Homework assignment #3

Due: Friday, May. 26, 23:59

Total points: **100**

Questions

1. You are trying to appreciate how important the principle of locality is in justifying the use of a cache memory, so you experiment with a computer having an L1 data cache and a main memory (you exclusively focus on data accesses). The latencies (in CPU cycles) of the different kinds of accesses are as follows: cache hit, 1 cycle; cache miss, 105 cycles; main memory access with cache disabled, 100 cycles. (40 points)

   a. When you run a program with an overall miss rate of 5%, what will the average memory access time (in CPU cycles) be? (10 points)

   b. Next, you run a program specifically designed to produce completely random data addresses with no locality. Toward that end, you use an array of size 256 MB (all of it fits in the main memory). Accesses to random elements of this array are continuously made (using a uniform random number generator to generate the elements indices). If your data cache size is 64 KB, what will the average memory access time be? (10 points)

   c. If you compare the result obtained in part (b) with the main memory access time when the cache is disabled, what can you conclude about the role of the principle of locality in justifying the use of cache memory? (10 points)

   d. You observed that a cache hit produces a gain of 99 cycles (1 cycle vs. 100), but it produces a loss of 5 cycles in the case of a miss (105 cycles vs. 100). In the general case, we can express these two quantities as G (gain) and L (loss). Using these two quantities (G and L), identify the highest miss rate after which the cache use would be disadvantageous. (10 points)

2. The LRU replacement policy is based on the assumption that if address A1 is accessed less recently than address A2 in the past, then A2 will be accessed again before A1 in the future. Hence, A2 is given priority over A1. Discuss how this assumption fails to hold when the a loop larger than the instruction cache is being continuously executed. For example, consider a fully associative 128-byte instruction cache with a 4-byte block (every block can exactly hold one instruction). The cache uses an LRU replacement policy. (30 points)

   a. What is the asymptotic instruction miss rate for a 64-byte loop with a large number of iterations? (5 points)

   b. Repeat part (a) for loop sizes 192 bytes and 320 bytes. (5 points)

   c. If the cache replacement policy is changed to most recently used (MRU) (replace the most recently accessed cache line), which of the three above cases (64-, 192-, or 320-byte loops) would benefit from this policy? (10 points)

   d. Suggest additional replacement policies that might outperform LRU. (10 points)

3. Consider a two-level memory hierarchy made of L1 and L2 data caches. Assume that both caches use write-back policy on write hit and both have the same block size. List the actions taken in response to the following events: (30 points)

   a. An L1 cache miss when the caches are organized in an inclusive hierarchy. (10 points)

   b. An L1 cache miss when the caches are organized in an exclusive hierarchy. (10 points)

   c. In both parts (a) and (b), consider the possibility that the evicted line might be clean or dirty (10 points)

Grading

1. You will be given 0 point for any kind of cheating.
2. Please give detailed process that how you solved it. Otherwise, no points will be charged.

Submission

1. Submit your homework (hardcopy) into HW box prepared near room #922 in N1 building.
2. Late submissions will not be accepted. Please keep the submission deadline.

If you have questions about this homework assignment, please send email to TA

(jw.park@kaist.ac.kr, jesung.kim@kaist.ac.kr)