matching of die, so it's really not an issue," Posner said. "But it will be shortly. To alleviate a malicious die getting into your supply chain, you have to have some form of authentication, and now that you're an authenticated system, what if someone's trying to do a hard hack where they've taken the package apart and somehow are getting access to the links between those two die? What sort of encryption or protection is put in place to go over those links?"

Others agree. "This mixed supplier ecosystem doesn't yet exist, i.e., an ecosystem in which suppliers are supplying some standard functions that can be reused in a lot of systems, and the primary vendor of the integrated package is buying those from one of several potentially

competitive suppliers to include into their package," said Mike Borza, Synopsys scientist. "You can imagine things like system controller chips and TPNs and all kinds of stuff like that being good examples of things that will be integrated in this way because they have well-defined functionality and there are protocols around them."

The more suppliers of chiplets, the harder it is to secure everything.

"You're going to have more vendors involved," said Steve Carlson, director/architect, aerospace and defense solutions at Cadence. "You have multiple chiplets, so you probably have a new interposer vendor, as well as new packaging and test folks everywhere along the supply chain. Even before that, with the IP blocks that go into the chiplets, there's potential for malfeasance or errors. Additionally, there are issues with the interoperability of the chips with regard to their safety protocols, and what are 'safe use' models. Those things aren't always communicated clearly, so when you connect two secure chiplets together, it may be done in a way that undoes the chips."

This is particularly true for interconnects.

"You have exposed those interconnects between them to be more easily exploited," Carlson said. "Along with these high-level considerations, there are the traditional issues, as well, having to do with Trojans. You could produce things that are out of spec with regard to the materials and the interconnect on the interposer, overproduction, counterfeits, reverse engineering, side channels, and recycling kinds of things."

Approaches to improving chiplet security vary greatly, but first and foremost there needs to be a focus on system-wide security.

"System-wide security review has now become an integral part of the development cycle of any product that has any interaction with customer data," said Rajesh Velegalati, senior security analyst at Riscure. "Certification entities like Common Criteria and EMVCo also play a crucial role in ensuring these products adhere to security standards."

However, if these security reviews are conducted in the latter phases of the design cycle, "even if vulnerabilities are found, fixing them might become an issue in itself," Velegalati said. "If the vulnerabilities are found in the hardware or the immutable ROM code (the first piece of code which the product executes on powering up), patching in mitigations would be difficult and may be downright impossible. In this scenario, the product might be released with vulnerabilities, and the manufacturers are only able to patch it in the next iteration of the product. The next iteration will depend on the product lifecycle, which could be a minimum of one to several years."

Complicating matters is that every customer has different concerns about security and trust, and has different approaches to address those concerns.

"DoD customers are all about security, and know more about security than we do, including encryption and decryption, and so on," said Geoff Tate, CEO of Flex Logix. "When an FPGA is embedded inside the chip, the customer can put a lot of security around the FPGA. The biggest concern that commercial customers have is people with packages can snoop and see data coming in and out of the package. But the DoD is worried about people even using sensors to detect stuff happening inside the package. Those guys are very concerned about security, but they don't need us to provide solutions. With commercial customers, people use embedded FPGA first from a security angle because they can reprogram stuff. So if they have an encryption algorithm, and it gets broken, they can reprogram the encryption algorithm if it's an embedded FPGA. They can't if it's hardwired."

The notion of secure communications between chips is something that has been studied with anti-snooping measures, unauthorized memory accesses, and other approaches.

"Those issues have been worked on, and are being worked on," Cadence's Carlson said. "Security in the past has been a one-off project per program endeavor. Now we're looking to standardize those things, and it's very difficult because with the DoD — whether it's a civilian system or not — when something gets exploited, they like to classify the data and make it secret. MITRE CWE is the rallying point that we can work on. It's an abstracted view of what the vulnerabilities are, and they can talk about some common mitigation techniques at that level. The hope is that will help drive things toward standardization more quickly as people agree upon what the threat vectors are that are most important to protect against."

Moving forward, Riscure's Velegalati said the best way to address this issue is to include security reviews in the design phase of the product itself. "In other words, finding and fixing vulnerabilities pre-silicon, so that after the chips are fabricated, hopefully they won't have any of the identified vulnerabilities. This might require a cultural change in the way a development cycle itself is executed."

There are several commercial tools available to detect vulnerabilities in software using static and dynamic analysis techniques.

"Deployment of security agents can help ensure system bring up integrity, operation and the detection and recovery from compromised events," said Siemens' Mastroianni. "Additionally, ensuring the source integrity and traceability of the chiplet components throughout the full supply chain will become a critical requirement for new and existing products. By ensuring the source and integrity of each chiplet can be verified and traced back to its source, it becomes easier to identify and respond to any security or trust issues that may arise."

Another way to deal with chiplet security is with a layered approach. Pim Tuyls, CEO of Intrinsic ID, points to three layers of chiplet

security. “One is that you say, ‘We are going to make sure that every part can be identified,’ and use a physical unclonable function (PUF) as an identifier. The advantage of that solution is it’s lightweight. Second, you say, ‘We need a little bit more. Namely, we need chiplets that can securely communicate with each other, and can authenticate each other. You can do that with symmetric crypto. Here, you implement the PUF on every chiplet with symmetric crypto attached to it. The advantage is that all those chiplets can create their own keys, and can run a key exchange protocol. You don’t have to program every chiplet. Third, you say, ‘We want the highest level of security, and the most sophisticated.’ That could be tough in a public key infrastructure (PKI) system.”

An open question is whether the chiplet security solutions will be in place when they are needed at a cost point that is acceptable.

“We have 100% confidence that the solutions will be available in the timeframe that they’re required,” said Rambus’ Best. “Even if I just look at our anti-counterfeiting experience, we still have customers showing up asking us for anti-counterfeiting solutions. So the techniques for protecting against this are known. The downside is that none of it comes for free. If it makes manufacturing 10 cents more expensive for a $5 chip, most commercial vendors will reluctantly agree to it. Others will negotiate until it’s only a 5 cent cost hit, or maybe less than a penny. ‘How much can I protect this for less than a penny, because I’m on razor-thin margins right now?’ One ends up dealing with somebody who’s responsible for procurement and manufacturing. The security architect has long since been taken out of this conversation. You can build a very secure product with a very secure supply chain. We absolutely know how to do that. We have this technology, we know how to deploy it, we know how to deploy it cost-effectively. What’s difficult is just the normal decisions about whether we really solve for all of the security problems first, or is ‘perfect’ the enemy of ‘good’ here? Can we just pursue good and see if that’s sufficient and tested?’”

Others point to similar experiences. “Security isn’t done for free, and until you’ve been bitten, you probably don’t want to spend money on it,” Carlson said. “It’s like insurance. The systems companies bear the brunt of the responsibility. In cell phones and automotive, as well as areas that have been public about hacks, they have to do something. They have been hurt financially by security-related issues. In product areas where they haven’t been hurt yet, they really aren’t doing everything they could. They’re not doing nothing, but they’re doing less than they could. Another issue is that you could spend $100 million on security and still get hacked. There isn’t a single, assured solution for that problem. Hackers are really inventive and coming up with new and crazy ways every day.”

Additionally, chiplets can be reverse-engineered. How do you detect whether a chiplet has been tampered with? Guiaume Boillet, senior director of product management at Arteris IP, said

solutions do exist, at least from some providers. "We work with them with our joint customers to help them implement those systems, especially in SoCs, to make sure the chiplets are properly identified," he said. "In different angles of attack, they all have one idea to address it, but they come with their own challenges. Most will require, for instance, that the two chiplets are really designed hand-in-hand. So when you consider those attacks with the man-in-the-middle, where the connection between the two chiplets could be probed, there is a way to fix that by only sending encrypted messages. But every time you talk about encryption, it needs to be encrypted on one side, decrypted on the other side, and needs to be the same piece of IP. This is quite involved, but it can be addressed if it's an in-house chip disaggregation, where the two chiplets come from the same vendor. However, if we start talking about an ecosystem where it's off-the-shelf parts, we're going to need standardization well beyond everything we have in terms of electrical compatibility. What is apparent beyond all the discussion we have with customers while doing chip disintegration is that it's the mechanical aspect of it. The mechanical and physical aspect are not really addressed yet."

On the standards front, a CDX working group in OCP covers security, and within the UCIe, a working group is dedicated to chiplet security, but it is still very early days in both cases.

Synopsys' Borza noted UCIe is coalescing around PCIe and CXL by and large. "That's an important part of the equation, and means they'll be able to coast on the coattails of the work that's being done in CXL and IDE," he said.

Also, in the realm of chiplet security standards are approaches like security protocol data model (SPDM) to do the mutual authentication of chiplets.

"Once SPDM has run, there are protocols that can use the fact that you've got an authenticated connection now to set up an encrypted link to address the other aspect, because integrated parts made of chiplets sit somewhere between chip-on-board and a fully integrated SoC in a single-die kind of solution," Borza explained. "They are more vulnerable than an SoC to physical probing attacks, but they're less vulnerable than a chip-on-board. But that vulnerability still exists because it's much easier to probe a chiplet-to-chiplet link than it is to probe something on a chip itself. It's not simple by any means, but it's easier. The way I see this shaking out is that we're actually going to be able to benefit from some of the other work that's being done, and we won't get it for free. It's expensive to have high-bandwidth links that are running at low latency that are fully encrypted, but it is possible. We're doing it now with CXL, and we'll be able to move along that path with chiplets."

The proposed standards from Chiplet Design Exchange (CDX) working group (part of Open Compute Project Foundation) recommend an optional security agent that comes with the chiplet to ensure the supply chain is trustworthy. The security agent may be hardware or software.

Assembly and test

One of the big concerns with multi-chiplet integration is yield, which is determined by quality of all the parts. Ironically, while early chiplets were a way of improving yield — it's easier to manufacture a small die than a larger one — the equation changes dramatically with multi-sourced chiplets assembled like LEGOs into a package. One bad chiplet renders the whole package unusable, and that costs more than a bad planar SoC.

It's no surprise, then, that the industry is scrutinizing at every manufacturing and assembly step to ensure these devices use known good chiplets, and that every step of the packaging is ironed out to ensure predictability and repeatability.

Assembly, interconnect

A simplistic rule is if the foundry made all the dies, it could do the heterogeneous integration. If chiplets come from multiple foundries, an OSAT would do the integration. "Heterogeneous integration most often involves assembling several different dies using 2.5D and 3D packaging technologies," writes Vik Chaudhry, senior director of product marketing and business development at Amkor Technology, in his blog "An OSAT Perspective On Semiconductor Market Trends." "Many foundries offer these packaging technologies, where they can integrate several of their own dies in single package. OSATs also have similar solutions, where they can assemble dies from different foundries. The added advantage of using a third-party OSAT is that they may already be working with different foundries and have established intellectual property (IP) protections in place."

But multiple components in a single package is complicated. It requires sophisticated PHYs, SerDes, and communications protocols to ensure the reliable transmission of data. "You need PHYs for high-speed interfaces in 2.5D because you're driving [signals] up to 2 millimeters," said Siemens' Mastroianni. "You have to worry about timing and synchronization and dealing with signal integrity issues. But with true 3D, since the logic is nanometers or microns away, you can just use regular gates, the regular standard cells. They do have special cells that have a little bit of ESD built in, but essentially you don't need those PHYs. Instead, you just have those logical interfaces talking through regular logic. You have to do some synchronization for clock but that's normal STA logic type stuff and timing optimization."

Test

Every manufacturing step becomes more complicated with chiplets, because they all need to be inspected and tested in the context of other chiplets. In addition, because not all of the leads are exposed, testing approaches need to be considered early in the design phase.

“There is normally pre-bonding and post-bonding tests that OSATs perform,” said Patwardhan. “Direct probing of these micro-bumps, which are less than 10 microns, may or may not be possible with today’s testing techniques. A lot of the testing happens through test paths that are defined across the two dies. They insert programmable e-fuses that can run open circuit tests. We have to make sure that whatever test insertion we do, from an EDA point of view, we follow the emerging IEEE 1838 standard and make sure all those checks are available through the full EDA flow. Testing will evolve as these hybrid bonds become more mainstream.”

Improved testability, coupled with more tests at more insertion points, are emerging as key strategies for creating reliable, heterogeneous 2.5D and 3D designs with sufficient yield. Today, nearly all of the chiplets being used in multi-die systems or packages are developed by a single company. But as chipmakers move to leverage third-party chiplets developed by a multitude of vendors, they will need a combination of new design tools, equipment, and methodologies.

Test engineers are turning to enabling approaches in the transition from monolithic SoCs to chiplet-based systems, including:

- Improved probes for single-site and multi-site testing;
- Design for test (DFT) schemes for connecting and stacking chiplets;
- Test-cost modeling to identify the best tests moments, and
- Manufacturing flow optimization to drive yield.

Multi-chiplet packages compound the risk of chip failure. They also have a testability problem — they are harder to test in the context of a system because of their physical complexity (and sometimes lack of ports to test with) and the false assumptions that all the chiplets are known good dies (KGDs).

Each individual chiplet must function as designed after integration into the final package. A single failed chiplet, whether that’s due to defects or supply chain missteps, can increase yield loss.” This is why design for test (DFT) concepts are starting to be adapted to heterogeneous integration.

“When you have dies on chiplets, those need to be known good dies. Then, when you package them up, you have to have access to all of these dies in order to test them,” said Vidya Neerkundar, technical marketing engineer at Siemens EDA. “That’s the biggest challenge we have in the testing industry.” Access to pins

can be a problem with 2.5D, depending on the configuration, but 3D is much more challenging.

In 2.5D you may have access to each of the individual dies or chiplets through minimum access pins, but in 3D the dies are stacked on top of each other. “So the only access mechanism is only through the bottom die, Neerkundar said. “There is no other way you can access the dies that are stacked on top of it.”

The IEEE 1838 standard describes a way to access stacked dies for testing. It defines a serial access through a primary tab and a secondary tab. Each die has the two tabs (primary and secondary) that can be serially accessed through the package.

High-density interfaces, such as UCIe, pave the way for connecting various chiplets together. But leveraging those for testing purposes is a work in progress.

“UCIe is a start,” said Ken Lanier, business development manager at Teradyne. “It allows you to get data from one die to another efficiently. I am not sure there is consensus on what you would do with all the data sent on the bus, though. There is some idea that there is a need to do a silicon check — like BiST or scan — using the bus. The challenge is getting two different silicon providers to agree on what happens when their devices are connected together to validate the complete assembly.”

IEEE’s Heterogeneous Integration Roadmap underwent revision in 2023, with an update of system-level test and data analytics as key drivers for chiplet-based testing. The test chapter (17) said 2.5D and 3D testing should involve “the complete package at an application level and diagnosing failures at the die and interconnect level.” The roadmap lists the challenges to test:

- Known good die (KGD) testing;
- Interposer testing;
- High-speed Interconnects and signal integrity;
- Impact of emerging technologies with respect to test;
- 3D TSV/interconnect testing, and
- 3D probing, 3D die stacks, 3D stack repair.

It also calls for:

- KGD DFT methods for high-quality wafer probe testing to reduce fallout at final test;
- Die-to-die communication standards for thorough testing at final test;
- Repair methods at final test to keep yield high;
- A standardized test and repair methodology that considers new trends in 3D interconnects;
- Yield prediction and analysis methods so fallout at all levels of testing is understood, and
- End-to-end data analytics capability that applies to all dies on the package.

"When testing individual die, it may be difficult to physically contact TSVs or other small structures. The major challenge going forward is the testability of completed assemblies and finding faults in a cost-effective way," said Lanier.

That roadmap increasingly is influenced by the density of interconnects and power. "It used to simply be Moore's Law and 5nm, 3nm technology and a 1,000-pin BGA," said Dave Armstrong, principal test strategist at Advantest. "Now, interconnect density per millimeter and power connect density per millimeter have become ways to address the range of what the industry is doing, and that's a big change. It allows us to make real progress."

With monolithic SoCs or chips, engineers typically perform two main tests — one at wafer probe, the other after assembly and packaging, or package test. However, with multi-chiplet ICs, the number of potential tests points can increase significantly.

"In a three-die stacked IC, prior to bonding, wafer test can be performed before thinning or after thinning. Wafer bonding is typically by thermo-compression, and so you want to make sure that your dies have survived the bonding," said Erik Jan Marinissen, scientific director at imec. "If you don't do any pre-bond test, you might have a bad die in your stack. And once they're stacked, there's no way to replace them, rendering the entire stack worthless. The key challenges are testing on small micro-bumps and testing on very thin wafers."

In imec's case, wafers as thin as 50µm are used. "For us, 50 microns is the sweet spot," Marinissen said. "We have thinned down to 10

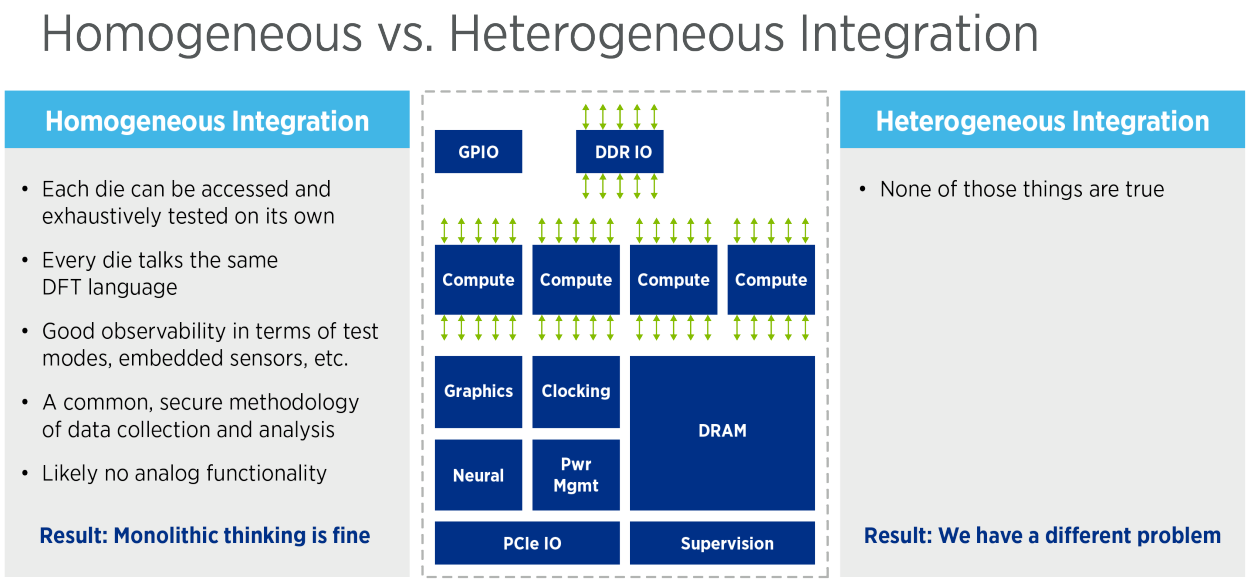


Fig. 12: Fundamental testing differences for heterogeneous integration requirements including DFT compatibility, observability of chips, data security. Source: Teradyne Inc.

microns, but then you start to affect the wafer, lose speed on your transistors, for example, and it becomes too thin to handle. If you go too thick, the TSVs will take a lot of time to form and they become too expensive."

"With chiplets, you have a reconstituted wafer, or they can use a die handler and test the parts. But they need to do some high-power testing, some at-speed testing, along with functional testing," said Advantest's Armstrong. "It's really beneficial to test the parts after thinning and after dicing, because you can get chips and cracks and other things. And being able to test at-temperature at the die or wafer level is advantageous."

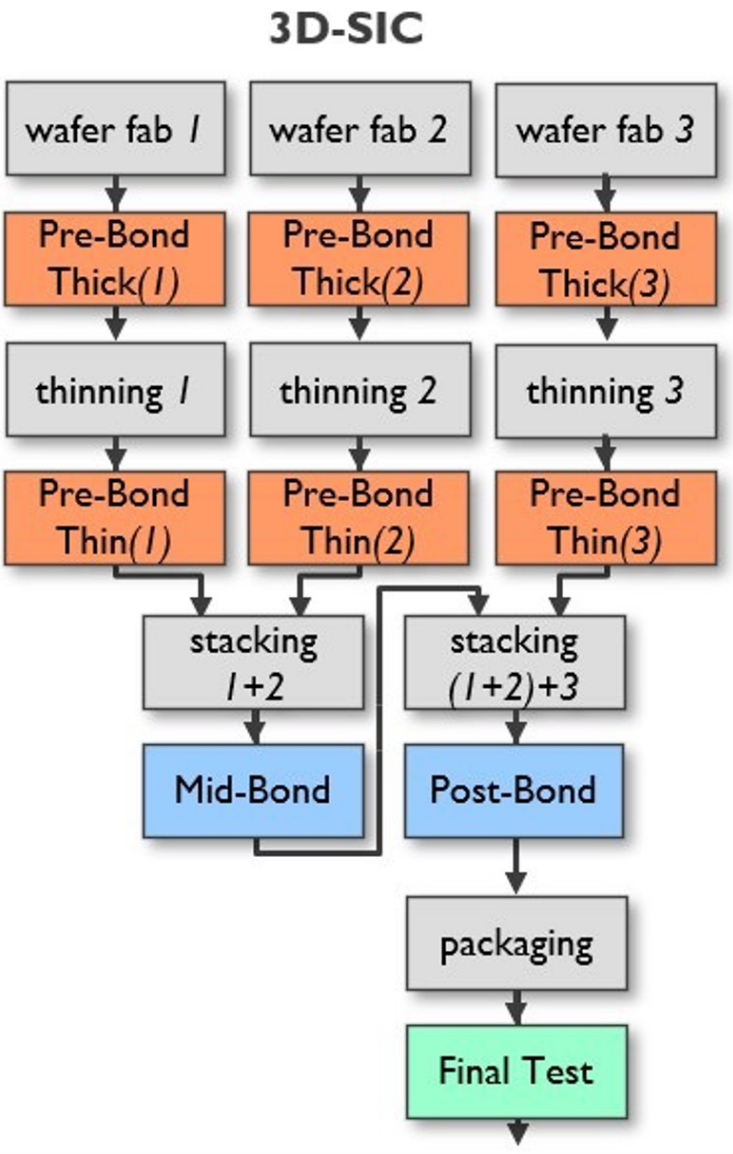
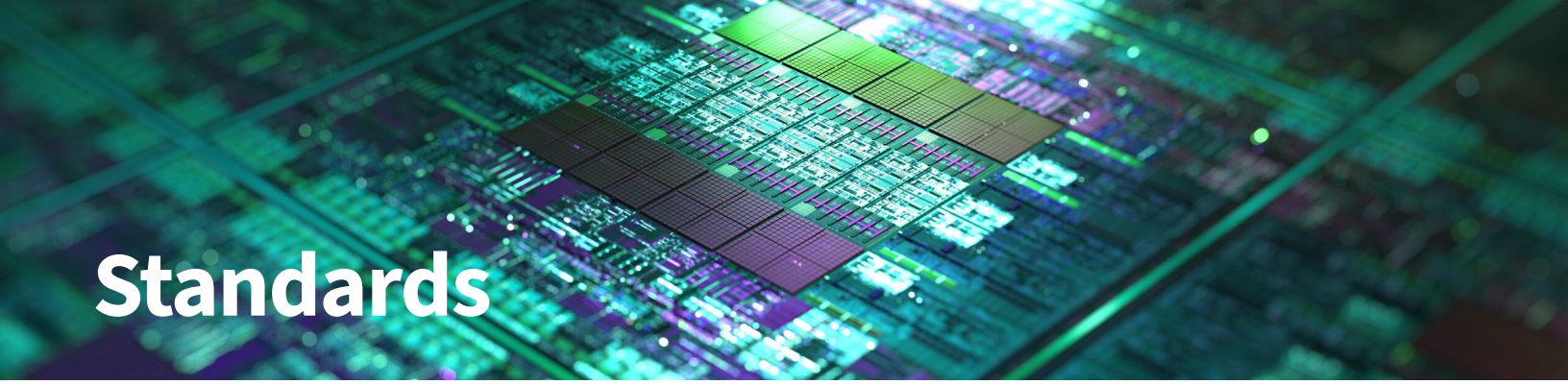


Fig. 13: Nine potential test moments for a three-chip stack. Source: imec



Standards

For most companies the shift to chiplets will happen slowly until proven standards are in place.

Interoperability and compatibility depend on many layers and segments of the supply chain coming to agreement. Unfortunately, fragmented industry requirements may lead to a plethora of solutions.

Standards always have enabled increasing specialization. In the early days of the chip industry, a company had to design, implement, and fabricate everything. For most companies, fabrication became separated from design by well-defined interfaces and models, such as PDKs, BSIM models, and libraries.

The emergence of the IP industry enabled companies to concentrate on the design of the system and the pieces that provided them with differentiation, but they still had to do most of the implementation themselves. Standards for interconnect and the models that were transferred between IP provider and consumer made this possible.

The semiconductor industry is transitioning from proprietary chiplets into one based on standards. Many of the proprietary solutions are being placed into the hands of standards bodies today. The industry is coming together to consolidate those solutions, but only a certain level of consolidation is possible or perhaps even desirable.

Use cases will drive the rate of adoption of the proposals, and if initial success is had, many more use cases will look to move in this direction. But they all may require variants of the standards. Flexibility and optimization are always tricky to balance.

"As fabless semiconductor companies begin to bring these disaggregated chiplets to market, their successful adoption requires the industry to standardize on a set of interface protocols in order to offer plug-and-play compatibility between different suppliers' chiplets," said the CDX working group in its standards recommendation document.

What needs to be standardized

The following table shows what needs to be standardarized and standards efforts. The list is not exhaustive.

Table 1: Standards. Source: Semiconductor Engineering

What needs to be standardized	Existing or proposed standards	Open or proprietary	Who is proposing/managing the standard
Design models			OCP/ODSA's CDX
Power models	Chip Power Model (CPM)		
Electrical properties	JEP30-E100		JEDEC
Thermal models	JEP30-T181		JEDEC
Chiplet connections	UCIe	open	AMD, Arm, ASE Group, Google Cloud, Intel, Meta, Microsoft, Qualcomm, Samsung, and TSMC
	AIB	Public on GitHub, opened to DARPA	Intel
	OpenHBI	open	OIF
	XSR		
	BoW		
Security	Optional security model	open	OCP/ODSA's CDX
Test — access for 3D stacked dies	IEEE 1838 standard		IEEE
Heterogeneous integration	IEEE roadmap		

Who is working on what standards

The Chiplet Design Exchange (CDX) has proposed thermal, physical, mechanical, I/O, behavioral, power, signal and power integrity, electrical properties, and test models for chiplets. Also included are optional traceability and operational security models.

CDX is a working group in the Open Compute Project (OCP) Foundation's sub-project Open Domain-Specific Architecture (ODSA). A recommendation document, "Proposed standardization of chiplet models for heterogeneous integration," was generated by Siemens EDA, Google, Palo Alto Electron, Cameron EDA, and Thrace Systems, with many more companies contributing comments, including Ansys, Applied Materials, Arm, ASE, Chipletz, Facebook, and Microchip.

"Successful deployment of chiplet-based devices requires the adoption of standardized chiplet

models to establish this emerging ecosystem," said the CDX working group in its standards proposal. The group said the models may need some tweaking, but most are usable in their current state and EDA vendors need to support the models in their tools and workflows. The group also calls for new standards and formats for 2.5 and 3D assembly rules. "EDA vendors and 2.5D/3D manufacturing vendors are currently collaborating on these new standards, which will hopefully be in place by the time the chiplet ecosystem becomes available to the broader market."

(See table 2 on page 46 for CDX's proposed models.)

IEEE Std 1838-2019 is the standard for 3D DFT that is being supported by the major EDA suppliers.

"Imec has made a test chip and ARM has reported on early application of the new standard," said Erik Jan Marinissen, imec's scientific director. "IEEE-SA started a new Study Group that is currently formulating a PAR (Project Authorization Request) to start a standard development working group into test and repair of 3D inter-die interconnects. This is driven by Intel. With their Ponte Vecchio, they have a product out in the market that has more than 75,000 of such inter-die interconnects, which makes effective and efficient testing and repair an economic necessity."

OpenHBI uses JEDEC's HBM3 electrical characteristics and I/O types.

Who is working on ecosystems — mini-consortia

Mini-consortia for chiplets are also springing up around the industry. These groups are loosely LEGO-like integration models for highly specific applications and end markets. All of these are starting small, because it's proving difficult to create a commercial marketplace for chiplets that can work across a wide variety of use cases.

It's one thing to connect chiplets using a standardized scheme, such as the Universal Chiplet Interconnect Express (UCIe), or using bridges developed by Intel or Samsung. It's quite another to expect them to work in heterogeneous devices under different compute loads and operating conditions.

Small consortia are the first signs of real progress. Foundries, packaging houses, and large chipmakers are working with partners with deep domain expertise using various arrangements to ensure various essential components work together and can be contextually characterized. A number of different approaches are on the table, from a handful of partners to more broad-based infrastructure, such as connectivity fabrics or standards coupled with stringent design rules. But for everyone involved, it's a cautious, learn-as-you-go process.

"We propose a set of standardized chiplet models to be provided by respective chiplet providers. It is strongly recommended that these

Table 2: Proposed standard models. Source: ODSA's Chiplet Design Exchange (CDX)/Siemens EDA

Model	Description
Thermal	ECXML – JEDEC JEP181
Physical, Mechanical, and IO	Library Exchange Format (LEF) GDSII or OASIS SPICE JEDEC JEP30-P101/ZEFXML Optional: Verilog to Physical Pin Mapping File (CSV)
Behavioral	SystemVerilog IEEE – 1800-2017 Recommended: Verilog-AMS 2.4 Optional: SystemC IEEE – 1666-2011 Optional: Bus Functional Model (BFM)
Power	Liberty (.LIB) IEEE2416 Standard for Power Modeling to Enable System Level Analysis Optional: Unified Power Format (UPF) – IEEE 1801-2018 or Chip Power Format (CPF) Optional: Verilog-AMS 2.4 Optional: SystemC IEEE – 1666-2011
Signal Integrity Analysis	IBIS/IBIS-AMI Optional: SPICE Netlist (for the IO driver and/or receiver) Optional: Channel Model
Power Integrity Analysis	Chip Power Model (CPM)
Electrical Properties	JEDEC JEP30-E101/ZEFXML
Test	Boundary Scan Description Language (BSDL) – IEEE 1149.1 BSDL – IEEE 1149.1/1149.6 ATPG Model - Primitive/UDP-based Verilog Internal JTAG (IJTAG) IEEE 1687 with PDL and ICL Optional: IEEE-1500 Core Test Language (CTL) Description for IP cores that require it Optional: IP Firmware (if applicable) Recommended: Gray-Box Level Netlist Full-Chip ATPG Vectors – STIL (IEEE1450.1) or Parallel WGL Full-Chip Memory BIST/Repair Vectors – STIL (IEEE1450.1) or Parallel WGL Optional: Unified Power Format (UPF) – IEEE 1801 or Chip Power Format (CPM)
Security	Optional: Security Agent
Documentation and Guidelines	General Chiplet Documentation SiP Physical Integration Guidelines SiP Test Guidelines Optional: Firmware (if applicable) Optional: Security

models are electronically readable for use in design workflows. The models should leverage existing industry standards that are readily available, with extensions to these standards and the addition of new standards to be defined as necessary," said a paper from Siemens EDA.

Others agree, and are putting that approach into action. "Samsung Electronics is building its own ecosystem," said Kevin Yee, director of IP and ecosystem marketing at Samsung Electronics. "Even in that, there's so much learning about things we hadn't thought about. From a technology perspective, chiplets are here and proven. But currently, they're all vertically integrated. So chiplets as a solution exists and is working. The challenge now is chiplets as a market or as a business. That's what everyone is working toward, and there's probably going to be an interim step. Everyone is going to start building their own micro-ecosystem first to make sure it works. So one company will do an I/O die, one will do the interconnect, one will do the data die. You will determine how to architect it and how to build something that will work. But the Holy Grail is when you can choose between 6 different compute dies, or 10 different I/O dies, select your memory dies, and then put it all together. We have a long way to go before that happens."

Likewise, Palo Alto Electron (PAe), which focuses on advanced packaging and chiplet design, is spearheading its own chiplet consortium that includes Promex Industries (system integration), Thrace Systems (power dissipation analysis), Palo Alto Electron (advanced packaging and chiplet design), iTest (reliability and failure analysis), Hyperion (system-level design, interconnects and advanced packaging), and Anemoi Software (thermal solvers).

"This is a minimum viable ecosystem, and all of these companies are in the United States," said Jawad Nasrulla, CEO of PAe. "We said, 'Okay, we're going to find small businesses and innovative startups that are attacking specific technical problems.' For example, Anemoi Software is focused on thermal modeling of chiplets. Power/thermal is a critical enabling technology for doing the design. Palo Alto Electron has experience building ICs, and in the last seven years we've been doing it with chiplets. Each of these companies is an expert in a certain area. Now we need to learn how to work with each other."

Dick Otte, CEO of Promex Industries, concurred. "Each one of these pieces is a multi-pronged thing. Each one has a capability, but everybody's got less than 100% of all possible capabilities. So the real question is going to be, when we have jobs rolling through here, how many of them will there be? And how frequently will we have to go find some support or added capability? So far, it hasn't been an issue."

Government gets in the act

In the United States, the CHIPS Act, signed into law in August 2022, will be pouring money into research. The new National Semiconductor

Technology Center will be looking at R&D, which will most likely include chiplets and heterogeneous integration.

In a position paper, the Semiconductor Alliance — a U.S. group of semiconductor industry and university members, formed in 2021 by MITRE Engenuity to guide the U.S. on investment strategies for all the money flowing from the new U.S. CHIPS Act, highlighted the possibility of chiplets as a breakthrough technology that should have investment.

The U.S. Department of Defense, meanwhile, began working with modular approaches to complex chips years before the CHIPS Act. In fact, the Defense Advanced Research Projects Agency (DARPA) has been working on chiplets since 2016, when it asked for bids from outside companies for its CHIPS program (not to be confused with the recent funding act passed by the U.S. Congress, many parts of which is managed through the Department of Commerce).

DARPA's CHIPS (Common Heterogeneous Integration and IP Reuse Strategies) goal is to devise a modular design and manufacturing flow for chiplets that the DoD can use to make advanced chips at a reasonable cost for the DoD, which isn't going to buy the huge volumes of chips that an automotive or mobile phone maker would. DARPA also plans to develop a large catalog of third-party chiplets for commercial and military apps. All told, the CHIPS flow is expected to lead to a 70% reduction in design cost and turn-around times.

DoD announced in early April 2023 that its State-of-the-Art (SOTA) Heterogeneous Integrated Packaging (SHIP) program delivered chiplet-based prototype devices for DoD uses. The prototypes, Intel's Multi-Chip Package (MCP-1) for SHIP Digital and Qorvo's Multi-Chip Module (MCM-1) for SHIP Radio Frequency (RF), will be tested demonstrated by BAE Systems, Inc. Both MCPs contain SOTA chiplets with advanced functionality, low power, smaller size, and cutting- edge performance, along with Intel Agilex advanced FPGA technology. Qorvo has a SHIP RF Design Center and Assembly & Test Center in Richardson, Texas, where their MCM-1 will replace legacy chip and wire design.

The SHIP Program still has more prototype devices in the works.

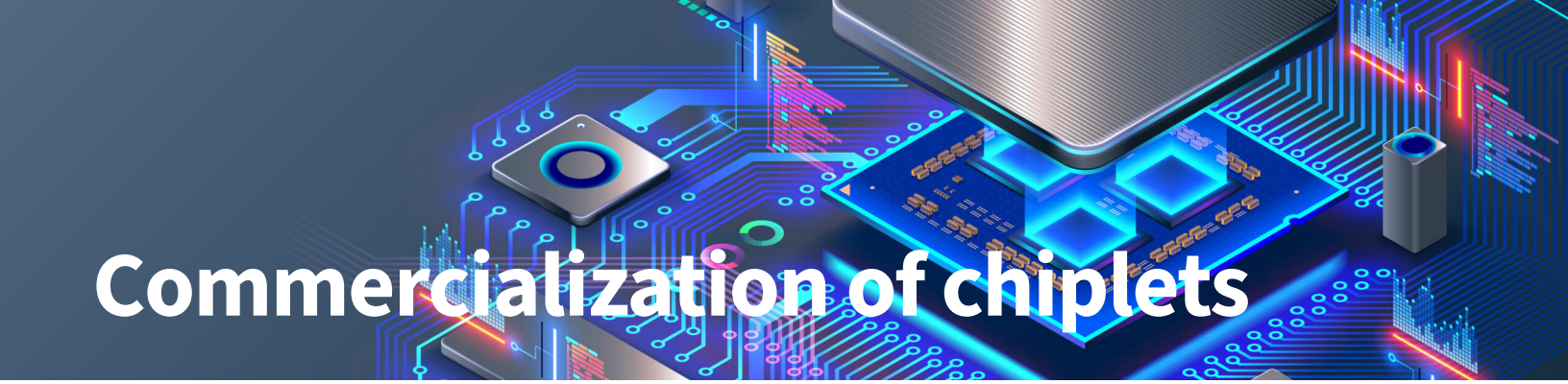
The U.S. DoD is far from alone. Governments across the world are getting involved in the semiconductor industry, and chiplets offer a way to both reduce design/development costs and to get involved with localized and specialized chiplets rather than building expensive SoCs.

In Europe, the European Union approved the European Chips Act in April 2023. The act will put €6 billion into European-based semiconductor research, production, and workforce. Private and additional public funding may put the investment at €15 billion. One of the EU's goals is to "build and reinforce capacity to innovate in the design, manufacturing and packaging of advanced chips." Part of the EU's Chips Act, the

Chips for Europe Initiative will invest in the EU's semiconductor technology and innovation capabilities. "It will ensure the deployment across Europe of advanced semiconductor design tools, pilot lines for prototyping the next generation of chips and testing facilities for innovative applications of latest chips technology." The Horizon program supports research.

Japan's METI (Ministry of Economy, Trade and Industry) is reworking its semiconductor strategy, with plans to invest 15 trillion yen ($108 billion) by 2030 for re-shoring chip production and investing in advanced technology. TSMC plans to build a fab in Japan. It remains to be seen what Japan will invest in chiplet production. Japan has always had an active packaging sector, including conferences produced by JIEP (the Japan Institute of Electronics Packaging) and SEMI Japan.

The U.S., Japan, South Korea, and Taiwan formed the Chip 4 Alliance, which aims to slow or stop advanced semiconductor manufacturing equipment and technology going to China, whose IC manufacturing industry is nascent.



Commercialization of chiplets

The idea of chiplets is not new. Chiplets have been proven to work extremely well in controlled situations and have been used for mostly for high performance computing. AMD and ASE have been collaborating on chiplets and graphics subsystems for a decade. Marvell has used chiplets since 2016. And Intel Foundry Services is customizing systems for its data center customers based on chiplets.

"In some ways, we are in a new era," said Bill Chen, fellow and senior technical advisor at ASE Group. "Optimization is a choice, and it's important for us to allow those choices. We also are in an era of opening up and sharing. With chiplets, SiP, and heterogeneous integration, we started with the most difficult area, which is high-performance computing. There are two reasons. One is this is the market sector that needs it the most. And second, this is where these products are being developed, and where the technology is there to address them. Then we can move on to communications and other markets, such as wearables and medical applications."

AMD was a pioneer of the chiplet. AMD's chiplets were introduced in 2015, for use in server market, as a way to distinguish itself from Intel's server chip products.

Marvell and Kandou Bus were working on the chiplet concept at about that time, as well. They announced a deal in 2016, under which Marvell would use Kandou's chip-to-chip interconnect technology to tie multiple chips together. Kandou is developing an ecosystem of small and midsize companies and agreed to give up some of its IP to others to jump-start this approach.

By 2022, Apple and AMD had processors with chiplets on the market. The companies were under production in high volumes. The use of these chiplets were and still are in large-volume devices. (Chiplets: Current Status).

"The advantage of systems consisting of many units is that the chiplet interface does not have to be standardized since large unit volumes are amenable to one-off development projects," said Fraunhofer's Andy Heinig. In a blog, he made the point that only large-volume applications will be using chiplets.

"The necessary ecosystem, consisting of interposer manufacturers as well as chiplet assembly on the interposers, is simpler to organize for large unit volumes since it is usually handled entirely by the chip maker," he said.

In another example, AMD developed its 3D V-Cache, a cache chiplet stacked on a processor. Both devices are based on TSMC's 7nm process.

In other developments:

- Fraunhofer IIS EAS will be using Achronix's field programmable gate arrays (FPGAs) to build chiplets. One application in this project will connect high-speed ADCs with Achronix's eFPGA IP for preprocessing in radar, wireless, and optical communication. The two entities are creating a demonstration platform that can be used in 5G/6G wireless infrastructure, ADAS, and high-performance test and measurement equipment.
- In January 2023, Intel launched its long delayed chiplet-based processor, code-named Sapphire Rapids, and Ponte Vecchio GPU graphics processor. Intel has plans to build a chiplet plant in Italy.
- Arm and TSMC made a 5nm chiplet test chip. Hauwei and Leti have both (separately) made AI processors using a chiplets.

The demand for chiplets is coming from several directions. "This is a real opportunity to help alleviate a lot of the challenges that companies see in this space," said Rob Mains, executive director for the CHIPS Alliance. "It requires a standardized interface. It needs a standardized PHY. It has to be instantiated for a particular chip process technology, or packaging technologies. And then it requires an EDA ecosystem that goes along with that. DARPA's vision is right, and it is a matter of getting the education level set with design teams globally. This will bring an understanding of the benefits and provide

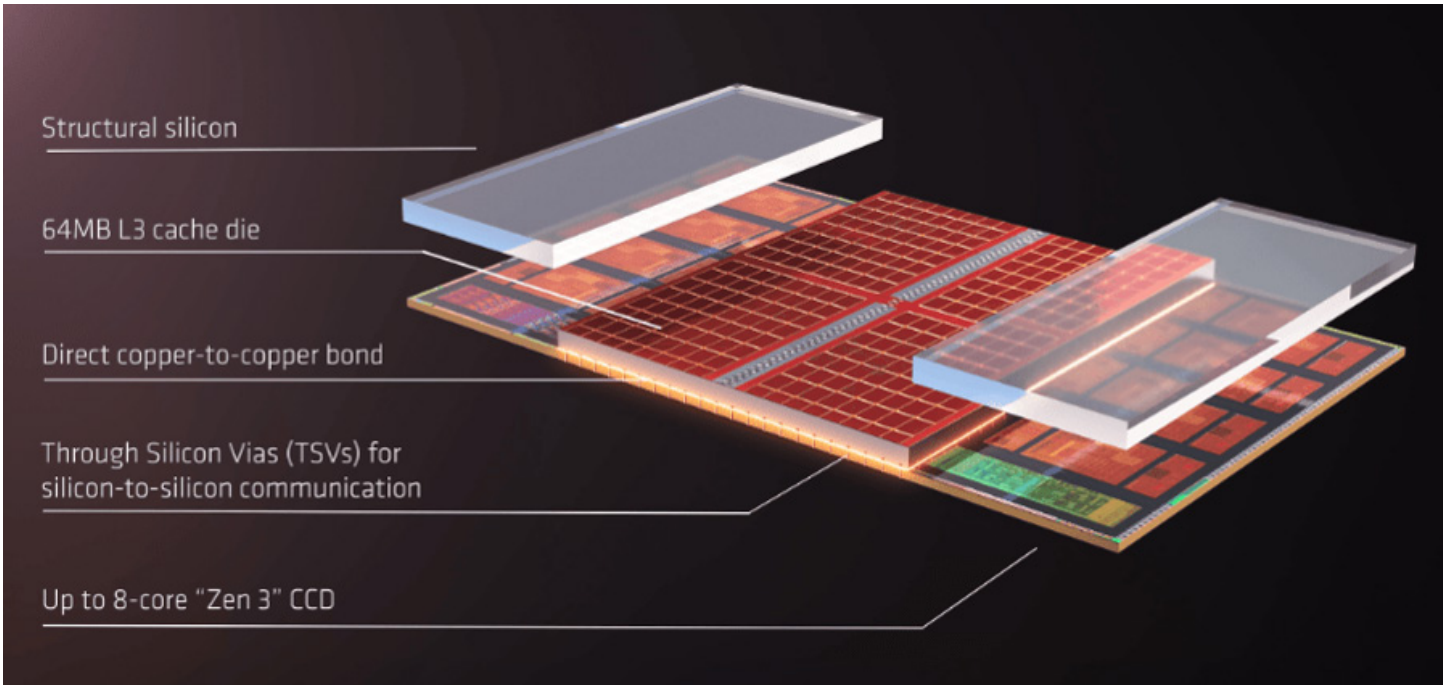


Fig. 14: AMD's 3D V-Cache stacks the cache on a processor. Source: AMD

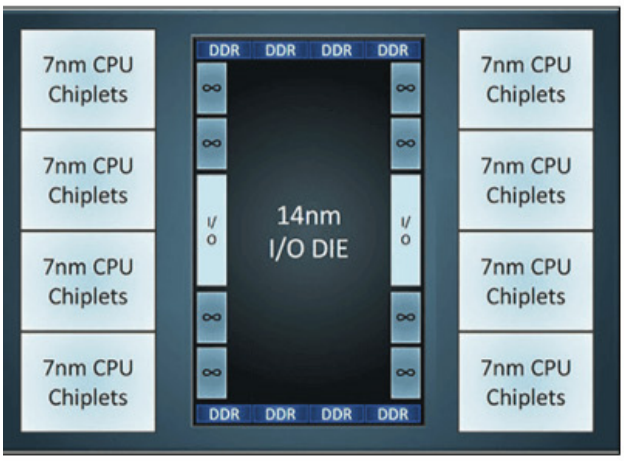
a level of assurance that it will yield effective results."

Today, we are on the cusp of another level of specialization, where a company will only design the system — and design and implement the pieces of the system that provide differentiation — without having to worry about the implementation or fabrication of commodity parts of the design. These would be available in the form of chiplets, which are fully implemented and fabricated pieces that can be assembled with custom silicon to form a system. To get there, the industry needs some new standards.

The pioneering work has been done by large systems companies, which own both the system and the chiplets. This enables them to make larger, or more modular offerings, and along the way to iron out many of the kinks. They have developed proprietary means to make these systems.

Unsurprisingly, there is significant variation in those solutions. "The industry is divided up into a range of offerings, simply because there was a need by ASIC companies," said Ketan Mehta, senior director of SoC IP product marketing at OpenFive. "They are building custom silicon, and they want solutions right away. They don't want to wait for the standards to be developed and evolve. So proprietary implementations are being developed and proven in all these companies."

Companies are also working on RISC-V chiplets. Tenstorrent is making RISC-V CPU chiplets for AI, destined for edge devices, edge servers, and cloud servers. The company adopted Arteris IP's Ncore and FlexNoC interconnect IP.



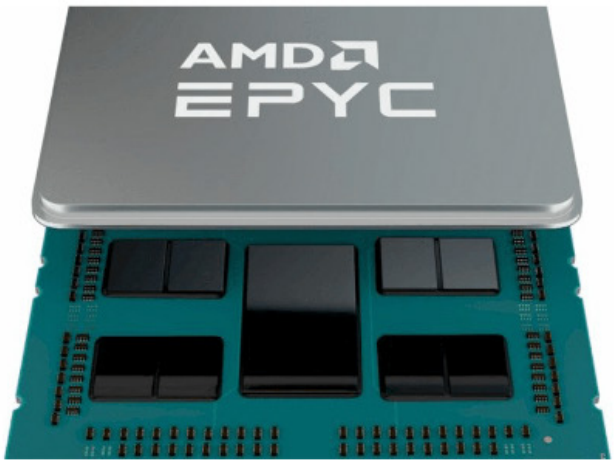


Fig. 15: AMD's EPYC architecture integrates different process technologies. Source: Synopsys/ AMD/Chiplet Summit

Startup funding

This list shows the chiplet and chiplet-related startups that investors funded from January 2022 through April 2023, as tracked by Semiconductor Engineering each month.

Chiplets

April 2023

JoinSilicon Microelectronics raised over CNY 100.0M (~$14.5M) in Series A funding led by **Aplus Capital**, joined by **ZJ Innopark** and **Z&Y Capital**. JoinSilicon provides ASIC design services, including front-end and back-end design, IP integration and customization, verification, package design, and ATE test design. It also develops a range of high-speed data and memory interface IP. Funds will be used for R&D of additional high-speed interface IP and chiplet products. Based in Nanjing, China, it was founded in 2021.

January 2023

M Square raised over CNY 100.0M (~$14.7M) in Series A funding from **Delta Capital**, **Oceanpine Capital**, and **C*Core Technology**. M Square develops interface IP and chiplets. Its products include USB, PCIe, SATA, SerDes, MIPI, DDR, HDMI, DisplayPort, and high-bandwidth memory (HBM). It is currently working on developing LPDDR5X, ONFI 5.1, and UCIe IP. Funds will be used for R&D. The startup is eyeing an IPO in 2025. Founded in 2021, it is based in Shanghai, China.

November 2022

Eliyan Corporation closed a $40.0M Series A round led by **Tracker Capital**, joined by **Celesta Capital** and strategic investors including **Intel Capital** and **Micron Ventures**. Eliyan develops chiplet interconnect technologies that work with organic substrates and don't require advanced packaging solutions such as silicon interposers. Its NuLink technology, which is a superset of Bunch of Wires (BoW) and Universal Chiplet Interconnect Express (UCIe), is a die-to-die (D2D) PHY to connect different functions in one package over any substrate. Eliyan claimed a recent 5nm tapeout achieved twice the bandwidth at less than half the power consumption of current interconnect methods. It has also developed technology for 2.5/3D implementations that enables mix and match of chiplets with different die-to-die interfaces in different processes, such as DRAM and SOI. "Our extensive background in developing bleeding-edge technologies in this space led us to focus on a key challenge: interconnect improvements for system-in-package and chip-to-memory architectures as the path to deliver performance scaling," said Eliyan CEO and co-founder Ramin Farjadrad. "Our approach supports and is compliant with the overall industry move toward chiplet-optimized interconnect protocols, including the UCIe standard as well as High Bandwidth Memory (HBM) protocols." Its first silicon is expected in the first quarter of 2023. Based in Santa Clara, California, USA, it was founded in 2021.

October 2022

Polar Bear Tech raised CNY 150.0M (~$20.9M) in angel+ financing from **Inno-Chip**, **iFlytek**, **China Fortune-Innovation Capital**, and others, following an earlier round in May. The startup provides design services for custom algorithm accelerator server chips. It takes a chiplet-based approach, which it says reduces the time and cost for a full design compared to traditional ASIC solutions. The company's first lightweight Hub Die chiplet has been taped out and will be released soon. It has also developed several NPU chiplets. Funds will be used for development on the next generation of its general-purpose Hub Die chiplet and related interface technologies. Founded in 2021, it is based in Xi'An, China.

September 2022

Sunlune raised tens of millions of dollars in Series A financing from **Intel Capital**. Sunlune makes 3D stacked high-throughput chips that integrate storage and processing. It targets big data and Web3 applications in data centers and cloud computing. Funds will be used for R&D, hiring, and global expansion. Based in Beijing, China, it was founded in 2009.

August 2022

Ventana Micro Systems raised $55.0M in new funding. Ventana is developing data center class high-performance RISC-V CPUs with extensible instruction set capability delivered in the form of multi-core chiplets. The company also offers a customizable SoC chiplet for cloud, enterprise data center, 5G, edge compute, and automotive applications. Funds will be used to productize its chiplets, and it expects tape out of its first UCIe (Universal Chiplet Interconnect Express) chiplet products in the second half of next year. Ventana's cores and chiplets will be available through the Intel Foundry Services ecosystem. Based in Cupertino, California, USA, it was founded in 2018.

Nuclei Technology raised hundreds of millions of yuan (CNY 100.0M is ~$14.7M) in new financing led by **Legend Capital** and joined by **Accurate Capital**, **Bluerun Ventures**, **C&D Emerging Investment**, **Xiaomi**, **Shanghai International Group**, **Shougang Fund**, **Sunic Capital**, **Xianghui Capital**, **Zhejiang Paradise**, and **Zhongguancun Development Frontier Enterprise Investment Fund**. Nuclei Technology develops RISC-V CPU IP. It offers a range of processors from low-power 32-bit, 2 stage pipeline cores for MCUs and IoT devices through to high-performance 64-bit, 9 state pipeline architectures capable of running Linux for data center, networking, and baseband applications. Funds will be used to develop IP targeted for security, automotive, and AIoT applications. It also plans to develop hard IP for chiplets and explore heterogeneous systems using its cores. Based in Shanghai, China, it was founded in 2018.

KiwiMoore raised CNY 100.0M (~$14.7M) in angel and seed funding led by **CasStar** and joined by **Fosun Capital**, **Junsan Capital**, **Step Fund**, and others. KiwiMoore designs 2.5D and 3D IC

chiplets and enabling technology, including a high-performance silicon interposer and chiplet software design platform for high-speed communication interfaces, distributed near memory compute, and high-efficiency power networks. The startup also offers an advanced packaging shuttle line, a library of chiplets, and design services. It targets chips designed for markets such as next-generation data centers, autonomous driving, and the metaverse. Funds will be used for R&D, hiring, and marketing. Based in Shanghai, China, it was founded in 2021.

Chiplego Technology raised nearly CNY 300.0M (~$44.1M) in angel and strategic funding rounds, which included **Matrix Partners China** and **Summitview Capital**. Chiplego is using chiplets to design high-performance chips for smart vehicles. The startup has developed its own high-bandwidth, low-latency interconnect technology and what it calls an embedded HPC architecture. Based in Shanghai, China, it was founded in November 2021.

July 2022

M2 Semi received more than CNY 100.0M (~$14.9M) in pre-Series A funding led by **Delta Capital**. M2 Semi develops high-performance CPUs using a chiplet-based design for the cloud server market. It has also designed die-to-die interconnect IP for chiplet integration that uses a fusion architecture and can support a variety of interface protocols, multiple package types, and mainstream process nodes. Based in Beijing, China, it was founded in 2021.

Chipuller received Series A investment from **Oriental Fortune Capital**. Chipuller has developed an active, reprogrammable interposer used to assemble chiplets into packages. The interposer provides reconfigurable routing and integrated power management to aid in power gating, including integrated power/load switches with slew rate control, high-efficiency low-Iq DC-DC converters, and support for multiple power domains and embedded level shifters. The company uses this interposer as the basis for its chiplet integration platform, which gives customers access to design software with floorplanning, schematic capture, and layout capabilities along with a library of chiplets including sensors, memories, microprocessors, and RF. It says its platform enables designs to be built quickly and with the flexibility to change the design during testing. Founded in 2021, it is based in Shenzhen, China.

June 2022

Monozukuri, also known as **MZ Technologies**, received a strategic investment from industrial holding company **MARCAP**. The investment gives MARCAP a 16% equity stake in Monozukuri. Monozukuri has developed tools for IC/package co-design that enable system design across multiple levels and components including die, chiplets, silicon interposer, package, and surrounding PCB for 2D, 2.5D, and 3D designs. It offers a single, consistent interconnect manager that represents and maintains the interconnect model of the entire system along with a cross-hierarchical 3D-aware pathfinding engine that

identifies the best interconnect in a 3D space. The latest version of its tool suite introduces early parasitic estimation and stack planning functionality to compare multiple floorplans. The tools integrate with existing flows and implementation platforms. "We've invested a lot in developing a technology platform with unique features and now we are ready to serve global markets, thanks to important contacts with major Silicon Valley companies. We will continue to enrich GENIO with new exclusive features, because chiplet technology represents the future of the microchip industry," said Anna Fontanelli, Monozukuri founder and CEO. Founded in 2015, the company got its start with the HIPER project under Europe's Horizon 2020 program. It is based in Rome, Italy.

April 2022

D-Matrix raised $44.0M in Series A funding led by **Playground Global**, joined by Microsoft's **M12**, **SK Hynix**, and existing investors **Nautilus Venture Partners**, **Marvell Technology**, and **Entrada Ventures**. D-Matrix has developed a digital in-memory computing (DIMC) architecture that targets Transformer AI data center inference workloads, along with accompanying ML tools, software, and algorithms. It also announced a chiplet based on DIMC technology, with another to be released soon. "We've developed a path-breaking compute architecture that is all-digital, making it practical to implement while advancing AI compute efficiency far past the memory wall it has hit today," said Sid Sheth, founder, president & CEO at D-Matrix. The funds will be used to build out its product roadmap and hire. Founded in 2019, it is based in Santa Clara, California, USA.

Ayar Labs raised $130.0M in Series C funding led by **Boardman Bay Capital Management** and joined by new investors including **Hewlett Packard Enterprise**, **Nvidia**, **Agave SPV**, **Atreides Capital**, **Berkeley Frontier Fund**, **IAG Capital Partners**, **Infinitum Capital**, **Nautilus Venture Partners**, and **Tyche Partners**, as well as existing strategic investors **Applied Ventures**, **GlobalFoundries**, **Intel Capital**, and **Lockheed Martin Ventures**. Ayar Labs offers a chip-to-chip optical interconnect solution it says eliminates the bottlenecks associated with system bandwidth, power consumption, latency, and reach for AI, HPC, cloud, telecommunications, aerospace, and remote sensing applications. The company has been working with GlobalFoundries on the development of its GF Fotonix platform. "The overall financing is much larger than we originally targeted, underscoring the market opportunity for optical I/O and Ayar Labs' leadership position in silicon photonics-based interconnect solutions," said Charles Wuischpard, CEO of Ayar Labs. "This financing allows us to fully qualify our solution against industry standards for quality and reliability and scale production starting this year." Based in Santa Clara, California, USA, it was founded in 2015.

February 2022

Ayar Labs received a strategic investment from **Hewlett Packard Pathfinder**. The company develops optical based interconnect chiplets

and multi-wavelength lasers to replace traditional electrical based I/O. As part of the agreement, Hewlett Packard Enterprise and Ayar Labs will partner on photonics research and commercial development, particularly high speed, high density, low power optical-based interconnects to target future generations of the HPE Slingshot high performance Ethernet fabric designed for HPC and AI solutions. Based in Emeryville, California, USA, it was founded in 2015.

Packaging

April 2023

SJ Semiconductor raised $340.0M in Series C+ financing from **Legend Capital**, **Goldstone Investment**, **Jade Stone Venture**, **Shang Qi Capital**, **Leafoison Capital**, **TCL Capital**, **China Fortune-Innovation Capital**, and **GLP-C&D Capital**, and existing shareholders **Oriza Rivertown** and **Oriza Hua Capital**. SJSemi is a 12-inch middle-end-of-line (MEOL) packaging foundry that specializes in advanced wafer bumping, 3D multi-die integration technology, wafer-level chip-scale packaging (WLCSP), wafer-level antenna-in-package, and testing. It offers a system-in-package platform for 3D heterogeneous integration of active and passive devices that comprises Cu vertical interconnect, multi-layer double sided fine pitch RDLs, wafer-level multi-layer precise aligned antennas, flip-chip, and SMT. Founded in 2014, it is based in Jiangyin, China, and has raised over $1B.

EMC Semi-Conductor Technology drew over CNY 100.0M (~$14.5M) in angel funding led by **FutureX Capital**, joined by **Wuzhong Financial Holdings**, **Co-Win Ventures**, and others. EMC manufactures multi-layer high-density glass fiber free substrates for semiconductor packaging. The startup says its coreless substrate improves plasticity and reliability while lowering costs. Its products include substrates for wire bond and flip-chip chip-scale packages (WB-CSP/FC-CSP); wire bond and flip-chip ball-grid array packages (WB-BGA/FC-BGA); system-level substrates for system-in-package, multi-chip modules, and chiplets; and an embedded substrate that can integrate active and passive components. Funds will be used in construction of a production line. Founded in 2022, it is based in Suzhou, China.

February 2023

Lux Semiconductors raised $2.3M in seed funding led by **Ultratech Capital Partners**, with participation from **AIN Ventures**, **Hemisphere Ventures**, and **Lockheed Martin Ventures**. Lux has developed a System-on-Foil advanced packaging platform designed for chiplet-based heterogeneous integration, system-in-packages, and multi-chip modules. The startup says System-on-Foil allows for many small die or chiplets to be closely connected through high-density, high-speed interconnects for reduced package size and weight while improving high frequency performance and thermal management compared to existing materials. The low warpage metal substrate provides

rigidity down to 100 µm thickness, with ultra-thin options under 25 µm available to support flexible applications. Founded in 2017, it is based in Albany, New York, USA.

September 2022

iSabers Materials raised nearly CNY 200.0M (~$28.6M) in a Series A++ round from **Beijing Integrated Circuit Advanced Chip Fund**, **SunGrow Power Supply**, **Beijing Futian Industrial Investment Holding Group**, **CCB Trust**, and **Wofo Capital**. iSabers Materials makes silicon carbide (SiC) substrates for advanced packaging and heterogeneous integration. Funds will be used for developing new production lines. Founded in 2020, it is based in Beijing, China.

July 2022

CAS Microelectronics Integration Technology, or **Casmeit**, raised over CNY 150.0M (~$22.4M) in Series B financing from **Xin Ding Capital**, **New Hope Capital**, and **Xiamen Hengxing Group**. Casmeit provides advanced packaging services, including fan-out, fan-in, flip-chip, 2.5D/3D stacked packaging, wafer-level chip scale packaging, and through-silicon vias. It also provides test services, including reliability, thermal, and failure analysis. Based in Xuzhou, China, it was founded in 2018.

June 2022

UnitySC raised €48.0M (~$51.4M) in venture funding led by **Jolt Capital**, the French State through **French Tech Souveraineté**, and **Supernova Invest**. UnitySC offers metrology and inspection equipment for the semiconductor industry. It makes measurement equipment for advanced packaging, including 3D IC, that combines automated optical inspection and 3D imaging with high depth of focus line scanning, temporal-mode interferometry, spectrometry, and phase shift analysis. It also offers a range of equipment dedicated to other types of high-end processes, including patterned and un-patterned defect inspection for compound semiconductors, transparent substrates, or specialty devices such as MEMS and power, automotive, and RF. A spin out from Fogale Nanotech in 2016, it is based in Montbonnot, France.

April 2022

Teyan Semiconductor raised “tens of millions of yuan” (CNY 10.0M is ~$1.6M) in a Series A round from **Vinno Capital**. Teyan Semiconductor provides laser, plasma, and sputter semiconductor manufacturing equipment for both wafers and substrate/package. Its offerings include laser marking, cutting, drilling, grooving, dicing, debonding, and solder bumping equipment. The company says its equipment is suitable for advanced packaging production lines such as SiP, fan-out, chiplet, and 3D. Founded in 2017, it is based in Shenzhen, China.

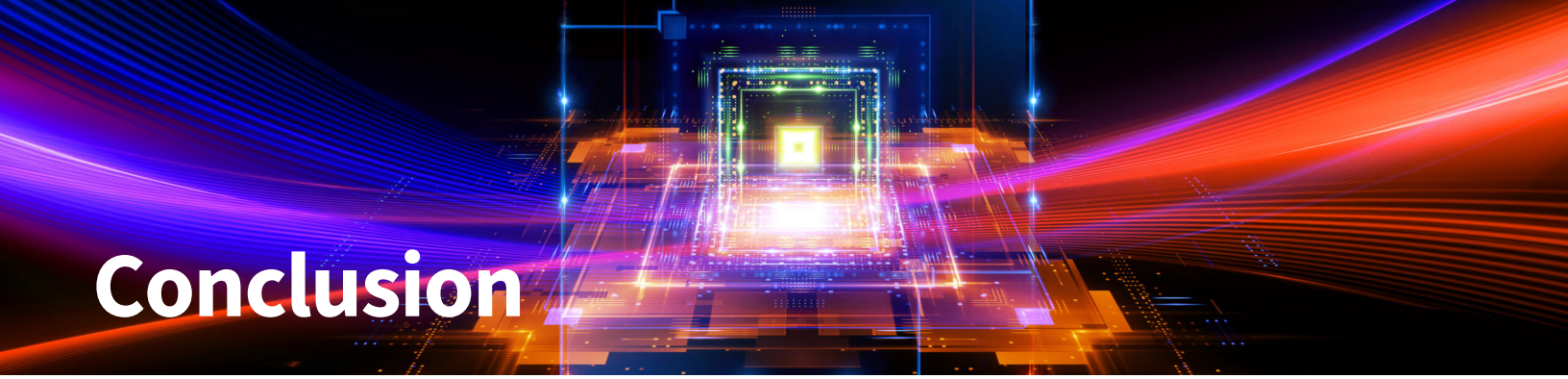
January 2022

National Center For Advanced Packaging China (**NCAP China**) also known as **Huajin Semiconductor**, raised over CNY 200.0M

(~$31.6M) in Series A financing from **Shenzhen Capital Group**, **Turing Ventures**, and others. NCAP China is an R&D and commercialization center for advanced packaging, testing, and system integration. Current research directions include 2.5D/3D technology and TSV, high density wafer level packaging, SiP product development capabilities, and advanced materials and equipment technologies for microelectronics packaging. Founded in 2012, it is based in Wuxi, China.

Table 3: Terminology used in system in packages (SiP) of which chiplets are a part. Source: Semiconductor Engineering

Company	Type of chiplet related products	Supersector	Sector	Subsector	HQ	City	Year Start-ed	$ Made	Fund Type	Month Funded 2022, 2023
Chiplego Technology	*Chiplet*	*Automotive*	*Auto Chips*	*ADAS/AV Chip*	*China*	*Shanghai*	*2021*	*$44.1*	*Angel*	*08 (22)*
Chipuller	*Integration*	*Chips*	*Processors, Logic*	*IP, Chiplet, Arch.*	*China*	*Shenzhen*	*2021*	*No info.*	*Series A*	*07 (22)*
M2 Semi	*Chiplet*	*Chips*	*Processors, Logic*	*CPU*	*China*	*Beijing*	*2021*	*$14.9+*	*Pre-A*	*07 (22)*
Nuclei Technology	*Chiplet*	*Chips*	*Processors, Logic*	*CPU*	*China*	*Shanghai*	*2018*	*$14.7+*	*Venture*	*08 (22)*
Sunlune	*3D Stacked IC*	*Chips*	*Processors, Logic*	*Data Center, Network*	*China*	*Xiamen*	*2009*	*$10.0+*	*Series A*	*09 (22)*
Ventana Micro Systems	*Chiplet*	*Chips*	*Processors, Logic*	*CPU*	*USA*	*Cupertino*	*2018*	*$55.0*	*Venture*	*08 (22)*
D-Matrix	*Chiplet*	*Chips*	*AI HW*	*Data Center AI*	*USA*	*Santa Clara*	*2019*	*$44.0*	*Series A*	*04 (22)*
Eliyan Corporation	*Integration*	*Chips*	*Processors, Logic*	*IP, Chiplet, Arch.*	*USA*	*Santa Clara*	*2021*	*$40.0*	*Series A*	*11 (22)*
JoinSilicon Microelectronics	*Chiplet*	*Chips*	*Processors, Logic*	*Design Services*	*China*	*Nanjing*	*2021*	*$29.4*	*Pre-A, Series A*	*04 (23), 07 (22)*
Polar Bear Tech	*Chiplet*	*Chips*	*Processors, Logic*	*Design Services*	*China*	*Xi'an*	*2021*	*$20.9*	*Angel*	*05, 10 (22)*
M Square	*Chiplet*	*Chips*	*AMS*	*Interface*	*China*	*Shanghai*	*2021*	*$14.7*	*Series A*	*01 (23)*
KiwiMoore	*Chiplet*	*Chips*	*Processors, Logic*	*IP, Chiplet, Arch.*	*China*	*Shanghai*	*2021*	*$14.7*	*Angel, Seed*	*08 (22)*
Monozukuri	*Chiplet EDA*	*EDA*	*EDA*	*Design*	*Italy*	*Rome*	*2015*	*No info.*	*Strategic*	*06 (22)*
Teyan Semiconductor	*Packaging*	*Manufacturing*	*Equipment*	*Package Equip*	*China*	*Shenzhen*	*2017*	*$1.6+*	*Series A*	*04 (22)*
UnitySC	*3D IC Inspection*	*Manufacturing*	*Equipment*	*Inspection, Metrology*	*France*	*Montbonnot*	*2016*	*$51.4*	*Venture*	*06 (22)*
iSabers Materials	*Packaging*	*Manufacturing*	*Materials*	*Packaging Substrate*	*China*	*Beijing*	*2020*	*$28.6*	*Series A+*	*09 (22)*
EMC Semi-Conductor Technology	*Packaging*	*Manufacturing*	*Materials*	*Packaging Substrate*	*China*	*Suzhou*	*2022*	*$14.5*	*Angel*	*04 (23)*
Ayar Labs	*Chiplet*	*Other Tech*	*Photonics*	*Optical Interconnect*	*USA*	*Santa Clara*	*2015*	*$130.0*	*Series C*	*02, 04 (22)*
Casmeit	*Packaging*	*Packaging & Test*	*Packaging*	*Adv. Packaging*	*China*	*Xuzhou*	*2018*	*$22.4+*	*Series B*	*07 (22)*
SJ Semiconductor	*Packaging*	*Packaging & Test*	*Packaging*	*Adv. Packaging*	*China*	*Jiangyin*	*2014*	*$340.0*	*Series C+*	*04 (23)*
NCAP China	*Packaging*	*Packaging & Test*	*Packaging*	*Adv. Packaging*	*China*	*Wuxi*	*2012*	*$31.6*	*Series A*	*01 (22)*
Lux Semiconductors	*Packaging*	*Packaging & Test*	*Packaging*	*Adv. Packaging*	*USA*	*Albany*	*2017*	*$2.3*	*Seed*	*02 (23)*



Conclusion

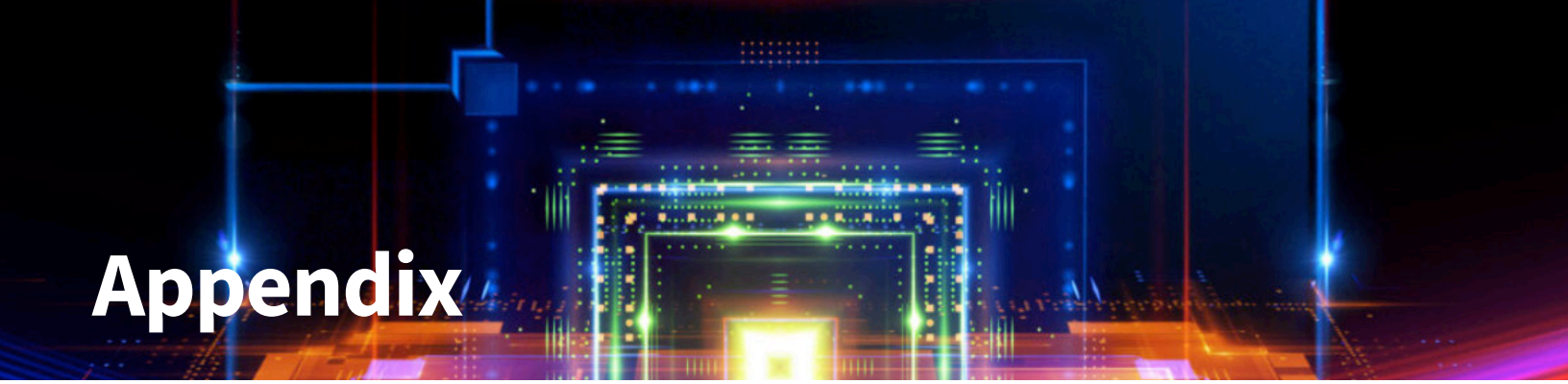
While there are numerous technical issues, work is underway to solve all of these. The bigger challenge for chiplets has always been the business side.

"From a technical point of view, it is not hard," said Kandou's Shokrollahi. "It really is a LEGO-like approach. There is, of course, the whole business side of it — creating the infrastructure and the ecosystems for it and getting companies to adopt it. All this has to come together in order to have this thing take off. But we have seen this in other industries. It has happened. Once we went from mainframes to PCs, that's exactly what happened. And that was enormously successful. And we think here is very similar."

The challenge now is to begin developing chiplets and chiplet-friendly architectures that can be sold commercially for any vendors' devices. But as companies at the leading edge of this shift will attest, sourcing chiplets from different vendors adds a whole new set of challenges. And even if the idea makes sense in theory, it never has been realized. Now small consortia and standards bodies are showing the first signs of real progress.

Monolithic die are being decomposed into functional blocks, and those functional blocks are being hardened. The challenges will be putting systems together in packages using those components in a standardized way, and then re-using those components in other systems as needed. This has been talked about for years, but the industry appears to be finally on board with this concept.

Improving testability is also important. "The fundamental problem is how to get the most optimal implementation that is manufacturable and testable with the least risk and in the most productive way," said Shekhar Kapoor, senior product director for multi-die systems at Synopsys. "There are existing tools and approaches that can be used to implement multi-die systems. However, to address the increased complexity of heterogeneous integration systems, and to do it efficiently while reducing the risks, requires a more holistic approach. It really boils down to having a comprehensive solution including very diligent architecture design, hardware-software partitioning, chiplet design planning, and integrating manufacturing testability. Other critical considerations include thermal and power delivery earlier in the design cycle."



Appendix

Terminology

Table 4: Terminology used in system in packages (SiP) of which chiplets are a part. Source: Semiconductor Engineering

Terms	Definitions
System in a Package (SiP)	*An umbrella term for a package with multiple dies, versus a system on chip (SoC) that has all the system blocks on one die. An SiP disaggregates all the blocks into separate dies/dielets/chiplets on the same substrate, which are connected to function as one IC.*
2.5D	*A type of SiP in which chiplets are placed side by side or vertically on an interposer. The packaging and EDA workflows for 2.5D designs used for HBM memories are relatively mature. Xilinx first began using a 2.5D approach in 2011.*
3D-IC	*A type of SiP in which chiplets are stacked vertically. The big issue with 3D historically has been thermal, but through proper placement and high-speed interconnects, this recently has proven to be less of an issue.*
Chiplet	*A component used on a SiP, the chiplet is a separate die/dielet/tile that is integrated with other die on a substrate or interposer.*
Multi-Chip Modules	*An early incarnation of advanced packaging, dating back to the 1990s, that integrated multiple chips.*
Heterogeneous Integration	*Heterogeneous integration refers to disaggregating chips — typically SoCs — into component parts and connecting them all together. Because the dies can be developed using different process nodes, they need to be fully characterized. These devices can be extremely complex, and they also can be quite large.*

Sources, references, further reading

The following sources were used in this book or are relevant content for further reading. Catch up on Semiconductor Engineering's latest chiplets content in the Knowledge Center.

Relevant staff-written articles on Semiconductor Engineering

Mutschler, Ann, "Chiplet Planning Kicks Into High Gear," Issues involving design, manufacturing, packaging, and observability all need to be solved before this approach goes mainstream for many applications. MAY 25TH, 2023. https://semiengineering.com/chiplet-planning-kicks-into-high-gear/

Mutschler, Ann, "Automotive Relationships Shifting With Chiplets," As the automotive ecosystem balances the best approaches for designing in increasingly advanced features, how companies interact is still evolving. MAY 24TH, 2023. https://semiengineering.com/automotive-relationships-shifting-with-chiplets/

Bailey, Brian, "True 3D-IC Problems," Stacking logic requires solving some hidden issues; concerns about thermal dissipation may be the least of them. APRIL 27TH, 2023, https://semiengineering.com/true-3d-ic-problems/

Appendix

Sperling, Ed, "Mini-Consortia Forming Around Chiplets," Commercial chiplet marketplaces are still on the distant horizon, but companies are getting an early start with more limited partnerships. MARCH 20TH, 2023, https://semiengineering.com/mini-consortia-forming-around-chiplets/

Peters, Laura, "Enabling Test Strategies For 2.5D, 3D Stacked ICs," Better standards, 3D DFT, and next-generation probes are a great start toward fully testing these complex systems. SEPTEMBER 8TH, 2022. https://semiengineering.com/enabling-test-strategies-for-2-5d-3d-stacked-ics/

LaPedus, Mark, "Piecing Together Chiplets," Changes that could push this packaging approach into the mainstream, and the challenges ahead. JULY 22ND, 2021. https://semiengineering.com/piecing-together-chiplets/

Bailey, Brian, "Many Chiplet Challenges Ahead," Assembling systems from physical IP is gaining mindshare, but there are technical, business and logistical issues that need to be resolved before this will work. APRIL 12TH, 2021. https://semiengineering.com/many-chiplet-challenges-ahead/

Bailey, Brian, "Waiting For Chiplet Standards," An ecosystem is required to make chiplets a viable strategy for long-term success, and ecosystems are built around standards. Those standards are beginning to emerge today. MARCH 25TH, 2021. https://semiengineering.com/waiting-for-chiplet-standards/

Bailey, Brian, "Designing 2.5D Systems," Connecting dies using an interposer requires new and modified processes, as well as organizational changes. MARCH 15TH, 2021. https://semiengineering.com/designing-2-5d-systems/

Mutschler, Ann, "Getting Serious About Chiplets," Issues involving known good die and test still remain, but this approach is getting a lot of interest. JANUARY 8TH, 2018. https://semiengineering.com/getting-serious-about-chiplets/

Mutschler, Ann, "Chiplets Gaining Steam," Hard IP could reduce time and cost for heterogeneous designs, but there are still challenges to solve. NOVEMBER 6TH, 2017. https://semiengineering.com/chiplets-gaining-steam/

Mutschler, Ann, "DARPA CHIPS Program Pushes For Chiplets," The U.S. government takes its own approach to heterogeneous integration. SEPTEMBER 14TH, 2017. https://semiengineering.com/darpa-chips-program-pushes-for-chiplets/

Appendix

Sperling, Ed, "The Chiplet Option," More companies are assessing pre-built and pre-verified circuits as a way of reducing time to market. OCTOBER 9TH, 2017. https://semiengineering.com/the-chiplet-option/

Blogs

Fraunhofer Institute of Integrated Circuits, Division of Engineering and Adaptive Systems.

Heinig, Andy, "Chiplets: More Standards Needed," Current chiplet interface standardization efforts fall short when it comes to handling analog signals and power. APRIL 13TH, 2023, https://semiengineering.com/chiplets-more-standards-needed/.

Heinig, Andy, "Chiplets: Current Status," Recent products show that even large and complex chiplet structures can be designed reliably and efficiently. MAY 12TH, 2022. https://semiengineering.com/chiplets-current-status/

Amkor Technology

Chaudhry, Vik, "An OSAT Perspective On Semiconductor Market Trends," Advanced packaging is expected to exceed the mainstream segment for the first time by 2027. https://semiengineering.com/an-osat-perspective-on-semiconductor-market-trends/

Achronix

ACHRONIX, "Speedchip FPGA Chiplets," Why integrating an FPGA chiplet and ASIC die in same package offers flexibility and high performance. JANUARY 9TH, 2019. https://semiengineering.com/speedchip-fpga-chiplets/

Siemens EDA

Medhat, Dina. "2.5/3D IC Reliability Verification Has Come A Long Way," A systematic methodology to verify latch-up prevention and ESD protection of 2.5D and 3D ICs. AUGUST 1ST, 2022. https://semiengineering.com/2-5-3d-ic-reliability-verification-has-come-a-long-way/

Videos

Kandou

Sperling, Ed; Shokrollahi, Amin, "The Case For Chiplets," Issues in advanced packaging. DECEMBER 10TH, 2018. https://semiengineering.com/the-case-for-chiplets/

ASE

Sperling, Ed; Chen, William, "System In Package (2017)," Tech Talk: William Chen, ASE fellow and senior technical advisor, talks with Semiconductor Engineering about advanced packaging options and why they are so important. JUNE 4, 2017. https://youtu.be/fPUEmw-fiG8

Appendix

Synopsys

Sperling, Ed; Posner, Michael, "Multi-Die Integration," Advantages and challenges for heterogeneous integration in advanced packaging. JANUARY 4TH, 2023. https://semiengineering.com/multi-die-integration/

Flex Logix

Sperling, Ed; Tate, Geoff. "EFPGAs Vs. FPGA Chiplets," Which approach works best where. JULY 22ND, 2020. https://semiengineering.com/efpgas-vs-fpga-chiplets/

Promex Industries

Sperling, Ed; Otte, Dick. "Heterogeneous Integration Issues And Developments," Why putting different chips and chiplets into a package is harder than it sounds. DECEMBER 14TH, 2022. https://semiengineering.com/heterogeneous-integration-issues-and-developments/

ClioSoft

Sperling, Ed; Pires, Pedro. "Managing IP In Heterogeneous Designs," How an interoperability layer is essential to tracking IP. NOVEMBER 14TH, 2022. https://semiengineering.com/managing-ip-in-heterogeneous-designs/

Relevant technical papers

See packaging section or the general technical paper home page.

Framework To Compile Quantum Programs Onto Chiplets (UCSB, Cisco)

Data-Centric Reconfigurable Array Chiplets (Princeton)

Design Considerations And Recent Advancements In Chiplets (UC Berkeley/ Peking University)

Chiplet Placer With Thermal Consideration For 2.5D ICs

Modeling And Thermal Analysis Of 3DIC

All-Digital MDL-Based Fast Lock Clock Generator For Low-Power Chiplet-Based SoC Design

An Arrangement Of Chiplets That Outperforms A Grid Arrangement (ETH Zurich / U. Of Bologna)

Hardware Trojans Target Coherence Systems In Chiplets (Texas A&M / NYU)

Cost Characteristics Of The 2.5D Chiplet-Based SiP System

Securing Heterogeneous Integration At The Chiplet, Interposer, And System-In-Package Levels (FICS-University Of Florida)

Delay-Based PUF For Chiplets To Verify System Integrity

Appendix

White papers

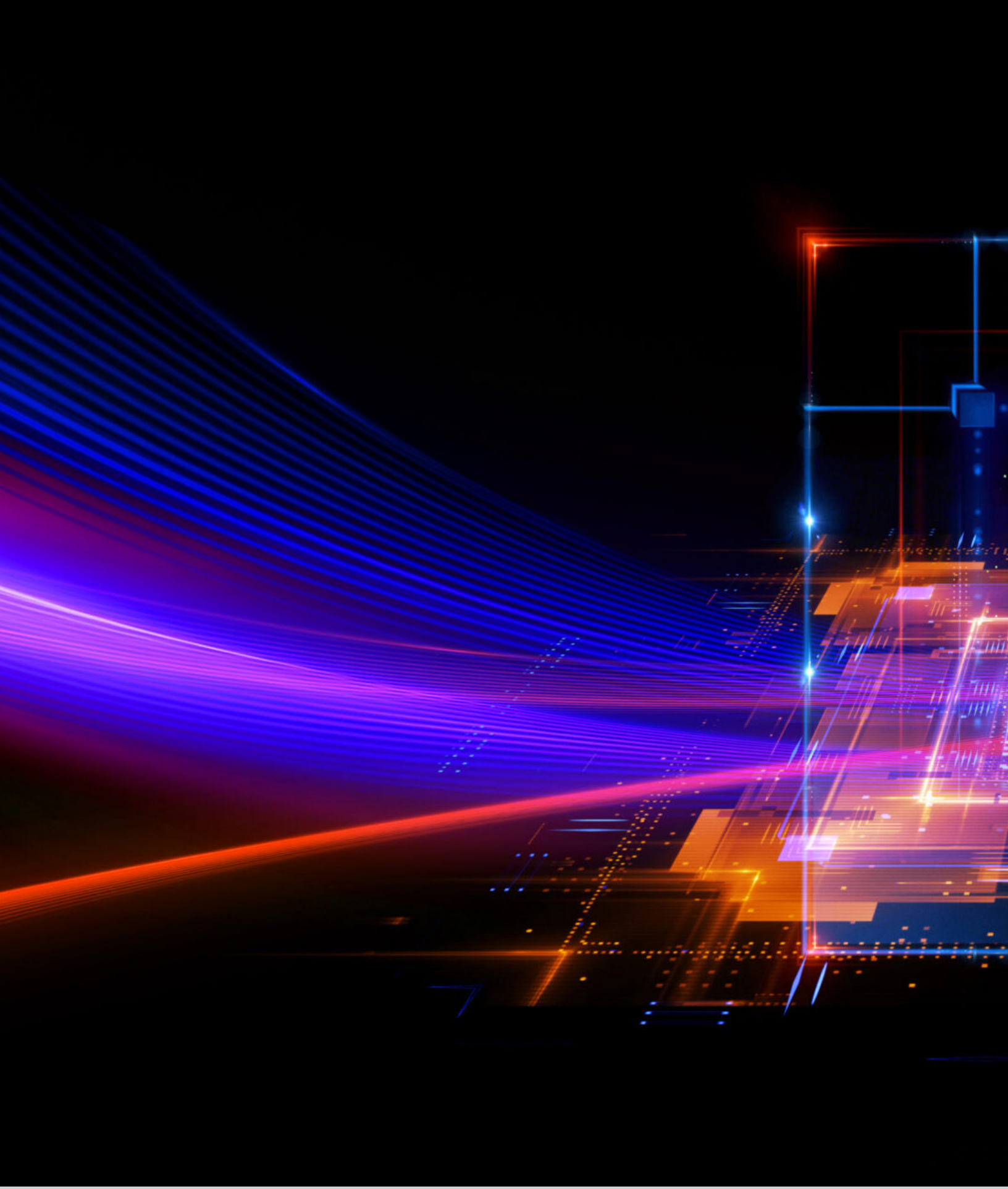
Promex Industries

"Heterogeneous Assembly Datasheet," How heterogeneous assembly works. DECEMBER 16TH, 2022. https://semiengineering.com/heterogeneous-assembly-datasheet/

Intel

Accelerating innovation through a standard chiplet interface: The Advanced Interface Bus (AIB). https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/accelerating-innovation-through-aib-whitepaper.pdf

See more white papers on Semiconductor Engineering's latest chiplets page in the Knowledge Center.



Semiconductor ENGINEERING
DEEP INSIGHTS FOR THE TECH INDUSTRY