浙江大学2003 —2004学年第 2 学期期终考试
《计算机体系结构》课程试卷

考试时间：__*120*_____分钟 开课学院___计算机_____任课教师_____

姓名_____ 学号_____班级 _____

一	二	三(1)	三(2)	三(3)	三(4)	Total

Answer table for Section One:

1	2	3	4	5	6	7	8	9	10
A	B	D	C	B	A	D	A	D	A
11	12	13	14	15	16	17	18	19	20
A	C	C	D	C/A	A	D	A	B	A
21	22	23	24	25	26	27	28	29	30
C	D	A	C	A	D	C	D	B	B
31	32	33	34	35	36	37	38	39	40
A	C	A	D	A	C	B	D	C	A

一.There are 40 questions or uncompleted statements in this section. Beneath every subject there are a few phases or statements marked A, B, C and D. Choose the statement which correctly answer the question, or the phrase that best completes the sentence. （40）（*Please write your answers in above answer table.*）

1. Amdahl's Law states that the _____ performance improvement to be gained from using some faster mode of execution is limited by the fraction of the time the _____ can be used. Amdahl's Law defines the speedup that can be gained by using a particular feature. The most speedup overall is limited by _____.
A. overall \ enhanced mode \ 1/(1-F) B. enhanced mode \ overall \ 1/(1-F)
C. faster mode \ overall \ 1/(1-F) D. overall \ 1/(1-F) \ enhanced mode

2. Which is the best to be used to evaluate a new computer system?
A. Kernel benchmarks B. Real workload C. Toy benchmarks D. Synthetic benchmarks

3. Which of the following architecture is not used in computers shipping today?
A. load-store B. register-register C. register-memory D. memory-memory

4. The goal is to provide a memory system with cost almost _____ level of memory and speed almost _____ level. The levels of the hierarchy usually subset one another. All data in one level is also found in the level below, and all data in that lower level is found in the one below it, and so on until we reach the bottom of the hierarchy.
A. as low as the cheapest \ as large as the fastest
B. as low as the fastest \ as fast as the cheapest
C. as low as the cheapest \ as fast as the fastest
D. as low as the fastest \ as low as the cheapest

5. Which of the following statements about two-level cache is NOT correct?
   A. With two-Level cache, we can decrease miss penalty.
   B. The first-level cache should be large enough to obtain a small miss rate.
   C. The second-level cache generally uses a bigger block size than that of the first-level cache.
   D. The second-level cache should be large enough and use higher association to catch almost all memory accesses in the second level.

6. All of the following are types of data hazards except _____.
   A. RAR B. WAW C. RAW D. WAR

7. Which of the following descriptions about "Computer architecture" is true?
   A. Computer architecture refers only to instruction set design.
   B. Computer architecture means the implementation of a machine, which has two components: organization and hardware.
   C. Computer architecture design doesn't need to provide support to compilers.
   D. Computer architecture is intended to cover all three aspects of computer design ---- instruction set architecture, organization and hardware.

8. A company's R&D belongs to _____
   A. gross margin B. direct cost C. component cost D. average discount

9. According to the structure of recent compilers, loop transformations belong to which of the following part?
   A. front end per language B. code generator C. global optimizer D. high-level optimization

10. CPU performance mostly refers to _____.
   A. user CPU time. B. Response time C. Elapsed time D. System CPU time

11. Which of the following views is correct?
   A. Since the focus is shifted from computation to communication and storage of information, the importance of I/O is increasing fast than ever.
   B. The I/O performance doesn't matter because the processor is so fast that it always need to wait for human' feedback.
   C. The I/O performance doesn't matter because the CPU will run another task when a process waits for a peripheral, the throughput does not descend.

12. Which of the following relationship is always correct?
   A. geometric mean ≤ arithmetic mean ≤ harmonic mean
   B. harmonic mean ≤ arithmetic mean ≤ geometric mean
   C. harmonic mean ≤ geometric mean ≤ arithmetic mean
   D. arithmetic mean ≤ harmonic mean ≤ geometric mean

13. Compared with the memory-memory architecture, the Register-register architecture has
   A. Higher codes density. B. Less instructions to complete a function.
   C. Lower CPI D. Large variation in instruction size.

14. In the following selections, which is NOT the measurement for reducing the cache miss rate ?
   A. Higher associativity B. Pseudo-associative C. Victim cache D. Write buffer

15. To solve the data hazard in the following instructions, we must_____.

LD R2, 0(R3):	IF	ID	①	EX	②	MEM	③	WB	
ADD R1, R5, R2:		IF	④	ID	⑤	EX	⑥	MEM	WB

   A. Bypassing from ③ to ⑤;
   B. Bypassing from ② to ⑤;
   C. Insert a stall in ⑤;
   D. Bypassing from ② to ⑥;

16. Which RAID level waste most of the storage?
   A. RAID 1 B. RAID 0 C. RAID3 D. RAID 5

17. As processes working, not all objects referenced by a program need to reside in main memory. If the computer has _____, then some objects may reside on ___. The address space is usually broken into fixed-size blocks, called __. At any time, each _ resides either in main memory or on __.
A. cache memory \ main memory \ blocks \ block \ main memory
B. cache memory \ disk \ blocks \ block \ disk
C. virtual memory \ main memory \ pages \ page \ main memory
D. virtual memory \ disk \ pages \ page \ disk

18. To reduce control hazards, we always bring the calculation of branch destination from ____ to _____.
A. EX, ID B. MEM, EX C. EX, IF D. MEM, ID

19. We often use a technique named_____ to solve register allocation problem.
A. colored paging B. graph coloring C. colored graph D. page coloring

20. Which is NOT the characteristic of RISC machine?
A. Powerful instruction functions. B. Use a reduced instruction set.
C. Simple memory addressing mode D. Use the Load/Store architecture.

21. The destination address of a control flow must be specified explicitly in the vast majority of cases, which of the following instruction is the major exception?
A. procedure call B. jump C. procedure return D. branch

22. Which of the following processor is RISC architecture?
A. IBM 360 B. 80x86 C. VAX D. PowerPC

23. The extension of MIPS pipeline to handle multi-cycle operation will bring about_____ hazard.
A. WAW B. WAR C. RAW D. RAR

24. We often use _____ to allocate dynamic objects.
A. register B. stack C. heap D. global data area

25. Which of the following policy will NOT improve virtual memory performance?
A. Write through B. full-associative map C. TLB cache D. LRU replacement

26. In the following selections, which is NOT a measurement for resolution of control hazard?
A. To calculate the branch destination address as in the earlier pipeline stages as possible.
B. Delayed branch.
C. To predict the branch untaken in case avoiding the untaken stalls when branch is really not taken.
D. Double bump.

27. Which of the following descriptions about the Average Selling Price (ASP) is NOT true?
A. The ASP means the component costs adding direct costs and gross margin.
B. If the average discount is cut from the list price, the left is ASP.
C. The Average selling price is just the list price.
D. The ASP is the money that comes directly to the company for each product sold.

28. To solve the control hazard in following instructions, ____ is the best choice to be put into the delay slot.

```
        ADD R3, R1, R2----------①
        BNEZ R1, DES
        < Delay Slot >
        SUB R5, R4, R6----------②
DES:    SUB R7, R9, R8----------③
```

A. It depends B. ③ C. ② D. ①. E. None

29. Source register fetch is completed in _____ clock cycle.
A. IF B. ID C. EX D.MEM E. WB

30. If one functional unit is not fully pipelined, it will lead to_____ hazard. And the division of instruction-memory and data-memory is aimed at solving _____ hazard.

A. Data, Structural B. Structural, Structural C. Control, Control D. Structural, Control

31. Which of the following statements about the causes of cache miss is correct?
A. The obvious way to reduce capacity misses is to increase capacity of the cache, while at the risk of longer hit time and higher cost.
B. The larger the block size is, the better to decrease the conflict misses, because larger size take better advantage of spatial locality.
C. Use larger block size can increase compulsory misses.
D. Higher associativity can be used to reduce conflict misses, and at the same time it decrease the average memory access time too.

32. Which method can NOT be used to reduce cache miss penalty?
A. Multi-level caches B. Victim cache C. Pipelined cache access D. Nonblocking cache

33. To solve the data hazard in the following instructions, the bypassing from _____ to _____ is needed.

ADD R2, R3, R5:	IF	ID	①	EX	②	MEM	③	WB
SUB R1, R4, R2:		IF	④	ID	⑤	EX	⑥	MEM WB

A. ② , ⑤ B. ① , ⑥ C. ③ , ⑥ D. ① , ⑤ E. ②, ⑥

34. The data hazards rises in pipelining due to
① The inherent data dependence among the instructions
② Insufficient function units of pipeline
③ Control instructions such as jump, branch, call or return.
④ The overlapped execution mode for pipelining.
A. ① B. ① and ② C. ① and ③ D. ① and ④

35. Which strategy can NOT be used to solve data hazard?
A. Splitting Cache B. Forwarding C. Pipeline interlock D. Insert stall

36. If a multi-cycle function unit (OP) has the characteristics as following:
latency =6 clock cycle, initial interval =1 clock cycle,
Then which of the following description is NOT correct?
A. If an instruction (J) follows an instruction (I) that use the function OP. And instruction J will use the result of instruction I, then instruction J should enter the pipeline at least 6 clock cycles later after the instruction I into the pipeline in case read the wrong result.
B. The function unit OP is a full-pipelined unit.
C. If an instruction (J) enter the pipeline just after instruction (I) and both instruction (I) and (J) need to use the function unit OP, then a structural hazard will occur.
D. It takes seven clock cycles for the function unit OP to finish its operation.

37. In a cache-memory hierarchy system, assume that the memory size is 256MB, with a 4KB write back cache in 2-way associative. The block size is 32B. Then the size of index field of physical memory address is
A. 5 bit B. 6 bit C. 7 bit D. 11 bit E. 17 bit

38. Assume there are M blocks in a cache, and every K blocks are grouped in one set, then which following description is NOT correct?
A. If K=1, then it's a direct mapped cache.
B. If K=1, then it's a one-way set associative cache.
C. If K=M, then it's a full-associative cache.
D. If K>1 and K<M, then it's a M/K-way set associative cache.

39. Computer pioneers correctly predicted that programmers would want unlimited amounts of fast memory. An economical and palmary solution to that desire is _____, which takes advantage of ___ and cost/performance of memory technologies.
A. the Amdahl's Law \ principle of locality
B. a memory interleaved organization \ principle of locality
C. a memory hierarchy \ principle of locality
D. a memory hierarchy \ the Amdahl's Low

40. Assume there is a code segment as following. And the elements in arrays are place in a row-and-row order.

```
for (j = 0; j< 100; j = j+1)
        for (i = 0; i < 5000; i = i+1)
                x[i][j] = x[i][j] + C;     /* C is a constant.  */
```

Some one suggests to optimize the above code by exchanging the nesting of the loops as following:

```
for (i = 0; i < 5000; i = i+1)
        for (j = 0; j< 100; j = j+1)
                x[i][j] = x[i][j] + C;     /* C is a constant.  */
```

Which of the following statements is correct ?

A. The optimization can decrease cache misses by improving the spatial locality.
B. The optimization can decrease cache misses by improving the temporal locality.
C. The optimization can decrease cache misses by improving both the temporal locality and spatial locality.
D. This measurement can not decrease cache misses at all.

二.Fill in the blanks (24, with each 2)

1. Suppose a computer spends 90 percent of its time handling a particular type of computation when running a given program, and its manufacturers make a change that improves its performance on that type of computation by factor of 10. The speedup is __5.26 (5.3)__.

2. Suppose the hardware implementation is the classic 5-stage RISC pipeline. Unconditional branch is resolved after the end of ID stage, while the branch-target address is known at ID too. But the branch condition is evaluated till the end of EX stage. The branch strategy is predict-taken. Then how many stalls must each type of instruction take?

Branch taken:__1__cycles; Branch untaken: __2__cycle

3. A cache has 64-KB capacity, 128-bytes/line, and is 4-way set-associative. The system containing the cache uses 32-bit addresses.

The cache has ___512___ lines. The cache has ____128____ sets.

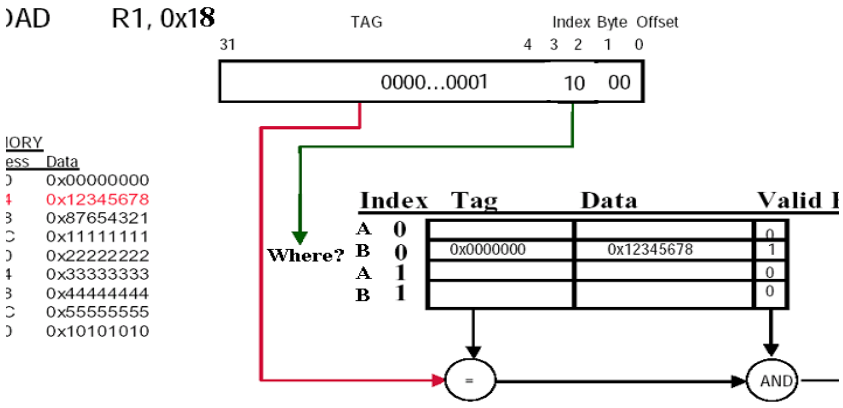
Tag information is__18__ bits.

4. In 2-way set-associative cache, assume cache has 4 blocks and each block is 1 word and 2 blocks per set. For instruction LOAD R1, 0x18, is memory access misses? If the access misses, will replacement occur? And where is the location that the new loaded block will be located?

Memory access to 0x18 will (miss or hit) ___miss_____ in the cache .

Whether there is a replacement (yes or no) ____No_____ .

Block will be place in Set __0__ and Block __A__ .



5. Assume the performance of the basic memory organization is:

4 clock cycles to send the address
56 clock cycles for the access time per word
4 clock cycles to transfer a word of data

Given a cache block of four words, and that a word is 8 bytes, the miss penalty is ___256_____(Calculation expression should be given out.) clock cycles, with a memory bandwidth of __1/8__ bytes per clock cycle.

三.Calculations（36）

1. (10)Your company is developing a program with high requirement on computation. You asked your R&D department to make some improvements on the execution time. After several months, they give you two solutions. The first one is to use a new hardware technology, by which 40% of the computation can be accelerated by 10 times. Another solution is focused on algorithm design, which can enhance 60% and 10% of the total computation by 2 and 20 times respectively.
Question:
a) What is the overall enhancement of the hardware solution?
b) What is the overall enhancement of the software solution?
c) Which one will you choose?

Answer:
a):

$$Speedup_{overall-hardware} = \frac{1}{(1-40\%)+\frac{40\%}{10}} = \frac{1}{0.64} = 1.5625$$

b):

$$Speedup_{overall-software} = \frac{1}{(1-60\%-10\%)+\frac{60\%}{2}+\frac{10\%}{20}} = \frac{1}{0.605} = 1.6529$$

c): Software solution is better.

2. (13)Within some **memory/cache** memory hierarchies, there are 2 words in a block. Access time form Cache is 8ns and for main memory miss penalty is 70ns. For the code of C language below, assumes that each element is one word in array (**A[i]**). Except array, another variables has be loaded to registers. While the C codes execute, please calculate and questions below:

```
for ( i=0; i<100; i++)
    s=s+A[i] ;
```

(1) What is the miss rate for data accesses?
(2) What is the average memory access time for data read?
(3) What is the overall CPI including memory access? Assume processor runs at 1.1GHz and has a CPI of 1 excluding memory accesses. Ignores instructions misses and data hazard and control hazard. Assumes assembler code is below:

```
                ……
LOOP:   LOAD    R2,   0(R1)
                ADD   R5,R1,#4
        ADD     R3,   R2,R3        ;s was stored in R3
        BNE     R5,   LOOP
                …….
```

Answer:
assumed condition:

Block	**1 word/block**
Access time of cache(hit time)	**8ns**
Access time of memory (miss penalty)	**70ns**
CPU clock rate	**1GHz**
Ideal CPI	**1**
All memory accesses	**100**
Clocks for one accesses	**time/T=time×f=70ns×1.1GHz=77**

(1) For data accesses
Misses Accesses for even elements: A[0],A[2],……

There are 50 misses accesses

Miss rate for data is 50/100=50%

(2) Average memory accesses time

AMAT＝hit time×(1-miss rate) ＋ miss rate×miss penalty

＝8×(1-50%) ＋50%×70

＝39ns

(3) Overall CPI

Number of instructions are 400.

CPI＝ideal CPI ＋ (Misses/instructions) × miss penalty

＝1＋50/400 × 77

＝9.625

3. (13) Consider the following pipeline. All instructions have five cycles but autoincrement addressing instruction, which is IF, ID, EX, MEM, WB. Branch will complete at the third cycle. The pipeline extended MIPS pipeline in autoincrement addressing which have seven pipe stages. The Fig 1 is an example in autoincrement addressing:

eg:	Add R1, (R2)+
means:	Regs[R2]←Regs[R2]+4
	Regs[R1]←Regs[R1] + Mem[Regs[R2]]

Fig 1: example instruction of autoincrement addressing

The register files can perform two reads and two write mode every clock cycle. To handle reads and writes to the same register, assume the register write in the first half of the clock cycle and the read in the second half.

Read the following code segment . The pipeline has forwarding path.

```
Loop:  LW R1, 4(R2)
       ADD R2, (R1)+
       ADDI R3, R3, #4
       SUB R4, R1, R2
       SW R2, 4(R4)+
       BNEZ R3, Loop
```

IF: Instruction fetch

ID: Instruction decode

ADDR: AutoIncrement Addressing

WB1: Write Result of ADDR to Register file

EX: Memory Reference: Calculate the absolute address

ALU Instruction: Calculate.

MEM: Memory Access

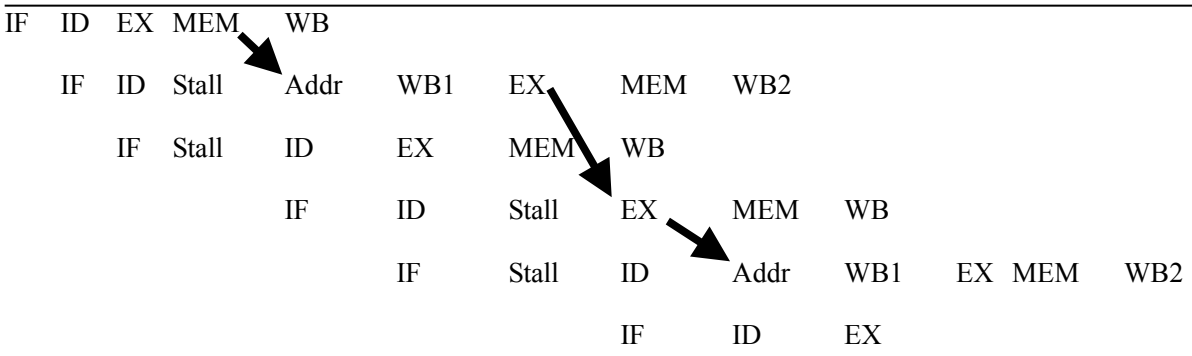
WB2: Write Result to Register file

Fig 2: seven pipeline stage of the autoincrement addressing

(question1)if the pipeline has no delay slot, find out how forwarding path work in every instruction? Draw the pipe stage diagram, mark the every forwarding path in diagram.

(question2)if the pipeline has one delay slot, how to adjust the code segment. Draw the pipe stage diagram.

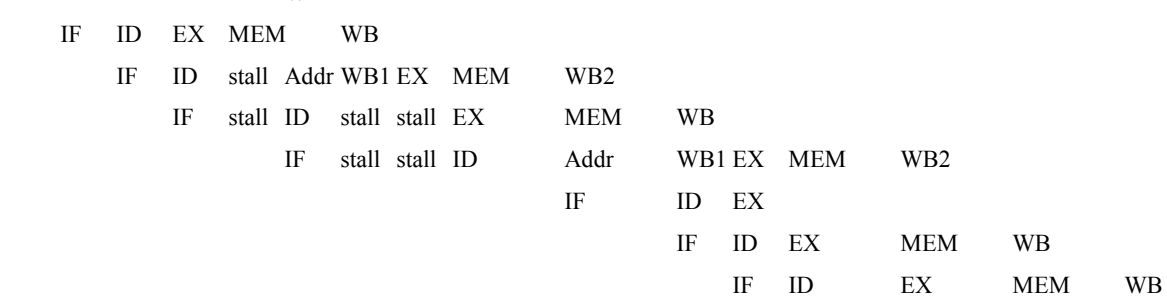
答案1：



评分标准：6分，每条指令0.5分，3个箭头各1分

答案2：

```
    ADDI R3, R3, #4
  Loop: LW R1, 4(R2)
      ADD R2, (R1)+
      SUB R4, R1, R2
      SW R2, 4(R4)+
      BNEZ R3, Loop
      ADDI R3, R3, #4   延时槽
    ADDI R3, R3, #-4
```



评分标准：延时槽内指令1分，循环前后预处理与后处理语句各0.5分。流水线状态图2分

如果有学生画成流水线时空图，看答案是否正确给分。