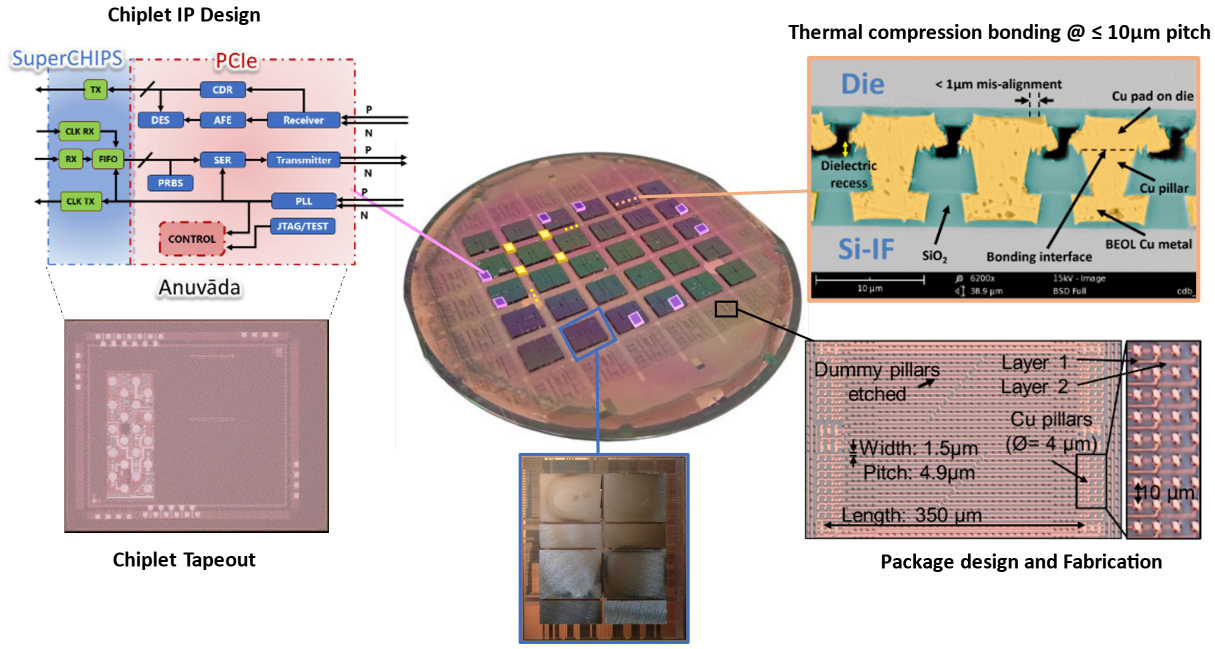
High – throughput wafer – scale dielet assembly schemes for Si-IF

The cornerstone of wafer-scale assembly on the Silicon Interconnect Fabric (Si-IF) is dielet-to-wafer thermal compression bonding (TCB). At UCLA-CHIPS, we have developed an innovative, scalable, and reliable bonding technique to integrate dielets with copper (Cu), gold (Au), or aluminum (Al) pads at sub-10 µm pitch onto copper pillars on the Si-IF package substrate. This bonding process utilizes optimized metal-to-metal thermal compression, achieving military-grade bonding strength and electrical reliability. The process supports throughputs of up to 1,100 units per hour, making it highly efficient for industrial applications.

Research within the Si-IF program at UCLA-CHIPS extends beyond bonding techniques to include the study of defect mechanisms and strategies to mitigate them, application of assembly passivation as well as JEDEC reliability standards for bonded assemblies. Furthermore, these insights are being applied to package design, signal integrity analysis, dielet-to-dielet protocol implementation, and chip design. This comprehensive approach addresses and resolves complex design-and-fabrication co-optimization challenges, paving the way for advancements in high-performance, wafer-scale packaging technologies.

More details of our work are found at our publication link: https://www.chips.ucla.edu/journals ; https://www.chips.ucla.edu/conferenceProceedings

PhD students: Krutikesh Sahoo, Vineeth Harish, Jui-Han Liu, Samuel Wang
Master's students: Tanmay Konnur

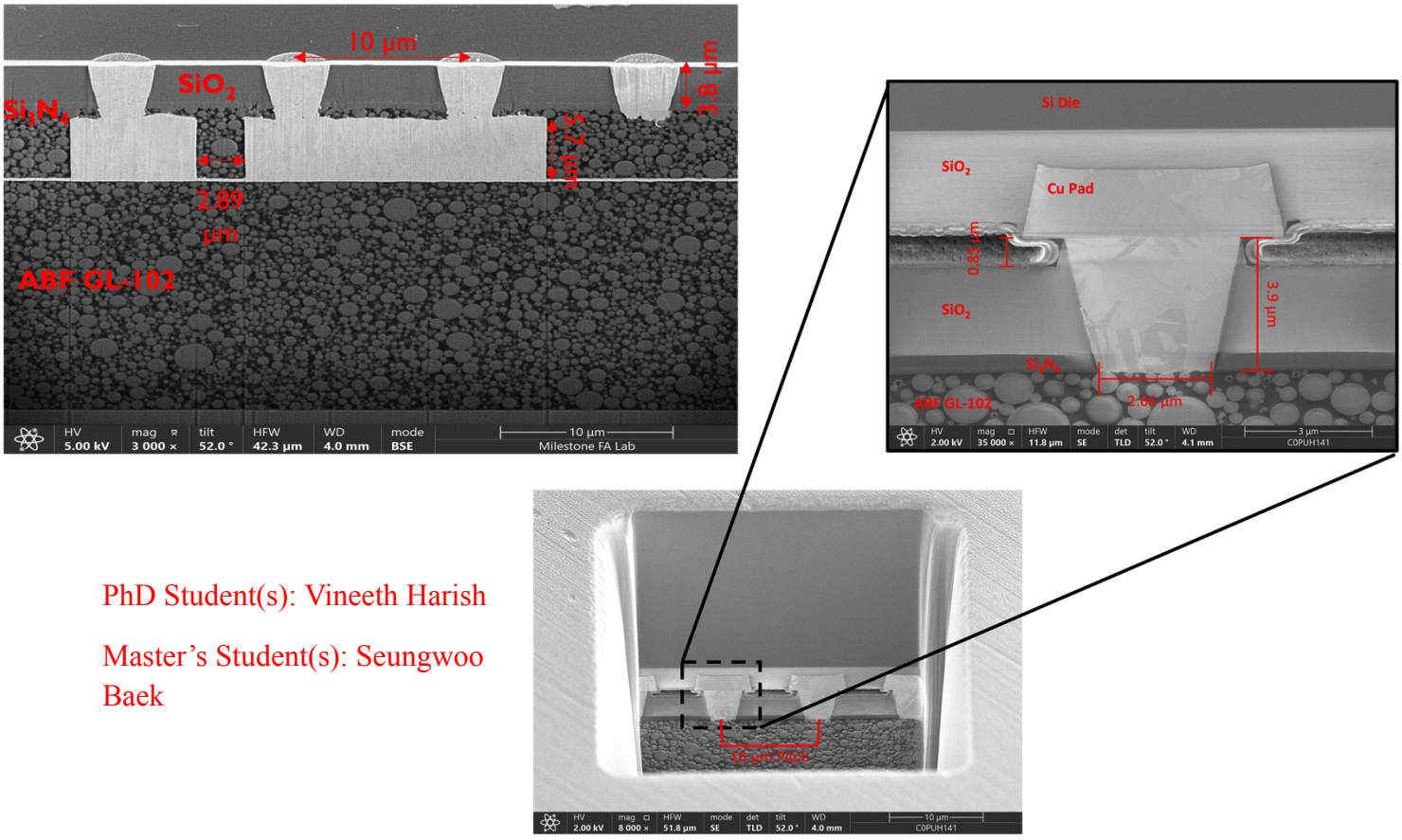


Heterogeneous dielet assembly on Silicon Interconnect Fabric (Si -IF)

Hybrid and Spin-on Polymer Wiring Layers for Si-IF

As the demands for high-performance computing increase, advanced packaging platforms have pushed to reduce the assembly pitch between the die and the substrate to increase the number of I/Os to meet bandwidth requirements. While organic polymer wiring layers have struggled to scale down below the 10 µm range, we achieved this reduction by utlizing CMOS back-end-of-line (BEOL) wiring techniques. By integrating inorganic dielectric layers (SiO_2) on top of organic polymer layers, we have enabled sub-10 µm pitch through damascene BEOL techniques in SiO_2 and solderless Cu-Cu thermal compression bonding (TCB) between the die and the substrate.

Alternative low loss organic polymer dielectrics materials are also being explored for the wiring layers of the Silicon-Interconnect Fabric (Si-IF). Instead of laminating a dielectric layer onto the substrate, spin-on polymer dielectric is utilized to enable precise adjustment of the dielectric material's thickness allowing for improvements in fabrication ultra fine dry etched vias as an alternative to coarser laser drilled vias. Ongoing research in this area includes developing layer transfer techniques to bond and release the fabricated polymer layers to another substrate to improve yield and manufacturability.



PhD Student(s): Vineeth Harish

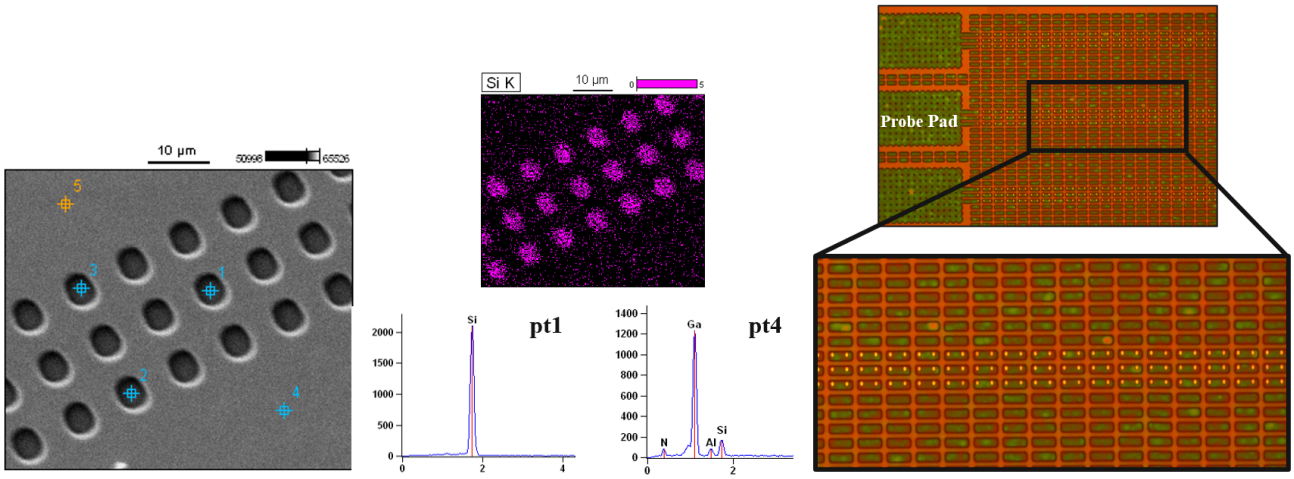
Master's Student(s): Seungwoo Baek

Through GaN Strata Via for 3D Heterogeneous Strata Integration

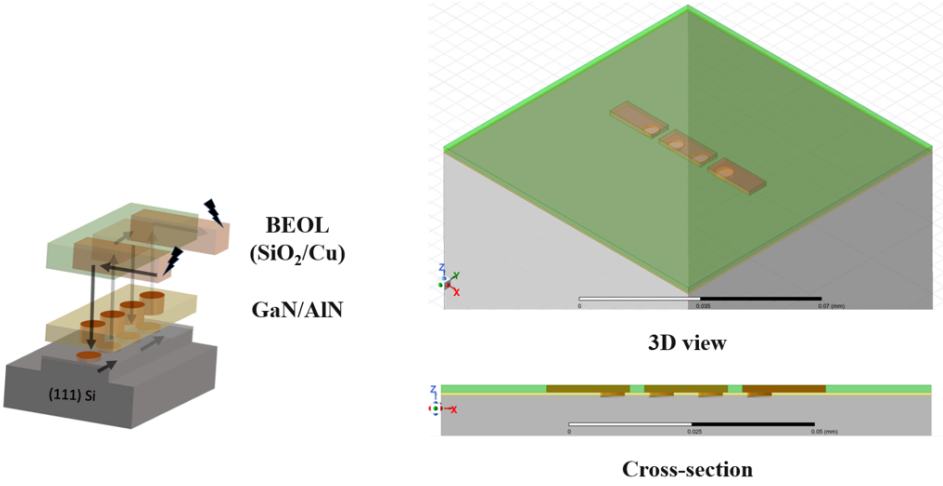
The development of a through-GaN/AlN via process offers a solution for enabling vertical connectivity in GaN and Si heterogeneous integration. GaN-on-Si platforms combine the high breakdown voltage and fast switching properties of GaN devices with the scalability and cost-effectiveness of silicon technology, making them ideal for high-power and high-frequency applications. However, achieving efficient interconnects across GaN/AlN strata layers poses significant challenges due to the material and process complexities involved.

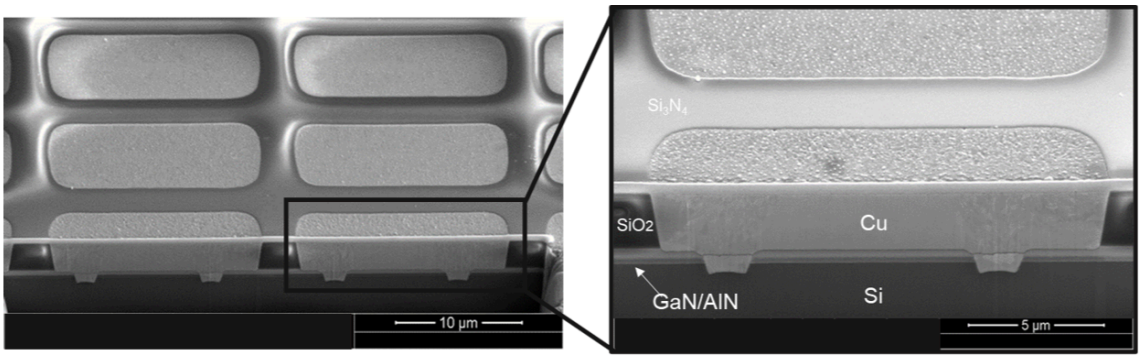
Our process utilizes inductively coupled plasma reactive ion etching (ICP-RIE) for precision etching of GaN and AlN layers, followed by a Cu damascene process to create robust vertical vias. This approach supports high interconnect density with minimum via diameters of 2 µm and pitches as small as 5 µm, ensuring reliable electrical connections between the Cu/SiO_2 back-end-of-line (BEOL) on the GaN/AlN strata and the underlying Si wafer. Process techniques, such as PECVD oxide liner deposition and chemical mechanical polishing using Si_3N_4 polish stops, enhance both the precision and scalability of the process. To validate the process, test structures featuring a 3D daisy chain network were fabricated and observed using focused ion beam scanning electron microscopy (FIB-SEM). The results demonstrate a uniform and conductive through-strata interconnect, making this process a promising pathway for integrating GaN and Si devices within next-generation heterogeneous systems.

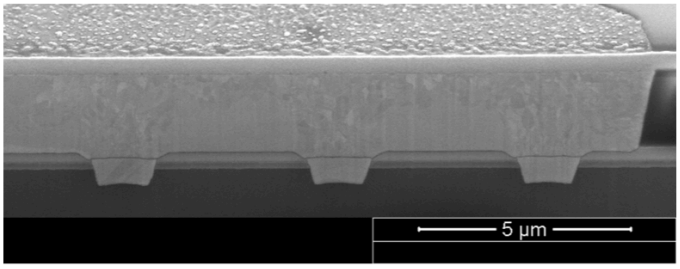
PhD student: Jui-Han Liu, Haoxiang Ren



SEM & EDS analysis of 5 µm diameter through GaN/AlN via, and top view image after Cu damascene with patterned PECVD SiO_2/Si_3N_4, before BEOL Cu filling.







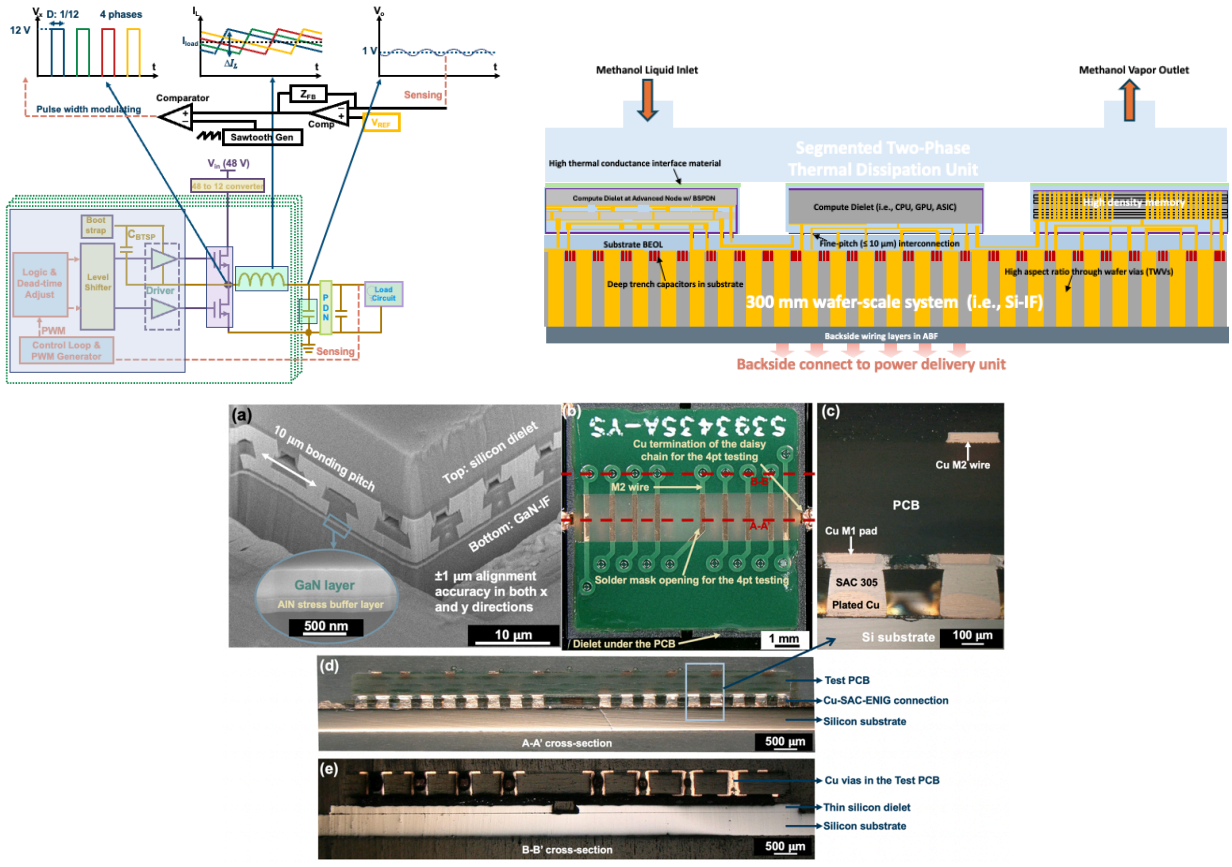
Schematic and model through-strata via test structure, and FIB-SEM cross-sectional image of 2 µm through-GaN-strata via, and the 2 µm via with 5 µm pitch.

Power Delivery and Power Management for Chiplet-based Large Scale Systems

Heterogeneously integrated, large interposer-based systems or wafer-scale systems have emerged as a promising solution for addressing the computational demands of high-performance computing (HPC) and artificial intelligence/machine learning (AI/ML) applications. These systems offer thousands of processing cores and hundreds of GB of memory capacity at moderate bisection bandwidths, making them the workhorse for today's Large Language Model (LLM) training machines and beyond. Their scalability enables the accommodation of hundreds of billions of parameters in compact systems. However, ensuring efficient, uniform, and cost-effective power delivery and power management to support these large-scale systems remains a significant challenge. This research delves into the power delivery complexities and aim to demonstrate solutions to enhance power efficiency, density, integrity, and granularity while at the same time providing for a simple heat dissipation scheme.

Sub-group: Demonstration of 12-1 V Heterogeneously Integrated Voltage Converter (HIVC) using high-voltage silicon or GaN w/ thermal solutions

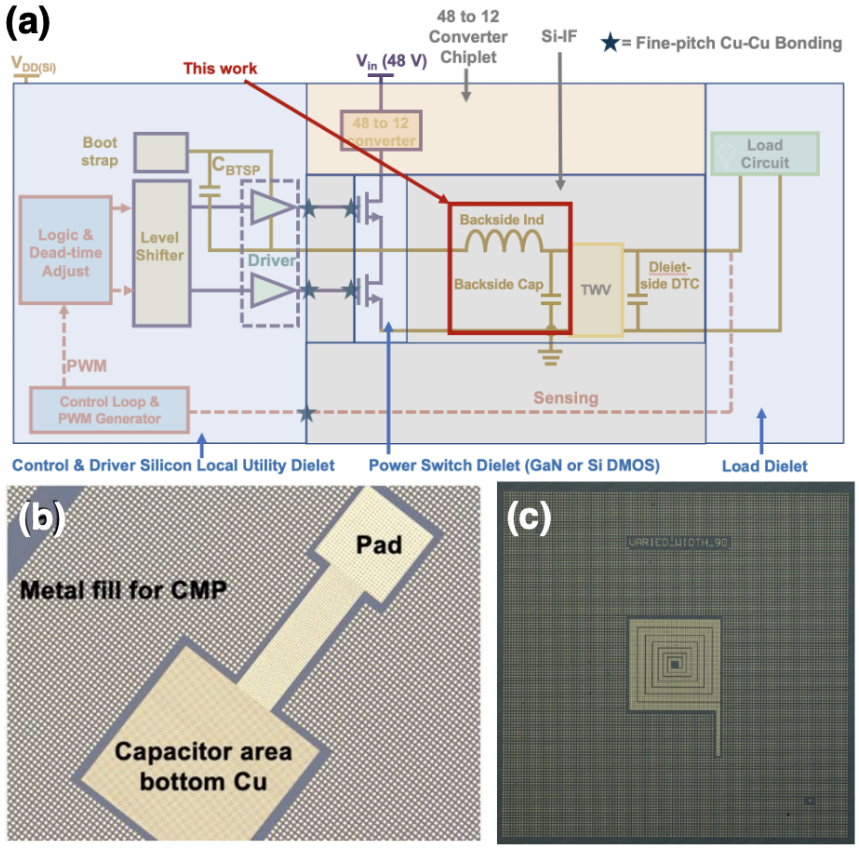
PhD students: Haoxiang Ren, Ziyi Guo, Krutikesh Sahoo, Naarendharan Meenakshi Sundaram
Master students: Ben Yang



Schematics of the HIVC Buck circuits (Top Left); Schematic view of the power and thermal architecture for large-scale systems (Top right); Technology demonstration of the heterogeneous integration at fine-pitch (Bottom)

Sub-group: Design and Fabrication of Integrated Passive Devices in the Si-IF

PhD students: Haoxiang Ren, Jui-Han Liu
Master students: Cheng-Ting Yang, Ben Yang
UG students: Zoe Chen

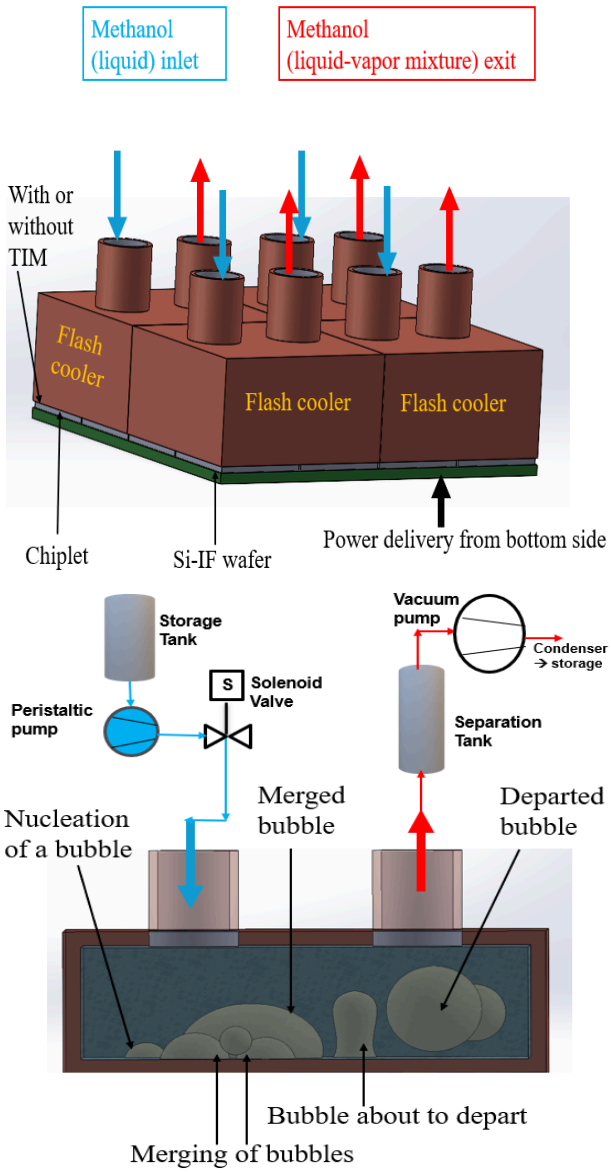


(a) Schematic design of the heterogeneously integrated voltage regulator (HIVR) on the Si-IF, and the IPDs are key components for backside high-density power delivery; (b) a fabricated MIM cap in the Si-IF; (c) a fabricated air-core spiral inductor in the Si-IF.

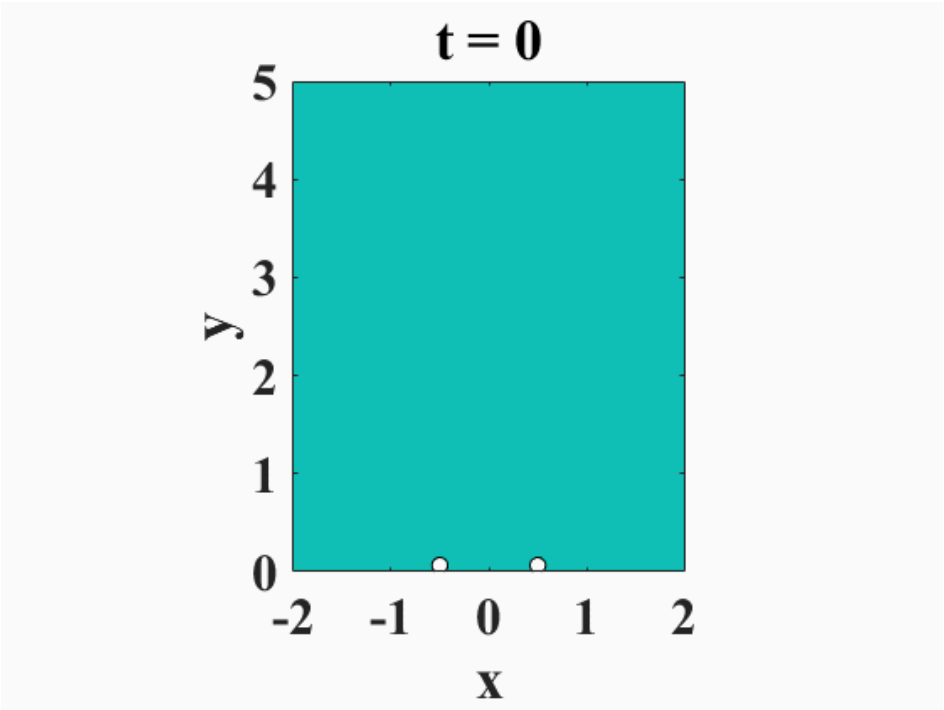
Dynamic Flash Boiling for Efficient Thermal Management of Si-IF

Advanced packaging systems, such as the Silicon Interconnect Fabric, have unique thermal challenges. Such heterogeneously integrated systems generate large amounts of heat that need to be dissipated effectively without any lateral heat spreading. We utilize two-phase boiling heat transfer as the cooling solution for the thermal management of Si-IF. A dynamic “Flash Cooling” system was developed. The flash cooling system utilizes a combination of temperature and pressure-driven boiling of methanol for the effective thermal management of Si-IF. Methanol is chosen as the working fluid due to its preferable boiling point and latent heat of vaporization. The TDU is maintained at sub-atmospheric pressure, and a pulsed flow of the methanol is enabled with the help of a solenoid valve. Once the methanol encounters low-pressure and high superheat conditions inside the TDU, it boils off, taking the heat away.

PhD student: Naarendhara Meenakshi Sundaram, Haoxiang Ren
Masters students: Jiseong Oh, Ben Yang



Section view of the flash cooler (Left), Schematic showing the flash cooler on top of the chiplets (Right); only a part of Si-IF wafer is shown



Pool boiling showing the heat removal by the cyclic formation and departure of the vapor at the heated wall