Chapter-5

MEMORY MANAGEMENT

IMPORTANT QUESTIONS

1. What is Compaction ?

- Compaction is a technique of combining all the free spaces together into a large block by pushing all process downwards as far as possible.

2.What is dynamic loading ?

- Better memory-space utilization can be done by dynamic loading. With dynamic loading, a routine is not loaded until it is called. All routines are kept on disk in a re-locatable load format. The main program is loaded into memory and is executed.

- The advantage of dynamic loading is that an unused routine is never loaded.

3.What is hit ratio ?

- The percentage value of the number of times a page is found in the TLB is called the hit ratio. A hit ratio of 85% implies that 85% of the time ,the required page is found in the TLB.

4.What is overlays?

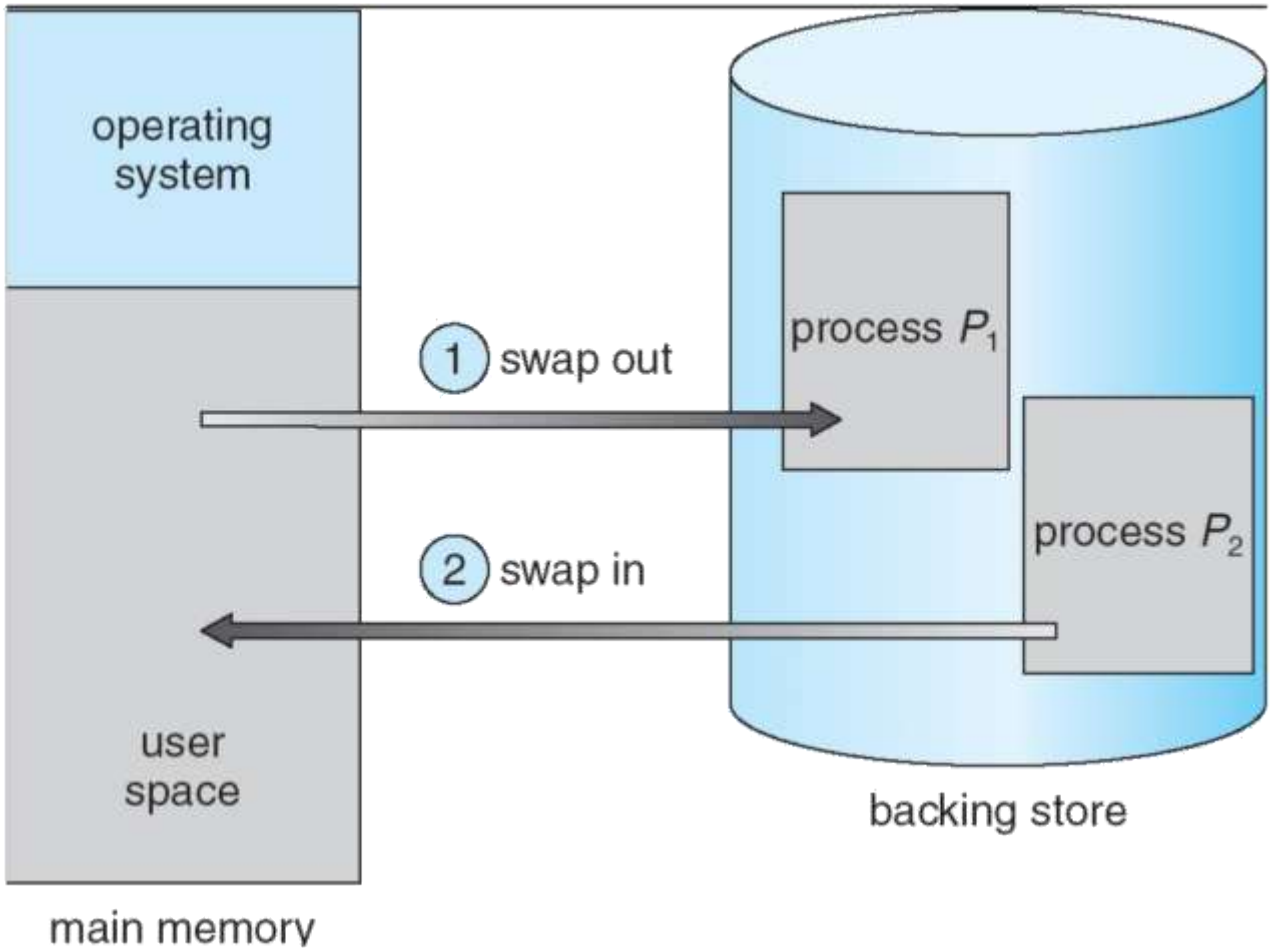
- Overlays are used to enable a process to be larger than the amount of memory allocated to it. The main objective of this scheme is to keep only those instructions and data in memory , which are required at that time. Other instructions are loaded into the memory whenever they are required.

5.What is the difference between **Logical and Physical Address Space ?**

- An address generated by the CPU is commonly referred to as a logical address, whereas an address seen by the main memory unit is commonly referred to as a physical address.

6.What do you mean by swapping?

- A variant of this swapping policy is used for priority-based scheduling algorithms. If a higher-priority process arrives and wants service, the memory manager can swap out the lower-priority process so that it can load and execute the higher priority process. When the higher priority process finishes, the lower-priority process can be swapped back in and continued. This variant of swapping is sometimes called rollout, roll in.


operating system
user space
main memory
1 swap out
2 swap in
process P_1
process P_2
backing store

7. Explain the first-fit , best-fit and worst-fit algorithms

- Three most common strategies of selecting free hole from set of available holes
- **First-fit:** Allocate the first hole that is big enough. Searching can start either at the beginning of the set of holes or where the previous first-fit search ended. We can stop searching as soon as we find a free hole that is large enough.

- **Best-fit:** Allocate the smallest hole that is big enough. We must search the entire list, unless the list is kept ordered by size. This strategy-produces the smallest leftover hole.

- **Worst-fit:** Allocate the largest hole. Again, we must search the entire list unless it is sorted by size. This strategy produces the largest leftover hole which may be more useful than the smaller leftover hole from a best-t approach.

8. What is Fragmentation?

- Fragmentation means wastage of memory.
- External fragmentation - It exists when enough to the memory space exists to satisfy a request, but it is not contiguous; storage is fragmented into a large number of small holes.
- Internal fragmentation - memory that is internal to partition, but is not being used

9. Explain the concept of Paging

Paging is a memory management scheme which permits the physical address of a process to non-contiguous. This paging scheme gives solution to external fragmentation

Paging

- Logical address space of a process can be noncontiguous
- Divide logical memory into blocks of same size called **pages.**
- Divide physical memory into fixed-sized blocks called **frames**
- When a process is executed, its pages are loaded into frames in main memory from disk

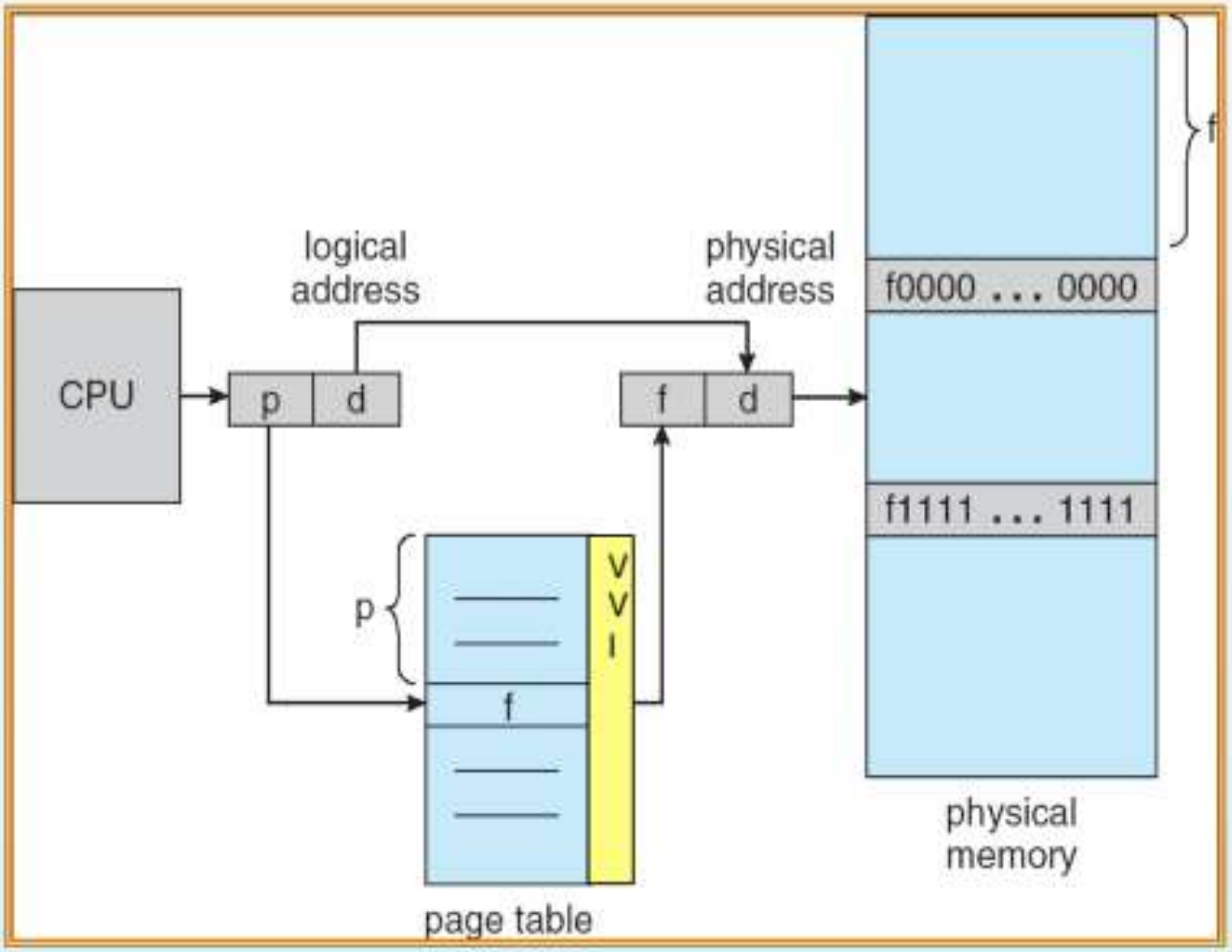
Address Translation Scheme

Address generated by CPU is divided into:

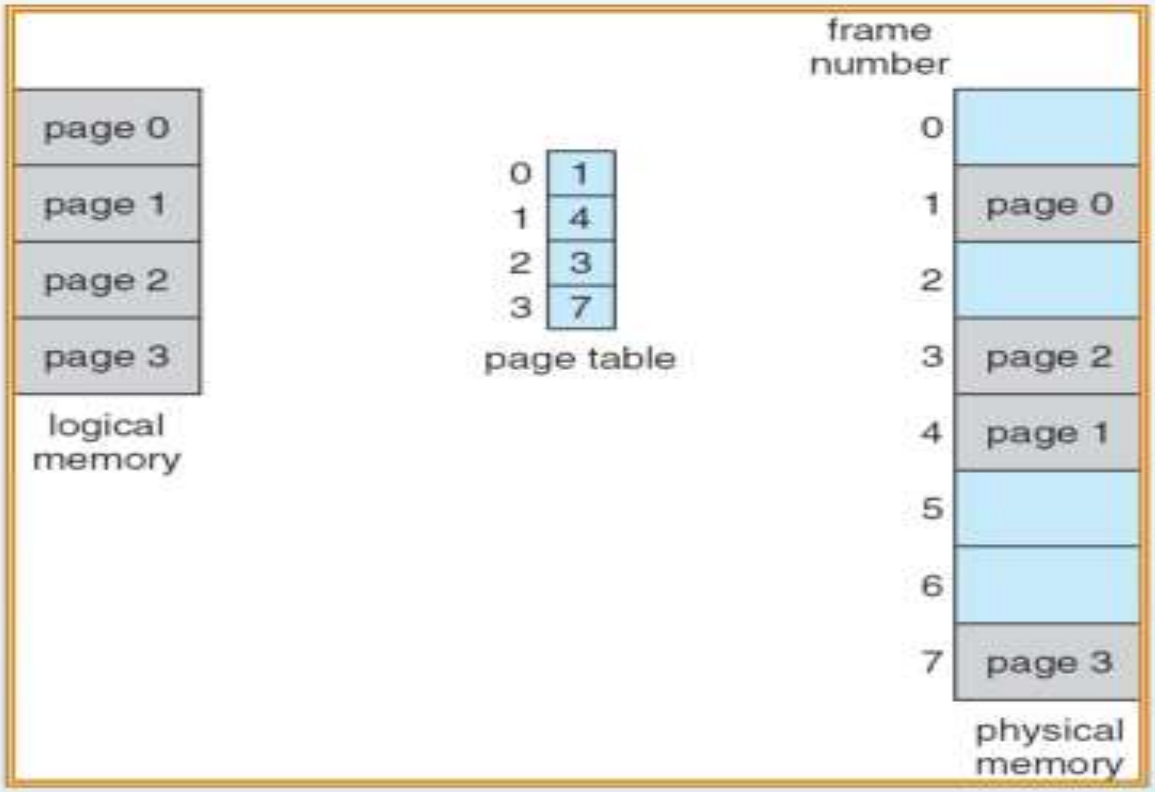
Page number (p) – used as an index into a page table which contains base address of each page in physical memory

Page offset (d) – combined with base address to define the physical memory address that is sent to the memory unit"

Address Translation Architecture



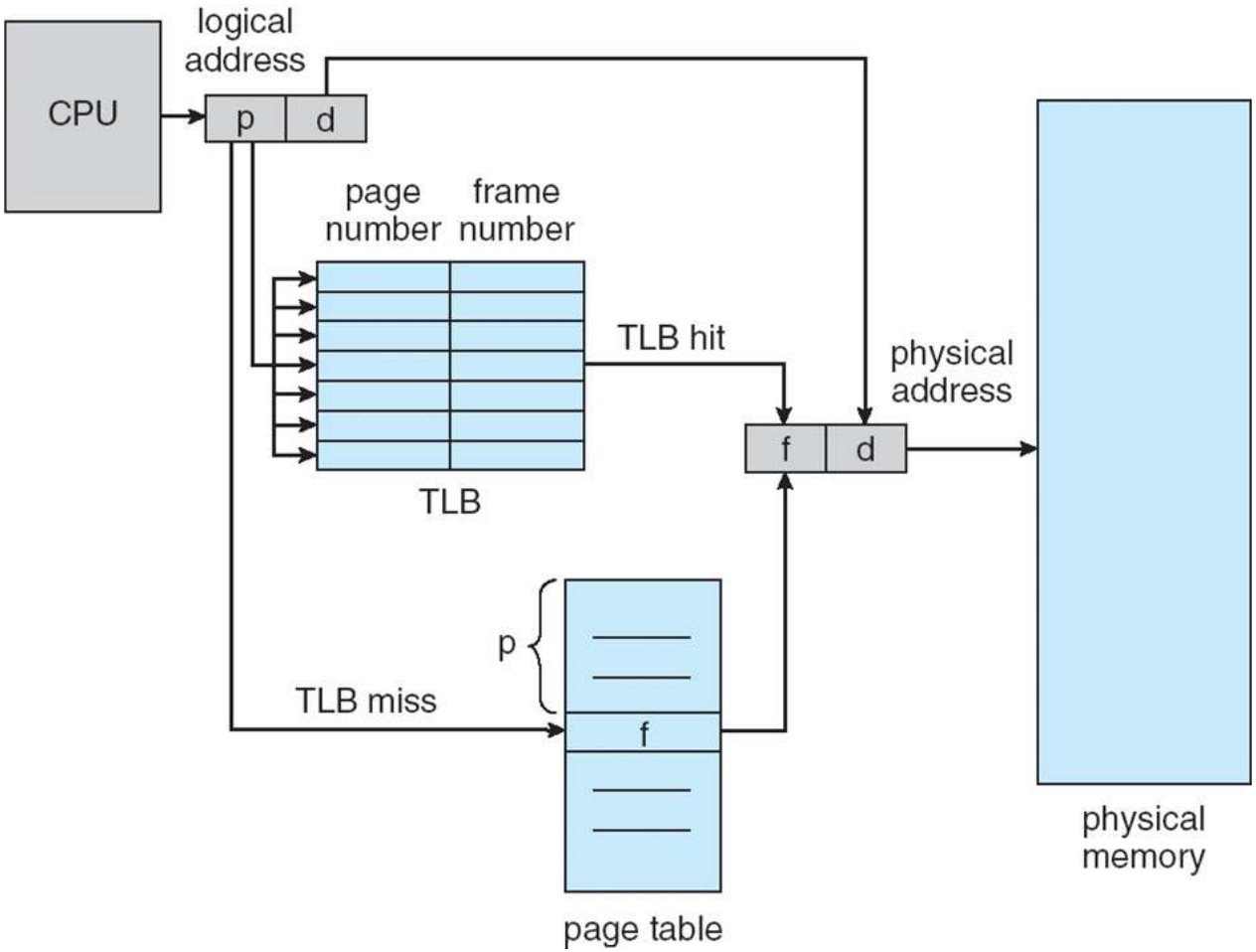
Paging Example



Implementation of Page Table

- Page table is kept in main memory
- *Page-table base register (*PTBR) points to the page table
- *Page-table length register* (PRLR) indicates size of the page table
- In this scheme every data/instruction access requires two memory accesses. One for the page table and one for the data/instruction.
- The two memory access problem can be solved by the use of a special fast-lookup hardware cache called **associative memory** or **translation look-aside buffers (TLBs)**

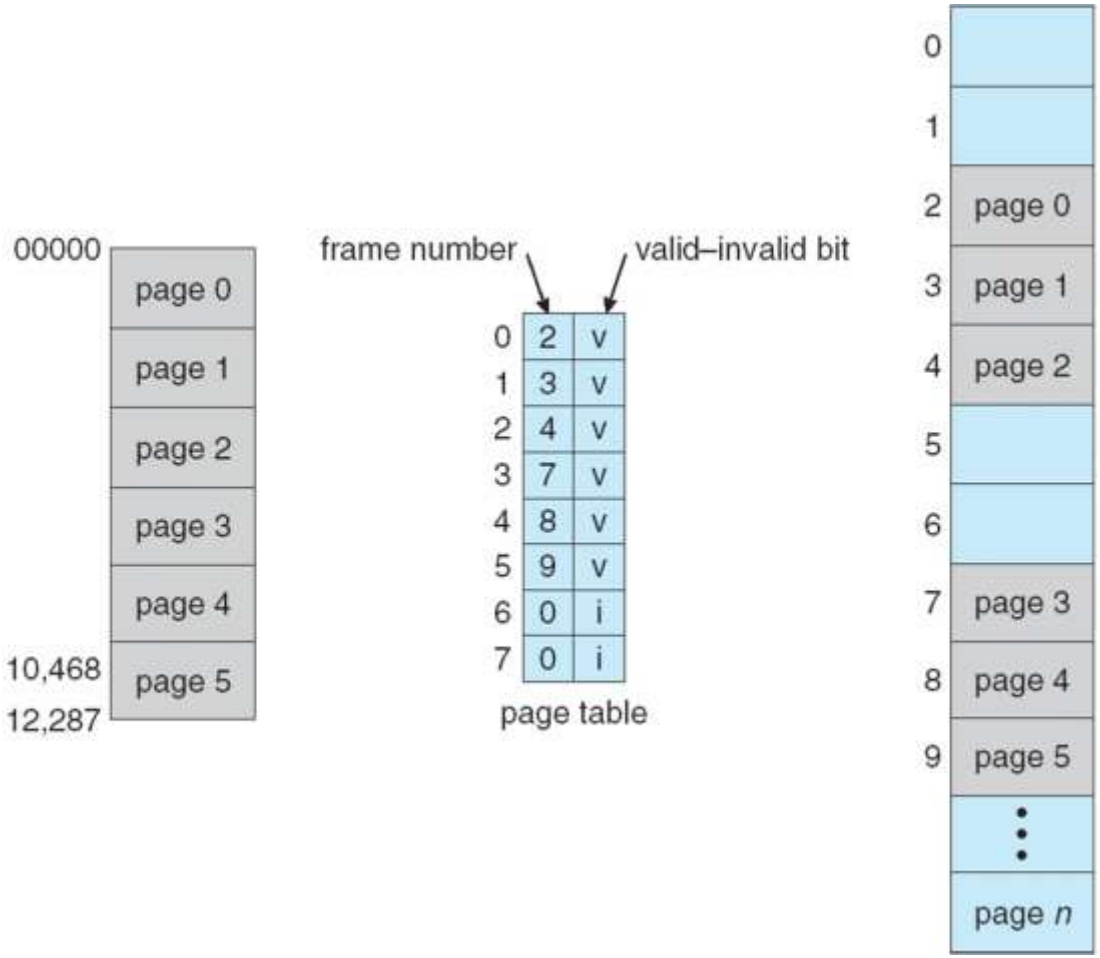
Paging Hardware With TLB



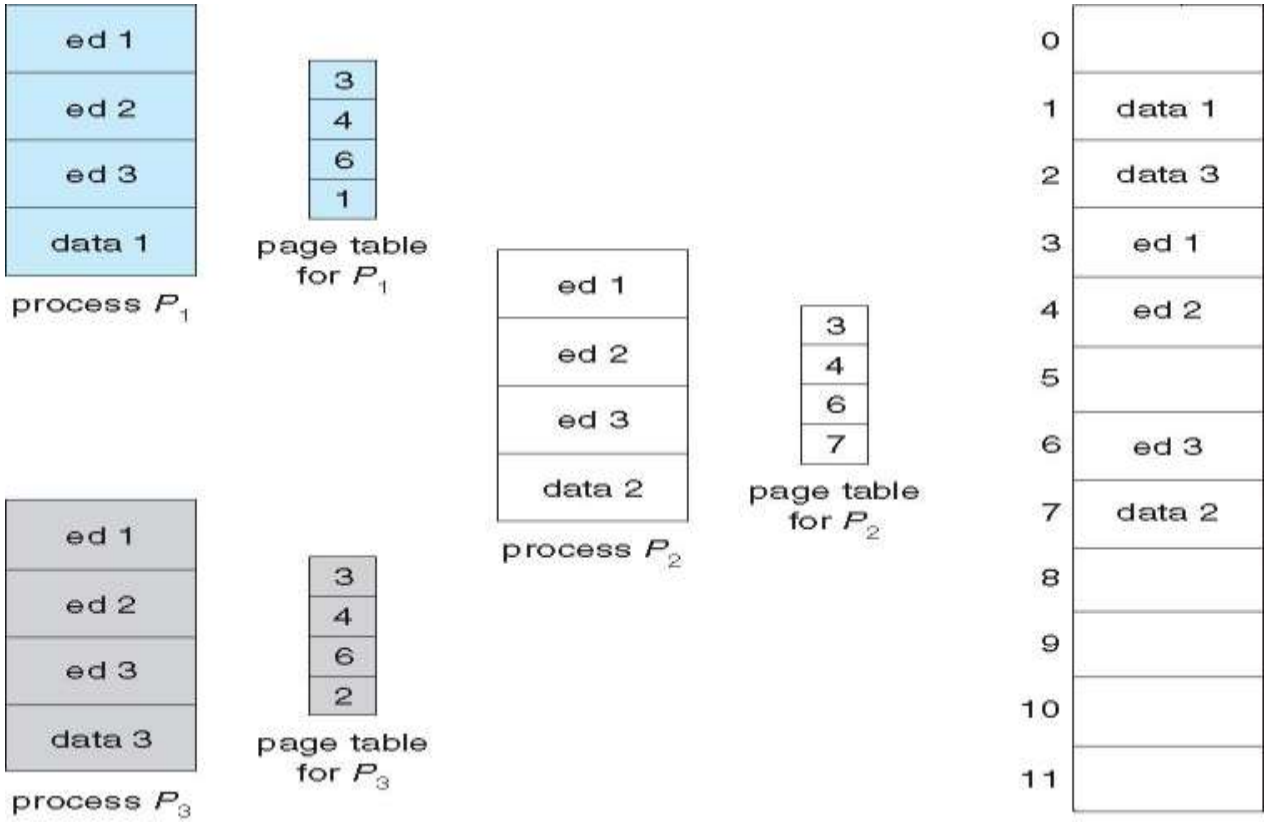
Memory protection

- Memory protection implemented by associating protection bit with each frame to indicate if read-only or read-write access is allowed
  - Can also add more bits to indicate page execute-only, and so on
- **Valid-invalid** bit attached to each entry in the page table:
  - "valid" indicates that the associated page is in the process' logical address space, and is thus a legal page
  - "invalid" indicates that the page is not in the process' logical address space

Valid (v) or Invalid (i) Bit In A Page Table



Shared Pages



Advantages of paging

- Paging supports time sharing system.
- It avoids external fragmentation.
- Compaction overheads are eliminated.
- Sharing of common code are possible.

Disadvantages of paging

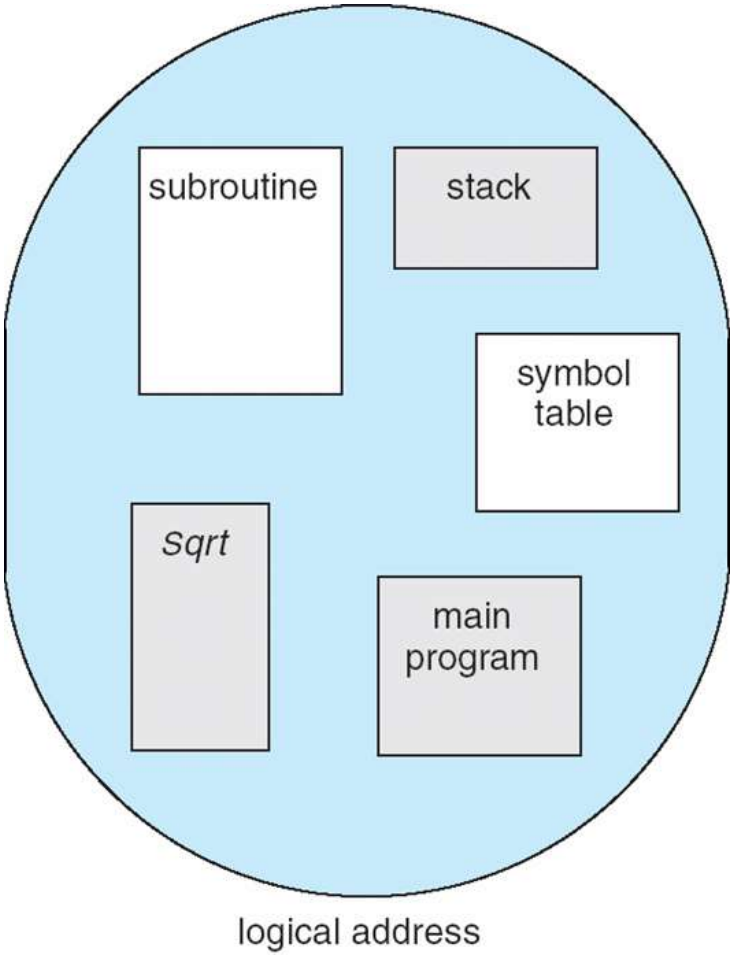
- When number of pages are large, it is difficult to maintain page tables.
- Hardware cost is high.

10. Explain the concept of Segmentation

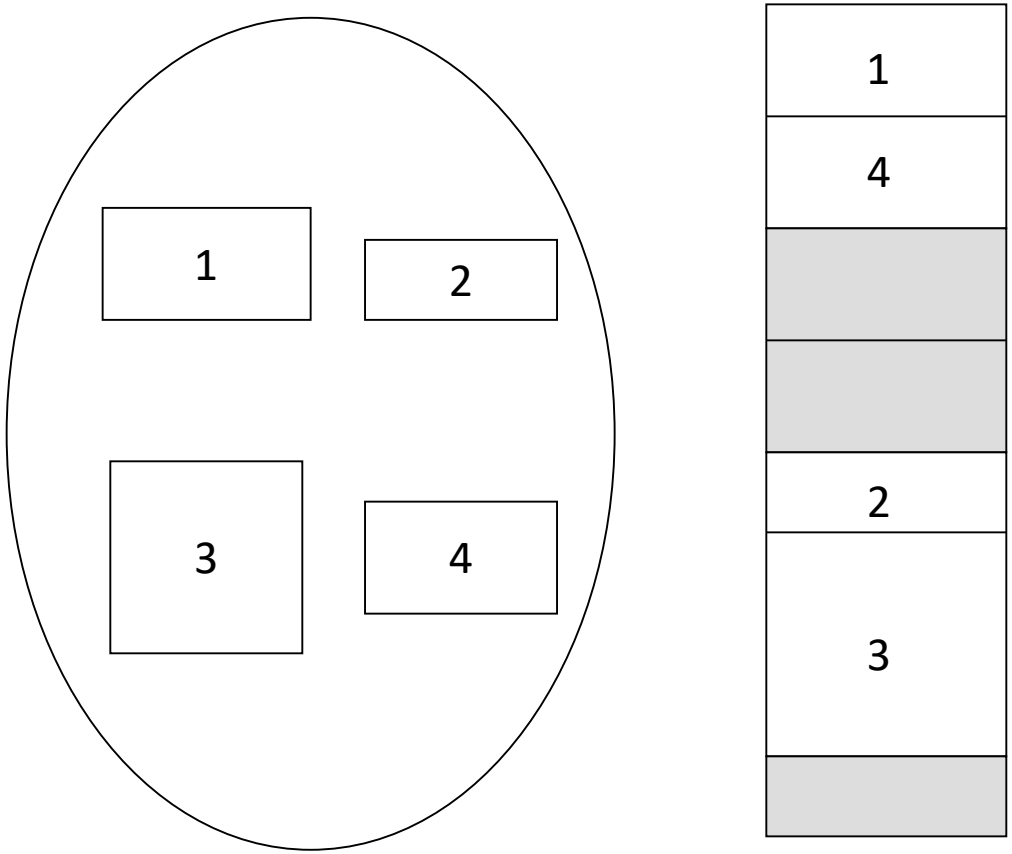
- Memory-management scheme that supports user view of memory
- A program is a collection of segments
  - A segment is a logical unit such as:

    main program
    procedure
    function
    method
    object
    local variables, global variables
    common block
    stack
    symbol table
    arrays

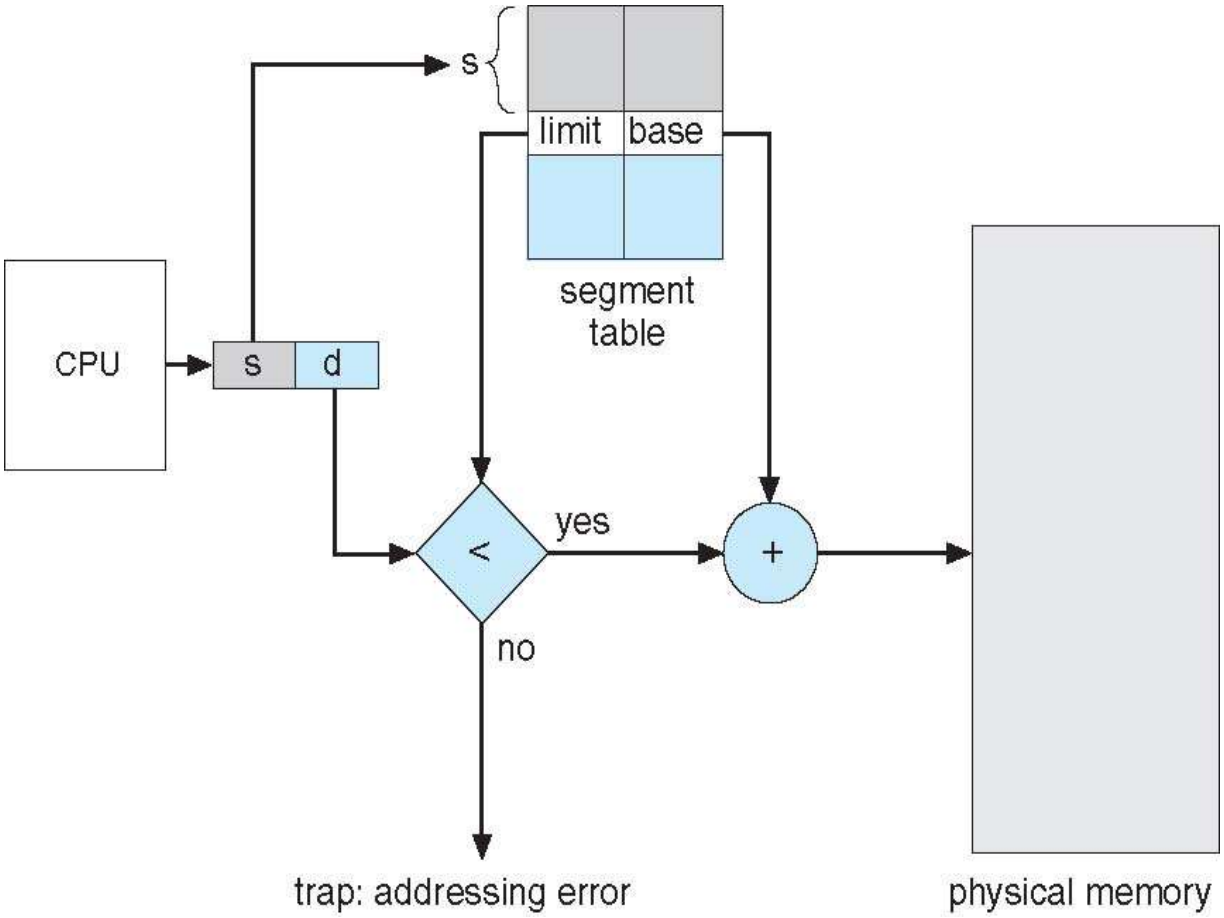
User view of the program



Logical view of segmentation



Segmentation hardware



Segmentation Architecture

- Logical address consists of a two tuple:

  <segment-number, offset>,

- **Segment table** – maps two-dimensional physical addresses; each table entry has:
  - **base** – contains the starting physical address where the segments reside in memory
  - **limit** – specifies the length of the segment
- **Segment-table base register (STBR)** points to the segment table's location in memory
- **Segment-table length register (STLR)** indicates number of segments used by a program;

Protection for segmentation

- Protection
  - With each entry in segment table associate:
    - validation bit = 0 $\Rightarrow$ illegal segment
    - read/write/execute privileges

- Protection bits associated with segments; code sharing occurs at segment level

Advantages of Segmentation

- Eliminates Fragmentation
- Protection mechanism is good
- Allows shared segments among users
- Dynamic loading and linking of segments is possible
- Allow dynamic growing of segments
- Supports modular programming

11.Explain the Difference between paging and segmentation

Paging	segmentation
The main memory is divided into fixed size partitions called pages	The main memory is divided into variable sized partitions called segments
OS maintains a page map table	OS maintains a segment table
Page table consists of two parts - page number and frame number	Segment table consists of three parts segment number , base address and limit
Processes use the page number and offset to calculate physical address	Processes use segment number and offset to calculate physical address
It does not support modular programming	It supports modular programming
It does not support dynamic growth of pages	It supports dynamic growth of segments