

What is paging? Why paging is used?

Paging is a memory management technique in which the memory is divided into fixed size pages. Paging is used for faster access to data. When a program needs a page, it is available in the main memory as the OS copies a certain number of pages from your storage device to main memory. Paging allows the physical address space of a process to be noncontiguous.

What is paging? Why paging is used?

OS performs an operation for storing and retrieving data from secondary storage devices for use in main memory. Paging is one of such memory management scheme. Data is retrieved from storage media by OS, in the same sized blocks called as pages. Paging allows the physical address space of the process to be non contiguous. The whole program had to fit into storage contiguously.

Paging is to deal with external fragmentation problem. This is to allow the logical address space of a process to be noncontiguous, which makes the process to be allocated physical memory.

A computer can address more memory than the amount physically installed on the system. This extra memory is actually called virtual memory and it is a section of a hard that's set up to emulate the computer's RAM. Paging technique plays an important role in implementing virtual memory.

Paging is a memory management technique in which process address space is broken into blocks of the same size called **pages** (size is power of 2, between 512 bytes and 8192 bytes). The size of the process is measured in the number of pages.

Similarly, main memory is divided into small fixed-sized blocks of (physical) memory called **frames** and the size of a frame is kept the same as that of a page to have optimum utilization of the main memory and to avoid external fragmentation.

Address Translation

Page address is called **logical address** and represented by **page number** and the **offset**.

Logical Address = Page number + page offset

Frame address is called **physical address** and represented by a **frame number** and the **offset**.

Physical Address = Frame number + page offset

A data structure called **page map table** is used to keep track of the relation between a page of a process to a frame in physical memory.

Advantages and Disadvantages of Paging

Here is a list of advantages and disadvantages of paging –

- Paging reduces external fragmentation, but still suffers from internal fragmentation.
- Paging is simple to implement and assumed as an efficient memory management technique.
- Due to equal size of the pages and frames, swapping becomes very easy.
- Page table requires extra memory space, so may not be good for a system having small RAM.

Page faults

When a program tries to reference a page not currently present in RAM, the processor treats this invalid memory reference as a **page fault** and transfers control from the program to the operating system. The operating system must:

1. Determine the location of the data on disk.
2. Obtain an empty **page frame** in RAM to use as a container for the data.
3. Load the requested data into the available page frame.
4. Update the **page table** to refer to the new page frame.
5. Return control to the program, transparently retrying the **instruction** that caused the page fault.

When all page frames are in use, the operating system must select a page frame to reuse for the page the program now needs. If the evicted page frame was **dynamically allocated** by a program to hold data, or if a program modified it since it was read into RAM (in other words, if it has become "dirty"), it must be written out to disk before being freed. If a program later references the evicted page, another page fault occurs and the page must be read back into RAM.

The method the operating system uses to select the page frame to reuse, which is its **page replacement algorithm**, is important to efficiency. The operating system predicts the page frame least likely to be needed soon, often through the **least recently used** (LRU) algorithm or an algorithm based on the program's **working set**. To further increase responsiveness, paging systems may predict which pages will be needed soon, preemptively loading them into RAM before a program references them.