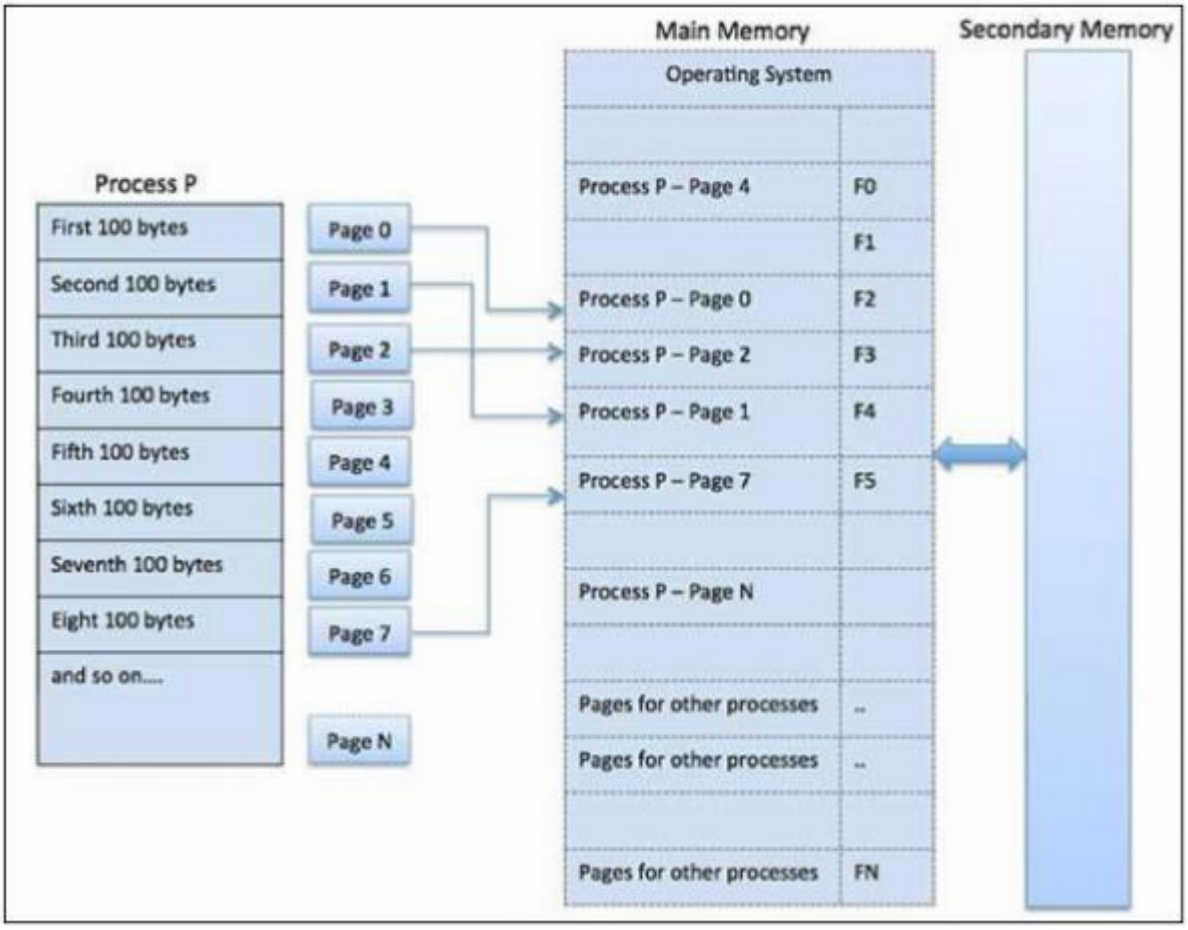
III PAGING:

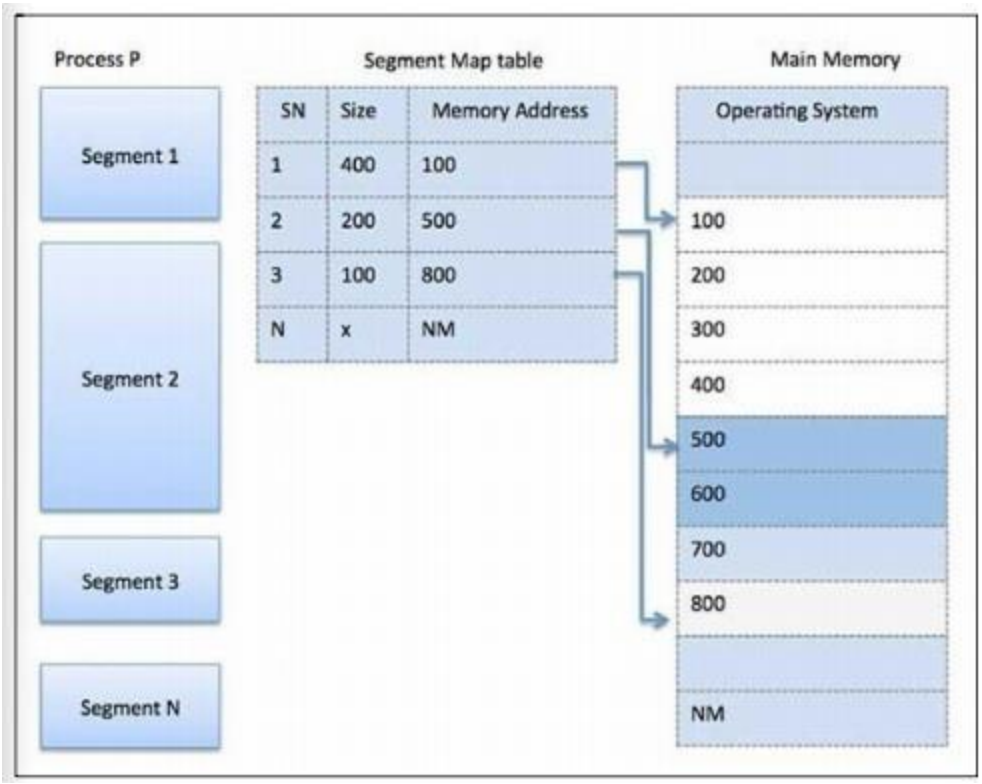
- The address space of a process is broken into fixed sized blocks.
- These fixed size blocks are known as pages.
- The operating system divides the memory blocks into pages.
- The size of the page is determined based on the memory available.
- This technique is quick in terms of memory access.
- It can cause internal fragmentation since some pages would not be utilized as much as the other pages.
- During the process of paging, a logical address is divided into page number and page offset.
- A page table is used to store the page data.This is how Paging works −



Segmentation

- In this method, the address space of a process is broken down into varying sized blocks.
- These varying sized blocks are known as sections.
  - A compiler is responsible in determining the size of the segment, the virtual address and the actual address.
  - The size of the section is determined by the user.
  - The process of segmentation is slower in comparison to paging.
  - It can result in external fragmentation since some memory blocks may not be used at all.
  - During this process, a logical address gets divided into a section number and a section offset.
  - A segmentation table can be used to store the segmentation data.

This is how Segmentation techniques works with Segment Map Table –



Both paging and swapping are important concepts in operating systems that place a process in the main memory for its execution, but they are quite different from each other in many aspects. Read this article to learn more about **paging** and **swapping** and their specific characteristics.

What is Paging?

In OS, **Paging** is a memory management strategy in which the process address space is divided into blocks of the same size, called **pages** (where the size of each page is power of 2, and is between 512 bytes and 8192 bytes). The size of the process is then measured in the number of pages.

In the same way, the main memory is divided into small blocks of fixed size called **frames**. The size of

each frame is kept the same as that of a page to have the optimum utilization of the main memory and to avoid the external fragmentation. Therefore, paging is basically a memory allocation technique. It utilizes non-contiguous memory management technique.

What is Swapping?

Swapping is a memory management technique in which an entire process is copied to another location. In other words, swapping is a technique in which an entire process is to be placed in the main memory for its execution. Also, swapping removes the inactive processes from the main memory of the system. Swapping helps to provide memory space for the operation of other processes. Hence, swapping impacts the performance of a system, as it helps in executing multiple large operations concurrently. Swapping can be done without using any memory management technique.

Now, let us discuss the differences between paging and swapping in detail.

Difference between Paging and Swapping in OS

The following are the important differences between paging and swapping in Os −

S.No.	Swapping	Paging
1.	It is process where the entire process is copied to another location.	It is a memory allocation technique.
2.	This process occurs when the entire process has been transferred to the disk.	This process occurs when a part of a process is transferred to the disk.
3.	Here, the data is swapped temporarily from the main memory to a secondary memory.	The contiguous block of memory is made non-contiguous but it consists of fixed size called frames known as pages.

4.	It can be done without using anymemory management methods.	It uses non-contiguous memory management technique.
5.	It can be done by processes that are inactive as well.	Only a process that is currently active canperform paging operation.
6.	It helps give a direction with respectto the solution.	There is no suggestion about the solution inthis technique.

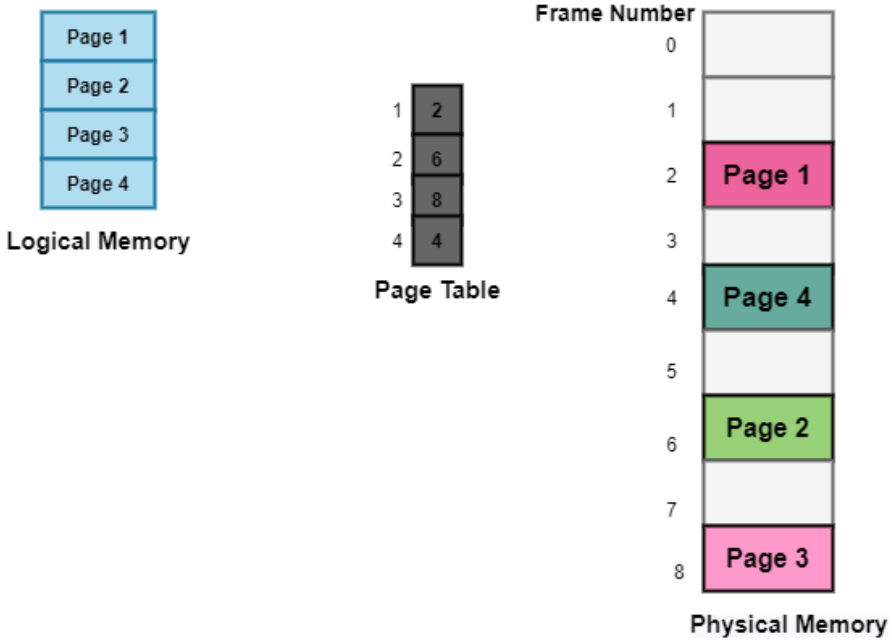
Structure of Page Table in Operating Systems

In this tutorial, we will cover some of the most common techniques used for structuring the Page table.

The data structure that is used by the virtual memory system in the operating system of a computer in order to store the mapping between physical and logical addresses is commonly known as **Page Table**.

As we had already told you that the logical address that is generated by the CPU is translated into the

- Thus page table mainly provides the corresponding frame number (base address of the frame)

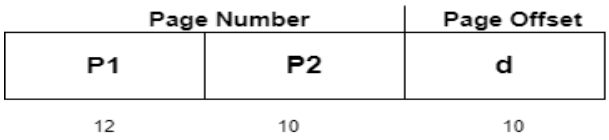


- physical address with the help of the page table.

The above diagram shows the paging model of Physical and logical memory.

Characteristics of the Page Table

Some of the characteristics of the Page Table are as follows:



- It is stored in the main memory.
- Generally; the Number of entries in the page table = the Number of Pages in which the processis divided.
- **PTBR** means page table base register and it is basically used to hold the base address for thepage table of the current process.
- Each process has its own independent page table.

Techniques used for Structuring the Page Table

Some of the common techniques that are used for structuring the Page table are as follows:

1. Hierarchical Paging
2. Hashed Page Tables
3. Inverted Page Tables

Let us cover these techniques one by one;

Hierarchical Paging

Another name for Hierarchical Paging is multilevel paging.

- There might be a case where the page table is too big to fit in a contiguous space, so we mayhave a hierarchy with several levels.
- In this type of Paging the logical address space is broke up into Multiple page tables.
- Hierarchical Paging is one of the simplest techniques and for this purpose, a two-level pagetable and three-level page table can be used.

Two Level Page Table

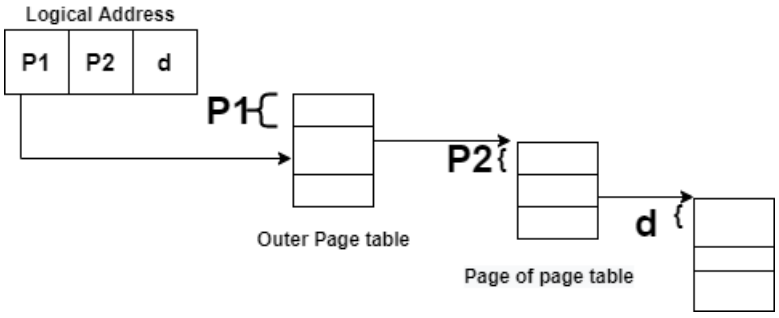
Consider a system having 32-bit logical address space and a page size of 1 KB and it is further divided into:

- Page Number consisting of 22 bits.
- Page Offset consisting of 10 bits.

As we page the Page table, the page number is further divided into :

- Page Number consisting of 12 bits.
- Page Offset consisting of 10 bits.

Thus the Logical address is as follows:



In the above diagram,

P1 is an index into the **Outer Page** table.

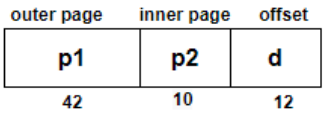
P2 indicates the displacement within the page of the **Inner page** Table.

As address translation works from outer page table inward so is known as **forward-mapped Page Table**.

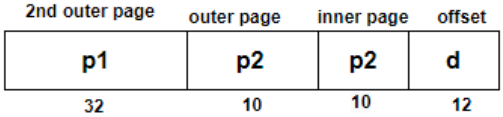
Below given figure below shows the Address Translation scheme for a two-level page table

Three Level Page Table

For a system with 64-bit logical address space, a two-level paging scheme is not appropriate. Let us suppose that the page size, in this case, is 4KB.If in this case, we will use the two-page level schemethen the addresses will look like this:



Thus in order to avoid such a large table, there is a solution and that is to divide the outer page table,it will result in a **Three-level page table:**



Hashed Page Tables

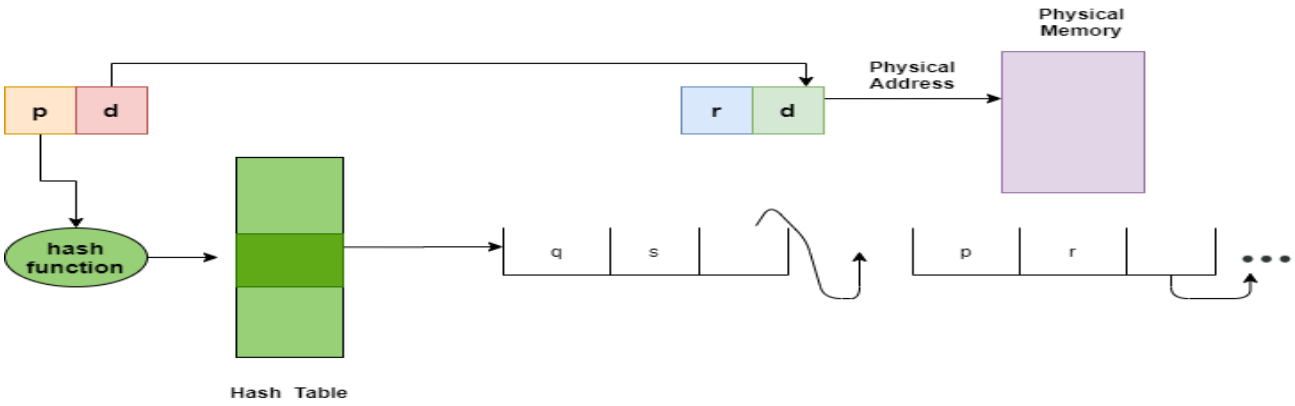
This approach is used to handle address spaces that are larger than 32 bits.

- In this virtual page, the number is hashed into a page table.
- This Page table mainly contains a chain of elements hashing to the same elements.

Each element mainly consists of :

1. The virtual page number
2. The value of the mapped page frame.
3. A pointer to the next element in the linked list.

Given below figure shows the address translation scheme of the Hashed Page Table:



The above Figure shows Hashed Page Table The Virtual Page numbers are compared in this chain searching for a match; if the match is found thenthe corresponding physical frame is extracted. In this scheme, a variation for 64-bit address space commonly uses **clustered page tables**

These are similar to hashed tables but here each entry refers to several pages (that is 16) ratherthan 1.

- Mainly used for sparse address spaces where memory references are non-contiguous andscattered

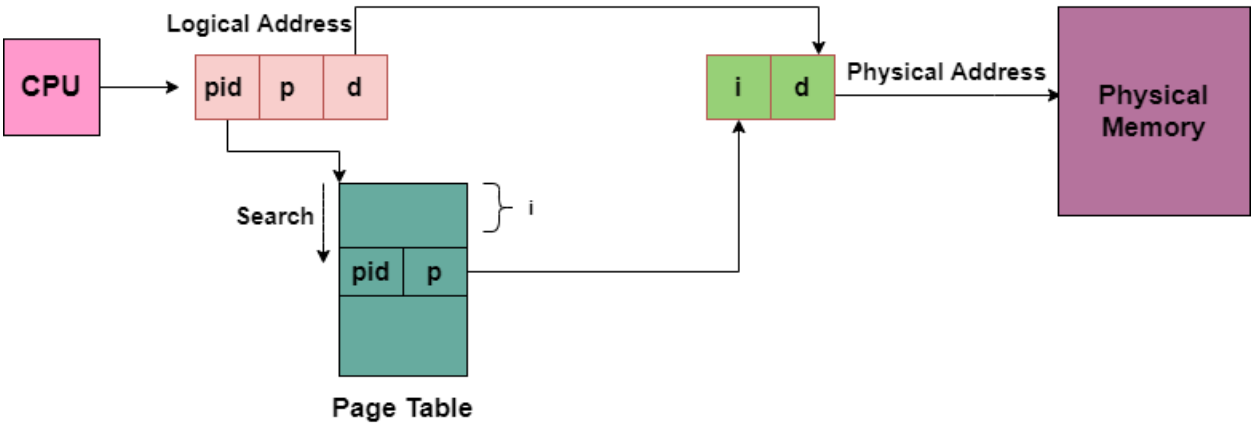
Inverted Page Tables

The Inverted Page table basically combines A page table and A frame table into a single data structure.

- There is one entry for each virtual page number and a real page of memory
- And the entry mainly consists of the virtual address of the page stored in that real memorylocation along with the information about the process that owns the page.
- Though this technique decreases the memory that is needed to store each page table; but it also

increases the time that is needed to search the table whenever a page reference occurs.

Given below figure shows the address translation scheme of the Inverted Page Table:



In this, we need to keep the track of process id of each entry, because many processes may have the same logical addresses.

Also, many entries can map into the same index in the page table after going through the hashfunction. Thus chaining is used in order to handle this.

Segmentation in Operating Systems

In this tutorial, we will be covering segmentation in the Operating System.

Segmentation is another way of dividing the addressable memory. It is another scheme of memory management and it generally supports the user view of memory. The Logical address space is basically the collection of segments. Each segment has a name and a length.

Basically, a process is divided into segments. Like paging, segmentation divides or segments the memory. But there is a difference and that is while the **paging** divides the memory into a **fixed size** andon the other hand, segmentation divides the **memory into variable segments** these are then loaded into logical memory space.

A Program is basically a collection of segments. And a segment is a logical unit such as:

- main program
- procedure
- function
- method
- object
- local variable and global variables.

- symbol table
- common block
- stack
- arrays

Types of Segmentation

Given below are the types of Segmentation:

- **Virtual Memory Segmentation** With this type of segmentation, each process is segmented inton divisions and the most important thing is they are not segmented all at once.
- **Simple Segmentation** With the help of this type, each process is segmented into n divisions andthey are all together segmented at once exactly but at the runtime and can be non-contiguous (that is they may be scattered in the memory).

Characteristics of Segmentation

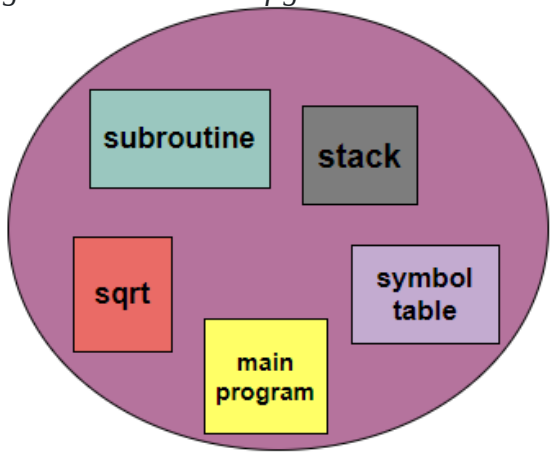
Some characteristics of the segmentation technique are as follows:

- The Segmentation partitioning scheme is **variable-size.**
- Partitions of the secondary memory are commonly known as **segments**.
- Partition size mainly depends upon the length of modules.
- Thus with the help of this technique, secondary memory and main memory are divided intounequal-sized partitions.

Need of Segmentation

One of the important drawbacks of memory management in the Operating system is the separation of the user's view of memory and the actual physical memory. and Paging is a technique that provides the separation of these two memories. User'view is basically mapped onto physical storage. And this mapping allows differentiation betweenPhysical and logical memory.

User's View of a Program - Given below figure shows the user's view of segmentation:



logical address

Basic Method

A computer system that is using segmentation has a logical address space that can be viewed as multiple segments. And the size of the segment is of the variable that is it may grow or shrink. As we had already told you that during the execution each segment has a name and length. And the address mainly specifies both thing name of the segment and the displacement within the segment.

Therefore the user specifies each address with the help of two quantities: segment name and offset.

For simplified Implementation segments are numbered; thus referred to as segment number ratherthan segment name.

Thus the logical address consists of two tuples:

<segment-number,offset>

where,

Segment Number(s): Segment Number is used to represent the number of bits that are required to represent the segment.

Offset(d) Segment offset is used to represent the number of bits that are required to represent the sizeof the segment.

Segment Table Base Register(STBR) The STBR register is used to point the segment table's location in the memory.

	Limit	Base
Segment 0 →	1400	1400
Segment 1 →	400	6200
Segment 2 →	1100	4400
Segment 3 →	1300	4800

Segment Table

Segmentation Architecture

Segment Table

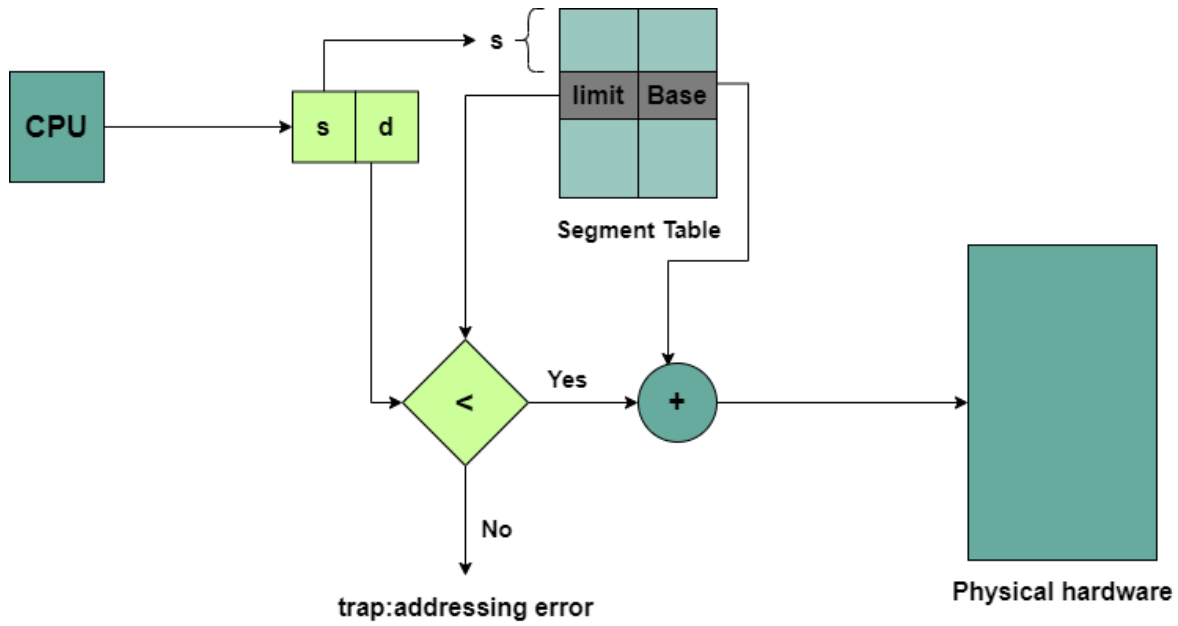
A Table that is used to store the information of all segments of the process is commonly known as Segment Table. Generally, there is no simple relationship between logical addresses and physical addresses in this scheme.

- The mapping of a two-dimensional Logical address into a one-dimensional Physical address isdone using the segment table.
- This table is mainly stored as a separate segment in the main memory.
- The table that stores the base address of the segment table is commonly known as theSegment table base register (STBR)

In the segment table each entry has :

1. **Segment Base/base address:** The segment base mainly contains the starting physical addresswhere the segments reside in the memory.
2. **Segment Limit:** The segment limit is mainly used to specify the length of the segment.

Given below figure shows the segmentation hardware :



The logical address generated by CPU consist of two parts:

1. Segment Number(s): It is used as an index into the segment table.
2. Offset(d): It must lie in between '0' and 'segment limit'.In this case, if the Offset exceeds the segmentlimit then the trap is generated.

Thus; **correct offset+segment base= address in Physical memory**

and segment table is basically an array of base-limit register pair.

Advantages of Segmentation

The Advantages of the Segmentation technique are as follows:

- In the Segmentation technique, the segment table is mainly used to keep the record of segments. Also, the segment table occupies less space as compared to the paging table.
- There is no Internal Fragmentation.
- Segmentation generally allows us to divide the program into modules that provide bettervisualization.
- Segments are of variable size.

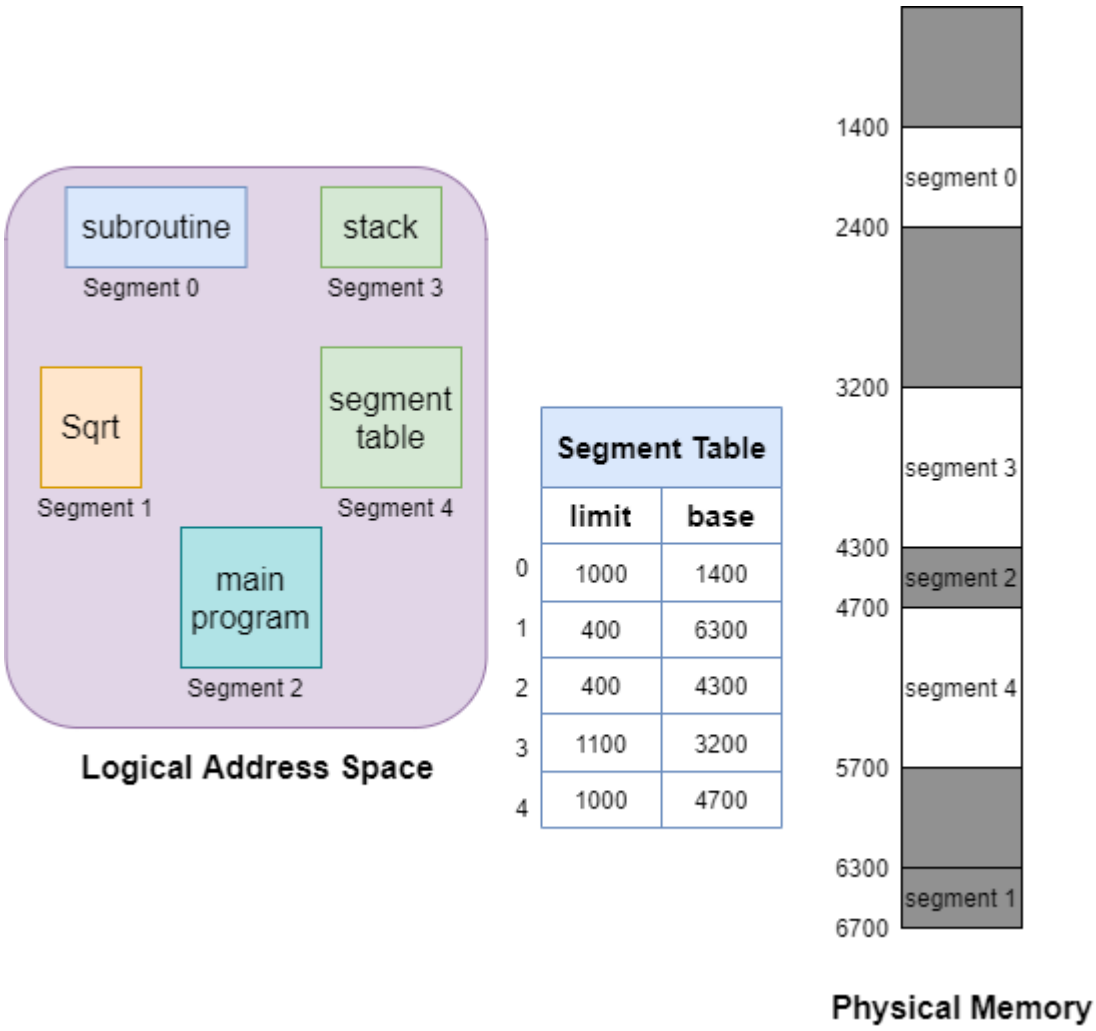
Disadvantages of Segmentation

Some disadvantages of this technique are as follows:

- Maintaining a segment table for each process leads to overhead
- This technique is expensive.
- The time is taken in order to fetch the instruction increases since now two memory accessesare required.
- Segments are of unequal size in segmentation and thus are not suitable for swapping.
- This technique leads to external fragmentation as the free space gets broken down into smallerpieces along with the processes being loaded and removed from the main memory then this will result in a lot of memory waste.

Example of Segmentation

Given below is the example of the segmentation, There are five segments numbered from 0 to 4. These segments will be stored in Physical memory as shown. There is a separate entry for each segment in the segment table which contains the beginning entry address of the segment in the physical memory(denoted as the base) and also contains the length of the segment(denoted as limit).



Segment 2 is 400 bytes long and begins at location 4300. Thus in this case a reference to byte 53 of segment 2 is mapped onto the location 4300 (4300+53=4353). A reference to segment 3, byte 85 is mapped to 3200(the base of segment 3)+852=4052.

A reference to byte 1222 of segment 0 would result in the trap to the OS, as the length of this segmentis 1000 bytes.

Difference between Paging and Segmentation

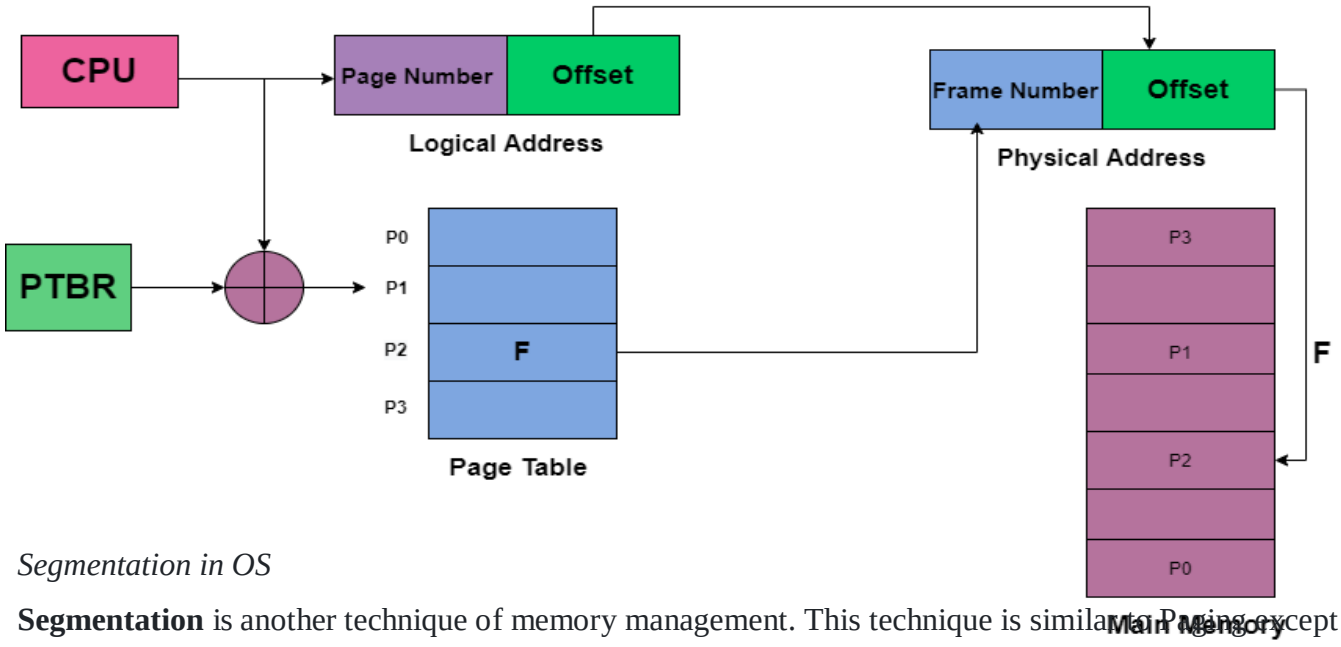
In this tutorial we will be covering the difference between Paging and Segmentation these are two different Memory management techniques in Operating systems.

Paging in OS

Paging is basically a technique of memory management in the operating system. In this technique, the

operating system retrieves the processes from the secondary memory into the main memory and this memory is in the form of pages. Paging is a logical concept.

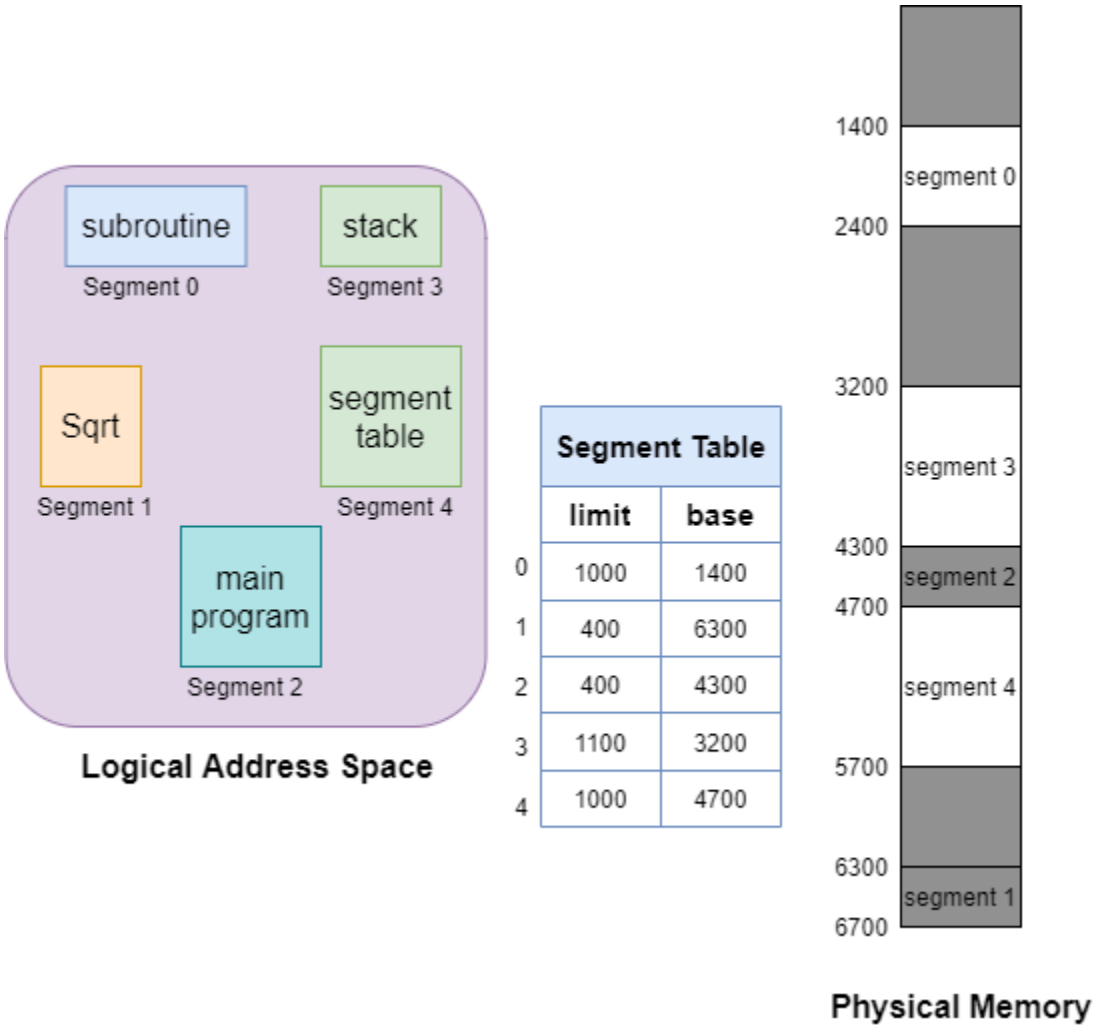
With the help of this technique, the main memory is split into the small blocks of physical memory that are commonly known as frames. In paging size of frames is fixed. In order to prevent external fragmentation and for the maximum usage of the main memory, the frame size must be the same as the page size. This technique helps to access the data faster.



Segmentation in OS

Segmentation is another technique of memory management. This technique is similar to Paging except for the fact that in segmentation the segments of a process are of variable length but in Paging the pages are of fixed size.

The memory is split into variable-length parts in segmentation. Each part is commonly known as segments. Information related to each segment is stored in a table and this table is commonly knownas **the segment table.** The **segment table** generally occupies less space as compared to the **paging table.**



Now, we will cover the differences between Paging and Segmentation in the table given

Paging	Segmentation
Paging is a memory management technique where memory is partitioned into fixed-sizedblocks that are commonly known as **pages.**	Segmentation is also a memory management techniquewhere memory is partitioned into variable-sized blocks that are commonly known as **segments**.

Paging	Segmentation
With the help of Paging, the logical address isdivided into a **page number** and **page offset**.	With the help of Segmentation, the logical address isdivided into **section number** and **section offset**.
This technique may lead to **Internal Fragmentation**.	Segmentation may lead to **External Fragmentation**.
In Paging, the page size is decided by thehardware.	While in Segmentation, the size of the segment isdecided by the user.
In order to maintain the page data, the page tableis created in the Paging	In order to maintain the segment data, the segment tableis created in the Paging
The page table mainly contains the base addressof each page.	The segment table mainly contains the segment numberand the offset.
This technique is faster than segmentation.	On the other hand, segmentation is slower than paging.
In Paging, a list of free frames is maintained by the	In Segmentation, a list of holes is maintained by the