Late Breaking Results: Thermal Feasibility of Backside Integrated LDOs in 2.5D/3D System-in-Package Using Nanosheet Technology

Yukai Chen, Subrat Mishra, Julien Ryckaert, Dwaipayan Biswas, James Myers
System Technology Co-Optimization Program, IMEC, Leuven, Belgium
Email: yukai.chen@imec.be

***Abstract*—Digital Low Dropout Regulators (LDOs) are an excellent candidate for area-efficient fine-grain power management in heterogeneous systems, leveraging integrated power switches. Relocating the power switches to the backside of the wafer in conjunction with the Backside Power Delivery Network (BSPDN) layer is envisaged as a System Technology Co-Optimization (STCO) booster for finer grain power management and reduced area/cost. We perform a detailed thermal analysis using power-switch-based LDOs enabling per-core DVFS for a high-performance server 3D computing chiplet in a Nanosheet CMOS (A10) technology node with BSPDN. While BSPDN introduces thermal penalties due to a lack of lateral heat spreading, our high-resolution thermal simulations explore the feasibility of moving LDOs to the backside. Increasing the LDO area from 5% to 50% of the backside die area effectively lowers the 2.5/3D System-in-Package (SiP) peak temperature, confirming that thermal concerns do not impede backside LDO integration. This study supports the cost-effective design of next-generation SiPs by demonstrating no adverse thermal impact for relocating power switches to the wafer backside in the nanosheet era.**

I. Introduction

System scaling continues leveraging compute density and integration schemes for higher performance, thanks to advanced CMOS technology like nanosheet (A10) and 2.5D/3D packaging innovations. Traditional System-on-Chip (SoC) designs are evolving into chiplet-based System-in-Package (SiP) [1], enhancing modularity in terms of technology choices for compute, memory, interconnect, and power management. One promising technique is the Backside Power Distribution Network (BSPDN) [2], utilizing the backside of the wafer for power routing, reducing front-side congestion, improving signal integrity, and helping as a scaling booster.

Low Dropout Regulators (LDOs) are an attractive choice for fine-grain on-chip power management when used as power switches and do not necessarily need to be designed in advanced logic nodes. Hence, there is an opportunity to move them to the backside of the wafer along with BSPDN, which helps drive multiple potential advantages: a) more area on the backside allows adding more power switches for better voltage regulation, b) helps in cost reduction by not manufacturing them in advanced logic and c) frees up Front End of Line (FEOL) resources which could be used for adding more advanced logic components and hence act as a scaling booster. The introduction of BSPDN induces a thermal penalty due to the reduced lateral spreading from silicon die thinning [2]. Linear regulators, such as LDOs, often face efficiency challenges when the input voltage significantly exceeds the regulated output, leading to substantial heat dissipation. Thus, integrating the LDO with the BSPDN layer may further compound thermal issues due to cross-heating.

Therefore, this study focuses on the thermal analysis of integrating LDOs into the BSPDN layer for a high-performance server system that has been disaggregated from a traditional 2D System-on-Chip (SoC) into a 2.5D/3D chiplet-based System-in-Package (SiP) using nanosheet A10 technology. The SiP comprises a 2D I/O chiplet and a 3D stacked Memory-on-Logic (MoL) computing chiplet. For our thermal analysis, we consider LDOs decoupled from the FEOL and relocated to the BSPDN layer of the 3D MoL computing chiplet. High-resolution thermal simulations are performed with the LDO occupying 5%, 25%, and 50% of the BSPDN layer area. Our results confirm that LDOs distributed on the backside over a large area do not exacerbate the thermal impact of the BSPDN and can, to a certain extent, function as a heat spreader.

This study, therefore, lays the foundation for functionalizing the wafer backside along with the power delivery network in an A10 CMOS node from a thermal perspective and shows a pathway for improved power management in complex SiPs of the future with multiple power domains.

II. System and Thermal Modeling

Figure 1.(b) illustrates the 2.5D chiplet-based SiP [3] investigated in this work, disaggregating from a classic 2D HPC server SoC shown in Figure 1.(a), and consisting of a 2D I/O chiplet and a 3D computing chiplet with MoL implementation indicated in Figure 1.(b). The 3D computing chiplet, the focus of this work, comprises compute clusters, having four cores each, with integrated LDOs on each core's top and bottom shown in Figure 1.(b). Figure 1.(c) shows the three active power layers: memory, logic, and BSPDN layers containing the LDO. LDO is integrated into the BSPDN of the 3D computing chiplet, forming the basis for reducing power density and conducting thermal analysis over the LDO area sweep (5% - 50%), as discussed in the previous section.

Figure 1.(d) depicts the thermal model of the SiP. Figure 1.(e) shows the detailed 3D chiplet stack layers, excluding the 2D I/O chiplet, which is the same in different cases examined in this work. The three cases are as follows:

Case 1: The core and LDO are integrated into the FEOL with a Front-side Power Distribution Network (FSPDN), using face-to-face (F2F) bonding, representing the baseline design.

978-3-9826741-0-0/DATE25/© 2025 EDAA

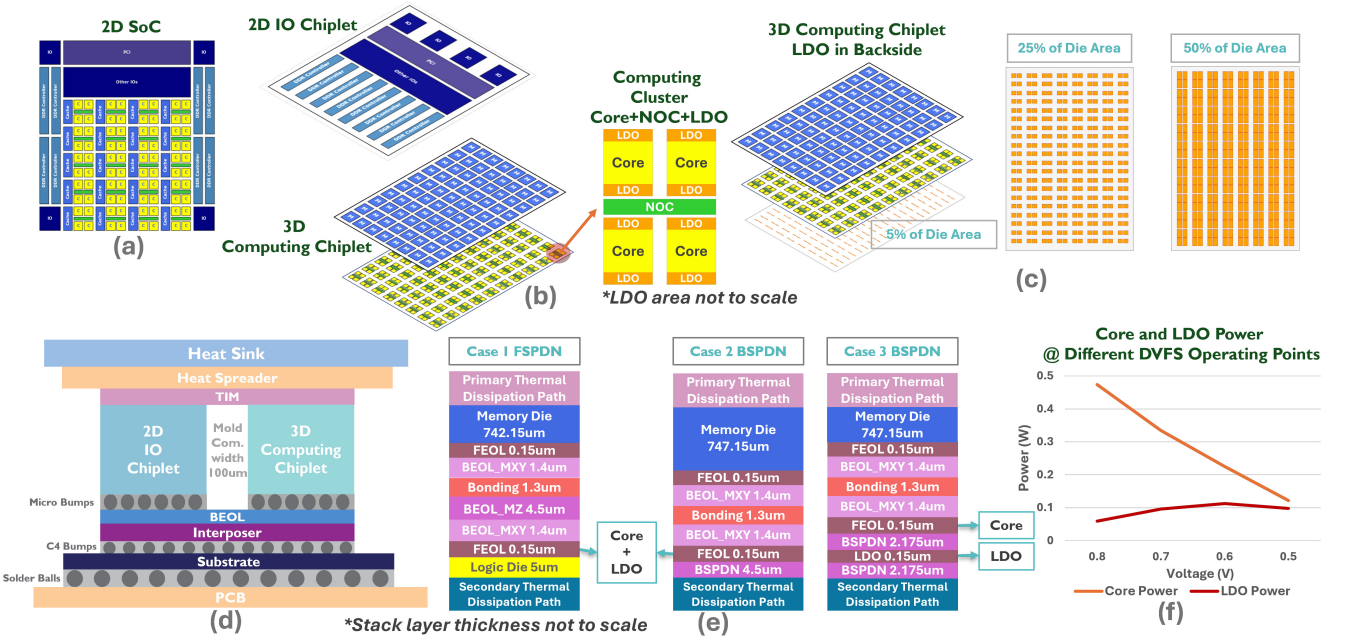

Figure 1. (a) 2D HPC Server SoC; (b) Two chiplet's floorplans in SiP, amplified a single computing cluster featuring core and LDO; (c) LDO decoupled from the core and relocated to the backside with area exploration; (d) Thermal model for the chiplet-based SiP; (e) Detailed stacked layers of the 3D computing chiplet from various case studies; (f) Power consumption of the core and LDO at different DVFS points.

Case 2: The F2F 3D bonding with BSPDN integration, but the core and LDO are still in the FEOL layer, to assess the impact of the BSPDN layer without relocating the LDO.

Case 3: The LDO is decoupled from the core and put in the middle of the BSPDN layer, occupying varying percentages of the die area, to evaluate the thermal effects of LDO allocation.

III. Results and Conclusion

A. Simulation Setup

We performed high-resolution thermal simulations with the core's non-uniform power maps through an adaptive global-local methodology leveraging HotSpot [4]. Material and stack properties were set to standard industry values [3]. The boundary conditions were consistent with 25 °C ambient temperature and convective heat transfer coefficients of $2500\,\mathrm{W/m^2K}$ for the heat sink and $200\,\mathrm{W/m^2K}$ for the PCB, replicating realistic cooling conditions in HPC environments.

The non-uniform core power map was obtained using commercial EDA tools and in-house A10 nanosheet technology PDK, from post-routed netlist annotated for the Dhrystone workload at different DVFS frontier points ([0.8V, 2.6GHz], [0.7V, 2.3GHz], [0.6V, 2.0GHz], [0.5V, 1.3GHz]). The LDO power was determined by assuming an external supply voltage of 0.9V and equal current flow through the LDO and core at different DVFS points as indicated in Figure 1.(f).

B. Thermal Simulation Outcomes and Conclusion

Thermal simulations at various DVFS points provide insights into the impact of LDO placement and BSPDN integration on peak temperatures. Table I summarizes the peak temperatures for the three cases described in the previous section. Introducing the BSPDN layer in Case 2 results in higher peak temperatures across all DVFS points w.r.t. Case 1 due to the additional thermal resistance impeding heat dissipation from the core and LDO, confirming BSPDN thermal penalties.

Table I
Peak Temperatures (°C) at Different DVFS Frontier Points

DVFS	Case1	Case2	Case3 5%	Case3 25%	Case3 50%
0.5V&1.3GHz	72.68	72.96	74.61	73.33	72.67
0.6V&2.0GHz	88.46	89.20	90.88	89.96	89.04
0.7V&2.3GHz	102.83	104.2	104.93	103.64	103.02
0.8V&2.6GHz	120.23	122.39	121.46	121.19	120.73

However, in Case 3, where the LDO is decoupled from the core and relocated to the BSPDN layer, the thermal issues due to BSPDN integration are effectively mitigated with increasing LDO area. At 5% LDO area allocation, peak temperatures in Case 3 are higher than Case 2 across most DVFS points; however, increasing to 50% reduces peak temperatures compared to Case 2 at all DVFS points. This demonstrates that allocating more area to the LDO helps mitigate BSPDN thermal issues. Enlarging the LDO's area decreases its power density, improves heat dissipation, and lowers peak temperatures. These findings support the initial hypothesis that having a large number of LDOs on the wafer backside, deemed beneficial for power management, does not further exacerbate thermal issues of BSPDN. Strategically placing LDOs in the BSPDN layer can help harness the advantages of backside processing for power and thermal management in future systems.

References

[1] G. Van der Plas and E. Beyne, "Design and technology solutions for 3D integrated high performance systems," in *VLSI*. IEEE, 2021.

[2] M. Lofrano *et al.*, "Block level and package level thermal assessment for back side power delivery network." in *ECTC*. IEEE, 2024.

[3] Y. Chen *et al.*, "Thermal implications in scaling high-performance server 3D chiplet-based 2.5D SoC from FinFET to Nanosheet," in *ISVLSI*. IEEE, 2024.

[4] R. Zhang, M. R. Stan, and K. Skadron, "Hotspot 6.0: Validation, acceleration and extension," in *University of Virginia, Tech. Rep*, 2015.