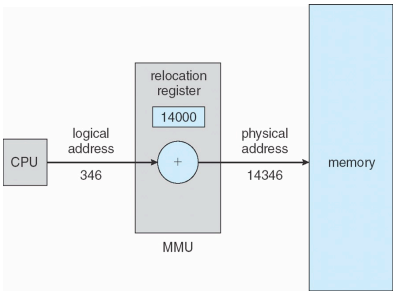
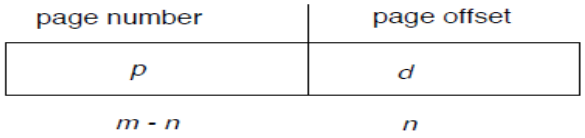
Ch.8 – Main Memory

- Cache sits between main memory and CPU registers
- Base and limit registers define logical address space usable by a process
- Compiled code addresses bind to relocatable addresses
  - Can happen at three different stages
    - Compile time: If memory location known a priori, absolute code can be generated
    - Load time: Must generate relocatable code if memory location not known at compile time
    - Execution time: Binding delayed until run time if the process can be moved during its execution
- Memory-Management Unit (MMU) device that maps virtual to physical address
- Simple scheme uses a relocation register which just adds a base value to address
- Swapping allows total physical memory space of processes to exceed physical memory
  - Def: process swapped out temporarily to backing store then brought back in for continued execution
- Backing store: fast disk large enough to accommodate copes of all memory images
- Roll out, roll in: swapping variant for priority-based scheduling.
  - Lower priority process swapped out so that higher priority process can be loaded



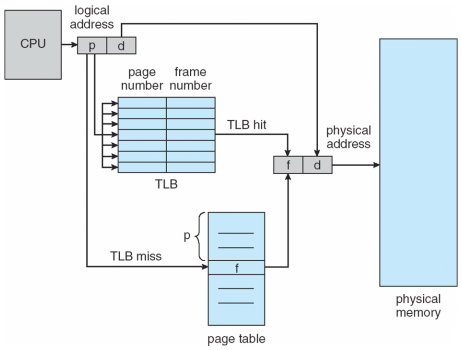
- Solutions to Dynamic Storage-Allocation Problem:
  - First-fit: allocate the first hole that is big enough
  - Best-fit: allocate the smallest hole that is big enough (must search entire list) → smallest leftover hole
  - Worst-fit: allocate the largest hole (search entire list) → largest leftover hole
- External Fragmentation: total memory space exists to satisfy request, but is not contiguous
  - Reduced by compaction: relocate free memory to be together in one block
    - Only possible if relocation is dynamic
- Internal Fragmentation: allocated memory may be slightly larger than requested memory
- Physical memory divided into fixed-sized frames: size is power of 2, between 512 bytes and 16 MB
- Logical memory divided into same sized blocks: pages
- Page table used to translate logical to physical addresses
  - Page number (p): used as an index into a page table
  - Page offset (d): combined with base address to define the physical memory address
- Free-frame list is maintained to keep track of which frames can be allocated



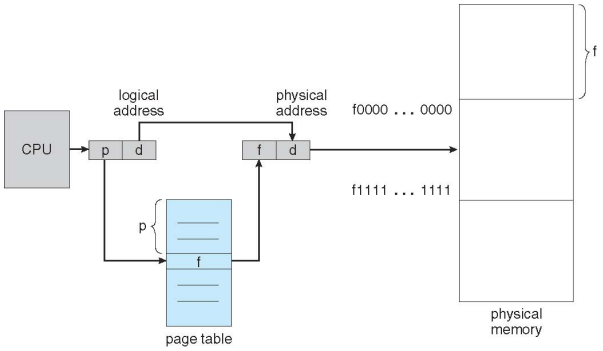
For given logical address space 2^m and page size 2^n

Ch.8 – Main Memory Continued

- Transition Look-aside Buffer (TLB) is a CPU cache that memory management hardware uses to improve virtual address translation speed
  - Typically small – 64 to 1024 entries
  - On TLB miss, value loaded to TLB for faster access next time
  - TLB is associative – searched in parallel

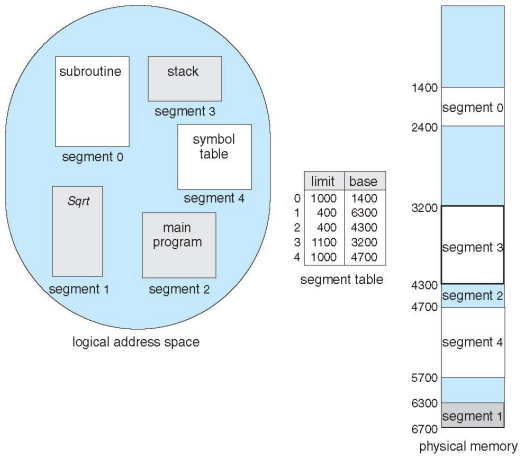


Paging with TLB

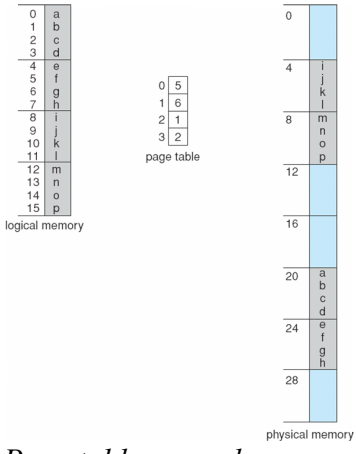


Paging without TLB

- Effective Access Time: EAT = $(1 + \varepsilon)\,\alpha + (2 + \varepsilon)(1 - \alpha)$
  - ε = time unit, α = hit ratio
- Valid and invalid bits can be used to protect memory
  - "Valid" if the associated page is in the process' logical address space, so it is a legal page
- Can have multilevel page tables (paged page tables)
- Hashed Page Tables: virtual page number hashed into page table
  - Page table has chain of elements hashing to the same location
  - Each element has (1) virtual page number, (2) value of mapped page frame, (3) a pointer to the next element
  - Search through the chain for virtual page number
- Segment table – maps two-dimensional physical addresses
  - Entries protected with valid bits and r/w/x privileges



Segmentation example



Page table example