

Daniel **Fitz**
(43961229)



University Of Queensland
COMP3301 – Operating Systems Architecture

ENGG2800 Lecture Notes



Table of Contents

Course Outline	2
Introduction	2
What is an operating system?	2
What Operating Systems Do	2
Operating System Definition	2
Common Functions of Interrupts	3
Interrupt Handling	3
Computer-System Architecture	3
Operating System Structure	3
Operating-System Operations	3
Transition from User to Kernel Mode	4
Process Management	4
Memory Management	4
Storage Management	4
IO Subsystem	4
Protection and Security	5
Open-Source Operating Systems	5

List of Tables

Placeholder for table of contents	0
-----------------------------------	---

List of Figures

Placeholder for table of contents	0
-----------------------------------	---

Course Outline

- Intro
 - What is an OS? Major roles/responsibilities. User/kernel interaction. Operating system structures
- Processes and Threads
 - Operations on processes, Process state, Program vs process, threads vs processes, IPC
- Scheduling
 - Concepts, criteria, algorithms, threads and scheduling
- Deadlock and synchronization
 - Prevention/avoidance/detection/recovery
- Memory management and virtual memory
 - The memory hierarchy, swapping, paging
 - Demand paging, copy-on-write, page replacement, mem-mapped files
- I/O subsystems
 - IO hardware, device models, drivers, interrupt handling, DMA
- Mass storage and filesystems
 - Disks, file systems, mounting, network file systems, disk IO scheduling
- Specialized OSes
 - Real time systems, multimedia systems, embedded systems
- Protection and Security

Introduction

What is an operating system?

- A program that act as an intermediary between a user of a computer and the computer hardware
- Operating system goals:
 - Execute user programs and make solving user problems easier
 - Make the computer system convenient to use
 - Use the computer hardware in an efficient manner

What Operating Systems Do

- Depends on the point of view
- Users want convenience, ease of use
 - Don't care about resource utilization
- But shared computer such as mainframe or minicomputer must keep all users happy
- Users of dedicate systems such as workstations have dedicated resources but frequently use shared resources from servers
- Handheld computers are resource poor, optimized for usability and battery life
- Some computers have little or no user interface, such as embedded computers in devices and automobiles

Operating System Definition

- OS is a resource allocator
 - Manages all resources
 - Decides between conflicting requests for efficient and fair resource use
- OS is a control program
 - Controls execution of programs to prevent errors and improper use of the computer
- No universally accepted definition
- "Everything a vendor ships when you order an operating system" is good approximation (But varies wildly)
- "The one program running at all times on the computer" is the kernel. Everything else is either a system program (ships with the operating system) or an application program

Common Functions of Interrupts

- Interrupt transfers control of the interrupt service routine generally, through the interrupt vector, which contains the addresses of all the service routines
- Interrupt architecture must save the address of the interrupted instruction
- A trap or exception is a software-generated interrupt caused either by an error or a user request
- An operating system is interrupt driven

Interrupt Handling

- The operating system preserves the state of the CPU by storing registers and the program counter
- Determines which type of interrupt has occurred:
 - polling
 - vectored interrupt system
- Separate segments of code determine what action should be taken for each type of interrupt

Computer-System Architecture

- Most systems use a single general-purpose processor (PDAs through mainframes)
 - Most systems have special-purpose processors as well
- Multiprocessors systems growing in use and importance
 - Also known as parallel systems, tightly-coupled systems
 - Advantages include:
 - 1) Increased throughput
 - 2) Economy of scale
 - 3) Increased reliability – graceful degradation or fault tolerance
- Two types:
 - 1) Asymmetric Multiprocessing
 - 2) Symmetric Multiprocessing

Operating System Structure

- Multiprogramming needed for efficiency
 - Single user cannot keep CPU and I/O devices busy at all times
 - Multiprogramming organizes jobs (code and data) so CPU always has one to execute
 - A subset of total jobs in system is kept in memory
 - One job selected and run via job scheduling
 - When it has to wait (for I/O for example), OS switches to another job
- Timesharing (multitasking) is logical extension in which CPU switches jobs so frequently that users can interact with each job while it is running, creating interactive computing
 - Response time should be <1 second
 - Each user has at least one program executing in memory -> process
 - If several jobs ready to run at the same time -> CPU scheduling
 - If processes don't fit in memory, swapping moves them in and out to run
 - Virtual memory allows execution of processes not completely in memory

Operating-System Operations

- Interrupt driven by hardware
- Software error or requests creates exception or trap
 - Division by zero, request for operating system service
- Other process problems include infinite loop, processes modifying each other or the operating system
- Dual-mode operation allows OS to protect itself and other system components
 - User mode and kernel mode

-
- Mode bit provided by hardware
 - Provides ability to distinguish when system is running user code or kernel code
 - Some instructions designated as privileged, only executable in kernel mode
 - System call changes mode to kernel, return from call resets it to user
 - Increasingly CPUs support multi-mode operations
 - i.e. virtual machine manager (VMM) mode for guest VMs

Transition from User to Kernel Mode

- Timer to prevent infinite loop / process hogging resources
 - Set interrupt after specific period
 - Operating system decrements counter
 - When counter zero generate an interrupt
 - Set up before scheduling process to regain control or terminate program that exceeds allotted time

Process Management

- A process is a program in execution. It is a unit of work within the system. Program is a passive entity, process is an active entity
- Process needs resources to accomplish its task
 - CPU, memory, IO, files
 - Initialization data
- Process termination requires reclaim of any reusable resources
- Single-threaded process has one program counter specifying location of next instruction to execute
 - Process executes instructions sequentially, one at a time, until completion
- Multi-threaded process has one program counter per thread
- Typically system has many processes, some user, some operating system running concurrently on one or more CPUs
 - Concurrency by multiplexing the CPUs among the processes / threads

Memory Management

- All data in memory before and after processing
- All instructions in memory in order to execute
- Memory management determines what is in memory when
 - Optimizing CPU utilization and computer response to users
- Memory management activities
 - Keeping track of which parts of memory are currently being used and by whom
 - Deciding which processes (or parts thereof) and data to move into and out of memory
 - Allocating and deallocating memory space as needed

Storage Management

- OS provides uniform, logical view of information storage
 - Abstracts physical properties to logical storage unit – file
 - Each medium is controlled by device (i.e. disk drive, tape drive)
 - Varying properties include access speed, capacity, data-transfer rate, access method (sequential or random)
- File-System management
 - Files usually organized into directories
 - Access control on most systems to determine who can access what
 - OS activities include
 - Creating and deleting files and directories
 - Primitives to manipulate files and dirs
 - Mapping files onto secondary storage
 - Backup files onto stable (non-volatile) storage media

IO Subsystem

-
- One purpose of OS is to hide peculiarities of hardware devices from the user
 - IO subsystem responsible for
 - Memory management of IO including buffering (storing data temporarily while it is being transferred), caching (storing parts of data in faster storage for performance), spooling (the overlapping of output of one job with input of other jobs)
 - General device-driver interface
 - Drivers for specific hardware devices

Protection and Security

- Protection – any mechanism for controlling access of processes or users to resources defined by the OS
- Security – defense of the system against internal and external attacks
 - Huge range, including denial-of-service, worms, viruses, identity theft, theft of service

Open-Source Operating Systems

- Operating system made available in source-code format rather than just binary closed-source
- Counter to the copy protection and Digital Rights Management (DRM) movement
- Started by Free Software Foundation (FSF), which has "copyleft" GNU Public License (GPL)
- Examples include GNU/Linux and BSD UNIX (including core of Mac OS X), and many more