Homework 01

CS307-Operating System (D), Chentao Wu, Spring 2020.

Name: 方泓杰(Hongjie Fang)   Student ID: 518030910150   Email: galaxies@sjtu.edu.cn

- (1.1) What are the three main purposes of an operating system?

  Solution. The three main purposes are as follows.

  (1) **To allocate the resource.** A computer system has many resources that may be required to solve a problem. The operating system acts as the manager of these resources.

  (2) **To control the devices and other programs.** An operating system is a control program which manages the execution of user programs to prevent error and improper use of the computer. It is especially concerned with the control of I/O devices.

  (3) **To provide an environment for other programs.** An operating system provides an environment within which other programs can do useful work.

  □

- (1.3) What is the main difficulty that a programmer must overcome in writing an operating system for a real-time environment?

  Solution. A real-time system has well-defined, fixed time constraints. So the main difficulty is the system should make **a correct response** to the users **in time**, or the real-time system may crash. Therefore, when writing a real-time operating system, the designer should make sure that the algorithms and technologies they use will not make the response time exceed the time limit. □

- (1.5) How does the distinction between kernel mode and user mode function as a rudimentary form of protection (security)?

  Solution. Some instructions can be executed, and some devices can be accessed, only in kernel mode. And only the system is in kernel mode, it can control the interrupts. The user mode provide an environment for user applications to be executed, but when the user application need a system service which is relatively dangerous, the system will switch to kernel mode to fulfill the request in a safer environment. □

- (1.6) Which of the following instruction should be privileged?

  a. Set value of timer.

  b. Read the clock.

  c. Clear memory.

  d. Issue a trap instruction.

  e. Turn off interrupts.

  f. Modify entries in device-status table.

  g. Switch from user to kernel mode.

  h. Access I/O device.

Solution. Instructions that relate to I/O control, timer management, interrupt management should be privileged. So the instructions **a,c,e,f and h** should be privileged, that is, be executed only in kernel mode. □

- (1.10) Give two reasons why caches are useful. What problems do they solve? What problems do they cause? If a cache can be made as large as the device for which it is caching (for instance, as large as a disk), why not make it that large and eliminate the device?

  Solution. The reasons are listed as follows.

  (i) Caches are useful when two or more components with different transferring speed need to exchange data.

  (ii) Caches are useful when we access some data frequently. Owing to their high speed, the data access can be efficient.

  The cache solve the **data exchange problem** of components with different transferring speed by acting as an **intermediary** and provides a buffer of intermediate speed between the components. If the faster device finds the data it needs in the cache, it can use them and do not have to wait for the slower device.

  As a result, the data in the caches should be updated when the original data is updated. And that cause the **data synchronization problem** in multiprocessor environment and distributed environment. When there are many sub-environments, the same data can be stored in several caches, or even in several computers. If different programs, or different computers use the data or update the data at the same time, it may cause the problem.

  If a cache can be made as large as the device for which it is caching, we can make it eliminate the device **only if it satisfies the following conditions**.

  (i) The expense of an cache is **affordable**.

  (ii) The cache and the device should have **equivalent state-saving capacity**. For instance, if we want to use an cache to eliminate a disk, then the data in the cache can not get lost when the power is off.

  □

- (1.11) Distinguish between the client-server and peer-to-peer models of distributed system.

  Solution. The client-server model distinguishes the roles of client and server. The client requests services from the server, and the server provides the services to the client.

  The peer-to-peer model doesn't have the strict roles of client and server. All the computers in the peer-to-peer model can be server, client, or both. A computer may request services from other computers, and may provide services to other computers. □