Computer Architecture – tutorial 5

Context, Objectives and Organization

The goal of the quantitative exercise in this tutorial is to explore qualitatively and quantitatively some hardware and software optimizations to improve cache performance.

E1: CAR September 2003 exam P2, groups of 2 − 35 min

Problem

Consider a computer system with a first-level data cache with the following characteristics: size: 16KBytes; associativity: direct-mapped; line size: 64Bytes; addressing: physical.

The system has a separate instruction cache and you can ignore instruction misses in this problem. This system is used to run the following code:

```
for (i=0; i<4096; i++)
  X[i] = X[i] * Y[i] + C
```

Assume that both X and Y have 4096 elements, each consisting of 4 bytes (single precision floating point). These arrays are allocated consecutively in physical memory. The assembly code generated by a naive compiler is the following:

```
loop: lw f2, 0(r1)      # load X[i]
      lw f4, 0(r2)      # load Y[i]
      multd f2, f2, f4  # perform the multiplication
      addd f2, f2, f0   # add C (in f0)
      sw 0(r1), f2      # store the new value of X[i]
      addi r1, r1, 4    # update address of X
      addi r2, r2, 4    # update address of Y
      addi r3, r3, 1    # increment loop counter
      bne r3, 4096, loop # branch back if not done
```

a. How many data cache misses will this code generate? Breakdown your answer into the three types of misses. What is the data cache miss rate?

b. Provide a software solution that significantly reduces the number of data cache misses. How many data cache misses will your code generate? Breakdown the cache misses into the three types of misses. What is the data cache miss rate?

c. Provide a hardware solution that significantly reduces the number of data cache misses. You are free to alter the cache organization and/or the processor. How many data cache misses will your code generate? Breakdown the cache misses into the three types of misses. What is the data cache miss rate?