〈共 4 頁〉

Department of Computer Science, National Tsing Hua University
Ph.D. Qualification Examination
Computer Architecture, Fall 2017

1. (20 pts) Cache Access

A processor has a 4-way set-associative L1 cache that can house 4 blocks in total. The access latency to this cache is 1 cycle. The replacement policy is true LRU. The processor is known to not employ any prefetching mechanism.

The processor also has a 16-way set-associative L2 cache that can house 128 blocks in total. The access latency to this cache is 20 cycles.

A programmer writes a test program that in a loop repeatedly accesses only the following data cache blocks (assume billions of iterations are run):

A, B, C, D, E, F

where A~F are different cache block addresses.

In the steady state (i.e., after the loop has executed for a few iterations), the programmer finds out that the average memory access time is 1 cycle.

Then, the programmer writes another program that in a loop repeatedly accesses only the following data cache blocks:

A, B, C, D, E, F, G, H

In the steady state (i.e., after the loop has executed for a few iterations), the programmer finds out that the average memory access time is 20 cycles.

(a) What can you say about this processor? (I.e., what is going on?) Please describe everything you can say, concretely, but be concise. **(5pts)**

(b) Based on the above information, what do you expect the average memory access time of yet another program that in a loop repeatedly accesses only the following data cache blocks? **(5pts)**

A, B, C, D, E

(c) Again, based on the above information, what do you expect the average memory access time of yet another program that in a loop repeatedly accesses only the following data cache blocks? **(5pts)**

A, B, C, D, E, F, G

(d) Finally, again, based on the above information, what do you expect the average memory access time of yet another program that in a loop repeatedly accesses only the following data cache blocks? **(5pts)**

A, B, C, D, E, F, G, H, I

2. (15 pts) Cache Implementation

(a) L1 Data Cache

A creative designer realizes one day that he can increase the L1 cache size of his design from 64KB to 2MB without affecting the number of cycles it takes to retrieve data from the cache and without reducing the processor frequency.

What can you definitively say about this design? **(5pts)**

(b) Sectored Cache vs. Smaller Blocks **(5pts)**

After mounds of coffee and a good number of high-level simulation cycles, you narrowed down the design choices for the L1 cache of the next-generation processor you are architecting to the following two:

Choice 1. A sectored 64KB cache with 64-byte blocks and 8-byte sub-blocks
Choice 2. A non-sectored 64KB cache with 8-byte blocks

Assume the associativity of the two caches are the same. Please give one definitive advantage of the sectored cache over the non-sectored one. Also, please give one definitive advantage of the non-sectored cache over the sectored one.

(c) FastCaches, Inc. **(5pts)**

Suppose a creative colleague of yours at FastCaches, Inc. proposed the following idea. Instead of directly using a number of bits to index into the cache, take those bits and form the index using a hash function implemented in hardware that randomizes the index.

What type of cache misses can this idea potentially reduce? Please explain why.

What is a disadvantage of the idea other than the additional hardware cost of the hash function?

3. (15%) Please describe the values of X and Y and please show how they can be calculated

Instruction	**address**						
Loop: sll $t1, $s3, 2	**3004**	**0**	**0**	**19**	**9**	**4**	**0**
add $t1, $t1, $s6	**3008**	**0**	**9**	**22**	**9**	**0**	**32**
lw $t0, 0($t1)	**3012**	**35**	**9**	**8**	**0**		
bne $t0, $s5, Loop	**3016**	**5**	**8**	**21**	**X**		
j Loop	**3024**	**2**	**Y**				

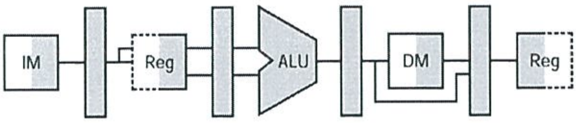
4. (5% each) Describe the following terms
   a. What is pseudo instruction? Please show MIPS of **move $t0, $t1**
   b. Compare the speed between DRAM, FF, disk, SRAM
   c. MIPS has no "branch on greater than." **if (g > h) goto GREATER;** Explain your reasoning
5. (5pts) Dynamic Branch Prediction: Draw a state machine of the 2-bit branch prediction scheme and explain its advantage over the 1-bit scheme.
6. (5pts) Pipelining exploits the potential parallelism among instructions, which is called instruction-level parallelism (ILP). Please explain the concept and limitation of the two primary methods for increasing the potential amount of ILP.
7. (15pts) Complete the pipeline schedule of the following code segment using the pipeline diagram.

```
lw   $s3, 20($s1)
sub  $s4, $s3, $s6
add  $s7, $s3, $s5
and  $s7, $s7, $s3
or   $s2, $s5, $s7
```

(a) Identify and explain all the dependences and hazards of the code segment.

(b) Complete the following graphical representation of the pipeline diagram with all the necessary forwarding and stall to solve the hazards in an optimized way.

`lw $s3, 20($s1)`



8. (10pts) How well would this loop be scheduled on a static two-issue pipeline for MIPS with loop unrolling four copies of the loop body? What is the resultant CPI?

```
Loop: lw      $t0, 0($s1)        # $t0=array element
      addu    $t0, $t0, $s2      # add scalar in $s2
      sw      $t0, 0($s1)        # store result
      addi    $s1, $s1, -4       # decrement pointer
      bne     $s1, $zero, Loop   # branch $s1!=0
```