



BITS Pilani
K K Birla Goa Campus

Operating Systems

Dr. Lucy J. Gudino
Dept. of CS and IS

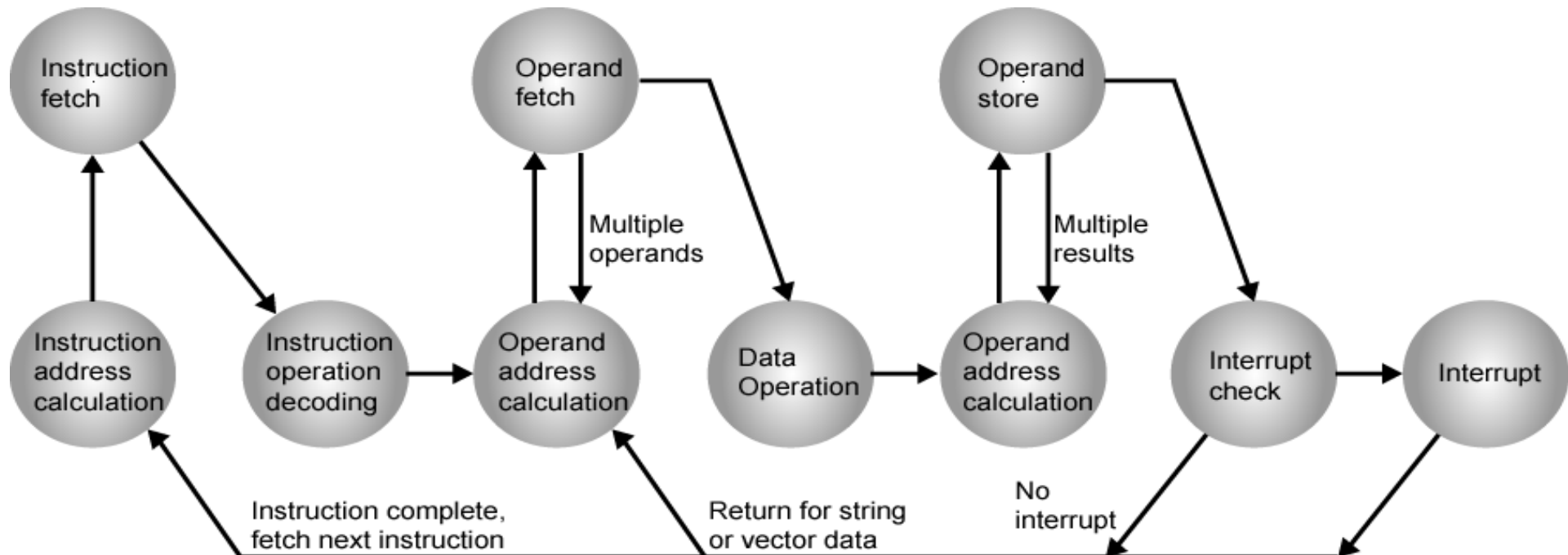
Memory Management

- Background
- Swapping
- Contiguous Allocation
- Paging

Introduction

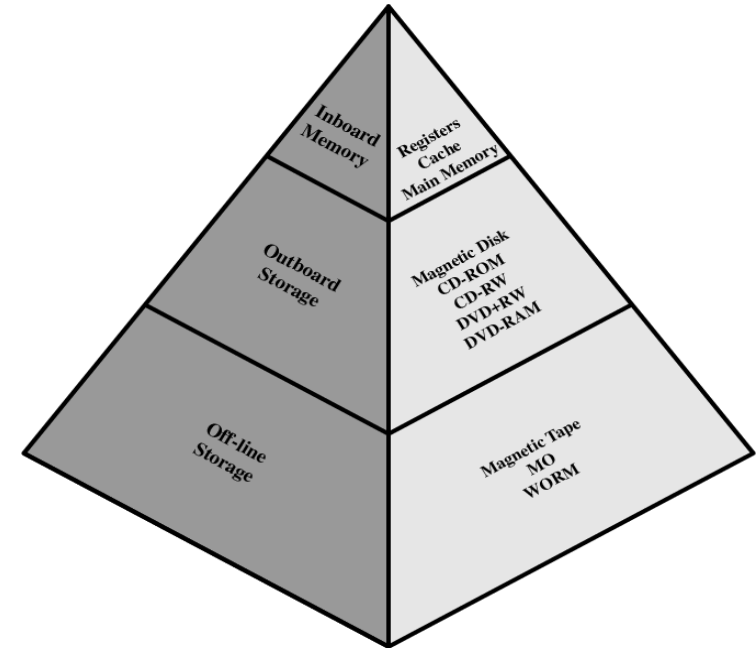


- Program must be brought into memory and placed within a process for it to be run.
- Fetch and execution of an instruction
- Instruction execution cycle
- stream of memory addresses



Background

- Speed, Access time and Capacity
 - Faster access time , **greater** cost per bit
 - Greater capacity, **smaller** cost per bit
 - Greater capacity, **slower** access time
- Main memory and registers are only storage CPU can access directly
- Register access in one CPU clock (or less)
- Main memory can take many cycles
- **Cache** sits between main memory and CPU registers
- Protection of memory required to ensure correct operation



Classification of information



- By role in program :
 - program instructions
 - constants
 - variables
- By protection status: Important for sharing data and/code
 - Read and write
 - Read only
- Addresses vs. Data
 - address change when program is relocated
- Uniprogramming or multiprogramming

Contd...



- Memory manager : **The part of the operating system that manages the memory hierarchy**
- Functions:
 - keep track of which parts of memory are in use and which parts are not in use,
 - allocation memory to processes and deallocate
 - swapping between main memory and disk

Memory Management Requirements



- Five requirements
 - Relocation
 - Protection
 - Sharing
 - Logical organization
 - Physical organization

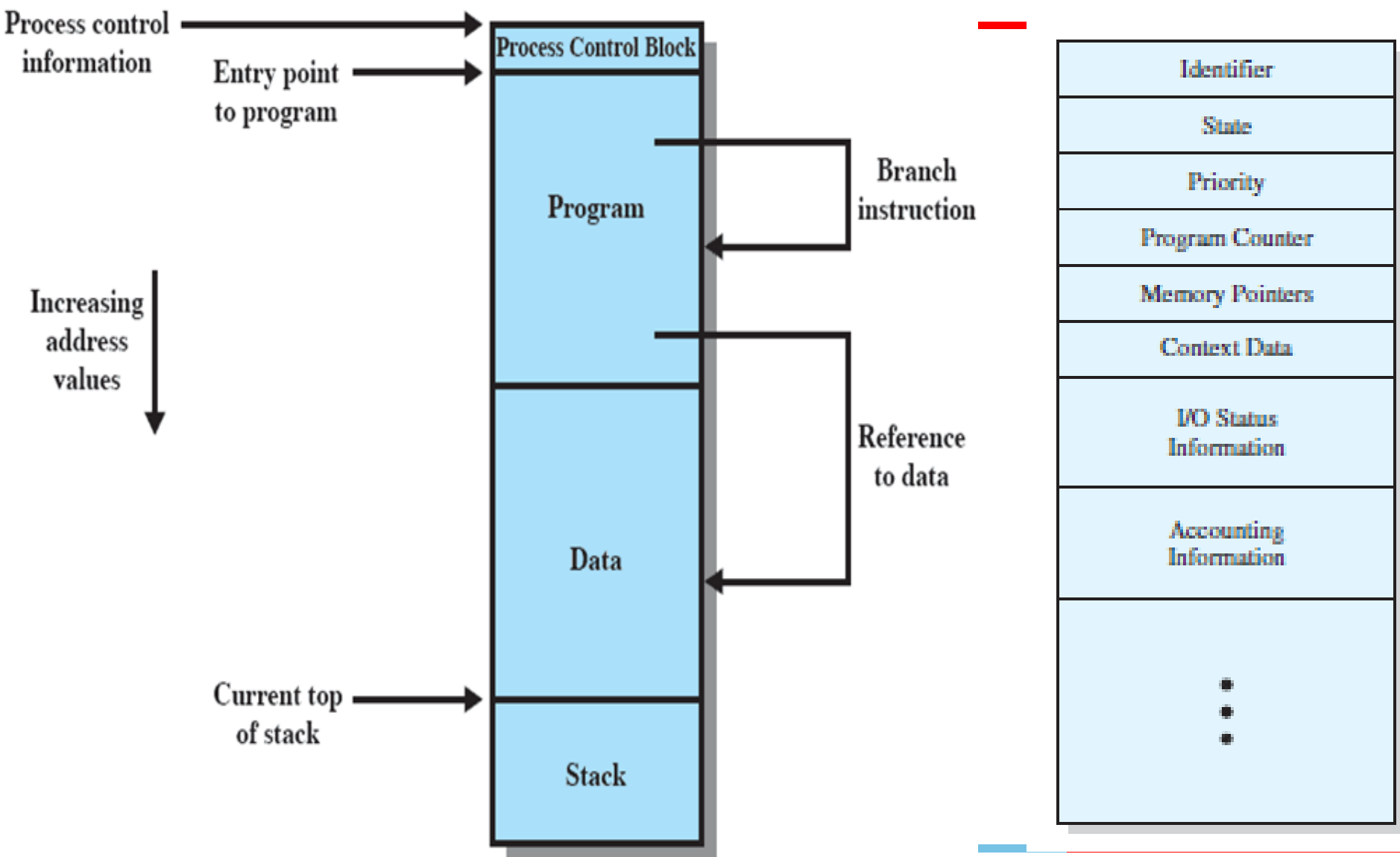
Relocation



- The programmer does not know where the program will be placed in memory when it is executed,
 - it may be swapped to disk and return to main memory at a different location (relocated)
- Memory references must be translated to the actual physical memory address

Term	Description
Frame	<i>Fixed</i> -length block of main memory.
Page	<i>Fixed</i> -length block of data in secondary memory (e.g. on disk).
Segment	<i>Variable-length</i> block of data that resides in secondary memory.

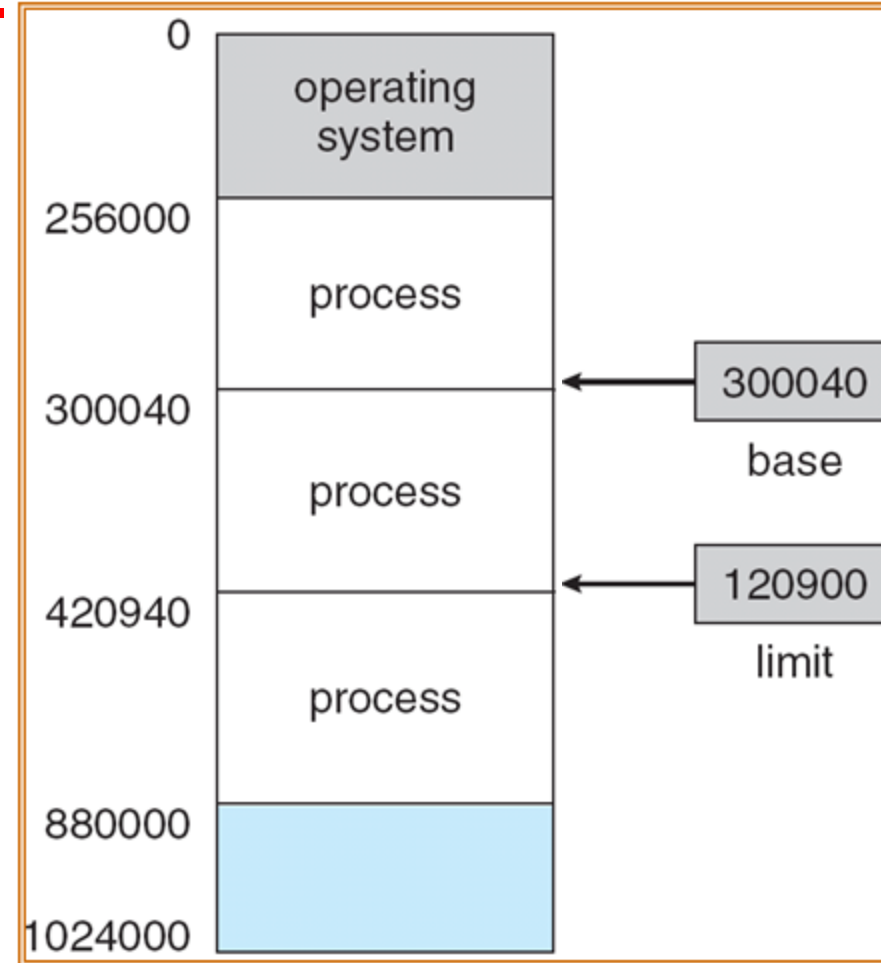
Contd...



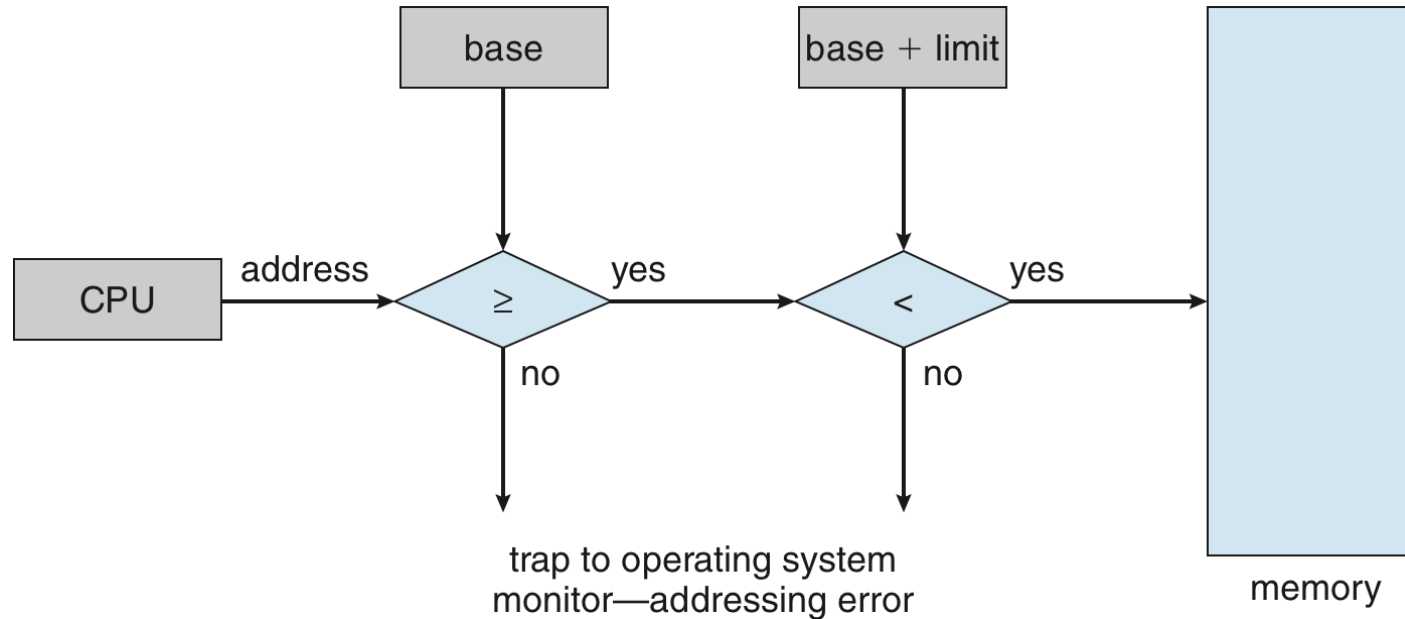
Protection



- Each process has a separate memory space
 - security ?
- provide security using 2 registers
 - base register
 - limit register
- The base register holds the smallest legal physical memory address; the limit register specifies the size of the range.
- Need a hardware for address comparison



Hardware Support for Relocation and Limit Registers



Points to be noted



- Illegal access results in trap to OS → fatal error
- Special privileged instructions are used to load base and limit registers
 - executed in kernel mode
 - OS
- The operating system has unrestricted access to both operating system and users' memory.

Sharing



- Allow several processes to access the same portion of memory
- Better to allow each process access to the same copy of the program rather than have their own separate copy

Logical Organization



- Memory is organized linearly (usually)
- Programs are written in modules
 - Modules can be written and compiled independently
- Different degrees of protection given to modules (read-only, execute-only)
- Share modules among processes
- Segmentation helps here

Physical Organization

- Two Level organization : Main memory and Secondary Memory
 - Main memory : Faster access at relatively high cost, volatile, smaller capacity
 - Secondary memory: Slower, cheaper, nonvolatile, large capacity
- Cannot leave the programmer with the responsibility to manage memory

Contd...



- Memory available for a program plus its data may be insufficient
 - Usage of Overlaying : replacement of a block of stored instructions or data with another
 - allows various modules to be assigned the same region of memory but is time consuming to program
- Programmer does not know how much space will be available