

# Homework Assignment 4 of Computer Architecture

College of Computer Science, Zhejiang University

Submission Deadline: June 30, 2015

(100 points)

1. (18 points) Consider the following description of a memory hierarchy.

Virtual address wide = 45 bits, Memory physical address wide = 38 bits, Page size = 4KB.

Cache capacity = 8KB. It is a write-back 2-way associative cache.

- How many bits are there in the fields of tag, index and block offset of the physical memory address.
- Draw a graph to show a cache line (including tag, data, and some other control bits) in the cache.
- Draw a graph to show if it is implemented in the way of virtually indexed and physically tagged cache.
- Please describe the access procedure to the memory hierarchy in (c) that when a CPU address (virtual address) is given to access the cache.

2. (22 points) Assume that we have two machines A and B. The only difference between A and B lies in their cache hierarchies:

Machine A: 64 KB level-one data cache with a  $8\text{ ns}$  access time and a miss rate of 8%

Machine B: 8 KB level-one data cache with a  $2\text{ ns}$  access time and a miss rate of 15%, and a 1 MB level-two cache with a  $20\text{ ns}$  access time and a miss rate of 10%.

Assume that both machines have an I-cache miss rate of 0%, a main memory access time of  $50\text{ ns}$ , and all the bus transfer time could be ignored. Which machine will have a better performance in memory access (AMAT)? Why?

3. (20 points) Suppose you own a computer which has the following properties:

- the pipeline can accept a new instruction every cycle
- the cache can provide data every cycle (i.e. no penalty for cache hits)
- the instruction cache miss rate is 2.5%
- the data cache miss rate is 3.5%
- 30% of instructions are memory instructions
- the cache miss penalty is 80 cycles.

Now you want to purchase a new computer. you can either

- purchase a machine with a processor and cache that is twice as fast as your current one (memory speed is the same as the old machine, though), or
- purchase a machine with a processor and cache that is the same speed as your old machine but in which the cache is twice as large.

Assume the cache miss rate will drop by 40% with this larger cache (although this is generally not true in the real world).

Which computer are you best off purchasing? Explain in detail, showing the relative performance of each choice.

4. (40 points) You are building a system around a processor with in-order execution that runs at 1.1 GHz and has a CPI of 0.7 excluding memory accesses. The only instructions that read or write data from memory are loads (20% of all instructions) and stores (5% of all 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 D-cache are direct mapped and hold 32KB each. The I-cache has a 2% miss rate and 32-byte blocks, and the D-cache is write through with a 5% miss rate and 16-byte blocks. There is a write buffer on the D-cache that eliminates stalls for 95% of all writes.

The 512KB write-back, unified L2 cache has 64-byte blocks and an access time of 15ns. 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 bus cycle. Of all memory references sent to the L2 cache in this system, 80% are satisfied without going to main memory. Also 50% of all blocks replaced are dirty.

The 128-bit-wide main memory has an access latency of 60ns, 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.

- a) (8points) What is the average memory access time for instruction accesses ?
- b) (8points) What is the average memory access time for data reads ?
- c) (8points) What is the average memory access time for data writes ?
- d) (8points) What is the overall CPI, including memory accesses ?
- e) (8points) You are considering replacing the 1.1GHz CPU with one that runs at 2.1GHz, but is otherwise identical. How much faster does the system run with a faster processor ? Assume the L1 cache still has no hit penalty, and that the speed of the L2 cache, main memory, and buses remains the same in absolute terms (e.g. the L2 cache has a 15n access time and a 266MHz bus connecting it to the CPU and L1 cache).