PART I: COMPUTER ARCHITECTURE

1. (10 pts) Answer the following questions briefly:

   (a) [2 pts] Typically one CISC instruction, since it is more complex, takes more time to complete than a RISC instruction. Assume that an application needs *N* CISC instructions and *2N* RISC instructions, and that one CISC instruction takes an average of *5T* ns to complete, and one RISC instruction takes *2T* ns. Which processor has the better performance?

   (b) [2 pts] Which of the following processors have a CISC instruction set architecture?

ARM	AMD Opteron
Alpha 21164	IBM PowerPC
Intel 80x86	MIPS
Sun UltraSPARC	

   (c) [6 pts] True & False questions: (2 pts each)
   (1) There are four types of data hazards: RAR, RAW, WAR, and WAW. (True or False?)
   (2) AMD and Intel recently added 64-bit capability to their processors because most programs run much faster with 64-bit instructions. (True or False?)
   (3) With a modern processor capable of dynamic instruction scheduling and out-of-order execution, it is better that the compiler does not to optimize the instruction sequences. (True or False?)

2. (5 pts) For commercial applications, it is important to keep data on-line and safe in multiple places.

   (a) [2 pts] Suppose we want to backup 100GB of data over the network. How many hours does it take to send the data by FTP over the Internet? Assume the average bandwidth between the two places is 1Mbits/sec.

   (b) [3 pts] Would it be better if you burn the data onto DVDs and mail the DVDs to the other site? Suppose it takes 10 minutes to burn a DVD which has 4GB capacity and the fast delivery service can deliver in 12 hours.

3. (10 pts) Suppose we have an application running on a shared-memory multiprocessor. With one processor, the application runs for 30 minutes.

   (a) [2pts] Suppose the processor clock rate is 2GHz. The average CPI (assuming that all references hit in the cache) on single processor is 0.5. How many instructions are executed in the application?

見背面

(b) [2 pts] Suppose we want to reduce the run time of the application to 5 minutes with 8 processors. Let's optimistically assume that parallelization adds zero overhead to the application, i.e. no extra instructions, no extra cache misses, no communications, etc. What fraction of the application must be executed in parallel?

(c) [3 pts] Suppose 100% of our application can be executed in parallel. Let's now consider the communication overhead. Assume the multiprocessor has a 200 ns time to handle reference to a remote memory and processors are stalled on a remote request. For this application, assume 0.2% of the instructions involve a remote communication reference, no matter how many processors are used. How many processors are needed at least to make the run time be less than 5 minutes?

(d) [3 pts] Following the above question, but let's assume the remote communication references in the application increases as the number of processors increases. With *N* processors, there are *0.02*(N-1)*% instructions involve a remote communication reference. How many processors will deliver the maximum speedup?

4. (10 pts) Number representation.

(a) What range of integer number can be represented by 16-bit 2's complement number?

(b) Perform the following 8-bit 2's complement number operation and check whether arithmetic overflow occurs. Check your answer by converting to decimal sign-and-magnitude representation.

```
    11010011
  - 11101100
```

5. (15 pts) Bus

(a) Draw a graph to show the memory hierarchy of a system that consists of CPU, Cache, Memory and I/O devices. Mark where memory bus and I/O bus is.

(b) Assuming system 1 have a synchronous 32-bit bus with clock rate = 33 Mhz running at 2.5V. System 2 has a 64-bit bus with clock rate = 66Mhz running at 1.8V. Assuming the average capacitance on each bus line is 2pF for bus in system 1. What is the maximum average capacitance allowed for the bus of system 2 so the peak power dissipation of system 2 bus will not exceed that of the system 1 bus?

(c) Serial bus protocol such as SATA has gained popularity in recent years. To design a serial bus that supports the same peak throughput as the bus in system 2, what is the clock frequency of this serial bus?

接次頁

PART II – OPERATING SYSTEMS

6. (10 pts) If you were to solve the synchronization problem such as double booking（一票雙賣）in high speed rail（高鐵）operation, what kind of centralized or distributed booking system would you design? Please describe the most important technology needed to solve the problem, list the advantages and disadvantages of centralized and distributed solutions, and why you choose your solution.

7. (15 pts) Explain the differences between a user-level thread, a kernel-level thread, and a process. (Hint: Not only explain each one.) Give three scenarios (exclude the following example) why and where one performs better than the other. E.g. user-level multi-threading performs better than single process even in single-processor system when doing some massive big matrix multiplication with proper programming of CPU (multiplication) and IO overlapping (input from and output to file) because the IO unit might be considered another processor to take advantages of multithreading on multiprocessor system.

8. (5 pts) Contiguous allocation of files leads to disk fragmentation because some space in the last disk block will be wasted in files whose length is not an integral number of blocks. Is this internal fragmentation or external fragmentation? Can you make an analogy with memory allocation?

9. (10 pts) In distributed systems, it is very often that a coordinator is needed to synchronize the work among the processes in the systems. Election algorithms such as bully and ring-based algorithms are designed to determine which process is the coordinator process. Assume that there are *n* processes in the system and each process has a unique label. Please answer the following questions:
   (a) In the conventional ring-based election algorithm, a unidirectional ring is used. What is the number of messages needed in worst case and in average to elect a new coordinator when only the coordinator process fails?
   (b) In order to reduce the size of sent messages, Bob decides to include only the largest process ID in the message, rather than all the IDs for the processes receiving election messages. Can Bob's algorithm still work? Please explain your answer.

10. (10 pts) Disk I/O has been a critical overhead in the operating systems. FCFS Scheduling, SSTF Scheduling, SCAN scheduling C-SCAN scheduling, and LOOK scheduling algorithms are five common disk scheduling algorithms. Please answer the following questions.
    (a) [6 pts] In time-sharing systems, fairness is an important metric. Which of the above algorithms are not fair? Please explain your answer.
    (b) [4 pts] Under different system workload patterns, we may have to choose different disk scheduling algorithms. Please choose a best scheduling algorithm for the following two workload patterns.
        i. Lightly loaded system that has occasional burst of disk IO accesses.
        ii. Heavily loaded system which rarely has an empty disk IO request queue.

試題隨卷繳回