

CSE 240A Fall 2013 Project 2

Name: Junchuan Wang

PID: A53047208

Due date: Tuesday, 10 Dec 0:00 Am

Cache Simulator

Design consideration:

(1) Explore the maximum parallelism.

L1, Victim Cache, L2 will be accessed in parallel.

So the hit latency one of the threes:

- $\text{Min}(\text{L1_Latency}, \text{Victim_Latency})$
- $\text{Or } \text{L1L2transferLatency} + \text{L2_Latency}$
- 350 ms

(2) Eviction policy is LRU

(3) L1 block will be swapped with Victim cache if Victim cache is a hit. So there can be only one of the two that will get a hit.

(4) The cache is inclusive so if there a block in L1, the same can be found on L2 to keep consistency.

The baseline performance is :

Cache Config:

L1 size=16 KB, assoc=2, LS=32 bytes, lat=3 cyc

Victim size=16384 bytes, assoc=4, LS=32 bytes, lat=5 cyc

L2 size=256 KB, assoc=4, LS=32 bytes, lat=20 cyc

Total size of all caches 294912 bytes

Total CPI=35.10

Average memory access time=89.47

L1 Miss Rate=0.25

Victim Local Miss Rate=1.00

L2 Local Miss Rate=0.98

L2 Global Miss Rate=0.25