

Course

Monday, 27 February 2017
11:26 pm

Operating Systems Internals

Assignments

- 20% each
- Week 6 and 12

Homework

- Do homework after lecture to solidify understanding
- Exam questions similar to homework questions

Exam

- 2 hours
- Closed book
- Will examine concepts and pseudocode

Textbooks

- Operating System Concepts
- Operating Systems: Internals and Design Principles

Motivation:

Make effective use of available hardware.

1 Computer System Overview

Wednesday, 8 March 2017

12:55 pm

Operating system is a software program

- Manages resources
 - CPU
 - Memory
 - Input output
 - Files
- Control execution of programs
- Control and allocate memory
- Control input and output
- Manage files
- Facilitating networking
- Abstraction

Operating system goals

- Use resources in efficient manner
- Level of abstraction between hardware and programmer
 - Programmer focuses on problem logic rather than where data is stored, or how to retrieve data

The one program running at all times is a **KERNEL**

Basic elements of hardware

- Processor
- Memory
- IO
- System bus

Stored programming concept

- CPU can only work with data in main memory
- Recall LOAD, STORE... instructions
- The CPU can then FETCH and EXECUTE

Interrupts

- Interrupt normal sequencing of the processor
- Makes use of CPU idle time -- IO is slower than processor
- There are several classes

Table 1.1 Classes of Interrupts

Program	Generated by some condition that occurs as a result of an instruction execution, such as arithmetic overflow, division by zero, attempt to execute an illegal machine instruction, and reference outside a user's allowed memory space.
Timer	Generated by a timer within the processor. This allows the operating system to perform certain functions on a regular basis.
I/O	Generated by an I/O controller, to signal normal completion of an operation or to signal a variety of error conditions.
Hardware failure	Generated by a failure, such as power failure or memory parity error.

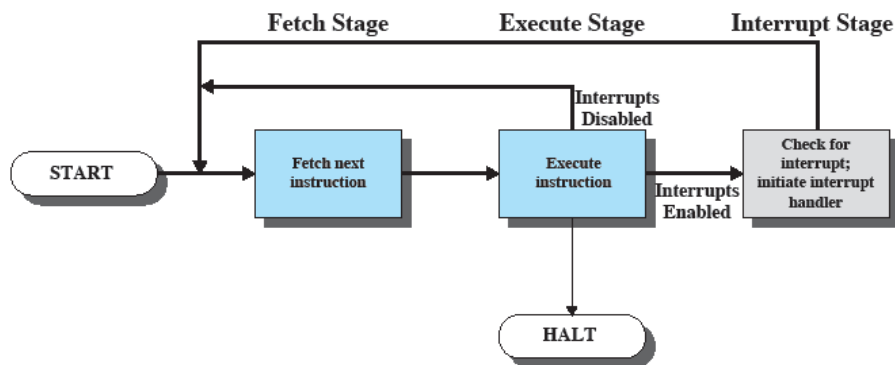


Figure 1.7 Instruction Cycle with Interrupts

Locality

- structure code so that you access data already stored in main memory
- Temporal and spatial

Main memory is volatile

Secondary memory is non-volatile

- used to store program and data files

Main memory

- Cache
- CPU checks cache first
- If data is there, use it
- If data is not there, copy from lower memory to cache
- Multicore processors can also have L2 and L3 caches

Homework 1

Wednesday, 8 March 2017

4:16 pm

List and briefly define the four main elements of a computer.

Processor	Performs operations
Memory	Storage of data and instructions
IO	Move information between computer and external devices
System bus	Allow communication between the above three

What is the purpose of interrupts?

- To stop current execution and handle an imminent task.
- Control is returned to before once interrupt is handled
- E.g. IO is slow and the CPU must wait during reads and writes. Interrupts are used during idle time to make efficient use of CPU.

Give one reason 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, a cache as large as a disk), why not make it that large and eliminate the device?

- Caches are useful when two (or more) components need to exchange data
- One does not need to wait for the other
- Cache must be updated if any data changes in memory (false sharing)
- They cannot be made as large as a disk because
 - fast storage is expensive, and
 - caches are volatile.

What characteristics distinguish the various elements of a memory hierarchy?

- Access (read) times
- Capacity
- Cost

What is the distinction between spatial locality and temporal locality?

- Spatial locality -- accessing memory locations which are close together
- Temporal locality -- accessing memory location which are used recently