Introduction to Computing

Lectured by: Dr. Pham Tran Vu

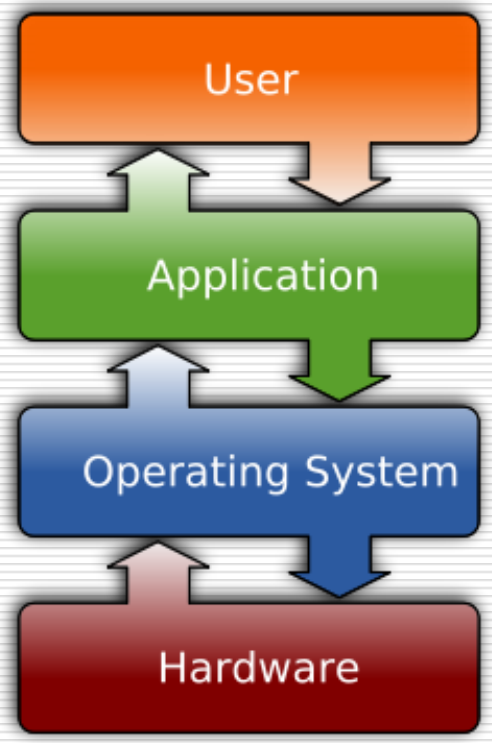
t.v.pham@cse.hcmut.edu.vn

Operating Systems

Definition

- “The software component of a computer system that is responsible for the management and coordination of activities and the sharing of the limited resources of the computer”

From Wikipedia



Computer Process and Program

- A process is a sequence of actions produced through the execution of program instructions
- When running, a process may create other processes

Resources Managed by OS

- CPU
- Main Memory
- Input/Output devices
- Backing storage
- Files

Other Functions

- Interpretation of command languages
- Error handling
- Protection of data files and programs from corruption and unauthorised use
- Accounting and logging of computer resources

Interrupt Handling

- Interrupts are necessary to notify the OS of system events, such as
  - Hardware failure
  - Program termination
  - Peripheral data transfer failure
  - Attempt to access to non-existent memory address
  - Completion of a time-slice in time-sharing
  - Program instruction error
  - External generated command from the operator

Error Handling and Trapping

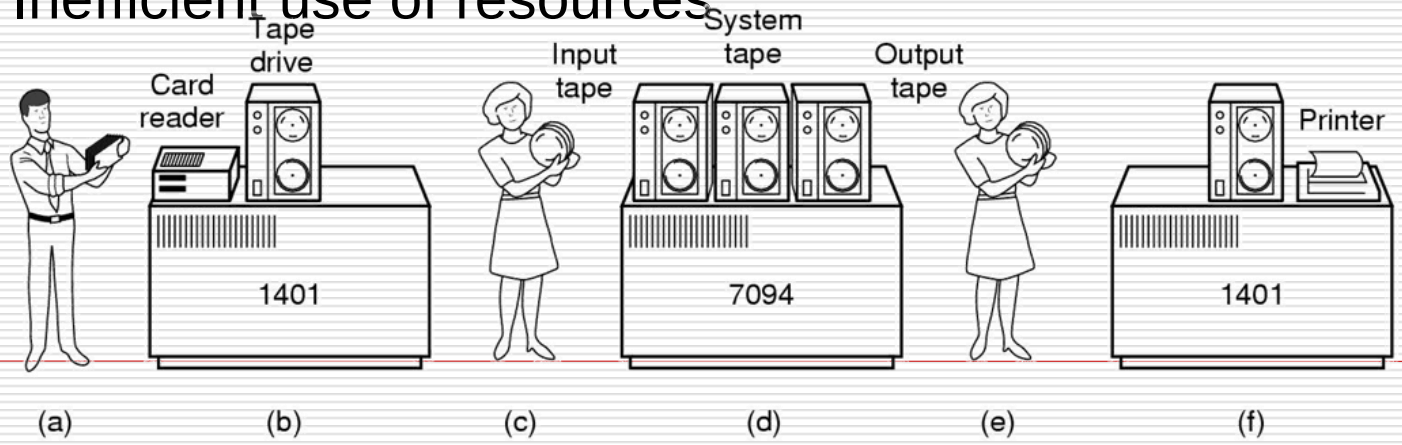
- To handle events that likely to crash the system in an order way
- Examples:
  - Illegal program instruction
  - Arithmetic errors (divide by zero)

File Management and Security

- Create and delete files
- Allocate space on storage media
- Identify and keep track of files on storage media
- Edit the contents of files
- Protecting files from hardware malfunction
- Protecting files from other applications and users
- Protecting files from unauthorised accesses

Categories of Operating System (1)

- Single stream
  - Can only handle one job at a time, e.g MS DOS
- Batch processing
  - Jobs are prioritised, scheduled and executed in sequence
  - Inefficient use of resources



Categories of Operating System (2)

- Batch Multi-Programming Systems
  - Running several job in main memory apparently simultaneously
  - Better use of resources in comparison to batch processing
  - Resources allocated dynamically
  - Jobs are prioritised based on I/O involved, types and speeds of I/O devices and processor time needed
  - Using high-level and low-level schedulers

Categories of Operating System (3)

- Time-sharing
  - Allocate processor time-slices to a number of programs
  - Time-slicing methods:
    - Round robin
      - Each process is given an equal slice of processor time
    - Priority
      - Allocate time slices to processes according to their priorities

Categories of Operating System (4)

- Multi-tasking
  - The system can accommodate several tasks in memory at one time
  - These tasks can be run concurrently by rapidly switching the processor's attention between them
- Multi-user
  - A number of different users can concurrently gain access to shared computing resources
  - The system needs to protect each user's files from access by other users

Categories of Operating System (5)

- Real-time
  - React to inputs at a sufficiently high speed to permit tight control of its environment

Memory Management

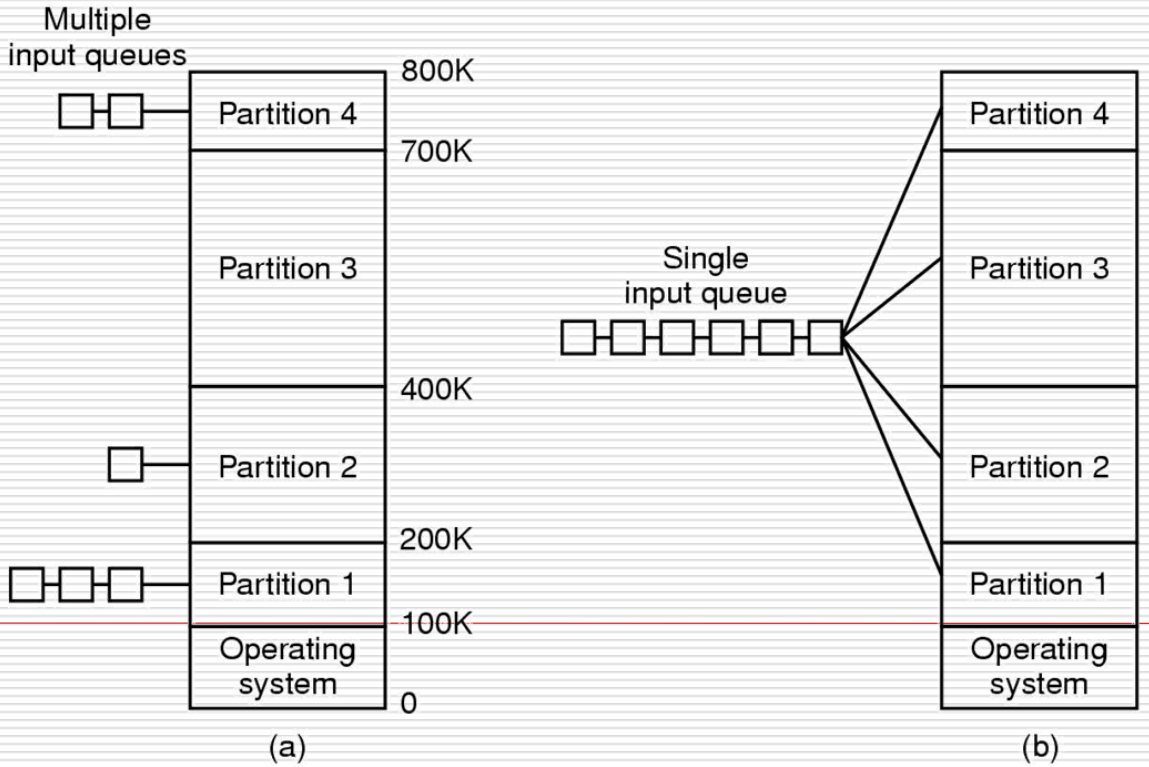
- In single stream system, only one application is resident at one time in memory -> the management in simple
- Multi-programming and multi-tasking systems require complex memory management complex
  - Memory requirements are dynamic
  - Require relocation of programs

Memory Partitioning and Relocating (1)

- The operating system needs to partition memory dynamically to meet the need of individual programs
- Programs and data in memory may be relocated to use available memory optimally

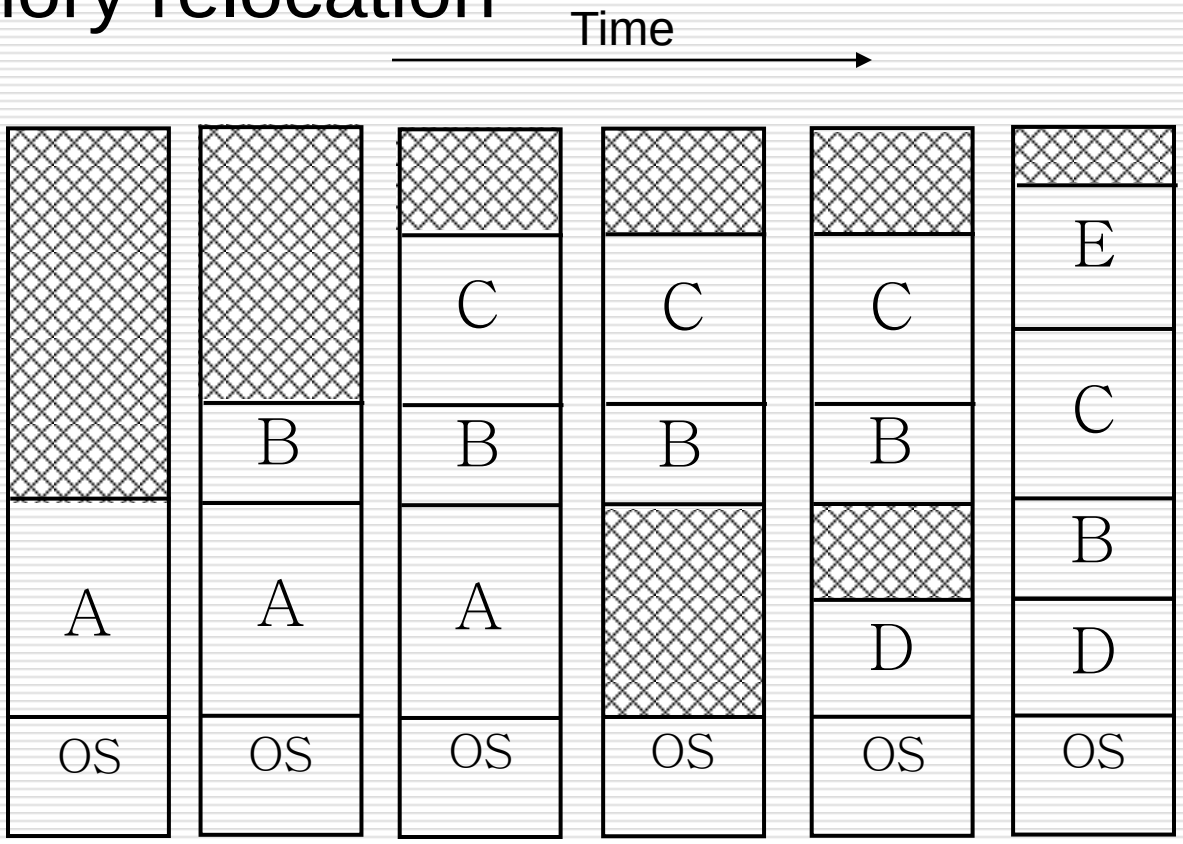
Memory Partitioning and Relocating (2)

- Two methods of allocating memory to partitions:
  - Using multiple queues (a)
  - Using single queue (b)



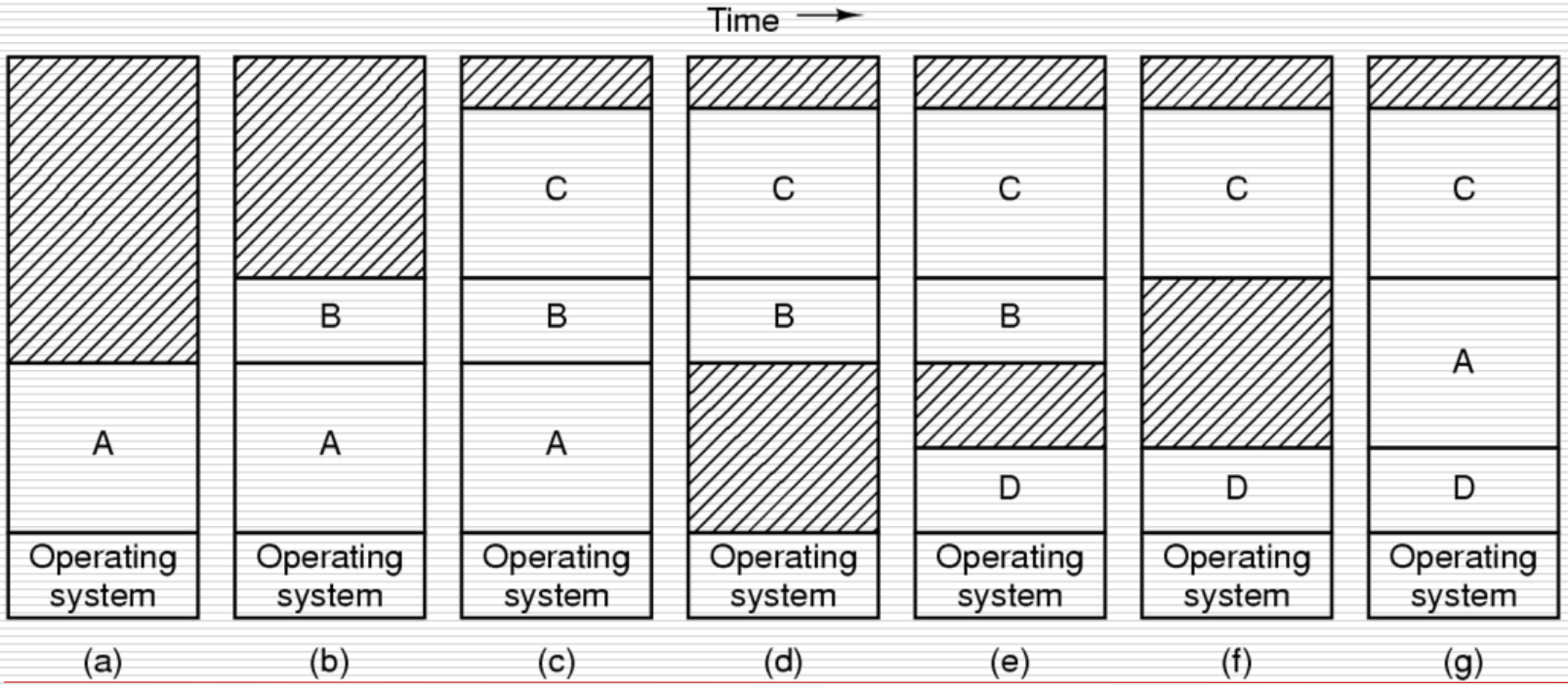
Memory Partitioning and Relocating (3)

- Memory relocation



Memory Swapping

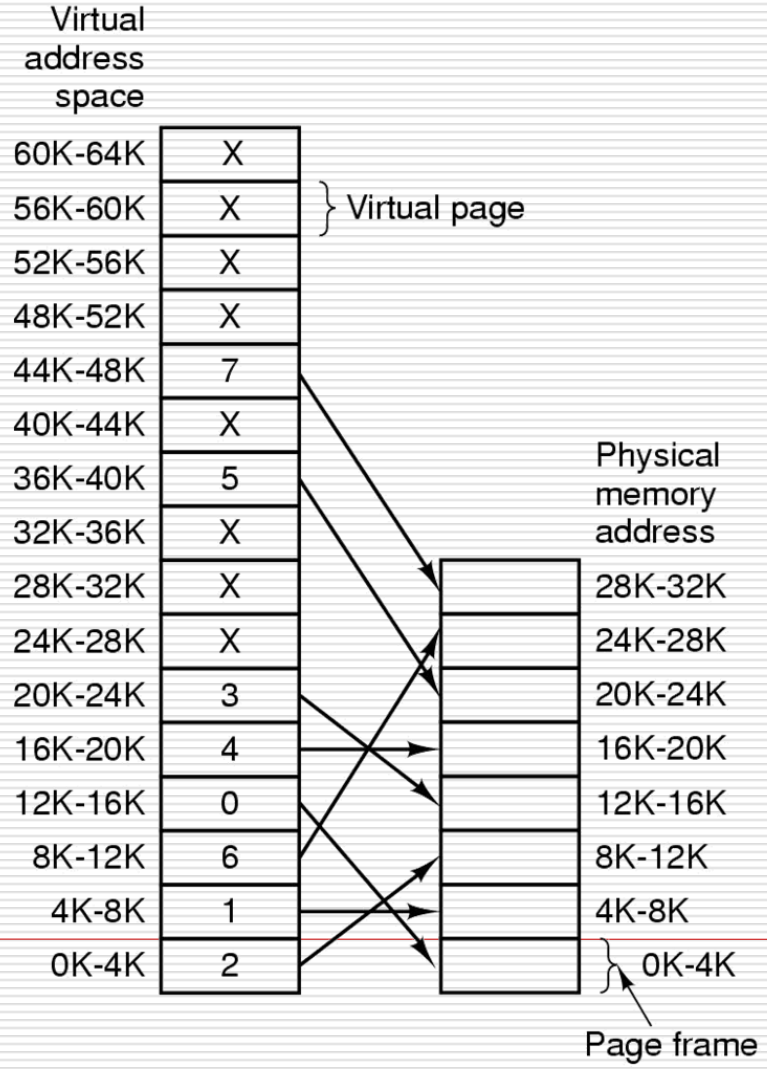
- In active program (A) is swapped into backing storage and swapped back in when needed



Virtual Memory (1)

- Main memory sometimes may not large enough to accommodate large programs
- Virtual memory:
  - use direct access backing storage as if it is main memory
  - Addressable memory space is beyond the physical capacity of the main memory
  - Programs are divided in to pages (unit of virtual storage)
  - Only parts of the program needed to be loaded to main memory during execution]
  - The virtual memory addresses are mapped to absolute address memory addresses during execution

Virtual Memory (2)



Typical Operating Systems

- MS-DOS
- Windows family: 3.1, 9x, NT, 2000, XP, Vista, Server 2003, etc
- OS/2
- Novell Netware
- Linux family
- Unix
- Mac OS family