

Evaluation of Thermal and Power integrity and its Impact on Performance for 3D Memory-on-Logic CPUs with FSPDN and BSPDN

Yumeng Wang †
School of Electronics
Peking University
Beijing, China
2501112234@stu.pku.edu.cn

Xincheng Liu †
School of Physics and Electronics
Hunan University, Hunan, China
School of Electronics
liuxincheng2000@163.com

Hu Zhou†
School of Electronics
Peking University
Beijing, China
2200012717@stu.pku.edu.cn

Linqiu Wang
School of Electronics
Peking University
Beijing, China
wanglinqiu@stu.pku.edu.cn

Haolan Yang
Hunan Institute of Advanced Sensing
and Information Technology
Xiangtan University, Hunan, China
hlyang0112@163.com

Zhuojun Chen*
School of Physics and Electronics
Hunan University
Hunan, China
zjchen@hnu.edu.cn

Zhiyong Zhang
Center for Carbon-based Electronics,
School of Electronics,
Peking University
Beijing, China
zyzhang@pku.edu.cn

Lianmao Peng
Center for Carbon-based Electronics,
School of Electronics,
Peking University
Beijing, China
lmpeng@pku.edu.cn

Rongmei Chen*
Center for Carbon-based Electronics,
School of Electronics,
Peking University
Beijing, China
crm@pku.edu.cn

Abstract—While three-dimensional (3D) Memory-on-Logic integration benefits high-performance computing (HPC), it faces critical bottlenecks in power delivery and thermal management. This paper presents a comprehensive power, performance, area, and thermal (PPAT) evaluation of a 3D Memory-on-Logic CPU utilizing Frontside Power Delivery Network (FSPDN) and Backside Power Delivery Network (BSPDN). Our analysis reveals a fundamental trade-off: while BSPDN significantly improves power integrity by reducing logic IR drop by $7.7\times$ (vs. 3D FSPDN CPU) and $12\times$ (vs. 2D CPU), the extreme substrate thinning required for backside connectivity severely impedes lateral heat dissipation, raising peak temperatures by $\sim 8^\circ\text{C}$ (vs. 3D FSPDN CPU) and $\sim 12^\circ\text{C}$ (vs. 2D CPU). By incorporating thermal-electrical coupling into a spatial-temperature-aware timing analysis, we demonstrate that unlike 3D FSPDN which yields negligible gains over 2D case due to through-silicon via bottlenecks, the superior power integrity of BSPDN decisively outweighs thermal penalties, achieving a net $\sim 30\%$ performance improvement over the 2D counterpart.

Keywords—BSPDN, Memory-on-Logic, PPAT, Electro-thermal Coupling

I. INTRODUCTION

High-performance computing (HPC) demands drive the adoption of 3D Memory-on-Logic integration [1]. While overcoming bandwidth limits, this architecture faces a severe "power-thermal wall": lengthened power delivery paths degrade power integrity, and the stacked structure impedes heat dissipation of the bottom die [2]. To mitigate the severe IR drop, Backside Power Delivery Network (BSPDN) has emerged as a transformative solution [3-4]. However, it complicates thermal management due to extreme substrate thinning, which eliminates the bulk silicon's lateral heat spreading, thereby intensifying localized hotspots [5-6].

In this paper, we perform a power, performance, area and thermal (PPAT) evaluation of a 7nm 3D Memory-on-Logic ARM Cortex-M7 CPU with Frontside Power Delivery

Network (FSPDN) and BSPDN. We employ a spatial-temperature-aware methodology to quantify the critical trade-off between BSPDN's superior power integrity and its thermal degradation. We demonstrate that with rigorous design, the superior power integrity of BSPDN can outweigh its thermal costs, achieving a net performance gain of $\sim 30\%$.

II. CHIP DESIGN AND EVALUATION SETUP

A. Architecture & Physical Design

The physical implementation is based on a 3D Memory-on-Logic architecture, featuring an 8 Mbit SRAM die stacked on a 7nm ARM Cortex-M7 logic die through high-density hybrid bonding. This architecture is realized using a pseudo-3D design methodology [7] to co-optimize the memory and logic modules across two tiers. The timing signoff flow incorporates the impact of coupled power-thermal effects, with the Worst Negative Slack (WNS) as a key metric.

B. Thermal Evaluation Framework

Thermal performance is evaluated via finite element analysis (FEA) utilizing an effective anisotropic thermal conductivity for multilayer stacks [8]. Heat sources are imported from standard-cell-granular power maps of the physical design. In 3D cases, high-thermal-conductivity Si dummy material ($k=140\text{ W/m}\cdot\text{K}$) are modeled to enhance lateral heat spreading. Realistic cooling boundaries are applied: single-phase jet cooling ($\text{HTC}=8\text{E}+4\text{ W/m}^2\cdot\text{K}$) at the top and package contact ($\text{HTC}=20\text{ W/m}^2\cdot\text{K}$) at the bottom.

C. IR Drop Evaluation Framework

We employ a calibrated in-house tool [4] to quantify power integrity. Interconnect resistance is iteratively updated based on 3D temperature profiles using the temperature-dependent model:

$$R_T = R_0 \times (1 + \alpha \times (T - T_0))$$

where $T_0=25^\circ\text{C}$, $\alpha=0.00393/^\circ\text{C}$ is the temperature coefficient of copper.

†Co-first authors with equal contributions.

*Corresponding authors.

Acknowledgement: This work was supported by the National Natural Science Foundation of China (Grant No. 62474007). Appreciate Huawei Cheng and Binze Ma from Huawei for their valuable technical discussion.

III. EVALUATION RESULTS AND ANALYSES

A. 2D and 3D CPU Physical Design

Fig. 1 depicts the schematics and floorplans of the 2D CPU and the 3D CPU with FSPDN and BSPDN. By customizing the SRAM layout for hybrid bonding and re-implementing the physical design, the 3D integration achieves an intrinsic 11% reduction in critical path delay compared to the 2D baseline. The timing impact of the additional memory die routing and hybrid bonds is negligible due to the central placement of critical SRAM cells.

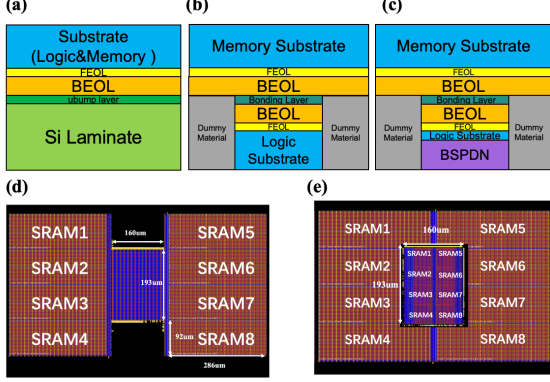


Fig. 1. (a)-(c) Cross-sectional schematics of 2D, 3D FSPDN, and 3D BSPDN CPUs. (d)-(e) Floorplans illustrating the 2D layout and the 3D Memory-on-Logic vertical stacking.

B. CPU Temperature Distribution

As shown in **Fig. 2**, compared to the 2D baseline, the 3D FSPDN configuration effectively limits peak temperature rise to $\sim 4^\circ\text{C}$ by leveraging the high-conductivity Si dummy. However, the 3D BSPDN introduces significant thermal challenges, increasing the peak temperature by $\sim 8^\circ\text{C}$ versus the 3D FSPDN, with a total increase of $\sim 12^\circ\text{C}$ compared to the 2D case. This is primarily due to the extreme substrate thinning (from $40.074\ \mu\text{m}$ to $1.28\ \mu\text{m}$) required for backside connectivity, which severely impedes lateral heat conduction. Specifically, the logic die in 3D BSPDN exhibits intensified hotspots with a $5\times$ increase in in-plane temperature variation ($\sim 10^\circ\text{C}$ vs. $\sim 2^\circ\text{C}$ in 3D FSPDN).

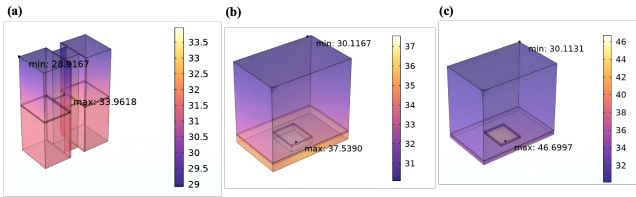


Fig. 2. Temperature distribution of (a) 2D CPU, (b) 3D FSPDN CPU, and (c) 3D BSPDN CPU.

C. Chip IR Drop Distribution

The 3D BSPDN logic die demonstrates superior power integrity, reducing the 98th percentile IR drop to $11.5\ \text{mV}$ —a $12\times$ improvement over the 2D case and a $7.7\times$ improvement over 3D FSPDN (**Fig. 3** a-c). For the top SRAM die (**Fig. 3** d-e), BSPDN also achieves a $2\times$ reduction in IR drop. Critically, thermal-electrical coupling exacerbates the worst-case IR drop by $\sim 9\%$ in 3D BSPDN (significantly higher than the $2\text{--}4\%$ observed in FSPDN). This reveals that neglecting thermal

effects leads to significant inaccuracies in BSPDN, necessitating coupled electro-thermal evaluation.

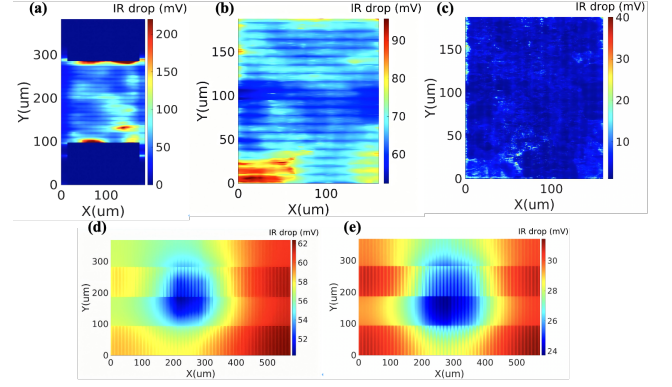


Fig. 3. IR drop distribution heatmaps at 25°C . (a)-(c) Logic die of 2D, 3D FSPDN, and 3D BSPDN CPUs. (d)-(e) SRAM die of 3D FSPDN and 3D BSPDN CPUs.

D. PPA Characterization Considering Thermal-Power Integrity Co-analysis

We employ a spatial-temperature-aware timing analysis that maps thermally coupled IR drops to standard cells on the critical path. As summarized in **Fig. 4**, our approach provides accurate evaluation results, while global IR statistics yield misleading projections (16.7% for average IR drop and 76.3% for 98th percentile IR drop). The 3D FSPDN configuration shows a negligible $\sim 5\%$ frequency gain over the 2D baseline, limited by TSV voltage loss. In contrast, 3D BSPDN achieves a significant 29.6% net improvement over the 2D baseline. This confirms that BSPDN's power integrity benefits decisively outweigh its thermal penalties, leading to a net performance gain.

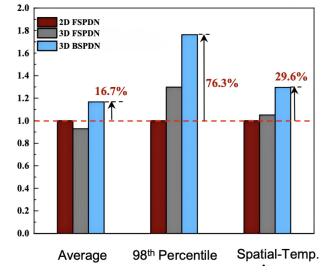


Fig. 4. Normalized CPU frequency improvements evaluated under different IR drop metrics, illustrating the disparity between global statistics (Average, 98th percentile) and the proposed Spatial-Temperature-Aware analysis.

IV. CONCLUSION

This work presents a comprehensive PPAT analysis of a realistic 3D Memory-on-Logic CPU with FSPDN and BSPDN. We quantified the inherent trade-off in BSPDN: while the extreme substrate thinning in BSPDN exacerbates thermal dissipation, its superior power integrity decisively outweighs these thermal penalties. By utilizing a rigorous spatial-temperature-aware timing analysis, we demonstrate that the 3D BSPDN architecture achieves a net frequency improvement of $\sim 30\%$ over the 2D baseline (contrast to the negligible gain in FSPDN). These findings validate BSPDN as a viable solution for overcoming 3D IC bottlenecks and highlight the necessity of coupled electro-thermal modeling in future design flows.

REFERENCE

- [1] M. Min and S. Kadivar, "Accelerating Innovations in the New Era of HPC, 5G and Networking with Advanced 3D Packaging Technologies," 2020 International Wafer Level Packaging Conference (IWLPC), San Jose, CA, USA, 2020, pp. 1-6.
- [2] V. F. Pavlidis, I. Savidis, and E. G. Friedman, *Three-dimensional Integrated Circuit Design*. Morgan Kaufmann, 2017.
- [3] R. Chen et al., "Design and optimization of SRAM macro and logic using backside interconnects at 2nm node," in Proc. IEEE Int. Electron Devices Meeting (IEDM), San Francisco, CA, 2021, pp. 22.4.1-22.4.4.
- [4] R. Chen et al., "Backside PDN and 2.5D MIMCAP to Double Boost 2D and 3D ICs IR-Drop beyond 2nm Node," 2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits), Honolulu, HI, USA, 2022, pp. 429-430.
- [5] R. Chen et al., "Power, Performance, Area and Thermal Analysis of 2D and 3D ICs at A14 Node Designed with Back-side Power Delivery Network," 2022 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2022, pp. 23.4.1-23.4.4.
- [6] F. F. Xie, S. H. Liu, and T. Wei, "Thermal analysis of dual-sided cooling for backside power delivery networks (BSPDN) on 2.5D glass/silicon interposer package," in Proc. IEEE Intersociety Conf. Therm. Thermomech. Phenom. Electron. Syst., 2024, pp. 1-6.
- [7] L. Bamberg et al. 2020. Macro-3D: A physical design methodology for face-to-face-stacked heterogeneous 3D ICs. In Proceedings of the Design, Automation and Testing Europe Conference.37-42.
- [8] L. Wang et al., "Power and Thermal Integrity Analysis of High Performance and Low Power CPUs at Sub-2nm Node Designed with Various Advanced Backside PDNs," 2024 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2024, pp. 1-4.