

UNIT - III

* Concept of Memory Management -

- The task of subdividing the memory among different processes is called memory management.
- Memory management is a method in the OS to manage operations b/w main memory & disk during process execution.
- Main aim of memory management is utilization of memory.

* Why Memory Management is required?

- (1) Proper utilization of main memory.
- (2) To minimize fragmentation issues.
- (3) To keep track of used memory space by processes.
- (4) To maintain data integrity while executing of process.
- (5) Allocate & deallocate memory before and after process execution.

* Logical & Physical Address Space -

- Logical Address Space - An address generated by CPU is called logical address.
 - * Also known as Virtual Address.
 - * It can be defined as the size of process.
 - * A logical address can be changed.

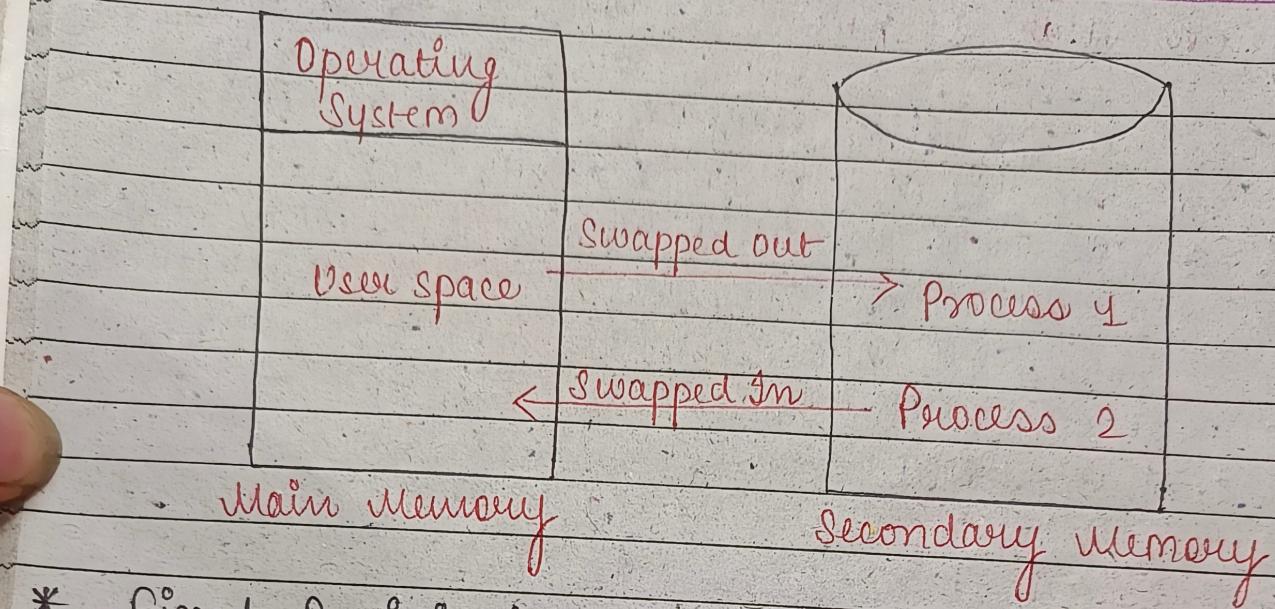
→ Physical Address Space: An address seen by the memory unit is known as physical address.

- * Also known as Real Address.
- * The set of all physical addresses corresponding to these logical addresses is called as Physical Address Space.
- * It is computed by Memory Management Unit (MMU).
- * The physical address always remain constant.

* Swapping -

Swapping is a process of swapping a process temporarily into a secondary memory from the main memory which is fast as compared to secondary memory.

- * Swapping allows more processes to be run & can be fit into memory at a time.
- * The transferred time & total time is directly proportional to the amount of memory swapped.
- * Swapping is also known as roll-out or roll because if a higher priority process arrives & wants service, the memory manager can swap out the lower priority process & then load & execute the higher priority process.
- * After finishing higher priority work, the lower priority work is swapped back in memory & continued to the execution process.



* Fixed Partitioning -

- In this, the memory is divided into fixed size partitions or blocks during system initialization.
- Each partition can accommodate exactly one process.
- When a process needs memory, it is allocated one of the available partitions based on its size. If there is no partition large enough to accommodate the process, it waits until one becomes available.
- Fixed Partitioning is relatively simple to implement but may lead to inefficient memory utilization, especially if the processes vary widely in size.
- The inefficiency in size or memory occurs because small processes may occupy large partitions, leaving unused space within those partitions.
- It suffers from fragmentation both internal (unused memory within a partition) & external (unused memory scattered between partitions).

* Dynamic Partitioning -

- It allows for the allocation of variable sized memory blocks to processes, based on their size requirements.
- When a process requests memory, the system searches for a suitable block of memory in the available free space. If a block is found that is large enough to accommodate the process, it is allocated to the process.
- Dynamic Partitioning doesn't define the partition sizes, allowing for better memory utilization, as memory is divided more flexibly.
- It introduces overhead in terms of managing and tracking free memory blocks.
- It also suffers from fragmentation, especially external fragmentation, where free memory blocks become scattered throughout memory, making it challenging to allocate contiguous blocks of memory to processes.

* First Fit Allocation -

In first-fit, the first available free hole fulfil the requirement of the process allocated.

Process A =	OS 20 KB USED	Hole = 40-25 = 15 KB USED
	15 KB USED	
	40 KB (*) USED	
	60 KB USED	
	25 KB	

* In this diagram, a 40 KB block is the first available free hole, as the first two was not sufficient.

* Best Fit Allocation -

In best fit allocation, the smallest hole that is big enough to process requirements. For this, we search the entire list, unless the list is ordered by size.

+ In this, memory utilization is maximum as compared to other techniques.

Process A = 25 KB

OS
20 KB
USED

15 KB
USED

40 KB
USED

60 KB
USED

25 KB (*)

HOLE = 25 - 25
= 0 KB

* Worst Fit Allocation -

In worst fit, it allocates the largest available hole to process. This process produces the largest leftover holes.

+ Inefficient memory utilization is a major issue in the worst fit.

Process A = 25 KB

OS
20 KB
USED

15 KB
USED

40 KB
USED

60 KB (*)
USED

HOLE = 60 - 25
= 35 KB

25 KB

* Paging

P₁ P₂ P₃ P₄

Given Requests from processes are 300K, 25K, 125K, 50K respectively, the above request could satisfy with -

First Fit

USED
P ₂
150KB ↑ 125KB ↓
P ₃
USED
P ₁
350KB ↑ 50 ↓
P ₄
USED

Best Fit

USED
P ₃
150KB
Unused
USED
P ₁
350KB
50
Unused
USED

Worst Fit

USED
P ₂
150KB ↑ 125KB ↓
P ₃
USED
P ₁
350KB
50
Unused
USED
P ₄
50 ↓

(P₄) cannot
be allocated.

Ques. Given six memory partitions of 300 KB, 600 KB, 350 KB, 200 KB, 750 KB, 125 KB in order, how would the first fit, best fit, worst fit algorithms place processes of size 115 KB, 500 KB, 358 KB, 200 KB, 375 KB in order in fixed size partitioning?

$$P_1 = 115, P_2 = 500, P_3 = 358, P_4 = 200, P_5 = 375 \text{ KB}$$

	First fit	Best fit	Worst fit
(185)	P1 Unused	300KB Unused	300KB Unused
(100)	P2 Unused	P2 Unused	600 KB Unused
(150)	P4 Unused	350KB Unused	350, KB Unused
392	Unused P3 Unused P5 Unused Unused	200 KB P4 P3 P5 Unused P1 Unused	200KB Unused 750 KB Unused 125KB Unused Unused

(P5) cannot be allocated.

* Paging -

- It is a storage mechanism used to retrieve processes from the secondary storage into the main memory in the form of pages.
- Main idea is to divide each process in the form of pages.
- Main memory is divided in form of frames.
- One page of the process is stored in one of the frames of memory.
- Pages of process are brought in main memory only when needed; otherwise stored in secondary memory.
- Each frame must be equal &
 $\text{Page Size} = \text{Frame Size}$

* Memory Management Unit -

- Main Purpose: convert logical address \rightarrow Physical Address.
- logical address is generated by CPU for every page & physical address is the actual address of the frame where each page will be stored.
- The logical address has 2 parts:

- ① Page Number
- ② Offset

MMU converts page number to the frame number.

* Segmentation -

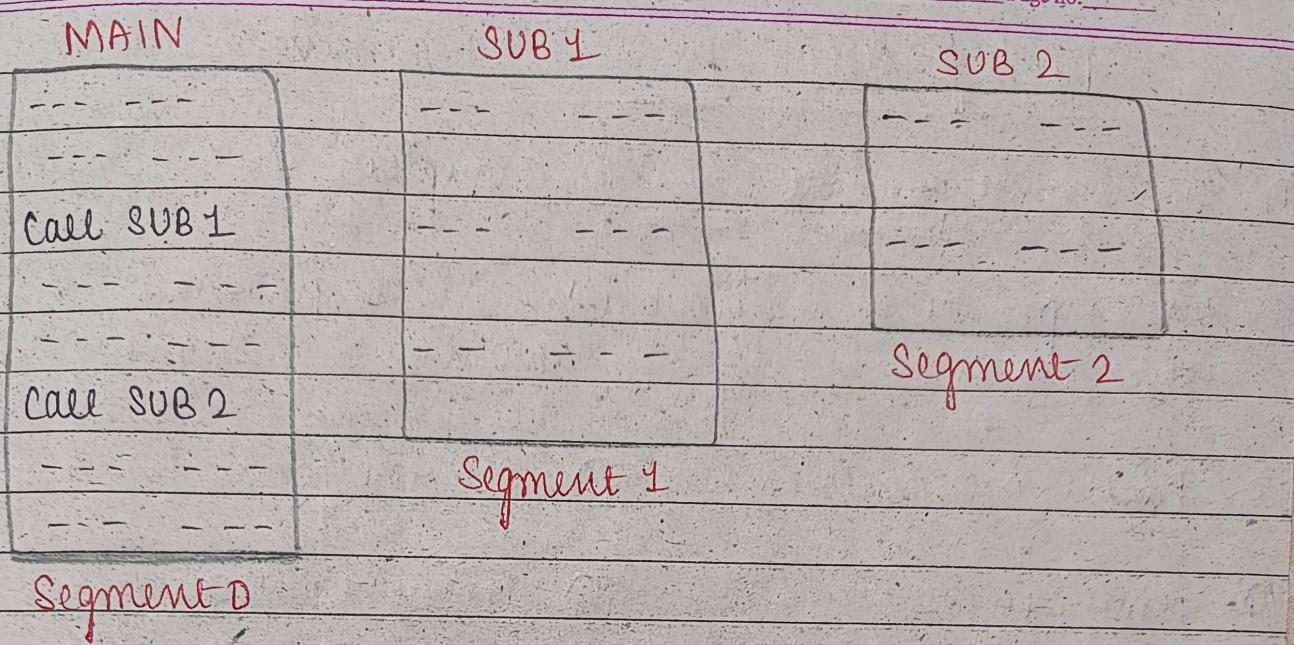
- It is a memory management technique in which memory is divided into variable size parts.
- Each part is known as Segment & can be allocated to a process.
- details about each segment is stored in Segment Table.
- Segment Table contains mainly two information about segments.

- ①. Base: base address of segment.
- ②. Limit: length of the segment.

* Need of Segmentation -

Paging technique is more close to the OS than the User. Paging divides the processes into the form of pages regardless of the fact that a process can have some relative parts of functions which need to be loaded into the same page. The OS doesn't care about User's view of process. It may divide the same function into diff. pages & those pages may or may not be loaded at the same time into the memory. It decreases the efficiency of system.

∴ Segmentation is required. Each segment contains the same type of functions such as the main functn can be included in one segment & the library functions can be included in the other segment.



* Advantages of Segmentation -

- ①. No internal fragmentation.
- ②. Less overhead.
- ③. Avg segment size is larger than Actual Page size.
- ④. It is easier to relocate segments than entire address space.
- ⑤. The segment table is of lesser size as compared to the page table in paging.

* Disadvantages of Segmentation -

- ①. It can have external fragmentation.
- ②. Costly Memory Management Algorithms.
- ③. It is difficult to allocate contiguous memory to variable sized partition.

Paging

①. OS is responsible.

②. Divides program into fixed size pages.

③. Faster process.

④. closer to OS.

⑤. suffers from internal fragmentation.

⑥. Logical address is divided into page no. & page offset.

⑦. Page table maintains page information.

⑧. Page table entry has the frame no. & some flag bits to represent details about pages.

Segmentation

①. Compiler is responsible.

②. Divides program into variable size segments.

③. Slower process.

④. closer to User.

⑤. suffers from external fragmentation.

⑥. Logical address is divided into segment no. & segment offset.

⑦. Segment table maintains segment information.

⑧. Segment table entry has the base address of the segment & some protection bits for the segment.

* Segmented Paging -

- Not very popular & not being used in many OS.
- Segmentation can be combined with Paging to get the best features out of both the techniques.
- In segmented Paging, the main memory is divided into variable size segments which are further divided into fixed size pages.
- Pages are smaller than segments.
- Each segment has a page table which means every program has multiple page tables.
- The logical address is represented as Segment No., Page No. & Page offset.

Segment No. : it points to appropriate Segment Number.

Page No. : it points to the exact page within Segment.

Page Offset : used as an offset within the page frame.

* Advantages -

- ①. It reduces memory usage.
- ②. NO external fragmentation.
- ③. It simplifies memory allocation.
- ④. Page table size is limited by segment size.
- ⑤. Segment table has only one entry corresponding to one actual segment.

* Disadvantages -

- ①. Internal fragmentation is there.
- ②. High complexity level as compared to Paging.
- ③. Page tables need to be contiguously stored in memory.