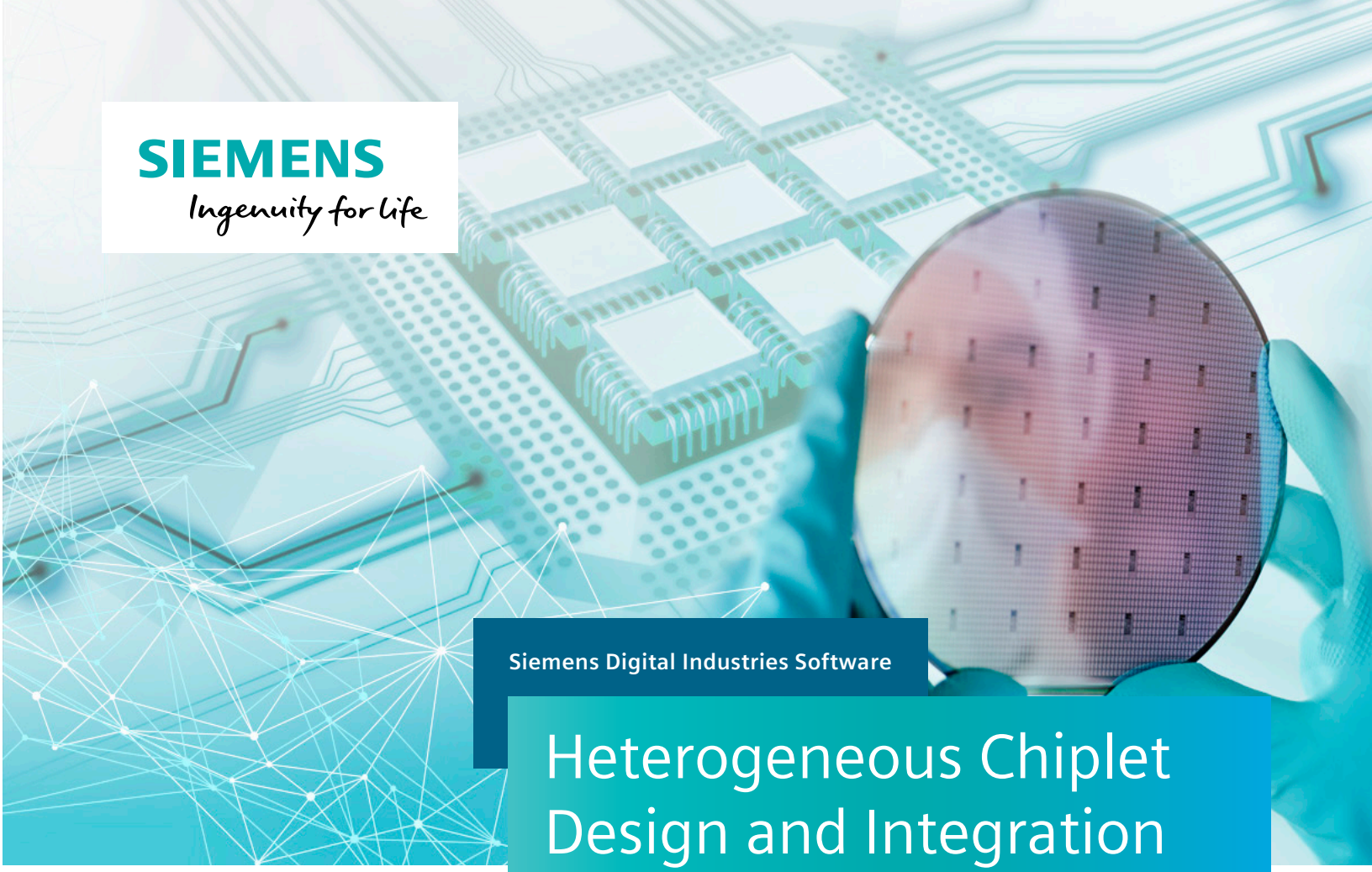


SIEMENS
Ingenuity for life

Siemens Digital Industries Software

Heterogeneous Chiplet Design and Integration

Bringing a New Twist to SiP Design

Executive summary

The semiconductor industry is facing an inflection point as higher cost, lower yield, and reticle size limitations drive the need for viable alternatives to traditional monolithic solutions, which have hit the limits of physics. This is driving an emerging trend to disaggregate what typically would be implemented as an SoC into solid, fabricated IP blocks, or chiplets, that typically include just a couple of functions, resulting in a multi-die heterogeneous integrated implementation.

Keith Felton, Anthony Mastroianni, Kevin Rinebold and Per Viklund

Introduction

A number of factors are converging and driving the chiplet design revolution. To start with the economic advantage of silicon scaling is slowing.

The semiconductor industry is facing an inflection point as higher cost, lower yield, and reticle size limitations drive the need for viable alternatives to traditional monolithic solutions, which have hit the limits of physics. A lot of these inflection point issues have been driven by a reduction in the performance and power benefits one can achieve by scaling below 10 nm, along with the expanding number of physics-related issues at the most advanced nodes, such as multiple types of noise, thermal effects, and electro migration.

This is driving an emerging trend to disaggregate what typically would be implemented as an SoC into solid, fabricated IP blocks, or chiplets, that typically include just a couple of functions, resulting in a multi-die heterogeneous integrated implementation.

For example, HPC, AI, and other extreme data processors reach or exceed reticle sizes when using a monolithic approach. So companies such as Intel, AMD, and Marvell already utilize a chiplet approach for many of their designs. Many of these are homogeneous approaches where what would have been a monolithic implementation is partitioned into multiple blocks interconnected with proprietary interfaces. However, there are efforts underway to standardize chiplet interfaces, opening the way to third-party chiplets.

Adopting heterogeneous and homogeneous integration offers a path to enhanced device functionality, faster time to market, and silicon yield resiliency. Multiple integration technology platforms have emerged that allow for cost, size, performance, and power optimizations that satisfy the need of multiple markets; such as mobile computing, automotive, HPC, AR/VR, AI, IoT, medical, aerospace, and 5G. Multiple integration technology is supported by approaches such as fan-out wafer-level packaging (FOWLP), which rose to popularity with Apple and TSMC, and with TSMC's Integrated Fan Out (InFO) technology. FOWLP was originally developed as a low-cost alternative to 2.5D and 3D-IC, but increased density, pillars, high-bandwidth memory, and faster interconnects have made FOWLP much more attractive to other market segments as well. Thus, 2.5D and 3D-ICs are becoming the platforms of choice at the high end of the performance/bandwidth market.

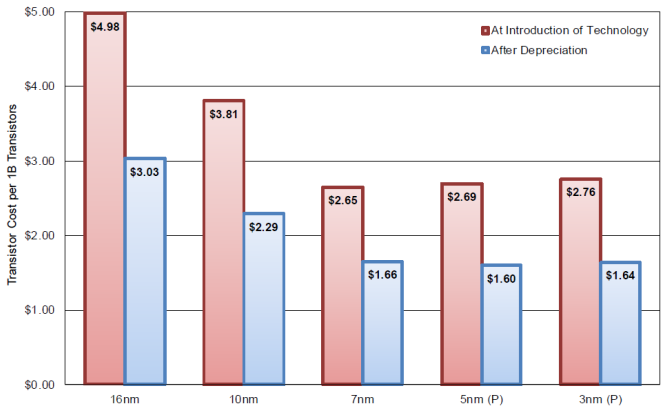


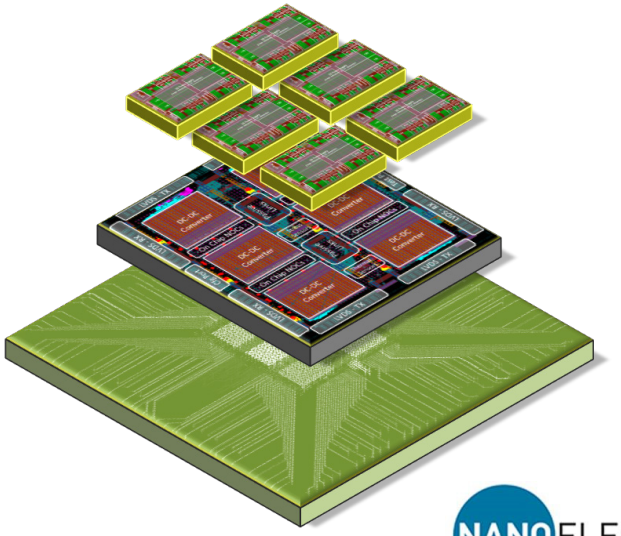
Figure 1: Cost of transistor at different process technologies. (Source: International Business Strategies 2020)

The supply chain is also gearing up support and technology offerings, with all the major foundries supporting advanced packaging. TSMC, UMC, Global Foundries, Samsung, and others now offer sophisticated advanced packaging options. Not to be left behind, the traditional OSATs have also responded to the trend with the top two, ASE and Amkor, offering robust high-volume integration platforms.

Is It a SiP or a chiplet?

"A system in package, or SiP, is a way of bundling two or more ICs inside a single package. This is in contrast to a system on chip, or SoC, where the functions on those chips are integrated onto the same die. SiP has been around since the 1980s in the form of multi-chip modules. Rather than put chips on a printed circuit board, they can be combined into the same package to lower cost or to shorten distances that electrical signals have to travel. Connections historically have been through wire bonds"

Semiengineering



A SiP design typically contains multiple fully functional dies, making them, in fact, similar in design concept to packaged ICs used on a PCB. So at a macro level, SiP can be thought of as a miniaturization platform, not an alternative to SoC implementation.

A chiplet can be thought of as a bare die specifically designed and optimized for operation in conjunction with other chiplets within a package.

Chiplets have significant chiplet-to-chiplet high performance interfaces and require close proximity to each other to reduce energy consumption and ensure data bandwidth performance.

The drive for chiplets comes from the semiconductor industry's move to smaller process nodes and the costs for yielding large dies continues to increase along with the manufacturing limitation of the reticle size. Additional drivers come from the NRE costs involved in designing ASIC's/SoC's for the latest process nodes making them viable for huge volume products only.

Enabling a broader adoption of this technology will require standardized interfaces, protocols, chiplet models, and design and test flows. Standardized interfaces are key if the Lego-like plug and play model is to be realized, where designers can simply select chiplets from an array of vendors and be confident they will function together as envisioned.

To date there has been good progress in the standardization of interfaces and protocols, including USR, XSR, and BoW serial interfaces and OpenHBI, HBM, and BoW Fine parallel interfaces. There are still several companies promoting their proprietary interfaces as standards, such as Intel's Advanced Interface Bus (AIB), which is now public and royalty free, and this list may grow, but overall it seems to be converging on standardization, which is a positive sign for broader chiplet adoption.

Heterogeneous chiplet design, not yesterday's SiP

At the macro level an assembled chiplet design looks similar to a traditional SiP. Side-by-side, the difference is noticeable as the die in a SiP are usually significantly bigger, spaced further apart, and fewer in number.

Chiplets are fully manufactured, solid IP building blocks that perform specific, focused tasks. Combined chiplets can perform or outperform a traditional monolithic SoC, as they can scale beyond the reticle size limit and leverage multiple process nodes that best suit their function.

Today there are two approaches to chiplet based design. There will likely be combinations of both approaches used on actual chiplet designs. The first is processor disaggregation, where a complex CPU, GPU, or perhaps AI processor is decomposed into plug and play modules assembled and interconnected with a silicon interposer or bridges. This approach will likely be offered by a small set of semiconductor vendors who use proprietary interfaces for their processor building blocks along with optional, general purpose interfaces and general purpose chiplet offerings.

The second approach is to use general purpose building block chiplets that are assembled, aggregated, and interconnected with a custom ASIC or ASICs. The general purpose chiplet approach will likely be offered by a broad and diverse set of chiplet providers. It has the potential to provide broad appeal with a Lego-like plug and play methodology, assuming the chiplet suppliers provide the relevant chiplet related design-in IP, which would ideally be provided in a standardized chiplet design kit. It also offers the ability to scale functionality and capability, which at the monolithic level are limited by the semiconductor manufacturer's reticle size limitations.

As an example, we will use a 5G base station with a digital baseband chip and an RF chip. Digital implementation is better (power, performance, etc.) at lower nodes, while RF/analog is better at larger nodes. The chiplet approach enables designers to do several highly valuable things: optimize functions to the process node, reduce power, and get a better overall form factor for the finished device. With this technique, functions of different technologies can be combined that would never or rarely be combined into a single ASIC.

Adapting to design with chiplets

Using a large array of heterogeneous chiplets requires multiple high-speed, high bandwidth, low latency data paths running between them, as well as a large, distributed power delivery network. To achieve this the chiplets must be integrated onto a carrier substrate that is capable of meeting the performance requirements of the interfaces.

One of the first design decisions is the selection of the carrier substrate technology and materials. This is typically influenced or constrained by the technologies, materials, and production capacity made available by the package fabricator and assemblers.

Let's look at technologies first. For example, is a silicon interposer the right choice, should embedded bridges be used in an organic or redistribution layer (RDL) substrate, could an RDL substrate, such as a FOWLP, satisfy the performance and power delivery goals on its own. There are many factors to consider, including cost, volume, manufacturing lead time and capacity, and performance. But interface performance and power delivery are not the only factors to consider when choosing the optimum carrier substrate. Chiplets can also add complex behaviors in terms of heat dissipation and the thermal interactions between chiplets and substrates; these must be understood, planned for, and optimized.

Early planning and predictive analysis

Designing with heterogeneous chiplets makes early planning and predictive analysis of the complete package assembly mandatory, and one of the first things that gets attention is thermal and power delivery considerations and analysis.

For thermal, this can start before any physical planning has started. Using heterogeneous multi-chiplet and multi-substrate packages often introduces new connectivity structures, such as 3D stacking, TSVs, bumps, hybrid bonding, and Cu pillars. These can cause unexpected device performance and reliability problems such as heat dissipation and thermal induced stresses, which can cause functional failures and critical reliability problems. Investigating these connectivity structures and evaluating the available material options will prevent design stage changes and even late-stage failures.

In the case of predictive power delivery analysis, this evaluation can take place early in the planning process, before detailed layout has even started. Once the chiplets (completed or not) have been placed—and assuming their power consumption information is known, however approximate—power delivery analysis can be performed. A typical approach here is to approximate the percentage of metal coverage per layer using a Monte Carlo-type sweep analysis to identify those parts of a circuit whose values have the greatest impact on performance so those values can be communicated to the layout designer to control downstream copper area creation.

Apart from thermal and power delivery, the other big challenge chiplet-based design brings is in how to develop and co-optimize the interposer, package, and chiplets asynchronously, as each domain's design team is likely to be on different schedules and possibly in different time zones. Waiting for the chiplet floorplan and bump matrix to be completed is one approach, but its serial nature delays planning and reduces the opportunity for co-optimization of the chiplets external interfaces and their assignment to the carrier substrate.

Hierarchical chiplet co-optimization

A far superior approach to the serial approach is to enable concurrent design and refinement using hierarchical planning of the chiplets. Such an approach must include a robust ECO mechanism to manage the asynchronous changes between the teams so that any changes do not get missed or overlooked. A hierarchical approach takes into consideration that, even though a chiplet is just a subset of functions typically found in an SoC, it is still floorplanned in functional blocks, with each block potentially designed by a different designer or team on a different schedule.

The package team needs the chiplet bump array and associated signals so they can plan out the package floorplan and connectivity. Waiting until the chiplet bump array is fully defined will delay package planning and limit the ability for co-optimization, as the chiplet design team is likely to resist any significant changes.

With a heterogeneous approach the package planning process can start even if the chiplet design is not yet started. The chiplet's bump array and signal assignments can be created at the package/interposer level and passed back to the IC design team, who can then iterate with the package/substrate team as the design progresses.

Chiplet interface management and design

A standardized interface is a key enabling characteristic of a chiplet. This is how the chiplet, in a predefined manner, communicates with a core design or other chiplets. Therefore, broad adoption of chiplets requires standardized interfaces and protocols, such as those discussed earlier: USR, XSR, and BoW serial interfaces and OpenHBI, HBM, and BoW Fine parallel interfaces. All these interfaces bring a new challenge for designers: how to rapidly describe the interfaces for new chiplets while interconnecting COTS or existing chiplets.

Current design approaches, such as graphical schematics or writing thousands of lines of HDL, make it challenging to capture, visualize, manage, and implement chiplet designs. A designer could look up the interface definition for each chiplet interface and manually create the required connectivity in accordance with the spec, then define electrical constraints to ensure correct package design. But this is a lot of manual work and introduces the risk of human generated mistakes that might not be easy to catch early in the design process. To avoid this risk, we will introduce a novel concept: *interface based design*.

Automating interface based design

Interface based design (IBD) is an exciting new approach to capturing, designing, and managing large numbers of complex interfaces that interconnect multiple chiplets. Since the chiplet has a known formal interface, the interface description can become part of the chiplet part model. When a designer places an instance of this chiplet, everything related to the interface is automatically put in place. This way we take the human out of the equation, ensuring that correct-by-design chiplet connectivity is established.

With an interface defined as an IBD object, the designer can focus on a higher level of connectivity abstraction. This facilitates more insightful chiplet floorplanning and chiplet-to-package or chiplet-to-interposer signal assignments, and it allows designers to explore, define, and visualize route planning without having to transition the design into a substrate place-and-route tool. IBD allows the designer to see both "the forest" and "the trees" during the design process by expanding or contracting the interface; thus providing visualization and manipulation at the appropriate level of interface expansion.

Thermal, stress, and reliability management

The close proximity of devices within heterogeneous packages necessitates understanding the effect they have on one another, also referred to as chip-package interactions (CPI). These could be electrical, thermal, or stressed related, and are not mutually exclusive.

Using a combination of chip-level and package/system thermal modeling, designers can generate power-aware thermal and stress device-level models that provide greater accuracy for thermal and mechanical simulations. The models can then be used to perform warpage, stress, and fatigue analysis.

When it comes to material choices, substrate stackup and device/chiplet proximity have considerable impact on thermal and stress performance. So it is very important not to wait until design of the package assemblies is completed. Instead, start with predictive analysis before or during the prototyping/planning phase. Starting analysis as far left in the process as possible allows for the most flexibility in making choices and tradeoffs and usually results in the minimum impact on the design.

Test and testability

The production test methodology used in digital, homogeneous designs has been established for many years: deploying structural design-for-test (DFT) logic implemented during the ASIC design process. DFT test tools are run on the inserted test logic to generate the automated-test-equipment (ATE) production test programs used for wafer and package level production testing. Additionally, boundary scan description language (BSDL) test patterns are generated for the design to be used for PCB level tests.

Heterogeneous chiplet design requires extensive changes and additions to the traditional, homogenous design. Since these designs include two or more ASIC/chiplet components, a production test program must be provided for each of the internal components. It is assumed that externally sourced chiplets will be wafer-sorted and delivered as a known-good-die but still need to be retested once they are assembled in the SiP device. Furthermore, these tests need to be run from the external package pins, most of which are not connected directly to the chiplet pins. In addition to the individual die testing, the interfaces between each component need to be functionally tested, preferably at speed for each of the die-to-die interfaces.

IEEE test standards are being developed to accommodate these 2.5D test methods. Different tool vendors may deploy different approaches in implementing these standards, which may cause test compatibility issues of components that use different DFT vendor tools. For board level testing, a composite BSDL file for each of the internal components is preferred, but not necessarily supported, by all DFT tool vendors, which further complicates the PCB level testing.

With the introduction of 3D heterogeneous designs, additional challenges are introduced as the die stacked above the base die may not be accessible through traditional BSDL/JTAG interfaces. There are additional emerging IEEE test standards being developed to accommodate 3D test methods as well. These methods deploy hierarchical test methods that can only test the stacked die after assembly. Just as with 2.5D, DFT vendor capability issues will likely arise in 3D stacked dies built with components using different vendor tools.

The inclusion of multi-die in a package may also dramatically increase the production test time and cost. New high-speed scan methods are being developed that will enable the use of very high-speed test interfaces, such as XSR, to replace the traditional slow-speed JTAG approach, which should significantly reduce the SiP level connectivity as well as reduced test time.

Since the test connectivity in 2.5D devices is implemented through an interposer, the package design is required for the planning and routing of the die-to-die test connectivity, which will require new package design and analysis flows.

Driving verification and signoff

It is critical for all verification to start in the planning process and continue throughout the layout process. It starts during initial planning where early assembly validation of device and bump placement can be performed along with IO pad ring validation, ERC checks, and ESD cell insertion determination. Such in-design validation provides early identification and resolution of manufacturing issues without running the full sign-off flow—which can be resource and time intensive and usually requires a different department's involvement.

When it comes to final design verification, more than just mask metal layer fabrication checking against the fabricators rules is involved. It is also very important to analyze various layout enhancements that will improve yield and reliability, such as analysis of thickness variations and planarity issues of the RDL. It is extremely important to release to manufacturing with confidence that all the devices and substrates work together as expected in order to avoid costly late-stage errors and delays.

Recommendations

There are five areas that deliver the most impactful effects on successfully implementing and designing with chiplets.

1. Chiplet design kits (CDK) provide a model of the chiplet for implementation and integration. A CDK can include interface protocols, IO models, ATE test methods, power characteristics, and thermal models such as BCI-ROM.
2. Heterogeneous planning and co-optimization should use a complete 3D digital model (aka digital twin) of the entire device assembly that drives all downstream aspects of design, analysis, and verification maintains a continuous digital thread.
3. Physical verification at every level of 3D assembly, from the substrate layer through design rule checks to assembly-level layout-versus-schematic.
4. Multi-domain testing starting with the individual die and continuing with die-to-die and across the entire package assembly.
5. Ecosystem interoperability including the ability to seamlessly share designs and data with suppliers, partners, foundries, and OSATs.

Siemens Digital Industries Software

Headquarters
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 972 987 3000

Americas
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 314 264 8499

Europe
Stephenson House
Sir William Siemens Square
Frimley, Camberley
Surrey, GU16 8QD
+44 (0) 1276 413200

Asia-Pacific
Unit 901-902, 9/F
Tower B, Manulife Financial Centre
223-231 Wai Yip Street, Kwun Tong
Kowloon, Hong Kong
+852 2230 3333

About Siemens Digital Industries Software

Siemens Digital Industries Software is driving transformation to enable a digital enterprise where engineering, manufacturing and electronics design meet tomorrow. Our solutions help companies of all sizes create and leverage digital twins that provide organizations with new insights, opportunities and levels of automation to drive innovation. For more information on Siemens Digital Industries Software products and services, visit siemens.com/software or follow us on LinkedIn, Twitter, Facebook and Instagram. Siemens Digital Industries Software – Where today meets tomorrow.

siemens.com/eda

83718-C1 05/21 D