Toward Advancing 3D-ICs Physical Design: Challenges and Opportunities

Xueyan Zhao[1,4], Weiguo Li[2], Zhisheng Zeng[2], Zhipeng Huang[3], Biwei Xie[1,2,4], Xingquan Li[2,3,✉], Yungang Bao[1,3,4]

[1]State Key Laboratory of Processors, Institute of Computing Technology, CAS, [2]Pengcheng Laboratory, [3]Beijing Institute of Open Source Chip, [4]University of Chinese Academy of Sciences

Email:lixq01@pcl.ac.cn,xiebiwei@ict.ac.cn,baoyg@ict.ac.cn

Abstract

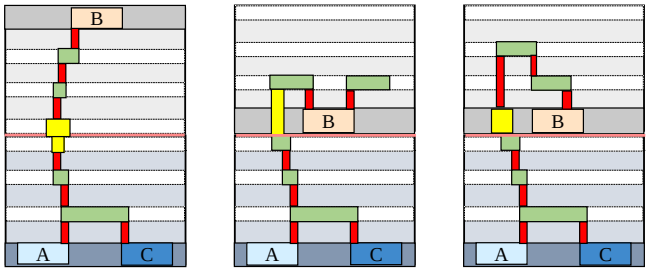
As the demand for higher integration density and performance efficiency continues to grow, 3D stacking has emerged as a promising solution. In 3D ICs, the complexity of physical design and the optimization space is significantly increasing. Therefore, researching high-quality 3D native instead of pesudo 3D physical design has become even more important. This paper reviews recent advancements and persistent challenges in 3D physical design, focusing on F2F bonding technologies. Then, this paper discusses several issues that still require further research and some overlooked problems, with the hope of helping researchers develop higher-quality 3D native physical design tools in the future.

ACM Reference Format:

Xueyan Zhao[1,4], Weiguo Li[2], Zhisheng Zeng[2], Zhipeng Huang[3], Biwei Xie[1,2,4], Xingquan Li[2,3,✉], Yungang Bao[1,3,4] . 2025. Toward Advancing 3D-ICs Physical Design: Challenges and Opportunities. In *30th Asia and South Pacific Design Automation Conference (ASP-DAC '25), January 20–23, 2025, Tokyo, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3658617.3703135

1 Introduction

Three-Dimensional Integrated Circuits (3D-ICs) have emerged as one of the most promising solutions for extending Moore's Law. In 3D-ICs, devices are distributed across multiple dies and interconnected, significantly reducing the length of interconnects within the chip, thus achieving additional performance gains. Typical application chips of 3D-ICs include Intel's Meteor Lake and AMD's Zen 4, both of which have demonstrated significant performance improvements and cost savings [45].



(a) F2F with Bondings (b) F2B with MIVs (c) F2B With TSVs

Figure 1: 3D stacking techniques for the inter-die connection.

Vertical interconnect technology plays a crucial role in the design and functionality of 3D-ICs. As 3D-ICs involve stacking multiple tiers of devices, efficient vertical connections between these tiers are essential for optimizing performance, power consumption, and data transmission speed. Currently, there are three main 3D integration technologies according to the vertical interconnection technology: Through Silicon Via (TSV) based 3D integration, Monolithic 3D integration (M3D), and Face-to-Face (F2F) hybrid bonding 3D integration.

- **T3D**: TSV-based 3D (T3D) is the most mature vertical interconnect technology in 3D IC design, where TSVs are fabricated across two or more tiers of devices. However, the relatively large pitch and parasitic characteristics limit their applicability.This is mainly used to design chip involved memory-to-logic or large logic-to-logic designs with a small number of global interconnects [38].
- **M3D**: An emerging alternative is Monolithic 3D integration (M3D), where tiers are manufactured sequentially and connected using Monolithic Inter-tier Vias (MIVs). Since wafer alignment is not required, the sizes of these MIVs are approximately the same as local vias. Overall, compared to TSV-based 3D-ICs, the small size of MIVs enables ultra-high integration density, significantly reducing silicon area and cost. However, M3D

ASPDAC '25, January 20–23, 2025, Tokyo, Japan

ACM ISBN 979-8-4007-0635-6/25/01

https://doi.org/10.1145/3658617.3703135

Table 1: Physical dimensions of inter-tier connections.

	MIVs	Bonding terminals	TSVs
Via size	$20-40nm$	$0.5-2\mu m$	$3-10\mu m$
Via pitch	$40-80nm$	$1-4\mu m$	$6-20\mu m$
Via height	$100-200nm$	$<1\mu m$	$30-100\mu m$

presents manufacturing challenges, as each subsequent tier is fabricated under strict thermal limitations to avoid damaging the existing tiers. [24]

- **F2F hybrid bonding 3D**: Face-to-Face (F2F) bonding is another approach, which bonds the two faces of chips together. F2F-bonded 3D-ICs do not require additional silicon area for 3D connections, eliminating the need to reserve empty space for vias and allowing for higher integration density. The zero silicon area overhead of F2F bonding provides significant advantages in many applications. Overall, F2F bonding terminals are much smaller and easier to manufacture than TSVs, allowing for higher interconnect density and lower costs in 3D integration [13].

As shown in Figure. 1, TSVs require significant silicon area allocation, impacting device density. MIVs reduce silicon area overhead but are constrained to homogeneous technology node integration. In contrast, hybrid bonding terminals (HBTs) enable heterogeneous technology stacking while minimizing silicon area consumption, providing a more efficient vertical integration solution. Table 1 presents the physical dimensions of these inter-tier interconnects. Additionally, MIVs faces significant yield challenges, making F2F hybrid bonding 3D integration the most effective technology until further technology advancements are made. There are some previous reviews [5, 27, 37, 60], while this paper aims to review the existing 3D physical design algorithms especially for F2F-bonded technology and explore some previously overlooked aspects to facilitate 3D-native physical design in the future.

The remainder of this paper is organized as follows. Section 2 reviews the existing 3D IC flow and discusses some of the current challenges. Section 3 presents potential future research directions for the 3D IC flow. Finally, Section 4 provides a summary of this paper.

2 Existing 3D Physical Design Solutions

Current methods for 3D-ICs can be broadly categorized into two main types: pseudo 3D flow and 3D-native flow, as illustrated in Figure 2. The remainder of this section reviews recent technologies within these two categories.

2.1 Pseudo 3D Physical Design Flow

Some of the earlier approaches to addressing 3D-IC physical design involved pseudo 3D flows, which completed 3D

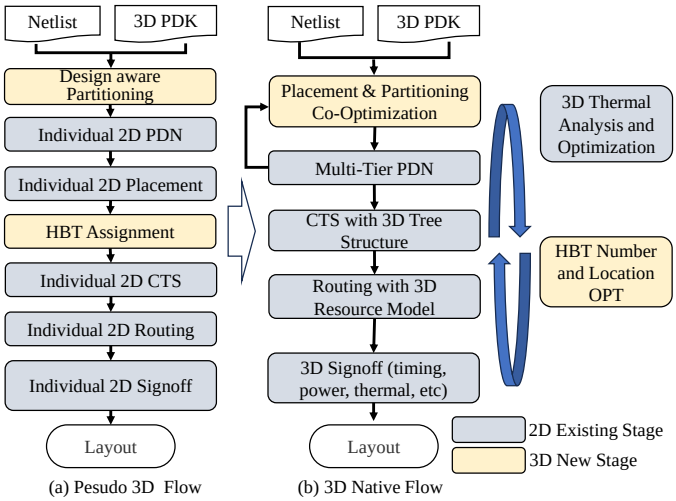


Figure 2: Pesudo 3D Flow and 3D Native Flow.

design by making certain 3D considerations and optimizations based on 2D IC tools. Methods [4, 28, 43, 48, 55] start by scaling down the standard cell sizes or chip area and then utilize commercial 2D tools to perform the cell placement. Following this, they use balanced hypergraph bipartitioning to assign cells to their respective dies. Finally, commercial 2D tools are used again for routing and the final sign-off process. Macro-3D [3] further enhances this flow by incorporating the simultaneous partitioning of macros and standard cells, demonstrating superior PPA optimization capabilities. Building on the Macro-3D approach, methods such as M3D-ADTCO [51] and Hier-3D [1] introduce considerations for SoC-level designs, including memory components.

2.2 Power Delivery Network

The design of a Power Delivery Network (PDN) is a constrained optimization problem. An optimal PDN must limit voltage drops caused by transistor switching activity, meet current density constraints imposed by electromigration limits, and use minimal metal resources to achieve design density goals [52].

Previous research efforts (e.g., [17, 20, 22, 40]) have primarily focused on TSV-based technologies, proposing various methods for constructing TSV topologies and developing IR drop analysis models. Zhu et al. [59, 61] have proposed PDN design approaches specifically targeting F2F-bonded technologies. They developed design models for F2F-bonded PDNs under specific technology assumptions and demonstrated their effectiveness in real-world applications, such as commercial CPUs.

2.3 Partitioning

Tier partitioning is also a critical step in the F2F-bonded 3D-IC design, as it determines the specific tier of each cell or macro. Previous work primarily utilized traditional heuristic

graph partitioning algorithms for tier partitioning, such as recursive partitioning [46] and bin-based min-cut partitioning with the FM partitioning algorithm in Shrunk-2D [41]. These partitioning algorithms consider basic constraints like area balance and iteratively optimize the cut size (i.e., number of 3D cross-die nets). However, they do not take into account other critical factors, such as cell coordinates, bonding terminal density, timing, or design hierarchy, which can lead to discrepancies between the partitioning results and actual design requirements. Additionally, Panth et al. [42] explored a partitioning method driven by routability using a mini-cost flow algorithm. enabling the cut of multi-pin nets during partitioning and resulting in partitioning outcomes that are more conducive to routing.

Another typical class of partitioning algorithms based on machine learning. TP-GNN [35] is a hierarchical partitioning framework based on Graph Neural Networks (GNN), which considers multiple design and technology factors, including design hierarchy and cell timing. These features are encoded into vectors and fed into the GNN for unsupervised learning. Finally, weighted K-means clustering is applied to generate the tier partitioning. Compared to the bin-based min-cut method, the TP-GNN framework can reduce the total wirelength of a 3D-IC by 7%.

2.4 Placement

3D-ICs offer significant improvements in integration density and performance but also introduce novel placement challenges. As the design space extends from 2D to 3D, die assignment for elements and the utilization of vertical interconnects have increasingly substantial impacts on subsequent design stages. Additionally, the expansion of the solution space adds complexity, making high-quality placement outcomes more challenging to achieve.

Early 3D placement research primarily extended traditional 2D models, adapting conventional 2D wirelength and density models into three-dimensional space. For example, the force-directed 3D placement method [14] and the non-linear 3D placement method [11] expanded the layout space by introducing continuous variables in the z-direction, with discretization handled during post-processing. The TSV-aware 3D placement method [18] further accounted for TSV space requirements, reserving TSV locations during global placement to reduce resource conflicts and routing complexity in later stages. Building on this, ePlace-3D [32] introduced an electrostatic model that balances density through electric field forces, achieving enhanced placement quality. Despite extensive research, existing 3D-IC placement techniques are generally unsuitable for F2F-bonded stacked 3D-IC designs with heterogeneous technology nodes [19]. To address this issue, researchers have proposed two categories of approaches:

Discrete Methods: Traditional "partition-first, then place" discrete placement techniques typically rely on methods like min-cut partitioning, which fail to fully leverage the advantages of 3D-ICs. These methods lack co-optimization, making it challenging to effectively control metrics such as wirelength and routability, and they are prone to local optima. Zhao et al. [57] proposed an alternating iterative method based on bilevel programming, which considers multiple objectives, including wirelength, cell density, and heterogeneous technology nodes, to achieve superior solutions.

Analytical Methods: Traditional analytical 3D placement algorithms use continuous optimization method but face two significant challenges in F2F-bonded 3D-ICs: a) The continuous density model in traditional 3D algorithms cannot effectively handle heterogeneous tiers; b) Previous algorithms use a simplified 3D HPWL model to smooth vertical connections, but this model is not suited to F2F-bonded 3D designs. To address these challenges, [8, 29] have proposed novel wirelength models, along with improvements to eDensity, enabling analytical methods to be effectively applied in this problem.

In addition to optimizing wirelength and reducing terminal count, recent studies consider more design constraints. For example, recent studies [9, 58] introduced targeted enhancements for 3D mixed-size placement. [47] proposes optimizing routability during bonding bumps and terminals legalization, while TA3D [23] achieves timing optimization in 3D-ICs by optimizing the critical path.

2.5 Clock Tree Synthesis

Clock tree synthesis (CTS) is a critical component in IC design, aiming to distribute the clock signal efficiently across the chip while minimizing skew, delay, and power consumption. Classic 2D clock tree designs, such as mesh [15], "fishbone" [2], and symmetric H-tree topologies, reduce clock skew and improve robustness. The GH-tree [16] extends the H-tree with a branching factor to optimize performance. Deferred-Merge Embedding (DME) and its variants [6, 10, 54] builds a topology via bottom-up merging and top-down embedding for effectively managing skew. In [7, 49, 53], buffer insertion is integrated with clock tree routing, allowing continuous delay updates for a balanced clock tree.

In 3D-ICs, CTS faces more challenges and opportunities due to the stacking of multiple dies and the direct connection of clock net pins to F2F bonding terminals. These are primarily reflected in the following aspects: a) increased complexity in timing calculations. b) convergence issues in heterogeneous designs. d) and increased difficulty in clock topology design and routability. To address these issues, current research in 3D CTS primarily focuses on the following directions:

Thermal-aware Methods: Aims to mitigate clock skew induced by temperature gradients. Utilizes thermal-sensitive algorithms for clock tree topology generation, with techniques

like symmetrical buffer insertion and grid-based thermal profiling to reduce thermal hotspots [36, 39, 50].

TSV-aware Methods: Focuses on minimizing TSV count and placement while ensuring zero-skew clock distribution across 3D tiers. Uses methods like DME and TSV-aware tier embedding to optimize TSV usage alongside wirelength and delay reduction [25, 26, 34, 56].

Power-efficient Methods: [30, 33] integrate clock gating into CTS and elaborately assign TSVs' locations to reduce dynamic power in 3D-ICs .

Fault-tolerant Methods: Incorporates redundancy mechanisms and TSV pairing algorithms to address TSV-related vulnerabilities, ensuring stable clock distribution under TSV faults [44].

Despite these advancements, current 3D CTS techniques are not fully equipped to address the unique challenges posed by F2F-bonded 3D-ICs. The cross-die clock topology in F2F designs requires innovative approaches to balance delay, skew, load, and power consumption while accommodating heterogeneous clock cell requirements. Accurate evaluation of clock cell delays and interconnect delays is crucial for convergence in clock tree synthesis. Therefore, there is a need for new methodologies that can effectively handle the complexities of CTS in F2F-bonded 3D-ICs.

2.6 Routing

Compared to 2D chip design, F2F-bonded 3D-ICs introduce an additional dimension and more routing space, which makes developing true 3D placers and routers more challenging. Current routers handle only 2D ICs, and the typical approach is to split the 3D design into individual designs for each tier, with each design being routed independently [28, 43]. This requires the positions of the bonding terminals to be known so that the authors use I/O pins to represent the bonding terminals for each tier. Once the partitioning of all cells is finalized, current bonding terminal-based placers perform a joint placement of terminals and cells to determine the final terminal positions. To ensure the legality of F2F-bonded 3D-IC terminal planning, Pentapati et al. [47] proposed a dual-assignment algorithm that legalizes the bonding terminals after the routing engine inserts them, aiming to minimize the total displacement of the terminals. However, considering legality at the post-processing stage can result in deviations from the optimized routing solution. Furthermore, improvements in legality are heavily constrained by the initial assignment positions.

In response to challenges with 3-D via overlap violations in modern technology nodes, Huang et al. [21] developed an innovative approach that integrates a new via legalization stage and a refinement phase during routing. They introduced two distinct methods for via legalization: a force-based algorithm and a bipartite-matching algorithm optimized with Bayesian techniques. To overcome the aforementioned limitations, BTAssign [31] explored a generalized form of the bonding terminal assignment problem. The proposed routability-driven assignment framework is integrated before the routing phase, allowing for greater flexibility and improvement. Within the BTAssign framework, an adaptive generalized assignment formulation is efficiently solved through an iterative divide-and-conquer algorithm.

3 Challenges and Opportunities

The previous section introduces several recent 3D physical design works. However, to achieve 3D native physical design flow, there remain several issues that have a critical impact on the quality of results but have not been given enough attention. This section introduces some opportunities and challenges from the perspective of methodologies in the EDA field, focusing on three aspects: new design objectives that need to be considered, new constraints introduced by emerging technologies, and new stages that should be integrated into the design flow.

3.1 Partitioning Algorithm

Apart from the challenges discussed in the Section 2.3, hypergraph balanced bi-partitioning remains a fundamental challenge in EDA. For 3D-ICs, the partitioning algorithm should address two distinctive challenges:

P1: Balanced Hypergraph Partitioning with Variable Vertex Weights. Existing partitioning works in 2D-IC design mainly focus on min-cut based partitioning. The integration of heterogeneous technologies results in varying cell sizes across different dies due to distinct process characteristics. This makes it critical to develop balanced bi-partitioning algorithms that accommodate variable weights.

P2: Advanced Objective Requirements. Beyond traditional balanced partitioning constraints, partitioning algorithms should address a broader range of objectives. Specifically, the partitioning process needs to account for critical considerations such as timing and thermal effects to ensure overall design quality.

3.2 Macro Placement

With the increasing complexity of modern integrated circuits, numerous Intellectual Property (IP) blocks must be incorporated to meet design requirements. Consequently, the strategic placement of these IP blocks becomes crucial, particularly in 3D-ICs. Two key challenges must be addressed for solving the macro placement problem in 3D-ICs:

P1: 3D Layout Representation. While conventional macro placement employs various data structures for layout representation, these traditional approaches lack adequate extensions for 3D spaces. Novel algorithms are required to bridge this fundamental gap in 3D representation.

P2: Data-flow Driven 3D Macro Placement. Traditionally, macro placement relies heavily on engineers' expertise. However, the 3D-IC design necessitates a paradigm shift in design methodology. This underscores the importance of developing automated macro placement algorithms that explicitly consider data flow patterns and dependencies.

3.3 Placement

Despite recent advances in addressing 3D placement challenges (discussed in Section 2.4), several critical research directions remain to be explored for achieving high-quality placement.

P1: Wirelength-driven Analytical 3D Placement. A fundamental placement framework is the wirelength-driven analytical placer, traditionally guided by the Half-Perimeter Wirelength (HPWL) model. Extending the conventional 2D placer to 3D-ICs introduces unique complexities due to vertical integration and heterogeneous technologies considerations. Specifically, this can be broken down into three specific reasons: 1) The introduction of the Z-dimension creates significant discontinuity, making it difficult for methods that relax the problem to find approximate solutions, often resulting in a large gap; 2) The 3D wirelength model needs to account for the sum of wirelength for each tier, and the elements required for wirelength calculation change with variations in the Z-coordinate; 3) The areas of heterogeneous technologies cells differ across tiers, leading to discontinuity in the density function along the Z-axis. These three factors together make it challenging to analytically model and solve the wirelength model and density model, necessitating the research of more advanced models and algorithms.

P2: Fusion of Partitioning and Placement. In 3D-IC design, partitioning and placement need to be more closely integrated than in traditional 2D flows. Unlike traditional TSV-based 3D-ICs, hybrid bonding allows for denser inter-die connections, which can significantly impact partitioning and placement decisions. For example, partitioning must account for the physical constraints imposed by the hybrid bonding technology and the variation due to different interconnect capacities between tiers. By integrating partitioning with placement, it is possible to achieve more efficient die-level area utilization and optimized overall wirelength. The approach allows for adjusting the partitioning of cells dynamically based on placement feedback, ensuring that timing-critical paths are kept short and communication-intensive cells are placed in close proximity. This integration also facilitates better management of vertical interconnect resources, improving routability and minimizing congestion in critical regions.

P3: Co-optimization of Hybrid Bonding, Cells, and Macros. The advanced F2F hybrid bonding technology requires a co-optimization approach that simultaneously addresses the placement of standard cells and macros across multiple dies. Similar to solving placement problems with multiple fence regions, there should be a simultaneous movement phase that addresses cell placement across different tiers and supports hybrid bonding. This step would allow for the comprehensive optimization of critical objectives such as wirelength, timing, and routability.

3.4 CTS

Besides the challenges discussed in Section 2.5, several critical factors must be considered to optimize clock distribution across multiple dies, especially when utilizing hybrid bonding technologies. Here are some techniques that may need to be developed further:

P1: Leveraging Hybrid Bonding Terminals to Build 3D Tree Structure. The structure of clock tree is vital because it directly affects skew, latency, wirelength and power. Hybrid bonding terminals provide greater flexibility in vertical interconnections, increasing the solution sapce of tree structure construction. By using hybrid bonding terminals, either bottom-up or top-down approaches can achieve a novel merging method, increasing the possibility of obtaining high-quality 3D CTS. Additionally, bonding terminals introduce potential challenges such as thermal and routability issues. Carefully designed algorithms that balance thermal, power, and routability metrics will become increasingly important in 3D tree structure.

P2: Challenges from Heterogeneity. Designing a 3D clock tree becomes more complex when dealing with different materials and technology nodes across dies. Variations in material properties like dielectric constants and thermal expansion coefficients can introduce delays, even in physically identical clock paths. Accurate delay modeling across different dies is necessary to maintain synchronization. Additionally, threshold voltage variations in different technology nodes can increase skew, requiring fine-grained clock skew modeling to effectively balance these disparities.

3.5 Routing

Besides the challenges discussed in Section 2.6, there are several critical problems to be solved.

P1: Multi-tier Net Topology. In 2D-ICs, interconnects are typically implemented using multiple metal layers. The cost associated with vias, which are the vertical connections between these layers, is generally negligible. This allows designers to use a rectilinear Steiner tree topology to efficiently achieve interconnect routing without significant cost concerns. However, the cost of vertical connections in 3D-ICs is significant and cannot be overlooked. This makes the development of efficient 3D interconnect topologies crucial for

optimizing performance and cost. Dewan et al. [12] propose an algorithm for constructing a routing topology database that enables the creation of all multilayer monolithic rectilinear Steiner minimum trees on the 3D Hanan grid. To demonstrate the database's versatility across different applications, they apply it to generate timing-driven 3D routing topologies and perform congestion-aware global routing on 3D designs.

P2: Cross-tier Routing Resource Model. Most competitive routing algorithms route net one by one, this process does not consider the routing resources required by subsequent nets. This approach often leads to congestion because it doesn't optimize the overall resource allocation across all nets. Since inter-layer connections are fixed before the routing stage, available routing paths become limited, increasing congestion. In addition, since the nets may be distributed across different dies, it is challenging for sequential routing algorithms to simultaneously optimize wirelength, timing and congestion across different dies. To address these challenges, a more holistic approach is necessary. We should simultaneously consider the connectivity requirements of all nets, constructing a comprehensive resource model that includes fixed connections. By pre-allocating resources based on this model and using the pre-allocation results to guide net routing, we can effectively alleviate congestion and improve the efficiency of the routing stage. Multi-step collaborative optimization is also needed to enhance the routability of preceding steps.

3.6 Thermal-driven Physical Design

In modern 3D-ICs physical design, thermal management has become increasingly critical for ensuring reliable system operation. For example, previous work in [61] presents a rapid analysis of power integrity and thermal flow, enabling early-stage voltage drop and thermal analysis, while co-optimizing these solutions throughout the design process. This section examines thermal-driven physical design from two essential perspectives:

P1: Fast Thermal Analytical Method. Thermal-driven physical design demands fast analytical models to evaluate and predict on-chip temperature distributions, allowing designers to identify potential thermal hotspots early in the process. Since traditional methods for computing thermal distributions are time-consuming, an approximate model is needed for fast estimation. As each design stage provides varying levels of detail, it is crucial to develop suitable models tailored to each stage.

P2: Optimization Algorithm. The thermal models can be integrated into multiple design stages, from floorplanning to detailed placement and routing. It's worth considering integrating thermal analysis results into objective functions and optimization algorithms during the solution process, creating a feedback mechanism for improved thermal management.

3.7 Hybrid Bonding Optimization

In F2F-bonded 3D-ICs, the number and location of bonding terminals serve as critical determinants of overall design quality. To achieve better performance, hybrid bonding terminals not only require standalone optimizations, but also need to be integrated into the entire design flow. Two principal research directions warrant investigation:

P1: Optimization of The Number of Terminals. The number of HBTs is a key metric that partitioning algorithms need to consider. For a cross net, at least one HBT is required to connect pins on different tiers. Notably, using more than one hybrid bonding terminal for a cross net may lead to improved net topology, thereby impacting the overall routing results [31]. Therefore, allocating the appropriate number of hybrid bonding terminals for nets is an often-overlooked yet crucial issue.

P2: Optimization of The Location of Terminals. The location of HBTs is another critical metric. In F2F hybrid bonding technology, HBTs must connect to the top metal layer and are functionally represented as pins on this highest metal layer. The position of HBTs affects the actual routing; therefore, once the number of HBTs is determined, they typically need to simultaneously move together with the cells to collaboratively optimize key metrics such as wirelength and routability. Additionally, since the top metal layers usually contain power grids, the placement of HBTs must avoid these power grids to minimize design rule violations. Therefore, developing algorithms that consider process constraints and routability objectives is also of great importance.

4 Conclusions

In conclusion, this paper reviews the advancements and ongoing challenges in 3D physical design, particularly focusing on F2F bonding technology's potential and limitations. As integration density and complexity grow, traditional design frameworks are increasingly inadequate, necessitating innovative solutions across multiple design stages. Despite notable progress, achieving robust 3D layouts that optimize timing, thermal, and power constraints remains challenging. Future research is needed to develop more native 3D IC algorithms, such as heterogeneous-aware partitioning algorithms, advanced thermal-aware placement algorithms, and enhanced hierarchical co-optimization techniques. Addressing these areas will be crucial to unlocking 3D ICs' full potential and achieving scalable, high-performance designs.

Acknowledgment

This work is supported in part by the Major Key Project of PCL (No. PCL2023A03), the National Natural Science Foundation of China (No. 62090021), the Natural Science Foundation of Fujian Province (No. 2024J09045).

References

[1] Anthony Agnesina, Moritz Brunion, Alberto García-Ortiz, Francky Catthoor, Dragomir Milojevic, Manu Komalan, Matheus Cavalcante, Samuel Riedel, Luca Benini, and Sung Kyu Lim. 2022. Hier-3D: A Hierarchical Physical Design Methodology for Face-to-Face-Bonded 3D ICs. In Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design. 1–6.

[2] Alexander Andreev, Andrey Nikishin, Sergey Gribok, Phey-Chuin Tan, and Choon-Hun Choo. 2014. Clock Network Fishbone Architecture for a Structured ASIC Manufactured on a 28 NM CMOS Process Lithographic Node.

[3] Lennart Bamberg, Alberto García-Ortiz, Lingjun Zhu, Sai Pentapati, Sung Kyu Lim, et al. 2020. Macro-3D: A Physical Design Methodology for Face-to-Face-Stacked Heterogeneous 3D ICs. In 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE, 37–42.

[4] Kyungwook Chang, Saurabh Sinha, Brian Cline, Raney Southerland, Michael Doherty, Greg Yeric, and Sung Kyu Lim. 2016. Cascade2D: A Design-Aware Partitioning Approach to Monolithic 3D IC with 2D Commercial Tools. In 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). IEEE, 1–8.

[5] Yao-Wen Chang. 2024. Physical Design Challenges in Modern Heterogeneous Integration. In Proceedings of the 2024 International Symposium on Physical Design. 125–134.

[6] Ting-Hai Chao, Yu-Chin Hsu, Jan-Ming Ho, and AB Kahng. 1992. Zero skew clock routing with minimum wirelength. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing 39, 11 (1992), 799–814.

[7] Rishi Chaturvedi and Jiang Hu. 2004. Buffered clock tree for high quality IC design. In International Symposium on Signals, Circuits and Systems. Proceedings, SCS 2003.(Cat. No. 03EX720). IEEE, 381–386.

[8] Yan-Jen Chen, Yan-Syuan Chen, Wei-Che Tseng, Cheng-Yu Chiang, Yu-Hsiang Lo, and Yao-Wen Chang. 2023. Late Breaking Results: Analytical Placement for 3D ICs with Multiple Manufacturing Technologies. In 2023 60th ACM/IEEE Design Automation Conference (DAC). IEEE, 1–2.

[9] Yan-Jen Chen, Cheng-Hsiu Hsieh, Po-Han Su, Shao-Hsiang Chen, and Yao-Wen Chang. 2024. Mixed-Size 3D Analytical Placement with Heterogeneous Technology Nodes. In ACM/IEEE Design Automation Conference (DAC).

[10] Jason Cong, Andrew B Kahng, Cheng-Kok Koh, and C-W Albert Tsao. 1998. Bounded-skew clock and Steiner routing. ACM Transactions on Design Automation of Electronic Systems (TODAES) 3, 3 (1998), 341–388.

[11] Jason Cong and Guojie Luo. 2009. A multilevel analytical placement for 3D ICs. In 2009 Asia and South Pacific Design Automation Conference. IEEE, 361–366.

[12] Monzurul Islam Dewan, Sheng-En David Lin, and Dae Hyun Kim. 2023. Construction of All Multilayer Monolithic RSMTs and Its Application to Monolithic 3D IC Routing. ACM Transactions on Design Automation of Electronic Systems 29, 1 (2023), 1–28.

[13] Paul D Franzon, Erik Jan Marinissen, and Muhannad S Bakir. 2019. Handbook of 3D Integration, Volume 4: Design, Test, and Thermal Management. John Wiley & Sons.

[14] Brent Goplen and Sachin Sapatnekar. 2003. Efficient thermal placement of standard cells in 3D ICs using a force directed approach. In 2003 IEEE/ACM International Conference on Computer Aided Design (ICCAD). IEEE, 86–89.

[15] Matthew R Guthaus, Xuchu Hu, Gustavo Wilke, Guilherme Flach, and Ricardo Reis. 2012. High-performance clock mesh optimization. ACM Transactions on Design Automation of Electronic Systems (TODAES) 17, 3 (2012), 1–17.

[16] Kwangsoo Han, Andrew B Kahng, and Jiajia Li. 2018. Optimal Generalized H-Tree Topology and Buffering for High-Performance and Low-Power Clock Distribution. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 39, 2 (2018), 478–491.

[17] Michael B Healy and Sung Kyu Lim. 2011. A novel TSV topology for many-tier 3D power-delivery networks. In 2011 Design, Automation & Test in Europe. IEEE, 1–4.

[18] Meng-Kai Hsu, Yao-Wen Chang, and Valeriy Balabanov. 2011. TSV-aware analytical placement for 3D IC designs. In Proceedings of the 48th Design Automation Conference. 664–669.

[19] Kai-Shun Hu, I-Jye Lin, Yu-Hui Huang, Hao-Yu Chi, Yi-Hsuan Wu, and Chin-Fang Cindy Shen. 2022. 2022 ICCAD CAD Contest Problem B: 3D Placement with D2D Vertical Connections. In Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design. 1–5.

[20] Gang Huang, Muhannad Bakir, Azad Naeemi, Howard Chen, and James D Meindl. 2007. Power Delivery for 3D Chip Stacks: Physical Modeling and Design Implication. In 2007 IEEE Electrical Performance of Electronic Packaging. IEEE, 205–208.

[21] Yen-Hsiang Huang, Sai Pentapati, Anthony Agnesina, Moritz Brunion, and Sung Kyu Lim. 2024. On Legalization of Die Bonding Bumps and Pads for 3D ICs. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (2024).

[22] Nauman H Khan, Syed M Alam, and Soha Hassoun. 2010. Power Delivery Design for 3-D ICs Using Different Through-Silicon Via (TSV) Technologies. IEEE Transactions on Very Large Scale Integration (VLSI) Systems 19, 4 (2010), 647–658.

[23] Donggyu Kim, Minjae Kim, Junseok Hur, Jakang Lee, Jinoh Cho, and Seokhyeong Kang. 2024. TA3D: Timing-Aware 3D IC Partitioning and Placement by Optimizing the Critical Path. In Proceedings of the 2024 ACM/IEEE International Symposium on Machine Learning for CAD. 1–7.

[24] Jinwoo Kim, Lingjun Zhu, Hakki Mert Torun, Madhavan Swaminathan, and Sung Kyu Lim. 2021. Micro-bumping, Hybrid Bonding, or Monolithic? A PPA Study for Heterogeneous 3D IC Options. In 2021 58th ACM/IEEE Design Automation Conference (DAC). IEEE, 1189–1194.

[25] Tak-Yung Kim and Taewhan Kim. 2010. Clock tree embedding for 3D ICs. In 2010 15th Asia and South Pacific Design Automation Conference (ASP-DAC). IEEE, 486–491.

[26] Tak-Yung Kim and Taewhan Kim. 2011. Clock Tree synthesis for TSV-based 3D IC designs. ACM Transactions on Design Automation of Electronic Systems (TODAES) 16, 4 (2011), 1–21.

[27] Johann Knechtel and Jens Lienig. 2016. Physical design automation for 3D chip stacks: challenges and solutions. In Proceedings of the 2016 on International Symposium on Physical Design. 3–10.

[28] Bon Woong Ku, Kyungwook Chang, and Sung Kyu Lim. 2018. Compact-2D: A Physical Design Methodology to Build Commercial-Quality Face-to-Face-Bonded 3D ICs. In Proceedings of the 2018 International Symposium on Physical Design. 90–97.

[29] Peiyu Liao, Yuxuan Zhao, Dawei Guo, Yibo Lin, and Bei Yu. 2023. Analytical Die-to-Die 3-D Placement With Bistratal Wirelength Model and GPU Acceleration. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (2023).

[30] Minghao Lin, Heming Sun, and Shinji Kimura. 2016. Power-efficient and slew-aware three dimensional gated clock tree synthesis. In 2016 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC). IEEE, 1–6.

[31] Siting Liu, Jiaxi Jiang, Zhuolun He, Ziyi Wang, Yibo Lin, Bei Yu, and Martin Wong. 2024. Routing-aware Legal Hybrid Bonding Terminal Assignment for 3D Face-to-Face Stacked ICs. In Proceedings of the 2024 International Symposium on Physical Design. 75–82.

[32] Jingwei Lu, Hao Zhuang, Ilgweon Kang, Pengwen Chen, and Chung-Kuan Cheng. 2016. ePlace-3D: Electrostatics based Placement for 3D-ICs. In Proceedings of the 2016 on International Symposium on Physical Design. 11–18.

[33] Tiantao Lu and Ankur Srivastava. 2014. Gated low-power clock tree synthesis for 3D-ICs. In Proceedings of the 2014 international symposium on Low power electronics and design. 319–322.

[34] Tiantao Lu and Ankur Srivastava. 2017. Low-Power Clock Tree Synthesis for 3D-ICs. ACM Transactions on Design Automation of Electronic Systems (TODAES) 22, 3 (2017), 1–24.

[35] Yi-Chen Lu, Sai Surya Kiran Pentapati, Lingjun Zhu, Kambiz Samadi, and Sung Kyu Lim. 2020. TP-GNN: A Graph Neural Network Framework for Tier Partitioning in Monolithic 3D ICs. In 2020 57th ACM/IEEE Design Automation Conference (DAC). IEEE, 1–6.

[36] Jacob Minz, Xin Zhao, and Sung Kyu Lim. 2008. Buffered clock tree synthesis for 3D ICs under thermal variations. In 2008 Asia and South Pacific Design Automation Conference. IEEE, 504–509.

[37] Gauthaman Murali and Sung Kyu Lim. 2021. Heterogeneous 3d ics: Current status and future directions for physical design technologies. In 2021 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE, 146–151.

[38] T Naito, T Ishida, T Onoduka, M Nishigoori, T Nakayama, Y Ueno, Y Ishimoto, A Suzuki, W Chung, R Madurawe, et al. 2010. World's first monolithic 3D-FPGA with TFT SRAM over 90nm 9 layer Cu CMOS. In 2010 Symposium on VLSI Technology. IEEE, 219–220.

[39] Deok Keun Oh, Mu Jun Choi, and Ju Ho Kim. 2019. Thermal-aware 3D Symmetrical Buffered Clock Tree Synthesis. ACM Transactions on Design Automation of Electronic Systems (TODAES) 24, 3 (2019), 1–22.

[40] Jun So Pak, Joohee Kim, Jonghyun Cho, Kiyeong Kim, Taigon Song, Seungyoung Ahn, Junho Lee, Hyungdong Lee, Kunwoo Park, and Joungho Kim. 2011. PDN Impedance Modeling and Analysis of 3D TSV IC by Using Proposed P/G TSV Array Model Based on Separated P/G TSV and Chip-PDN Models. IEEE Transactions on Components, Packaging and Manufacturing Technology 1, 2 (2011), 208–219.

[41] Shreepad Panth, Kambiz Samadi, Yang Du, and Sung Kyu Lim. 2014. Placement-driven partitioning for congestion mitigation in monolithic 3D IC designs. In Proceedings of the 2014 on International symposium on physical design. 47–54.

[42] Shreepad Panth, Kambiz Samadi, Yang Du, and Sung Kyu Lim. 2015. Placement-Driven Partitioning for Congestion Mitigation in Monolithic 3D IC Designs. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 4, 34 (2015), 540–553.

[43] Shreepad Panth, Kambiz Samadi, Yang Du, and Sung Kyu Lim. 2017. Shrunk-2-D: A Physical Design Methodology to Build Commercial-Quality Monolithic 3-D ICs. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 36, 10 (2017), 1716–1724.

[44] Heechun Park and Taewhan Kim. 2014. Synthesis of TSV Fault-Tolerant 3-D Clock Trees. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 34, 2 (2014), 266–279.

[45] Heechun Park, Bon Woong Ku, Kyungwook Chang, Da Eun Shim, and Sung Kyu Lim. 2021. Pseudo-3D Physical Design Flow for Monolithic 3D ICs: Comparisons and Enhancements. ACM Transactions on Design Automation of Electronic Systems (TODAES) 26, 5 (2021), 1–25.

[46] Mohit Pathak, Young-Joon Lee, Thomas Moon, and Sung Kyu Lim. 2010. Through-silicon-via management during 3D physical design: When to add and how many?. In 2010 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). IEEE, 387–394.

[47] Sai Pentapati, Anthony Agnesina, Moritz Brunion, Yen-Hsiang Huang, and Sung Kyu Lim. 2023. On Legalization of Die Bonding Bumps and Pads for 3D ICs. In Proceedings of the 2023 International Symposium on Physical Design. 62–70.

[48] Sai Surya Kiran Pentapati, Kyungwook Chang, Vassilios Gerousis, Rwik Sengupta, and Sung Kyu Lim. 2020. Pin-3D: A physical synthesis and post-layout optimization flow for heterogeneous monolithic 3D ICs. In Proceedings of the 39th International Conference on Computer-Aided Design. 1–9.

[49] Anand Rajaram and David Z Pan. 2006. Variation tolerant buffered clock network synthesis with cross links. In Proceedings of the 2006 International Symposium on Physical Design. 157–164.

[50] Yang Shang, Chun Zhang, Hao Yu, Chuan Seng Tan, Xin Zhao, and Sung Kyu Lim. 2013. Thermal-reliable 3D clock-tree synthesis considering nonlinear electrical-thermal-coupled TSV model. In 2013 18th Asia and South Pacific Design Automation Conference (ASP-DAC). IEEE, 693–698.

[51] Sebastien Thuries, Olivier Billoint, Sylvain Choisnet, Romain Lemaire, Pascal Vivet, Perrine Batude, and Didier Lattard. 2020. M3D-ADTCO: Monolithic 3D Architecture, Design and Technology Co-Optimization for High Energy Efficient 3D IC. In 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE). IEEE, 1740–1745.

[52] Aida Todri-Sanial, Sandip Kundu, Patrick Girard, Alberto Bosio, Luigi Dilillo, and Arnaud Virazel. 2013. Globally Constrained Locally Optimized 3-D Power Delivery Networks. IEEE Transactions on very large scale Integration (VLSI) Systems 22, 10 (2013), 2131–2144.

[53] Jeng-Liang Tsai, Tsung-Hao Chen, and CC-P Chen. 2004. Zero skew clock-tree optimization with buffer insertion/sizing and wire sizing. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 23, 4 (2004), 565–572.

[54] Chung-Wen Albert Tsao and Cheng-Kok Koh. 2002. UST/DME: a clock tree router for general skew constraints. ACM Transactions on Design Automation of Electronic Systems (TODAES) 7, 3 (2002), 359–379.

[55] Pruek Vanna-Iampikul, Chengjia Shao, Yi-Chen Lu, Sai Pentapati, and Sung Kyu Lim. 2021. Snap-3D: A Constrained Placement-Driven Physical Design Methodology for Face-to-Face-Bonded 3D ICs. In Proceedings of the 2021 International Symposium on Physical Design. 39–46.

[56] Sying-Jyan Wang, Cheng-Hao Lin, and Katherine Shu-Min Li. 2013. Synthesis of 3D clock tree with pre-bond testability. In 2013 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2654–2657.

[57] Xueyan Zhao, Shijian Chen, Yihang Qiu, Jiangkao Li, Zhipeng Huang, Biwei Xie, Xingquan Li, and Yungang Bao. 2023. iPL-3D: A Novel Bilevel Programming Model for Die-to-Die Placement. In 2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD). IEEE, 1–9.

[58] Yuxuan Zhao, Peiyu Liao, Siting Liu, Jiaxi Jiang, Yibo Lin, and Bei Yu. 2024. Analytical Heterogeneous Die-to-Die 3D Placement with Macros. arXiv preprint arXiv:2403.09070 (2024).

[59] Lingjun Zhu, Chanmin Jo, and Sung Kyu Lim. 2022. Power Delivery Solutions and PPA Impacts in Micro-Bump and Hybrid-Bonding 3D ICs. IEEE Transactions on Components, Packaging and Manufacturing Technology 12, 12 (2022), 1969–1982.

[60] Lingjun Zhu and Sung Kyu Lim. 2021. Physical Design Challenges and Solutions for Emerging Heterogeneous 3D Integration Technologies. In Proceedings of the 2021 International Symposium on Physical Design. 127–134.

[61] Lingjun Zhu, Tuan Ta, Rossana Liu, Rahul Mathur, Xiaoqing Xu, Shidhartha Das, Ankit Kaul, Alejandro Rico, Doug Joseph, Brian Cline, et al. 2021. Power Delivery and Thermal-Aware Arm-Based Multi-Tier 3D Architecture. In 2021 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED). IEEE, 1–6.