ABP : Predictor Based Management of DRAM Row Buffers

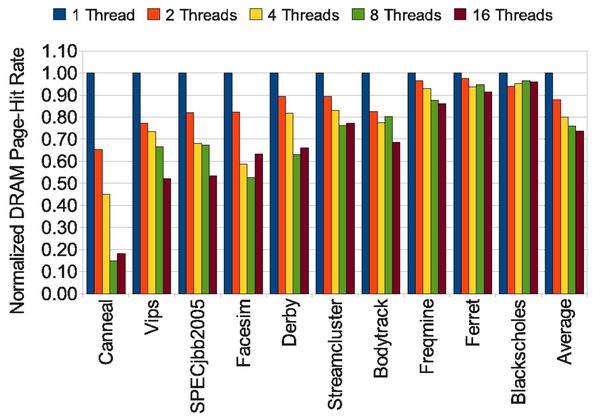
Manu Awasthi, David Nellans, Kshitij Sudan, Rajeev Balasubramonian and Al Davis

UTAH ARCH

THE UNIVERSITY OF UTAH

Background

- DRAM accesses are costly, especially in multicore systems.
- Future CMPs will run a mixed load of workloads/threads.
- Destructive interference at memory controller, spatio-temporal locality lost!



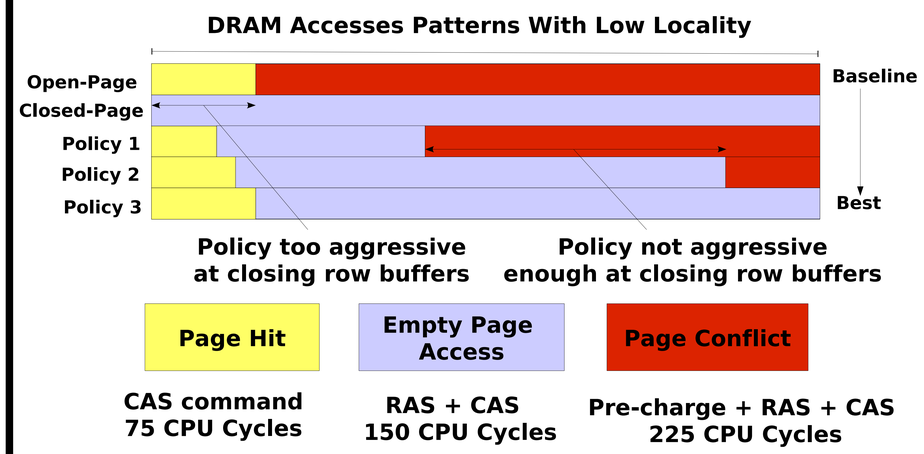
Motivation

- DRAM row-buffer **hits** are least expensive, row-conflicts are most.
- Randomized memory access patterns render traditional row-buffer management policies useless.
- Most commercial CMPs have no buffer management policies implemented [1].
- Timer based policies [2,3] are too coarse-grained to be effective.

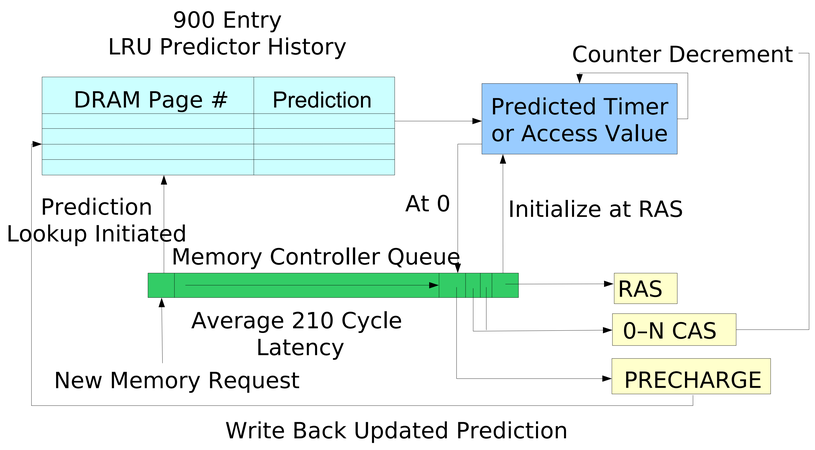
Proposal

- Rather then **time**, focus on **access patterns**.
- Access patterns are predictable, a predictor can accurately predict the number of accesses for which the row-buffer be kept open.
- In any case, can't do worse than a static policy.

Hybrid Row Buffer Management Policies



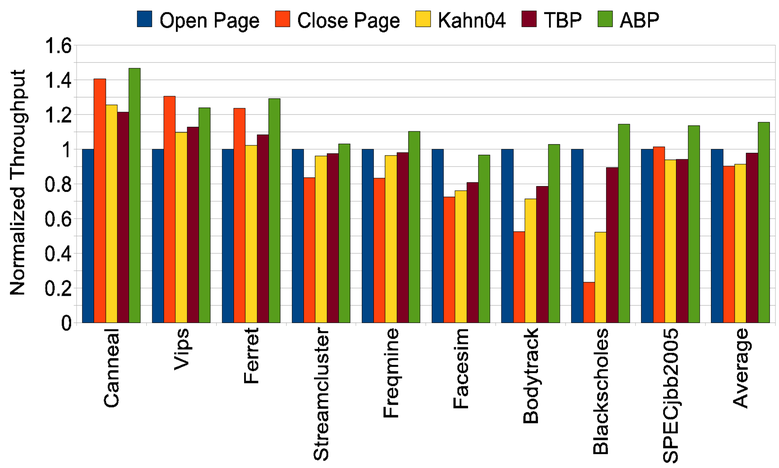
How ABP Works



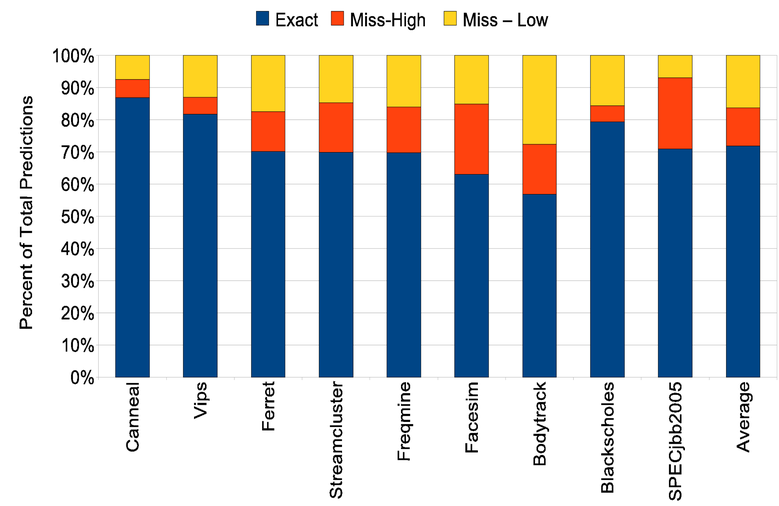
Overheads

- History table is 2048 entry/4 way cache, 20 KB of total storage.
- 1 cycle access latency, 1.39 pJ per access.
- NOT on the critical path.

Results



Throughput Comparison, 8 Core CMP



ABP Accuracy, 8 Core CMP

References

1. Intel Nehalem datasheet.
2. B. Fanning. **Method for Dynamically Adjusting a Memory System Paging Policy,** 2003. United States Patent, Number 6604186-B1.
3. O. Kahn and J. Wilcox. **Method for Dynamically Adjusting a Memory Page Closing Policy**, 2004. United States Patent, Number 6799241-B2