Physical Design Challenges and Solutions for Interposer-Based 3D Systems

Sergii Osmolovskyi and Jens Lienig
Dresden University of Technology, Dresden, Germany
sergii.osmolovskyi@tu-dresden.de, jens@ieee.org

Abstract

Three-dimensional (3D) chip integration ("More than Moore") is a promising alternative to traditional transistor scaling ("More Moore"). However, its industrial application is notably restricted by numerous design challenges, amplified by a lack of physical design tools. In order to exploit the advantages of 3D integration, layout designers and tool developers need to be fully aware of these challenges. We first investigate the variety of 3D architecture options and show that interposer-based systems are the most cost-effective candidate for heterogeneous chip design at present. Next, we review the system-level physical design challenges of interposer-based 3D ICs and outline possible solutions. Focusing on placement challenges, we propose a novel algorithm for optimal die placement on the interposer.

1 Introduction

Three-dimensional integrated circuits (3D ICs) are a promising option to overcome the manufacturing limits of transistor scaling [1] and to meet future chip performance, functionality and power consumption requirements [2–7]. In contrast to classical ICs, where all gates are located on one horizontal layer, gates in 3D ICs are separated between several vertically stacked dies (Fig. 1). Dense packaging of gates in three dimensions along with short vertical interconnects enable shorter total wirelengths and smaller footprint areas for such systems. These, in turn, improve performance and reduce power consumption.

The commercial application of 3D ICs, however, has fallen short of expectations. While a significant applicability growth of 3D ICs has been anticipated in the last few years, only a few products[1] are commercially available at this time [8, 9]. One of the reasons are the higher manufacturing costs over 2D solutions due to the additional works involved, such as TSV manufacturing, backside metallization, stacking, bonding, and testing. The high capital outlay coupled with moderate performance benefits is another deterrent—not to mention the slow pickup of 3D systems in industry due to design challenges. Expanding into the third dimension makes design tasks more complex as it affects not only the system architecture but all physical design stages.

Hence, different design constraints need to be optimized and multiple 3D design goals achieved. Classical design tasks, like floorplanning, placement, routing, etc., must be upgraded; and several novel, 3D-specific challenges addressed to ensure a smooth design flow, ultimately, leading to sound practical 3D systems. Among these challenges are TSV planning [10], die arrangements on the interposer, and advanced thermal management [11, 12].
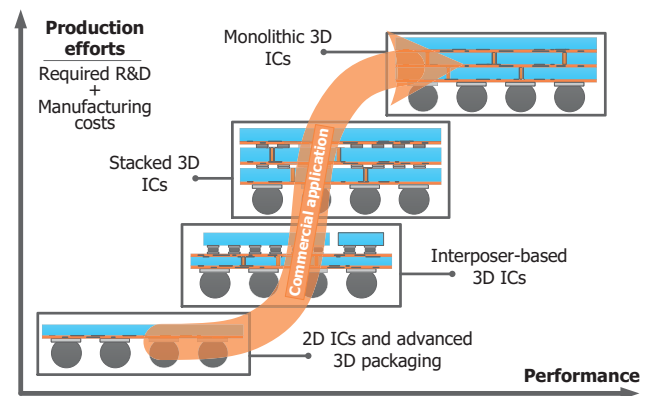

Figure 1 Different 3D technologies arranged according to manufacturing costs and expected performance benefits. The orange arrow indicates the expected commercial market entry.

In this paper, we compare various 3D design options with the purpose of identifying the most promising approach for a cost-efficient commercial application: interposer-based 3D circuits. Then we investigate the challenges facing the design community posed by these integrated circuits. We also outline solutions to some of these challenges, notably the *optimal* die placement on an interposer.

Our contributions in this paper can be summarized as follows:

1. We investigate the variety of available 3D stacking technologies and show that interposer-based systems are one of the most economically viable candidates for next-generation 3D chip design.

2. We outline the physical design challenges for interposer-based 3D ICs, formulate the major optimization objectives and elaborate on innovate ways of addressing these challenges.
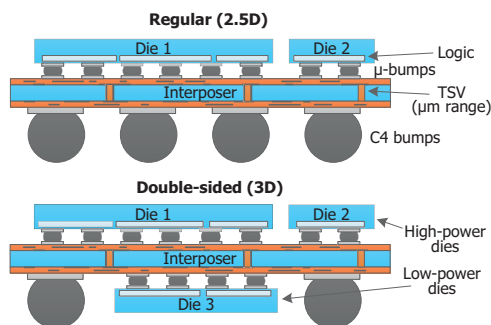
---

[1] AMD's Fury GPU [8] and the Virtex-7 FPGA from Xilinx [9]—both interposer-based 3D systems—are examples of such products.

Figure 2 Variants of interposer-based 3D ICs: (top) one-sided mounting of dies (2.5D integration); (bottom) double-sided placement (3D integration).

3. We propose an optimal approach to solve one of the major challenges posed by these systems—that is, the optimization of die placement on the interposer.

4. We prognosticate advanced design challenges that should be investigated by the EDA community in order to support the industrial application of interposer-based 3D systems.

2 Survey of 3D design alternatives

3D technologies can be classified according to integration level: package-level (e.g., package-on-package), wafer-level (e.g., wafer-level chip-scale package) and silicon-level (e.g., 2.5D and true 3D ICs). In this work, we will skip package- and wafer-level 3D technologies as they have been quite exhaustively investigated and widely applied in current chip production.

Silicon-level 3D integration, however, is still lagging in its commercial application. It can be divided into three design variants: (*i*) interposer-based 2.5D and 3D ICs, (*ii*) stacked 3D ICs and (*iii*) monolithic 3D ICs (see Fig. 1). The last two are often referred to as *true 3D ICs* [13].

Interposer-based 3D IC is a configuration where dies are mounted one-sided or double-sided on a thin silicon carrier, the so-called interposer (Fig. 2). Metallization layers on both interposer surfaces and through-silicon vias (TSVs) between them enable interconnectivity between dies and to the package. Dies are attached to the interposer using μ-bumps (~10 μm in diameter) instead of the flip-chip bumps (~100 μm in diameter) used in package-level integration, affording high interconnection and integration densities. Thus, interposer-based systems significantly boost performance while offering better heat dissipation than true 3D ICs (stacked and monolithic 3D ICs) [14]. In contrast to true 3D ICs, TSVs are restricted to the interposer (i.e., the dies have no TSVs). This means existing die designs can be used, rendering the interposer the most promising integration platform for heterogeneous design [15]. There are yield advantages, as well, as it is more profitable to manufacture and test several separate dies as opposed to a single large 2D chip [4, 6].

Despite having been first developed as a bridge technology to true 3D ICs, interposer-based integration has matured to become a cost-efficient 3D option with further growth expected over the next few years (Sec. 3 investigates the reasons for this). However, the lack of physical design tools has blocked this technology from going mainstream (Sec. 4 discusses the major challenges) [4, 16, 17].

Stacked 3D ICs refer to an integration option where several dies are placed on top of each other. Each chip here is manufactured separately (including logic and metallization layers) followed by stacking together and bonding (Fig. 3–top). While redistribution layers (RDLs) enable intra-die connections, dies employ TSVs to allow vertical communication. Depending on the bonding technology, TSVs can either be inserted into the wafers before bonding or implemented in the stack after bonding.

Stacked 3D ICs offer higher integration levels and shorter wirelengths compared to interposer systems, but have worse heat dissipation and, hence, require advanced stack-level thermal management. Die designs for this type of integration should be developed exclusively for each specific system due to, among other things, the required co-design of interconnect pads, TSV and deadspace arrangements in active and metal layers, and mechanical stress management (keep-out zones). Pre-manufactured chip designs are therefore unsuited for use in this context. That said, heterogeneous integration is supported for some bonding scenarios, and allows the use of dies manufactured with different technological processes and materials.

In addition to the aforementioned thermal management issues, several other aspects should be addressed in order to facilitate the commercial application of stacked 3D ICs. First, new technological-oriented constraints, such as the handling of thermomechanical stress induced by TSVs in active layers, must be considered during physical design [18]. Second, novel system-level and physical design challenges need to be solved. For instance, the initial choice of the die count and bonding option as well as cell-level partitioning between dies significantly impacts the complexity, functionality and manufacturability of the final system. We would like to point out that existing 2D-oriented tools for classical design steps such as floorplanning, placement, routing and timing analysis are inadequate when it comes to stacked 3D ICs; hence, they need to be significantly upgraded or completely different approaches need to be adopted. Additionally, some novel challenges arise from the use of TSVs: their locations and configurations should be planned at the die and stack levels [10]. While many academic studies, such as [6, 7, 12], tackle these design challenges, the commercial acceptance of stacked 3D IC is still low.
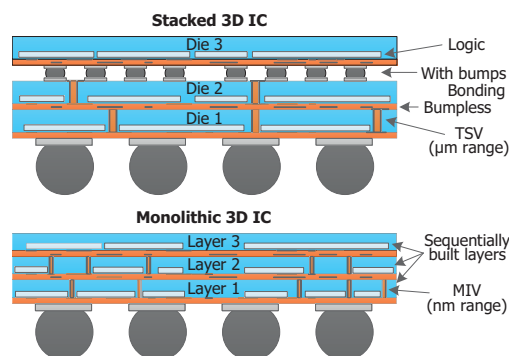

Figure 3 Variants of so-called true 3D ICs: (top) stacked 3D IC; (bottom) monolithic 3D IC.

Monolithic 3D ICs offer the highest transistor and interconnect density of all silicon 3D integration styles. Rather than stacking dies, monolithic 3D ICs encompass a single silicon layer which is further sequentially covered with active and metal layers using conventional manufacturing processes (Fig. 3–bottom). Vertical interconnects take the form of monolithic inter-tier vias (MIVs) in the nanometer scale (the TSVs mentioned earlier are on the order of micrometers) [19]. Together with very tightly stacked layers, this favors a very dense integration. Monolithic 3D ICs are beset by yield issues and heterogeneous integration is not an option with them because the entire IC is produced in a single fab process.

Since monolithic 3D integration is a relatively young technology and is still under development, several technological issues must be solved prior to its launch in the marketplace. While conceptually similar to stacked 3D ICs, monolithic integration is the outcome of drastically different manufacturing processes and has, as well, vastly different parameters, such as via arrangements, thermal management, and routing resources. Thus, (physical) design issues require more focused attention and algorithms need to be modified: (*i*) thermal properties remain an urgent problem but differ from those of stacked 3D ICs due to different via sizes and altered inter-layer thermal conductivity; (*ii*) routing is more complex due to tight placement and increased congestion; (*iii*) floorplanning and placement require the consideration of the various active layers and their interdependencies. Conventional 2D tools can however be adapted for classical challenges like placement and routing within one active layer [20, 21]. The accepted view is that monolithic 3D ICs will not become commercially available for several years to come.

3 Advantages of interposer-based 2.5/3D ICs

Next, we examine multiple manufacturing and design aspects which are crucial for the commercial applicability of 3D technologies: integration density (performance), costs, thermal properties, yield and reuse of existing designs.

Interposer-based systems have the lowest **integration density** among the previously discussed 3D options; obviously, true 3D ICs allow more compact cell packing and, thus, better performance. However, the performance gap between interposer-based and stacked 3D ICs is relatively moderate [14]. The performance benefits of true 3D over interposer-based ICs have not emerged—even with short vertical interconnects, true 3D ICs require significant resources in active and routing layers for vias. This may compromise any performance gain, make routing more complex and affect manufacturing costs. On the other hand, interposer integration does not require TSVs in active layers and offers high interconnect density and additional routing resources[2].

There are two basic **cost factors** that define the practicability of the particular 3D technology: (*i*) actual manufacturing costs (including wafer, TSV formation, bonding costs, etc.) and (*ii*) R&D expenditure (design tools, development time, fab line modification, etc.).

As for **manufacturing costs**, both interposer and stacked 3D ICs are considered as cost-effective solutions [3, 14, 22], but it is not clear which one is the most economical. Interposer wafer and TSV implementation generate extra costs in the case of interposer-based ICs [14][3]. Conversely, stacked 3D ICs have TSVs in logic layers, thus, the total TSV count and die areas (both directly impact manufacturing costs) are higher than for an interposer architecture. TSVs in active layers also decrease die yield due to possible defects during their production [22]. Moreover, there are hidden yield losses for stacked 3D ICs due to additional technological steps such as back-side metallization, die thinning, bonding and debonding between a supporting wafer and active dies [23].

Monolithic ICs have higher manufacturing costs due to their sequential fabrication process on a single semiconductor wafer. Testing of each separate layer is becoming more complex. There are also yield disadvantages because of the increased number of processing steps and the fact that possible defects in any layer will render the entire IC useless.

Interposer-based ICs have the lowest barrier to entry in the industry for **R&D costs**: the fabrication process is well understood; sufficiently diverse prefabricated chip designs are available; and thermomechanical management does not require increased attention (in contrast to true 3D ICs). As a consequence, there are several commercial examples available on the market [8, 9]. Interposer-based 3D systems are designed nowadays using conventional 2D tools [9, 15], as well as some degree of manual input [4, 16]. True 3D integration requires more R&D: some process steps still need (better) design solutions, like thin wafer handling, thermal management, and test; design tools for planning, implementation and verification must be improved and upgraded, as well.

Combining manufacturing and R&D costs, we can conclude that interposer-based systems are more affordable and cost-preferable than the other 3D stacking options presented above.

Interposer-based 3D ICs are the absolute favorite among all stacking scenarios in terms of their **thermal properties**. While true 3D systems are known for their demanding thermal management requirements due to their very high power density and problematic heat transfer from deep-covered active layers [24, 25], the interposer keeps thermal properties at a level similar to that of 2D ICs [14]. The interposer concept grants power-critical dies, such as CPUs, DSPs and other high-loaded computational logic, easier access to a heat sink, as the sink can be mounted directly on the die surfaces. A conventional passive interposer can itself serve as a supplementary heat spreader,

---

[2]The performance difference between interposer and true 3D ICs depends on system complexity, specification and especially on the implementation (number of dies, kind of system—CPUs, sensors, memory, etc.) and cannot be determined by abstracting from the particular system.

[3]Please note that there is a conceptual mistake in reference [14]—the authors increase the area of the active dies to cater for TSVs in 2.5D integration in their cost models, however, TSVs are implemented only in the interposer. The argument that interposers give rise to additional manufacturing costs is correct.

99

thus supporting better heat transfer. Moreover, integrated cooling solutions, such as embedded fluidic microchannels, are much easier implemented into a passive silicon carrier than into active dies [26].

Besides better heat transfer, **yields** are higher with interposer-based systems than with other alternatives. As there are no TSVs in the active dies, designs can be tested and verified, thus boosting the yield and eliminating risks associated with TSV-induced defects in logic dies. Interposer-based solutions have no yield disadvantages due to bonding, die thinning, and back-side metallization as against stacked 3D ICs. As partially explained in the paragraph about costs, monolithic designs deliver the lowest yield since the entire manufacturing process takes place on a single silicon substrate. A defect occurring in a late process step, while growing a new layer, for example, would completely destroy the IC.

Existing dies and designs can be reused only with the interposer approach. This apparently improves yield and cuts R&D costs; development times are shorter; existing dies from different vendors can be integrated in the IC; and off-the-shelf and newly designed dies can be combined into one system. We will discuss in depth the reuse of existing dies and heterogeneous designs in Sec. 3.1.

Hence, interposer designs promise a range of benefits: a high level of integration, shorter wirelengths and improved performance, matching true 3D ICs; they also offer several further advantages, such as higher yield, more routing capabilities, the reuse of existing dies, better heat dissipation, and lower production efforts. All in all, interposer-based 3D ICs are the most promising option for large-scale 3D integration and will most likely become the favorite 2.5/3D integration solution for the next-generation of chip design.

3.1 Heterogeneous integration

An important feature of interposer design is that it supports heterogeneous integration (Fig. 4): this refers to the integration of separately manufactured components, such as bio/imaging/environmental sensors, MEMS, RF and optical transducers, processors and memory dies. One of the major driving forces behind heterogeneous integration is IoT, which requires multifunctional, compact devices with high performance and low energy consumption. The IoT concept may require up to ten chips to be integrated into one assembly. Note that these chips are in many cases provided by different vendors (often IP protected) and produced with different technologies.

Integrating dies produced with different technologies and materials is not the only purpose of heterogeneous integration: dies fabricated with conventional semiconductor manufacturing processes under different technology nodes should also be capable of integration. All modules do not have to be fabricated with advanced technology nodes: only performance-critical parts need to be produced with costly and effortful advanced nodes (7–10nm); other parts can be produced with proved conventional nodes (>45nm), profiting from lower costs and higher yields.

Summarizing, interposer-based 3D ICs (*i*) are considered an efficient integrator for heterogeneous designs, and (*ii*) they will typically carry up to 10 dies according to the IoT paradigm.
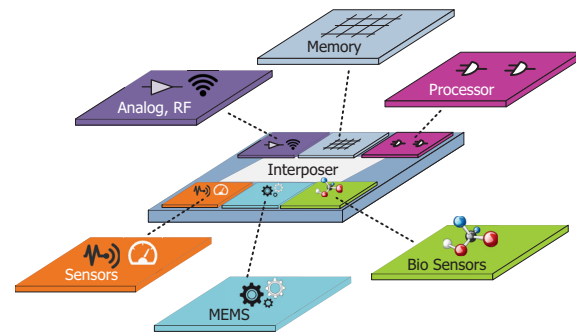

Figure 4 The interposer architecture supports heterogeneous integration: dies produced with different technologies and base materials can be integrated into one system.

3.2 Interposer system implementations

The interposer concept is for a versatile chip integrator that encompasses a range of system designs; they can be categorized as follows:

- Carrier material: silicon, glass [27] or organic substrates [28].
- Interconnect technology: electrical (TSVs), optical [29] or optofluidical [30].
- Type of interposer: fully passive, with active components, or with embedded structures such as microfluidic channels [26] and optical transmitters/receivers [31].
- Mounting technique: one/double-sided placement (see Fig. 2), distributed high/low-power die allocation [32].
- Chip design: integration of prefabricated heterogeneous dies (e.g., AMD Fury [8]) or partitioning of a homogeneous die into several smaller dies (e.g., Xilinx Virtex-7 [9]).

4 Physical design challenges and solutions for interposer-based ICs

The physical design of interposer-based 3D ICs occurs on two conceptually different levels: (*i*) the die-level and (*ii*) the interposer-level. Usually, each level is designed independently: dies are designed and manufactured separately and then mounted on a silicon interposer. While die-level design challenges (classical 2D architecture) can be solved using conventional EDA tools, the physical design on the interposer-level involves addressing novel, technology-related challenges, such as die placement on the interposer, interposer routability estimation, pin and TSV assignment, and system-level thermal management. Designing interposer-based ICs currently require some notable manual intervention for these challenges. This motivates our investigation towards promising and effective solutions.

In the following, we will discuss major interposer-level challenges and their solutions, neglecting die-level design issues, as they have been exhaustively investigated and can be solved using available CAD tools.

4.1 Die placement on an interposer

When integrating several dies on an interposer, the die arrangement should be optimized to shorten system-level interconnects. Excessively long wires on the interposer significantly increase interconnect length and power consumption, and they reduce bandwidth. Die placement may also impact interposer routability and thermal properties.

The peculiarity of interposer placement is that a small number of rotatable chips with many external pins should be placed on the interposer with the shortest interconnect length (in contrast to gate placement: many small standard cells with few pins). This challenge resembles classical floorplanning. The difference is that floorplanning often deals with "soft" blocks and optimizes the area/shape of the resulting layout, as well, while in interposer placement, dies and interposer are mostly pre-designed and/or have fixed sizes/shapes.

4.1.1 Prior art of interposer placement

Current research into interposer placement is typically focused on randomized algorithms such as *simulated annealing* [33, 34], often supporting flexible pin assignment [33–35]. Liu et al. in [35] run an *enumerative search* to obtain block positions before calling a pin assignment routine. Despite the fact that the authors develop various heuristics for efficient computation, their approach does not scale well for more than eight dies (six for optimal placement). Mao et al. in [36] propose an algorithm for placing FPGA modules/dies on the interposer based on a B*-tree representation and force-directed placement.

As for optimal die placement, the related challenge of *optimal floorplanning* has been addressed in [37] and [38]. However, these studies were either limited to six modules [37] or the rotation of modules was neglected [38], thus not assuring minimal interconnect length.

4.1.2 Optimal die placement

As interposers typically accommodate only a few dies (recall Sec. 3.1), their optimal arrangement is accessible. The problem of optimally placing dies on the interposer is *NP-hard* [4]. This means that with increasing die count, the solution space for this problem will grow dramatically (also referred to as combinatorial explosion). Since existing approaches can find an optimal placement only for instances with up to six dies due to NP-hardness, we propose a placement framework to scale up this limit.

Two basic components are required to place dies optimally on an interposer: (*i*) a mathematical representation of the layout, and (*ii*) an optimal solution-search algorithm. We extended and upgraded the *constraint-satisfaction problem (CSP)* formalism from [39] with die rotations to handle the first component. Based upon this layout encoding, we propose a highly efficient branch-and-bound (B&B) method with several speed-up techniques to search for an optimal placement. Key features of the latter are early identification and discarding of unpromising configurations (this is crucial to evade the combinatorial explosion). In practice, our proposed B&B algorithm can dramatically reduce the search time over previous propos-
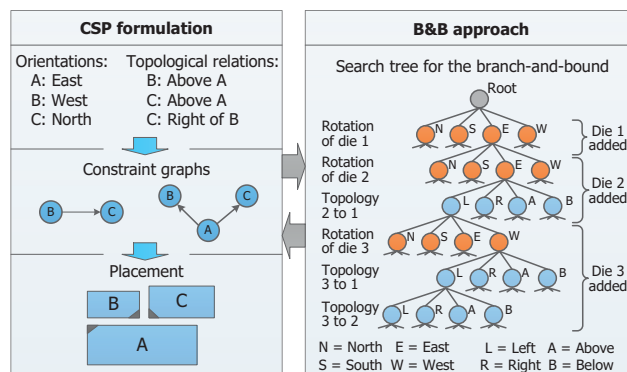

Figure 5 Two main components of the proposed placement framework: (*i*) the constraint-satisfaction problem formalism serves for layout representation, and (*ii*) the search for an optimal placement is conducted using the branch-and-bound method.

als and is capable of optimally placing up to eleven dies, while state-of-the-art tools top off at six [35, 37].

Layout representation. To encode die placement as a CSP, we define two properties: the orientation of each die and topological relations between each pair of dies (Fig. 5–left). The *orientation* for each die can take one of four values: north (0°), west (90°), south (180°) and east (270°). The *topological relation* for each pair of dies has four possible cases: die_1 is left of/right of/above/below die_2. CSP representations can be converted then into an actual placement by constructing the horizontal and vertical constraint graphs, and tracing their directed paths (see Fig. 5–left).

Placement approach. To address the combinatorial explosion, we resort to a branch-and-bound method. During the B&B, we incrementally construct placement by adding dies one at a time to the search tree, as shown in Fig. 5–right. After assigning a die, we sequentially select the orientation of this die (orange nodes in Fig. 5–right) and its topological relations to other dies (blue nodes in Fig. 5–right). Thus, we form partial configurations (placements with an incomplete number of dies, e.g., five out of eight) that could be estimated and discarded if they are unpromising, i.e., they cannot lead to optimal solutions. To evaluate whether a partial configuration is promising, we check the lower bound wirelength estimation of the partial configuration against the wirelength of the current best known solution. If the former is higher, then this partial configuration cannot produce a better solution than the best one found so far and can be discarded.

Speed-up techniques. The deficiencies mentioned above and associated with state-of-the-art methods for optimally placing up to six dies [35, 37] may be overcame by the B&B approach outlined earlier by adding several accelerating techniques, such as: (*i*) an appropriate branching schedule [37, 38], where dies are added to the search tree according to their impact; (*ii*) quickly finding an initial good-quality solution; (*iii*) checking in advance whether the assignment of a new node in the search tree will make the partial configuration unpromising in future, namely forward checking [39]; (*iv*) estimating how dies yet unassigned to the search engine will increase the wirelength; and (*v*) dominance techniques to rule out some unpromising configurations which appear much worse than others.

---

[4]NP-hardness can be derived from the *rectangle packing problem* [39] by accounting for the interconnects between dies.

4.2 Interposer routing and routability

The interposer serves primarily as an interconnect platform for system-level chip integration. Routing congestion is likely to occur due to the high number of external die connections and limited routing resources on the interposer. The congestion, in turn, may cause the interposer to be unroutable or increase the total wirelength due to the detours required (likely with timing degradation). Hence, interposer systems require an estimation of whether any particular die placement is routable; a placement is then adjusted based on this information.

When prefabricated dies with fixed pin positions are integrated, the routing of the passive interposer is not fundamentally different from the routing of individual dies and can be done with conventional tools or algorithms[5]. Thus, current research on interposer routing often tackles various technology-specific aspects.

A global routing algorithm for systems-on-a-package (SOPs) was examined in [41] and can be applied to interposer systems for routing or routability estimation purposes, as well. An IR-aware routing for the interposer and redistribution layers (RDLs) of each die, along with simultaneous micro-bumps planning/signal assignment was studied in [42]. The proposed approach defines a requisite number of micro-bumps for each chip, assigns them to I/O buffers and routes the RDLs/interposer. The authors of [43] suggest early estimation and minimization of required metal layers for interposer-based ICs. Their algorithm is based on a routability estimation which then derives the minimum number of metal layers.

In summary, a fast routability estimation algorithm for the interposer should be part of the suite of placement tools. This prevents, among others things, a placement approach finding a global TWL optimum that subsequently turns out to be unroutable or requires excessive detours.

4.3 Pin and TSV assignment

Dies may change their relative positions and orientations during the placement procedure. This affects the total interconnect length due to varying paths from I/O buffers in logic layers to escape points on the interposer. To facilitate a somewhat optimized bump/TSV/pin assignment, interposer placement algorithms often include pin assignment as a post-placement routine.

For example, a network-flow algorithm is utilized in [35] to establish the connections between I/O buffers and micro bumps. The authors of [34] use an integer linear programming (ILP) formulation for the same purpose; bipartite matching is deployed in [33]. Alternatively, pin assignment can be integrated into the design flow as a pre-routing process. Fang et al. in [42] assign the I/O buffers within dies to the placed bumps before starting with RDLs/interposer routing. Their approach covers IR-drop constraints, as well.

The related issue of TSV planning in the interposer has been poorly addressed so far. Current studies that aim for routing and TSV assignment often neglect TSV-placement optimization. The authors of [35], for example, utilize a uniformly pre-placed set of TSVs in the interposer. This problem requires further research in order to (*i*) achieve shorter total WL by signal-oriented planning of the TSVs; (*ii*) improve thermal properties by inserting thermal TSVs; and (*iii*) minimize electrical coupling and thermomechanical stress (induced by TSVs) using proper simulation.

4.4 Planning optical interconnects

Along with electrical connections, modern interposer designs may incorporate optical interconnect technologies, serving as a fast system-level communication between dies [44]. This optical infrastructure requires a variety of photonics components, such as waveguides, couplers, switchers, optical TSVs and various transceivers, to be integrated on the interposer. Hence, additional design effort is needed for the simultaneous placement and routing of optical and electrical elements.

Photonic structures are very different to conventional electrical (metal) interconnects [45, 46]. Design approaches and optimization objectives for photonic structures consequently differ very greatly from common place&route algorithms. E.g., the routing geometry of the waveguide plays a greater role (requiring a low number of crossings and bends), and both placement and routing steps aim to minimize the total signal loss instead of shortening the waveguide length. Although these systems are still manually designed, the process will have to be automated at some stage to meet the increasing level of integration of optical components. This therefore calls for novel algorithms to automate the design of optical structures on the interposer.

Several studies have tackled the placement and routing problems of photonics components [46–49]. The placement of switching elements (aka routers) for 3D Networks-on-Chip (NoCs) was investigated in [47]. The proposed method uses nonlinear programming to minimize the signal loss. A scalable algorithm based on the force-directed approach was proposed in [48] for the same purpose. Studies [46, 49] focus on the routing of waveguides, assuming that all optical components have been placed; both works aim to minimize the signal loss. While the former optimizes on-chip optical interconnects using ILP formulation, the latter deploys the rip-up-and-reroute algorithm with several additional techniques for SOPs.

Although all proposed techniques may be used or may, at least, serve as guideline approaches for interposer-based systems, there still is a dearth of dedicated interposer tools. The interposer specifics, such as double-sided waveguide routing, the ability to build transceivers directly on the carrier and the use of electrical TSVs for optical transmission [29] should be taken into account when developing algorithms for optical design. Another key aspect, which has been neglected in previous works, is the co-optimization of electrical and optical elements based on their high codependency.

4.5 Thermal management

Obviously, an interposer architecture provides generally better heat transfer and more flexibility for thermal management than true 3D ICs. However, wrongly partitioned

---

[5]Note that the integration of large-scale, and complex, network-on-chips using active interposer may require further research on routing [40].

designs or inattentive placement of high-power dies can create local hotspots. Hence, interposer designs require strict thermal management in order to control thermomechanical stability and reliability constraints.

Although several thermal-aware floorplanning and placement algorithms for 3D ICs have been published, dedicated solutions for interposer systems are thin on the ground. The first promising thermal models and thermal-aware design methodologies are presented in [32, 50–52]. Most of them [32, 51, 52] employ an accurate, yet time-consuming finite element method (FEM), only the authors of [50] adopt a coarse thermal-network modelling approach for speed-up reasons. However, they are all stand-alone, separately developed thermal models. To be useful for interposer design, they need to be adapted and integrated into the early stages of the physical design flow, such as floorplanning/partitioning or die placement.

To facilitate the integration of thermal management into the design flow, interposer systems, additionally, need appropriate data structures for their specific requirements. Typical requirements are efficient and fast data transfer from representations into actual geometry and into the applied thermal model; storing easily accessible information about physical and mechanical stress properties; as well as the consideration of thermal heat transfer characteristics.

5 Advanced challenges of chip/interposer co-design

While the current design challenges of interposer-based ICs can be partially addressed using conventional CAD tools and methodologies (see Sec. 4), there are still several advanced, system-level tasks requiring chip/interposer co-design. This co-design combines both die-level and interposer-level design challenges with the goal of a multi-objective optimization of the entire 3D system.

5.1 Simultaneous chip-interposer design

The ultimate desire is a simultaneous design of dies and interposer within the same flow, targeting the optimization of key system parameters, like wirelength, timing, routability, and thermomechanical stability. Only such a global approach can guarantee best possible matching between all system parts concerning local (gate) and global (die) characteristics, such as placement and thermal properties.

Excluding the case when all dies are pre-designed (and only interposer-level design tasks remain), there are two basic concepts of chip-interposer co-design: (*i*) when many dies have been prefabricated and only one or a few more dies have still to be designed, and (*ii*) when the entire system has to be designed from scratch. The former concept calls for a classical design approach—where die manufacturing is followed by die placement on the interposer—with additional constraints resulting from the prefabricated dies. The latter concept is more complex as it requires different design tasks and constraints to be handled simultaneously. Proper partitioning, finding the trade-off for the die count, considering all technology-related constraints, and a cost-related choice of an appropriate technology node are the key considerations here.
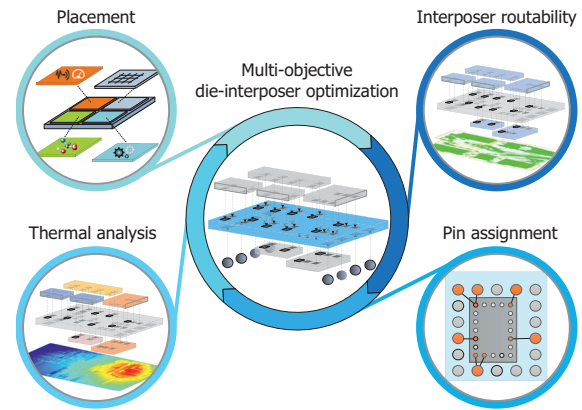

Figure 6 Multi-objective optimization: simultaneous placement, pin assignment, interposer routability estimation, and thermal modeling.

First attempts to simultaneously design chip and interposer have been published in [33–35], combining die placement (interposer-level) and pin assignment (die-level).

5.2 Multi-objective optimization

Physical simulations and additional optimization goals need to be applied during the early stages of 3D physical design. Ideally, thermomechanical simulations, pin assignment, power domain clustering as well as routability estimation should be done during the (early) floorplanning and placement stages of the design flow (Fig. 6). To do so, fast and accurate models need to be developed and associated with the design flow. Please refer to [10–12] for nascent approaches towards this idea.

5.3 Double-sided interposer placement

As mentioned in Sec. 3.2, dies can be mounted not only on the top side of the interposer (as is done in 2.5D ICs), but they can be also placed on both sides, benefiting from shorter interconnects and more effective area utilization. As a general rule of thumb for this type of placement operation, dies are allocated to a particular side according to their power consumption: high-power dies are arranged on the top side, closer to the heat sink, while low-power ones are placed on the bottom side.

The algorithms that consider both interposer sides during placement optimization do not exist yet and need to be developed. The framework proposed in Sec. 4.1.2 could serve as a starting point for such placement optimization.

5.4 System-level data structures

Generally speaking, data structures are an abstract model of a design problem. We need novel, system-level data structures to model interposer-based ICs. They must not only encompass layout characteristics, but also be able to transform efficiently the representation into various simulation environments; they must also contain the physical characteristics of the entire system, consider a multitude of constraints and interact smoothly with design flows. An overview of several data structures for 3D ICs, which includes candidate solutions for interposer-based systems, has been presented in [53].

6 Summary

To fully exploit the advantages of the extra dimension in 3D integrated circuits, layout designers and tool developers need to be aware of the major design options and challenges resulting from this new technology.

First, we introduced several criteria to evaluate the advantages and commercial applicability of different, currently available 3D technologies. Our review showed that interposer-based systems are the most promising option for today's heterogeneous 3D integration: despite their slightly lower integration level than other 3D technologies, interposer-based systems offer higher yield, more routing capabilities, the reuse of existing dies, better heat dissipation, and lower production efforts.

Die placement on the interposer, pin and TSV assignment, interposer routing/routability estimation, and system-level thermal management are some of the novel, technology-related challenges that need to be addressed today by physical design at the interposer-level. We proposed an optimal placement algorithm to solve the particular challenge of die placement on the interposer. Our layout representation as a constraint-satisfaction problem (CSP) and optimization strategy of using a smart branch-and-bound (B&B) method can be leveraged to other 3D design challenges as well.

Finally, we outlined advanced, system-level design initiatives that pursue the goal of simultaneous die/interposer co-design. These challenges should be further investigated by the EDA community in order to support the commercial application of interposer-based 3D systems.

7 Literature

[1] A. B. Kahng, "Lithography-induced limits to scaling of design quality," in *SPIE Adv. Lithography*, vol. 9053, 2014.

[2] R. S. Patti, "Three-dimensional integrated circuits and the future of system-on-chip designs," *Proc. IEEE*, vol. 94, no. 6, pp. 1214–1224, 2006.

[3] J. H. Lau, "The most cost-effective integrator (TSV interposer) for 3D IC integration system-in-package (SiP)," in *Proc. ASME InterPACK*, 2011, pp. 53–63.

[4] A. Kannan *et al.*, "Enabling interposer-based disintegration of multi-core processors," in *Proc. Int. Symp. Microarch.*, 2015, pp. 546–558.

[5] P. Batude *et al.*, "Advances, challenges and opportunities in 3D CMOS sequential integration," in *Proc. Int. Elec. Devices Meeting*, 2011, pp. 7.3.1–7.3.4.

[6] D. Stow *et al.*, "Cost analysis and cost-driven IP reuse methodology for SoC design based on 2.5D/3D integration," in *Proc. Int. Conf. Comp.-Aided Des.*, 2016, pp. 56:1–56:6.

[7] T. Lu *et al.*, "TSV-based 3D ICs: Design methods and tools," *Trans. Comp.-Aided Des. Integ. Circ. Sys.*, vol. PP, no. 99, pp. 1–1, 2017.

[8] J. Macri, "AMD's next generation GPU and high bandwidth memory architecture: FURY," in *Hot Chips Symp.*, 2015, pp. 1–26.

[9] P. Dorsey, "Xilinx stacked silicon interconnect technology delivers breakthrough FPGA capacity, bandwidth, and power efficiency," Xilinc, Inc., Tech. Rep., 2010.

[10] J. Knechtel *et al.*, "Planning massive interconnects in 3-D chips," *Trans. Comp.-Aided Des. Integ. Circ. Sys.*, vol. 34, no. 11, pp. 1808–1821, 2015.

[11] J. Cong and Y. Ma, "Thermal-aware 3D floorplan," in *Three Dimensional Integrated Circuit Design*, 2010, ch. 4, pp. 63–102.

[12] S.-K. Ryu *et al.*, "Impact of near-surface thermal stresses on interfacial reliability of through-silicon vias for 3-D interconnects," *Proc. Dev. and Mater. Reliab.*, vol. 11, no. 1, pp. 35–43, 2011.

[13] R. Fischbach *et al.*, "From 3D circuit technologies and data structures to interconnect prediction," in *Proc. Int. Workshop Sys.-Level Interconn. Pred.*, 2009, pp. 77–84.

[14] C. Zhang and G. Sun, "Fabrication cost analysis for 2D, 2.5D, and 3D IC designs," in *Proc. 3D Sys. Integ. Conf.*, 2012, pp. 1–4.

[15] J. Knechtel *et al.*, "Large-scale 3D chips: Challenges and solutions for design automation, testing, and trustworthy integration," *IPSJ Transactions on System LSI Design Methodology*, vol. 10, pp. 45–62, 2017.

[16] D. Milojevic *et al.*, "Design issues in heterogeneous 3D/2.5D integration," in *Proc. Asia South Pac. Des. Autom. Conf.*, 2013, pp. 403–410.

[17] B. Martin *et al.*, "A path finding based SI design methodology for 3D integration," in *Proc. Elec. Compon. Tech. Conf.*, 2014, pp. 2124–2130.

[18] K. H. Lu *et al.*, "Thermo-mechanical reliability of 3-D ICs containing through silicon vias," in *Proc. Elec. Compon. Tech. Conf.*, 2009, pp. 630–634.

[19] S. Panth *et al.*, "High-density integration of functional modules using monolithic 3D-IC technology," in *Proc. Asia South Pac. Des. Autom. Conf.*, 2013, pp. 681–686.

[20] S. Panth *et al.*, "Shrunk-2D: A physical design methodology to build commercial-quality monolithic 3D ICs," *Trans. Comp.-Aided Des. Integ. Circ. Sys.*, no. 99, pp. 1–1, 2017.

[21] K. Chang *et al.*, "Cascade2D: A design-aware partitioning approach to monolithic 3D IC with 2D commercial tools," in *Proc. Int. Conf. Comp.-Aided Des.*, 2016, pp. 130:1–130:8.

[22] X. Dong *et al.*, "Fabrication cost analysis and cost-aware design space exploration for 3-D ICs," *Trans. Comput.-Aided Des. Integr. Circuits Sys.*, vol. 29, no. 12, pp. 1959–1972, 2010.

[23] J. Lau, "TSV manufacturing yield and hidden costs for 3D IC integration," in *Proc. Elec. Compon. Technol. Conf.*, 2010, pp. 1031–1042.

[24] P. Jain *et al.*, "Thermal and power delivery challenges in 3D ICs," in *Three Dimensional Integrated Circuit Design*, 2010, ch. 3, pp. 33–61.

[25] K. Puttaswamy and G. H. Loh, "Thermal analysis of a 3D die-stacked high-performance microprocessor," in *Proc. Great Lakes Symp. VLSI*, 2006, pp. 19–24.

[26] G. Y. Tang *et al.*, "Integrated liquid cooling systems for 3-D stacked TSV modules," vol. 33, no. 1, pp. 184–195, 2010.

[27] G. Kumar *et al.*, "Ultra-high I/O density glass/silicon interposers for high bandwidth smart mobile applications," in *Proc. Elec. Compon. Tech. Conf.*, 2011, pp. 217–223.

[28] T. G. Lenihan and E. J. Vardaman, "Challenges to consider in organic interposer hvm," in *TechSearch Intern. for iNEMI Subst. & Pack. Workshop*, 2014.

[29] S. Killge *et al.*, "Optical through-silicon vias," in *3D Stacked Chips*, 2016, pp. 221–234.

[30] M. Odeh *et al.*, "Gradient-index optofluidic waveguide in polydimethylsiloxane," *in Applied Optics*, vol. 56, no. 4, pp. 1202–1206, 2017.

[31] S. Hosseini *et al.*, "Integrated optical devices for 3D photonic transceivers," in *3D Stacked Chips*, 2016, pp. 235–253.

[32] J. H. Lau *et al.*, "Thermal-enhanced and cost-effective 3D IC integration with TSV (through-silicon via) interposers for high-performance applications," *Proc. Int. Mech. Engin. Cong. & Expos.*, pp. 137–144, Jan. 2010.

[33] Y.-K. Ho and Y.-W. Chang, "Multiple chip planning for chip-interposer codesign," in *Proc. Des. Autom. Conf.* ACM, 2013, pp. 27:1–27:6.

[34] D. Seemuth *et al.*, "Automatic die placement and flexible I/O assignment in 2.5D IC design," in *Proc. Int. Symp. Quality Elec. Des.*, 2015, pp. 524–527.

[35] W.-H. Liu *et al.*, "Floorplanning and signal assignment for silicon interposer-based 3D ICs," in *Proc. Des. Autom. Conf.*, 2014, pp. 5:1–5:6.

[36] F. Mao *et al.*, "Modular placement for interposer based multi-FPGA systems," in *Proc. Great Lakes Symp. VLSI*, 2016, pp. 93–98.

[37] H. Onodera *et al.*, "Branch-and-bound placement for building block layout," *El. and Commun. in Japan (Part III: Fund. El. Science)*, vol. 76, no. 7, pp. 15–26, 1993.

[38] J. Funke *et al.*, "An exact algorithm for wirelength optimal placements in VLSI design," *Integration*, vol. 52, pp. 355–366, 2016.

[39] R. E. Korf *et al.*, "Optimal rectangle packing," *Annals of Oper. Research*, vol. 179, no. 1, pp. 261–295, 2008.

[40] G. H. Loh *et al.*, "Interconnect-memory challenges for multi-chip, silicon interposer systems," in *Proc. MEMSYS*, 2015, pp. 3–10.

[41] J. R. Minz and S. K. Lim, "Block-level 3-D global routing with an application to 3-D packaging," *Trans. Comp.-Aided Des. Integ. Circ. Sys.*, vol. 25, no. 10, pp. 2248–2257, 2006.

[42] E. J. W. Fang *et al.*, "IR to routing challenge and solution for interposer-based design," in *Proc. Asia South Pacific Des. Autom. Conf.*, 2015, pp. 226–230.

[43] W.-H. Liu *et al.*, "Metal layer planning for silicon interposers with consideration of routability and manufacturing cost," in *Proc. Des. Autom. Test Europe*, 2014, pp. 359:1–359:6.

[44] G.-K. Chang *et al.*, "Chip-to-chip optoelectronics SOP on organic boards or packages," *Trans. on Adv. Packaging*, vol. 27, no. 2, pp. 386–397, 2004.

[45] D. A. B. Miller, "Device requirements for optical interconnects to silicon chips," *Proc. of the IEEE*, vol. 97, no. 7, pp. 1166–1185, 2009.

[46] D. Ding *et al.*, "O-Router: An optical routing framework for low power on-chip silicon nano-photonic integration," in *Proc. Des. Autom. Conf.*, 2009, pp. 264–269.

[47] A. Boos *et al.*, "PROTON: An automatic place-and-route tool for optical networks-on-chip," in *Proc. Int. Conf. Comput.-Aided Des.*, 2013, pp. 138–145.

[48] A. von Beuningen and U. Schlichtmann, "PLATON: A force-directed placement algorithm for 3D optical networks-on-chip," in *Proc. Int. Symp. Phys. Des.*, 2016, pp. 27–34.

[49] J. R. Minz *et al.*, "Optical routing for 3-D system-on-package," *Trans. Compon. Pack. Tech.*, vol. 30, no. 4, pp. 805–812, 2007.

[50] Y. Hoe *et al.*, "Effect of TSV interposer on the thermal performance of FCBGA package," in *Proc. Elec. Pack. Tech. Conf.*, 2009, pp. 778–786.

[51] A. Heinig *et al.*, "Thermal analysis and optimization of 2.5D and 3D integrated systems with wide I/O memory," in *Proc. Therm. Thermomech. Phen. Elect. Syst. Conf.*, 2014, pp. 86–91.

[52] S.-T. Wu *et al.*, "Thermal and mechanical design and analysis of 3D IC interposer with double-sided active chips," in *Electronic Components and Technology Conference (ECTC), 2013 IEEE 63rd*, 2013, pp. 1471–1479.

[53] R. Fischbach *et al.*, "Investigating modern layout representations for improved 3D design automation," in *Proc. Great Lakes Symp. VLSI*, 2011, pp. 337–342.