

CS 308 — Homework #5

Due Monday, May 3, in class

1. From the textbook, question 7.20 on page 375. Note that when the book refers to a “32 KB” cache it is referring to the amount of storage for actual data. Space for tag fields, etc. is not included in the 32 KB figure.
2. A cache system has a 95% hit ratio, an access time of 100 ns. on a cache hit, and an access time of 800 ns. on a cache miss (i.e. the delay between issuing the address to the cache and receiving the data from main memory is 800 ns.). What is the effective (average) access time?
3. A cache is being designed for a computer with 2^{32} bytes of memory. The cache will have 2K slots (i.e. it will hold 2K lines of data) and use a 16 byte block (i.e. the line size is 16 bytes).
 - (a) Compute, for both a fully associative cache and a direct-mapped cache, how many total bits the cache will use, in tag, validity, and data bits.
 - (b) Show how the main memory address is partitioned for each of these cases.
4. Better performance is usually found for caches with longer cache lines, say 128 bits, than for caches with the same capacity but with shorter cache lines, like 32 bits. Why is this?? Write down all the reasons you can think of that would explain this behavior.

Pledge that you did this homework on your own.