

## CS 3853 Fall 2009 Midterm Exam

Mark your answer for each question on the SCAN-TRON form provided. Mark only one answer per question. If you believe the right answer is not among the options given, choose the closest answer. Write your name on the SCAN-TRON. You must turn in your question paper along with your SCAN-TRON form.

### True (A) or False (B)

1. In the classic 5-stage pipeline, forwarding eliminates all stalls due to data hazards.
2. In the MIPS ISA, pipeline registers are named with "F" while double-precision floating point registers are named with "D."
3. A direct-mapped cache will have a lower hit time than a fully-associative cache with the same capacity.
4. A larger block size in a cache tends to increase compulsory misses.
5. The Java statement `block_addr = addr % 4096;` would find the block address of the memory block containing address `addr` in a cache with 64 sets and 64-byte block size.
6. A page fault is an example of an exception.
7. Dividing by zero will cause an interrupt.
8. TLB entries contain dirty bits.
9. A branch history table (BHT) predictor with one-bit entries can be more accurate than a static branch predictor.
10. WAW hazards arise from true dependences.

### Multiple Choice

#### 11. Amdahl's law implies that:

- (a) The speedup of a program is directly proportional to its performance.
- (b) Decreasing stall cycles by executing branches in the decode stage can also eliminate the effects of structural hazards caused by loads.
- (c) Performance is inversely proportional to execution time.
- (d) The effect of an optimization is limited by the portion of the program to which it can be applied.

#### 12. The speedup of one machine over another for a given program is:

- (a) Old CPI divided by new CPI.
- (b) Old number of cycles divided by new number of cycles.
- (c) Old execution time divided by new execution time.

- (d) All of the above.

#### 13. Suppose machine A has a clock rate of 1 GHz and CPI of 1.0, while machine B has a clock rate of 2 GHz and a CPI of 1.5. Which machine is faster, and by how much?

- (a) Machine A is faster by 50%.
- (b) Machine B is faster by 33%.
- (c) Machine B is faster by 50%.
- (d) Machine B is faster by a factor of 2.

#### 14. Consider a cache with 32-bit addresses, 768 blocks, and a block size of 128 bytes. Tags are 17 bits. How many sets are there, and what is the associativity of the cache?

- (a) 128 sets, 6-way set associativity.
- (b) 256 sets, 3-way set associativity.
- (c) 128 sets, 8-way set associativity.
- (d) There is not enough information given to answer the question.

#### 15. Consider a 64KB cache with 48-bit addresses, a block size of 8 bytes, and an associativity of 16. What is the size of a tag?

- (a) 35 bits.
- (b) 36 bits.
- (c) 20 bits.
- (d) There is not enough information given to answer the question.

#### 16. Consider a virtual memory system with 40-bit virtual addresses. Pages are 1 MB. How many virtual pages can there be in a process?

- (a)  $2^{28}$ .
- (b)  $2^{10}$ .
- (c)  $2^{20}$ .
- (d) There is not enough information given to answer the question.

#### 17. Consider a virtual memory system with 48-bit virtual addresses and 32-bit physical addresses. Pages are 16KB. The cache has a block size of 32 bytes, is 16-way set associative, and uses physical tags. What is the minimum possible size of a tag in the cache?

- (a) 34 bits.
- (b) 22 bits.

- (c) 18 bits.
- (d) There is not enough information given to answer the question.

- (b) 8192 misses.
- (c) 4096 misses.
- (d) 256 misses.

18. Assume the classic 5-stage pipeline, with registers being written in the first half of WB and read in the last half of ID. How many cycles does the following code take, assuming no forwarding?

```
DADD  R1, R7, R8
DSUB  R2, R2, R9
DADD  R3, R1, R2
```

- (a) 11 cycles.
- (b) 8 cycles.
- (c) 6 cycles.
- (d) 14 cycles.

19. What would be the answer to the previous problem if we allow forwarding from the EX stage to the EX stage of the next cycle?

- (a) 10 cycles.
- (b) 7 cycles.
- (c) 6 cycles.
- (d) 5 cycles.

20. Consider a direct-mapped data cache with a 16 blocks of 32 bytes each. How many cache misses will result from the following C code to sum together all of the elements of a two-dimensional array of 32-bit integers?

```
void sum (int A[256][256]) {
    int i, j, sum = 0;
    for (i=0; i<256; i++)
        for (j=0; j<256; j++)
            sum += A[i][j];
    return sum;
}
```

- (a) 65536 misses.
- (b) 8192 misses.
- (c) 4096 misses.
- (d) 256 misses.

21. In the previous problem, suppose the two for loops are switched so that for i loop comes second and the for j loop comes first. How many cache misses will result?

- (a) 65536 misses.

22. Suppose an instruction set has 32-bit instructions. Every instruction has an 8-bit opcode and a 12-bit immediate operand. Some instructions have three register operands. Every instruction that uses registers must be able to specify any of the registers. How many integer registers can this instruction set support?

- (a) 32.
- (b) 64.
- (c) 16.
- (d) There is not enough information given to answer the question.

23. In the RISC ISA assignment, suppose there is a branch instruction at address 0x1000 that jumps to the instruction at address 0x1018. What is the value in the immediate field of the branch instruction?

- (a) 0x18
- (b) 0x14
- (c) 5
- (d) 6

24. Which of the following is NOT an example of spatial locality?

- (a) A loop that goes through every element of an array in sequence.
- (b) A program with branches that are mostly not taken.
- (c) Storing adjacent virtual pages in adjacent disk blocks.
- (d) Computing the sum of the integers from 1 to 100.

25. Why can dynamic branch prediction be more accurate than static branch prediction?

- (a) Because static branch prediction requires an extra profiling step that programmers are unlikely to use.
- (b) Because, on average, more branches tend to be taken than not taken.
- (c) Because the branch history table has a limited number of entries, while a static branch predictor can make a different prediction for every branch instruction.
- (d) Because dynamic branch predictors can adapt to changing branch behavior.

26. What is the optimal cache replacement pol-

icy?

- (a) Evict the least recently used block.
- (b) Evict the least frequently used block.
- (c) Evict the block that will next be used farthest in the future.
- (d) There is no such thing as an optimal replacement policy.

27. **What are valid bits?**

- (a) A valid bit is true if the corresponding cache block was loaded from memory.
- (b) With the write-back policy, if the valid bit is true it means that the block has been modified and should be written back.
- (c) The valid bit is used in the least-recently-used replacement policy to indicate that a block is the most recently used and should not be replaced.
- (d) The valid bits are used to distinguish between memory blocks in a set.

28. **How many valid bits are needed in a cache?**

- (a) One.
- (b) Number of sets.
- (c) Number of sets  $\times$  associativity.
- (d) Associativity  $\times$  block size.

29. **What is an example of a structural hazard?**

- (a) The EX stage of an ADD instruction and the ID stage of a branch instruction occurring in the same cycle on a machine with a single adder where branches are executed in ID.
- (b) A branch that must stall waiting for a register value to become available in the register file.
- (c) A load that forwards its result to the EX stage of the next instruction but must still stall for one cycle because MM is the fourth stage and EX is the third stage.
- (d) A cache miss in the IF stage.

30. **Which of the following accurately computes the set index for a cache with 8-byte blocks and 32 sets?**

- (a)  $\text{index} = (i / 8) \% 32;$

- (b)  $\text{index} = (i \gg 3) \& 31;$
- (c)  $\text{index} = (i \& 0xf8) / 8;$
- (d) All of the above.

31. **Suppose an integer array starts at virtual address 0x1000. What is the virtual address of the fifth element of that array?**

- (a) 0x1005
- (b) 0x1004
- (c) 0x1014
- (d) 0x1010

32. **The L1 cache usually has lower associativity than the L2 cache. Why is that?**

- (a) The L2 cache trades a higher hit time for a lower miss rate. The L1 cache is focused on lower latency.
- (b) The L1 cache has fewer blocks than the L2 cache, so it naturally has lower associativity.
- (c) The L1 cache must be small. The extra tags that higher associativity entails cause it to be larger.
- (d) The L1 cache's associativity is constrained by the need for physical tags. The L2 cache has no such constraint.

33. **We speak of direct-mapped, set-associative, and fully-associative placement of blocks in caches. However, we don't usually speak of a choice of placement policy of pages in virtual memory. Why?**

- (a) Because the only choice is direct-mapped. Since replacement is handled by software, it would be too costly to implement associativity because that would involve a time-consuming search for the LRU block.
- (b) Because the placement policy is constrained by the page size, which in turn constrains the number of sets in the cache. So the virtual memory naturally has the same placement policy and associativity as the cache.
- (c) The placement policy is implemented by the TLB, which is fully-associative with replacements being handled in software. (d) Because the only choice is fully-associative. The miss rate of other placement schemes due to conflict would lead to an unacceptable number of page faults, which are far more costly than cache misses.