Abstract

Vertical integration (3D ICs) has demonstrated the potential to reduce inter-block wire latency through flexible block placement and routing. However, there is untapped potential for 3D ICs to reduce intra-block wire latency through architectural designs that can leverage multiple silicon layers in innovative ways. Furthermore, it is particularly challenging to simultaneously explore the physical design space and microarchitectural space for vertical integration. The physical design space typically has no information on the microarchitectural impact of latency optimization, and the microarchitectural space has no information on the physical design impact of different architectural alternatives.

We make the following contributions in this paper: (1) the introduction of port partitioning, a new approach to constructing multi-layer blocks, (2) the extension of a microarchitectural exploration tool to include the ability to model multi-layer blocks and to consider these blocks as alternative implementations of single layer architectural blocks on the fly, within a single floorplanning run, and (3) the evaluation of vertical integration on a design driver using this framework.

For this design driver, we see an average 36% improvement in performance (measured in BIPS) over a single layer architecture, and a 29% improvement in performance over a multi-layer architecture with single layer blocks. The on-chip temperature is kept below 40 °C.