COE 308 – Spring 2006
Computer Architecture

Assignment 6: Memory Hierarchy

Due: Saturday, May 27, 2006

1. **(4 pts)** Suppose a computer's address is k bits (using byte addressing), the cache data size is S bytes, the block size is $B = 2^b$ bytes, and the cache is m-way set-associative. Figure out what the following quantities are in terms of S, B, m, b, and k.

a) The number of sets in the cache.

b) The number of index bits in the address

c) The number of tag bits in the address

d) The total number of bits required to store all the valid and tag bits in the cache

2. **(4 pts)** Consider a processor with a 2 ns clock cycle, a miss penalty of 20 clock cycles, a miss rate of 0.05 misses per instruction, and a cache access time (hit time) of 1 clock cycle. Assume that the read and write miss penalties are the same.

a) **(1 pt)** Find the average memory access time (AMAT).

b) **(1 pt)** Suppose we can improve the miss rate to 0.03 misses per instruction by doubling the cache size. However, this causes the cache access time to increase to 1.2 cycles. Using the AMAT as a metric, determine if this is a good trade-off.

c) **(2 pts)** If the cache access time determines the processor's clock cycle time, which is often the case, AMAT may not correctly indicate whether one cache organization is better than another. If the processor's clock cycle time must be changed to match that of a cache, is this a good trade-off? Assume that the processors in part **(a)** and **(b)** are identical, except for the clock rate and the cache miss rate. Assume 1.5 references per instruction (for both I-cache and D-cache) and a CPI without cache misses of 2. The miss penalty is 20 cycles for both processors.

3. **(3 pts)** Consider one of the three organizations for main memory: one-word-wide, four-word-wide, and interleaved memory with 4 banks. If the main memory latency for a new access is 10 memory bus cycles and the transfer time is 1 memory bus cycle, what are the miss penalties for each of these organizations?

4. **(2 pts)** Assume a memory system that supports interleaving either four reads or four writes. Given the following memory addresses in order as they appear on the memory bus: 3, 9, 17, 2, 51, 37, 13, 4, 8, 41, 67, 10, which ones will result in a bank conflict?

5. (5 pts) Consider three processors with three cache configurations:

Processor 1: Direct-mapped i-cache and d-cache with one-word blocks
Instruction miss-rate = 4%, data miss-rate = 6%
Processor 2: Direct-mapped i-cache and d-cache with four-word blocks
Instruction miss-rate = 2%, data miss-rate = 4%

Processor 3: Two-way set associative i-cache and d-cache with four-word blocks
Instruction miss-rate = 2%, data miss-rate = 3%

a) (3 pts) For these processors, 50% of the instructions contain a data reference. Assume that the cache penalty is 6 + Block size in words. Determine which processor spends the most cycles on cache misses.

b) (2 pts) The cycle time is 420 ps for Processor 1 and 2, and 310 ps for the third processor. Determine which processor is the fastest and which one is the slowest.

6. (2 pts) Consider a virtual memory system with the following properties:

40-bit virtual byte address
16 KB pages
36-bit physical byte address

What is the total size of the page table for each process on this processor, assuming that the valid, protection, dirty, and use bits take a total of 4 bits and that all the virtual pages are in use? (Assume that disk addresses are not stored in the page table).