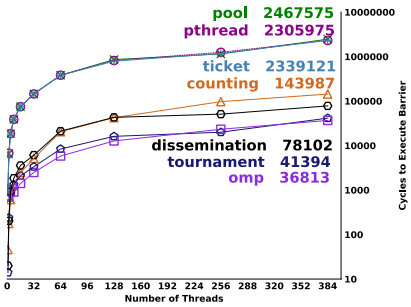
Paths to Fast Barrier Synchronization on the Node

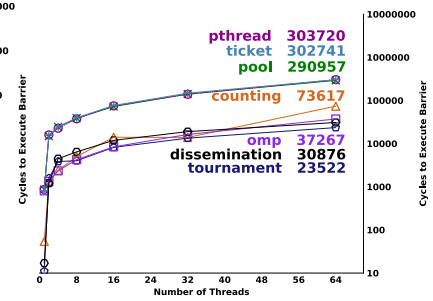
Conor Hetland, Georgios Tziantzioulis, Brian Suchy, Michael Leonard, Jin Han, John Albers, Nikos Hardavellas, and Peter Dinda

Barriers Are Slow

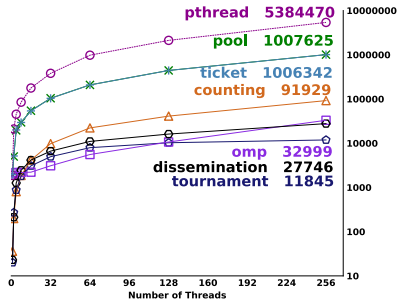
8 Socket NUMA



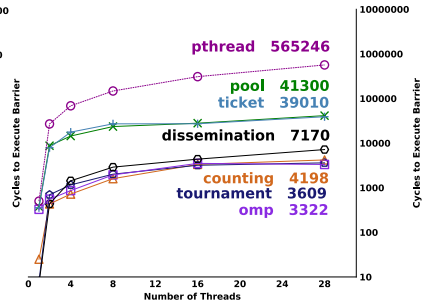
4 Socket NUMA



Phi



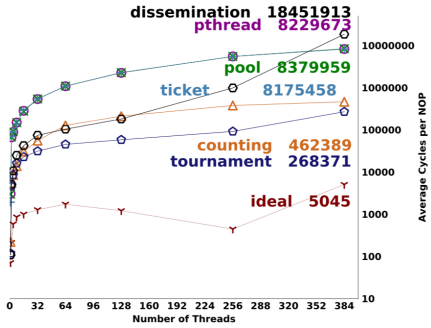
HARP



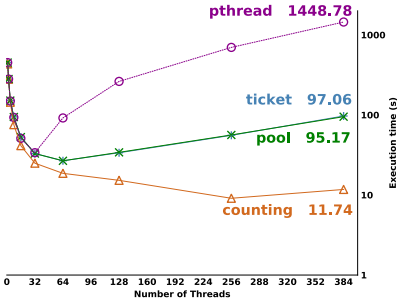
- Software barrier latency is on the order of the tens-of-thousands of cycles
- This is really slow
- A barrier is functionally the logic of an AND gate plus communication; this should take on the order of hundreds of cycles

Barriers Matter

NESL (VCODE) Interpreter



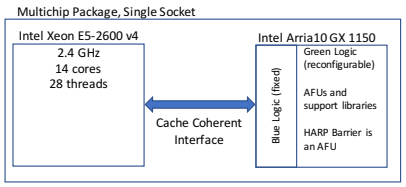
Streamcluster



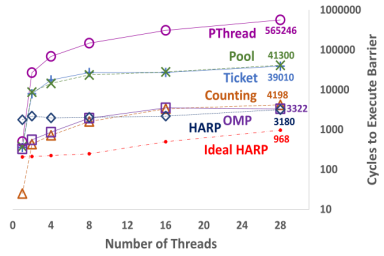
Runtimes and applications with fine-grained parallelism rely heavily on barrier performance

Intel HARP Barrier

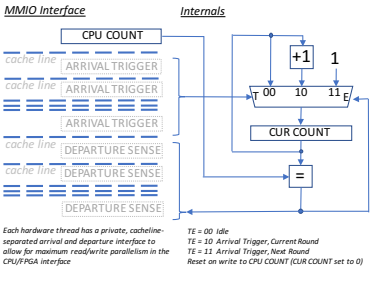
Diagram of the HARP



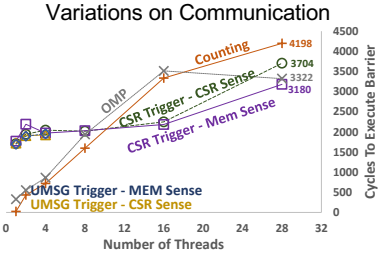
Barrier Performance



Design Outline

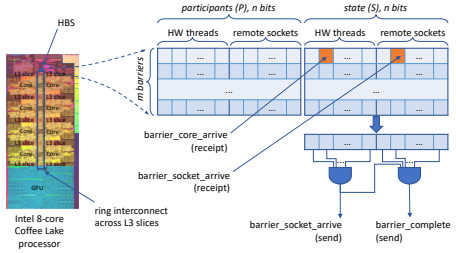


Variations on Communication



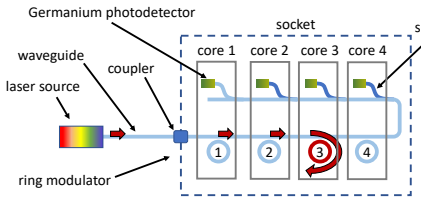
Upon hitting the barrier, each hardware thread in the CPU issues an MMIO Write to an individual cache line on the FPGA. An arrived thread then polls on an individual cache line to see if all threads have arrived at the barrier. Each FPGA cycle, the FPGA adds all reported barrier arrivals to an internal register. When this internal register holds the number of threads allocated to the barrier, the internal register containing the count is reset, and all polled cache lines are written to signal the waiting threads to continue. The hardware can then be reused. The system alternates between two copies of this logic.

Proposed x86 Extension



- Utilizes existing cache coherence network
- 22K barriers in only 0.01% of the Xeon Platinum 8180's die
- Adds two instructions to the ISA: barinit and barwait
- Security maintained with a kernel-administered timeout

Silicon Photonic Barrier



Future silicon photonic interconnects could provide fast barrier synchronization essentially for free

C. Hetland, G. Tziantzioulis, B. Suchy, M. Leonard, J. Han, J. Albers, N. Hardavellas, P. Dinda, "Paths to Fast Barrier Synchronization on the Node", *Proceedings of the 28th ACM Symposium on High-performance Parallel and Distributed Computing (HPDC 2019)*, June, 2019.