International Symposium on Physical Design

# Invited: Physical Design for Advanced 3D ICs: Challenges and Solutions

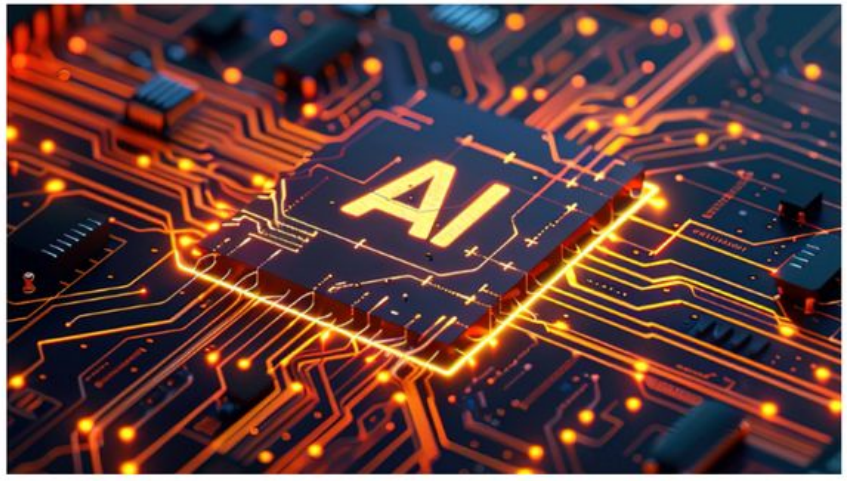
Yuxuan Zhao, Lancheng Zou, Bei Yu

Department of Computer Science and Engineering
The Chinese University of Hong Kong

# Outline

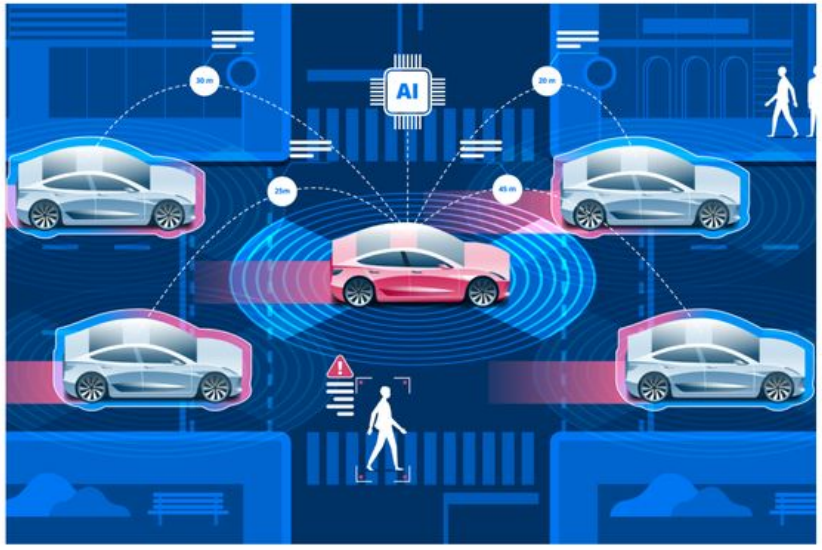
# Applications

## Artificial Intelligence



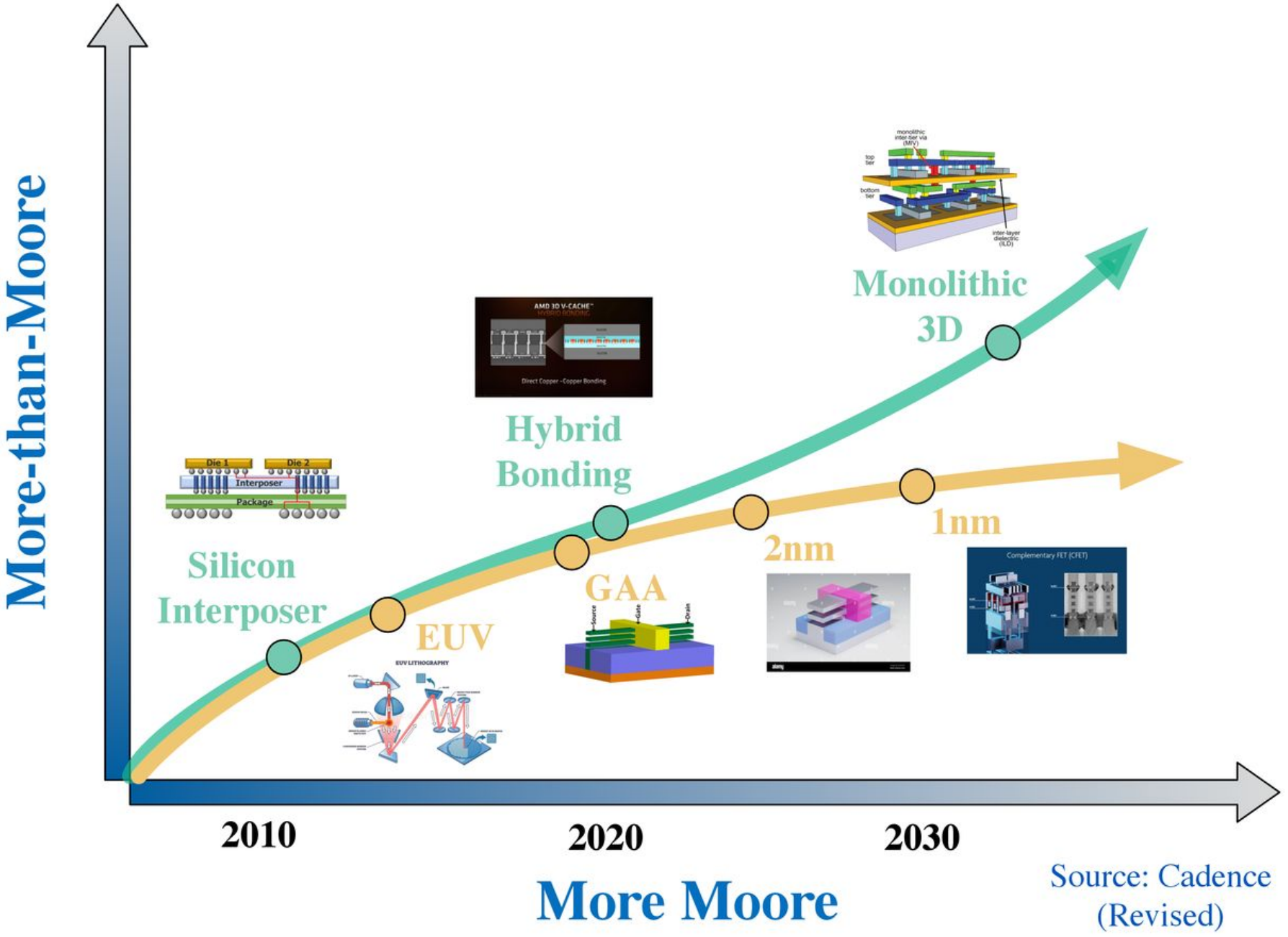

## High Performance Computing

## Autonomous Driving



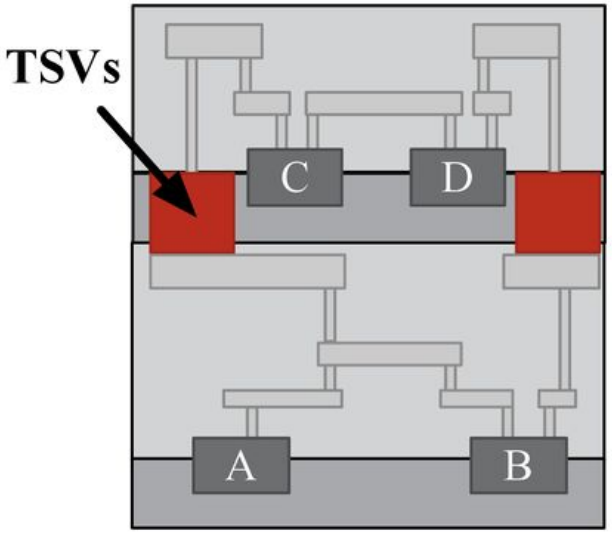

## Internet of Things

# More Moore and More-than-Moore



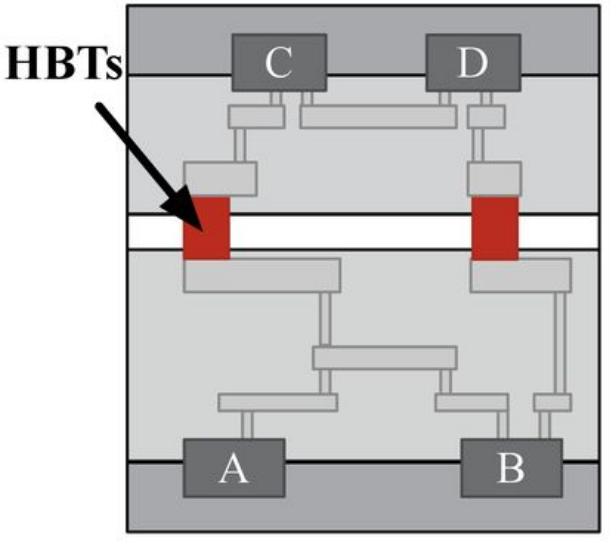

Source: Cadence (Revised)

# 3D Integration Technologies



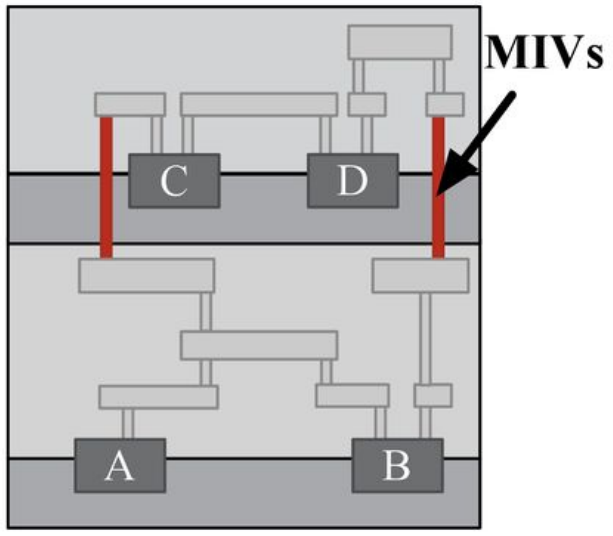

**TSV based**

- Commercially viable
- High area overhead
- High RC parasitics




**Hybrid bonding**

- Low fabrication cost
- No area overhead
- Coupling issues



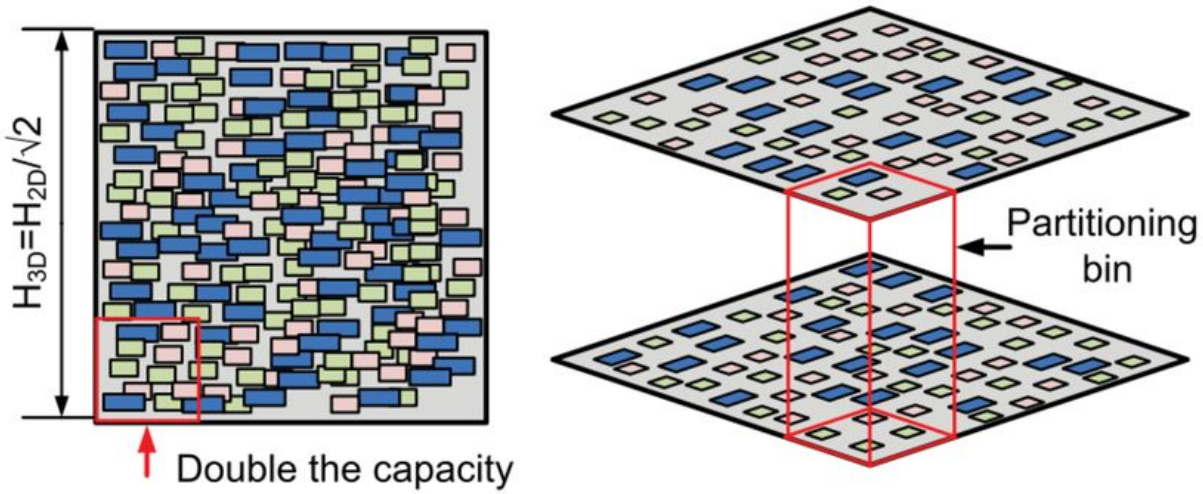

**Monolithic**

- High fabrication cost
- High integration density
- Small RC parasitics

# Physical Design Challenges

# Tier Partitioning

- Most existing approaches employ min-cut partitioners designed for 2D ICs.
  - Most design flows overlook critical technology and design-related parameters during partitioning.

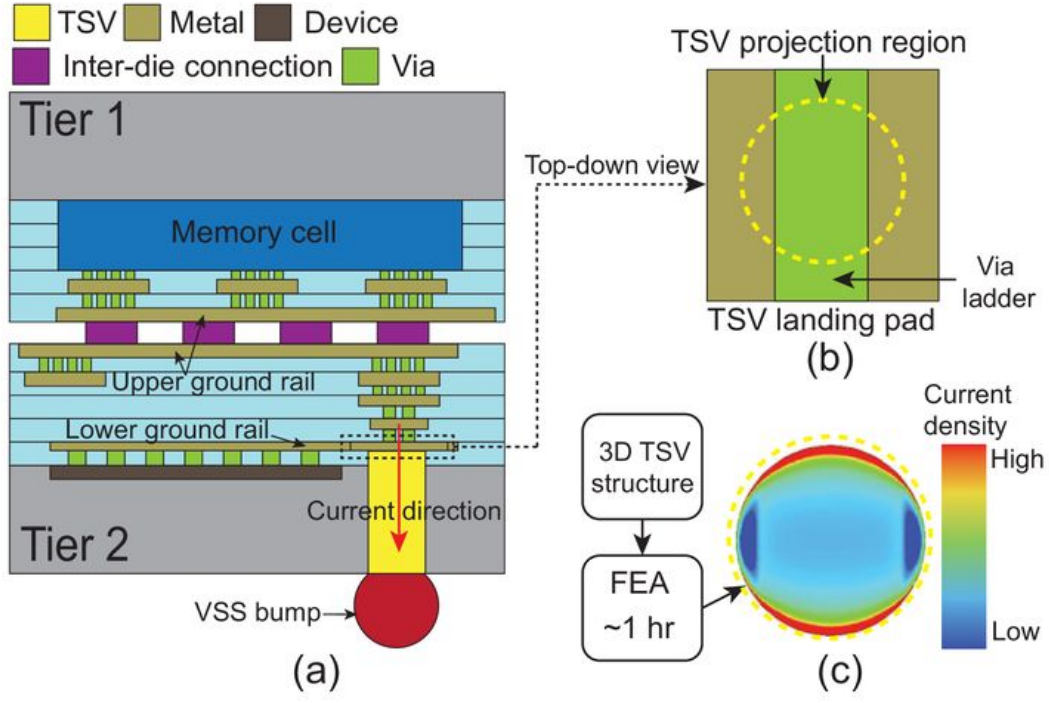



Bin-based FM partitioning [ISPD'14][1] is unaware of technology and design-related information.

[1] Shreepad Panth et al. (2014). "Placement-driven partitioning for congestion mitigation in monolithic 3D IC designs". In: *Proc. ISPD*, pp. 47–54.

# 3D Power Delivery

- Increased power density and larger current within PDNs and TSVs result in elevated IR drop and EM degradation.
  - Existing simulation methods are time-consuming.
  - 3D PDN optimization is essential for performance and power integrity.

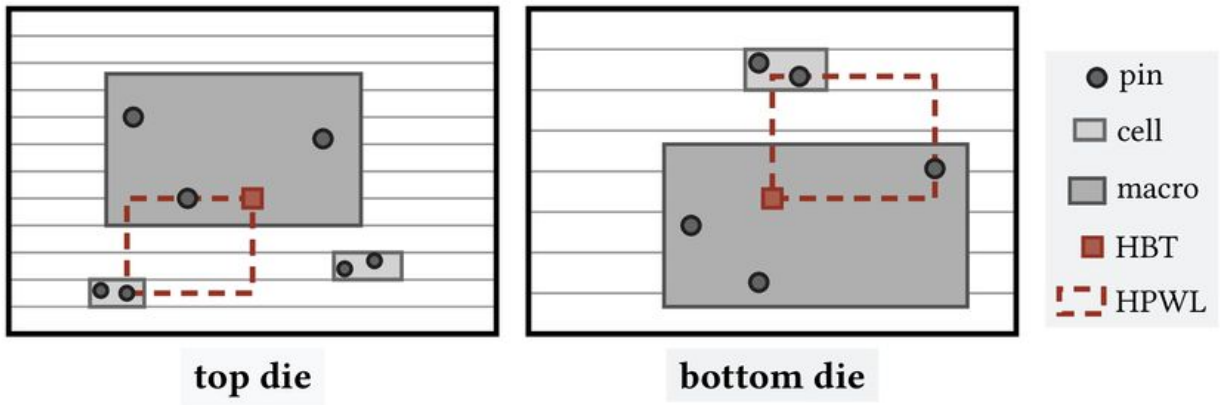



Face-to-face 3D IC PDN structure [ISPD'25][2]. FEA simulation is time-consuming.

[2]Zheng Yang et al. (2025). "ML-based fine-grained modeling of DC current crowding in power delivery TSVs for Face-to-Face 3D ICs". In: *Proc. ISPD*.

# 3D Placement

- Pseudo-3D placement is limited by the initial partitioning and overlook the 3D via locations.
  - Traditional HPWL model does not account for the locations of 3D vias.
  - $z$-direction in 3D placement is highly discrete.

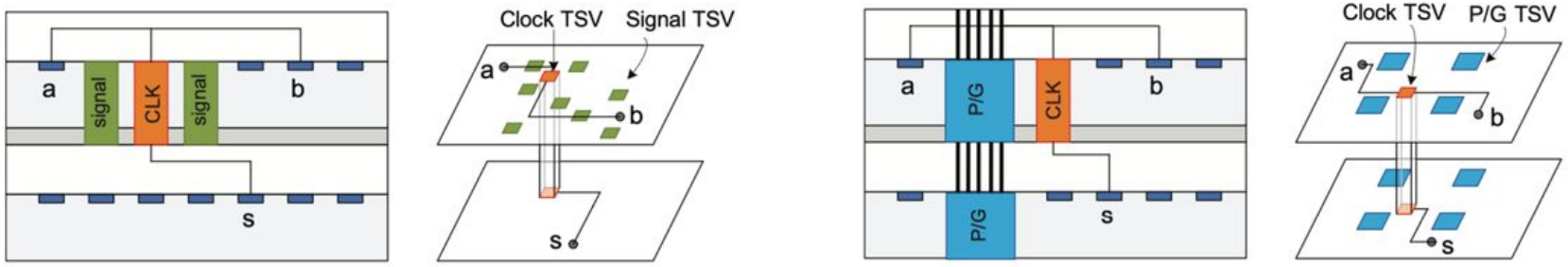



Die-to-die HPWL [ICCAD'23][3] of a 3D net is the sum of the wirelength of the top and bottom nets.

[3] Kai-Shun Hu et al. (2023). "2023 ICCAD CAD Contest Problem B: 3D Placement with Macros". In: *Proc. ICCAD*. IEEE, pp. 1–6.

# 3D Clock Delivery

- 3D clock tree design should minimize clock skew while ensuring power efficiency and signal integrity.
  - Carefully account for the placement and utilization of 3D clock vias.
  - The placement of sequential elements, the partitioning strategy, and the 3D clock tree topology critically influence the performance.




(a) signal TSVs as placement obstacle

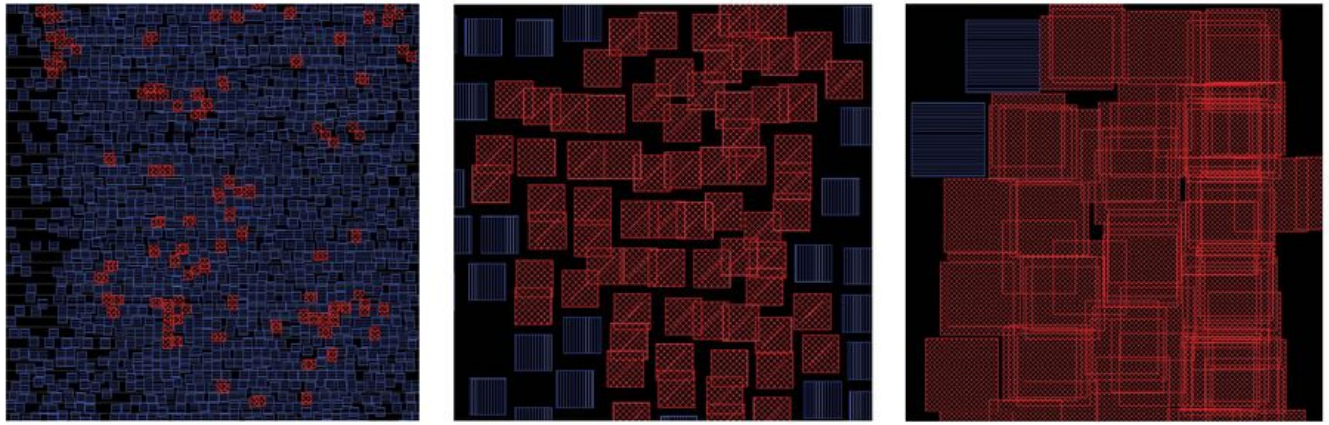
(b) P/G TSVs as placement and routing obstacle

Clock TSVs should be aware of the locations of the signal TSVs and P/G TSVs [ASPDAC'12][4].

[4] Xin Zhao and Sung Kyu Lim (2012). "Through-silicon-via-induced obstacle-aware clock tree synthesis for 3D ICs". In: *Proc. ASPDAC*. IEEE, pp. 347–352.

# 3D Routing

- 3D routing can be executed using conventional 2D routers with specialized considerations for inter-die connections.
  - 3D rectilinear steiner minimum tree (RSMT) construction.
  - 3D via legalization to reduce violations while maintaining minimal impact on wirelength.



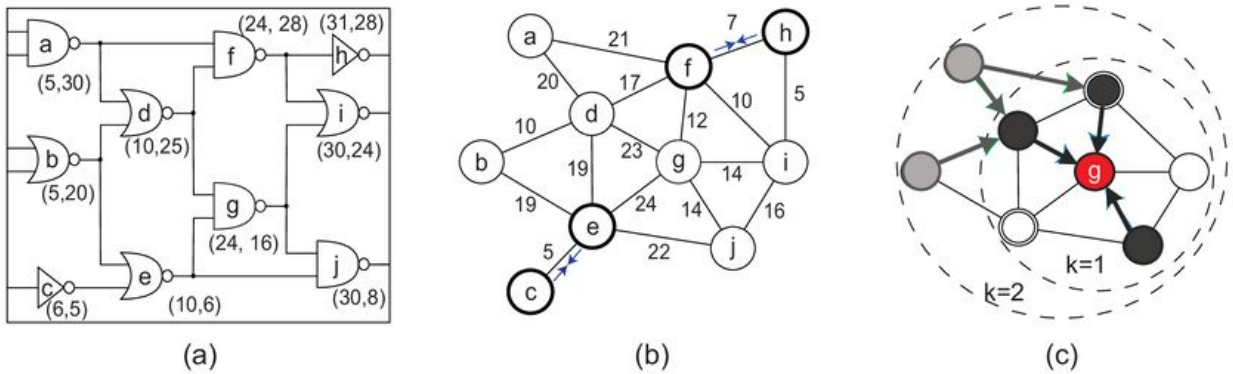

3D via overlaps shown in red at various pitch values [ISPD'23][5].

[5] Sai Pentapati, Anthony Agnesina, et al. (2023). "On legalization of die bonding bumps and pads for 3D ICs". In: *Proc. ISPD*, pp. 62–70.

# Physical Design Solutions

# Tier Partitioning

- Design-aware partitioners: logic-on-memory partitioning [DATE'20][6] and RTL level partitioning [ICCAD'16][7].
- GNN-based partitioning [DAC'20][8] to consider technology and design-related information.




The framework of TP-GNN [DAC'20].

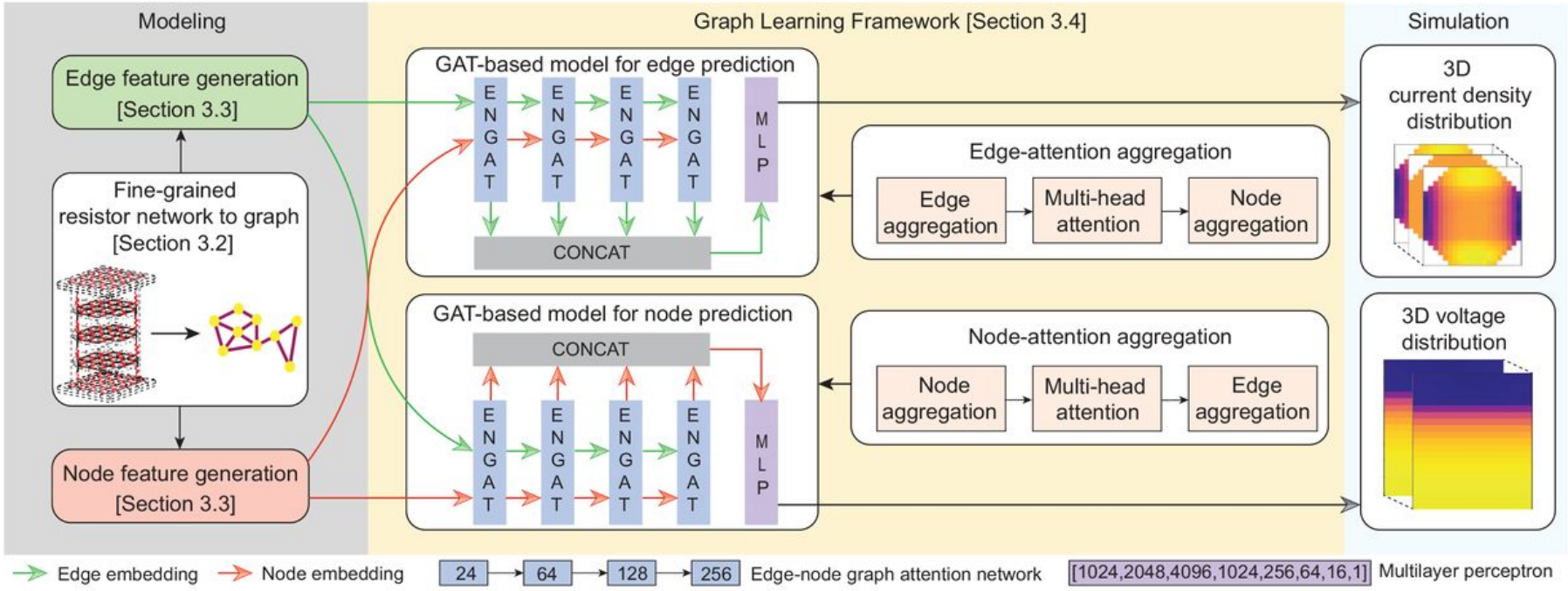
[6]Lennart Bamberg et al. (2020). "Macro-3D: A physical design methodology for face-to-face-stacked heterogeneous 3D ICs". In: *Proc. DATE*. IEEE, pp. 37–42.

[7]Kyungwook Chang et al. (2016). "Cascade2D: A design-aware partitioning approach to monolithic 3D IC with 2D commercial tools". In: *Proc. ICCAD*. IEEE, pp. 1–8.

[8]Yi-Chen Lu et al. (2020). "TP-GNN: A graph neural network framework for tier partitioning in monolithic 3D ICs". In: *Proc. DAC*. IEEE, pp. 1–6.

# 3D Power Delivery

- Fast DC current crowding modeling in P/G TSVs via GAT-based model [ISPD'25].
- Based on the Gaussian process regression, Hetero-3D [ISLPED'24][9] minimizes total wire resistance while observing IR drop constraints.

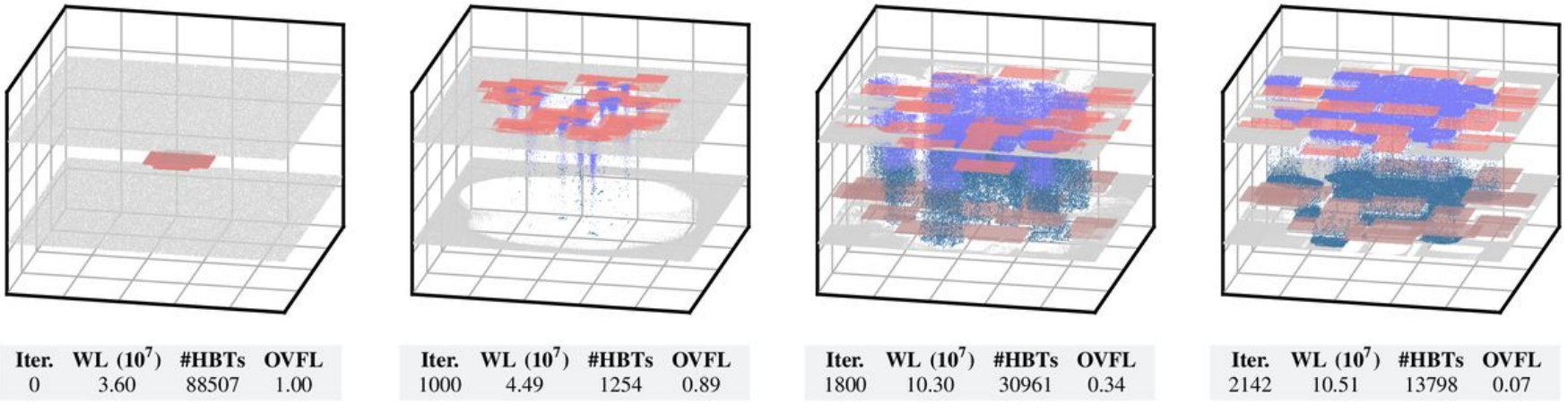



GAT-based analysis framework [ISPD'25].

[9] Lingjun Zhu et al. (2024). "Hetero-3D: Maximizing performance and power delivery benefits of heterogeneous 3D ICs". In: *Proc. ISLPED*, pp. 1–6.

# 3D Placement

- Bistratal wirelength model [TCAD'23][10] for die-to-die HPWL.
- 3D analytical mixed-size placement framework [TCAD'24][11] with 3D density model and bistratal wirelength model.



| Iter. | WL ($10^7$) | #HBTs | OVFL |
|---|---|---|---|
| 0 | 3.60 | 88507 | 1.00 |

| Iter. | WL ($10^7$) | #HBTs | OVFL |
|---|---|---|---|
| 1000 | 4.49 | 1254 | 0.89 |

| Iter. | WL ($10^7$) | #HBTs | OVFL |
|---|---|---|---|
| 1800 | 10.30 | 30961 | 0.34 |

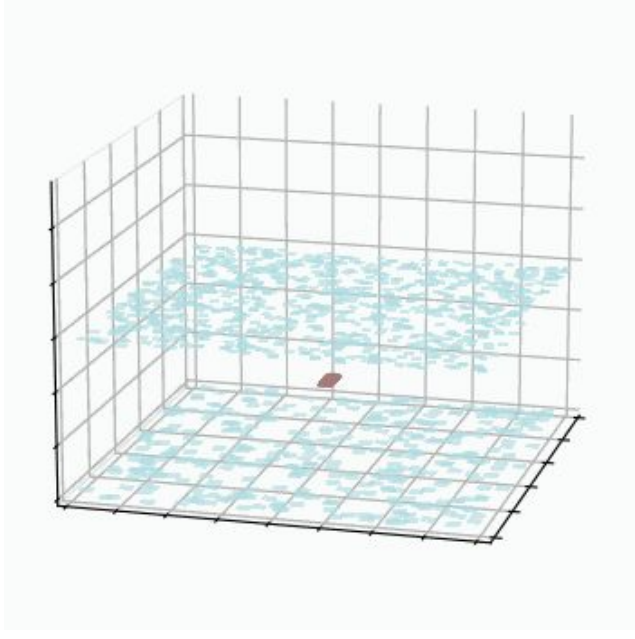
| Iter. | WL ($10^7$) | #HBTs | OVFL |
|---|---|---|---|
| 2142 | 10.51 | 13798 | 0.07 |

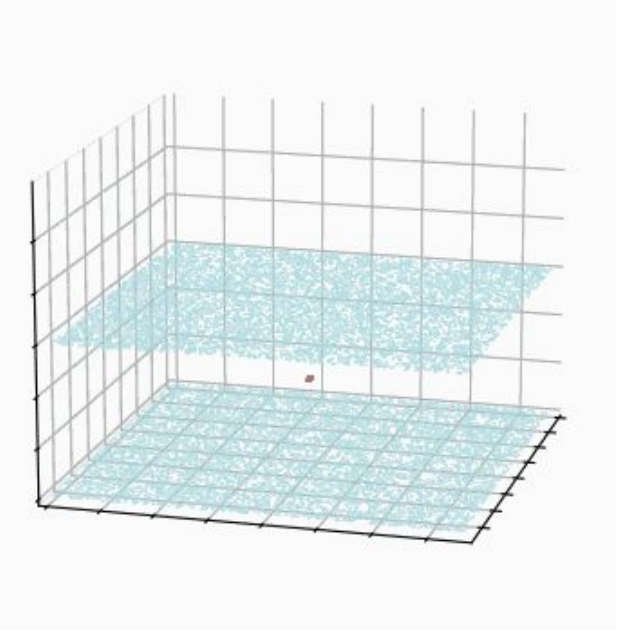
3D mixed-size global placement [TCAD'24].

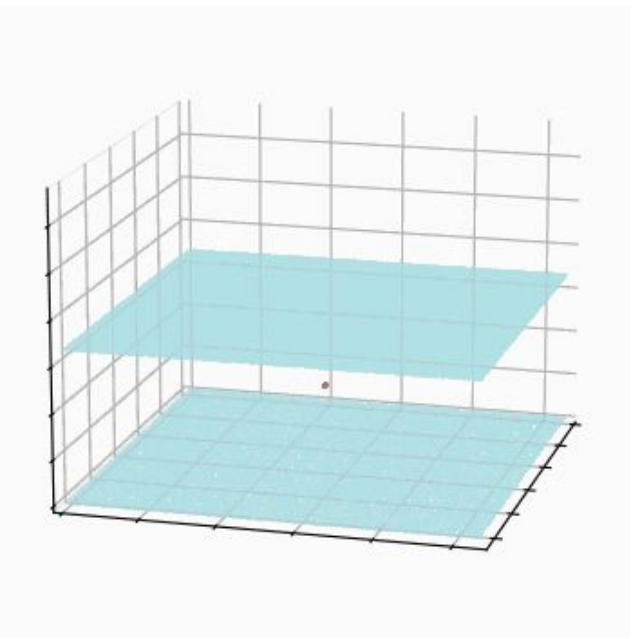
[10] Peiyu Liao et al. (2023). "Analytical Die-to-Die 3D Placement with Bistratal Wirelength Model and GPU Acceleration". In: *IEEE TCAD*.

[11] Yuxuan Zhao et al. (2024). "Analytical Heterogeneous Die-to-Die 3D Placement with Macros". In: *IEEE TCAD*.

# Concurrent Partitioning & Placement

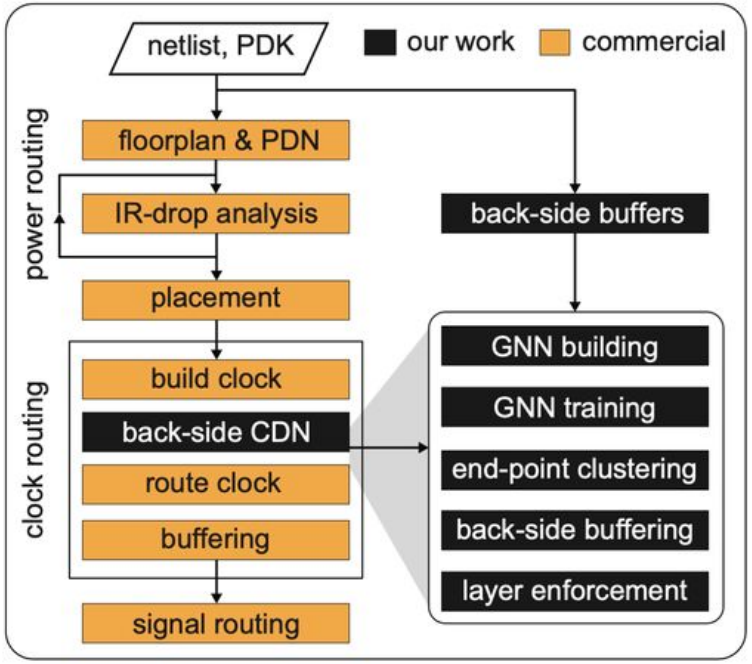






# 3D Clock Delivery

- 3D-MMM [TCPMT'10][a] & 3D-DME [TODAES'11][b] construct 3D clock tree.
- Tier partitioning and flip-flop relocation [ISLPED'19][c] for 3D clock tree optimization.
- GNN-based multi-bit flip-flop clustering [TODAES'23][d].




Back-side clock routing [DAC'24]

[a] Xin Zhao, Jacob Minz, and Sung Kyu Lim (2010). "Low-power and reliable clock network design for through-silicon via (TSV) based 3D ICs". In: *IEEE TCPMT* 1.2, pp. 247–259.

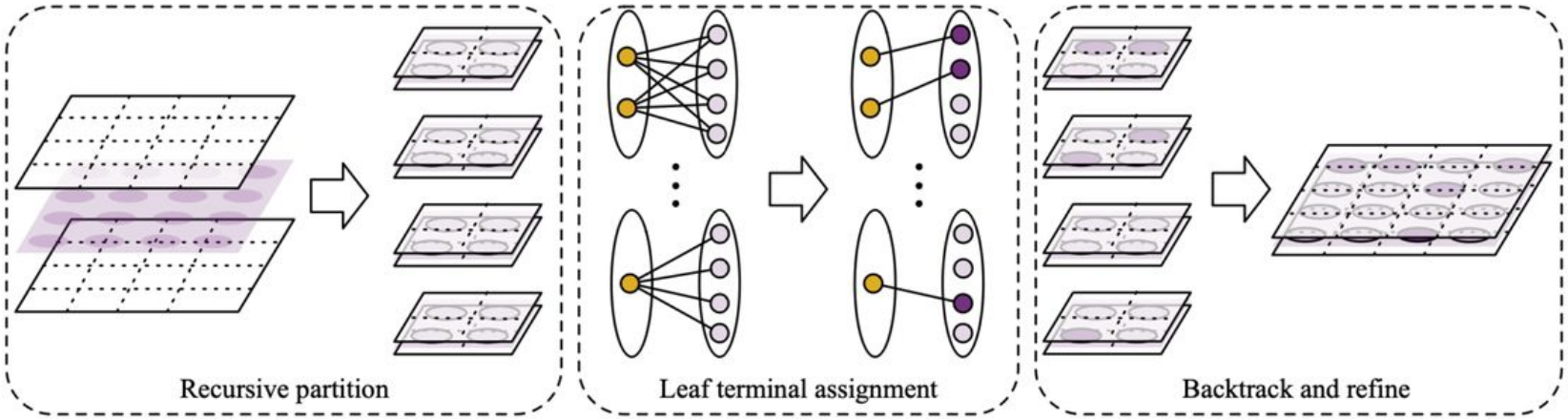
[b] Tak-Yung Kim and Taewhan Kim (2011). "Clock tree synthesis for TSV-based 3D IC designs". In: *ACM TODAES* 16.4, pp. 1–21.

[c] Sai Pentapati, Jeehyun Lee, et al. (2019). "Tier partitioning and flip-flop relocation methods for clock trees in monolithic 3D ICs". In: *Proc. ISLPED*. IEEE, pp. 1–6.

[d] Pruek Vanna-Iampikul et al. (2023). "GNN-based multi-bit flip-flop clustering and post-clustering design optimization for energy-efficient 3D ICs". In: *ACM TODAES* 28.5, pp. 1–26.

# 3D Routing

- FLUTE-like 3D RSMT construction [TODAES'23][12].
- 3D via legalization [ISPD'23] using force-based solver and bipartite-matching at post-routing stage.
- BTAssign [ISPD'24][13] utilizes an iterative hierarchical bipartite-matching algorithm to assign 3D vias to legal positions before routing.




Bonding terminal assignment framework with Quad-tree [ISPD'24].

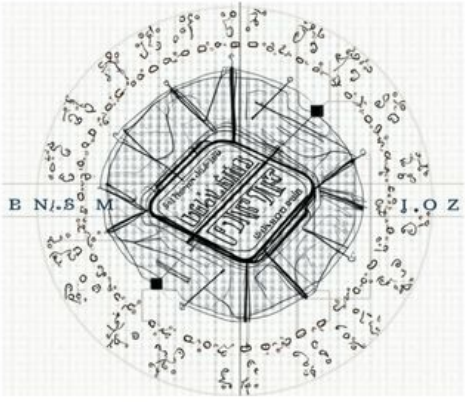
[12]Monzurul Islam Dewan, Sheng-En David Lin, and Dae Hyun Kim (2023). "Construction of all multilayer monolithic rsmts and its application to monolithic 3D IC routing". In: *ACM TODAES* 29.1, pp. 1–28.

[13]Siting Liu et al. (2024). "Routing-aware legal hybrid bonding terminal assignment for 3D face-to-face stacked ICs". In: *Proc. ISPD*, pp. 75–82.

# Conclusion

# Future Studies

- Machine learning approaches, particularly multimodal fusion techniques, show great promise in partitioning and clock delivery problems.
- 3D PDN design should be considered with other design stages.
- 3D placement and routing should account for other constraints, including blockages, routability, and timing.
- Additionally, large language models (LLMs) have the potential to streamline the control of the physical design process.



# Conclusion

- 3D ICs have shown promising enhancements in PPA, independent of the costly scaling of transistors.
- The development of large-scale native 3D EDA tools is crucial for fully realizing the potential of this promising technology frontier.