CSE 560 – Practice Problem Set 7 Solution

We are interested in assessing the performance impact of two different data cache organizations on the performance of a single-issue, in-order processor with a 5-stage pipeline that uses 32-bit addresses for memory references. Assume that the CPI with a perfect cache is 2.0, the clock cycle time is 1 ns (i.e., 1 GHz clock rate), and there are 0.2 memory references per instruction (20% of instructions are loads or stores).

The size of the two caches is 64 KB, and both have a block size of 64 bytes. One cache (cache A) is direct mapped and the other (cache B) is two-way set associative.

To accommodate the additional complexity of a set-associative cache, the processor cycle time must be stretched 1.25 times its value with a perfect cache. The direct mapped cache does not impact the processor's clock rate. The cache miss penalty is 75 ns for either cache organization. Assume the hit time for both caches is 1 cycle, the miss rate of the direct-mapped cache is 1.4%, and the miss rate of the set-associative cache is 1.0%.

(a) If the memory is byte-addressable, how many address bits represent the offset within the cache block for cache A? For cache B?

$\log_2 64 = 6$ bits

(b) How many blocks fit (total) in cache A? In cache B?

64KB/64 = 1024 blocks in both A and B

(c) How many address bits are used to select the index for cache A? For cache B?

$\log_2 1024 = 10$ bits for A   1 fewer (9 bits) for B, because it is 2-way set associative

(d) How many tag bits are required per cache block in cache A? In cache B?

32 – (10+6) = 16 tag bits for A   32 – (9+6) = 17 tag bits for B

(e) What is the average memory access time (in ns) for cache A? For cache B?

$t_{AVG} = t_{HIT} + \%miss \times t_{MISS}$

$t_{AVG,A} = 1ns + (0.014)(75ns) = 2.05$ ns

$t_{AVG,B} = 1.25ns + (0.01)(75ns) = 2$ ns

(f) What is the CPI for the machine with cache A? With cache B?

$CPI = CPI_{BASE} + (\%load/store) \times (\%miss) \times (clocks/miss)$

$CPI_A = 2 + (0.2)(0.014)(75ns/1ns) = 2.21$

$CPI_B = 2 + (0.2)(0.01)(75ns/1.25ns) = 2.12$

(g) For an application with 10^9 instructions, what is the total execution time for the machine with cache A? With cache B?

T = (instruction count) x (CPI) x (tCLK)

T_A = 10^9 x 2.21 x 1ns = 2.21 sec

T_B = 10^9 x 2.12 x 1.25ns = 2.65 sec

(h) Which machine completes the application faster?

A is faster