Applied Operating System

Chapter 3: Memory Management

Prepared By:
Amit K. Shrivastava
Asst. Professor
Nepal College Of Information Technology

Concept

. Memory management is the functionality of an operating system which handles or manages primary memory. Memory management keeps track of each and every memory location either it is allocated to some process or it is free. It checks how much memory is to be allocated to processes. It decides which process will get memory at what time. It tracks whenever some memory gets freed or unallocated and correspondingly it updates the status.

- Address binding of instructions and data to memory addresses can happen at three different stages
  - •Compile time -- When it is known at compile time where the process will reside, compile time binding is used to generate the absolute code.
  - •Load time -- When it is not known at compile time where the process will reside in memory, then the compiler generates re-locatable code.
  - •Execution time -- If the process can be moved during its execution from one memory segment to another, then binding must be delayed to be done at run time

Logical vs. Physical Address Space

- The concept of a logical address space that is bound to a separate Physical address space is central to proper memory management
  - Logical address – generated by the CPU; also referred to as virtual address
  - Physical address – address seen by the memory unit
- Logical and physical addresses are the same in compile-time and load-Time address-binding schemes; logical (virtual) and physical addresses differ in execution-time address-binding scheme.
- The run-time mapping from virtual to physical address is done by the memory management unit (MMU) which is a hardware device. MMU uses following mechanism to convert virtual address to physical address.
  - The value in the base register is added to every address generated by a user process which is treated as offset at the time it is sent to memory. For example, if the base register value is 10000, then an attempt by the user to use address location 100 will be dynamically reallocated to location 10100.
  - The user program deals with virtual addresses; it never sees the real physical addresses.

Swapping

- A process can be swapped temporarily out of memory to a backing store, and then brought back into memory for continued execution.
- Backing store – fast disk large enough to accommodate copies of all memory images for all users; must provide direct access to these memory images.
- Roll out, roll in – swapping variant used for priority-based scheduling algorithms; lower-priority process is swapped out so higher-priority process can be loaded and executed.
- Major part of swap time is transfer time; total transfer time is directly proportional to the amount of memory swapped. Let us assume that the user process is of size 100KB and the backing store is a standard hard disk with transfer rate of 1 MB per second. The actual transfer of the 100K process to or from memory will take
- 100KB / 1000KB per second = 1/10 second = 100 milliseconds

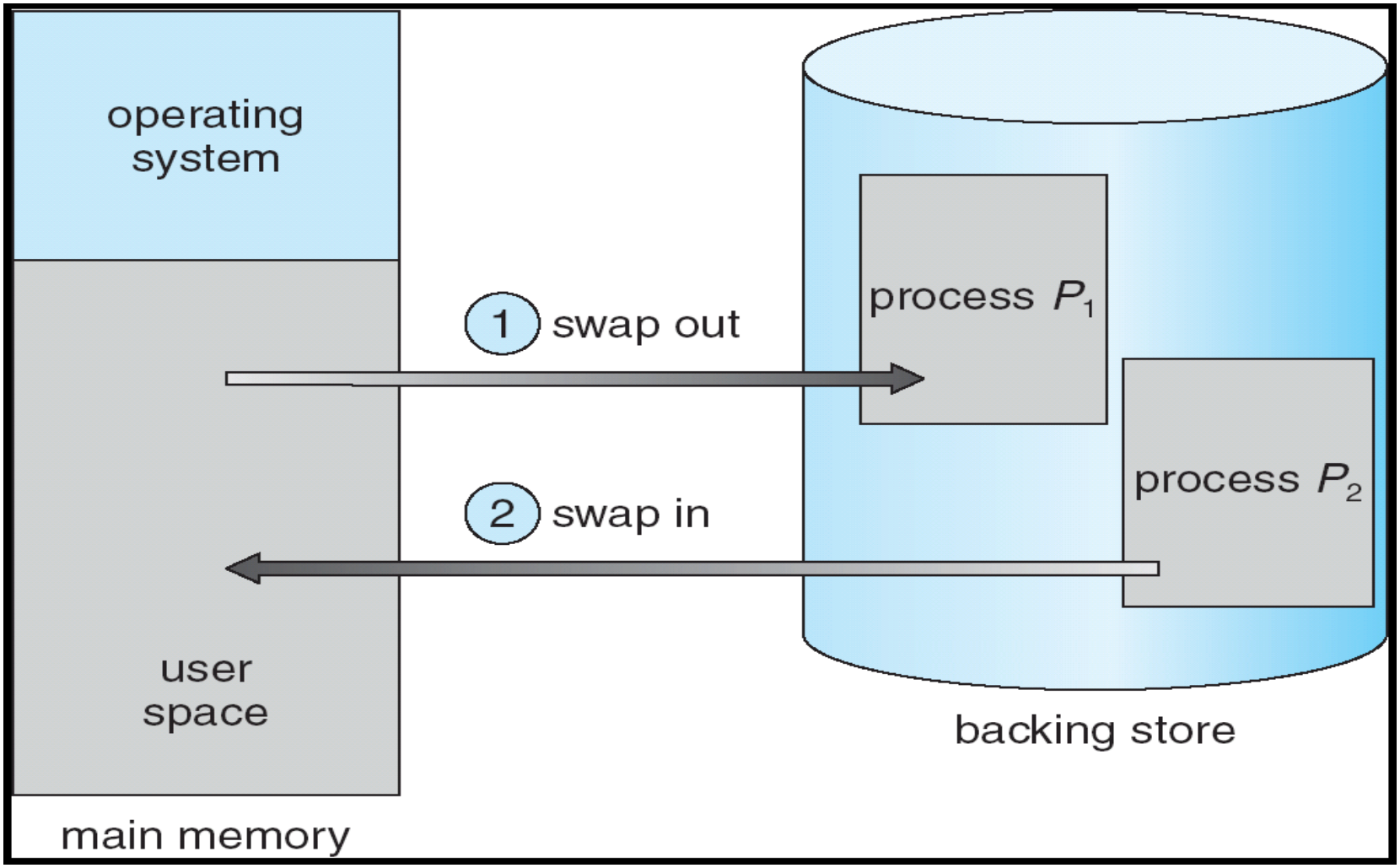


Fig: Schematic View of Swapping

Storage Hierarchy:

In 1960 storage hierarchy is extended by one more level, i.e the cache which is a high-speed storage that is much faster than main storage

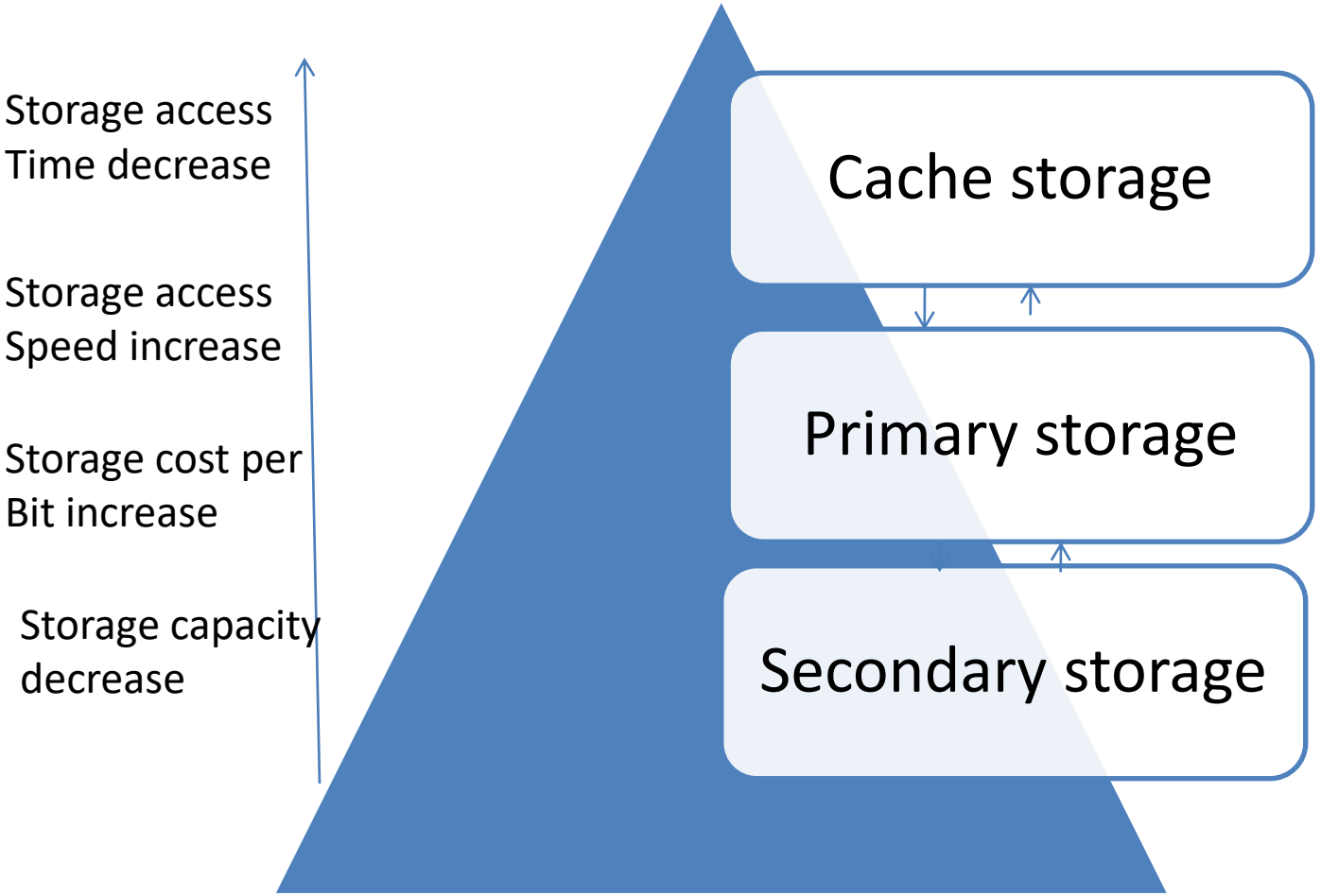


Fig: Hierarchical storage organization

Storage Management Strategies:

- Storage Management strategies are generated for obtaining the best possible use of the main storage. They are divided into following categories:
- 1. Fetch Strategies
  - a) Demand
  - b) Anticipatory

  2. Placement Strategies

  3. Replacement Strategies

- Fetch Strategies are concerned with when to obtain the next piece of program or data for transfer to main storage from secondary storage.
- Placement Strategies are concerned with determining where in main storage to place an incoming program. E.g. first-fit, best-fit, and worst-fit
- Replacement Strategies are concerned with determining which piece of program or data to displace to make room for incoming programs.

Contiguous Memory Allocation

- In this type of allocation, relocation-register scheme is used to protect user processes from each other, and from changing operating-system code and data. Relocation register contains value of smallest physical address whereas limit register contains range of logical addresses. Each logical address must be less than the limit register.
- **Multiple-Partition allocation:** In this type of allocation, main memory is divided into a number of fixed-sized partitions where each partition should contain only one process. When a partition is free, a process is selected from the input queue and is loaded into the free partition. When the process terminates, the partition becomes available for another process.

Fragmentation:

•**External Fragmentation –** total memory space exists **to** satisfy a request, but it is not contiguous.

• **Internal Fragmentation –** allocated memory may be slightly larger than requested memory; this size difference is memory internal to a partition, but not being used.

• Reduce external fragmentation by compaction

- ✦ Shuffle memory contents to place all free memory together in one large block.
- ✦ Compaction is possible *only if relocation is dynamic, and is* done at execution time.
- ✦ I/O problem
  - ✔ Latch job in memory while it is involved in I/O.
  - ✔ Do I/O only into OS buffers.

Fixed Partition Multiprogramming

In a multiprogramming environment, several programs reside in primary memory at a time and the CPU passes its control rapidly between these programs. One way to support multiprogramming is to divide the main memory into several partitions each of which is allocated to a single process. Depending upon how and when partitions are created, there may be two types of memory partitioning: (1) Static(fixed) and (2) Dynamic(variable). Static partitioning implies that the division of memory into number of partitions and its size is made in the beginning (during the system generation process) and remain fixed thereafter. The basic approach here is to divide memory into several fixed size partitions where each partition will accommodate only one program for execution. The number of programs (i.e. degree of multiprogramming) residing in memory will be bound by the number of partition When a program terminates, that partition is free for another program waiting in a queue.

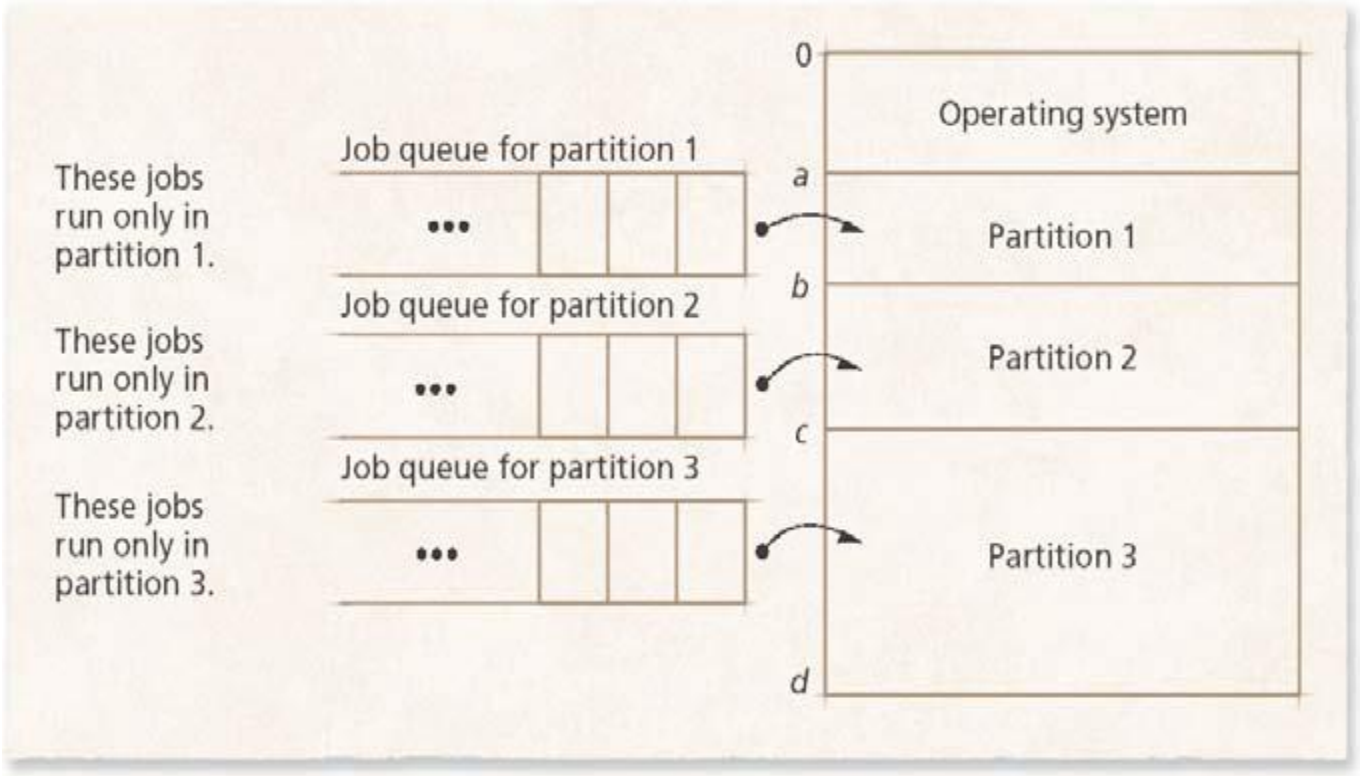


Figure : Fixed-partition multiprogramming with absolute translation and loading.

Drawbacks to fixed partitions

- Early implementations used absolute addresses
  - If the requested partition was full, code could not load
  - Later resolved by relocating compilers

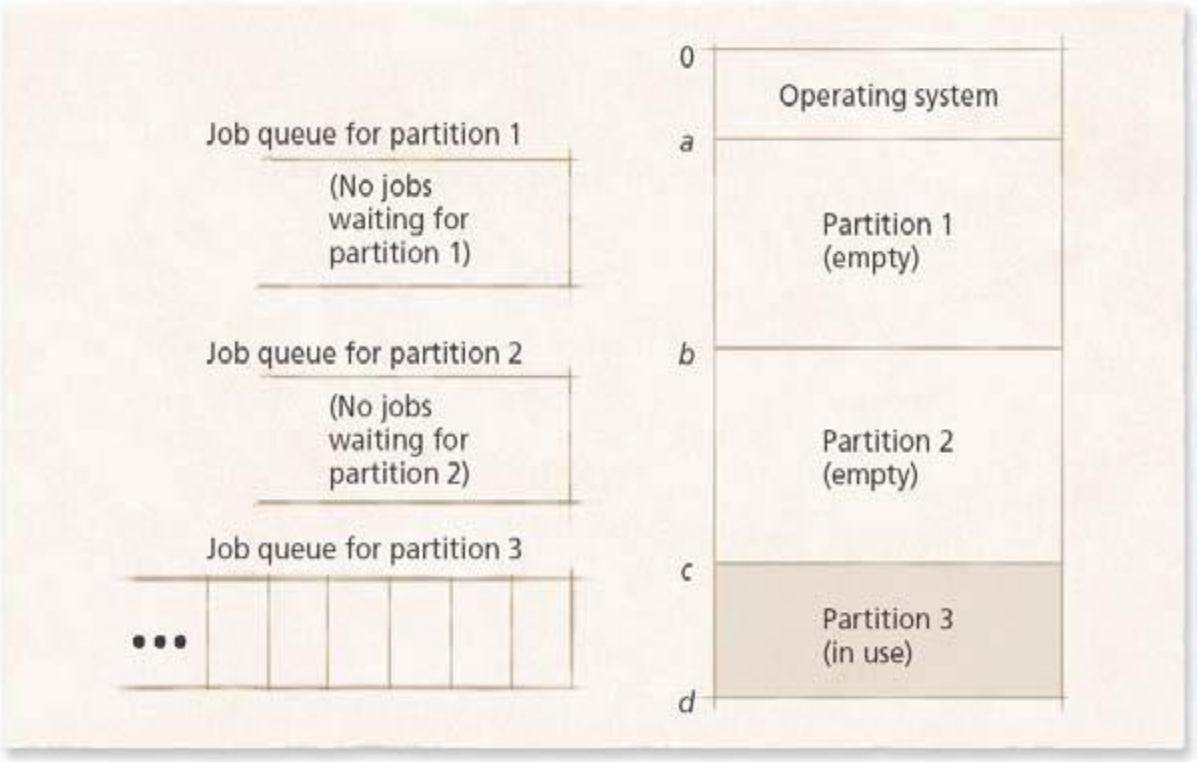


Figure : Memory waste under fixed-partition multiprogramming with absolut translation and loading.

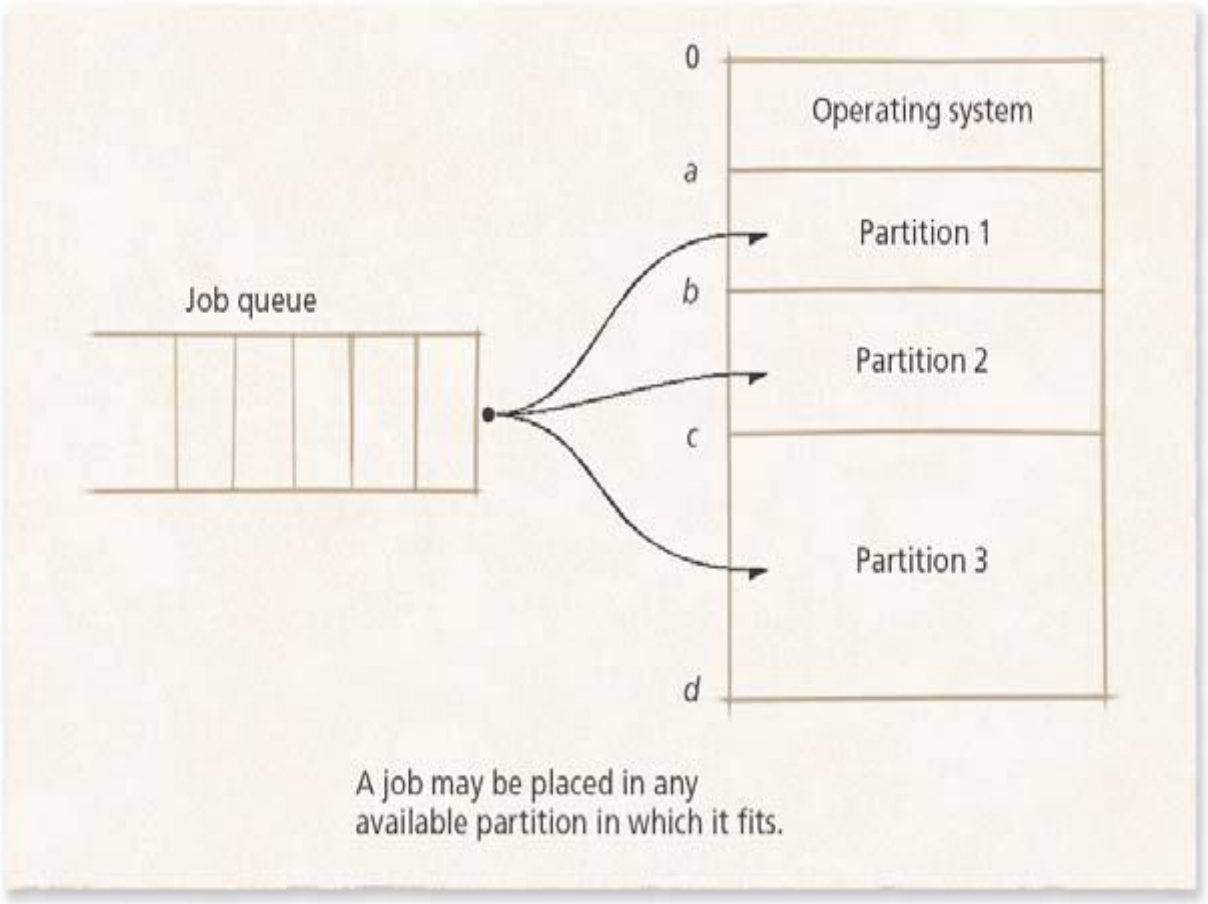


Figure : Fixed-partition multiprogramming with relocatable translation and load

Variable-Partition Multiprogramming

System designers found fixed partitions too restrictive

- Internal fragmentation
- Potential for processes to be too big to fit anywhere
- Variable partitions designed as replacement

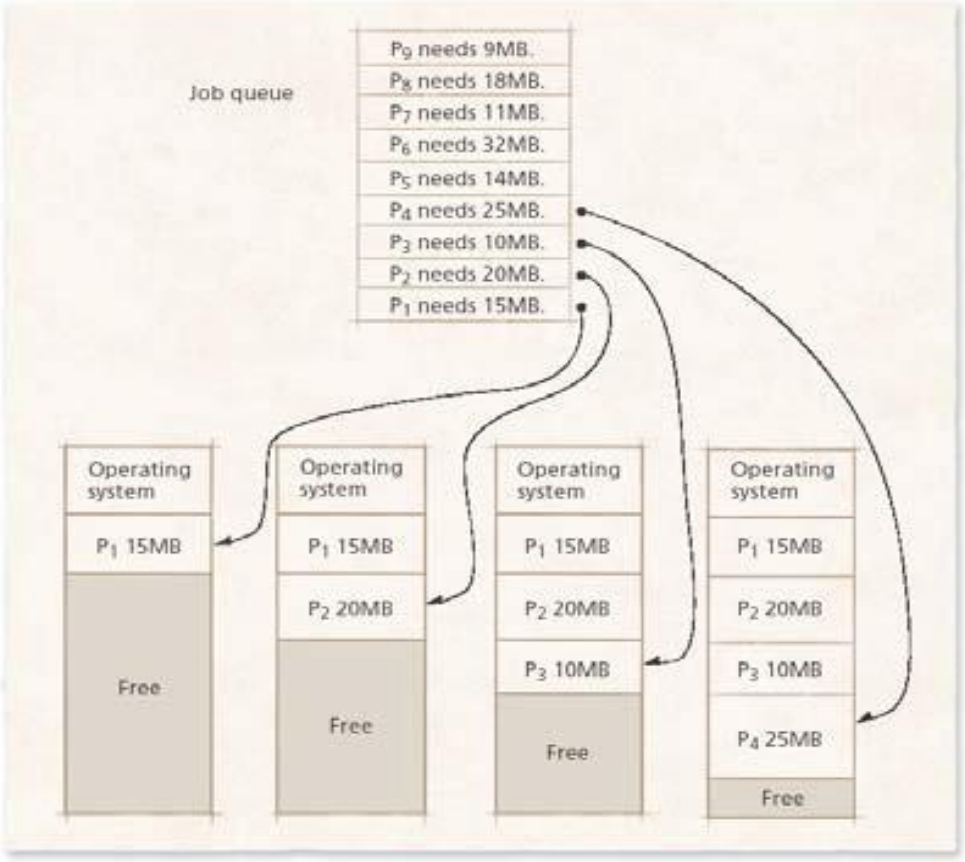


Figure : Initial partition assignments in variable-partition programming.

Variable-Partition Characteristics

Jobs placed where they fit

- No space wasted initially
- Internal fragmentation impossible
  - Partitions are exactly the size they need to be
- External fragmentation can occur when processes removed
  - Leave holes too small for new processes
  - Eventually no holes large enough for new processes

Variable-Partition Characteristics

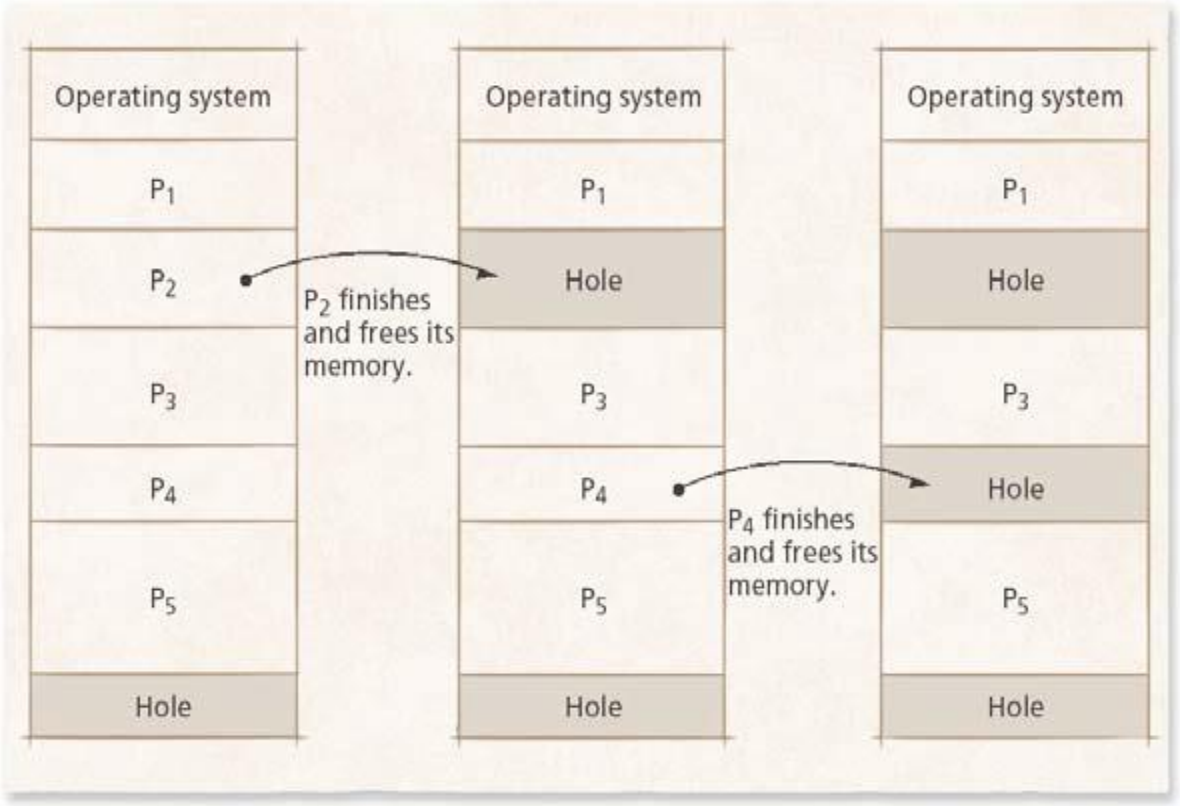


Figure : Memory “holes” in variable-partition multiprogramming.

Variable-Partition Characteristics

- Several ways to combat external fragmentation
  - Coalescing
    - Combine adjacent free blocks into one large block
    - Often not enough to reclaim significant amount of memory
  - Compaction
    - Sometimes called garbage collection (not to be confused with GC in object-oriented languages)
    - Rearranges memory into a single contiguous block free space and a single contiguous block of occupied space
    - Makes all free space available
    - Significant overhead

Variable-Partition Characteristics

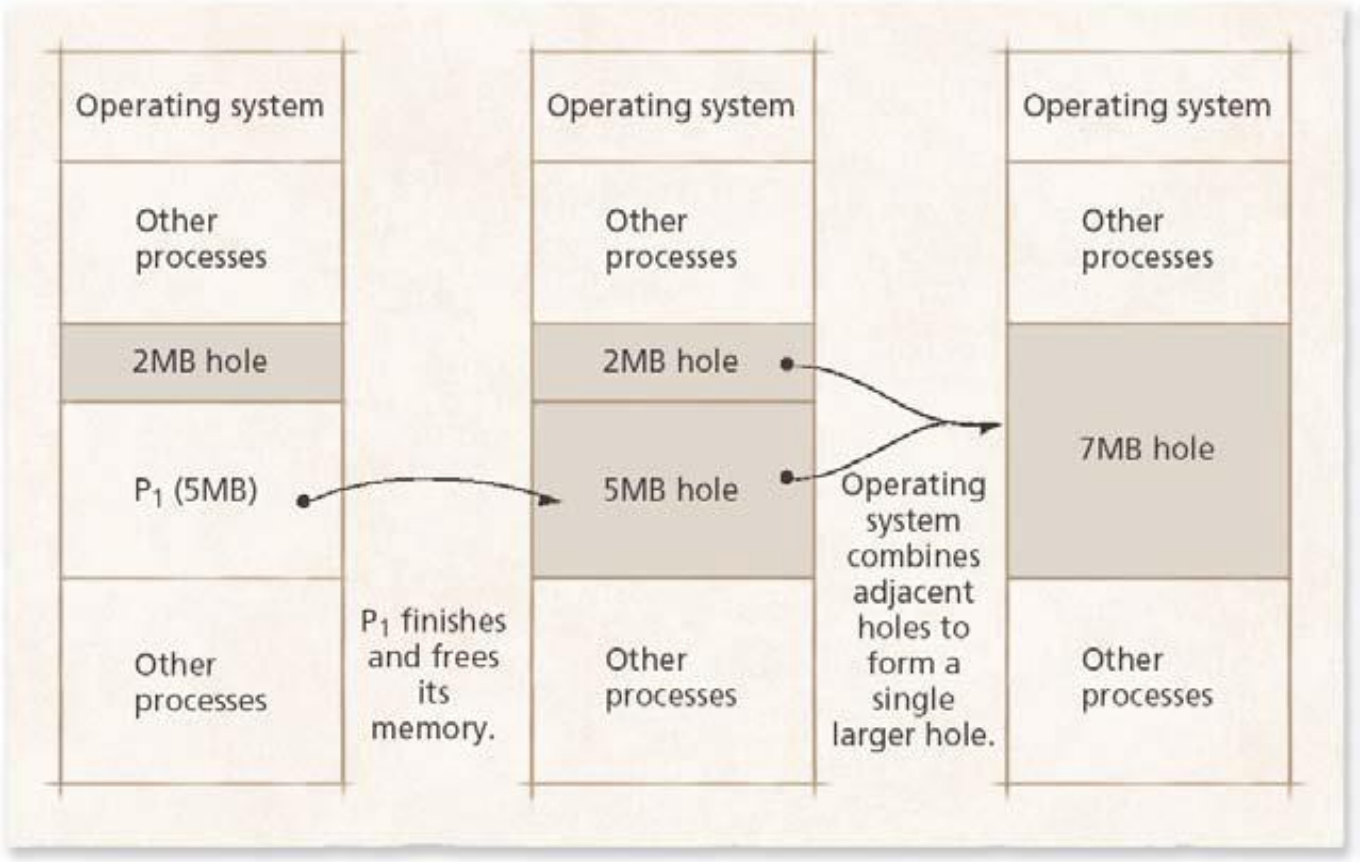


Figure : Coalescing memory “holes” in variable-partition multiprogramming

Variable-Partition Characteristics

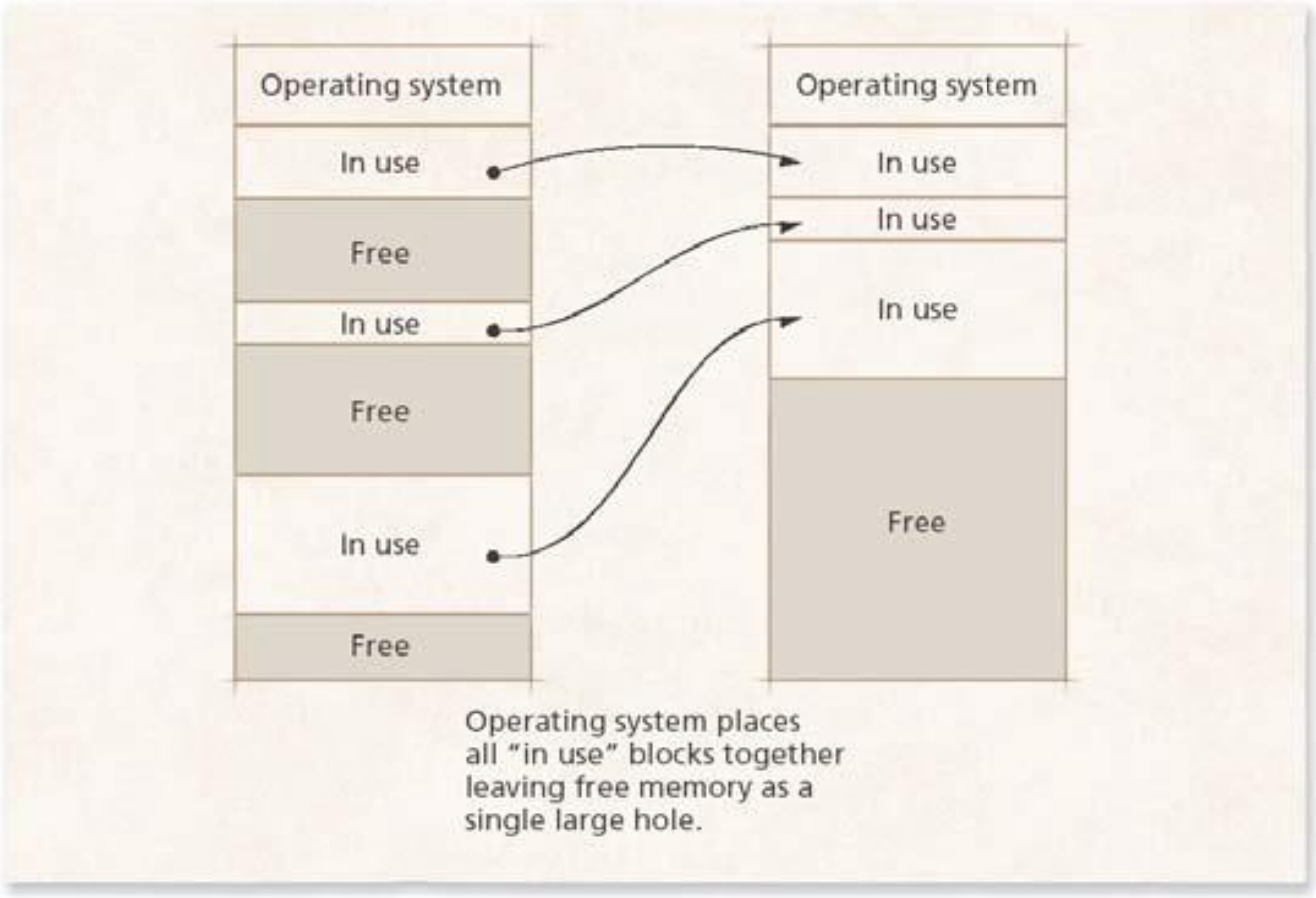


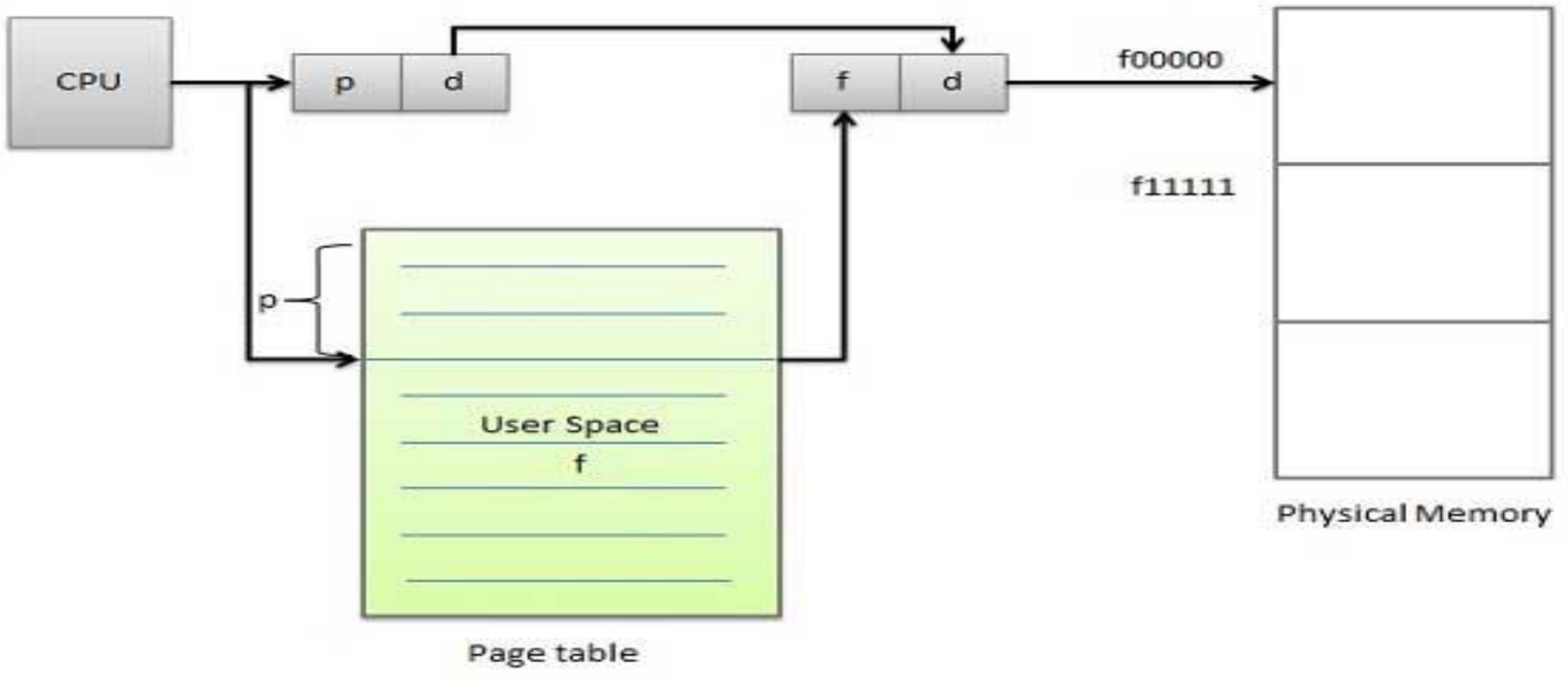
Figure : Memory compaction in variable-partition multiprogramming

Paging:

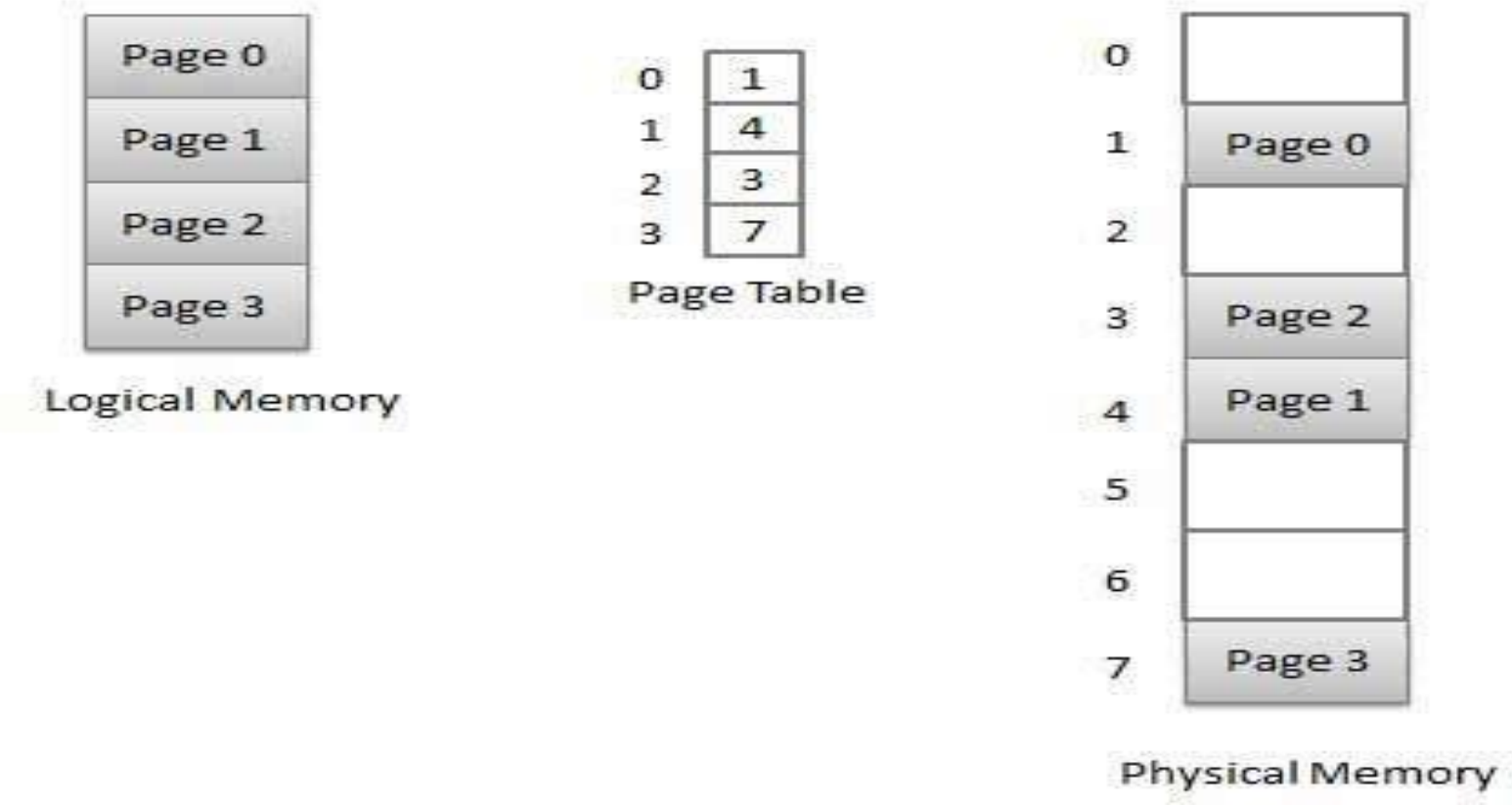
•Logical address space of a process can be noncontiguous; process is allocated physical memory whenever the latter is available.
• Divide physical memory into fixed-sized blocks called **frames**
•(size is power of 2, between 512 bytes and 8192 bytes).
• Divide logical memory into blocks of same size called **pages.**
• Keep track of all free frames.
• To run a program of size *n pages, need to find n free frames* and load program.
• Set up a page table to translate logical to physical addresses.
• Internal fragmentation

Paging(contd..)

- Address generated by CPU is divided into
  - Page number (p) -- page number is used as an index into a page table which contains base address of each page in physical memory.
  - Page offset (d) -- page offset is combined with base address to define the physical memory address.



Following figure show the paging table architecture



Page Table Structure

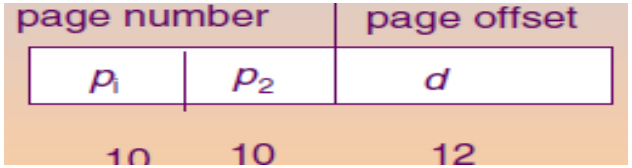
- Hierarchical Paging
- Hashed Page Tables
- Inverted Page Tables

Hierarchical Page Tables

- Break up the logical address space into multiple page tables.
- A simple technique is a two-level page table.

Two-Level Paging Example

- A logical address (on 32-bit machine with 4K page size) is divided into:
  - ✦ a page number consisting of 20 bits.
  - ✦ a page offset consisting of 12 bits.
- Since the page table is paged, the page number is further divided into:
  - ✦ a 10-bit page number.
  - ✦ a 10-bit page offset.
- Thus, a logical address is as follows:

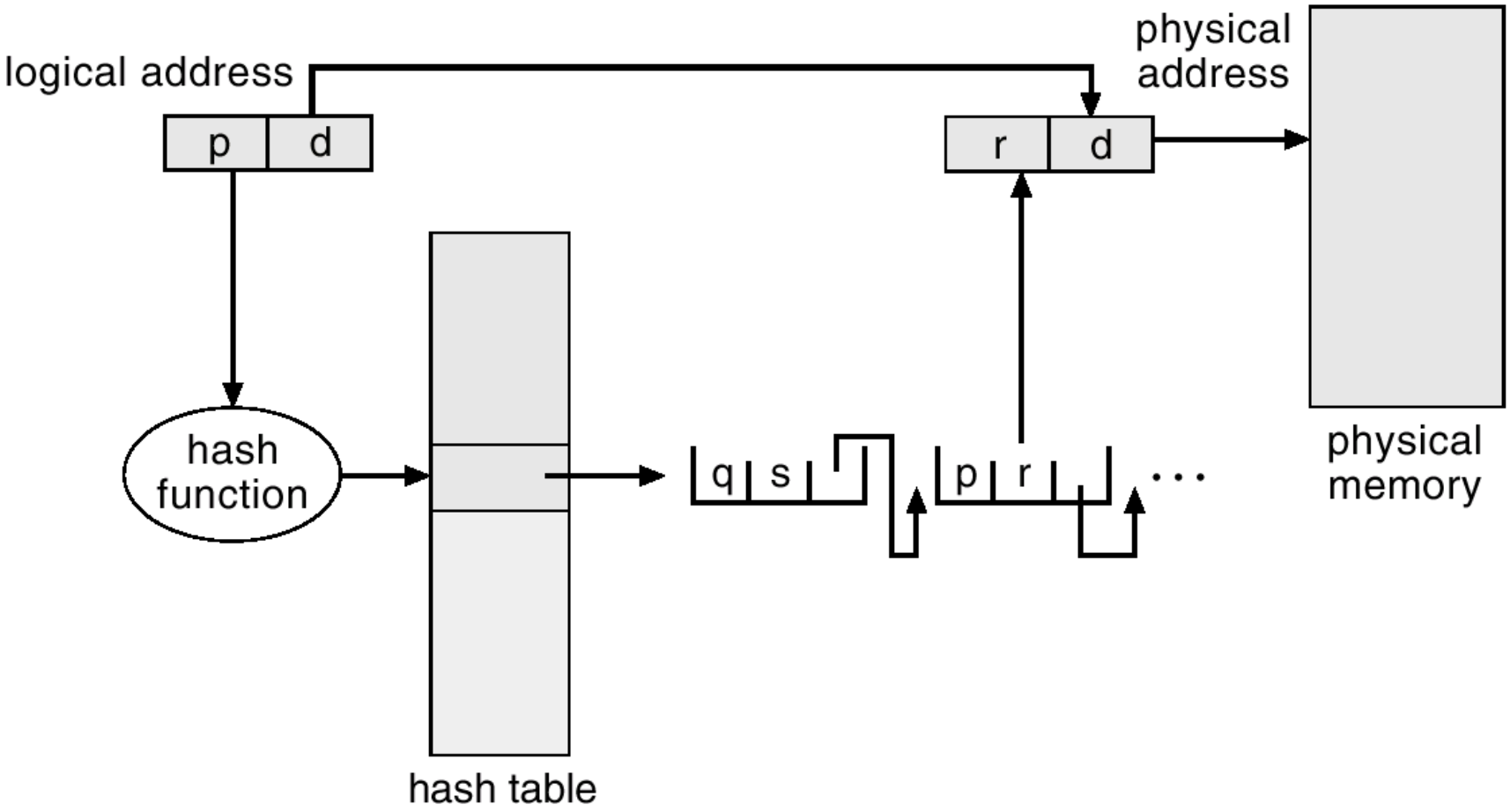


- where *pi is an index into the outer page table, and p2 is the* displacement within the page of the outer page table

Hashed Page Tables

- Common in address spaces > 32 bits.
- The virtual page number is hashed into a page table. This page table contains a chain of elements hashing to the same location.
- Virtual page numbers are compared in this chain searching for a match. If a match is found, the corresponding physical frame is extracted.

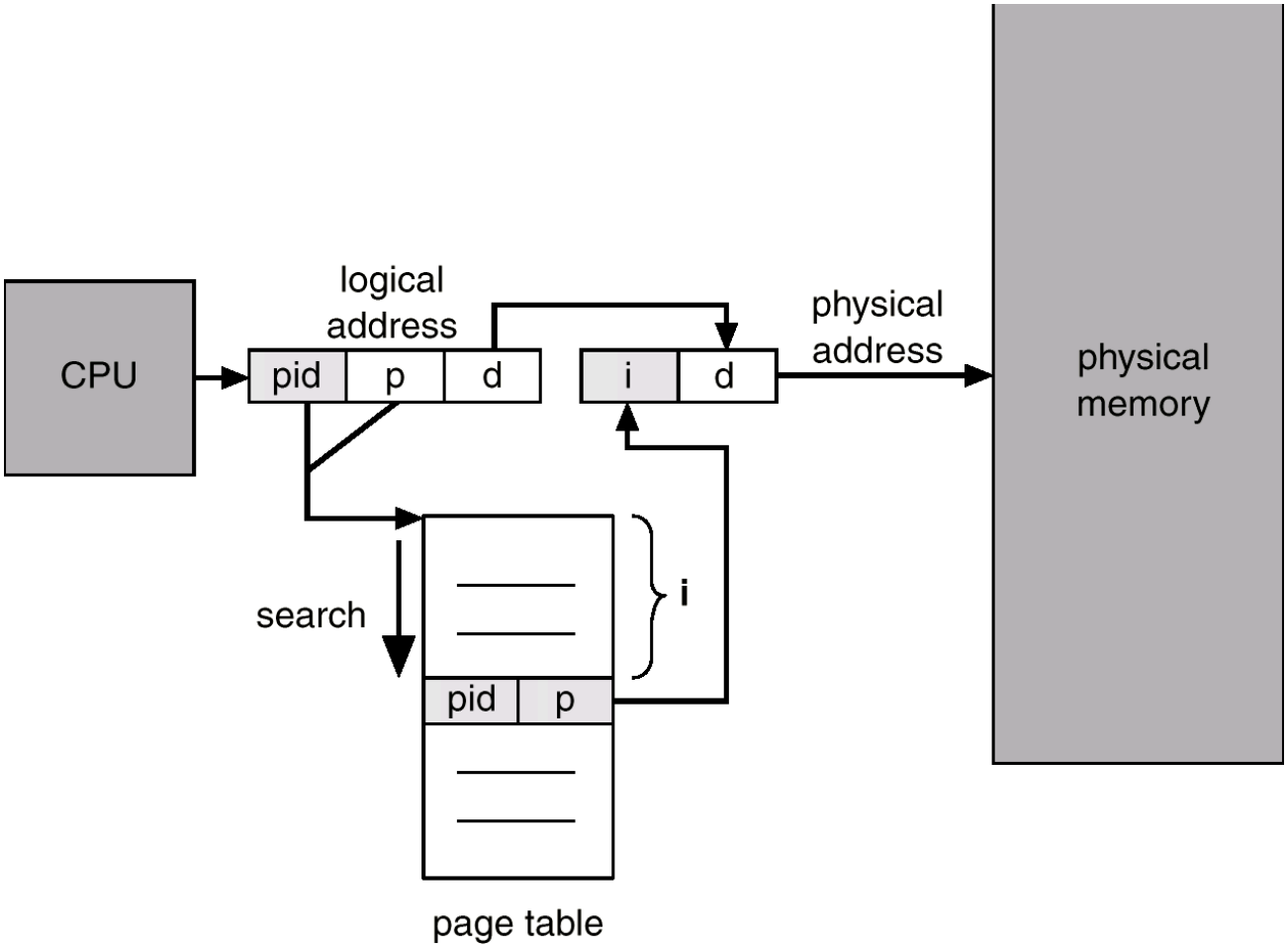
Hashed Page Table



Inverted Page Table

•One entry for each real page of memory.
•Entry consists of the virtual address of the page stored in that real memory location, with information about the process that owns that page.
• Decreases memory needed to store each page table, but increases time needed to search the table when a page reference occurs.
• Use hash table to limit the search to one — or at most a few — page-table entries.

Inverted Page Table Architecture



TLB(Translation Look aside Buffers)

A translation look aside buffer(TLB) is a memory cache that stores recent translations of virtual memory to physical addresses for faster retrieval. Each entry in the TLB consists of two parts: a key (or tag) and a value. When the associative memory is presented with an item, the item is compared with all keys simultaneously. If the item is found, the corresponding value field is returned. The search is fast; a TLB lookup in modern hardware is part of the instruction pipeline, essentially adding no performance penalty. To be able to execute the search within a pipeline step, however, the TLB must be kept small. It is typically between 32 and 1,024 entries in size. The TLB is used with page tables in the following way. The TLB contains only a few of the page-table entries. When a logical address is generated by the CPU, its page number is presented to the TLB. If the page number is found, its frame number is immediately available and is used to access memory. As just mentioned, these steps are executed as part of the instruction pipeline within the CPU, adding no performance penalty compared with a system that does not implement paging.

TLB(Translation Look aside Buffers(contd…)

- If the page number is not in the TLB (known as a **TLB miss), a memory** reference to the page table must be made. Depending on the CPU, this may be done automatically in hardware or via an interrupt to the operating system. When the frame number is obtained, we can use it to access memory (Figure in next slide). In addition, we add the page number and frame number to the TLB, so that they will be found quickly on the next reference. If the TLB is already full of entries, an existing entry must be selected for replacement.

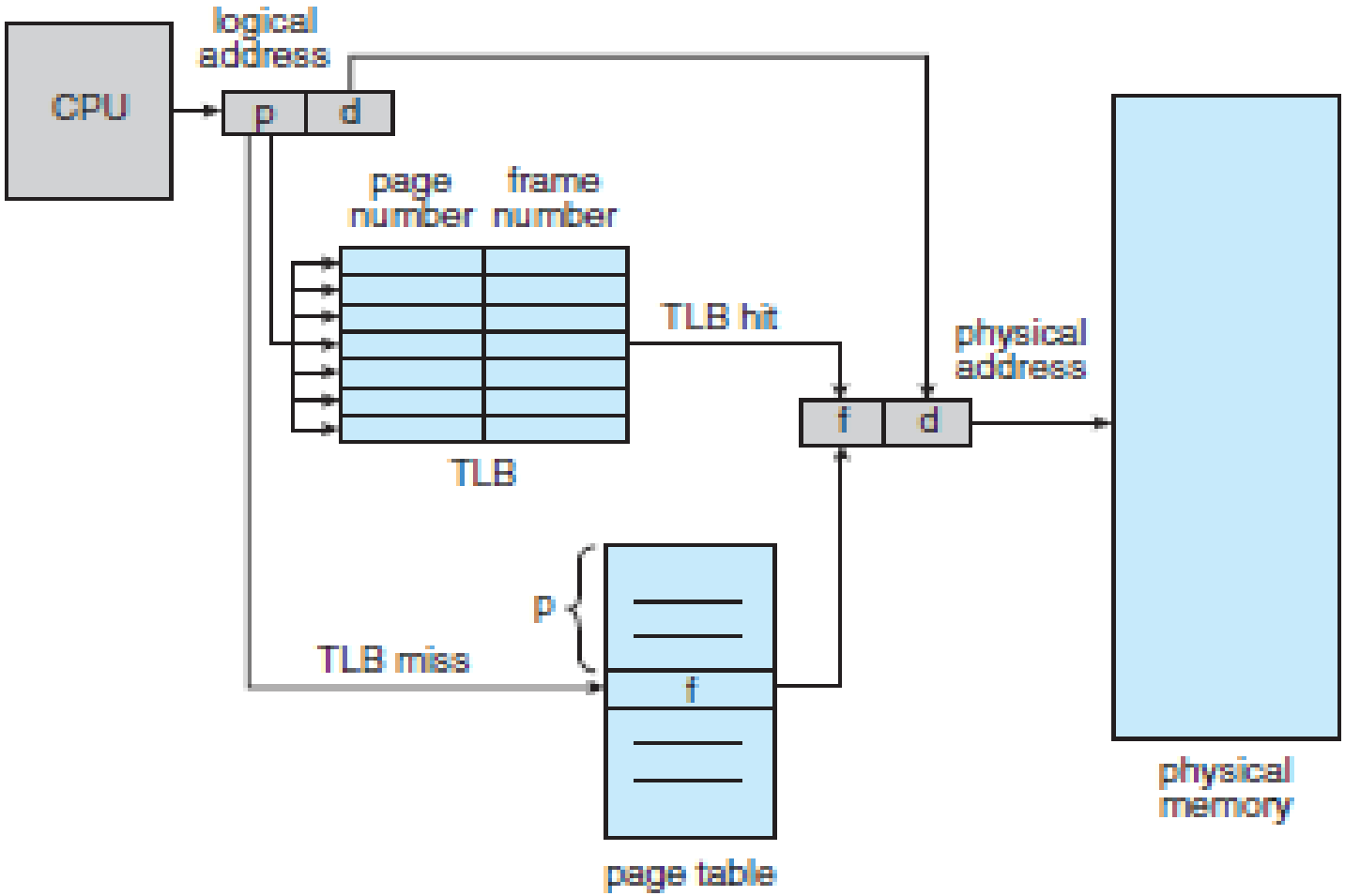


Fig: Paging hardware with TLB

Segmentation

- **Segmentation is a** Memory-management scheme that supports user view of memory.
- A program is a collection of segments. A segment is a logical unit such as: main program, procedure, function, method, object, local variables, global variables, common block,stack, symbol table, arrays

Segmentation Architecture

•Logical address consists of a two tuple:

<segment-number, offset>,

• *Segment table – maps two-dimensional physical* addresses; each table entry has:

✦ base – contains the starting physical address where the segments reside in memory.

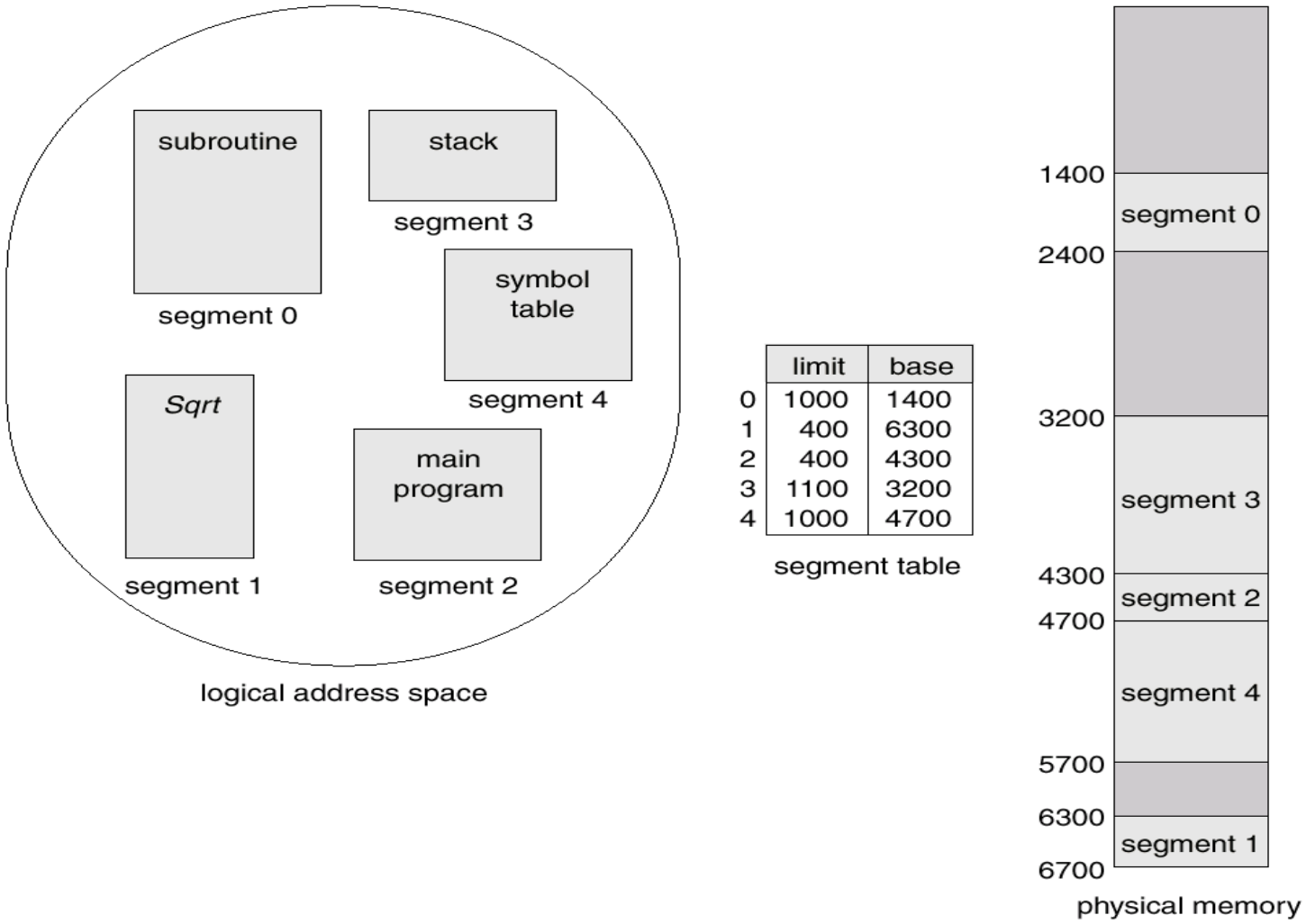
✦ *limit – specifies the length of the segment.*

• *Segment-table base register (STBR) points to the* segment table's location in memory.

• *Segment-table length register (STLR) indicates number of* segments used by a program;

segment number *s is legal if s < STLR.*

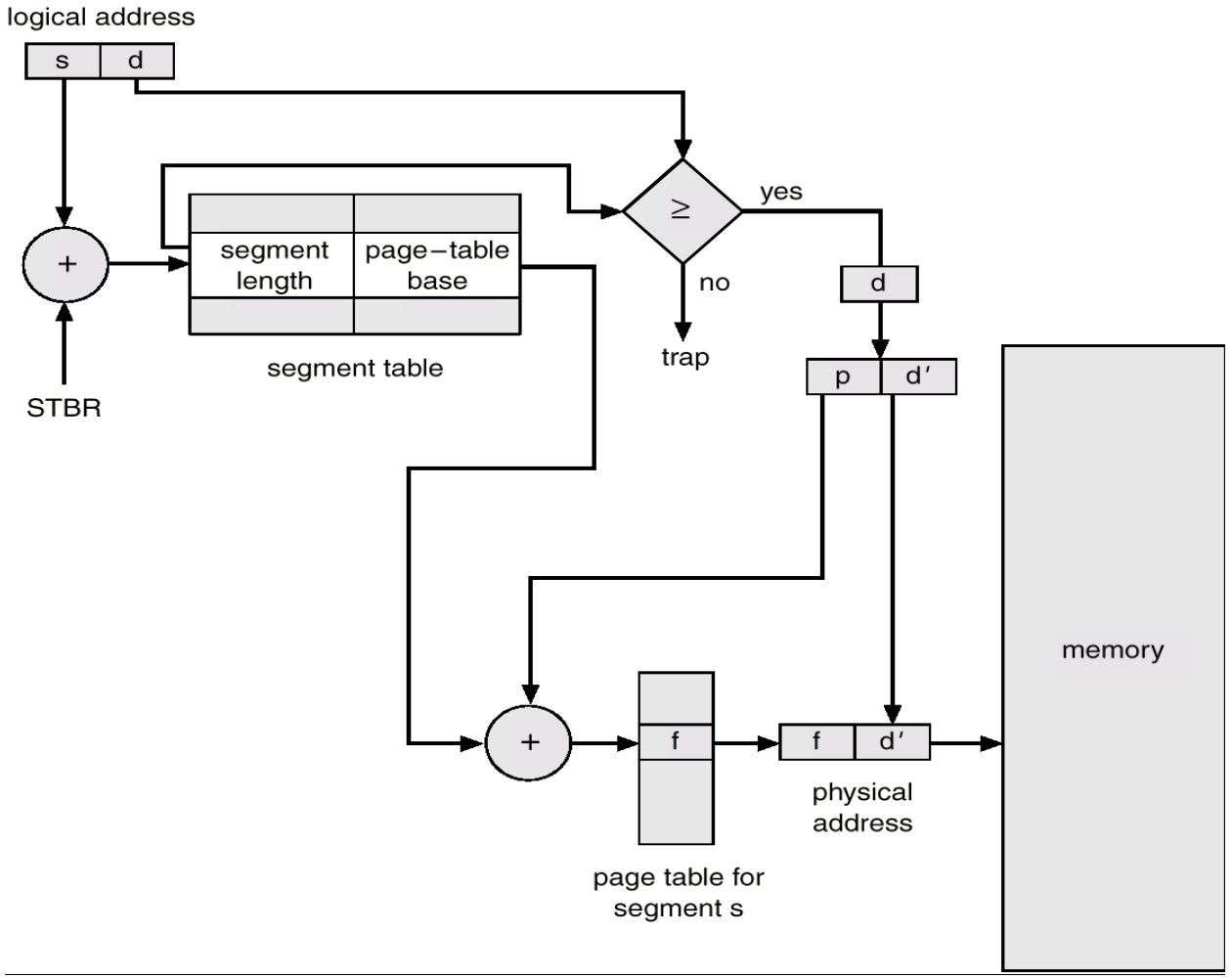
Example of Segmentation



Segmentation with Paging – MULTICS

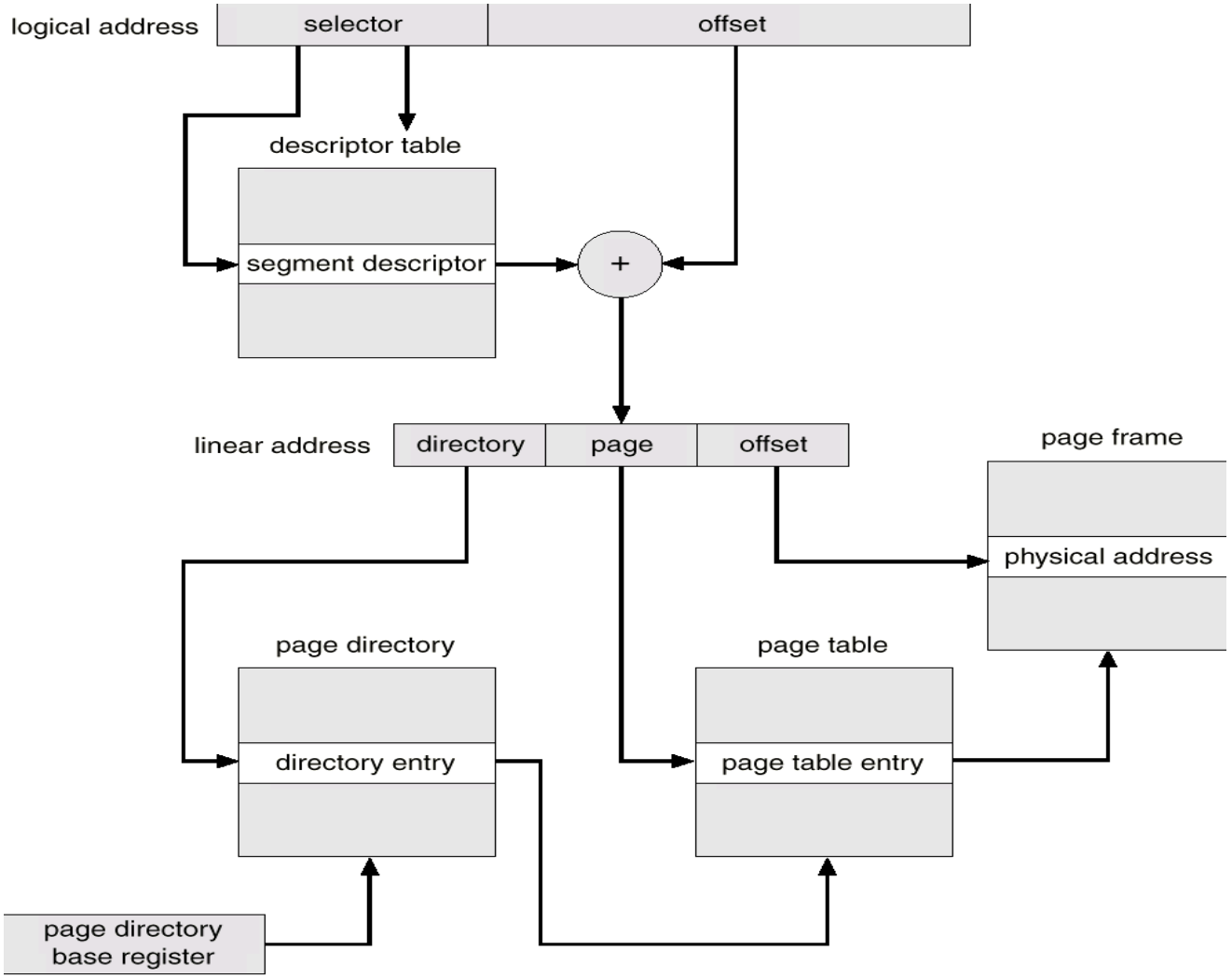
- The MULTICS system solved problems of external fragmentation and lengthy search times by paging the segments.
- Solution differs from pure segmentation in that the segment-table entry contains not the base address of the segment, but rather the base address of a *page table for* this segment.

MULTICS Address Translation Scheme



Segmentation with Paging – Intel 386

As shown in the following diagram, the Intel 386 uses segmentation with paging for memory management with a two-level paging scheme.



Relocation and Protection

•Relocation

In systems with virtual memory, programs in memory must be able to reside in different parts of the memory at different times. This is because when the program is swapped back into memory after being swapped out for a while it can not always be placed in the same location. The virtual memory management unit must also deal with concurrency. Memory management in the operating system should therefore be able to relocate programs in memory and handle memory references and addresses in the code of the program so that they always point to the right location in memory.

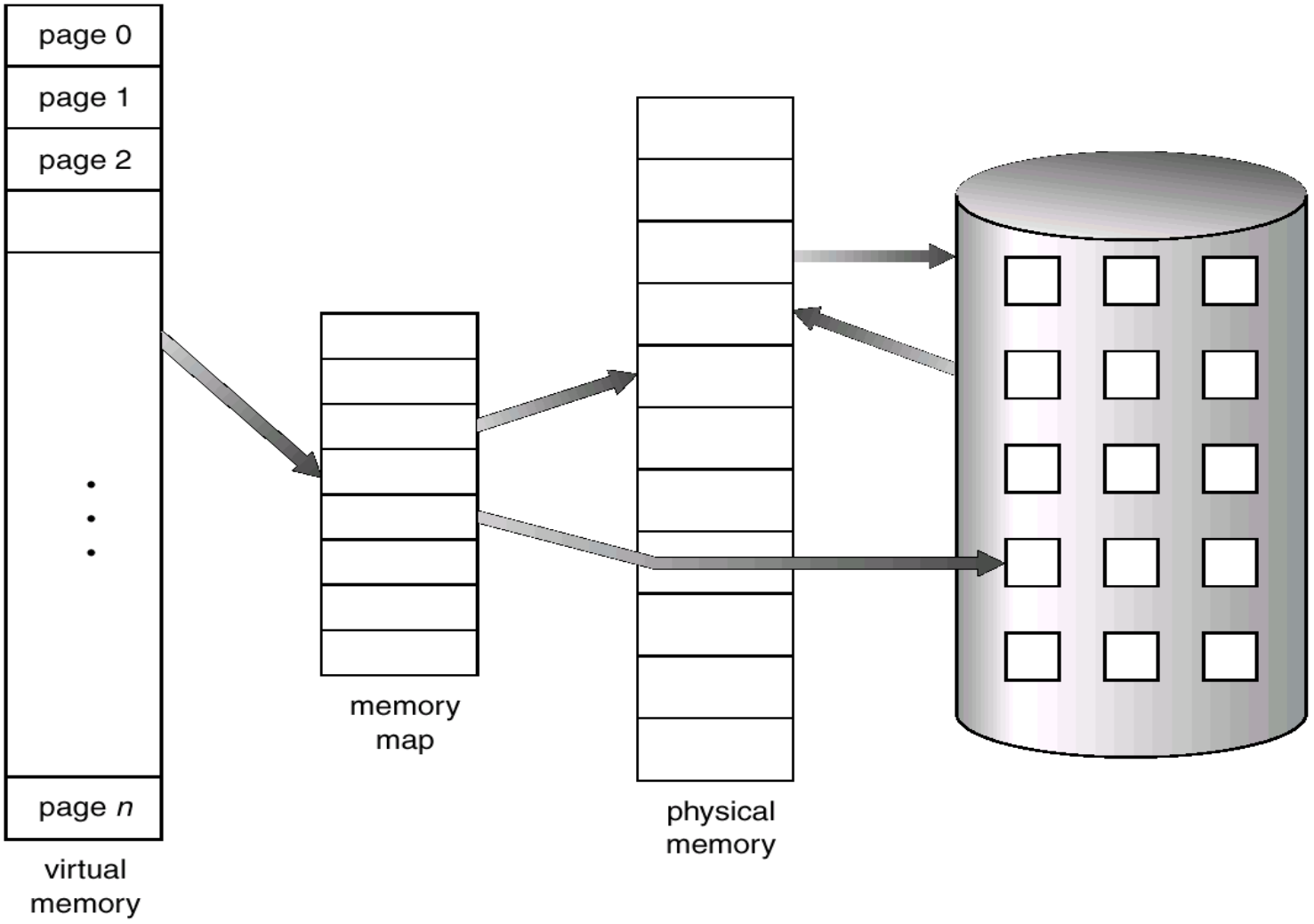
•Protection

Processes should not be able to reference the memory for another process without permission. This is called memory protection, and prevents malicious or malfunctioning code in one program from interfering with the operation of other running programs.

Virtual Memory

- Introduction:
  - Solves problem of limited memory space
  - Creates the illusion that more memory exists than is available in system
  - Two types of addresses in virtual memory systems
    - Virtual addresses
      - Referenced by processes
    - Physical addresses
      - Describes locations in main memory
  - Memory management unit (MMU)
    - Translates virtual addresses to physical address

Virtual Memory That is Larger Than Physical Memory



Virtual Memory: Basic Concepts

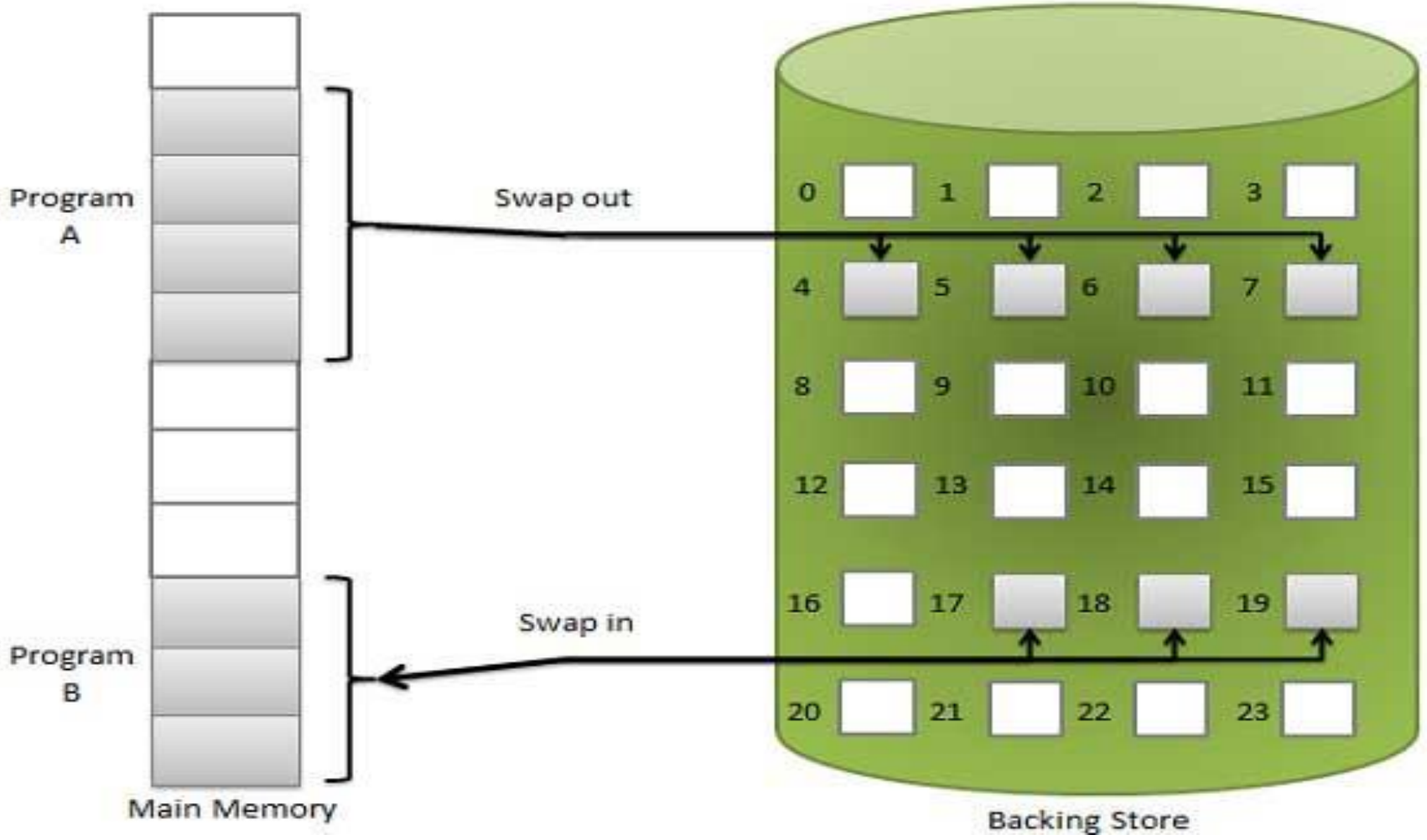
- Virtual address space, *V*
  - Range of virtual addresses that a process may reference
- Real address space, *R*
  - Range of physical addresses available on a particular computer system
- Dynamic address translation (DAT) mechanism
  - Converts virtual addresses to physical addresses during program execution
- |*V*| *is often much greater than* |*R*|
  - OS must store parts of *V for each process outside of main* memory
- Two-level storage
  - OS shuttles portions of *V between main memory (and caches) and* secondary storage

Demand Paging

A demand paging system is quite similar to a paging system with swapping. When we want to execute a process, we swap it into memory. Rather than swapping the entire process into memory, however, we use a lazy swapper called pager.

When a process is to be swapped in, the pager guesses which pages will be used before the process is swapped out again. Instead of swapping in a whole process, the pager brings only those necessary pages into memory. Thus, it avoids reading into memory pages that will not be used in anyway, decreasing the swap time and the amount of physical memory needed.

Hardware support is required to distinguish between those pages that are in memory and those pages that are on the disk using the valid-invalid bit scheme. Where valid and invalid pages can be checked by checking the bit. Marking a page will have no effect if the process never attempts to access the page. While the process executes and accesses pages that are memory resident, execution proceeds normally



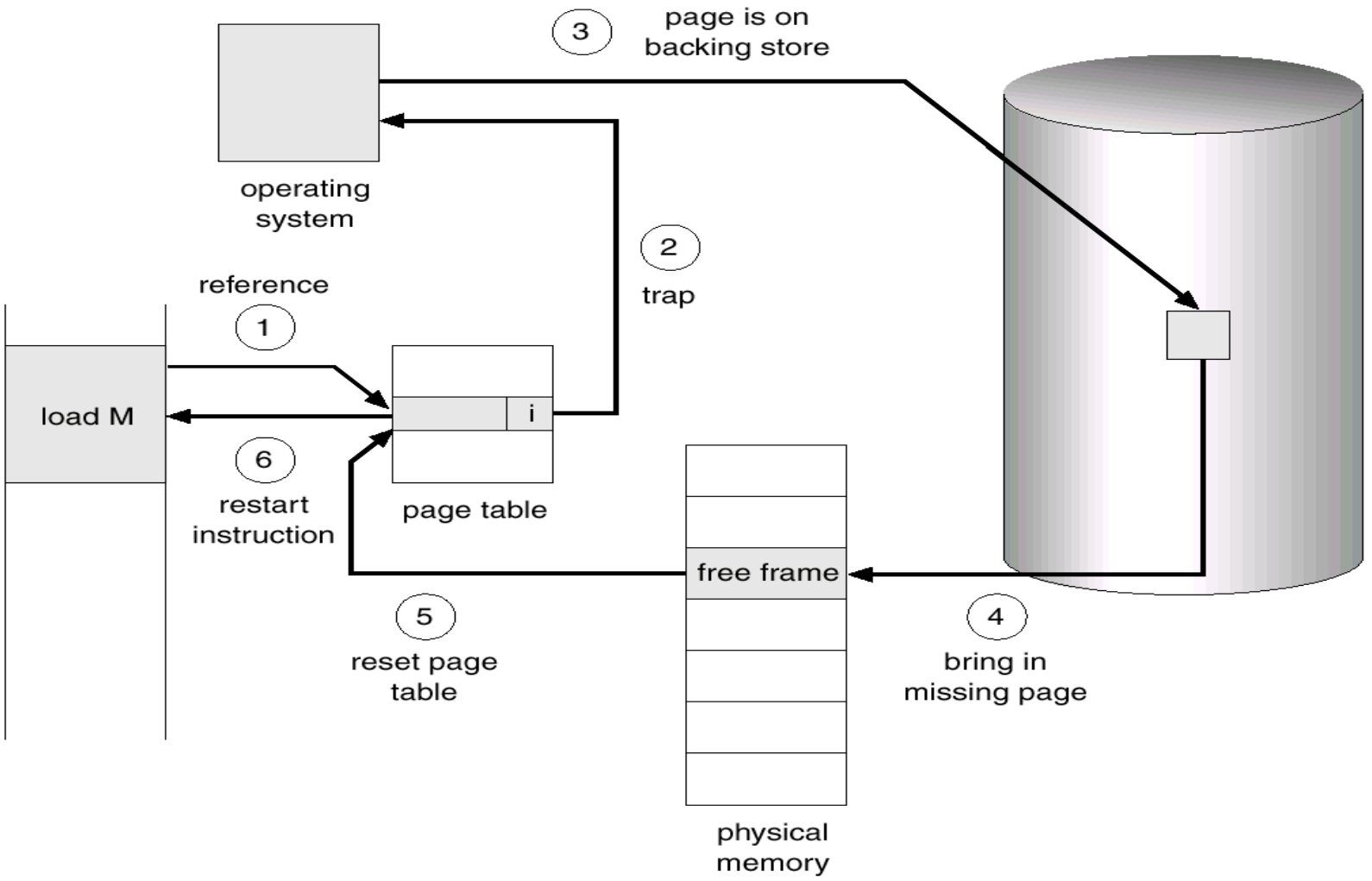
Transfer of a Paged Memory to Contiguous Disk Space

Page Fault

•If there is ever a reference to a page, first reference will trap to OS→page fault

- OS looks at another table to decide:
  - ✦ Invalid reference→ abort.
  - ✦ Just not in memory.
- Get empty frame.
- Swap page into frame.
- Reset tables, validation bit = 1.
- Restart instruction: Least Recently Used
  - ✦ block move
  - ✦ auto increment/decrement location

Steps in Handling a Page Fault



What happens if there is no free frame?

• Page replacement – find some page in memory, but not really in use, swap it out.

- ✦ algorithm
- ✦ performance – want an algorithm which will result in minimum number of page faults.

• Same page may be brought into memory several times.

Page Replacement

- Prevent over-allocation of memory by modifying page fault service routine to include page replacement.
- Use *modify (dirty) bit to reduce overhead of page* transfers – only modified pages are written to disk.
- Page replacement completes separation between logical memory and physical memory – large virtual memory can be provided on a smaller physical memory.

Basic Page Replacement

•Find the location of the desired page on disk.

• Find a free frame:

- If there is a free frame, use it.

- If there is no free frame, use a page replacement

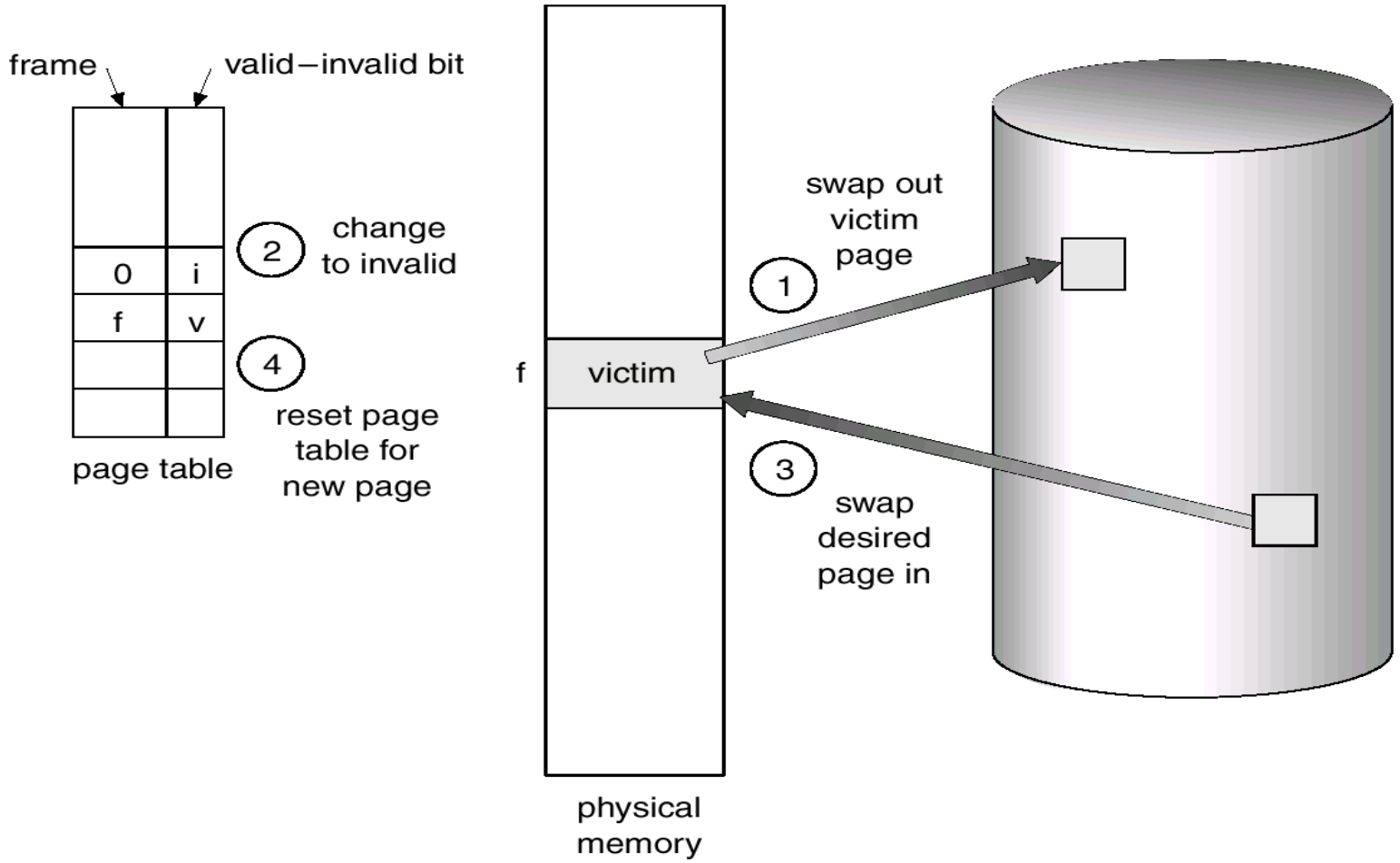
algorithm to select a *victim frame.*

• Read the desired page into the (newly) free frame

. Update the page and frame tables.

• Restart the process.

Page Replacement



Page Replacement Algorithms

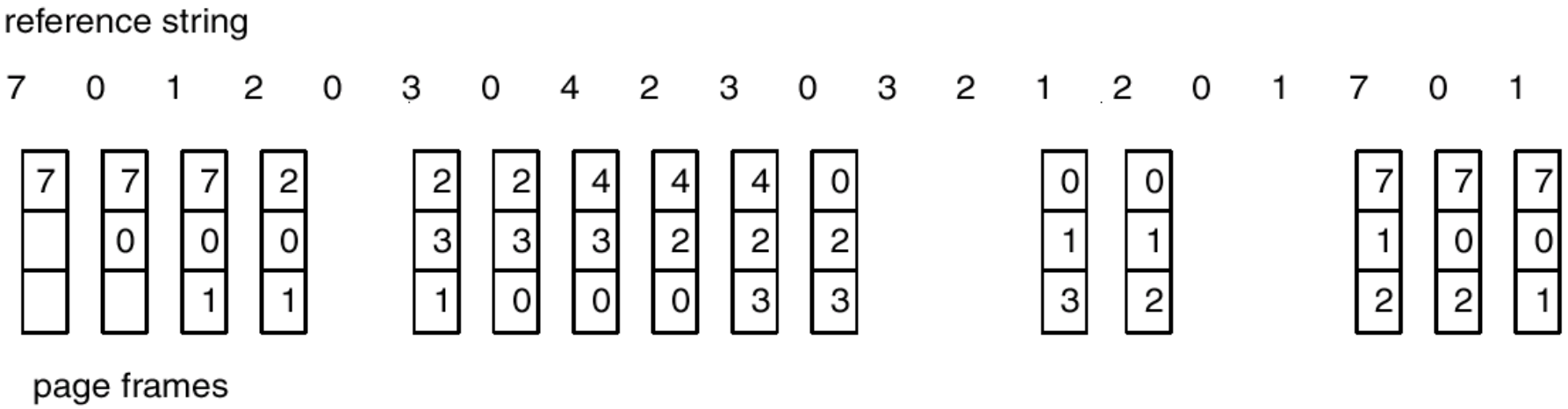
- Want lowest page-fault rate.
- Evaluate algorithm by running it on a particular string of memory references (reference string) and computing thenumber of page faults on that string.
- In all our examples, the reference string is

7,0,1,2,0,3,0,4,2,3,0,3,2,1,2,0,1,7,0,1

First-In-First-Out (FIFO) Page Replacement

- FIFO page replacement
  - Replace page that has been in the system the longest
  - Likely to replace heavily used pages
  - Can be implemented with relatively low overhead
  - Impractical for most systems

FIFO Page Replacement



3 frames are used (3 pages can be in memory at a time per process)

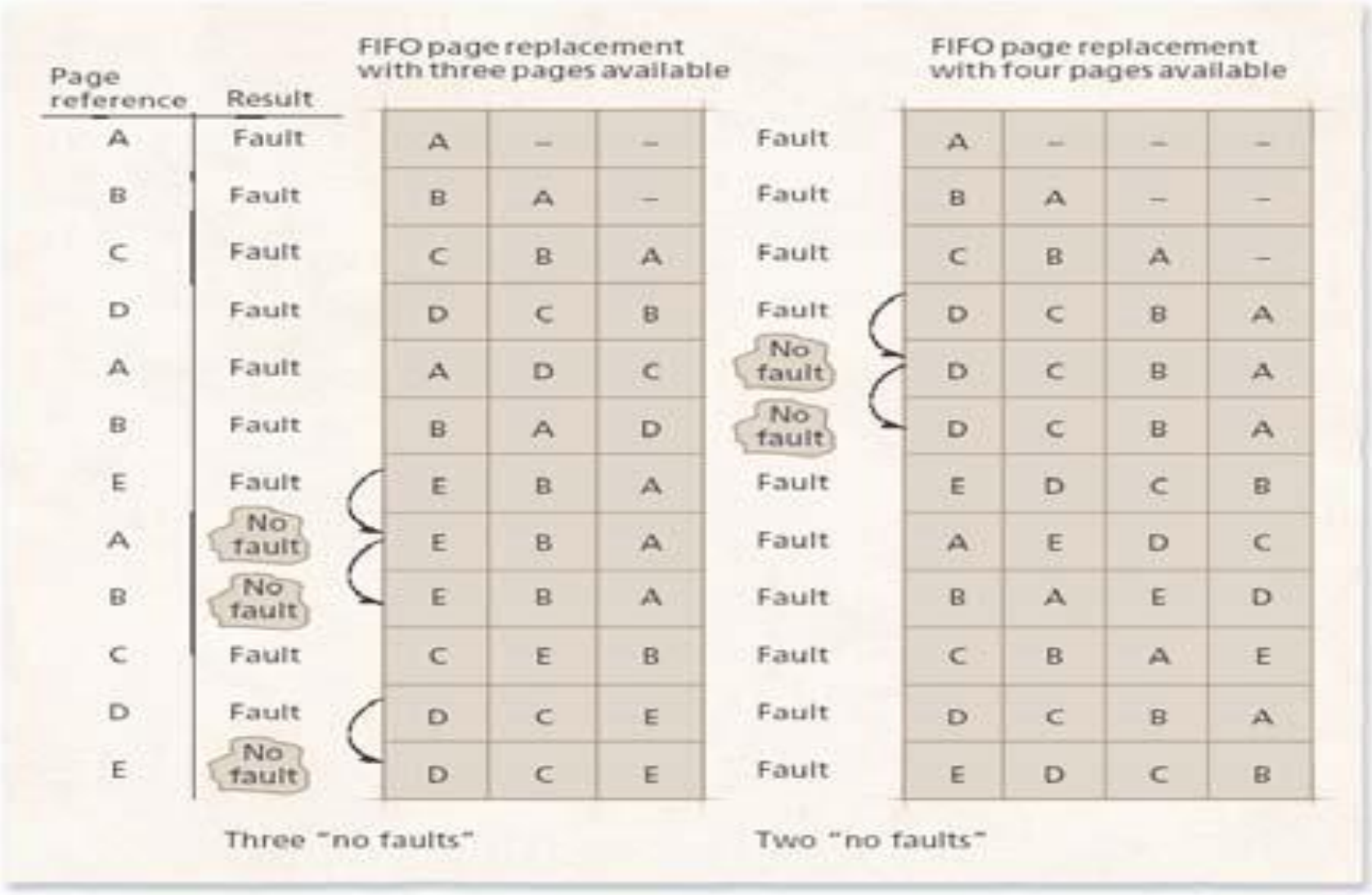
FIFO Page Replacement

•For our example reference string, our three frames are initially empty. The first three references (7,0,1) cause page faults, and are brought into these empty frames. The next reference (2) replaces page 7, because page 7 was brought in first. Since *0 is the next reference and 0 is already in memory, we have no fault* for this reference. The first reference to **3 results in page *0 being replaced, since*** it was the first of the three pages in memory (0, 1, and 2) to be brought in. Because of this replacement, the next reference, to 0, will fault. Page 1 is then replaced by page 0. This process continues as shown in Figure of previous slide. Every time a fault occurs, we show which pages are in our three frames. There are 15 faults altogether.

FIFO Anomaly

- Belady's (or FIFO) Anomoly
  - Certain page reference patterns actually cause more page faults when number of page frames allocated to a process is increased

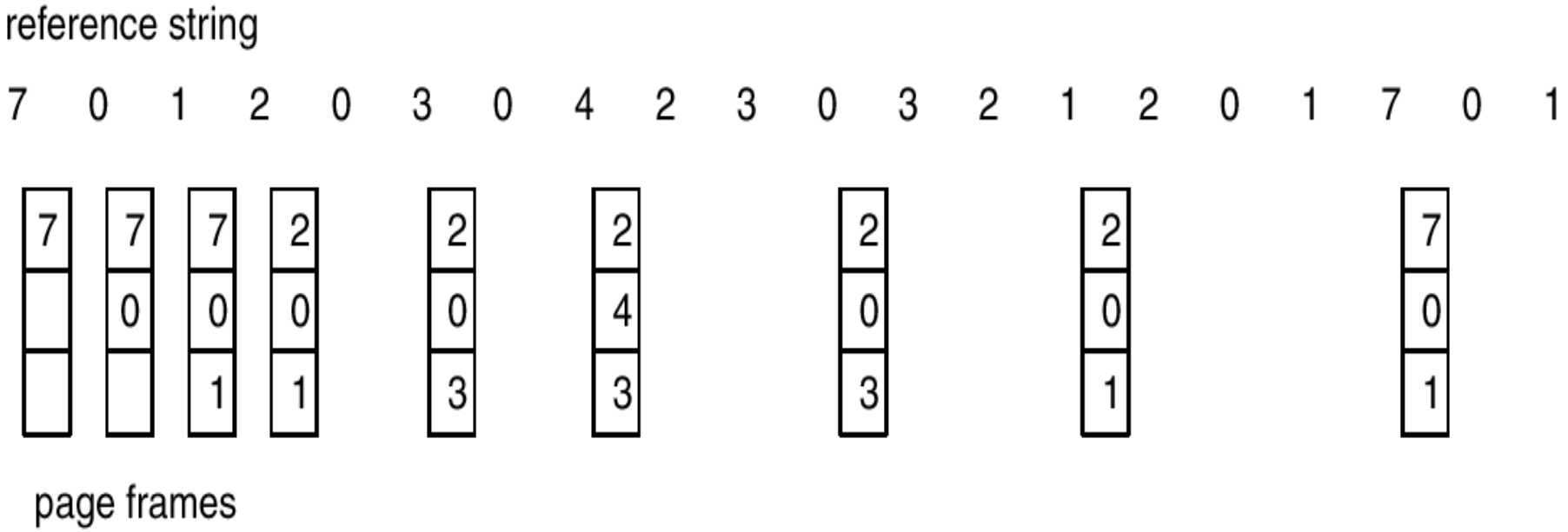
FIFO Anomaly



Optimal Page Replacement

- Replace page that will not be used for longest period of time.
- Use of this page replacement algorithm guarantees the lowest possible page rate for a fixed number of frames.
- How do you know this?
- Used for measuring how well your algorithm performs.

Optimal Page Replacement



Optimal Page Replacement

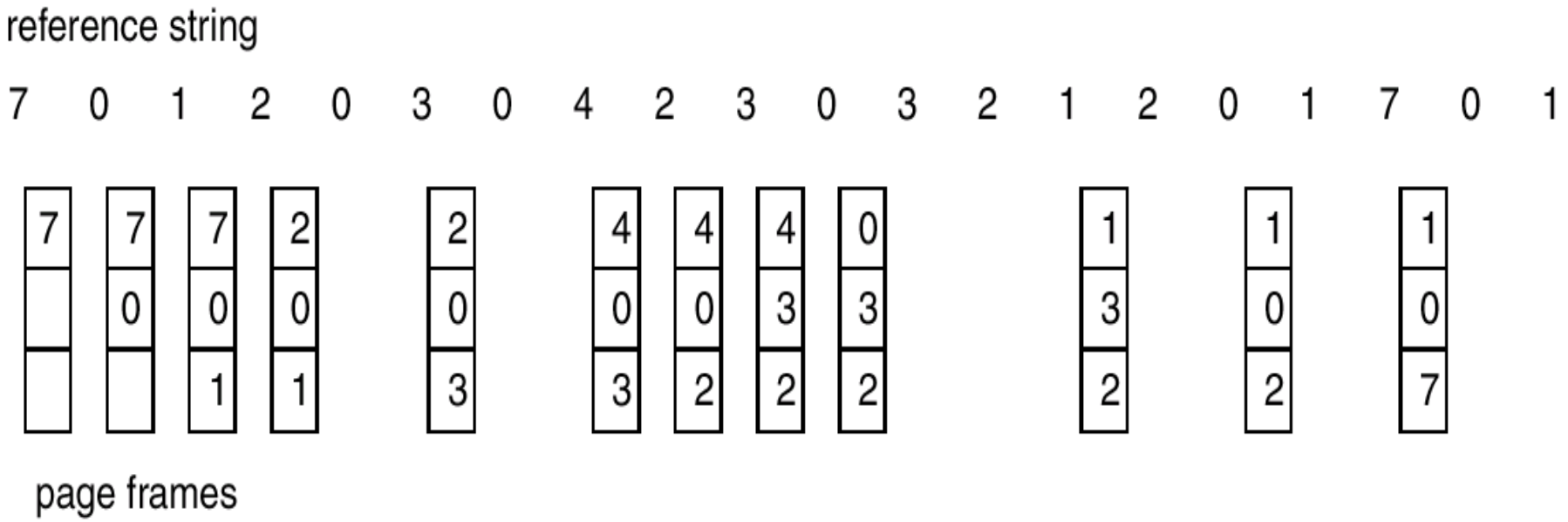
•For example, on our sample reference string, the optimal page-replacement algorithm would yield nine page faults, as shown in Figure of previous slide. The first three references cause faults that fill the three empty frames. The reference to page 2 replaces page 7, because 7 will not be used until reference 18, whereas page 0 will be used at 5, and page 1 at 14. The reference to page 3 replaces page 1, as page 1 will be the last of the three pages in memory to be referenced again. With only nine page faults, optimal replacement is much better than a FIFO algorithm, which had 15 faults.

•Unfortunately, the optimal page-replacement algorithm is difficult to implement, because it requires future knowledge of the reference string.

Least Recently Used (LRU) Algorithm

- LRU page replacement
  - Exploits temporal locality by replacing the page that has spent the longest time in memory without being referenced
- – Can provide better performance than FIFO
- – Increased system overhead
- – LRU can perform poorly if the least-recently used page is the next page to be referenced by a program that is iterating inside a loop that references several pages

LRU Page Replacement



LRU Page Replacement

•The result of applying LRU replacement to our example reference string is shown in Figure of previous slide. The LRU algorithm produces 12 faults. Notice that the first five faults are the same as the optimal replacement. When the reference to page 4 occurs, however, LRU replacement sees that, of the three frames in memory, page 2 was used least recently. The most recently used page is page 0, and just before that page **3 was used. Thus, the LRU algorithm replaces page** 2, not knowing that page 2 is about to be used. When it then faults for page 2,the LRU algorithm replaces page **3 since, of the three pages in memory {0,3,4},** page **3 is the least recently used. Despite these problems, LRU replacement with** 12 faults is still much better than FIFO replacement with 15.

LRU Algorithm (Cont.)

•Counter implementation

✦ Every page entry has a counter; every time page is referenced through this entry, copy the clock into the counter.

✦ When a page needs to be changed, look at the counters to determine which are to change.

•Stack implementation – keep a stack of page numbers in a double link form:

•✦ Page referenced:

✔ move it to the top. So the most recently used page is always at the top of the stack and the least recently used page is always at the bottom.

✦ No search for replacement

Not Recently Used Page Replacement Algorithm

-Approximates LRU with little overhead by using referenced bit and modified bit to determine which page has not been used recently and can be replaced quickly

– Can be implemented on machines that lack hardware referenced bit and/or modified bit

Group	Referenced	Modified	Description
Group 1	0	0	Best choice to replace
Group 2	0	1	[Seems unrealistic]
Group 3	1	0	
Group 4	1	1	Worst choice to replace

Figure : Page types under NRU.

Second-Chance Page Replacement Algorithm

•The basic algorithm of second-chance replacement is a FIFO replacement algorithm. When a page has been selected, however, we inspect its reference bit. If the value is 0, we proceed to replace this page. If the reference bit is set to 1, however, we give that page a second chance and move on to select the next FIFO page. When a page gets a second chance, its reference bit is cleared and its arrival time is reset to the current time. Thus, a page that is given a second chance will not be replaced until all other pages are replaced (or given second chances). In addition, if a page is used often enough to keep its reference bit set, it will never be replaced.

•**Clock page replacement**

– Similar to second chance, but arranges the pages in circular list instead of linear list

Working Set Page Replacement Algorithm

•The set of pages that a process is currently using is known as its working set. If the entire working set is in memory, the process will run without causing many faults until it moves into another execution phase. If the the available memory is too small to hold the entire working set, the process will cause many page faults and run slowly.

•To implement the working set model, it is necessary for the operating system to keep track of which pages are in working set. Having this information also immediately leads to a possible page replacement algorithm: when a page fault occurs, find a page not in the working set and evict it.

Allocation Of Frames

.There were two important tasks in virtual memory management: a page-replacement strategy and a frame-allocation strategy.

Minimum Number of Frames

.The absolute minimum number of frames that a process must be allocated is dependent on system architecture, and corresponds to the worst-case scenario of the number of pages that could be touched by a single (machine) instruction.
If an instruction (and its operands) spans a page boundary, then multiple pages could be needed just for the instruction fetch.
Memory references in an instruction touch more pages, and if those memory locations can span page boundaries, then multiple pages could be needed for operand access also.

.The worst case involves indirect addressing, particularly where multiple levels of indirect addressing are allowed. Left unchecked, a pointer to a pointer to a pointer to a pointer to a . . . could theoretically touch every page in the virtual address space in a single machine instruction, requiring every virtual page be loaded into physical memory simultaneously. For this reason architectures place a limit (say 16) on the number of levels of indirection allowed in an instruction, which is enforced with a counter initialized to the limit and decremented with every level of indirection in an instruction - If the counter reaches zero, then an "excessive indirection" trap occurs . This example would still require a minimum frame allocation of 17 per process.

Allocation of frames(contd..)

Allocation Algorithms

- Equal Allocation - If there are m frames available and n processes to share them, each process gets m / n frames, and the leftovers are kept in a free-frame buffer pool.
- Proportional Allocation - Allocate the frames proportionally to the size of the Process relative to the total size of all processes. So if the size of process i is S_i, and S is the sum of all S_i, then the allocation for process P_i is a_i = m * S_i / S. Variations on proportional allocation could consider priority of process rather than just their size.

Global versus Local Allocation

- One big question is whether frame allocation (page replacement) occurs on a local or global level.
- With local replacement, the number of pages allocated to a process is fixed, and page replacement occurs only amongst the pages allocated to this process. With global replacement, any page may be a potential victim, whether it currently belongs to the process seeking a free frame or not.
- Local page replacement allows processes to better control their own page fault rates, and leads to more consistent performance of a given process over different system load levels. Global page replacement is overall more efficient, and is the more commonly used approach.

Thrashing

•If the process does not have sufficient number of frames for the pages in active use, it will quickly page fault. At this point, it must replace some page. However, since all its pages are in active use, it must replace a page that will be needed again right away. Consequently, it quickly faults again, and again, and again. The process continues to fault, replacing pages for which it then faults and brings back in right away.
•This high paging activity is called **thrashing. A process is thrashing if it is** spending more time paging than executing.