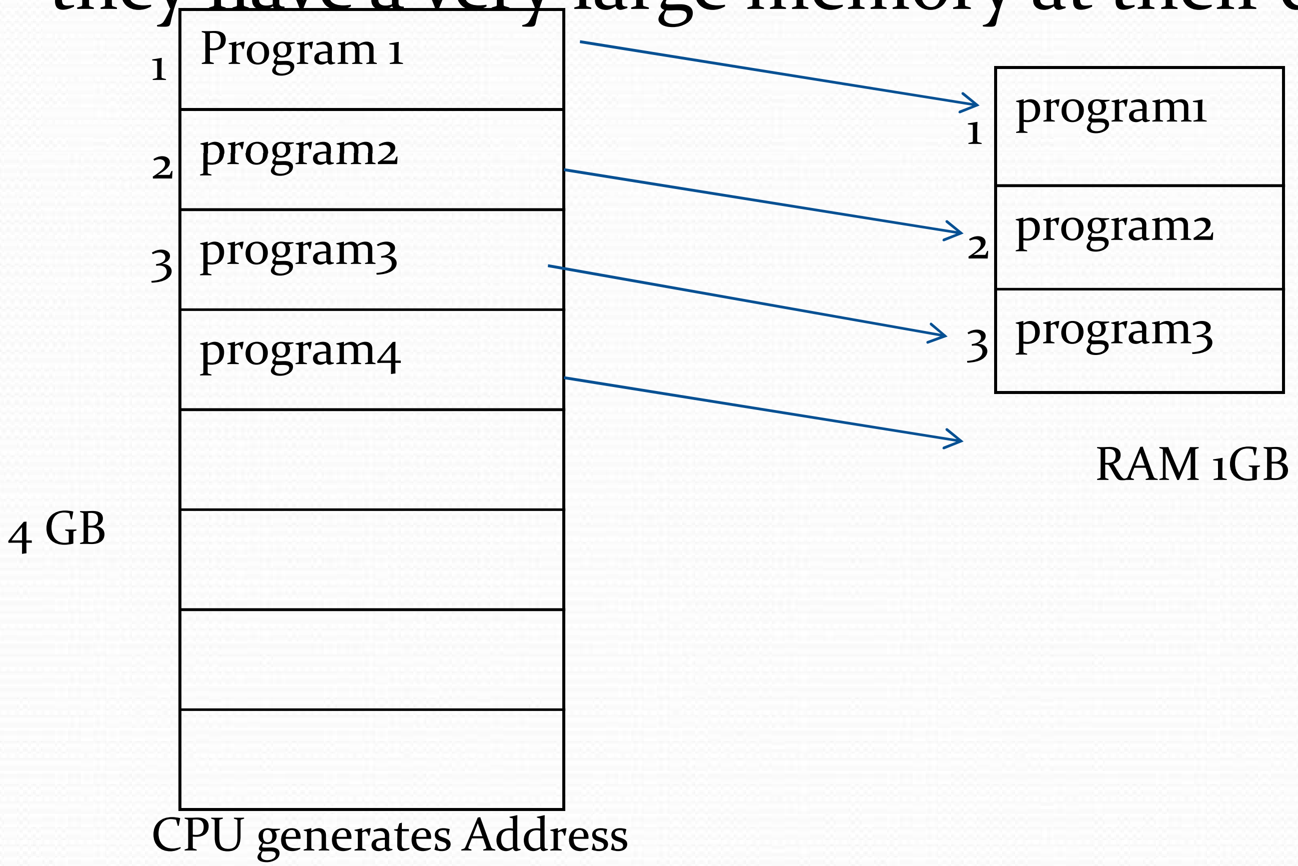
Virtual Memory

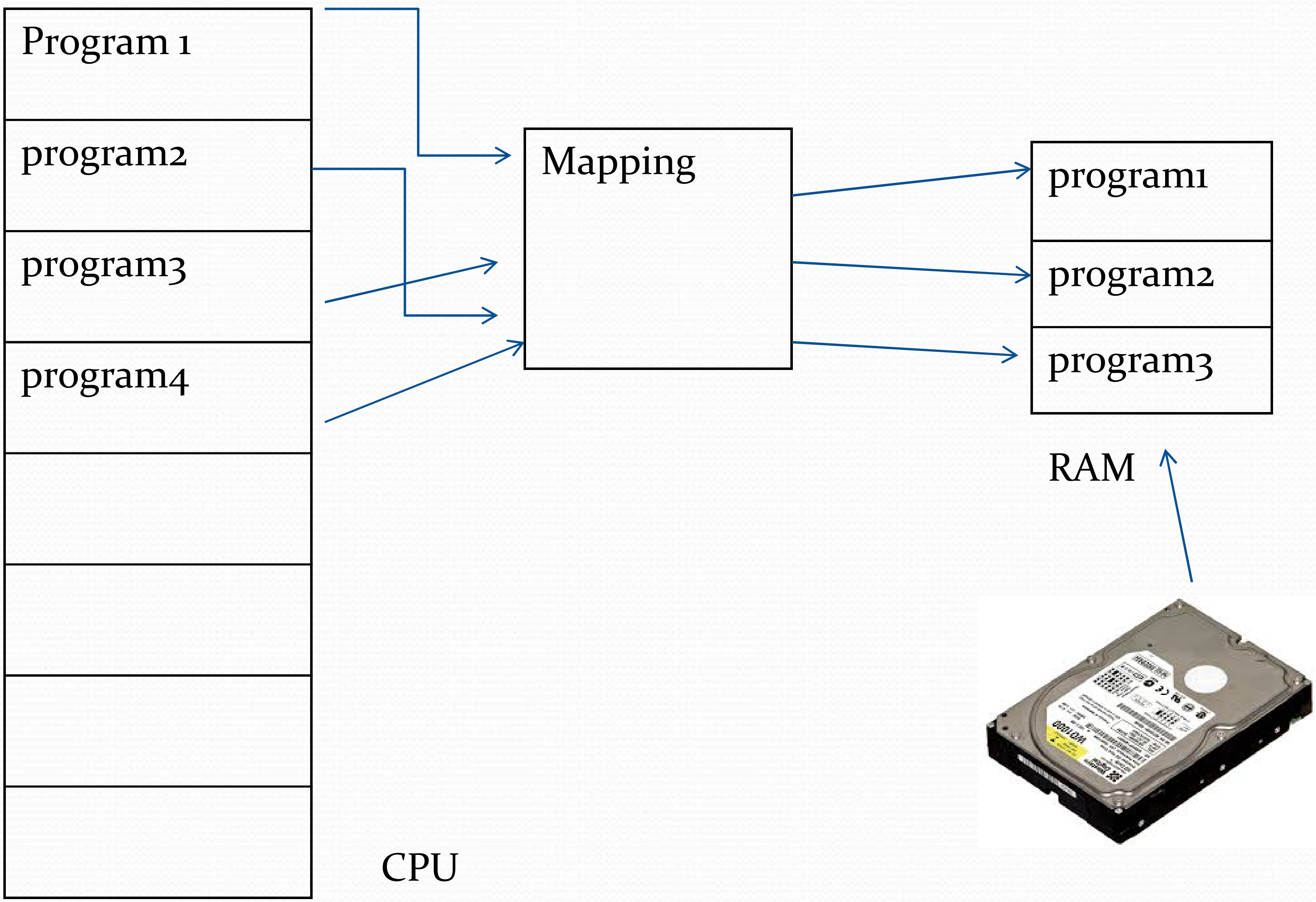
Jestin James M
Assistant Professor, Dept of Computer Science
Little Flower College, Guruvayoor

Virtual Memory

- It is used to give programmers the illusion that they have a very large memory at their disposal



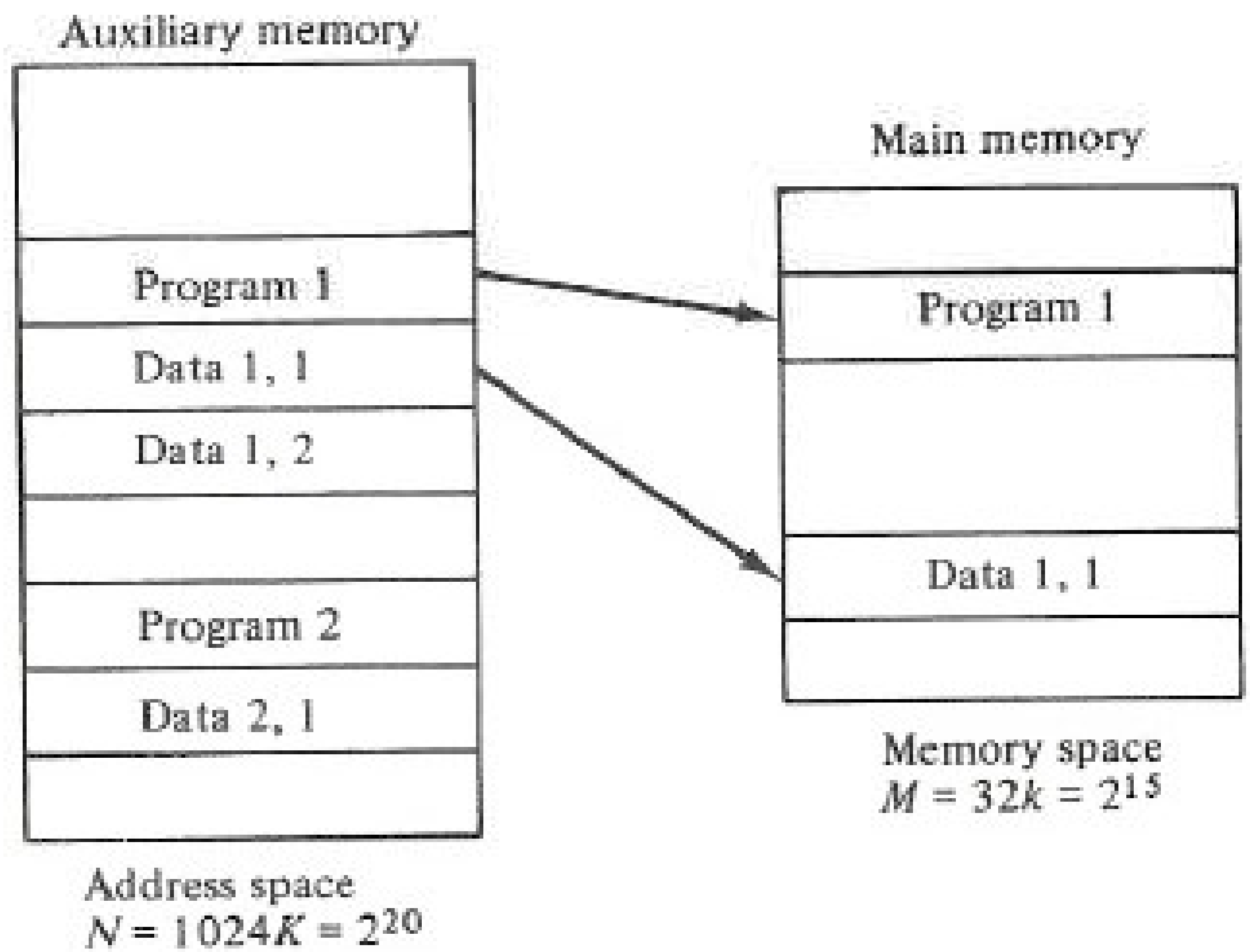
Virtual Memory



Address Space and Memory Space

- An address used by a programmer will be called a *virtual address,*
- Set of such addresses the *address space*
- An address in main memory is called a location or physical address .
- The set of such locations is called the memory space

Example

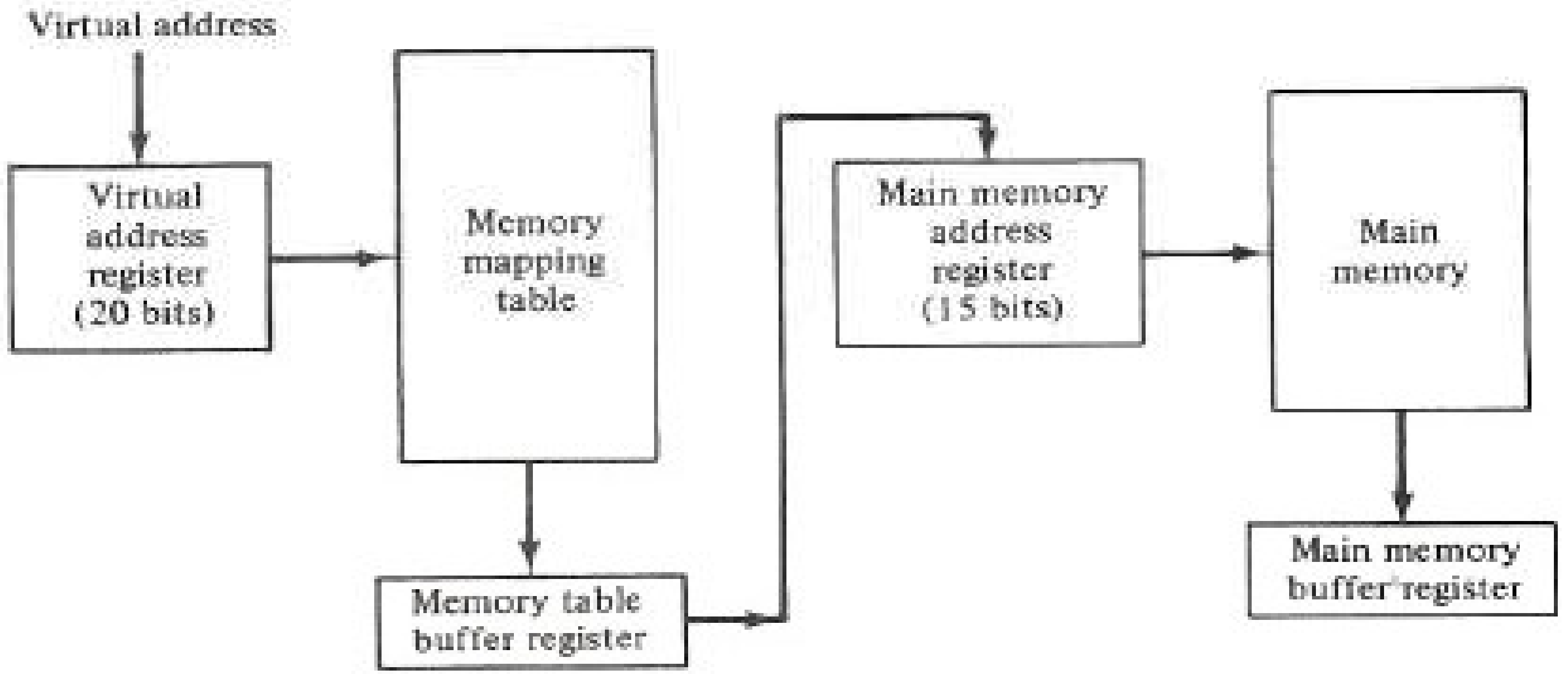


CPU generates 20 bit address

Mapping 20 bit to 15 bit

Memory only 15 bits

Virtual Memory



Page table in Separate Memory or in memory

Virtual Memory

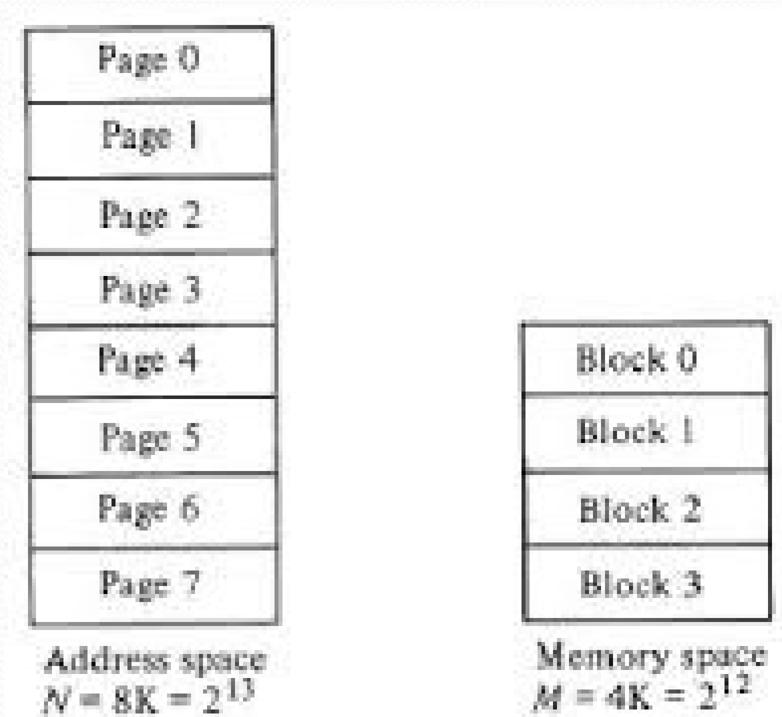
- In separate memory one extra memory access time
- In main memory table takes space from main memory. Two access to memory required
- Third associative memory

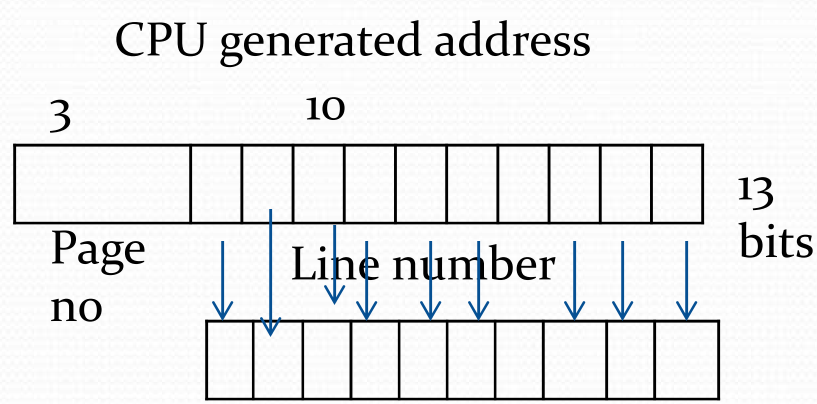
Address Mapping Using Pages

- The physical memory is broken down into groups of equal size called blocks,
- Page frame
- The term page refers to groups of address space of the same size.
- A page and block consists of 1k words
- 1024 pages and 32 blocks

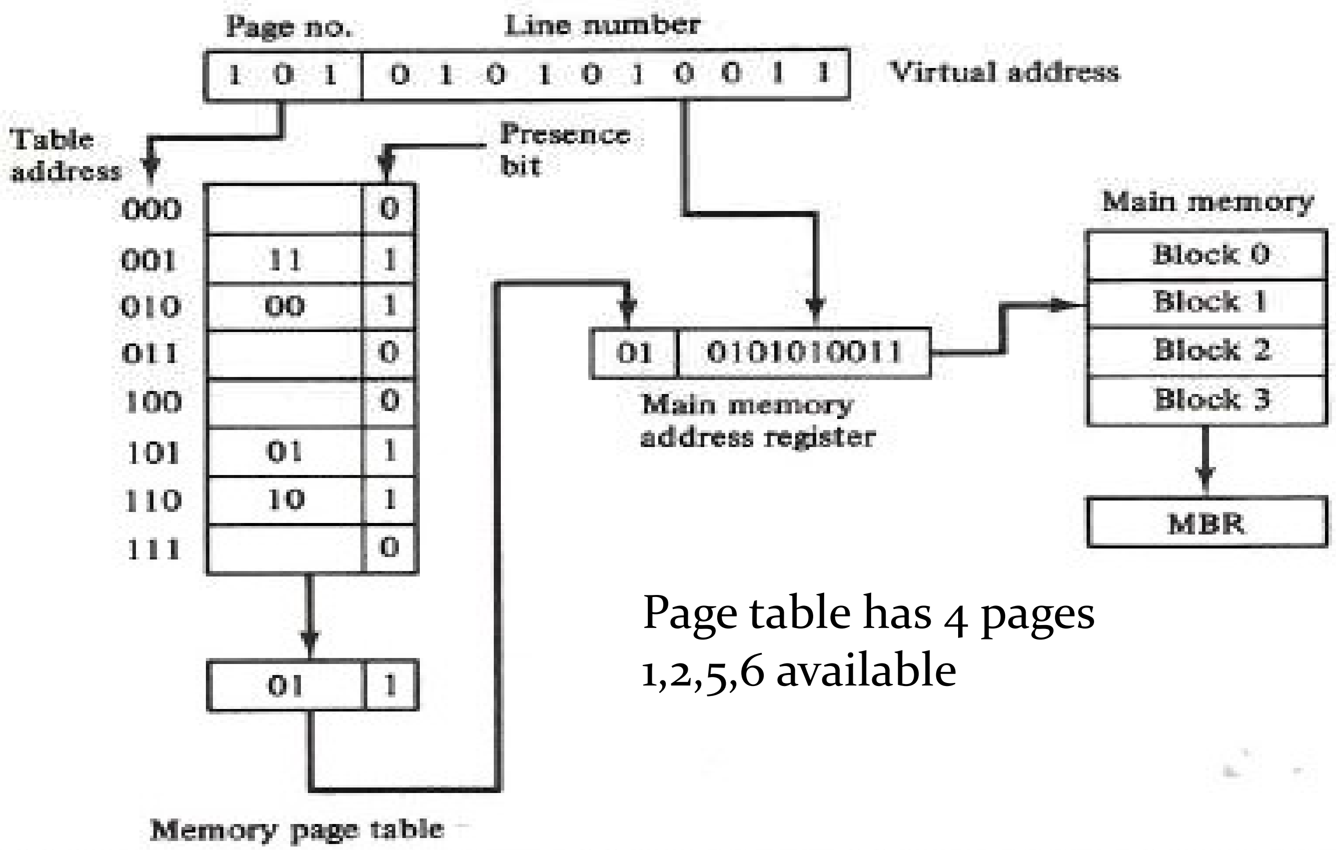
Example

computer with an address space of 8K and a memory space of 4K.





Example continue



Page table has 4 pages
1,2,5,6 available

The memory-page table consists of eight words, one for each page

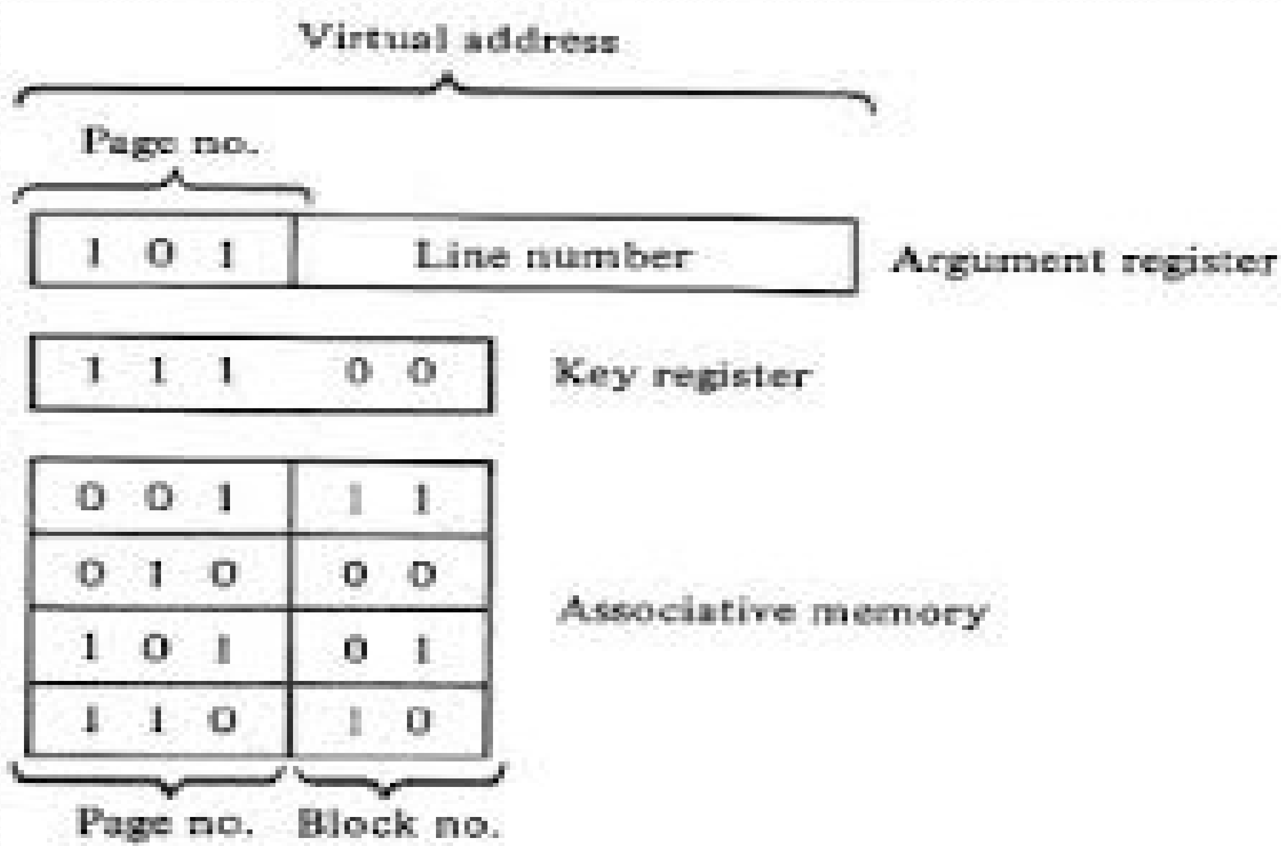
Content denotes block number

Associative memory Page table

- At any time only 4 page full . 4 page empty or not in use
- to organize the page table would be to construct it with a number of words equal to the number of blocks in main memory
- This method can be implemented by associative memory

Associative memory Page table

Word=page number +block number



Page Replacement

- A virtual memory system is a combination of hardware and software techniques.
- Software must decide

(1)which page in main memory ought to be removed to make room for a new page

(2) when a new page is to be transferred from auxiliary memory to main memory

(3) where the page is to be placed in main memory

Page Replacement

- The program is executed from main memory until it attempts to reference a page that is still in auxiliary memory. -page fault.
- loading a page from auxiliary memory to main memory is basically an I/O operation, the operating system assigns this task to the I/O processor
- In the meantime, control is transferred to the next program in memory that is waiting to be processed in the CPU

Page Replacement

- When a page fault occurs in a virtual memory system and
- If main memory is full, it would be necessary to remove a page from a memory block to make room for the new page.
- The policy for choosing pages to remove is determined from the replacement algorithm that is used.

Page Replacement

- First-In, First-Out (FIFO):
- algorithm selects for replacement the page that has been in memory the longest time

Page Replacement

- Least Recently Used (LRU): The LRU algorithm can be implemented by associating a counter with every page that is in main memory. When a page is referenced, its associated counter is set to zero.
- At fixed intervals of time, the counters associated with all pages presently in memory are incremented by 1

Page Replacement

- The least recently used page is the page with the highest count. The counters are often called aging registers

Memory Management Hardware

- The basic components are

1. A facility for dynamic storage relocation that maps logical memory references into physical memory addresses
2. A provision for sharing common programs stored in memory by different users
3. Protection of information against unauthorized access between users and preventing users from changing operating system functions

Memory Management Hardware

- A segment is a set of logically related instructions or data elements associated with a given name.
- Segments may be generated by the programmer or by the operating system.
- Examples of segments are a subroutine, an array of data, a table of symbols, or a user's program.

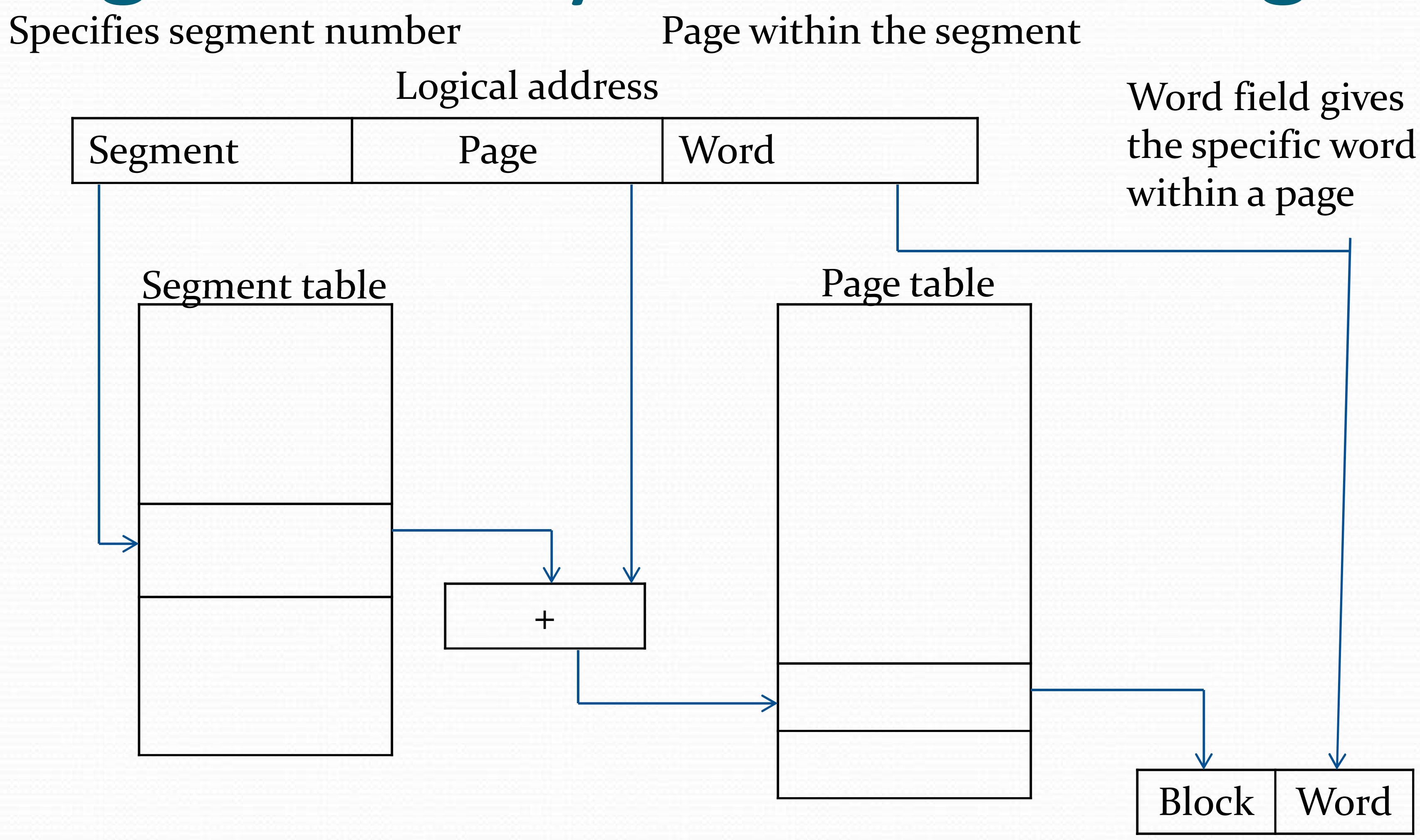
Memory Management Hardware

- The address generated by a segmented program is called a logical address
- logical address space is associated with variable-length segments rather than fixed-length pages
- Logical address may be larger ,equal or smaller than length of physical memory

Segmented-Page Mapping

- The length of each segment is allowed to grow and contract according to the needs of the program being executed
- specifying the length of a segment is by associating with it a number of equal-size pages.

Logical to Physical addressing



How it works

- The segment number of the logical address specifies the address for the segment table.
- The entry in the segment table is a pointer address for a page table base
- The page table base is added to the page number given in the logical address.
- sum produces a pointer address to an entry in the page table
- The value found in the page table provides the block number in physical memory.

How it works

- The concatenation of the block field with the word field produces the final physical mapped address.
- The two mapping tables may be stored in two separate small memories or in main memory.

How it works

- In either case, a memory reference from the CPU will require three accesses to memory:
- one from the segment table, one from the page table, and the third from main memory.
- This will slow the system

How it works

- To avoid this speed penalty, a fast associative memory is used to hold the most recently referenced table entries.
- This type of memory is sometimes called a translation lookaside buffer, abbreviated ***TLB.***

How TLB works

- The first time a given block is referenced, its value + the corresponding segment + page numbers are entered into the associative memory

Segment	Page	Block

Numerical Example

4	8	8
Segment	Page	Word

The 4-bit segment number specifies one of 16 possible segments.

The 8-bit page number can specify up to 256 pages,

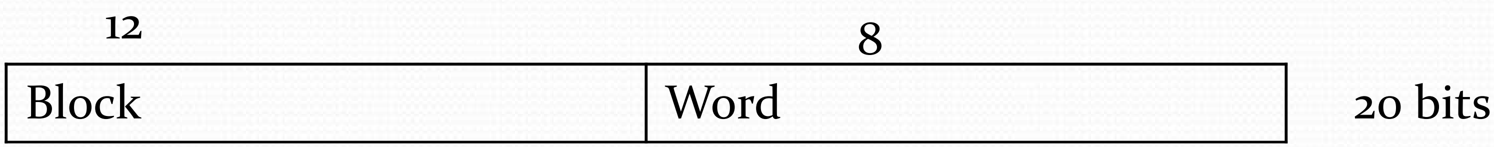
The 8-bit word field implies a page size of 256 words.

Numerical Example

- This configuration allows each segment to have any number of pages up to 256.
- The smallest possible segment will have one page or 256 words
- The largest possible segment will have 256 pages, for a total of 256 x 256 = 64K words.

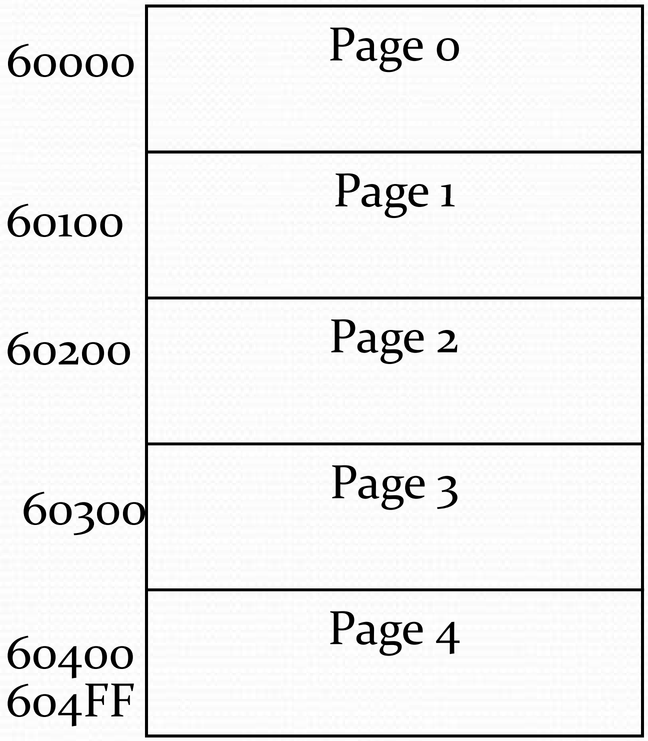
Numerical Example

- The 20-bit address is divided into two fields: a 12-bit block number and an 8-bit word number.
- Thus, physical memory is divided into 4096 blocks of 256 words each



Numerical Example

Program requires 5 pages. OS assign segment 6

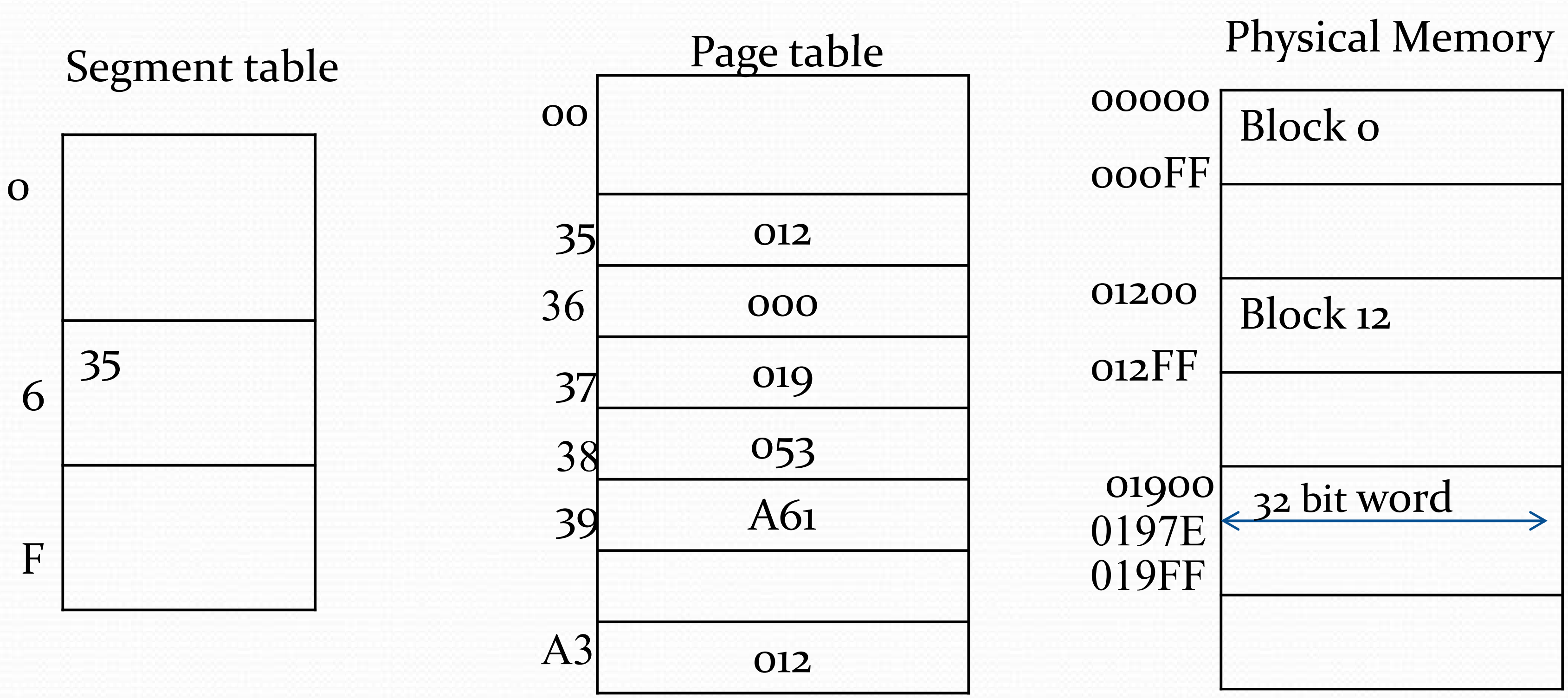


Segment	Page	Block
6	00	012
6	01	000
6	02	019
6	03	053
6	04	A61

Numerical Example

Logical Address in Hexa Decimal

6	02	7E



Segment	Page	Block
6	02	019
6	04	A61