Prashant Singh Chouhan

Simulating and evaluating cache prefetching policies

Study goals:

My main goal is to develop a prefetcher on top of a cache simulator. It will take application's memory trace as an input and then run its simulation. I will use stream buffer to keep prefected blocks from L2 cache. My simulator will implement an application adaptive prefetch buffer depth algorithm which adapts the number of prefetch blocks according to the application's current execution phase. At last,I will evaluate different prefetch policies and plot the results.

Assumptions:

1. Two level of caches- L1 and L2 Cache
2. Prefetcher will fetch from L2 cache to the stream buffer cache.
3. L1 has different Instruction and Data cache while L2 is unified.
4. L1 will be a direct mapped cache

Tools:

I will use the Pin tool to generate the memory and instruction trace for the benchmark. And I will build my prefetcher on top of dineroIV simulator.

Experiments:

I plan to evaluate my simulator for always and miss prefetching algorithm with an adaptive prefetch distance. I will choose 3 different benchmarks and for each benchmark, I will vary the block size, total size and associativity of L2 and L1 (L1 will always be directly mapped) caches to produce 6 different configurations in total.

Evaluation metric:

I will evaluate Prefetching algorithms based on:
1. Prefetch Coverage = Cache misses eliminated by prefetching / Total cache misses
2. Prefetch Accuracy = Cache misses eliminated by prefetching / Total prefetches

Prashant Singh Chouhan

Grading Scheme:

Grade	Implementation task	# of Benchmarks	# of Cache Configs
A	2 prefetch policy	3	6
A+	1 prefetch policy	3	6
B	1 prefetch policy	3	3
B-	1 prefetch policy	2	6
C	1 prefetch policy	2	3

References:

[1] https://en.wikipedia.org/wiki/Cache_prefetching
[2] https://www.usenix.org/legacy/event/fast07/tech/full_papers/gill/gill_html/node11.html
[3] https://compas.cs.stonybrook.edu
[4] https://www.archive.ece.cmu.edu
[5] Jouppi, Norman P. "Improving direct-mapped cache performance by the addition of a small fully associative cache and prefetch buffers." *ACM SIGARCH Computer Architecture News*. Vol. 18. No. 2SI. ACM, 1990.