

Operating System Structures



Practice Exercises

- 2.1 What is the purpose of system calls?
Answer: System calls allow user-level processes to request services of the operating system.
- 2.2 What are the five major activities of an operating system with regard to process management?
Answer: The five major activities are:
- The creation and deletion of both user and system processes
 - The suspension and resumption of processes
 - The provision of mechanisms for process synchronization
 - The provision of mechanisms for process communication
 - The provision of mechanisms for deadlock handling
- 2.3 What are the three major activities of an operating system with regard to memory management?
Answer: The three major activities are:
- Keep track of which parts of memory are currently being used and by whom.
 - Decide which processes are to be loaded into memory when memory space becomes available.
 - Allocate and deallocate memory space as needed.
- 2.4 What are the three major activities of an operating system with regard to secondary-storage management?
Answer: The three major activities are:
- Free-space management.
 - Storage allocation.
 - Disk scheduling.

2.5 What is the purpose of the command interpreter? Why is it usually separate from the kernel?

Answer: It reads commands from the user or from a file of commands and executes them, usually by turning them into one or more system calls. It is usually not part of the kernel since the command interpreter is subject to changes.

2.6 What system calls have to be executed by a command interpreter or shell in order to start a new process?

Answer: In Unix systems, a *fork* system call followed by an *exec* system call need to be performed to start a new process. The *fork* call clones the currently executing process, while the *exec* call overlays a new process based on a different executable over the calling process.

2.7 What is the purpose of system programs?

Answer: System programs can be thought of as bundles of useful system calls. They provide basic functionality to users so that users do not need to write their own programs to solve common problems.

2.8 What is the main advantage of the layered approach to system design? What are the disadvantages of using the layered approach?

Answer: As in all cases of modular design, designing an operating system in a modular way has several advantages. The system is easier to debug and modify because changes affect only limited sections of the system rather than touching all sections of the operating system. Information is kept only where it is needed and is accessible only within a defined and restricted area, so any bugs affecting that data must be limited to a specific module or layer.

2.9 List five services provided by an operating system, and explain how each creates convenience for users. In which cases would it be impossible for user-level programs to provide these services? Explain your answer.

Answer: The five services are:

- a. **Program execution.** The operating system loads the contents (or sections) of a file into memory and begins its execution. A user-level program could not be trusted to properly allocate CPU time.
- b. **I/O operations.** Disks, tapes, serial lines, and other devices must be communicated with at a very low level. The user need only specify the device and the operation to perform on it, while the system converts that request into device- or controller-specific commands. User-level programs cannot be trusted to access only devices they should have access to and to access them only when they are otherwise unused.
- c. **File-system manipulation.** There are many details in file creation, deletion, allocation, and naming that users should not have to perform. Blocks of disk space are used by files and must be tracked. Deleting a file requires removing the name file information and freeing the allocated blocks. Protections must also be checked to assure proper file access. User programs could neither ensure ad-

herence to protection methods nor be trusted to allocate only free blocks and deallocate blocks on file deletion.

- d. **Communications.** Message passing between systems requires messages to be turned into packets of information, sent to the network controller, transmitted across a communications medium, and reassembled by the destination system. Packet ordering and data correction must take place. Again, user programs might not coordinate access to the network device, or they might receive packets destined for other processes.
 - e. **Error detection.** Error detection occurs at both the hardware and software levels. At the hardware level, all data transfers must be inspected to ensure that data have not been corrupted in transit. All data on media must be checked to be sure they have not changed since they were written to the media. At the software level, media must be checked for data consistency; for instance, whether the number of allocated and unallocated blocks of storage match the total number on the device. There, errors are frequently process-independent (for instance, the corruption of data on a disk), so there must be a global program (the operating system) that handles all types of errors. Also, by having errors processed by the operating system, processes need not contain code to catch and correct all the errors possible on a system.
- 2.10 Why do some systems store the operating system in firmware, while others store it on disk?
- Answer:** For certain devices, such as handheld PDAs and cellular telephones, a disk with a file system may not be available for the device. In this situation, the operating system must be stored in firmware.
- 2.11 How could a system be designed to allow a choice of operating systems from which to boot? What would the bootstrap program need to do?
- Answer:** Consider a system that would like to run both Windows XP and three different distributions of Linux (e.g., RedHat, Debian, and Mandrake). Each operating system will be stored on disk. During system boot-up, a special program (which we will call the **boot manager**) will determine which operating system to boot into. This means that rather than initially booting to an operating system, the boot manager will first run during system startup. It is this boot manager that is responsible for determining which system to boot into. Typically boot managers must be stored at certain locations of the hard disk to be recognized during system startup. Boot managers often provide the user with a selection of systems to boot into; boot managers are also typically designed to boot into a default operating system if no choice is selected by the user.

