COMP25212

Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

System Architecture

Date: Tuesday 26th May 2015

Time: 14:00 - 16:00

Please answer any THREE Questions from the FOUR questions provided

Use a SEPARATE answerbook for each SECTION

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

[PTO]

Section A

1. **Caches**

a) Explain the role of CPU caches in modern computer systems. (2 marks)

b) Explain the differences between the two forms of locality exploited by caches and how they relate to key characteristics of CPU caches. (5 marks)

c) Give one example for each of the two types of locality occurring in programs. (2 marks)

d) Explain why modern computer systems frequently rely on multiple levels of cache. (4 marks)

The elements of multidimensional arrays are stored contiguously in memory, for example, a 2-dimensional array A of size N*N and containing 4-byte integers will be stored one row after another so that the address of element A[i][j] can be computed as the address of A[0][0] plus (i*N*4 + j*4).

Assume that a CPU has a 4-way set associative L1 data cache of 16KB where the cache line size is 64 bytes and the replacement policy is LRU, and that the size of the array is N = 8192.

e) How many cache lines are in each set of this CPU's L1 data cache? (1 mark)

f) How many lines within the L1 data cache will be used by a program repeatedly traversing a single column of the array (i.e., A[0][k], A[1][k], A[2][k] … A[N-1][k] where k is constant)? (3 marks)

g) What type of cache misses will this program experience? (1 marks)

h) If the CPU cache is fully associative rather than 4-way set associative, how many cache lines will be used and what type of cache misses will occur? (2 marks)

[PTO]

2. **Virtualization and Storage**

a) What is "System Virtualisation"? (1 mark)

b) What are the major goals of System Virtualisation? (2 marks)

c) Explain what “live migration” means and how it can be implemented using System Virtualisation. (4 marks)

d) A server disk drive is specified as having a mean “seek time” of 4 milliseconds, a “rotation speed” of 10,000 revolutions per minute and a “transfer rate” of 200 megabytes per second.

i) Explain what each term in this specification means for the performance of disk operation. (3 marks)

ii) How long on average would a transfer of 4 kilobytes take from a random position on the disk? And for 4 megabytes? (4 marks)

e) Explain the two main reasons, aside from increasing storage capacity, why more than one disk drive may be used in a single system and how these goals may be achieved. (6 marks)

[PTO]

Section B

3. **Pipelining**

Consider the following application implementing a simple cash till program:

```
total=0;
for (i=total_items; i>0; i--) {
  cost[i]=quantity[i]*price[i];
  total+=cost[i]*taxes[i];
}
total=total*discount;
```

And the following assembly code generated by the compiler:

```
 1  init  adr r9, quantity     ; put addresses of arrays
 2        adr r10, price       ; in registers r9-r12
 3        adr r11, cost
 4        adr r12, taxes
 5        mov r6, #0           ; total (kept in register)
 6        ldr r0, total_items
 7  loop  ldr r1, [r9,r0]      ; quantity[i]
 8        ldr r2, [r10, r0]    ; price[i]
 9        mul r3, r1, r2       ; cost[i] (in register)
10        str r3, [r11,r0]     ; now in memory
11        ldr r4, [r12,r0]     ; taxes[i]
12        mul r5, r4, r3       ; cost[i]*taxes[i]
13        add r6, r5, r6       ; update total (in reg.)
14        sub r0, r0, #1       ; i--
15        cmp r0, #0
16        bgt loop             ; iterate while i>0
17        ldr r7, discount
18        mul r6, r6, r7       ; total * discount
19        str r6, total
```

a) Identify the possible hazards this code could suffer when run in a classic 5-stage in-order pipeline. Discuss what measures could be taken to avoid (or minimize) the penalties due to these hazards. You do not need to do the discussion in a case by case basis. (14 Marks)

b) Draw the dependence graph of the instructions within the main body of the loop (instructions 7-16) and discuss how suitable this code would be for being executed in a 4-way superscalar processor. Assume structural hazards are never a problem. (6 Marks)

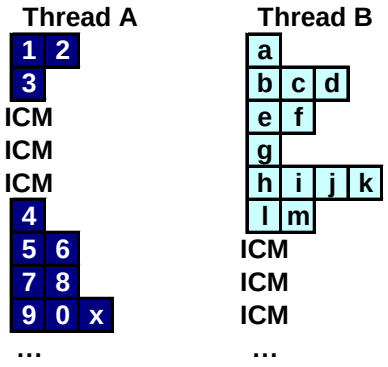
[PTO]

4. **Multithreading / Multicore**

a) Superscalar, multithreading and multicore are three different techniques used to improve the performance of processors. All of them exploit different forms of parallelism. Discuss the main differences between them and whether they are compatible with each other. (8 Marks)

b) Explain why instruction reordering can be useful to improve the performance of processors and discuss the benefits and limitations of doing it in the compiler or in hardware. (6 Marks)

c) Consider a 4-way superscalar processor with 2-way multithreading and 2 simple programs, A and B, with the instructions issued as per the diagrams below. ICM stands for instruction cache miss and means no instruction can be issued that cycle.



Draw a diagram showing the execution flow (issuing only) of the two threads if the processor used:

i) Coarse-grain multithreading
ii) Fine-grain multithreading
iii) Simultaneous multithreading

Assume the processor starts issuing instructions from A. (6 Marks)

END OF EXAMINATION