第一大題單擇題和第二大題多選題考生應作答於『答案卡』(請勿作答於試卷之選擇題作答區)，未作答於答案卡者，本兩大題不予計分；非選擇題則請作答於『試卷』之非選擇題作答區。

一、單選題 (每題 5 分，共 25 分)

1. Consider a program P that originally runs on a single processor. Assume we have 6 processors now. What is the minimum percentage of P that can be parallelizable to attain 3X speed-up using 6 processors?
   a. 75%
   b. 50%
   c. 80%
   d. 66%
   e. None of the above is correct
2. There are a primary cache L1 and a primary memory RAM. The access times of L1 and RAM are 1ns and 100ns, respectively. The miss rate of L1 is 0.05%, so the average access time is 6ns. To speed up the memory system, we add a secondary cache L2 with access time 10ns. For a 2X speed-up, what should the maximum miss rate of L2 be?
   a. 50%
   b. 20%
   c. 30%
   d. 15%
   e. None of the above is correct
3. Consider an 8-way set-associate cache. Assume that a physical address has 32 bits, the size of a block is 16 words, and the cache contains 64K bytes data. How many bits does the cache at least need? (Remember the valid bits)
   a. 28672
   b. 29696
   c. 34816
   d. 20480
   e. None of the above is correct
4. Consider a processor with an instruction cache and a memory cache. Assume that the basic CPI (i.e., for an ideal cache) is 1.5, the miss rate of the instruction cache is 2%, the miss rate of the memory cache is 5%, the miss penalty is 80 cycles, and 40% of instructions access the memory. What is the effective CPI?
   a. 7.1
   b. 3.74
   c. 6.14
   d. 4.7
   e. None of the above is correct
5. Consider the following assembly program. There are one input parameter n and one output parameter b. We assume that n is a positive integer. The input parameter n is stored in the address 0(x1) before the program starts, and the output parameter b needs to be stored in the address 4(x1) after the program ends.

```
1        ld     x5,   0(x1)
2        addi   x6,   x0,   1
3        add    x7,   x0,   x0
4  LOOP: bge    x6,   x5,   END
5        add    x6,   x6,   x6
6        addi   x7,   x7,   1
7        j      LOOP
8  END:  sd     X7,   4(x1)
```

見背面

Which of the following statements is correct?

a. If n = 3, b= 1.
b. If n = 5, b= 4.
c. $b = \lceil \log_2 n \rceil$, i.e., b is a round-up for the binary logarithm of n.
d. If n = 9, b=3.
e. None of the above is correct

二、多選題 (每題 5 分，共 25 分)

6. Consider three instructions and two programs that uses the three instructions as follows.

Instruction	A	B	C
CPI	3	1	2

	A	B	C
P1	2	2	1
P2	1	4	3

Which of the following statements is/are true?

a. CPI of P1 is 2
b. CPI of P2 is 1.6
c. CPI of P2 is 1.625
d. P1 is 1.3X faster than P2
e. P2 is 1.3X faster than P1

7. Consider the floating point numbers defined by IEEE 754-1985. A floating point number X is represented as follows:

$$X = (-1)^{Sign} \times (1 + Fraction) \times 2^{Exponent-Bias}.$$

In single precision, Exponent has 8 bits, Fraction has 23 bits and Bias is 127; In double precision, Exponent has 11 bits, Fraction has 52bit and Bias is 1023. In the corresponding binary string, Sign lies in the left, Exponent lies in the middle, and Fraction lies in the right. Remember that Exponents 0000...00 and 1111...11 are reserved, i.e., in single precision, Exponent is from 1 to 254, and in double precision, Exponent is from 1 to 2046.
Which of the following statements is/are true?

a. In double precision, the largest positive number is $(2 - \frac{1}{2^{52}}) \times 2^{1023}$.
b. In double precision, the smallest normalized positive number is 2^{-1022}.
c. In single precision, the binary string of -0.3125 is 10111110101000000000000000000000.
d. In single precision, the binary string 00111111011100000000000000000000 represents 0.875.
e. In double precision, the smallest denormalized positive number is 2^{-1023}.

8. A simplified RISC datapath consists of five stages: Instruction Fetch, Instruction Decode, Execution, Memory Access and Write Back. The processing time for these five stages are 0.5 ns, 0.35 ns, 0.25 ns, 0.5 ns and 0.4 ns.
Which of the following statements is/are true?

a. The clock rate is at most 500 MHZ.
b. If we pipeline the datapath, the clock rate is at most 2 GHZ
c. If we pipeline the datapath, the clock rate is at least 4 GHZ
d. The speed-up by pipelining the datapath is at most 4 times.
e. The speed-up by pipelining the datapath is at least 2 times

接次頁

9. Consider a memory system in which a virtual address has 40 bits, a physical address has 32 bits, a page has 16K bytes, and a physical address stores a byte.
   a. There are 2^23 virtual pages.
   b. There are 2^26 virtual pages.
   c. The page offset has 14 bits.
   d. There are 2^18 physical pages.
   e. There are 2^15 physical pages.

10. Consider the following simple assembly code:

```
1  ld    x5,  0(x1)
2  addi  x5,  x5,  2
3  ld    x6,  4(x1),
4  addi  x7,  x5,  x6
5  sd    x7,  8(x1)
```

We assume that the pipelined datapath consists of five stages: Instruction Fetch, Instruction Decode, Execution, Memory Access and Write Back. Although some instructions do not perform the fifth stage, please also count the fifth stage Which of the following statements is/are true?

   a. Without data forwarding, this code requires 13 cycles.
   b. With data forwarding, this code requires 11 cycles.
   c. Re-ordering this code can totally avoid data hazard.
   d. Re-ordering this code and forwarding data can attain 9 cycles
   e. Re-ordering this code and forwarding data only can attain 10 cycles

三、問答題 (每題 5 分，共 50 分) ※ 注意：請於試卷內之「非選擇題作答區」依序作答，並應註明作答之部份及題號。

11. SRTF is the Shortest Remaining Time First scheduling algorithm. Please write down a set of processes to explain how SRTF can result in starvation.

12. Continuing from problem 11, let us consider the Highest Response Ratio Next (HRRN) scheduling algorithm in which the priority of a process at a timepoint is defined as (accumulated waiting time + remaining time)/remaining time. Please explain how the HRRN algorithm can avoid starvation to the set of processes in your answer to problem 11.

13. Please write down a short program in pseudo code to explain data parallelism.

14. Please write down a short program in pseudo code to explain task parallelism.

15. Please explain how a thread library manage to implement a multi-thread program in a single kernel thread.

16. Please write down a page-reference sequence to explain how the LFU (least frequently used) algorithm can perform worse than the reference-bit algorithm.

17. Please explain NUMA in memory management and its design and function in a symmetric multiprocessing system.

18. Please write down a multithread program in pseudo code to explain race condition.

19. Please write down a multithread program in pseudo code to explain deadlock.

20. Please explain the principle of locality and how it affects the performance of paging systems.

試題隨卷繳回