CSCI 3323 (Principles of Operating Systems), Fall 2013

Homework 6

Credit: 30 points.

1 Reading

Be sure you have read (or at least skimmed) Chapters 4, 5, and 6.

2 Problems

Answer the following questions. You may write out your answers by hand or using a word processor or other program, but please submit hard copy, either in class or in my mailbox in the department office.

1. (5 points) The textbook describes more than one strategy for keeping track of free blocks in a file system (free blocks, bitmaps, and FATs). All of these strategies rely on information that is kept both on disk and in memory, sometimes with the most-current information only in memory. What would happen if the copy on disk of whatever data structure is used to keep track of free blocks was lost or damaged because of a system crash — is there a way to recover, or do you have to just reformat the disk and hope you backed up any really important files? Answer separately for MS-DOS FAT-16 (which uses a FAT) and UNIX V7 filesystems (which uses one of the other strategies).

2. (5 points) Consider a UNIX filesystem (as described in section 4.5.3) in which each i-node contains 10 direct entries, one single-indirect entry, one double-indirect entry, and one triple-indirect entry. If a block is 1KB (1024 bytes) and a disk addresses is 4 bytes, what is the maximum file size, in KB? (*Hint:* Use the blocksize and size of disk addresses to determine how many entries each indirect block contain.)

3. (5 points) Consider the following two I/O devices. For each device, say whether you think programmed I/O or interrupt-driven I/O makes the most sense, and justify your answer. (*Hint:* Consider the time required for interrupt processing versus the time needed for the actual input/output operation.)

   (a) A printer that prints at a maximum rate of 400 characters per second, connected to a computer system in which writing to the printer's output register takes essentially no time, and using interrupt-driven I/O means that each character printed requires an interrupt that takes a total of 50 microseconds (i.e., 50×10^{-6} seconds) to process.

   (b) A simple memory-mapped video terminal (output only), connected to a system where interrupts take a minimum of 100 nsec to process and copying a byte into the terminal's video RAM takes 10 nsec.

4. (5 points) The textbook divides the many routines that make up an operating system's I/O software into four layers. In which of these layers should each of the following be done? Why? (Assume that in general functionality should be provided at the highest level at which it makes sense — e.g., in user-level software rather than device-independent software.)

(a) Converting floating-point numbers to ASCII for printing.

(b) Computing the track, sector, and head for a disk read operation.

(c) Writing commands to a printer controller's device registers.

(d) Detecting that an application program is attempting to write data from an invalid buffer address. (Assume that detecting an invalid buffer address can only be done in supervisor mode.)

5. (5 points) Suppose at a given point in time a disk driver has in its queue requests to read cylinders 10, 22, 20, 2, 40, 6, and 38, received in that order. If a seek takes 5 milliseconds (i.e., 5×10^{-3} seconds) per cylinder moved, and the arm is initially at cylinder 20, how much seek time is needed to process these requests using each of the three scheduling algorithms discussed (FCFS, SSF, and elevator)? Assume that no other requests arrive while these are being processed and that for the elevator algorithm the initial direction of movement is outward (toward larger cylinder numbers).

6. (5 points) Suppose you are designing an electronic funds transfer system, in which there will be many identical processes that work as follows: Each process accepts as input an amount of money to transfer, the account to be credited, and the account to be debited. It then locks both accounts (one at a time), transfers the money, and releases the locks when done. Many of these processes could be running at the same time. Clearly a design goal for this system is that two transfers that affect the same account should not take place at the same time, since that might lead to race conditions. However, no problems should arise from doing a transfer from, say, account A to account B at the same time as a transfer from account C to account D, so another design goal is for this to be possible. The available locking mechanism is fairly primitive: It acquires locks one at a time, and there is no provision for testing a lock to find out whether it is available (you must simply attempt to acquire it, and wait if it's not available). A friend proposes a simple scheme for locking the accounts: First lock the account to be credited; then lock the account to be debited. Can this scheme lead to deadlock? If you think it cannot, briefly explain why not. If you think it can, first give an example of a possible deadlock situation, and then design a scheme that avoids deadlocks, meets the stated design goals, *and uses only the locking mechanism just described.*