浙江大学 2020－2021 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 大小的手写资料和计算器入场

考试日期： 2021 年 1 月 23 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 (60 points)

1. For a memory unit of 512 kB, what's the total size of address bits and data bits? (C)

A. 17 B. 19 C. 27 D. 36

2. For X = -0.0011, Y = -0.0101, what's the 2's complement of (X+Y)? (D)

A. 1.1100 B. 1.1010 C. 1.0101 D. 1.1000

3. _____ is generally used to increase the apparent size of physical memory. (B)

A. Secondary memory B. Virtual memory C. Hard disk D. Disks

4. The time delay between two successive initiations of memory operation is _____. (A)

A. Memory access time B. Memory search time

C. Memory cycle time D. Instruction delay

5. When performing a looping operation, the instruction gets stored in the _____. (B)

A. Registers B. Cache C. System heap D. System stack

6. The instruction "lw x5, 40(x6)" does _____. (C)

A. Loads the value of x5 and stores it x6

B. Loads the value of x5 and stores it in Memory[x6 + 40]

C. Loads the value in Memory[x6 + 40] and stores it in x5

D. Loads the value of x6 and stores it in Memory[x5 + 40]

7. The addressing mode which makes use of both register file and memory is _____. (C)

A. Immediate addressing B. Register addressing
C. Base addressing D. PC-relative addressing

8. In a system which has 64 registers, the register id is _____ long. (D)
A. 32-bit B. 8-bit C. 5-bit D. 6-bit

9. The processor keeps track of the result of its operations using flags called _____. (A)
A. Conditional code flags B. Test output flags
C. Type flags D. None of the mentioned

10. The wrong statement/s regarding interrupts and subroutines among the following is/are _____. (D)
i) The subroutine and interrupts have a return statement
ii) Both of them alter the content of the PC
iii) Both are software oriented
iv) Both can be initiated by the user
A. i, ii, and iv B. ii and iii
C. iv D. iii and iv

11. The execution of the following two instructions may have the _____. (A)
ld x5,0(x6)
sd x5,0(x6)
A. RAW (Read after Write) B. WAW (Write after Write)
C. WAR (Write after Read) D. No data dependency.

12. For a 32-bit cache-memory system, a 32KB, 4-way set-associative cache has 2 words cache line size, how many bits are there in such cache's tag? (A)
A: 19 B: 21 C: 23 D: 25

13. In a cache-memory system, assume that hit rate is 95%, and penalty to access the cache and main memory() to be 1ns and 10ns respectively. We can infer that the average memory access time is _____ ns. (C)
A: (1 + 10) / 2 B: 10×5% + 1×95%
C: (10 + 1)×5% + 1×95% D: 10×95% + 1×5%

14. Which of the following cache designer guideline is not valid? (B)
A. Fully associative caches have no conflict misses.
B. In reducing misses, associativity is more important than capacity.
C. The higher the memory bandwidth, the larger the cache block.
D. The shorter the memory latency, the smaller the cache block.

15. For the following memory access pattern: 1,5,1,6,3, what's the content of a 4-entries, direct-mapped cache, assuming cache is vacant at the beginning. (A)

A:

	1	6	3

B:

1	6	3	5

C:

5	6	3	

D:

	6	3	

II. TRUE OR FALSE (10 points)

1. The miss penalty can be reduced by improving the mechanisms for data transfer between the different levels of hierarchy. (T)
2. For forwarding you need only look at the data available in the WB stage. (F)
3. In a system where multiple programs are running, the physical address space must be larger than the total size of the virtual address spaces. (F)
4. In set associative and associative mapping there exists less flexibility. (F)
5. When using the Big Endian assignment to store a number, the sign bits of the number is stored in the lower order byte of the word. (F)
6. The order in which the return addresses are generated and used is First-In-First-Out. (F)
7. Cache performance is of less importance in faster processors because the processor speed compensates for the high memory access time. (F)
8. Pipelining improves performance by increasing instruction throughput. (T)
9. A write-through cache typically requires more bus bandwidth than a write-back cache. (T)
10. An executable binary file that can run on a RISC-V CPU may not be able to execute on another RISC-V CPU. (T)

Write the answers here: ___, ___, ___, ___, ___, ___, ___, ___, ___, ___.

III. Assembly and Pipeline (10 points)

Consider the following RISC-V assembly code, where $result, $A, $B, and $C represent the registers that are used to store the addresses of the variables result, A, B, and C.

```
lw x1, 0($C)
lw x2, 0($A)
mul x3, x1, x2
lw x4, 0($B)
mul x5, x1, x4
sub x6, x5, x3
lw x7, 0($result)
add x8, x7, x6
sw x8, 0($result)
```

The table below gives the instruction latencies for a simple pipeline implementation of an in-order issue, out-of-order completion CPU (an instruction can be issued as long as its operands are ready and the previous instruction is issued, do not worry about forwarding, cache misses and hits):

Assembly instruction formats	Instruction latency (# of cycles)
lw x1, #offset(x2)	8
sw x1, #offset(x2)	8
add x1, x2, x3	6
sub x1, x2, x3	6
mul x1, x2, x3	10

1. Transform the assembly code into C code.

```
// A, B, C, result are global variables
int *A, *B, *C, *result;

void foo() {
    // write your code here
    *result = *result + *C * (*B - *A);


}
```

2. What's the execution time of this assembly code?

51 cycles

3. Optimize the assembly code to minimize the execution time. Write down the minimal execution time and the assembly code after optimization. You should rename the registers to make sure new registers appear in ascending order. (e.g. x1, x2, x3, ...) The table below may help you.

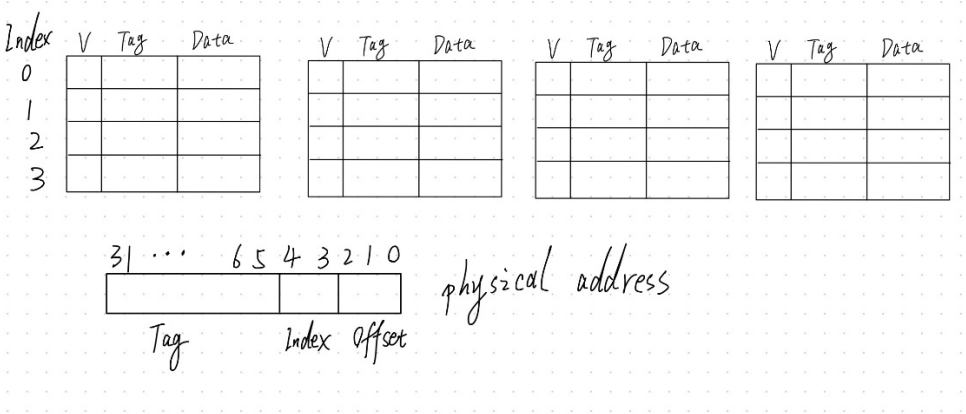
Instruction	Start cycle	End cycle
lw x1, 0($A)	1	8
lw x2, 0($B)	2	9
lw x3, 0($C)	3	10
lw x4, 0($result)	4	11
sub x5, x2, x1	10	15
mul x6, x3, x5	16	25
add x7, x4, x6	26	31
sw x7, 0($result)	32	39

39 cycles

IV. Cache Mapping (10 points)

A computer system has a 128-byte cache. It uses four-way set-associative mapping with 8 bytes in each block. The physical address size is 32 bits, and the smallest addressable unit is 1 byte.

1. Draw a diagram showing the organization of the cache and indicating how physical addresses are related to the cache addresses.



2. To what block frames of the cache can the address 0x000010AF be assigned?

Any block frame in set 1

3. If the addresses 0x000010AF and 0xFFFF7AXY can be simultaneously assigned to the same cache set, what values can the address digits X and Y have?

X: 0, 2, 4, 6, 8, A, C, E
Y: 8, 9, A, B, C, D, E, F

V. Virtual Memory (10 points)

Suppose that a system has a 32-bit (4GB) virtual address space. It has 4GB of physical memory, and uses 4kB pages. Assume each page table entry is 32-bit long and there are 100 programs running in the system at the same time.

1. If this system uses a single-level page table, calculate the total size of the page tables.

   4 MB * 100 = 400 MB

2. If this system uses a two-level page table, what's the **smallest possible** size of all page tables in the memory? In such case, how many bits are there in the 1st-level and 2nd-level page number?

   (4 kB + 4 kB) * 100 = 800 kB
   10 bits and 10 bits