Summary – PhD Dissertation Defense

Presenter: Gaurav Kothari

Advisor: Prof. Kanad Ghose

Title: Chiplet Based High-Performance Processing Systems

Objective:

Methodologies for exploiting the performance benefits of 3D integration in thermally-aware manner that minimizes high-temperature induced performance throttling.

Proposed Work:

Discussed how it is a difficult challenge to stack multi-core chiplets due to chiplets centers that heat up a lot. The heat generated in the chiplet results in thermal throttling and degrade performance.

Discussed how chiplets are stacked and explained terminologies like TSV (through-silicon vias), bump connections via a ball-grid array (BGA) to the socket. Silicon interposer (SI) provides connectivity to other chiplets in the package.

Evaluated and presented various stacking multi-core x86 chiplets to scale up core counts.

Proposed solution where they first identify CORE and UNCORE sections. The CORE section includes the processor datapath and private L1 and L2 caches. The UNCORE section included the last level cache and an NOC router for communication with other core tiles.'

Processor model and floorplan based on Intel Skylake-SP server and used to demonstrate the three stacking techniques. Used different tools for power modeling, floorplanning.

Three different stacking techniques:

- Naïve stacking
- Thermally-aware multi-core chiplet stacking (TMS)
    - Coarse-grained TMS (TMS-CG)
    - Fine-grained TMS (TMS-FG)

Able to reduce the overall die area by 47% with TMS-CG and with better or equal performance when compared to 2D configuration. 38% for TMS-FG.

Smaller die sizes significantly improve the chiplet yield and reduce overall costs.