nature reviews electrical engineering

https://doi.org/10.1038/s44287-024-00078-x

Review article

Check for updates

System technology co-optimization for advanced integration

Saptadeep Pal[1], Arindam Mallik[2] & Puneet Gupta[3]

Abstract

Advanced integration and packaging will drive the scaling of computing systems in the next decade. Diversity in performance, cost and scale of the emerging systems implies that system technology co-optimization (STCO) would be essential to develop these integration technologies for future systems. Such STCO would need to comprehend not only integration technology, circuits, architectures and software but also their interactions with the power delivery, cooling and system costs. In this Review, we present a perspective on what would be needed from these STCO approaches with exemplar case studies covering the current state of the art and the future outlook.

Sections

[1]Etched.ai, Cupertino, CA, USA. [2]IMEC, Leuven, Belgium. [3]Department of Electrical and Computer Engineering, University of California, Los Angeles, CA, USA. e-mail: puneetg@ucla.edu

Key points

- Connectivity, scale, cost and form factor are the main drivers for use of advanced integration techniques in emerging computing systems.
- Support for technology heterogeneity, advanced power delivery and cooling within the advanced packaging are key system enablers.
- System technology co-optimization (STCO) approaches can be classified as link level, component level or cross-stack system level. Automated, fast cross-stack STCO frameworks are still in their infancy but are essential to guide high-value technology development.

Introduction

Traditionally, dimensional scaling has been the primary driver for dramatic improvements in the power, performance, form factor and cost of electronic integrated systems. Aggressive scaling of complementary metal-oxide semiconductor (CMOS) silicon and wiring minimum features by more than 1,000 times for more than four decades, enabled by advancements in patterning technologies, coupled with performance boosters such as the adoption of copper wiring, strained silicon and FinFETs have delivered on the promise of Moore's law. Unfortunately, this scaling has come at exponentially increasing cost[1–3]. Further scaling is becoming increasingly untenable as we approach physical limits. This is forcing the semiconductor industry to take a careful look at the 'system on chip' (SoC) trend, enabled by technology scaling, of the past few decades.

A chip is rarely the whole system. It is packaged and bonded to a printed circuit board (PCB), with a 'fan-out' at each level (that is, the size of interconnect at each level from chip to package to board increases). Although the dimensions within the chip have been scaled by more than three orders of magnitude in the last five decades, the dimensions of package/PCB input/output (I/O) (ball grid array) bumps have scaled barely five times[4]. As a result, multi-package systems on a PCB suffer in all aspects from power and performance to area, and cost, which drove the industry towards SoCs. With chip scaling becoming more difficult, there is a new focus on advanced packaging to scale inter-chip connectivity. This approach has the potential to reduce the cost of large systems, improving communication overheads and enabling new types of systems with intimately connected heterogeneous components. Therefore, advanced integration is expected to be a system-scaling driver in the coming decade.

The semiconductor industry has long relied on separating design and manufacturing. Several abstraction aids, such as design rules and compact device models, have been developed to preserve the clean abstraction of technology available to circuit designers. This has made design and technology development largely independent of each other. Unfortunately, difficulty in scaling has blurred these boundaries and made the co-optimization of design approaches and technology development essential. This has resulted in a strong interest in design technology co-optimization (DTCO), especially in the development of device technology[5–9] and lithographic patterning[10–12]. The eventual choice of patterning scheme at any technology node has as much been dictated by design considerations such as ease of design, availability of design automation tools and block-level power/performance/area metrics as by the complexity of the technology itself. Over time, DTCO approaches have become increasingly sophisticated, ranging from the earlier manual design of small benchmarks[12] to elaborate stitched electronic design automation (EDA) tool flows[12,13] to principled and fast frameworks[10,14].

Integration and packaging are going through a renaissance and are poised to see significant innovation in the coming years. Limiting DTCO to a single die is no longer sufficient and, therefore, evaluating an entire system that can consist of several chips integrated together using packaging technology would be essential[15,16]. This system technology co-optimization (STCO) is needed to guide innovation in the right direction. STCO approaches are still in their infancy owing to the lack of automated frameworks. Eventually, STCO would need to account for multiple facets, such as within-chip technologies (device, patterning, interconnect), heterogeneous system component technologies (for example, memory types), ways of connecting chips (2.5D or 3D integration), power delivery and cooling infrastructure, and architecture and software applications running on hardware. The future of system scaling is highly dependent on cross-layer optimization of different abstraction layers of computing systems (Fig. 1). Traditionally, the semiconductor industry has scaled logic, memory and interconnects separately and largely independently of the systems being constructed using these components. The future trend would be to optimize system functions or modules using the process technology best suited to it. In practice, this means building each module on its own chiplet manufactured with the appropriate process technology. An advanced packaging scheme, such as advanced 3D stacking, would then bind those together such that all of the functions act as if they were on the same piece of silicon (Fig. 2).

In this Review we discuss STCO in the context of packaging for computing systems. We provide recent industry examples of different choices of packaging approaches driven by the system context and highlight the system drivers and the enablers for advanced integration as well as how points on the multidimensional Pareto frontier of these drivers/enablers could dictate viable integration technologies. Next, we discuss some of the emerging approaches for DTCO/STCO for advanced packaging and integration and, finally, conclude with ideas for future directions of work.

Design-dependent choice of advanced packaging

With the proliferation of applications demanding high performance such as artificial intelligence (AI), the requirement for larger silicon systems has grown exponentially. Over the past decade, advanced packaging technologies have improved the scale, performance and energy efficiency of silicon systems. It is now possible to build large chips by dense integration of multiple silicon dies inside a package. These technologies have different trade-offs between integration density, scale and cost. For example, organic interposers are cheaper but allow a lower interconnect density between adjacent dies compared with silicon interposers. Therefore, depending on the target application and market, the right advanced packaging technology needs to be chosen, and the architecture must be co-optimized.

Recent developments have showcased a trend towards the design-dependent co-optimization of system architecture and advanced packaging across various products.

The field-programmable gate array (FPGA) industry is one of the earliest adopters of silicon interposers[17]. In the late 2000s, because of their easier reconfigurability and quick turnaround time, FPGAs gained popularity, and larger systems based on FPGAs began to develop. FPGAs have 20–40 times lower compute density. Therefore, FPGA silicon started to be as large as a full reticle[17] and systems were regularly built

using multiple FPGAs on a board. Reticle-sized silicon is yield-limited and, therefore, costly. In addition, multi-FPGA solutions often exhibit poor performance. To alleviate these issues, Xilinx used silicon interposers to build large FPGAs. Silicon interposers allow the integration of multiple known-good dies at a high interconnect density, allowing for lower-cost FPGA products. Moreover, it allows FPGAs to be built with integrated high-bandwidth memory (HBM), thus making them viable alternatives to building application-specific integrated circuits. For instance, Microsoft adopted FPGAs as the de facto platform to build custom accelerators[18].

Similarly, manufacturing yield concerns for building large core-count monolithic central processing units (CPUs) pushed AMD to adopt a chiplet-based architecture. Disintegrating a large monolithic processor into smaller chiplets allowed AMD to build processors with known-good dies and save on cost, often as much as 2.1 times[19]. Moreover, AMD leveraged the cost benefits of heterogeneous integration by integrating external I/O circuitry into an I/O chiplet on a lower-cost 12 nm node, as opposed to the core chiplets fabricated on an expensive 7 nm node. Cost constraints forced AMD to use organic substrates for chiplet integration rather than the expensive silicon interposers used by FPGAs and graphics processing units (GPUs). This was enabled by the co-design of the architecture with the packaging substrate characteristics, and the fact that the inter-chiplet bandwidth required was only a few hundred gigabits per second in the more general-purpose computing architecture as opposed to the HBM connections needed in the NVIDIA example below. Additionally, a chiplet-based methodology provides flexibility for building multiple product lines by altering the number of chiplets. AMD and Xilinx leveraged this flexibility to save non-recurring engineering costs and improve the time to market for different product lines. For example, the AMD 9654P high-performance computing (HPC) product has 12 compute core chiplets and 1 I/O chiplet, whereas the mainstream enterprise product 9224 has 4 compute core chiplets and 1 I/O chiplet. Sharing the chiplet designs across different product lines saves both non-recurring engineering design and manufacturing costs.

The demand for HPC and AI applications is driving the adoption of very high-bandwidth in-package integration technologies such as silicon interposers and silicon bridges. These applications are highly parallel and primarily run on accelerators such as general-purpose GPUs and Google tensor processing units. These accelerators are highly parallel (for example, 14,592 FP32 cores in NVIDIA H100) with a large amount of computing throughput, often more than one PFLOP of compute per die. Large computing throughput requires higher memory bandwidth[20]. Consequently, accelerator architectures rely on on-package dynamic random access memory (DRAM) to provide the required bandwidth (for example, 3 TB s^{-1} on an NVIDIA H100 GPU[21]). Multiple HBM devices are integrated with the accelerator compute die within the package[22]. HBMs use wide memory interfaces (for example, 16× double data rate (DDR) channels per device), and each pin supports a data rate of <10 Gb s^{-1} to maintain low I/O energy and area overhead. Integration technologies using silicon for inter-die links can accommodate a ten times higher density of signal pins and traces. Consequently, accelerators such as general-purpose GPUs and tensor processing units use technologies such as TSMC's chip-on-wafer-on-substrate technology (CoWoS-S[23], CoWoS-L[24]) and Intel's embedded multi-die interconnect bridge (EMIB)[25,26] instead of organic substrates for inter-chiplet connectivity. Beyond 2.5D integration using chiplets, 3D integration of two active dies on top of each other is gaining steam. Certain HPC, gaming and multimedia workloads benefit from larger caches[27]. However, static random access memory (SRAM) cost and area scaling has been underperforming compared with logic scaling over the past few technological nodes[28–30]. AMD introduced 3D integration of a cache die on top of a CPU die in their V-cache technology. This is a clever and elegant co-design of architecture and packaging. 3D integration using hybrid bonding can provide 25 times I/O density[27,31] and a shorter interconnect distance and lower energy than 2.5D integration. Therefore, it can provide the on chip-like bandwidth needed by the cache subsystem with minimal energy overhead. In one incarnation, the bottom CPU die is built in an expensive 5 nm node, whereas the cache die is built in a relatively cheaper 7 nm node optimized for SRAM, thus improving the overall cost of the system.

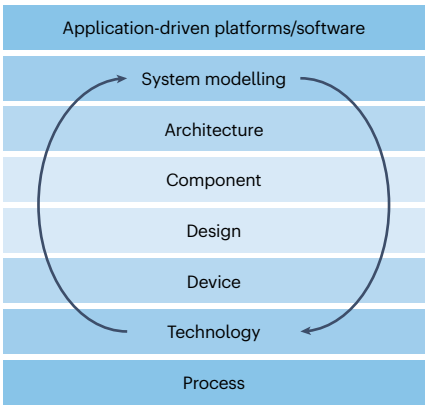


Fig. 1 | Cross-stack system technology co-optimization. Cross-layer optimization paves the way for system scaling[142], with recent computing systems from AMD, Xilinx and Nvidia being examples of such cross-layer optimizations in that they leverage different advanced integration schemes depending on the system needs.

These case studies show how careful co-design of the chiplet-based system architecture and integration scheme can lead to optimized product solutions. Recent commercial products such as the NVIDIA GH100 (refs. 23,32), second-generation AMD EPYC[19,33] and third-generation AMD EPYC with V-cache[27,34] show different characteristics (Fig. 3) depending on the respective integration schemes. We argue that STCO is critical to the success of next-generation products when the cost benefits of moving to newer technology nodes are dwindling. In addition, with the recent surge in demand for AI[35,36] and other HPC workloads[37], custom application-specific integrated circuits are becoming mainstream. STCO frameworks are required to guide the choice of both architecture and technology selection to extract the most value from these systems.

Key drivers for advanced integration and their design interactions

System drivers

What are the primary drivers behind the surging demand for advanced packaging technologies? The need for high-performance and energy-efficient connectivity between components inside a package is growing in scale. Additionally, the need for cost optimization and form-factor minimization is driving the development of advanced packaging.

Connectivity. Connectivity is the primary driver behind the development of advanced packaging solutions. Poor scaling of off-package links becomes a barrier to system performance and power scaling

when integrating multiple packaged chips on a PCB. Integrating chiplets inside a package is driven by the increased inter-die connectivity that can be achieved inside a package. Today's HPC and AI workloads demand multiple terabits per second of bandwidth between different compute and memory chiplets. Therefore, the development of advanced packaging technologies is geared towards enabling high inter-chiplet bandwidths at low energy overheads. This is accomplished by reducing I/O pitch (<20 µm versus >200 µm for off-package I/Os)[4,38], interconnect wiring pitch (<5 µm versus >50 µm) and length (<1 mm versus >10 cm)[38], which enables efficient highly parallel interfaces. Furthermore, this reduces the need for power-hungry high-speed serializer–deserializer circuitry that is needed to drive high data rates over individual interconnects in I/O-constrained designs. Today, greater than ten times bandwidth at equivalent interconnect power can be achieved between chiplets integrated on a package compared with chips interconnected over a PCB. For example, up to 6 TB s^{-1} of memory bandwidth can be achieved using six HBM3 (ref. 39) modules at approximately 160 W of inter-chiplet interconnect power. This is a bandwidth an order of magnitude higher than that achieved using off-package memories over the DDR interface[40,41] at iso-power. Similarly, 3D integration enables another step function improvement in I/O density (>15 times) and energy efficiency (>3 times)[42].

Scale. Improved connectivity facilitates system scaling within a package. As new workloads and data processing techniques demand increasingly parallel hardware, this scaling becomes essential. Compute requirements for machine learning workloads alone have far outpaced gains from Moore's law (Fig. 4a). As evident from several recent trends[24,43], silicon area per chip is growing fast to meet this seemingly insatiable demand (Fig. 4b). This is driving enormous research and development efforts for future advanced packaging technologies. As discussed before, newer advanced packaging technologies such as CoWoS-L are being developed to integrate up to 5,000 mm^2, that is, six reticles worth of silicon[24], in a single package. At the extreme, wafer-scale integration technologies are being developed commercially[44,45] and in academia[43,46] to build systems that are large as an entire 300 mm wafer. For some classes of applications, these technologies would enable systems that can provide an order of magnitude performance gain over systems built using conventional packages[46,47].

Cost. Although advanced packaging provides us with newer platforms for more connected and scaled systems, the primary driver behind the acceptance of a new technology is cost or, often, cost per performance. Given the manufacturing complexities of advanced packaging, can it offer economic advantages for the next generation of electronic systems? The traditional path for improving the cost of digital systems through silicon CMOS scaling is becoming increasingly difficult[1–3]. Chiplets are best thought of as an alternative design methodology to monolithic chips in a world where Moore's law has largely stopped being an economic benefit. A chiplet approach can help improve yield and reduce costs by allowing manufacturers to use smaller, more specialized chiplets rather than a single, monolithic chip for certain tasks[48,49]. AMD has demonstrated the economics of the chiplet approach to building its Ryzen client processors. A 16-core Ryzen chip, such as the Ryzen 9 5950X, built on a monolithic 7 nm die, would have cost AMD 2.1 times more in comparison with its chiplet-based approach of using two 8-core 80 mm^2 core complex dies paired with a cheaper 12 nm I/O die[19]. By modularizing the system based on chiplets, it can be customized for each market segment by simply adding or removing more chiplets. This approach saves cost and simultaneously enables faster design and time to market. The overall benefit can be seen in the total cost of ownership of the hardware[50]. Hence, chiplets are driving innovation within the semiconductor industry based on a flexible and cost-effective economic model.

Form factor. Consumer electronics devices such as laptops, mobile phones and smartwatches have been driving several packaging and integration technologies over the past couple of decades to maximize miniaturization and energy efficiency. Packaging technologies such as integrated fan-out wafer level packaging (InFO)[51], package on package[52], wire-bonded chip scale packages and flip-chip system in package (SiP) allow systems to be built with minimal area and volumetric footprint.

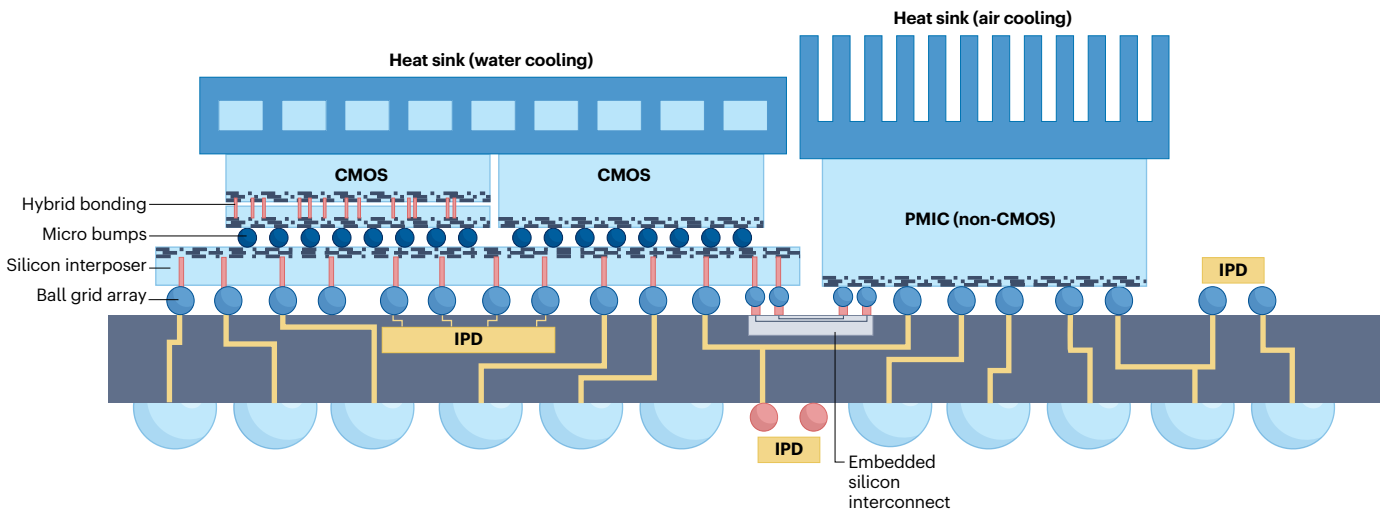
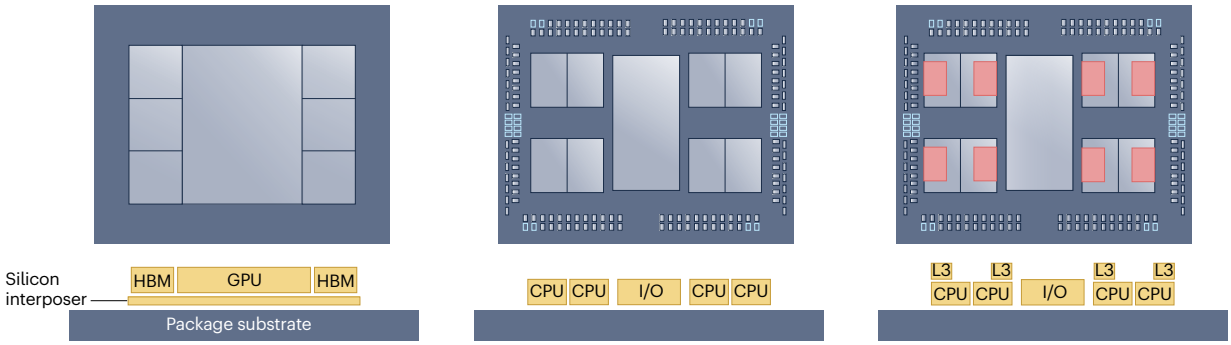


Fig. 2 | Cross-sectional view of a multi-chiplet packaged system. A diversity of chiplet integration technologies alongside power delivery and thermal management components are tightly integrated to realize the full potential of such a system. CMOS, complementary metal-oxide semiconductor; IPD, integrated passive device; PMIC, power management integrated circuit.



	NVIDIA GPU	AMD CPU	AMD CPU with 3D V-cache
Interconnect substrate	Silicon interposer	Organic substrate	Organic substrate
Bump pitch	45 µm (die-to-interposer)	130–150 µm (die-to-substrate)	130–150 µm (die-to-substrate) 9 µm (3D die-to-die)
Inter-chiplet I/O type and width	Parallel (1024 data + ~80 control	SerDes (72 data + 8 control)	SerDes (72 data + 8 control) Parallel (10,000+ for 3D die-to-die)
Total inter-die data bandwidth	512 GB s^{-1} (GPU-HBM)	~1 TB s^{-1} (core complex die–I/O die	~1 TB s^{-1} (core complex die–I/O die) 2.5 TB s^{-1} (3D L3–core complex die)
Cost	High	Low	Medium

Fig. 3 | Differing integration schemes for different system needs. Various integration schemes provide different interconnect characteristics and integration density. NVIDIA GH100 (refs. 23,32), second-generation AMD EPYC[19,33] and third-generation AMD EPYC with V-cache[27,34] and their characteristics are examples of specific integration schemes. CPU, central processing unit; GPU, graphics processing unit; HBM, high-bandwidth memory; I/O, input/output.

For example, smartwatches and mobile phones integrate power management IC and memory chips with the SoC using package-on-package and SiP techniques. Similarly, Apple's new M-series processors integrate low-power DDR (LPDDR) memory packages with a processor SoC die on the same package substrate. These technologies improve the form factor of these devices by as much as 50%[53,54]. These examples show that advanced packaging has a key role in enabling different use cases which would not have been possible with traditional single-chip packaging technologies.

System enablers

In a future system, platform heterogeneity of technology nodes, better connectivity and co-integration of specialized components could help provide a step function improvement in performance, cost and form factor of systems (Fig. 2).

Technology heterogeneity

Chipletization opens a major avenue for improved functional integration: intimate connection of disparate process technologies. In the past, the trend in the semiconductor industry has been towards a 'siliconification' of all functions due to cost, form factor and short-hop connectivity to the silicon CMOS compute fabric (that is, the SoC trend). Advanced integration (both 2.5D and 3D) allows system designers to buck this trend with possible gains in power and performance. Some examples of such technological heterogeneity include the following:

- Intimately connected memories. HBMs that use a DRAM process are now connected at very short distances (<5 mm) to the compute substrate with very high bandwidth[55–57]. This has improved performance, especially for memory-bottlenecked machine learning workloads. One can envision similar tight integration with other types of memory and storage technologies such as Flash.
- Intimately connected off-package interconnect. High-bandwidth, low-energy, low-latency photonic interconnect[58] has been another representative example of leveraging chiplet heterogeneity, which would otherwise have required much worse pluggable optics or electrical links.
- Intimately connected power delivery infrastructure. Efficient integrated voltage regulators (for example, using gallium nitride (GaN) technology transistors[59,60]) and within-package or within-interposer passives (capacitors and inductors) can dramatically improve power delivery efficiencies for large high-power systems[61].

Although multi-chip modules[62,63] and SiPs[64,65] of the past also allowed heterogeneous integration, the proximity of the different chiplets was more than one or two orders of magnitude worse (approximately 1 cm versus approximately 100 µm).

Power delivery. Advanced packaging enables systems with higher power density in a package. As a result, power integrity challenges in these systems need to be addressed by holistically looking at the integration technology. Novel techniques (architecture, design) and technologies (materials, in-substrate capacitors) are being developed, and more is needed to provide power reliably. Recently, TSMC has started embedding deep-trench capacitors in the silicon interposer. Similarly, newer versions of CoWoS (CoWoS-R[66] and CoWoS-L[24]) are being developed with integrated passive devices (IPDs) for better power integrity[67]. GraphCore[68] used 3D integration based on wafer-to-wafer bonding to integrate a deep-trench capacitor die alongside the compute die, resulting in approximately 40% higher performance. To build a SiP solution with CPU, GPU, accelerator and memory dies on an interposer, the platform voltage regulator needs to be integrated on the interposer close to the logic dies. This could be enabled using

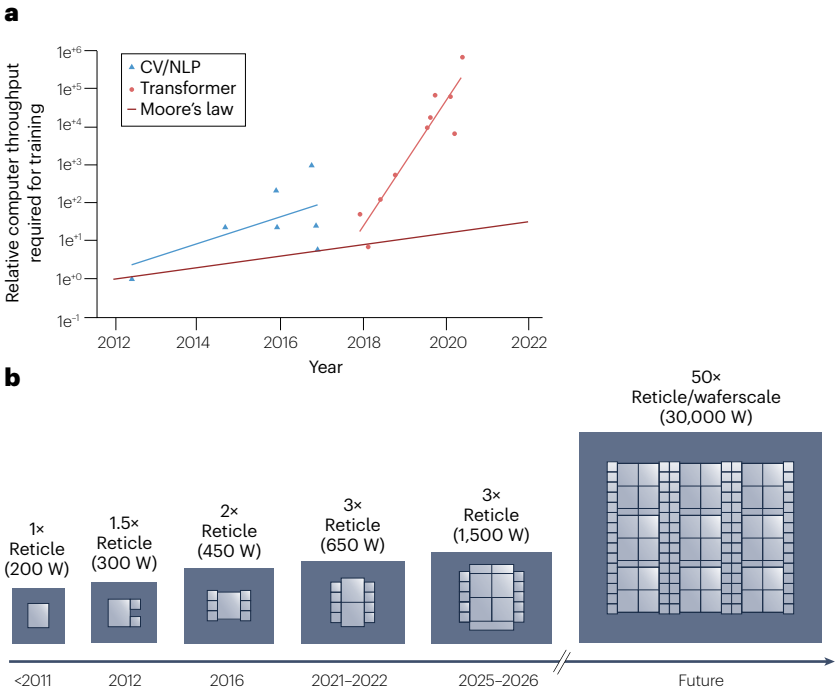


Fig. 4 | Hardware scaling. a, Comparison of hardware scaling versus computing demand scaling. Computing demand for artificial intelligence (AI) workloads is orders of magnitude higher than what Moore's law can provide[36]. Representative workloads consisting of computer vision (CV), natural language processing (NLP) and transformer neural network architecture-based large language models are plotted. **b**, Physical size scaling of compute hardware. Driven by the extreme growth of computing demand in the high-performance computing (HPC) and AI workloads, advanced packaging technologies are evolving to integrate large amounts of silicon in a single package. This alone is insufficient; co-optimization is necessary to extract maximum performance from the silicon area.

high-voltage complementary GaN (CGaN) devices with inductors embedded in the package using high-frequency, high-permeability materials[59,61].

Stable power supply to the microprocessor is important to ensure optimal performance. As technology nodes shrink, power density and voltage drop increase, challenging designers to maintain the 10% margin that is allowed for the power loss between the voltage regulator and the transistors. The development of a high-efficiency, dense integrated voltage regulator will be critical to meet the requirements of future high-performance microprocessors[61]. Alternatively, a backside power delivery network (BSPDN) decouples the power delivery network from the signal network by moving the entire power distribution network to the backside of the silicon wafer (Fig. 5). This approach promises to benefit the voltage drop, improve the power delivery performance, reduce routing congestion in the back end of line and allow standard cell height scaling[69–71]. BSPDN looks promising for the performance improvement of 3D SoCs[72]. For both 2D and 3D designs, the concept of exploiting the free backside of the wafer can potentially be expanded by adding specific devices to the backside, such as I/Os or electrostatic discharge devices[73].

Thermal management. The rise of hyperscaled data centres and AI computing has already increased the rack power density from 10–20 kW per rack to more than 30 kW per rack. In the near future, this number is expected to double. Increased power density exacerbates the thermal problem in a system. This necessitates advanced cooling technologies such as liquid cooling and phase-change cooling, and even techniques such as immersion cooling[74,75].

With heterogeneous packaging, there is a power density disparity across the total area of the package. This corresponds to a higher temperature gradient across the whole package, which can be addressed by novel heat spreader methodologies[76]. At the same time, the challenge of dissimilar heights of individual chiplets, for example, a logic die chiplet versus an HBM module, needs varying cavity depth to use an integrated heat spreader[77]. On the positive side, chipletization benefits thermal performance because heat-generating components are spread apart, thus reducing their thermal cross-talk[78]. Additionally, it helps the reliability of thermally sensitive components in the package, as well as overall system-level reliability[79–81].

The novel utilization of features specific to 2.5D or 3D integration such as through-silicon vias (TSVs) for heat dissipation and management is an interesting aspect. Thermal-aware floor planning can manage heat loads by optimizing the distribution of circuit components and TSVs, effectively reducing junction temperatures across the die[82–84]. Multiple pieces of research have been carried out to co-optimize thermal and electrical design challenges[85]. TSVs have been used as a heat-removal mechanism[86]. In addition, co-design approaches that couple TSVs with microfluidic cooling[87], silicon micropin fins[88] or air gaps[89] have been reported recently.

Overall, thermal management challenges in advanced packaging are closely related to electrical performance and manufacturing. These coupled phenomena often present critical trade-offs and constraints that must be correctly recognized and accounted for, through STCO.

STCO methodologies and frameworks

Advanced packaging innovation could enrich SiP technology, helping the semiconductor industry continue to benefit from Moore's law but at a system scale. Moore's law has enabled the production of less expensive semiconductors, that dissipate less power and have higher performance. This has led to a large demand for semiconductor systems with a wide range of integrated functionalities on a single die. As Moore's law scaling is slowing down and with Dennard's law being out of consideration, building high-performance, low-power and cost-effective silicon systems is no longer just about realizing a design in

one semiconductor manufacturing process. The monolithic SoC way of designing electronic systems is losing its viability as a cost-efficient, functional option for system integration. SiP, however, opens the door to the design of a nearly limitless variety of complex systems. SiP provides opportunities as well as new challenges across the entire stack that encompasses technology development, design, manufacturing, testing and system software.

Recent examples of such co-optimization have been emerging both in industry and in academia. Cerebras[44] addresses the problem of accelerating large AI workloads to run across multiple chips in a compute system. Instead of dicing a wafer into multiple dies to make traditional chips, they carve out a larger square within the round 300 mm wafer. That is a total of 84 dies, with 850,000 cores, all on a single piece of silicon. The Cerebras architecture enables running large machine learning models on a single chip without portioning, enabling scaling to become easy and natural. This required the researchers to rethink system architecture. New packaging technology, power delivery techniques and cooling systems were co-developed alongside the wafer-scale architecture to realize a massively parallel system for AI and HPC workloads.

GraphCore was faced with a problem of dynamic voltage droop in the package causing performance loss. They used a wafer-on-wafer hybrid bonding technology to 3D stack the accelerator die on top of a power delivery die[90]. This allowed them to improve AI workload performance by 40%. CMOS process technology scaling alone has stopped providing such leaps in performance, whereas the use of clever design, integration and manufacturing techniques can help realize the true performance potential of a system.

Another work[49] attempted to understand what the minimum size of a chiplet should be to minimize the overall cost. The authors showed that the cost of high-performance 2.5D substrates, inter-chiplet I/O overheads, assembly yield issues and cost of the die-to-substrate bonding can out-strip the yield and system composability benefits that chipletization of large silicon systems offers. The results reveal that for microprocessor class chiplets, the minimum size of chiplets would be around 40 mm^2, and 200 mm^2 for random logic. Therefore, bring-your-own hardened intellectual property (IP) business models may not be feasible as 40 mm^2 is very large real estate and would require multiple IPs in a chiplet. Selection of the right IPs requires an understanding of the diverse set of applications such chiplets would be targeted towards.

NVIDIA showed how careful optimization of architecture, design and packaging technology can be leveraged to target GPUs for different markets such as HPC, AI and so on. They propose a composable on-package GPU architecture[91] to provide domain specialization. In one incarnation, an additional cache layer is realized by either 3D integrating a cache die beneath the GPU die or 2.5D integrating multiple cache dies between the GPU and the HBM devices on the package. Each of these options offers different performance, power and physical size trade-offs, and just by leveraging packaging constructs with architectural optimizations, the paper showed that the same training performance can be achieved with a 50% fewer number of GPU instances.

These examples show that STCO can unleash the true potential of SiPs. To enable this, we need frameworks, methodologies and tools for STCO. Although industrial organizations have internal methodologies and frameworks, they are not publicly available and are largely ad hoc. With some effort in generalizable STCO frameworks, this has changed in recent years. We categorize the set of STCO frameworks into three categories: link level, component level and cross-stack system level (Fig. 6). All three levels of STCO can provide useful information about power, performance, cost and form-factor metrics but at different levels of abstraction and detail. Further, link-level and component-level modelling can feed into the true cross-stack STCO.

Link-level STCO

Advanced packaging technologies bridge the large gap between on-chip and off-package interconnects. Several frameworks have been built in the past to model and optimize the inter-die link characteristics. Recent works[38,92,93] analyse how different parameters of silicon substrates, such as interconnect length, inter-layer dielectric material, µbump pitch, inter-die spacing, electrostatic discharge (ESD) capacitance and so on, affect inter-chiplet bandwidth and energy efficiency. Using these tools, one can figure out which of these parameters should be improved upon or invested in. For example, it has been shown that scaling µbump pitches below 20 µm would not provide meaningful bandwidth or energy efficiency gains unless the I/O ESD requirement is reduced[38]. Signalling figures of merit have been developed as well[94]. These frameworks rely on simple I/O circuits to perform the design

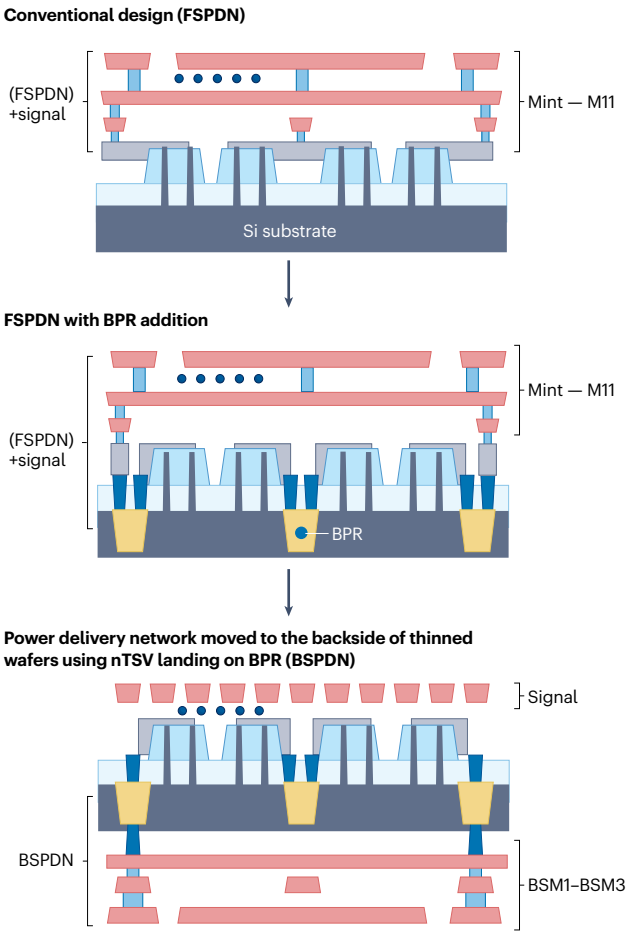


Fig. 5 | Backside power delivery network (BSPDN) enablement for power delivery. A conventional frontside power delivery network (FSPDN) is first augmented with backside power rails (BPRs) to relieve local interconnect congestion followed by moving the entire power delivery network to the backside using backside metal (BSM) layers[143]. This eliminates power delivery overheads from frontside signal routing as well as reducing voltage droop. TSV, through-silicon via.

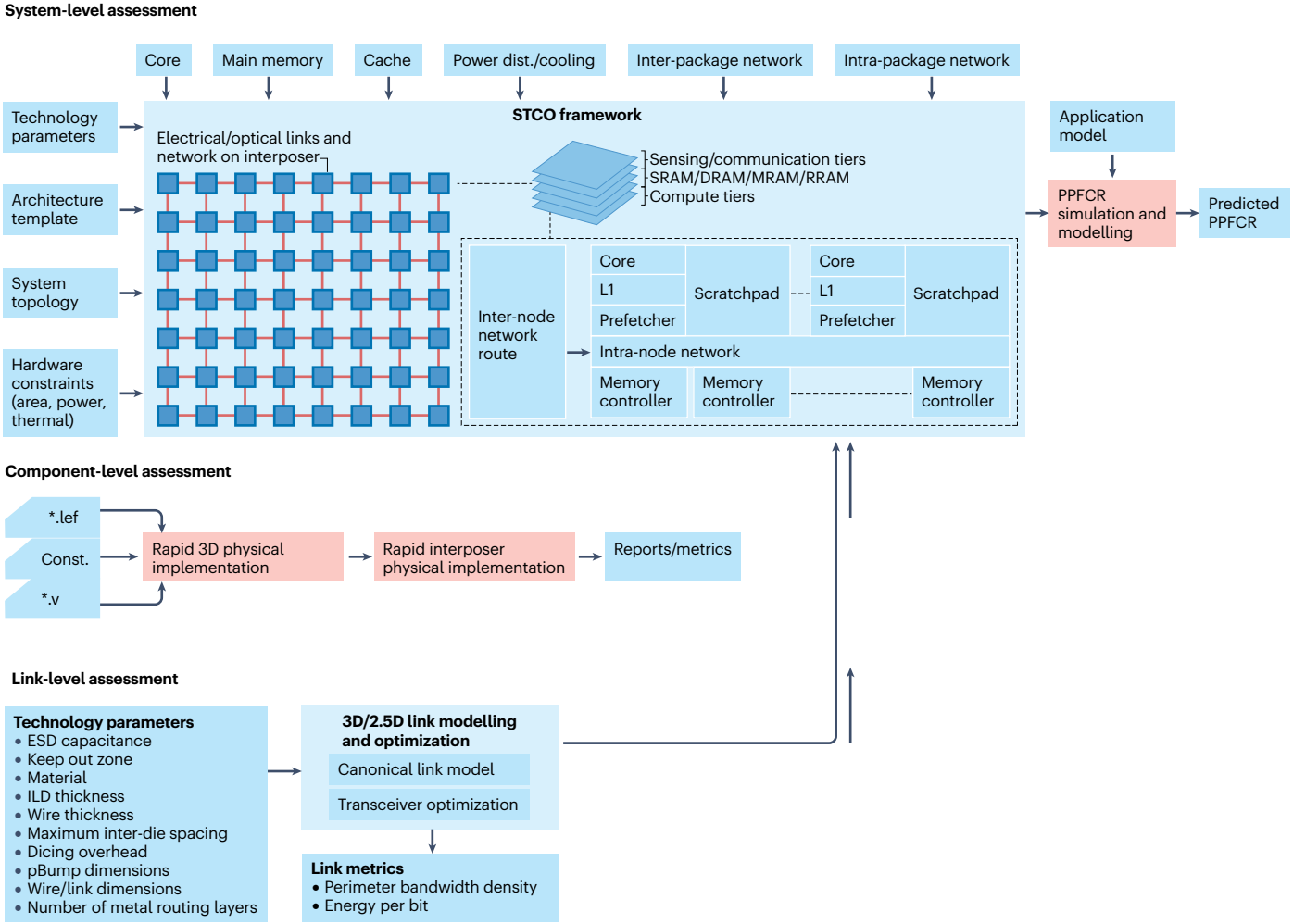


Fig. 6 | Overview of a system technology co-optimization (STCO) framework to predict system-level power, performance, form factor, cost and reliability (PPFCR). Link-level assessment primarily deals with inter-chiplet connectivity and evaluation of efficiency and performance of the link as a function of integration and input/output (I/O) technology. Component-level abstraction evaluates multi-chiplet 2.5D/3D systems by estimating power and performance area by rapid physical implementation of system. Finally, system-level assessment with cross-stack evaluation accounting of hardware, software, cooling and power delivery in one self-consistent framework. DRAM, dynamic random access memory; ESD, electrostatic discharge; ILD, inter-level dielectric; MRAM, magnetoresistive random access memory; RRAM, resistive random access memory; SRAM, static random access memory.

space exploration, which is suitable for integration technologies where the links are very short. On the other hand, organic substrates could be suitable for cost reasons[19,25] and, therefore, it is possible to co-optimize the I/O circuit design for organic substrates with longer links (five to ten times) and less density than that of silicon interposers.

Link-level STCO, however, does not cover the system-level implications of link characteristics. For example, a two times reduction in link energy efficiency may affect total power by a couple of percentage points while improving significantly reliability and cost. Although past works have laid a foundation, more comprehensive tools are needed to explore the design space of different integration schemes and their impact on the overall system. First, characteristics of substrate technologies such as their material properties affecting cost and reliability, as well as interconnect (wiring and bump) characteristics, need to be available. Process design kits and models in standardized formats need to be available to chip designers for simulation even during early phases of technology development, similar to early process design kits made available for advanced CMOS nodes. Second, details and requirements for the ESD protection circuitry are rarely available, and often designers over-design ESD circuitry and rely on post-silicon statistics to understand the impact of ESD events. Therefore, standardized ESD requirements based on the manufacturing environment should be available to the designers. Third, I/O circuits are often over-designed and made available as IPs. These IPs are usually used as is or with minor tweaks inside a chiplet, thus leading to suboptimal system-level power–performance–area characteristics. Therefore, I/O circuit generators alongside compact analytical models should be developed such that end-to-end interconnect characteristics, including the receiver and transmitters, could be evaluated and their impact on the overall area, power and performance of the chip architecture could be analysed early on. This could enable better co-optimization of the I/Os on the chiplets alongside the parameters of the integration technology. Future

link STCO research and development should address these shortcomings and develop tools and models using standard EDA and design tools for us to fully leverage today's integration technologies and drive the next generation of these technologies. Link-level STCO tools should generate abstract final models of the links and I/Os, which can then be used in higher-level tools such as component-level STCO tools. This could enable to evaluate the true impact of the interconnect technology at the system level.

Component-level STCO

The second class of STCO approaches is a natural extension of DTCO methodologies and leverages commercial and academic physical implementation EDA tools. These approaches take one or more benchmark designs (usually modestly sized) and go through an entire chip and system realization flow (placement, routing, power distribution and so on) for multiple chiplets integrated into a system (2.5D or 3D). Such component-level STCO approaches have been used to compare interposer types[95–97], assess backside power delivery[98–101], evaluate benefits of monolithic[9,102,103] or other 3D integration[104] and so on. The primary advantage of such approaches is the accuracy of the analyses performed and their ability to expose the design enablement challenges of new integration technologies. Unfortunately, there are several limitations, especially in the context of STCO for advanced integration. First, these approaches are not scalable to real systems that can have gate counts exceeding 100 M spanning several chiplets. Such component implementation approaches worked reasonably well for DTCO in the context of patterning[10,14,105–107] where results from small design blocks could be generalized to larger SoCs. However, generalization is difficult for large multi-chiplet packages. For example, inter-chiplet signalling overhead can look much worse for small chiplets[49] (that is, I/O cells and bumps can occupy a much larger fraction of chiplet area) whereas thermal and power delivery problems can look easier (that is, complex power delivery and cooling schemes are unnecessary for small chiplets with low power consumption, whereas they are a major challenge in large high-power systems which have been the primary user of advanced integration). Second, these strategies require evaluation of the underlying integration approaches to be mature enough to have tool-usable models such as process design kits, assembly design kits and so on, which for early technology exploration are rarely available. Third, most assessments rely on new EDA capability development. For example, a pseudo-3D design implementation flow using 2D tools[104,108] is necessary to do any 3D integration STCO. Although this has the benefit of simultaneous design enablement of the technology, it severely limits the pathfinding space in STCO. Finally, such component-level approaches ignore the system trade-offs which are only visible when the system architecture and the application workload running on it are accounted for.

Some of these shortcomings can be addressed by future research. Scalability issues can be partly dealt with by abstracting the physical implementation to block level rather than gate level, which should give one or two orders of magnitude speed up at the cost of hiding some detail (that is, solve block-level partitioning, macro placement and so on as a quicker predictor than detailed gate-level place)[109]. Further, block-level flows, which currently are not available, could be fed by automated system-level benchmark generators built on top of architecture design space exploration tools[110–112]. To better connect the component-level STCO with applications and architectures, analytical performance/power macro models can be developed to be used in conjunction with physical estimates, for example to constrain the physical implementation.

Cross-stack emerging STCO approaches

Advanced chiplet integration technologies are platforms for building large systems inside a package. However, traditional DTCO approaches and piecemeal STCO approaches at the link and component levels are not suitable for understanding the system-level impact of these technologies. Such traditional approaches often do not model the entire system and do not allow to understand the impact of decisions at the lower levels on application performance and power. Therefore, newer cross-stack frameworks are needed where the interplay of technology, design hardware and software architecture can be explored by evaluating the impact of the choices made at each layer of the stack at the application level. Recently, a set of efforts to build cross-stack STCO tools have emerged. On the 3D integration front, several hardware/software co-synthesis frameworks have been proposed[109,113–115] to explore the 3D SoC design space. For interposer-based designs, Floorplet[116] is a framework that can optimally partition a fixed SoC design into chiplets based on yield and reliability, generate the chiplet design, optimize the interposer floorplan and perform cycle accurate performance simulation to optimize the entire system. DeepFlow[110], on the other hand, allows co-optimization of the chip and the scale-out of distributed system architecture alongside the software parallelization strategy for a given machine learning workload. Low-level technology parameters such as area, power and performance of building blocks (arithmetic logic units (ALUs), SRAMs, DRAMs, interconnects) and physical constraints (power, thermal, pin density and so on) are provided as inputs. The framework can automatically search the architecture design space, model the performance and, using gradient-descent search approaches, explore the vast space of hardware–software co-design. These approaches are critical to understanding the end-to-end impact of advanced technology development on application-level performance. Unfortunately, these tools suffer from shortcomings. As rich cross-stack frameworks need to comprehend and search over an impossibly large parameter space (technology, design, architecture, software and so on), these tools are built around simplified abstractions and assumptions that render them useful for limited subsets of the design space. For example, Floorplet works with a given design/architecture and therefore is not able to show what the changes in technology parameters could lead to if one had to re-architect the SoC. DeepFlow targets deep learning workloads and targets exploration of a vast design space. The architecture generation and performance simulation portions of the tool are designed to be workload-specific for runtime efficiency reasons and therefore lack generality.

These tools, however, provide a solid foundation for future research. As is evident, system-level performance modelling is critical for cross-stack STCO tools, but these models need to be fast, relatively accurate and scalable for it to be useful when doing large design space exploration. This requires building composable analytical models for different types of architectural blocks such as CPU cores, SIMD cores, accelerators, registers and memory blocks. However, these analytical models need to be coupled with abstract workload modelling where the characteristics of the key kernels can be abstracted, and application dataflow graphs can be input to the simulation model. On the architecture generation front, link-level and component-level tools can help guide the generation of feasible and realizable architectures and SoC designs by providing abstract power, performance and area models of the different hardware components. Besides, several novel search techniques (for example, genetic algorithms, particle swarm optimization, data-driven machine learning-assisted techniques) can

be used to assess the design space and co-optimize across the stack of technology, chip and system architecture, and software strategies.

Future STCO contexts and directions

Apart from looking at advanced integration from a conventional computing lens, packaging could enable completely new sensing/computing paradigms. Flexible computing systems[117,118], biocompatible electronics[119,120] and heterogeneously co-integrated sensor and compute[121–123] are few examples of such emerging areas of research. Therefore, it is equally important to consider STCO in these contexts.

Lastly, we want to emphasize a few important system metrics that we have not discussed but are becoming increasingly important[124], especially in use contexts such as automotive[125]. The choice of materials in integration can have a substantial impact on thermo-mechanical stresses[126–128] but this needs to be balanced against cost and performance considerations. Advanced packaging can both help with supply-chain security[129–131] and expose more challenges in system security[132–135]. Environmentally sustainable manufacturing and reducing the life-cycle carbon and waste footprint of electronics have become critical[136–141]. Packaging is a big part of this footprint and system design using chiplets can open novel ways of looking at the sustainability problem as well as, potentially, additional carbon footprint. For example, any trade-off between the recyclability of packaging materials and their performance implication is part of STCO.

Conclusions

Advanced packaging is seen to enable 'more than Moore' scaling. Including integration/packaging as part of the performance, energy and total cost of ownership optimization requires expanding the scope of DTCO to include the system. Such STCO extends to aspects of logic/memory chip design/manufacturing as well as heterogeneously integrated power delivery, integrated cooling approaches and off-package interconnect. In this Review, we have discussed some of the existing approaches to STCO. Furthermore, to truly harness the full value of the technology, we argue that one needs to expand the scope of STCO to be cross-stack and account for micro-architecture and software/algorithms as well. Such materials-to-software frameworks and methodologies that could allow true STCO are still in their infancy and are likely to be domain-specific to bound the problem to be tractable.

STCO for advanced integration has the potential to become a vibrant, high-impact area of research and development in the coming decade and we encourage researchers to take a cross-disciplinary software–hardware–technology cross-stack approach to it.

Published online: 2 September 2024

References

1. Mallik, A. et al. Maintaining Moore's law: enabling cost-friendly dimensional scaling. In *Proc. Volume 9422, Extreme Ultraviolet (EUV) Lithography VI* 531–542 (SPIE, 2015).
2. Doug O'Laughlin, S. The rising tide of semiconductor cost. *SemiWiki* https://semiwiki.com/semiconductor-services/308018-the-risingtide-of-semiconductor-cost/ (2022).
3. Mallik, A. et al. The impact of sequential-3D integration on semiconductor scaling roadmap. In *2017 IEEE International Electron Devices Meeting (IEDM)* 32.1.1–31.1.4 (IEEE, 2017).
4. Iyer, S. S. Heterogeneous integration for performance and scaling. *IEEE Trans. Compon. Packag. Manuf. Technol.* **6**, 973–982 (2016).
   This paper makes a case that packaging dimensions have scaled much more poorly so far than within-chip dimensions but heterogeneous integration will be the backbone of sustaining Moore's law in the years ahead.
5. Zhang, J., Patil, N., Philip Wong, H.-S. & Mitra, S. Overcoming carbon nanotube variations through co-optimized technology and circuit design. In *2011 International Electron Devices Meeting* 4.6.1–4.6.4 (IEEE, 2011).
6. Gupta, S. K. & Roy, K. Device-circuit co-optimization for robust design of FinFET-Based SRAMs. *IEEE Des. Test. Comput.* **30**, 29–39 (2013).
7. Zhang, Z. et al. New-generation design-technology co-optimization (DTCO): machine-learning assisted modeling framework. In *2019 Silicon Nanoelectronics Workshop (SNW)* 1–2 (IEEE, 2019).
8. Wang, S., Pan, A., Chui, C. O. & Gupta, P. PROCEED: a pareto optimization-based circuit-level evaluator for emerging devices. *IEEE Trans. Very Large Scale Integr. Syst.* **24**, 192–205 (2016).
9. Wang, W.-C. & Gupta, P. Efficient layout generation and design evaluation of vertical channel devices. *IEEE Trans. Comput. Des. Integr. Circuits Syst.* **35**, 1449–1460 (2015).
10. Ghaida, R. S. & Gupta, P. DRE: a framework for early co-evaluation of design rules, technology choices, and layout methodologies. *IEEE Trans. Comput. Des. Integr. Circuits Syst.* **31**, 1379–1392 (2012).
11. Ryckaert, J. et al. Design Technology co-optimization for N10. In *Proc. IEEE 2014 Custom Integrated Circuits Conference* 1–8 (IEEE, 2014).
12. Yeric, G. et al. The past present and future of design-technology co-optimization. In *Proc. IEEE 2013 Custom Integrated Circuits Conference* 1–8 (IEEE, 2013).
13. Capodieci, L., Gupta, P., Kahng, A. B., Sylvester, D. & Yang, J. Toward a methodology for manufacturability-driven design rule exploration. In *Proc. 41st annual Design Automation Conference* 311–316 (ACM, 2004).
14. Kahng, A., Kahng, A. B., Lee, H. & Li, J. PROBE: a placement, routing, back-end-of-line measurement utility. *IEEE Trans. Comput. Aided Des. Integr. Circ. Syst.* **37**, 1459–1472 (2017).
15. Collaert, N. Future scaling: where systems and technology meet. In *2020 IEEE International Solid-State Circuits Conference - (ISSCC)* 25–29 (IEEE, 2020).
16. Samavedam, S. B. et al. Future logic scaling: towards atomic channels and deconstructed chips. In *2020 IEEE International Electron Devices Meeting (IEDM)* 1.1.1–1.1.10 (IEEE, 2020).
17. Lenihan, T. G., Matthew, L. & Vardaman, E. J. Developments in 2.5D: The role of silicon interposers. In *2013 IEEE 15th Electronics Packaging Technology Conference (EPTC 2013)* 53–55 (IEEE, 2013).
18. Chiou, D. The Microsoft catapult project. In *2017 IEEE International Symposium on Workload Characterization (IISWC)* 124–124 (IEEE, 2017).
19. Naffziger, S. et al. Pioneering chiplet technology and design for the AMD EPYC™ and Ryzen™ processor families : industrial product. In *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)* 57–70 (2021).
    This paper details the technology challenges that motivated AMD to use chiplets in their product families.
20. Zhu, M., Zhuo, Y., Wang, C., Chen, W. & Xie, Y. Performance evaluation and optimization of HBM-enabled GPU for data-intensive applications. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017* 1245–1248 (IEEE, 2017).
21. Elster, A. C. & Haugdahl, T. A. Nvidia hopper GPU and grace CPU highlights. *Comput. Sci. Eng.* **24**, 95–100 (2022).
22. Lee, C.-C. et al. An overview of the development of a GPU with integrated HBM on silicon interposer. In *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)* 1439–1444 (IEEE, 2016).
23. Huang, P. K. et al. Wafer level system integration of the fifth generation CoWoS®-S with high performance Si interposer at 2500 mm2. In *2021 IEEE 71st Electronic Components and Technology Conference (ECTC)* 101–104 (IEEE, 2021).
24. Hu, Y.-C. et al. CoWoS architecture evolution for next generation HPC on 2.5D system in package. In *2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)* 1022–1026 (IEEE, 2023).
25. Mahajan, R. et al. Embedded multi-die interconnect bridge (EMIB) — a high density, high bandwidth packaging interconnect. In *2016 IEEE 66th Electronic Components and Technology Conference (ECTC)* 557–565 (IEEE, 2016).
26. Duan, G., Kanaoka, Y., McRee, R., Nie, B. & Manepalli, R. Die embedding challenges for EMIB advanced packaging technology. In *2021 IEEE 71st Electronic Components and Technology Conference (ECTC)* 1–7 (IEEE, 2021).
27. Wuu, J. et al. 3D V-Cache: the implementation of a hybrid-bonded 64MB stacked cache for a 7nm x86-64 CPU. In *2022 IEEE International Solid-State Circuits Conference (ISSCC)* 428–429 (IEEE, 2022).
28. Chang, J. et al. A 3nm 256Mb SRAM in FinFET technology with new array banking architecture and write-assist circuitry scheme for high-density and low-VMIN applications. In *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)* 1–2 (IEEE, 2023).
29. Liu, J. et al. A reliability enhanced 5nm CMOS technology featuring 5th generation FinFET with fully-developed EUV and high mobility channel for mobile SoC and high performance computing application. In *2020 IEEE International Electron Devices Meeting (IEDM)* 9.2.1–9.2.4 (IEEE, 2020).
30. Lapedus, M. 5nm vs. 3nm. *Semiconductor Engineering* https://semiengineering.com/5nm-vs-3nm/ (2023).
31. Chia, H.-J. et al. Ultra high density low temperature SoIC with sub-0.5 µm bond pitch. In *2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)* 1–4 (IEEE, 2023).
32. TechPowerUp. NVIDIA H100 SXM5 96 GB. *Techpowerup.com* https://www.techpowerup.com/gpu-specs/h100-sxm5-96-gb.c3974 (2024).
33. Institute for Cyber-Enabled Research. Cluster amd20 with AMD CPUs. *ICER* https://docs.icer.msu.edu/Cluster_amd20_with_AMD_CPUs/ (2023).
34. Kennedy, P. AMD Milan-X Delivers AMD EPYC Caches to the GB-era. *Serve The Home* https://www.servethehome.com/amd-milan-x-delivers-amd-epyc-caches-to-the-gb-era/ (2022).

35. Sevilla, J. et al. Compute trends across three eras of machine learning. In *2022 International Joint Conference on Neural Networks (IJCNN)* 1–8 (IEEE, 2022).
36. Gholami, A. AI and memory wall. *Medium* https://medium.com/riselab/aiand-memory-wall-2cb4265cb0b8 (2023).
37. Shaw, D. E. et al. Anton 3: twenty microseconds of molecular dynamics simulation before lunch. In *Proc. International Conference for High Performance Computing, Networking, Storage and Analysis* 1–11 (ACM, 2021).
38. Pal, S. & Gupta, P. Pathfinding for 2.5D interconnect technologies. In *Proc. Workshop on System-Level Interconnect: Problems and Pathfinding Workshop* 1–8 (ACM, 2020).
39. Park, M.-J. et al. A 192-Gb 12-high 896-GB/s HBM3 DRAM with a TSV auto-calibration scheme and machine-learning-based layout optimization. *IEEE J. Solid State Circ.* **58**, 256–269 (2023).
40. Park, S. J. et al. Industry's first 7.2 Gbps 512GB DDR5 module. In *2021 IEEE Hot Chips 33 Symposium (HCS)* 1–11 (IEEE, 2021).
41. Park, S. & Huddar, V. A. Design and analysis of power integrity of DDR5 dual in-line memory modules. In *2022 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)* 1–3 (IEEE, 2022).
42. AMD. 3D V-Cache™ technology. https://www.amd.com/en/products/processors/technologies/3d-v-cache.html (2024).
43. Pal, S. et al. Designing a 2048-chiplet, 14336-core waferscale processor. In *2021 58th ACM/IEEE Design Automation Conference (DAC)* 1183–1188 (IEEE, 2021).
44. Lie, S. Cerebras architecture deep dive: first look inside the HW/SW co-design for deep learning : Cerebras Systems. In *2022 IEEE Hot Chips 34 Symposium (HCS)* 1–34 (IEEE, 2022).
45. Talpes, E., Williams, D. & Sarma, D. D. DOJO: the microarchitecture of Tesla's exa-scale computer. In *2022 IEEE Hot Chips 34 Symposium (HCS)* 1–28 (IEEE, 2022).
46. Pal, S. et al. Architecting waferscale processors - a GPU case study. In *2019 IEEE International Symposium on High Performance Computer Architecture (HPCA)* 250–263 (IEEE, 2019).
47. Rocki, K. et al. Fast stencil-code computation on a wafer-scale processor. In *Proc. International Conference for High Performance Computing, Networking, Storage and Analysis* 1–14 (IEEE, 2020).
48. Feng, Y. & Ma, K. Chiplet actuary: a quantitative cost model and multi-chiplet architecture exploration. In *Proc. 59th ACM/IEEE Design Automation Conference* 121–126 (ACM, 2022).
49. Graening, A., Pal, S. & Gupta, P. Chiplets: how small is too small? In *2023 60th ACM/IEEE Design Automation Conference (DAC)* 1–6 (IEEE, 2023).
50. Peng, H., Davidson, S., Shi, R., Song, S. L. & Taylor, M. Chiplet cloud: building AI supercomputers for serving large generative language models 2023. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2307.02666 (2024).
51. Liu, C. C. et al. High-performance integrated fan-out wafer level packaging (InFO-WLP): Technology and system integration. In *2012 International Electron Devices Meeting* 14.1.1–14.1.4 (IEEE, 2012).
52. Lujan, A. P. Comparison of package-on-package technologies utilizing flip chip and fan-out wafer level packaging. In *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)* 2089– 2094 (IEEE, 2018).
53. Shah, M. et al. Module/SiP packaging trends. In *2019 Electron Devices Technology and Manufacturing Conference (EDTM)* 82–84 (IEEE, 2019).
54. Octavo Systems. SiP technology. https://octavosystems.com/sip-technology/ (2023).
55. Choquette, J. & Gandhi, W. NVIDIA A100 GPU: performance & innovation for GPU computing. In *2020 IEEE Hot Chips 32 Symposium (HCS)* 1–43 (IEEE, 2020).
56. Macri, J. AMD's next generation GPU and high bandwidth memory architecture: FURY. In *2015 IEEE Hot Chips 27 Symposium (HCS)* 1–26 (IEEE, 2015).
57. Sodani, A. et al. Knights landing: second-generation Intel Xeon Phi product. *IEEE Micro* **36**, 34–46 (2016).
58. Wade, M. et al. TeraPHY: a chiplet technology for low-power, high-bandwidth in-package optical I/O. *IEEE Micro* **40**, 63–71 (2020).
59. Ren, H., Sahoo, K., Xiang, T., Ouyang, G. & Iyer, S. S. Demonstration of a power-efficient and cost-effective power delivery architecture for heterogeneously integrated wafer-scale systems. In *2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)* 1614–1621 (IEEE, 2023).
60. Desai, N. et al. A 32-A, 5-V-input, 94.2% peak efficiency highfrequency power converter module featuring package-integrated lowvoltage GaN nMOS power transistors. *IEEE J. Solid-State Circ.* **57**, 1090–1099 (2022).
61. Radhakrishnan, K., Swaminathan, M. & Bhattacharyya, B. K. Power delivery for high-performance microprocessors—challenges, solutions, and future trends. *IEEE Trans. Compon. Packag. Manuf. Technol.* **11**, 655–671 (2021).
62. Hagge, J. State-of-the-art multichip modules for avionics. *IEEE Trans. Compon. Hybrids Manuf. Technol.* **15**, 29–42 (1992).
63. Rinne, R. & Barbour, D. Multi-chip module technology. *Electrocompon. Sci. Technol.* **10**, 31–49 (1982).
64. Sun, P. et al. Development of a new package-on-package (PoP) structure for next-generation portable electronics. In *2010 Proceedings 60th Electronic Components and Technology Conference (ECTC)* 1957–1963 (IEEE, 2010).
65. Fontanelli, A. System-in-package technology: opportunities and challenges. In *9th International Symposium on Quality Electronic Design (isqed 2008)* 589–593 (IEEE, 2008).
66. Jeng, S.-P. & Liu, M. Heterogeneous and chiplet integration using organic interposer (CoWoS-R). In *2022 International Electron Devices Meeting (IEDM)* 3.2.1–3.2.4 (IEEE, 2022).
67. Roth, A. et al. Heterogeneous power delivery for 7nm high-performance chiplet-based processors using integrated passive device and in-package voltage regulator. In *2020 IEEE Symposium on VLSI Technology* 1–2 (IEEE, 2020).
68. Moore, S. K. 3 paths to 3D processors. *IEEE Spectrum* **59**, 24–29 (2022).
69. Cline, B., Prasad, D., Beyne, E. & Zografos, O. Power from below: buried interconnects will help save Moore's law. *IEEE Spectr.* **58**, 46–51 (2021).
70. Kobrinsky, M. et al. Novel cell architectures with back-side transistor contacts for scaling and performance. In *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)* 1–2 (IEEE, 2023).
71. Hafez, W. et al. Intel PowerVia pechnology: backside power delivery for high density and high-performance computing. In *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)* 1–2 (IEEE, 2023).
72. Sisto, G. et al. IR-drop analysis of hybrid bonded 3D-ICs with backside power delivery and μ- & n- TSVs. In *2021 IEEE International Interconnect Technology Conference (IITC)* 1–3 (IEEE, 2021).
73. Chen, R. et al. Backside PDN and 2.5D MIMCAP to double boost 2D and 3D ICs IR-drop beyond 2nm node. In *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)* 429–430 (IEEE, 2022).
74. Siricharoenpanich, A., Wiriyasart, S., Srichat, A. & Naphon, P. Thermal management system of CPU cooling with a novel short heat pipe cooling system. *Case Stud. Therm. Eng.* **15**, 100545 (2019).
75. Pambudi, N. A. et al. The immersion cooling technology: current and future development in energy saving. *Alex. Eng. J.* **61**, 9509–9527 (2022).
76. Elliott, J., Lebon, M. & Robinson, A. Optimising integrated heat spreaders with distributed heat transfer coefficients: a case study for CPU cooling. *Case Stud. Therm. Eng.* **38**, 102354 (2022).
77. Nelson, C. Thermal management implications for heterogeneous integrated packaging. *SemiconductorEngineering* https://semiengineering.com/thermal-management-implications-for-heterogeneous-integrated-packaging/ (2022).
78. Zhou, M., Li, L., Hou, F., He, G. & Fan, J. Thermal modeling of a chipletbased packaging with a 2.5-D through-silicon via interposer. *IEEE Trans. Compon. Packag. Manuf. Technol.* **12**, 956–963 (2022).
79. Lin, S.-C. & Banerjee, K. Cool chips: opportunities and implications for power and thermal management. *IEEE Trans. Electron Dev.* **55**, 245–255 (2008).
80. Eris, F. et al. Leveraging thermally-aware chiplet organization in 2.5D systems to reclaim dark silicon. In *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)* 1441–1446 (IEEE, 2018).
81. Chen, X. et al. Improving the thermal reliability of photonic chiplets on multicore processors. *Integration* **86**, 9–21 (2022).
82. Luo, G., Shi, Y. & Cong, J. An analytical placement framework for 3-D ICs and Its extension on thermal awareness. *IEEE Trans. Comput. Des. Integr. Circuits Syst.* **32**, 510–523 (2013).
83. Goplen, B. & Sapatnekar, S. Efficient thermal placement of standard cells in 3D ICs using a force directed approach. In *ICCAD-2003. International Conference on Computer Aided Design* 86–89 (IEEE, 2003).
84. Ma, Y., Delshadtehrani, L., Demirkiran, C., Abellan, J. L. & Joshi, A. iTAP-2.5D: a thermally-aware chiplet placement methodology for 2.5D systems. In *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)* 1246–1251 (IEEE, 2021).
85. Zhang, Y. et al. Coupled electrical and thermal 3D IC centric microfluidic heat sink design and technology. In *2011 IEEE 61st Electronic Components and Technology Conference (ECTC)* 2037–2044 (IEEE, 2011).
86. Lau, J. H. & Yue, T. G. Thermal management of 3D IC integration with TSV (through silicon via). In *2009 59th Electronic Components and Technology Conference* 635–640 (IEEE, 2009).
87. Shi, B., Srivastava, A. & Bar-Cohen, A. Hybrid 3D-IC cooling system using micro-fluidic cooling and thermal TSVs. In *2012 IEEE Computer Society Annual Symposium on VLSI* 33–38 (IEEE, 2012).
88. Zhang, Y., Dembla, A. & Bakir, M. S. Silicon micropin-fin heat sink with integrated TSVs for 3-D ICs: tradeoff analysis and experimental testing. *IEEE Trans. Compon. Packag. Manuf. Technol.* **3**, 1842–1850 (2013).
89. Zhang, Y., Zhang, Y. & Bakir, M. S. Thermal design and constraints for heterogeneous integrated chip stacks and isolation technology using air gap and thermal bridge. *IEEE Trans. Compon. Packag. Manuf. Technol.* **4**, 1914–1924 (2014).
90. IEEE Spectrum. Graphcore uses TSMC 3D chip tech to speed AI by 40%. *IEEE Spectrum* https://spectrum.ieee.org/graphcore-ai-processor (2022).
91. Fu, Y., Bolotin, E., Chatterjee, N., Nellans, D. & Keckler, S. GPU domain specialization via composable on-package architecture. *ACM Trans. Architecture Code Optim.* **19**, 1–23 (2022).
92. Pantano, N. et al. Technology optimization for high bandwidth density applications on 3D interposer. In *2016 6th Electronic System-Integration Technology Conference (ESTC)* 1–6 (2016).
93. Jangam, S. et al. Latency, bandwidth and power benefits of the SuperCHIPS integration scheme. In *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)* 86–94 (IEEE, 2017).
94. Jangam, S. & Iyer, S. S. A signaling figure of merit (s-FoM) for advanced packaging. *IEEE Trans. Compon. Packag. Manuf. Technol.* **10**, 1758–1761 (2020).
This paper proposes a simple signalling figure of merit for inter-chiplet interconnect and compares various link+ packaging technologies using it: an example of link-level STCO.

95. Stow, D., Xie, Y., Siddiqua, T. & Loh, G. H. Cost-effective design of scalable high-performance systems using active and passive interposers. In *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)* 728–735 (IEEE, 2017).
96. Kim, J. et al. Architecture, chip, and package codesign flow for interposer-based 2.5-D chiplet integration enabling heterogeneous IP reuse. *IEEE Trans. Very Large Scale Integr. Syst.* **28**, 2424–2437 (2020).
97. Stow, D., Akgun, I. & Xie, Y. Investigation of cost-optimal network-on-chip for passive and active interposer systems. In *2019 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)* 1–8 (IEEE, 2019).
98. Zhu, L., Jo, C. & Lim, S. K. Power delivery solutions and PPA impacts in micro-bump and hybrid-bonding 3D ICs. *IEEE Trans. Compon. Packag. Manuf. Technol.* **12**, 1969–1982 (2022).
99. Lanzillo, N. A. et al. Benchmarking power delivery network designs at the 5-nm technology node. *IEEE Trans. Electron. Devices* **69**, 7135–7140 (2022).
100. Choi, S. et al. PROBE3.0: a systematic framework for design-technology pathfinding with improved design enablement. *IEEE Trans. Comput. Aided Des. Integr. Circ. Syst.* **43**, 1218–1231 (2023).
101. Sisto, G. et al. System-level evaluation of 3D power delivery network at 2nm node. In *Proc. Volume 12495, DTCO and Computational Patterning II* 207–217 (SPIE, 2023).
102. Abdi, D. B. et al. 3D SRAM macro design in 3D nanofabric process technology. *IEEE Trans. Circ. Syst. I Regul. Pap.* **70**, 2858–2867 (2023).
103. Liebmann, L., Smith, J., Chanemougame, D. & Gutwin, P. CFET design options, challenges, and opportunities for 3D integration. In *2021 IEEE International Electron Devices Meeting (IEDM)* 3.1.1–3.1.4 (IEEE, 2021).
104. Agnesina, A. et al. Power, performance, area, and cost analysis of face-to-face-bonded 3-D ICs. *IEEE Trans. Compon. Packag. Manuf. Technol.* **13**, 300–314 (2023).
Together with Zhu et al. (2022), this work is an example of component-level STCO approaches which evaluate technologies using physical implementation of benchmark designs.
105. Liebmann, L. et al. Overcoming scaling barriers through design technology Cooptimization. In *2016 IEEE Symposium on VLSI Technology* 1–2 (IEEE, 2016).
106. Ryckaert, J. et al. DTCO at N7 and beyond: patterning and electrical compromises and opportunities. In *Proc. Volume 9427, Design-Process-Technology Co-optimization for Manufacturability IX* 101–108 (SPIE, 2015).
107. Kagalwalla, A. A., Lam, M., Adam, K. & Gupta, P. EUV-CDA: Pattern shift aware critical density analysis for EUV mask layouts. In *2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC)* 155–160 (IEEE, 2014).
108. Ku, B. W., Chang, K. & Lim, S. K. Compact-2D: a physical design methodology to build commercial-quality face-to-face-bonded 3D ICs. In *Proc. 2018 International Symposium on Physical Design* 90–97 (ACM, 2018).
109. Priyadarshi, S. et al. Pathfinder 3D: A flow for system-level design space exploration. In *2011 IEEE International 3D Systems Integration Conference (3DIC)* 1–8 (IEEE, 2012).
110. Ardalani, N., Pal, S. & Gupta, P. DeepFlow: a cross-stack pathfinding framework for distributed AI systems. *ACM Trans. Des. Autom. Electron. Syst.* https://doi.org/10.1145/3635867 (2023).
This work is one of the earliest attempts at developing algorithms for a technology cross-stack STCO framework in the context of distributed training of large neural networks.
111. Rashidi, S., Sridharan, S., Srinivasan, S. & Krishna, T. ASTRA-SIM: enabling SW/HW co-design exploration for distributed DL training platforms. In *2020 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)* 81–92 (IEEE, 2020).
112. Parashar, A. et al. Timeloop: a systematic approach to DNN accelerator evaluation. In *2019 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)* 304–315 (IEEE, 2019).
113. Zou, Q., Chen, Y., Xie, Y. & Su, A. System-level design space exploration for three-dimensional (3D) SoCs. In *2011 Proceedings of the Ninth IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)* 385–388 (IEEE, 2011).
114. Agrawal, P. et al. System level comparison of 3D integration technologies for future mobile MPSoC platform. *IEEE Embedded Syst. Lett.* **6**, 85–88 (2014).
115. Siozios, K., Papanikolaou, A. & Soudris, D. A method and tool for early design/technology search-space exploration for 3D ICs. In *IFIP/IEEE International Conference on VLSI-SoC* 359–364 (2008).
116. Chen, S. et al. Floorplet: performance-aware floorplan framework for chiplet integration. *IEEE Trans. Comput. Aided Des. Integr. Circ. Syst.* **43**, 1638–1649 (2024).
117. Biggs, J. et al. A natively flexible 32-bit arm microprocessor. *Nature* **595**, 532–536 (2021).
118. Bleier, N. et al. FlexiCores: low footprint, high yield, field reprogrammable flexible microprocessors. In *49th Annual International Symposium on Computer Architecture* 831–846 (ACM, 2022).
119. Jahanshahi, A., Salvo, P. & Vanfleteren, J. Stretchable biocompatible electronics by embedding electrical circuitry in biocompatible elastomers. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 6007–6010 (IEEE, 2012).
120. Wang, M. et al. Artificial skin perception. *Adv. Mater.* **33**, 2003014 (2021).
121. Lindsay, M. et al. Heterogeneous integration of CMOS sensors and fluidic networks using wafer-level molding. *IEEE Trans. Biomed. Circ. Syst.* **12**, 1046–1055 (2018).
122. Xie, L. et al. Heterogeneous integration of bio-sensing system-on-chip and printed electronics. *IEEE J. Emerg. Sel. Top. Circ. Syst.* **2**, 672–682 (2012).
123. Haruta, T. et al. 4.6 A 1/2.3inch 20Mpixel 3-layer stacked CMOS image sensor with DRAM. In *2017 IEEE International Solid-State Circuits Conference (ISSCC)* 76–77 (IEEE, 2017).
124. Liu, C. et al. Reliability challenges in advanced technology node: from transistor to circuit (invited). In *2020 IEEE 15th International Conference on Solid-State & Integrated Circuit Technology (ICSICT)* 1–4 (IEEE, 2020).
125. Sham, M.-L., Gao, Z., Leung, L. L.-W., Chen, Y.-C. & Chung, T. Advanced packaging technologies for automotive electronics. In *2007 8th International Conference on Electronic Packaging Technology* 1–5 (IEEE, 2007).
126. Iyer, S. S. & Bajwa, A. A. Reliability challenges in advance packaging. In *2018 IEEE International Reliability Physics Symposium (IRPS)* 4D.5–1–4D.5–4 (IEEE, 2018).
127. Chase, N. S., Irwin, R., Yang, Y. T., Ren, H. & Iyer, S. S. Reliability considerations for wafer scale systems. In *2021 IEEE 71st Electronic Components and Technology Conference (ECTC)* 84–89 (IEEE, 2021).
128. Yip, L., Lin, R., Lai, C. & Peng, C. Reliability challenges of high-density fan-out packaging for high-performance computing applications. In *2022 IEEE 72nd Electronic Components and Technology Conference (ECTC)* 1454–1458 (IEEE, 2022).
129. Xie, Y., Bao, C. & Srivastava, A. in *Hardware Protection through Obfuscation* 291–314 (Springer, 2017).
130. Gu, P. et al. Leveraging 3D technologies for hardware security: opportunities and challenges. In *2016 International Great Lakes Symposium on VLSI (GLSVLSI)* 347–352 (IEEE, 2016).
131. Imeson, F., Emtenan, A., Garg, S. & Tripunitara, M. Securing computer hardware using 3D integrated circuit (IC) technology and split manufacturing for obfuscation. In *Proc. 22nd USENIX conference on Security* 495–510 (USENIX Association, 2013).
132. Nabeel, M. et al. 2.5 D root of trust: secure system-level integration of untrusted chiplets. *IEEE Trans. Comput.* **69**, 1611–1625 (2020).
133. Xie, Y. et al. Security and vulnerability implications of 3D ICs. *IEEE Trans. Multi Scale Comput. Syst.* **2**, 108–122 (2016).
134. Knechtel, J. & Sinanoglu, O. On mitigation of side-channel attacks in 3D ICs: Decorrelating thermal patterns from power and activity. In *2017 54th ACM/EDAC/IEEE Design Automation Conference (DAC)* 1–6 (IEEE, 2017).
135. Dofe, J. et al. Security threats and countermeasures in three-dimensional integrated circuits. In *Proc. Great Lakes Symposium on VLSI 2017* 321–326 (ACM, 2017).
136. Wang, Y.-C., Chen, T.-C. T. & Wang, L.-C. Simulating a semiconductor packaging facility: sustainable strategies and short-time evidences. *Procedia Manuf.* **11**, 787–795 (2017).
137. Harland, J., Reichelt, T. & Yao, M. Environmental sustainability in the semiconductor industry. In *2008 IEEE International Symposium on Electronics and the Environment* 1–6 (IEEE, 2008).
138. Kerbusch, J. Why the electronics industry must address sustainability. In *EASS 2022; 11th GMM-Symposium* 1–3 (VDE, 2022).
139. Gandhi, A. et al. Metrics for sustainability in data centers. In *Proc. 1st Workshop on Sustainable Computer Systems Design and Implementation (HotCarbon 2022)* (2022).
140. Eeckhout, L. Towards sustainable computer architecture: a holistic approach. *HiPEAC Vision 2023* https://doi.org/10.5281/zenodo.7461989 (2023).
141. Bardon, M. G. et al. DTCO including sustainability: power-performance-area-cost-environmental score (PPACE) analysis for logic technologies. In *2020 IEEE International Electron Devices Meeting (IEDM)* 41–44 (IEEE, 2020).
142. IMEC. Getting the most out of your system. https://www.imec-int.com/en/articles/getting-most-out-your-system (2021).
143. Veloso, A. et al. Scaled FinFETs connected by using both wafer sides for routing via buried power rails. *IEEE Trans. Electron. Devices* **69**, 7173–7179 (2022).

Acknowledgements
P.G. discloses support for the research of this work from DARPA/SRC CHIMES JUMP 2.0 Center and National Science Foundation.

Competing interests
The authors declare no competing interests.

Additional information
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.