York University
Dept. of Computer Science and Engineering
CSE4201
HW 3

Problem 1

Consider 2 machines with the same processor and main memory but different cache organizations. Assume that both processors runs at 2GHz, have a CPI of 1, and have a cache miss time of 100 ns. Further assume that writing a 2-bit word to main memory requires 100 ns. (For the write through cache), and that writing a 32-byte block (for the write back cache). The caches are unified – They contain both instructions and data, and each cache has a total capacity of 64KB, not including tags and status bytes.

The cache on system A is a two-way set associative and has 32-byte blocks. It is written through and does not allocate a block on a write miss.

The cache on system B is direct mapped and has 32-byte blocks. It is a write back and does allocate a block on write miss.

1. Describe a program that makes system A runs as fast as possible relative to system B's speed.
2. Describe a program that makes system B runs as fast as possible relative to system A's speed
3. How much faster is the program in (1) on system A compared to system B
4. How much faster is the program in (1) on system B compared to system A

Problem 2

You are building a system around a processor with in-order execution that runs on 1.1 GHz and has a CPI of 0.7 excluding memory reference. The only instructions that read or write to the memory are loads (20% of total instructions) and stores (5% of total instructions).

The memory system for this computer is composed of a split L1 cache that imposes no penalty on hits. Both the I-cache and the D-cache are direct mapped and have 32KB each. The I-cache has a 1% miss rate, and a 32-byte blocks, and the D-cache is write through with 5% miss rate and 16-byte blocks. There is a write buffer on the D-cache that eliminates stalls for 95% of the writes.

The 512 KB write-back, unified L2 cache has a 64-byte blocks and an access time of 15 ns. It is connected to the L1 cache by a 128-bit data bus that runs at 266MHz and can transfer one 128-bit word per clock cycle. Of all memory reference sent to the L2 cache in this system, 80% are satisfied without going to the main memory. Also 50% of all blocks replaced are dirty.

The 128-bit wide main memory has an access latency of 60 ns, after which any number of bus words may be transferred at the rate of one per cycle on the 128-bit wide 133 MHz main memory bus.

1. What is the average memory access time for instruction access
2. What is the average memory access time for data reads
3. what is the average memory access time for data writes
4. What is the overall CPI including memory references
5. You are considering replacing the 1.1 GHz CPU with a 2.1 GHz, but otherwise identical. How much faster does the system runs with the new CPU assuming the the L1 cache still has no hit penality, and the L2 and main memory remain the same in absolute terms.