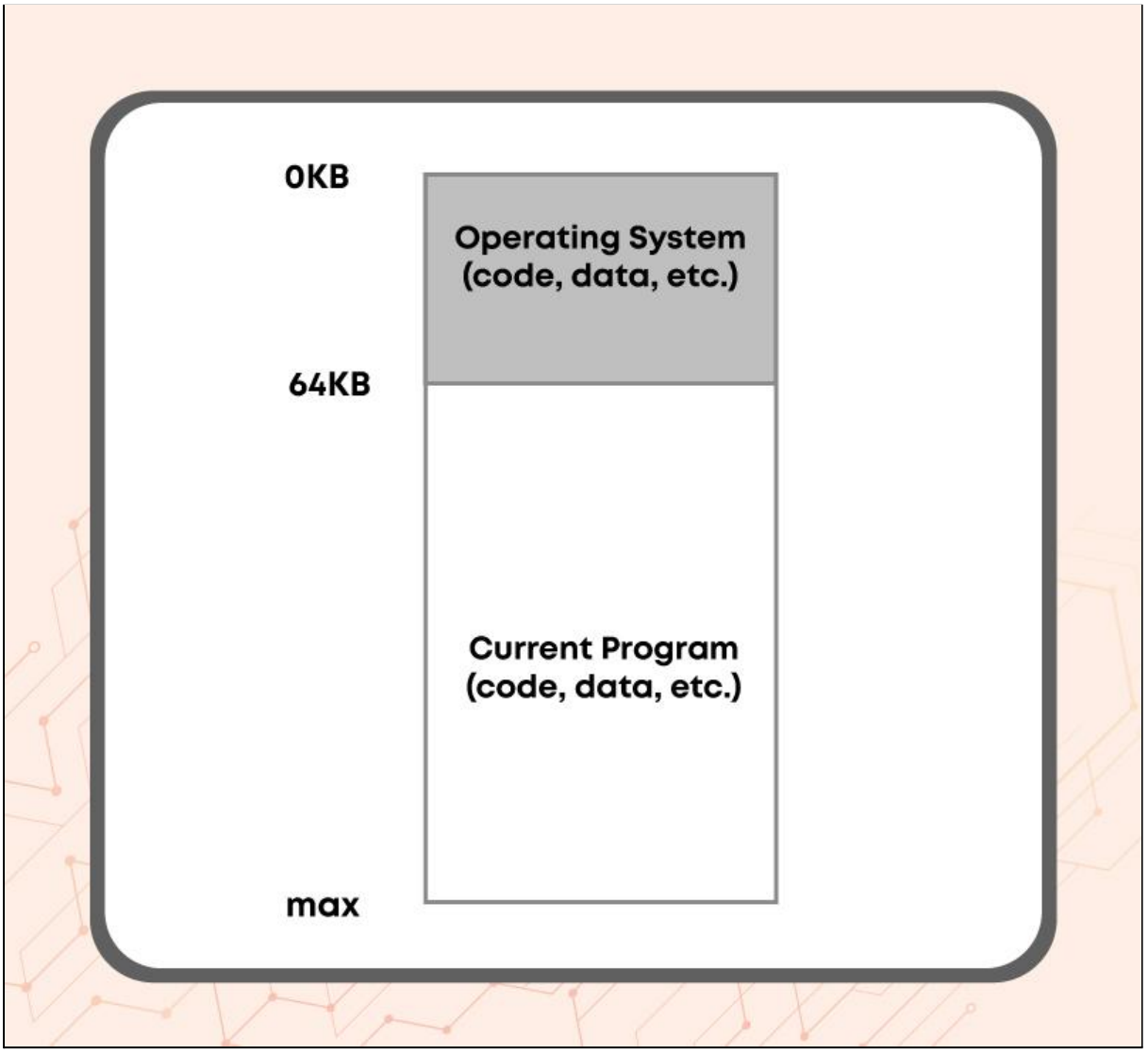
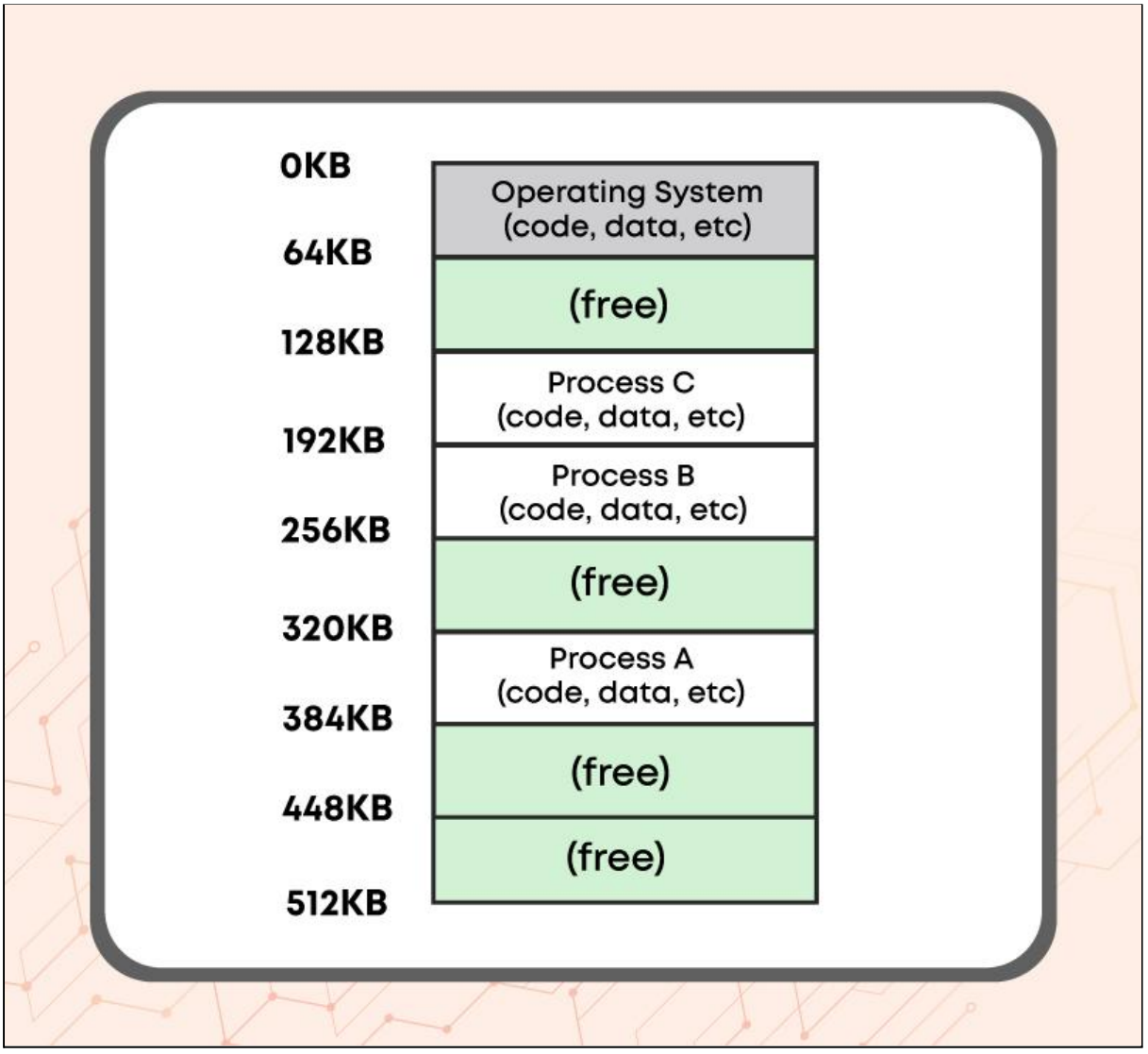
Summary Note: Introduction to Memory Management

Introduction

In the early days, the OS was added in the main memory first and then the process was allocated space in memory. Only one process was added to main memory at a time and therefore, the process used to have access to the entire memory.



As time passed, requests for more efficient utilization of memory came and therefore, we added multiple processes simultaneously in main memory.



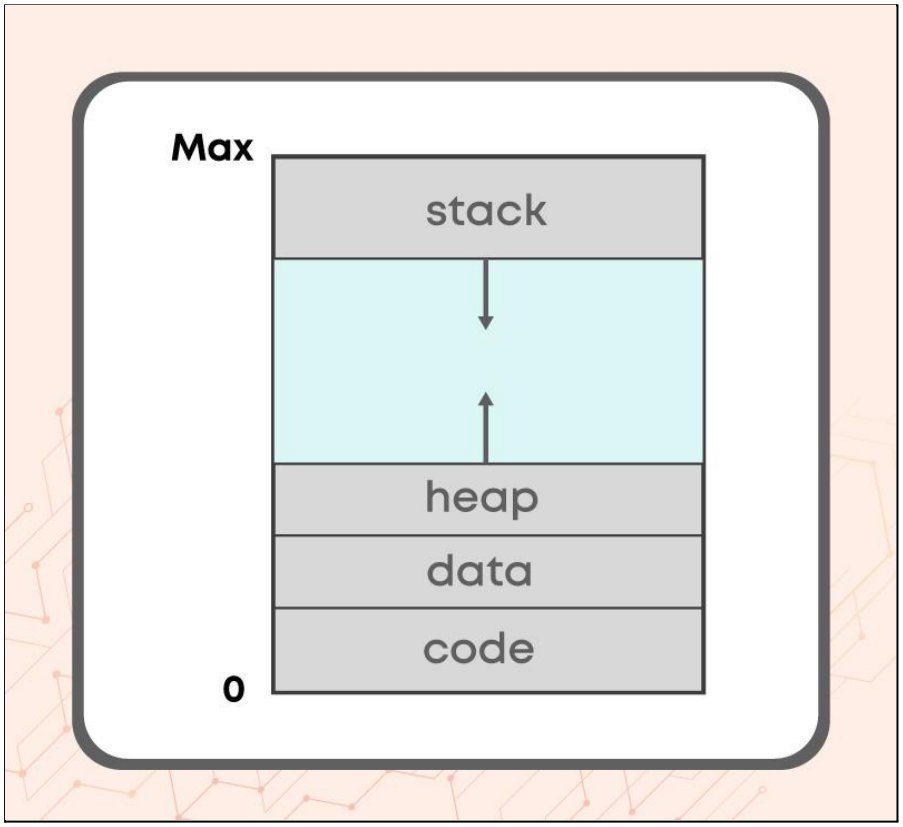
This leads to the need of providing isolation and protection to processes from one another. To counter this requirement, virtual private address space was provided to each process. The addresses of virtual space were mapped to physical addresses to establish isolation and protection of processes.

Virtual and Physical Address Space

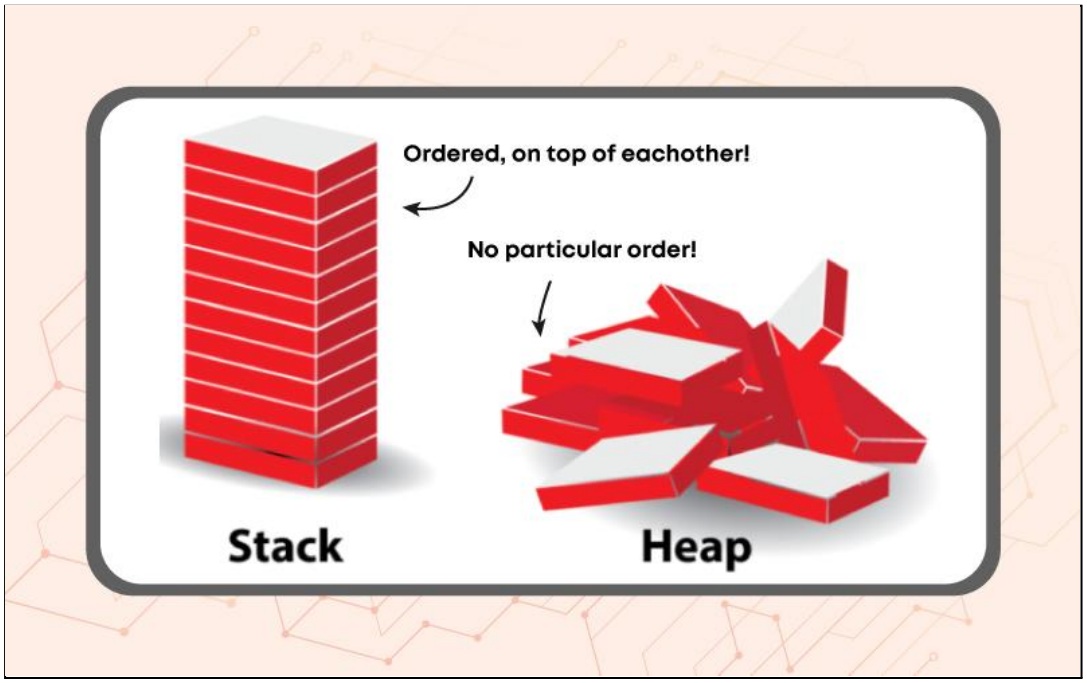
Virtual address is the one that is generated by CPU, also referred to as logical address. The program perceives this address space. Virtual address space is the set of all virtual addresses generated by a program.

Physical address is the actual address understood by computer hardware. Physical address space is the set of all physical addresses generated by a program.

A virtual address space has code, data, stack and heap.



We can see two types of memory in virtual address space: Stack and Heap. Stack memory is used for storing local variables, function arguments, and return variables, whereas Heap is used for storing dynamically allocated data structures. The following diagram illustrates the difference between them.



Moreover, there are two ways virtual address space can be allocated to main memory: Contiguous and Non-Contiguous.

Storage Allocation and Management Techniques

The Storage allocation can be of two types:

(i) Contiguous storage allocation.

(ii) Non-contiguous storage allocation.

In this lecture, we will discuss contiguous storage allocation techniques.

Contiguous Storage Allocation

► Contiguous storage allocation implies that a program's data and instructions are assured to occupy a single contiguous memory area.

► It is further subdivided into fixed - partition storage allocation strategy and dynamic or variable - partition storage allocation strategy.

Contiguous: Divide the memory into fixed-sized blocks

1. **Fixed-partition contiguous storage allocation**

The processes with small address space use small partitions and processes with large address space use large partitions. This is known as fixed partition contiguous storage allocation.

2. **Variable - partition contiguous storage allocation**

This notion is derived from parking of vehicles on the sides of streets where the one who manages to enter will get the space. Two vehicles can leave a space between them that cannot be used by any vehicle. This means that whenever a process needs memory, a search for the space needed by it is done. If contiguous space is available to accommodate that process, then the process is loaded into memory.

External Fragmentation

This phenomenon of entering and leaving the memory can cause the formation of unusable memory holes (like the unused space between two vehicles). This is known as External Fragmentation.