Introduction

Based on original slides by
Silberschatz, Galvin and Gagne

1.1

What is an Operating System?

- A program that acts as an intermediary between a user of a computer and the computer hardware.
- Operating system goals:
  - Provide an environment for executing user programs and making solving user problems easier.
  - Make the computer system convenient to use.
  - Use the computer hardware in an efficient manner.

1.2

Abstract View of System Components

1.3

Computer System Components

- Hardware
  provides basic computing resources (CPU, memory, I/O devices).
- Operating system
  controls and coordinates the use of the hardware among the various application programs for the various users.
- Applications programs –
  define the ways in which the system resources are used to solve the computing problems of the users (compilers, database systems, video games, business programs).
- Users
  (people, machines, other computers).

1.4

Operating System Definitions

- Resource allocator
  - manages and allocates resources.
- Control program
  - controls the execution of user programs and operations of I/O devices
- Kernel
  - the one program running at all times (all else being application programs).

1.5

Goals of an Operating System

Depend on the particular system

- Make the computer easy to use (e.g. PC)
- Optimize the computational resources (e.g. mainframe)
- Optimize shared resources (e.g., distributed systems)
- Make the computer easy to use and optimize energetic resources (e.g., handheld computers)
- …

Convenience vs. efficiency

1.6

Mainframe Systems

- Batch systems
  - Job processing without any interaction with the user
  - Mono-programmed
  - Multi-programmed
- Transaction-oriented systems
  - Process a large amount of small tasks
  - e.g. reservations of an airline company
- Time-sharing systems
  - Many remote users executes their jobs on a computer
  - Queries to a large DB system

Mainframe systems often supports all above service

1.7

Mainframe Systems

- Reduce setup time by batching similar jobs
- Automatic job sequencing – automatically transfers control from one job to another. First rudimentary operating system.
- Resident monitor
  - initial control in monitor
  - control transfers to job
  - when job completes control transfers pack to monitor

1.8

Memory Layout for a Simple Batch System

1.9

Multiprogrammed Batch Systems

Several jobs are kept in main memory at the same time, and the CPU is multiplexed among them.

1.10

OS Features Needed for Multiprogramming

- Memory management – the system must allocate the memory to several jobs.
- CPU scheduling – the system must choose among several jobs ready to run.
- Allocation of devices to concurrent processes

1.11

Time-Sharing Systems–Interactive Computing

- Limits of Batch systems:
  - no interaction between program and user
- Time-sharing systems extend the concept of multitasking
- Allow interaction between user and program or system
  - Users executes commands and wait for a response
  - Response time is usually short
- CPU is multiplexed among several jobs
  - Each user thinks to have the CPU dedicated to it
- Time-sharing system features
  - Memory management
  - CPU scheduling (typically Round Robin)

1.12

Desktop Systems

- Personal computer
  - computer system dedicated to a single user.
- User convenience and responsiveness
- Evolution of operating systems for desktop systems
  - From mono-user monoprogrammed systems (DOS) …
  - To multi-user, multiprogrammed, with protection features
- Most popular operating systems
  - Windows, MacOS. UNIX, Linux

1.13

Server

- Provide Services to Many Users
  - Web server
  - Mail Server
  - File server
  - ….
- PC, Workstations, Mainframe
  - Abundance of resources
- OS used for servers
  - Windows Server
  - UNIX
  - Linux

1.14

Parallel Systems

- Multiprocessor systems with more than one CPU
- Tightly coupled system
  - processors share memory and a clock
  - communication usually takes place through shared memory.
- Advantages of parallel system:
  - Increased *throughput*
  - Economical
  - Increased reliability

1.15

Parallel Systems (Cont.)

- *Symmetric multiprocessing (SMP)*
  - Each processor runs and identical copy of the operating system.
  - Many processes can run at once without performance deterioration.
  - Most modern operating systems support SMP
- *Asymmetric multiprocessing*
  - Each processor is assigned a specific task;
  - master processor schedules and allocated work to slave processors.
  - More common in extremely large systems

1.16

Distributed Systems

- A set of computers connected by communication links
- Loosely coupled system
  - each processor has its own local memory and clock;
  - processors communicate with one another through communication lines
    - high-speed LANs, telephone lines, etc.
    - …
- Network Operating Systems
  - Include support for networking, file sharing, inter-process communication
  - Autonomy between nodes
- Distributed Operating Systems
  - The collection of nodes appear as a single system to the user

1.17

Clustered Systems

- Clustered systems usually share mass storage and communicate through a LAN
- Provides high *availability*.
- *Asymmetric clustering*
  - one server runs the application while other servers are in standby.
- *Symmetric clustering*
  - all N hosts are running the application.

1.18

Real-Time Systems

- Systems with fixed-time constraints
  - operation are subject to time constraints
- Often used for control device in dedicated applications
  - industrial control systems
  - controlling scientific experiments
  - medical imaging systems
  - …
- Real-Time systems may be
  - *hard* real-time
  - *soft* real-time.

1.19

Real-Time Systems (Cont.)

- Hard real-time:
  - Secondary storage limited or absent, data stored in short term memory, or read-only memory (ROM)
  - Conflicts with time-sharing systems, not supported by general-purpose operating systems.
- Soft real-time
  - Limited utility in industrial control of robotics
  - Useful in applications (multimedia, virtual reality) requiring advanced operating-system features.

1.20

Embedded Systems

- Systems embedded for controlling systems
  - The entire system is not a computing system
    - Wash machine
    - Cellular telephone
- Dedicated operating system
  - Usually resides in ROM
- May have real-time features

1.21

Handheld Systems

- Personal Digital Assistants (PDAs)
- Cellular telephones
- Issues:
  - Limited memory
  - Slow processors
  - Small display screens
  - Limited energetic resources
- OS tries to optimize energy consumption
  - By switching off devices when not necessary
    - Display
    - Wireless interface
    - CPU
  - By scaling down CPU clock frequency

1.22