Assignment 1

1. (10 pts) A machine has a 2.5 GHz system clock. If primary memory takes 24 ns to access a word, and the processor can execute an instruction every 2 clock cycles, how many could be executed in the time it takes the memory to access a single word?

2. (10 pts) A machine has a base CPI of 2 clock cycles. Measurements obtained show that the instruction miss rate is 8% and the data miss rate is 10%, and that on average, 30% of all instructions contain one data reference. The miss penalty for the cache is 16 cycles. What is the effective CPI of this machine when these factors are taken into account?

3. (20 pts) A direct-mapped cache has blocks consisting of 16 eight-byte words. It has a capacity of 2 MB.

   (a) Write the mathematical function that computes the cache index from a given byte address β.

   (b) Write the mathematical function that computes the word offset of β in its cache block.

   (c) How many bits are in the tag of a cache block?

   (d) How many bits in total are needed in this cache, assuming it is a write-through cache?

4. (20 pts) A program uses a C++ structure type named `Record` whose total size is 16 bytes. It uses an array `A` of such `Record` structures declared as

   ```
   Record   A[480];
   ```

   which is stored in consecutive locations in primary memory starting at byte address 0 when the program is run. During its execution, the program runs through the following loop:

   ```
   for ( int i = 0; i < 480; i = i+3)
       sum = sum + A[i];
   ```

   Assume that `sum` is stored in a register throughout this loop and answer the following two questions.

   (a) A machine with 4-byte words has a split instruction and data cache. The data cache is a direct-mapped cache with 8-word blocks and a capacity of 512 bytes. Suppose that the above loop is executed on this machine and that none of the array is in the cache when the loop is reached. How many data misses are there?

   (b) Assume the cache in the machine described in part (a) is replaced by one with ***16-word*** blocks and calculate the total number of data misses with this cache. ***(THIS IS THE CHANGE.)***

5. (12 pts) Assume that in a particular machine, two-dimensional arrays are stored in row-major order. Assume that arrays X and Y are declared as follows:

   ```
   int X[10000][500];
   int Y[500][10000];
   ```

   Assume that these arrays have been initialized before the following code is reached.

   ```
   for ( i = 0; i < 10000; i++)
       for ( j = 0; j < 500; j++)
           X[i][j] = Y[j[i];
   ```

   (a) Is there any spatial locality exhibited by the process executing this code? If so describe which variable references exhibit it and why. References to which variables exhibit spatial locality?

(b) Is there any temporal locality exhibited by the process executing this code? If so describe which variable references exhibit it and why.

(c) If the inner and outer loops are swapped, what changes about your answers to parts (a) and (b)?

6. (10 pts) A machine with a two level cache has a base CPI of 2.0 when all references hit the primary cache. Given the following characteristics:

   - a clock rate of 1GHz
   - memory access time of 80ns
   - miss rate at the primary cache of 6%
   - secondary cache access time of 20ns
   - miss rate at secondary cache of 2%

   what is the total CPI?

7. (18 pts) The following table gives the parameters of various different caches. For each cache, fill in the missing fields in the table. The legend is as follows: m = number of bits in a physical memory address; C is the capacity of the cache in bytes; B is the block size in bytes; S is the number of cache sets; A is the associativity; T is the number of tag bits; s is the number of bits in a set index; b is the number of bits in the block offset, which in this case means the number of bits to address a single byte of a block.

Cache	m	C	B	A	S	T	s	b
1	32	1024	4	4				
2	32	1024	4	256				
3	32	1024	8	1				
4	32	1024	8	128				
5	32	1024	32	1				
6	32	1024	32	4				