Review of Operating System Principle

Li-Xudong

Software College, Nankai Univ.

leexudong@nankai.edu.cn

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- Input/Output
- Advanced Topics
- Operating System Design
- OS Case Studies

leexudong@nankai.edu.cn

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- Input/Output
- Advanced Topics
- Operating System Design
- OS Case Studies

leexudong@nankai.edu.cn

What is an operating system?

leexudong@nankai.edu.cn

OS Overview

- Computer hardware overview
- What is an operating system?
- The evolution of operating systems
- Operating system concepts
  - Process, address spaces, file, input/output
  - Protection, shell
  - **System calls**
- The operating system categories
  - Batch, Interactive, Real-Time
- Operating system structures
- Operating system Evaluation

leexudong@nankai.edu.cn

Key to Computer System Overview

- Computer hardware components
- John von Neumann Architecture
- CPU, Instruction
  - Kernel Mode: Privileged instruction
  - User Mode: Unprivileged instruction
- Memory Storage Architecture
- Cache Memory
- Bus
- Clock
- I/O Device
  - Disk, Tape, IDE, SATA, USB, SCSI
- Booting Operating System

leexudong@nankai.edu.cn

Key to Operating System Overview

Cont.,

- Viewpoints of OS
  - Extended Machine
  - Resource Manager
- OS development driving force
- Multi-programming, Time-sharing, Concurrent, Parallel, Multi-cores, Multi-processors
- Operating System Functions
- OS Characteristics
  - Concurrent, Shareable, Virtualization, Asynchronous
- System call, Trap, Library

OSP Contents

- OS Overview
- *Processes And Threads*
- Memory Management
- File Systems
- Input/Output
- Advanced Topics
- Operating System Design
- OS Case Studies

Processes And Threads

- Process Model
- Thread Model
- Process/Thread Scheduling
- Inter-Process Communication(IPC)
- Deadlock

Key to Process/Thread Model

- Process Running
  - Sequential Processes
  - Parallel Processes
  - Directed Acyclic Graph (DAG)
- Process Model
  - What is Process?
  - Process Characteristics
    - Dynamic, Concurrent,
    - Independence, Asynchronous
  - Process Creation, Termination, State
- Process Hierarchies
- Process Control Block (PCB), PID

Key to Process/Thread Model

Cont.,

- Thread Model
  - What is Thread?
  - Thread Creation, Termination, State
- Thread Hierarchies
- Thread Control Block, TID, Private Stack
- Implementing Threads
  - User-level Thread, Kernel-level Thread
  - Hybrid Thread
  - Comparison of User-level and Kernel-level Thread
- Scheduler Activations
- Pop-up Threads
- *Making Single-Threaded Code Multithreaded*
- Comparison of processes and threads

Key to Scheduling

- 3+1 Scheduling
  - High Level Scheduling (Job)
  - Intermediate-Level Scheduling (Memory)
  - Low Level Scheduling (Process/thread)
- Process/Thread Scheduling
  - Process Behavior: Compute-bound, I/O-bound
  - Scheduler
  - Scheduling algorithm
    - preemptive scheduling, non-preemptive scheduling
  - When to Schedule?
- Scheduling Algorithm Goals
  - Fairness, Policy enforcement, Balance
  - Throughput, Turnaround Time, CPU utilization
  - Response Time, Proportionality
  - Meeting deadlines, Predictability

Key to Scheduling

Cont.,

- Scheduling in Batch Systems
  - First Come First Served
  - Shortest Job First, Shortest Remaining Time Next
- Scheduling in Interactive Systems
  - Round-Robin Scheduling
  - Priority Scheduling, Multiple Queues
  - Shortest Process Next, Guaranteed Scheduling
  - Lottery Scheduling, Fair-Share Scheduling
- Scheduling in Real-Time Systems
  - Hard real time, soft real time
  - Periodic, aperiodic
- Policy versus Mechanism
- Thread Scheduling

Key to IPC

- Race Conditions
- Critical Regions
- Mutual Exclusion with Busy Waiting
  - Disabling Interrupts
  - ~~Lock Variables~~
  - Strict Alternation
  - Peterson's Solution
  - TSL Instruction
  - Swap Instruction
- Primitives(原语):
  - Sleep and Wakeup
- Priority Inversion Problem(优先权倒置)

Key to IPC

Cont.,

- Mutual Exclusion
  - Semaphore, Semaphores Set
  - Mutex
  - Event Counter
  - Condition Variable
  - Monitor
  - Message Passing
- Synchronization
  - Barriers
- Data Communication
- Pipe, Shared Memory, Message,Socket, Remote Procedure Call

Key to IPC

Cont.,

- Classical IPC Problems
  - The Producer and Consumer Problem
    - (The Bounded-Buffer Problem)
  - The Dining Philosophers Problem
  - The Readers and Writers Problem
    - Readers Have Priority
    - Writers Have Priority
  - The Sleeping Barber Problem

Key to Deadlock

- What is Deadlock?
- Resource
  - Preemptable and Nonpreemptable Resource
- Four ~~Necessary~~ Conditions for Deadlock
  - Mutual exclusion, Hold and wait,
  - No preemption, ***Circular wait***
- Deadlock Modeling: Resource allocation Graphs
- The Ostrich(鸵鸟) Algorithm
- Deadlock Detection
  - Detection with One Resource of Each Type
  - Detection with Multi-Resources of Each Type
- Deadlock Recovery
  - Preemption, Rollback, Killing Process

Key to Deadlock

Cont.,

- Deadlock Avoidance
  - Resource Trajectories
  - Safe State, Unsafe State
  - Banker Algorithm
- Deadlock Prevention
  - Attacking One of the Four Conditions
    - SPOOLing, Order resources numerically
- Other Issues
  - Two-Phase Locking
  - Non-resource Deadlock: semaphore
  - Starvation
    - Comparison of Deadlock and Starvation
  - Live Lock

OSP Contents

- OS Overview
- Processes And Threads
- *Memory Management*
- File Systems
- Input/Output
- Advanced Topics
- Operating System Design
- OS Case Studies

Memory Management

- Program Loading and Linking
- Memory hierarchy
- Memory Address Space
- Basic Memory Management
- Virtual Memory Management
- Paging System
  - Page Replacement Algorithms
  - Thrashing(抖动): Causes, How to Avoidance
  - Design Issues for Paging System
  - Implementation Issues for Paging System
- Segmentation System

Key to Basic Memory Management

- Partition-Based Memory Management
  - Fixed Partition, Dynamic Partition
  - Relocation, Protection
  - *Dynamic Partition Allocation Algorithm
  - **Free Partitions Management**: Linked List
- Fragmentation
- Memory Compaction
- Swapping(交换), Overlay(覆盖)
- Basic Paging Management
  - Page table, TLB, Multi-Level Page table
  - Inverted Page Table
- Basic Segmentation Management
  - Segmentation table

Key to Dynamic Partition Allocation Algorithm

- First Fit Allocation Algorithm
- Next Fit Allocation Algorithm
- Best Fit Allocation Algorithm
- Worst Fit Allocation Algorithm
- Buddy(伙伴式)Allocation Algorithm
- Quick Fit Allocation Algorithm

Key to Virtual Memory Management

- Principle of Locality
  - Time Locality
  - Space Locality
- What is virtual memory?
  - Required Paging
  - Required Segmentation
- MMU: Memory Management Unit
- Virtual Memory Characteristics
  - Discrete,Many times, Swap in out, Virtual
  - Required Paging System
  - Required Segmentation System

Key to Page Replacement Algorithms

- The Optimal Page Replacement Algorithm
- The Not Recently Used Page Replacement Algorithm
- The First-In, First-Out Page Replacement Algorithm
- The Second-Chance Page Replacement Algorithm
- The Clock Page Replacement Algorithm
- The **Least Recently Used**(LRU) Page Replacement Algorithm
- Simulating LRU in Software
- The **Working Set** Page Replacement Algorithm
- The WSClock Page Replacement Algorithm

- **Belady's Anomaly**
- Stack-like Page Replacement Algorithm

Key to Design Issues For Paging Systems

- Local versus Global Allocation Policies
- Load Control
- Page Size
- Separate Instruction and Data Spaces
- Shared Pages
- Shared Libraries
- Mapped Files
- Cleaning Policy
- Virtual Memory Interface

Key to Implementation Issues For Paging System

- Operating System Involvement with Paging
- Page Fault Handling
- Instruction Backup
- Locking Pages in Memory
- Backing Store
- Separation of Policy and Mechanism

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- *File Systems*
- Input/Output
- Advanced Topics
- Operating System Design
- OS Case Studies

File Systems

- Files
  - Logical Structure, Physical Structure(Index-node)
  - File Control Block(FCB), file descriptor, handle
- Directories
  - Tree-Structured Directory
  - Physical Structure
- File Share, File Protection, File Confidentiality
- File System Implementation
  - VFS, Log-Structured File System
- Storage Management
  - DISK, CD-ROM
- File System Backup
- File System Reliability
  - RAID: RAID0, RAID1, RAID5, RAID6, RAID10
- File System Performance
- Example File Systems

Key to Storage Management

- Block Size
- Disk Quotas
  - Soft file limit
  - Hard file limit
- Free Space Management
  File Using Free Block
  - Bitmap
  - Free Block linked
  - **Group Free Blocks linked**（成组链接法）

Key to Disk

- Disk Hardware
- RAID
  - RAID 0,1,5,1+0,...
- Disk Formatting
- **Disk Arm Scheduling Algorithms**
  - First-Come First-Served (FCFS)
  - Shortest Seek First (SSF)
  - Elevator(电梯) Algorithm (SCAN)
- Error Handling

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- *Input/Output*
- Advanced Topics
- Operating System Design
- OS Case Studies

Input/Output

- I/O Management Task
- Principles of I/O hardware
  - I/O Architecture, I/O Devices, Device Controllers
- Principles of I/O Software
  - Goals of the I/O Software
  - Ways of I/O Controlled
- I/O Software Layers
- ~~(Disks)~~
- Clocks
- User Interfaces
- Power Management

Key to Goals of I/O Software

- device independence
- uniform naming
- error handling
- synchronous vs. asynchronous
- buffer
- shareable vs. dedicated devices

Key to Ways of I/O Controlled

- Programmed I/O
- Interrupt-Driven I/O
- I/O Using DMA
- I/O Using Channel

- Comparison of Above Four Ways

Key to I/O Software Layers

- Interrupt Handlers
- Device Drivers
  - Block Device, Character Device
- Device-Independent I/O Software
  - Uniform Interfacing for Device Drivers
  - Buffering
  - Error Reporting
  - Allocating and Releasing Dedicated Devices
- User-Space I/O Software
  - SPOOLing

Key to Interrupt

- Interrupt Concept
  - Interrupt Controller
  - Interrupt Event
    - Outer Interrupt
    - Inter Interrupt (trap)
  - Interrupt Signal, PSW, Interrupt Vector
  - Precise Interrupt, Imprecise Interrupt
- Interrupt Processing
  - Disabled Interrupt
  - Interrupt Mask
  - Interrupt Priority
  - Interrupt Handler

Key to Clock

- Clock Hardware
  - Crystal Oscillator
    - One-shot mode
    - Square-wave mode
  - UTC: Universal Coordinated Time
- Clock Software
  - Clock functions
- Soft Timer

Key to User Interfaces

- Input Software
  - Keyboard Software
    - Canonical mode
    - Non-canonical mode
  - Mouse Software
- Output Software
  - Text Windows
  - GUI
    - The X Window System
    - Microsoft Windows
- Terminals
- THIN Clients

Key to Power Management

- Hardware Issues
  - Disposable, Rechargeable
- Operating System Issues
  - The Display
  - The Hard Disk
  - The CPU、GPU
  - The Memory
  - Wireless Communication
  - Thermal(热) Management
  - Battery Management
- OS Driver Interface
  - Advanced Configuration and Power Interface, ACPI
- Application Program Issues
  - Degrade Performance to Save Energy

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- Input/Output
- *Advanced Topics*
- Operating System Design
- OS Case Studies

Advanced Topics: Security

- The Security Environment
  - Threats, Intruders, Accidental Data Loss
- Basics of Cryptography
- Protection Mechanisms
- Authentication
- Insider Attacks
  - Logic Bombs, Trap Doors, Login Spoofing
- Exploiting Code Bugs
  - Buffer Overflow Attacks, Format String Attacks,...
- Malware(恶意软件)
  - Trojan Horses, Viruses, Worms, Spyware
  - Rootkits
- Defenses

Security: Key to Protection Domain

- Protection Domain
  - Object
  - Access Right
  - Domain
- Principle
  - Read to Know
- Access Matrix
  - Access Control List(ACL, 访问控制表)
  - Access Capabilities(访问权限表)

Advanced Topics: Virtualization and Cloud

- Requirements for Virtualization
- Sensitive Instructions
- Type 1 and Type 2 Hypervisors
- Key Technologies of Virtualization
- Virtual Appliances
- Cloud as a Service
- Cloud Computing
- ...

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- Input/Output
- Advanced Topics
- *Operating System Design*
- OS Case Studies

Operating System Design

- Goals of OS Design
- Why is it hard to design an OS?
- Operating System Standards
- Interface Design
- Implementation
- Performance
- Evaluation of the operating system
- Project Management
- Trends in Operating System Design

Key to Operating System Design

- Design & Runtime Architecture
  - Monolithic Systems
  - Layered Systems
  - Micro kernels
  - Client-Server Model
  - Object-Oriented
  - Exokernels
  - Monitor
  - Virtual Machines
  - ...

Key to Operating System Design

- Implementation
  - Mechanism versus Policy
  - Static Versus Dynamic Structure
- Performance
  - Space-Time Trade-offs
  - Caching
- Evaluation of the operating system
  - Performance
  - Reliability, Availability, Maintainability
    - MTBF 平均故障时间
    - MTTR 平均故障修复时间
  - Convenience
  - Portability

Key to Operating System Design

Cont.,

- Trends in OS Design
  - Multi-core
  - Virtualization
  - Large Address Space, Network
  - Parallel and Distributed Systems
  - Multimedia
  - Battery-Powered Computers
  - Embedded Systems
  - Sensor Nodes

Key to Operating System Design

Cont.,

- Project Management
  - The Mythical Man Month
  - No Silver Bullet 银弹

OSP Contents

- OS Overview
- Processes And Threads
- Memory Management
- File Systems
- Input/Output
- Advanced Topics
- Operating System Design
- *OS Case Studies*

OS Case Studies

- Linux
- Windows Research Kernel
- Open Solaris
- FreeBSD
- MacOS
- Android
- iOS
- ...