

PARALLELISM-AWARE BATCH SCHEDULING

by
Swathy Mudhagouni
Dyuti Tadepalli

Krug Hall, Room 5, April 13,2009, 5:30 PM

ABSTRACT

The main memory system is the major limiter of computer system performance. In modern multicore systems, as the number of on-chip cores increases, the pressure on the DRAM system increases, as does the interference among threads sharing the system. This uncontrolled interthread interference in DRAM scheduling can destroy the Memory level Parallelism (MLP) and serialize requests of individual threads, leading to significant degradation in single-thread and system performance in multicore and multithreaded systems. These problems can be solved by using the Parallelism- Aware Batch Scheduler (PAR-BS), a memory controller that controls and limits interthread interference. A conventional Scheduler can service requests in their arrival order where as a PAR-BS services each thread's concurrent requests in parallel. It preserves single thread memory-level parallelism and follows shortest-job-first with-in batch scheduling which helps improve its performance. It implements the idea of thread ranking and rank-based Scheduling to preserve individual threads' MLP. Batching and thread ranking together result in the large fairness improvements. Even when unfairness isn't a problem, preserving MLP significantly improves single-thread and system performance. PAR-BS is effective with multiple memory controllers, even without any coordination between the different controllers. Also the idea of request batching from I/O systems to provide fairness and starvation-freedom to threads sharing the DRAM system is adapted.