※ 注意：請用 2B 鉛筆作答於答案卡，並先詳閱答案卡之「畫記說明」。全部題目均為複選題。

1. (2 points) Please answer which of the following are performance-enhancing techniques of operating systems.
   (a) Switching execution order of machine instructions.
   (b) Switching load/store operation orders from cache to main memory.

2. (2 points) Please answer which of the following are performance-enhancing techniques of operating systems.
   (a) Cache line replacement.
   (b) Multi-core load balance.
   (c) Inter-procedural analysis.

3. (2 points) In deadlock avoidance, we may use Banker's algorithm to check whether a state is safe or not. Suppose that we have the following state of resource allocation.

	Allocation	Max	Available
	A B C D	A B C D	A B C D
P0	1 1 1 0	3 2 6 0	3 2 3 3
P1	0 0 1 1	2 2 3 2	
P2	2 2 2 3	5 3 2 3	
P3	3 0 3 2	3 3 4 3	

   Please answer which of the following statements are correct!
   (a) A request by P0 of (0,1,1,0), if granted, will make the system unsafe.
   (b) A request by P0 of (2,1,3,0), if granted, will deadlock the system.

4. (2points) Continuing from the last question, please answer which of the following statements are correct ?
   (a) A request by P1 of (0,1,1,0), if granted, will make the system unsafe.
   (b) A request by P2 of (3,1,0,0), if granted, will deadlock the system.
   (c) A request by P3 of (0,1,1,0), if granted, will make the system unsafe.

5. (2 points) Please answer which of the following statements about NUMA are correct!
   (a) Modern storage hierarchy results in NUMA.
   (b) Cache technology results in NUMA.

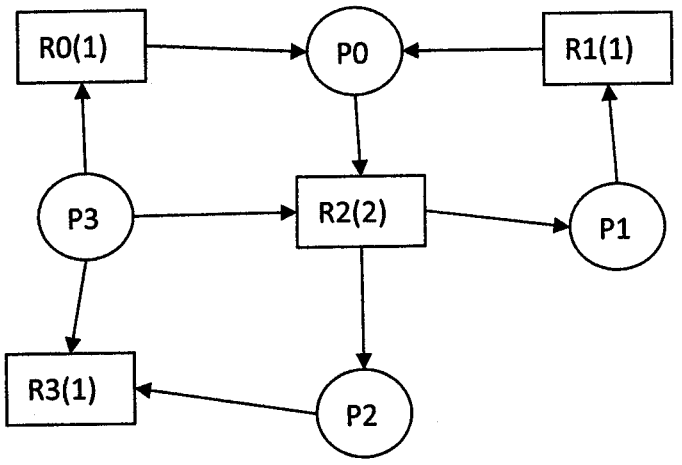
6. (2 points) Continuing from the last question, please answer which of the following statements about NUMA are correct!
   (a) It is intrinsic in von Neumann's computer model.
   (b) Pipeline results in NUMA.
   (c) It is an optimization technique of compiler.

7. (2 points) Please answer which of the following statements about Byzantine generals' problem are correct!
   (a) It is a problem in cryptography.
   (b) It is a problem in fault-tolerance computing.

8. (2 points) Please answer which of the following statements about Byzantine generals' problem are correct!
   (a) It is a problem in distributed computing.
   (b) It is a problem in deadlock avoidance.
   (c) It is a problem in disk scheduling.

見背面

9. (2 points) Assume that we have the following resource allocation graph of a state in a dining philosopher problem with four philosophers: P0, P1, P2, and P3.



There are four non-preemptive resource types: R0, R1, R2, and R3. The notations in the box Rj(k) means resource type j has k copies. For a philosopher to complete her task, she need a copy of every resource type that has an edge connected to her. Which of the following statements are correct ?

(a) There is a deadlock in this state.

(b) The state is safe according to the Banker's algorithm.

10. (2 points) Continuing from the last question, please answer which of the following statements are correct.

(a) If philosophers with lower indices have resource allocation priority, then in this system, there should be no deadlock.

(b) If philosophers request the resources in the ordering of the resource type indices, then in this system, there should be no deadlock.

(c) The reader-writer spin-locks of Linux can prevent deadlock in this system.

11. (2 points) Please answer which of the following are correct statements.

(a) L1 cache is closer to CPU than L2 cache.

(b) Compilers use the total cache size to tune the machine code performance.

12. (2 points) Please answer which of the following are correct statements.

(a) Compilers use the cache line size to tune the machine code performance.

(b) Compilers use cache hit ratio to tune machine code performance.

(c) Compilers use cache total store order to tune machine code performance.

13. (2 points) Assume a demand-paging system with TLB access time = 1 nanosecond, memory access time = 26 nanoseconds, and disk backing store access and transfer time = 1 milliseconds. Which of the following statements are true of the system?

(a) When TLB miss ratio is 1% and the total page fault ratio is 0.0001%, the effective memory access time is 27.27 nanoseconds with a 1-level page table.

(b) When TLB miss ratio is 0.1% and the total page fault ratio is 0.001%, the effective memory access time is 27.072 nanoseconds with a 2-level page table.

接次頁

14. (2 points) Continuing from the last question, which of the following statements are true of the system ?
   (a) When TLB miss ratio is 0.2% and the total page fault ratio is 0.002%, the effective memory access time is 27.276 nanoseconds with a 3-level page table.
   (b) When TLB miss ratio is 0.4% and the total page fault ratio is 0.0005%, the effective memory access time is 27.496 nanoseconds with a 4-level page table.
   (c) When TLB miss ratio is 0.3% and the total page fault ratio is 0.0009%, the effective memory access time is 27.525 nanoseconds with a 5-level page table.

15. (2 points) Which of the following statements are correct of cloud computing and services?
   (a) IaaS supports on-demand access to commercial information services.
   (b) PaaS supports a development environment to application development teams.

16. (2 points) Which of the following statements are correct of cloud computing and services?
   (a) According to NIST's definition, SaaS supplies on-demand system resources, like CPU cycles, memory capacities, and network storage to application development teams.
   (b) Auto-scaling means automatically adapting resources supplied by a cloud service according to the workload.
   (c) In cloud service, we use virtual machines that run a container for each thread of client applications.

17. (2 points) Consider that there are two threads with global variable A and B initialized to zero and a binary semaphore S also initialized to zero. Which of the following statements are correct? For convenience, we assume that the print-out value is (x,y) where x and y are the print-out va of thread 0 and 1 respectively. Also assume that cache load/store ordering adheres to their issue ordering.
   (a) If the two threads run the following code, the print-out can never be (0,0).

Thread 0	*Thread 1*
A = 1	B = 1
Print B	Print A

   (b) If the two threads run the following code, the print-out is (0, 1).

Thread 0	*Thread 1*
A = 1	Wait(S)
Print B	B = 1
Signal(S)	Print A

見背面

18. (2 points) Continuing from the last question, which of the following statements are correct?
(a) If the two threads run the following code, the print-out can be (1,0).

Thread 0	Thread 1
Wait(S)	B = 1
A = 1	Wait(S)
Print B	Print A
Signal(S)	Signal(S)

(b) If the two threads run the following code, the print-out can be (1,1).

Thread 0	Thread 1
Signal(S)	B = 1
A = 1	Wait(S)
Wait(S)	Print A
Print B	

(c) If the two threads run the following code, the print-out can be (0,1).

Thread 0	Thread 1
A = 1	B = 1
Signal(S)	Wait(S)
Print B	Print A
Wait(S)	Signal(S)

19. (2 points) Please answer which of the following statements are correct.
(a) malloc() may cause deadline missing because of memory compaction.
(b) malloc() may cause deadline missing because of garbage collection.

20. (2 points) Please answer which of the following statements are correct.
(a) malloc() may cause deadline missing because of hyper-threading.
(b) Virtual memory may cause deadline missing because of page replacement.
(c) Virtual memory may cause deadline missing because of out-of-order execution.

21. (2 points) According to Amdahl's law, please answer which of the following statements are correct. Assume that in the source-code level, P is the percentage of parallelizable execution time over the total execution time of a system.
(a) If P = 30%, then the speed-up ratio of parallelization is at most 10/7.
(b) If P = 60%, then the speed-up ratio of parallelization is at most 10/6.

22. (2 points) Continuing from the last question, which of the following statements are correct?
(a) If P = 80% and there are four cores, then the speed-up ratio of parallelization is at most 3.
(b) If the parallelization does not reach the speed-up upper-bound predicted by Amdahl's law, the reason could be locking of shared variables.
(c) If the parallelization does not reach the speed-up upper-bound predicted by Amdahl's law, the reason could be out-of-order execution arranged by the compiler.

23. (2 points) Please answer which of the following statements are correct.
(a) GPU uses pipelines.
(b) GPU is for task parallelism.

接次頁

24. (2 points) Please answer which of the following statements are correct.
    (a) GPU is for MIMD parallelism.
    (b) May batch groups of threads to minimize incoherent memory access.
    (c) GPU consists of stream multi-processors.

25. (2 points) Please answer which of the following statements are correct.
    (a) Physical caches are implemented inside physical main memory.
    (b) Physical caches do not need flush at context switching.

26. (2 points) Please answer which of the following statements are correct.
    (a) Pure physical caches are in general slower than pure virtual caches.
    (b) Virtual caches do not need address translation in at a cache hit.
    (c) Modern physical caches carry out address translation in parallel with cache accesses.

27. (2 points) Please answer which of the following statements are correct.
    (a) A superscalar processor can also be a pipeline processor.
    (b) Superscalar processors gain their performance by specific hardware in calculating variables of super sizes.

28. (2 points) Please answer which of the following statements are correct.
    (a) Memory-level parallelism uses data duplication in main memory to support parallel read/write to the main memory.
    (b) Memory-level parallelism uses out-of-order loads and stores between caches and main memory.
    (c) Memory-level parallelism uses memory data prefetching.

29. (2 points) Please answer which of the following statements are correct.
    (a) Out-of-order execution in the cache level may fail and result in revocation of previous operations.
    (b) Compilers use the technology of heap analysis to generate machine code with out-of-order execution.

30. (2 points) Please answer which of the following statements are correct.
    (a) Compilers use the technology of pointer analysis to generate machine code with out-of-order execution.
    (b) Compilers use the technology of alias analysis to generate machine code with out-of-order execution.
    (c) Superscalar processors detect instruction-level parallelism and carry out out-of-order execution.

31. (8 points) Assume that we have three classes of instructions with the following CPI values.

    Class A: CPI = 1, Class B: CPI = 2, Class C: CPI = 4

    For convenience, given a program, we use (a,b,c) to represent the numbers of CPU-issued instructions in class A, B, and C respectively of the program. Assuming that all instructions are executed without superscalar and pipeline technology, which of the following statements are true?
    (a) A program with (200,3000,800) has CPI = 2.30.
    (b) A program with (100,4000,600) has CPI= 2.05.
    (c) The size of a program with (200,3000,800) is 9400 bytes.
    (d) The size of a program with (100,40000,600) is 82500 bytes.
    (e) The execution of a program with (300,600,10) consumes 910 cycles.

見背面

32. (8 points) Suppose that we have a memory system with 48-bit addresses and a 1024 Kilobyte cache. The size of cache line (block) is 64 bytes. Which of the following statements are true of the system?
(a) When the cache is 8-way associative, there are 4K cache lines.
(b) When the cache is 8-way associative, the index field is 12 bits long.
(c) When the cache is 32-way associative, the index field is 10 bits long.
(d) When the cache is direct-mapped, the tag field is 42 bits long.
(e) When the writing policy is `write-through', the dirty data are written to the main memory only until the corresponding cache line becomes a victim in the replacement policy.

33. (8 points) Please answer which of the following are correct statements of context switching.
(a) Dispatch latency is an overhead in context switching.
(b) Dispatch latency can be reduced by cache-threading.
(c) Invoking operating system service may incur context switching.
(d) The first thing to do in context switching is checking the cause of the context switching.
(e) Context switching from a kernel process to a user process are finalized by executing two machine instructions: the first jumps to the user program machine instruction and the second set the CPU mode to user mode.

34. (8 points) Please answer which of the following are correct statements of hyper-threading.
(a) It is a superscalar technology.
(b) Hyper-threading performance is related to the number of duplicate ALUs in CPU.
(c) Hyper-threading performance is related to the number of stages in CPU pipeline.
(d) Hyper-threading can speed up cache line replacement.
(e) Hyper-threading can speed up context-switching.

35. (8 points) Which of the following statements are correct of compiler optimization techniques ?
(a) Inter-procedural analysis uses table lookup to skip procedure execution if a return value is already recorded in the table.
(b) Live variable analysis saves memory usage by freeing the space of those variables that will not be used any more.
(c) Changing while-loop to repeat-loop can lower the possibility of pipeline flush.
(d) Inline replacement can save push and pop operations.
(e) Strength reduction means replacing an expensive piece of code with an inexpensive but equivalent alternative.

試題隨卷繳回