浙江大学 2022–2023 学年秋冬学期

《计算机组成与设计》课程期末考试试卷

课程号：67190020，开课学院：信息与电子工程学院

考试试卷：√A 卷、B 卷（请在选定项上打√）

考试形式：√闭、开卷（请在选定项上打√），

允许带 1 张 A4 纸大小的手写资料和计算器入场

考试日期：2023 年 1 月 6 日，考试时间：120 分钟

诚信考试，沉着应考，杜绝违纪。

考生姓名：_____学号：_____所属院系（专业）：_____

题序	一	二	三	四	五	六	七	八	总 分
得分									
评卷人									

I. CHOICE (Only one correct answer, 60 points)

1. What is the range of exponent of IEEE 754 single precision? ( )C

A. 1~254

B. -128~126

C. -126~127

D. -127~128

2.The SRAMs are basically used as _____ ( ) B

A. register

B. cache

C. main memory

D. disk

3. What is the decimal product of the binary number 1.00×2^{-1} and -1.11×2^{-2}? () C

A. -0.4375

B. 0.0625

C. -0.21875

D. -0.0625

4. Consider the following C code:

```
typedef unsigned char *pointer; // sizeof(unsigned char) = 1 byte

void show_bytes(pointer start, size_t len) {
   for (int i = 0; i < len; i++)
        printf("0x%x\n", start[i]);
}

int main() {
   int a = 0x11223344;
   show_bytes((pointer) &a, sizeof(int));
}
```

If this C code runs on a little-endian machine, what will we get on the **third** line of the terminal output? () B

A. 0x11

B. 0x22

C. 0x33

D. 0x44

5. The reason for the implementation of the cache memory is _____. ()B

A. to increase the internal memory of the system

B. the difference in speeds of operation of the processor and memory

C. to reduce the memory access time

D. all of the mentioned

6. Which of the following allows simultaneous write and read operations? ()C

A. ROM

B. EROM

C. RAM

D. None of the above

7. The copy-back protocol is used _____. ()B

A. to copy the contents of the memory onto the cache

B. to update the contents of the memory from the cache

C. to remove the contents of the cache and push it on to the memory

D. none of the mentioned

8. Consider a virtual memory system with 32 bit virtual byte address, 4KiB/page, 32 bits each entry. The physical memory is 512MiB. Then, the total size of page table needs (____) .B

A. 1MiB

B. 4MiB

C. 8MiB

D. 16MiB

9. The function of assembler is _____. ()C

A. Transforming high level language to binary language

B. Transforming binary language to high level language

C. Transforming assembly language to machine code

D. Transforming high level language to assembly language

10. The temporal aspect of the locality of reference means (_____). C

A. That the recently executed instruction won't be executed soon

B. That the recently executed instruction is temporarily not referenced

C. That the recently executed instruction will be executed soon again

D. None of the mentioned

11. A given application written runs 15 seconds on a desktop processor. A new compiler is released that requires only 0.6 as many instructions as the old compiler. Unfortunately, it increases the CPI by 1.1. How fast can we expect the application to run using this new compiler? B

A.8.2sec

B.9.9sec

C.27.5sec

D.22.7sec

12. Given the following RISC-V assembly code (and assuming all registers start at 0):

```
        addi    t1, x0, 10
        add     t2, t1, t1
repeat: addi    t2, t2, -4
        add     t3, t2, t2
        addi    t1, t1, -2
        bne     x0, t1, repeat
```

What is the final value of register t3? A

A.0

B.-4

C.8

D.4

13. A multilevel page table is preferred in comparison to a single level page table for translating virtual address to physical address because_____()B

A. it reduces the memory access time to read or write a memory location

B. it helps to reduce the size of page table needed to implement the virtual address space of a process

C. it is required by the translation lookaside buffer

D. it helps to reduce the number of page faults in page replacement algorithms

14. You have a two-way set-associative cache with 8B blocks and a total size of 32B. Suppose use a least-recently used replacement policy and begin from power on. Given this sequence of byte addressed accesses, what is the hit rate? (_____) C

0, 4, 128, 24, 224, 88, 134, 92

A. 12.5%

B. 25.0%

C. 37.5%

D. 50.0%

15. Which one is not one of the five classic components of a computer? B

A. Input

B. Bus

C. Memory

D. Output

16. What is the range of 32-bit instructions that can be reached from the current PC using a J-Format jump instruction? D

A. $[-2^{21}, 2^{21} - 1]$

B. $[-2^{20}, 2^{20} - 1]$

C. $[-2^{19}, 2^{19} - 1]$

D. $[-2^{18}, 2^{18} - 1]$

17. Calculate AMAT (Average Memory Access Time) for a machine with the following specs: L1 cache with hit time = 1 cycle and miss rate = 5%, L2 cache with hit time = 5 cycles, miss rate = 15% and miss penalty = 200 cycles. A

A. 2.75 cycles

B. 2 cycles

C. 1.665 cycles

D. None of the above.

18. Assume a 5-stage pipelined RISC-V CPU with no forwarding. How many stalls would there need to be in order to fix all data hazards (assuming that we can read and write to the RegFile on the same cycle)? C

```
sub     t1, s0, s1
or      s0, t0, t1
sw      s1, 100(s0)
add     s2, s0, s2
lw      t1, 104(s0)
```

A. 2

B. 3

C. 4

D. 5

19. The bit width of PC is determined by ( ).B

A. Memory word length

B. Memory capacity

C. Instruction word length

D. Bit width of general-purpose registers

20. The fastest data access is provided using (_____). D

A. Caches

B. DRAM's

C. SRAM's

D. Registers

II. TRUE OR FALSE (10 points)

1. _T _ The LRU can be improved by providing a little randomness in the access.

2. _T _ The associative mapping is costlier than direct mapping.

3. _F _ In the memory hierarchy, as the speed of operation increases the memory size also increases.

4. _F _ The pipeline bubbling is a method used to prevent data hazard and control hazards.

5. _T _ Virtual memory allows a single program to expand its address space beyond the limits of main memory.

6. _T _ Cache block size (B) can affect both miss rate and miss latency.

7. _F _ There is no way to reduce compulsory misses.

8. _T _The CPI of superscalar processors can be less than one.

9. _T _The higher the memory bandwidth, the larger the cache block.

10. _ F _The page table is stored in the disk.

III. Assembly and Pipeline (10 points)

Consider the following sequence of instructions:

```
loop:   add     t0, t1, t2
        lw      t3, 10(t0)
        lw      t4, 14(t0)
        sub     t5, t4, t3
        sw      t5, 18(t0)
        addi    t2, t2, 4
        slti    t6, t2, 200
        bne     t6, x0, loop
```

Assume each datapath stage requires the following amount of time to complete:

- Instruction fetch (IF): 30 ns

- Instruction decode (ID): 20 ns
- Execute / address calculation (EX): 25 ns
- Memory access (MEM): 30 ns
- Register write back (WB): 20 ns

a) How long will a single iteration of this loop take in a single-cycle datapath? ___1000____ns **(2 points)**
b) If we assume ideal pipelining (i.e., no hazards and therefore no stalls), how long will one loop iteration take in a pipelined datapath? ___360____ns **(2 points)**
c) If we now assume a pipelined datapath with forwarding, how many data hazard(s) cannot be solved with forwarding? (number only)
__1____ **(3 points)**
d) If we now assume a pipelined datapath with forwarding, how long will one iteration take?
__390_____ns **(3 points)**

IV. CACHE (10 points)

You are trying to reverse-engineer the characteristics of a cache in a system, so that you can design a more efficient, machine-specific implementation of an algorithm you are working on. To do so, you have come up with two sequences of memory accesses to various bytes in the system in an attempt to determine the following four cache characteristics:

- Cache block size (16, 32, 64, or 128 B).
- Cache associativity (1-, 2-, 4-, or 8-way).
- Cache size (4 or 8 KiB).
- Cache replacement policy (LRU).

The only statistic that you can collect on this system is cache hit rate after performing each sequence of memory accesses. Here is what you observe:

	Addresses Accessed (Oldest → Youngest)								**Hit Rate**
1.	0	23	128	73	8192	255	16384	196	1/2
2.	127	4096	8192	32768	196	16384	0	512	3/8

Assume that the cache is initially empty at the beginning of the first sequence, but *not* at the beginning of the second sequence. The sequences are executed back-to-back, i.e., no other accesses take place in between the two sequences. Thus, at the beginning of

the second sequence, the contents are the same as at the end of the first sequence.

Based on what you observe, what are the following characteristics of the cache? Choose your answer.

a) Cache block size (16, 32, 64, or 128 B)? ____D_____**(3 points)**

A.16B B.32B C.64B D.128B

b) Cache associativity (1-, 2-, 4-, or 8-way)? ____C_____**(3 points)**

A.1-way B.2-way C.4-way D.8-way

c) To identify the cache size, you execute the following sequence right after sequence 2 (i.e., the contents are the same as at the end of the second sequence) and measure the cache hit rate:

Addresses Accessed (Oldest → Youngest): 8192 → X → Y

Which addresses should you use for X and Y?

X: ___A____ **(2 points)**

A.1024 B.2048 C. 4096 D. None of above

Y: ___C____ **(2 points)**

A.8192 B.16384 C. 32768 D. None of above

Solution:

a) Cache hit rate is 1/2 in sequence 1. This means that there are 4 hits. Depending on the cache block size, we can group addresses that belong to the same cache block as follows:

- 8–32 B: {0}, {32}, {128}, {73}, {8192}, {255}, {16384}, {196}. ∴ Number of possible hits = 0.
- 64 B: {0, 32}, {128}, {73}, {8192}, {255, 196}, {16484}. ∴ Number of possible hits = 2.
- 128 B: {0, 32, 73}, {128, 255, 196}, {8192}, {16384}. ∴ Number of possible hits = 4.

Therefore, we can know that the cache block size is 128 B

b) Cache hit rate is 3/8 in sequence 2, which means that there are 3 hits.

We already know that the cache block size is 128 B. Thus, there are 7 offset bits.

The access to address 196 in sequence 2 would hit because the cache block would not

be replaced.

The access to address 512 in sequence 2 would miss because address 512 does not belong to any cache block previously accessed.

Therefore, the accesses to addresses 0, 127, 4096, 8192, 16834 and 32768 in sequence 2 would hit 2 times.

Regardless of cache size, those addresses will never hit when the cache were 1-way or 2-way.

If the cache were 8-way, those addresses would all map to set 0. With 8 ways, addresses 127, 8192, 16384 would not be replaced, so the three addresses would hit.

Therefore, the cache is 4-way associative

c) X = 1024

Y = 32768

If the cache is 4-KiB, all addresses that are multiples of 1024 would map to set 0. If the cache is 8-KiB, all addresses that are multiples of 2048 would map to set 0.

After the access to 8192 in sequence 3, the LRU address in set 0 is 32768.

If the cache is 4-KiB, access to X = 1024 would replace 32768, so access to 32768 would miss. If the cache is 8-KiB, such access would not replace 32768, so access to 32768 would hit.

V. VIRTUAL MEMORY (10 points)

Consider a processor that includes a 40-bit virtual address, an MMU that supports 4096 (2^{12}) bytes per page, 2^{32} bytes of physical memory, and a large Flash memory that serves as a disk. The MMU and the page fault handler implement an LRU replacement strategy.

1. Please calculate the following parameters relating to the size of the page table. You may assume each page entry contains a valid bit and a dirty bit. **(3 points)**

Number of entries in the page table: __ 2^{28}____

Size of page table entry (in bits): ___22___

Size of the page table (in bits): ___$22*2^{28}$___

2. A program running on the processor is halted right before executing the following instruction located at address 0x056C:

```
lw      x3, 0(x5)        //   x5 = 0x2800
```

```
sw        x31,0(x6)        //    x6 = 0x4200
```

The first 8 locations of the page table, just before executing this test program, are shown below; the least-recently-used page ("LRU") and next least-recently-used page ("next LRU") are as indicated. If pages must be brought in from disk, increment the next largest page number.

Page Table

VPN	D	V	PPN
00	1	1	0x7
01	0	1	0x5
02	0	1	0x3
LRU→ 03	1	1	0x1
04	--	0	--
05	0	1	0x0
06	0	1	0x2
Next LRU→ 07	0	1	0x6

This processor also has a 4 element, fully associative, Translation Lookaside Buffer (TLB) that caches translations from VPN to PPN.

TLB

	Tag	D	V	PPN
LRU→	0x3	1	1	0x1
	0x2	0	1	0x3
	0x6	0	1	0x2
Next LRU→	0x1	1	1	0x5

For each virtual address in the chart below, please indicate the VPN, whether or not the access results in a TLB Miss, whether or not the access results in a page fault, the PPN, and the physical address. Please write all numerical values in hexadecimal. (5 points)

Virtual Address	VPN	TLB Miss? (Y/N)	Page Fault? (Y/N)	PPN	Physical Address
0x2800	0x2	N	N	0x3	0x3800

0x4200	0x4	Y	Y	0x8	0x8200

3. What is the physical address of the **sw** instruction? (2 points)

Physical address of sw instruction: 0x___ 7570_____