Module 4: Memory Management

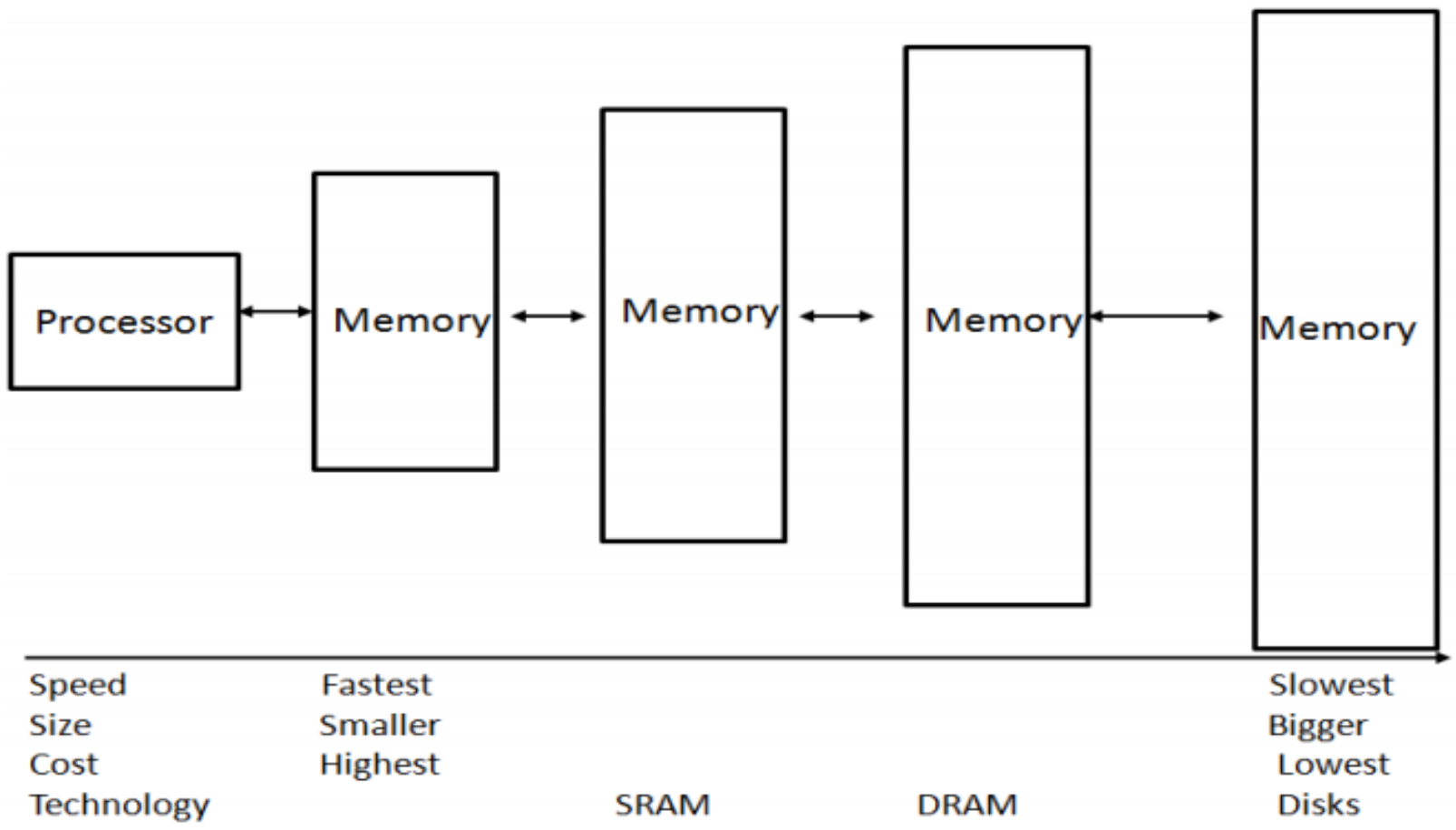


Figure: The basic structure of a memory hierarchy

Memory hierarchy :

1. Multiple levels of memories,
2. Distance from the processor,
3. More the distance or Less the distance,
4. Size of the memories,
5. Access time.

Memory Technologies

1. STATIC Random Access Memory - Static RAM,
2. DYNAMIC Random Access Memory - Dynamic RAM,
3. Flash Memory,
4. Disk Memory,

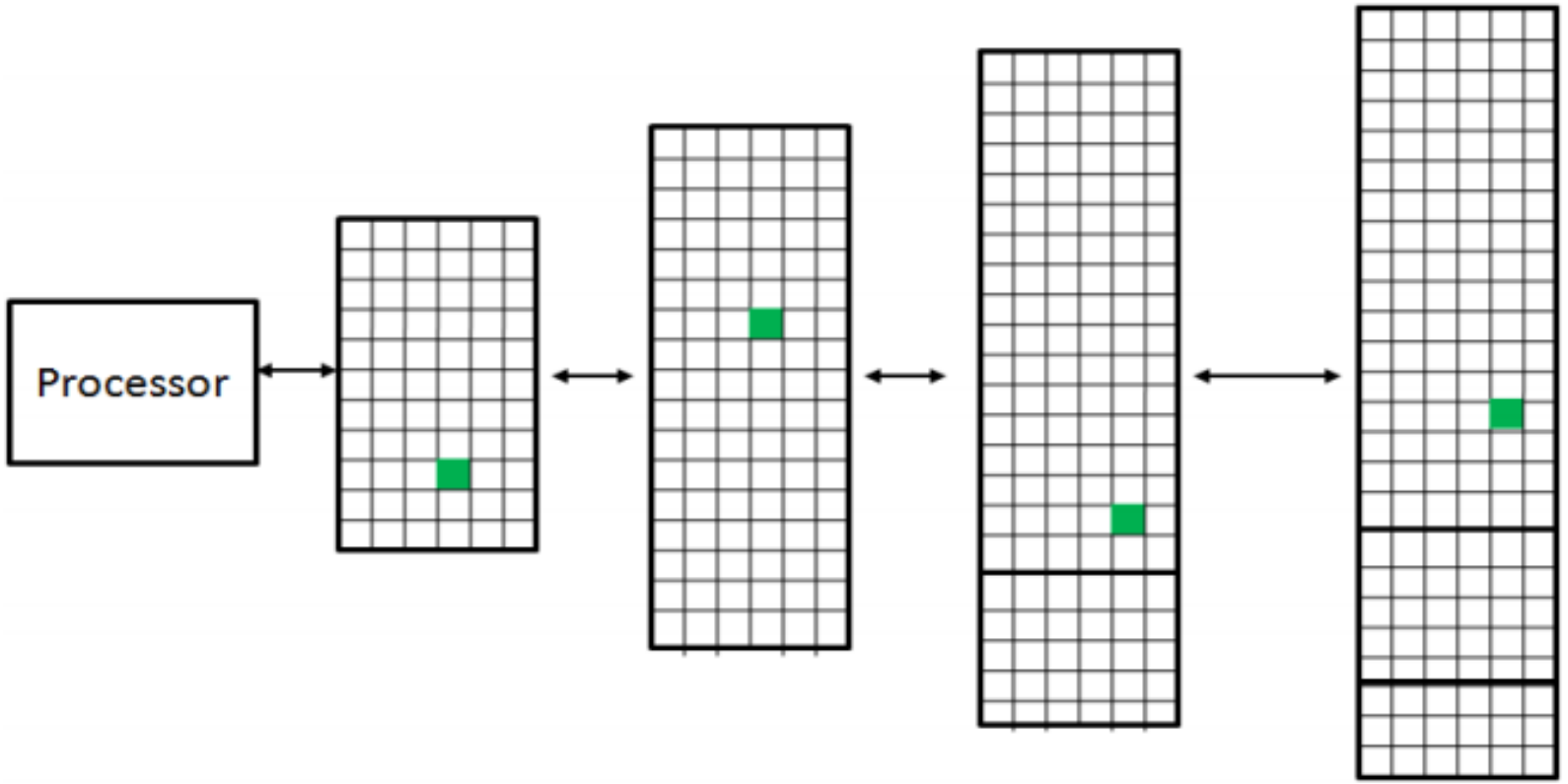
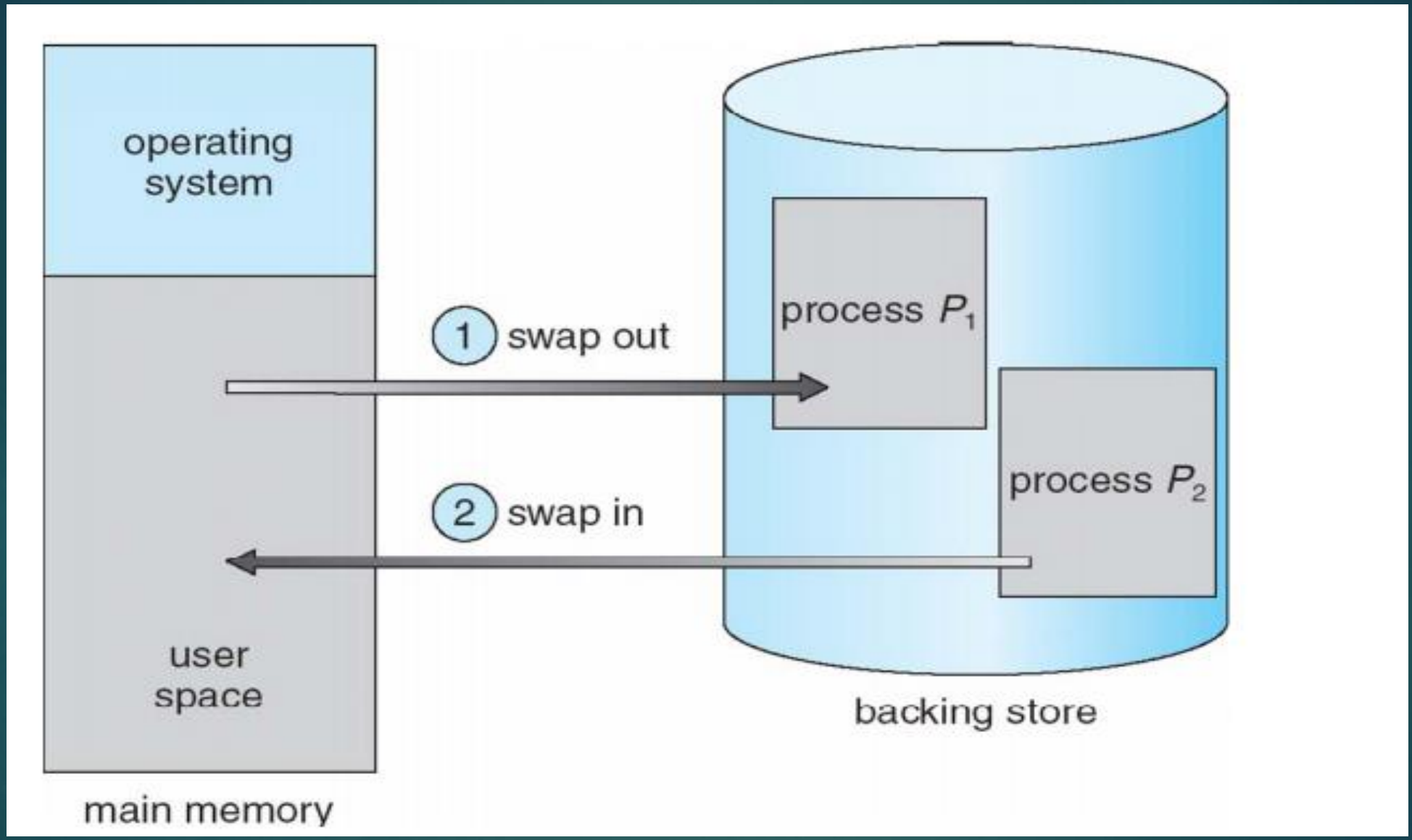


Figure: Information movement between the memory levels.1


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

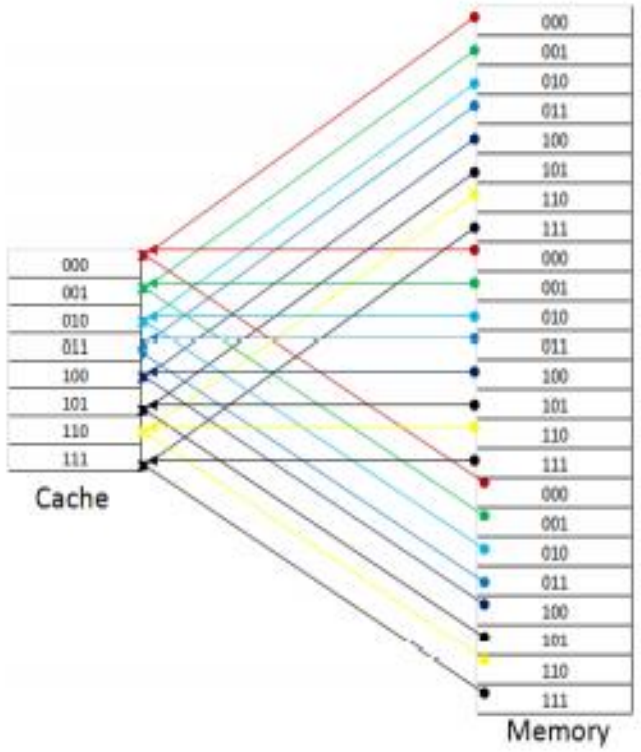
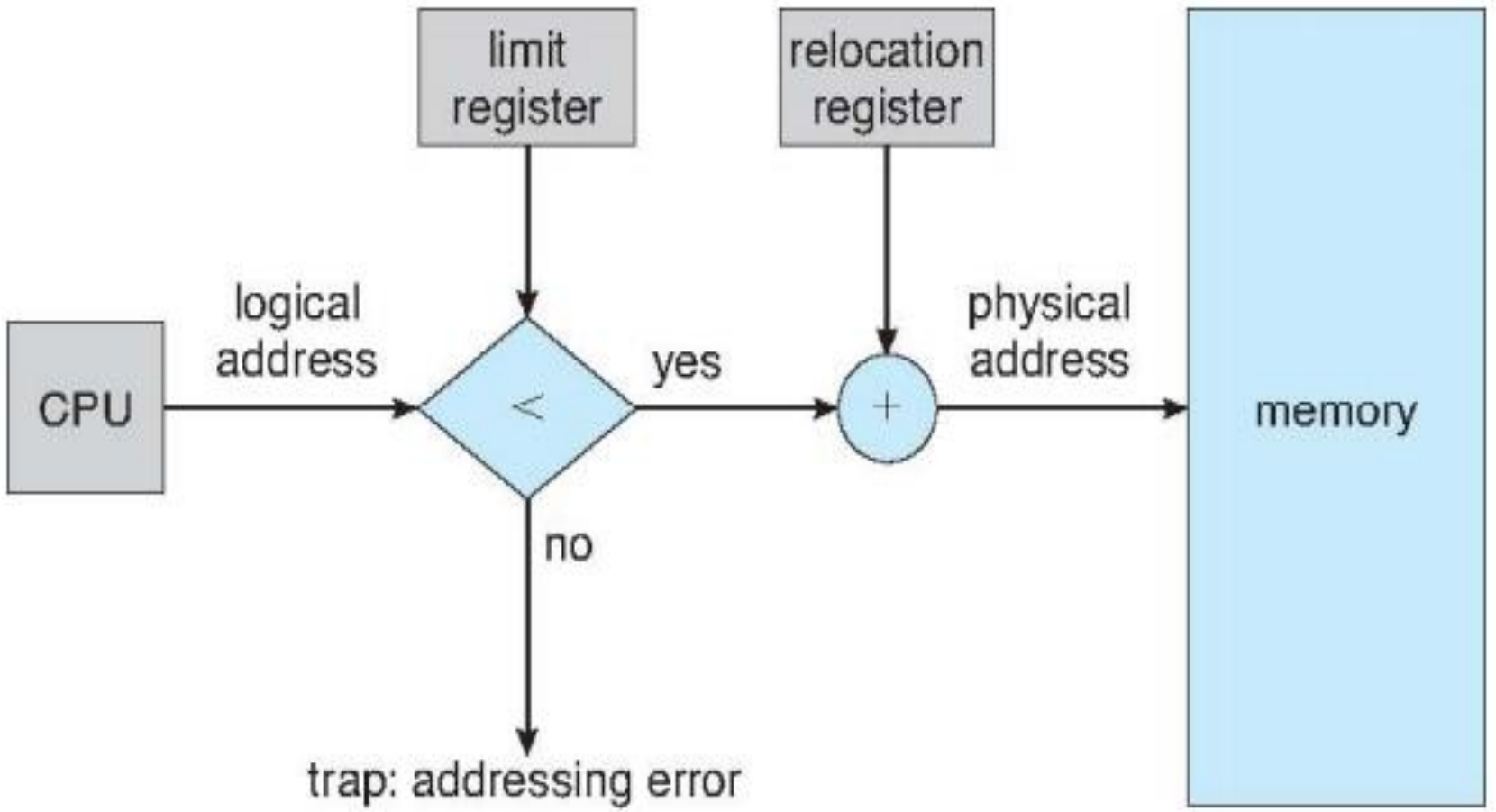


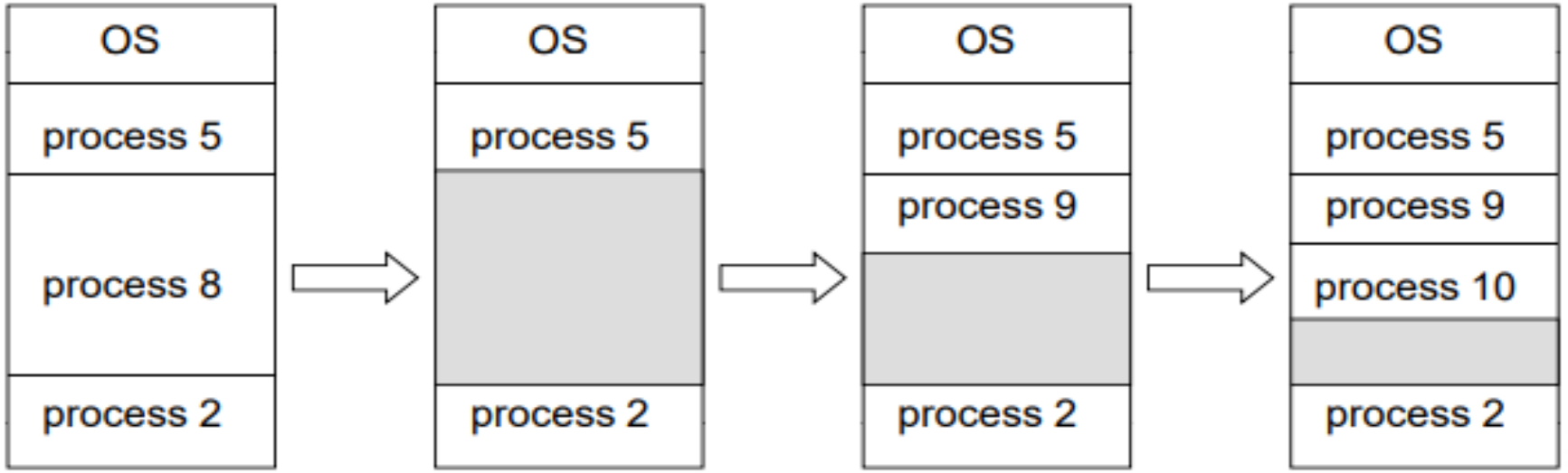
Figure: Memory to Cache : To a single location

Contiguous Allocation

- Main memory usually into two partitions:
  - Resident operating system, usually held in low memory with interrupt vector
  - User processes then held in high memory
- Relocation registers used to protect user processes from each other, and from changing operating-system code and data
  - Base register contains value of smallest physical address
  - Limit register contains range of logical addresses – each logical address must be less than the limit register
  - MMU maps logical address *dynamically*


limit register
relocation register
CPU
logical address
<
yes
+
physical address
memory
no
trap: addressing error

- Multiple-partition allocation
  - Hole – block of available memory; holes of various size are scattered throughout memory
  - When a process arrives, it is allocated memory from a hole large enough to accommodate it
  - Operating system maintains information about:
    a) allocated partitions b) free partitions (hole)

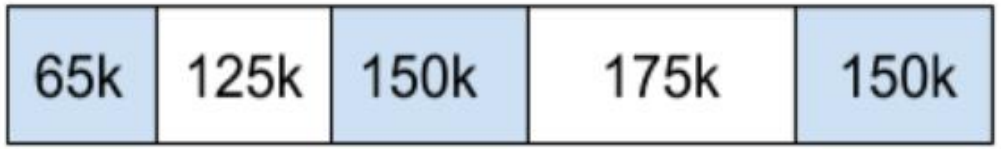


How to satisfy a request of size *n* from a list of free holes

- **First-fit**: Allocate the *first* hole that is big enough
- **Best-fit**: Allocate the *smallest* hole that is big enough; must search entire list, unless ordered by size
  - Produces the smallest leftover hole
- **Worst-fit**: Allocate the *largest* hole; must also search entire list
  - Produces the largest leftover hole

First-fit and best-fit better than worst-fit in terms of speed and storage utilization

Consider the following memory map using multiprogram with partition model. Blue represent memory in use while white represent free memory as shown in the figure below.



Request for memory follows the following order : 100k, 25k, 125k, 50k. Which of the following allocation satisfies the above request?
1)Best Fit 2)First Fit 3)Worst Fit

- 1,2,3
- 1,2
- 2,3
- None of these

Consider the following memory map using the partition model, with Blue representing memory in use and White representing free memory..

A new process Pnew is of size 10k. In which partition is Pnew placed for

a)Best Fit : _____k
b)First Fit : _____k
c)Worst Fit : _____k

- 10, 11, 11
- 10, 10, 11
- 11, 12, 40
- 11, 10, 48

Basic Terms

- **Page Frames or Frames:**

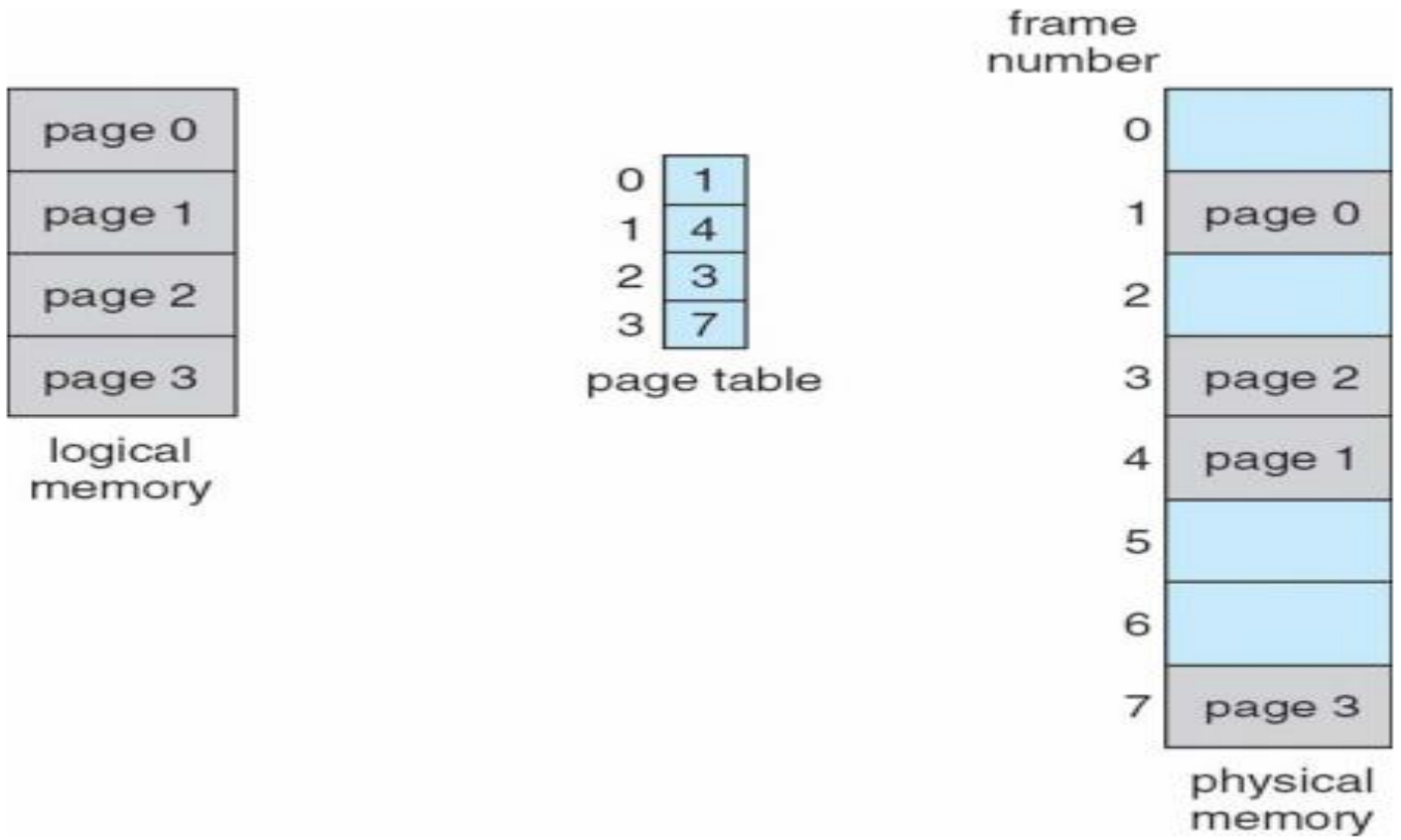
  Main memory is divided into fixed size of small blocks.

- **Pages:**

  Logical memory is divided into small and same size of blocks.

- **Swap Space:**

  The virtual machine usually creates the space on flash memory or disk for all the pages of a process when it creates the process. This space is called the swap space.


page 0
page 1
page 2
page 3
logical memory
0 1
1 4
2 3
3 7
page table
frame number
0
1 page 0
2
3 page 2
4 page 1
5
6
7 page 3
physical memory

Demand Paging Method

In the demand paging method, where processes are stored in secondary memory, and pages are loaded on demand only, not before.

When a context switch happens, after loading the first page, the computer system starts running the new program and fetches the pages of that program as they are referenced.

Advantages:

- Minimum I/O operations.
- Minimum Memory required.
- Quick response.
- Allow more user programs.

Basic Terms (Continued…)

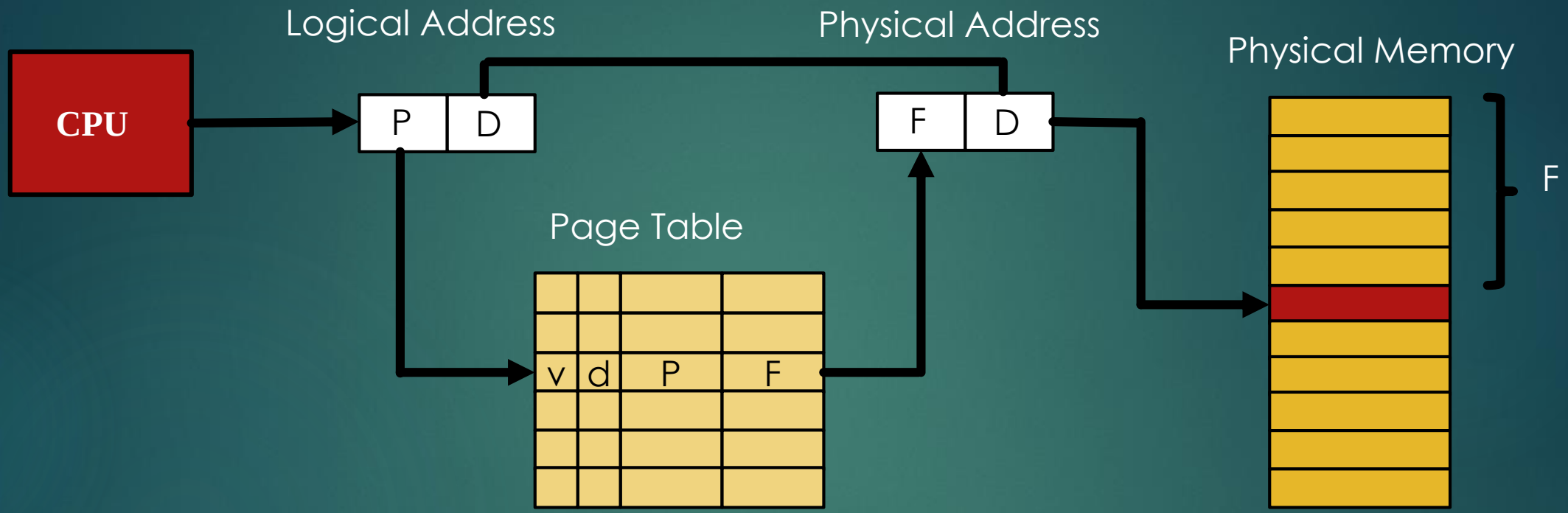
- **Dirty bit:**

  To track whether a page has been written since it was read into the memory, a dirty bit is added to the page table. The dirty bit is set when any word in a page is written. If the operating system chooses to replace the page, the dirty bit indicates whether the page needs to be written out before its location in memory can begiven to another page. Hence, a modified page is often called a dirty page.

- **Valid bit:**

  If the valid bit is on, the page table supplies the physical page number (i.e., the starting address of the page in memory) corresponding to the virtual page. If the valid bit is off , the page currently resides only on disk, at a specified disk address..

Memory Mapping



Memory Mapping (Continued…)

- CPU generated logical address, which consist of a page number “p” and a page offset “d”.
- Page table contains page number (is act as an index) and page frame number.
- Physical address is generated by combining page frame number “f” with page offset “d”.
- If logical address is m bits and page size is 2^n words then the logical address space is 2^m and (m-n) bits are assigned to page number.

Example Problem: 1

- 1) If virtual address is **32 bits** and size of page is **4kB** then find how many bits are required to define page number and page offset?

- 2) If a computer hardware has a page with **4096 bytes** and a process required **73,655 bytes** then find the number of frames need to store the process in memory?

- 3) If a computer system has a **48-bits** logical address, 8 bytes page table entry and **4 kB** pages, then find the size of page table.

Drawback:

- If page table is very large then implementation of fast registers for most modern computers are not feasible.

- With the use of Page Table Base Register (PTBR), we can reduce context switching time. Since this scheme needs 2 memory accesses, to overcome this issue a speedy lookup cache called Translation Lookaside Buffer (TLB) was proposed.