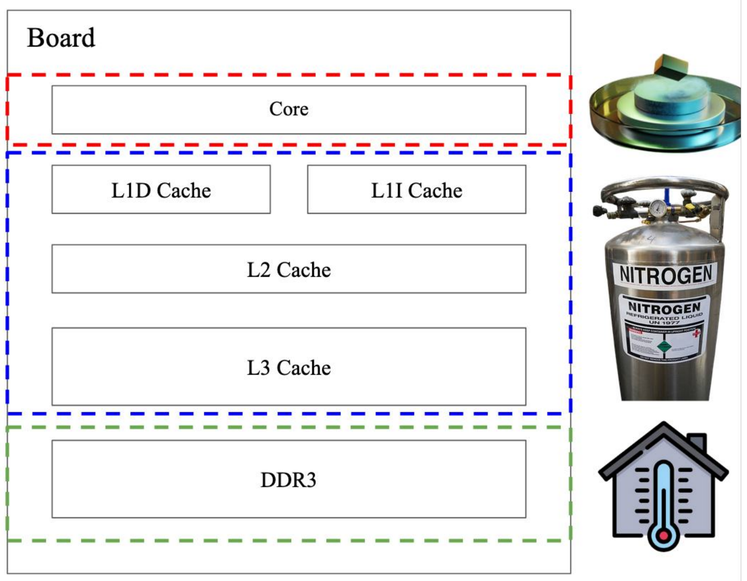
Potential and Limitation of High-Frequency Cores and Caches

Kunal Pai, Anusheel Nand, Jason Lowe-Power
University of California, Davis

Motivation



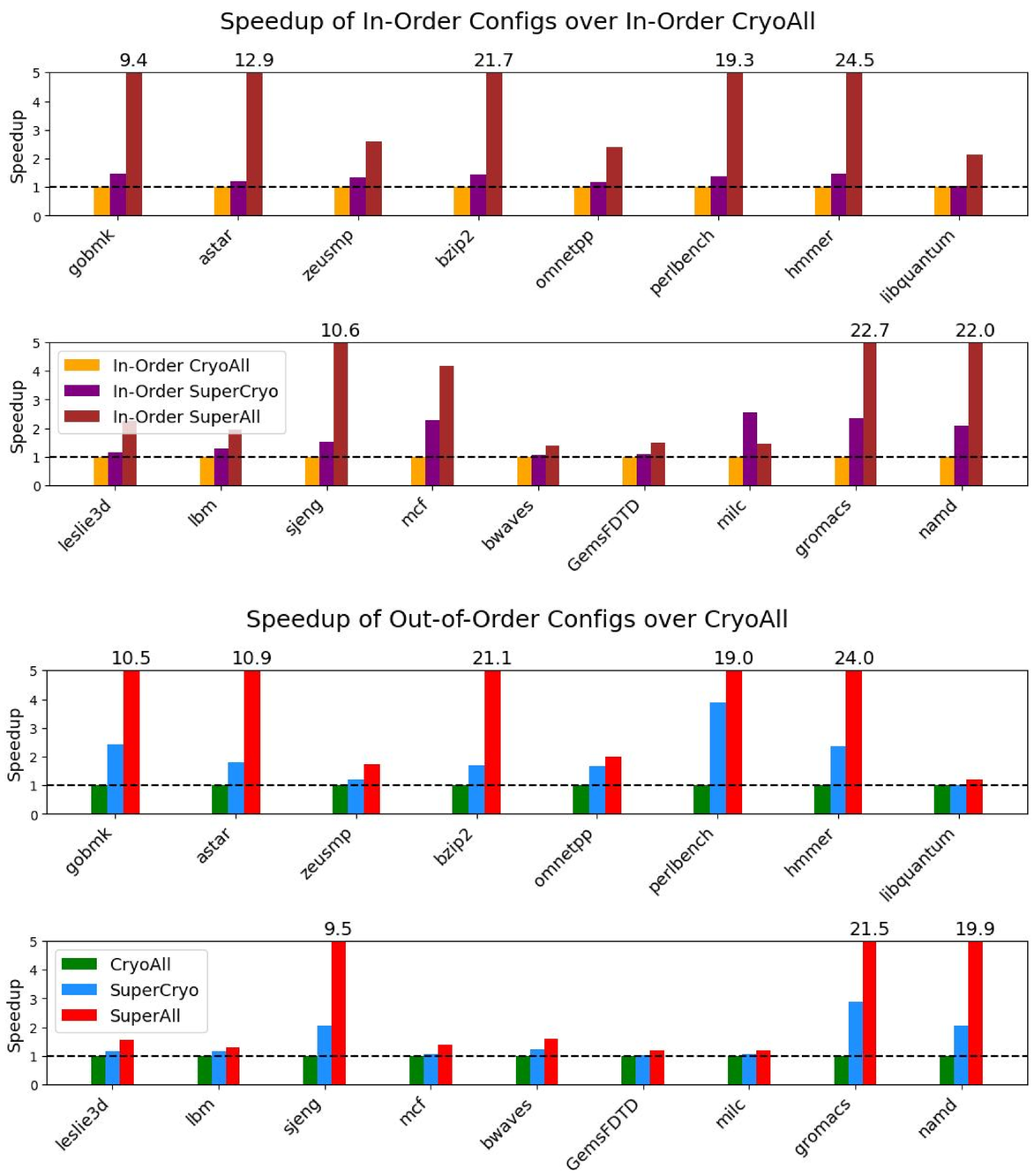
- **GOAL**: Model emerging tech with very high clocks.
- **Cryogenic computing & superconducting circuits**
- **gem5**: viable modeling framework.

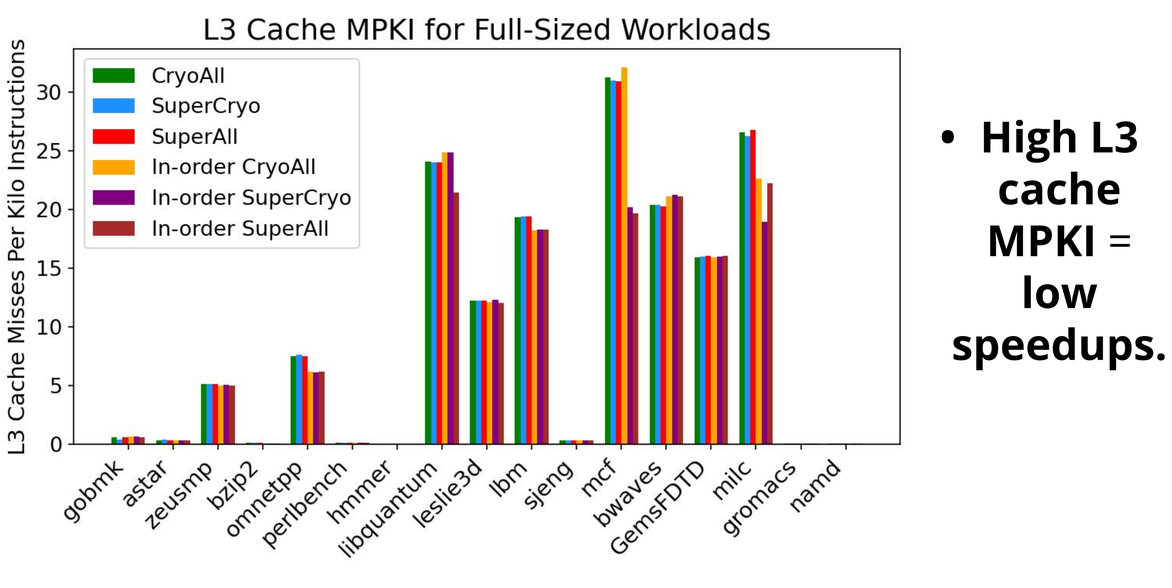
Experimental Set-Up

- Cryogenic components in gem5: BOOM (OOO), HiFive Unmatched (In-Order), 3-level cache hierarchy, all at **4 GHz.**
- Superconducting components: same μarch as cryogenic components, all at **100 GHz.**
- Configs: **CryoAll**, **SuperCryo**, **SuperAll**, **In-Order CryoAll**, **In-Order SuperCryo**, **In-Order SuperAll**.
- **Full-sized, realistic workloads**: SimPoints - **SPEC 2006** ref size.
- **Potential**: high speedups.
- **Limitation**: cost of data movement.

Potential Speedup

- **Bottlenecks**: Room temp. DRAM and low freq. caches.
- Latency hiding less important.
- In-order speedups over In-Order CryoAll > out-of-order speedups over CryoAll.

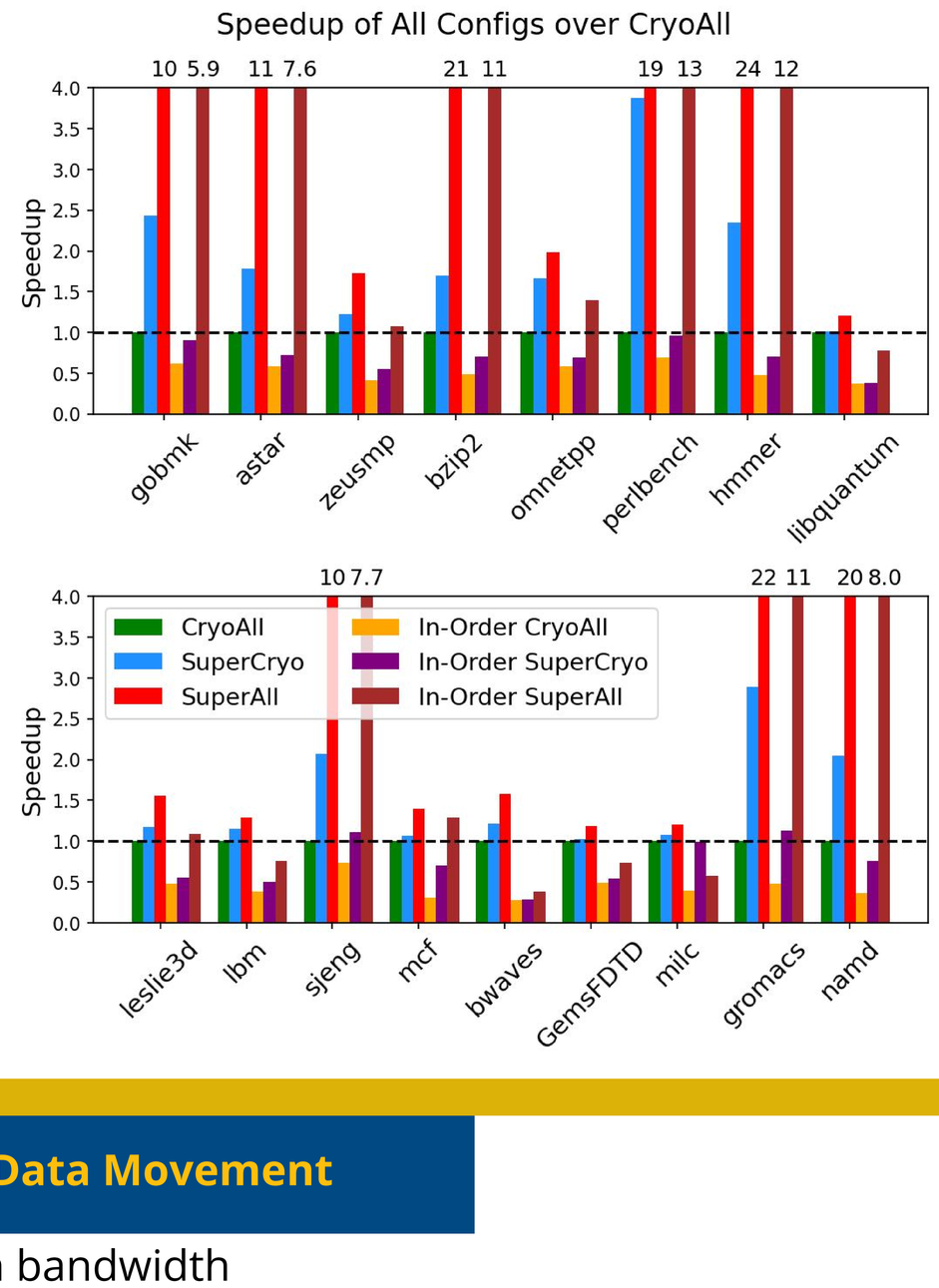




- **High L3 cache MPKI = low speedups.**

In-Order vs Out-of-Order

- Over CryoAll, out-of-order: more speedup.
- CryoAll outperforms In-Order CryoAll and In-Order SuperCryo.
- Big potential benefits, but only for some workloads, need for superconducting **accelerator with low memory traffic.**



Data Movement

- High bandwidth to the cache to actualize speedups.
- **Max. bandwidth**: out-of-order - 500 GB/s ; in-order - 130 GB/s, both for hmmer.