Chapter 8: Main Memory

1. Background

- Typical instruction/execution cycle
  - Fetches an instruction from memory
  - Decode
  - Operand fetched from memory
  - Result stored back in memory
- Memory units see **ONLY** a stream of memory addresses
  - Does not know
    - how they are generated (PC, indirection, indexing, literal addresses…)
    - What they are for (data, instruction)
- Accordingly, we can ignore **HOW** a program generates a memory address
  - We are interested only in the sequence of memory addresses generated by the running program

Chapter 8: Memory Management

Objectives:

- Detailed description of various ways of organizing memory hardware
- Various memory-management techniques, including paging and segmentation
- Detailed description of the Intel Pentium, which supports both pure segmentation and segmentation with paging

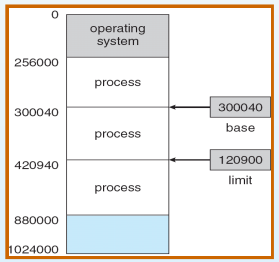
1. Background
2. Swapping
3. Contiguous Memory Allocation
4. Paging
5. Structure of the Page Table
6. Segmentation
7. Example: The Intel Pentium

Background (cont)

- Program must be brought (from disk) into memory and placed within a process for it to be run
- Main memory and registers are only storage CPU can access directly
- Register access in one CPU clock (or less)
- Main memory can take many cycles
- **Cache** sits between main memory and CPU registers
- Protection of memory required to ensure correct operation
  - Each process has a separate memory space

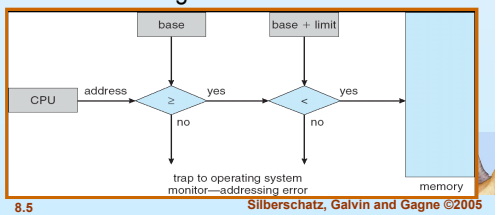
Base and Limit Registers

- A pair of **base** and **limit** registers define the logical address space



Loaded only by the OS
→special privilege instruction
→ Unrestricted access to the OS

- The CPU hardware compares **EVERY** address generated in user mode with the register

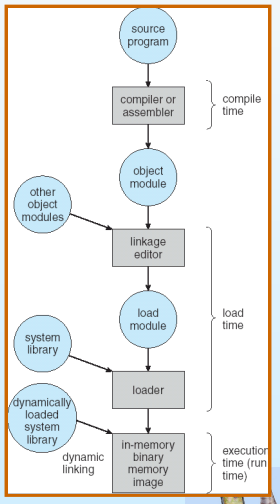


Logical vs. Physical Address Space

- 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 / loaded into the memory-address register
- 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

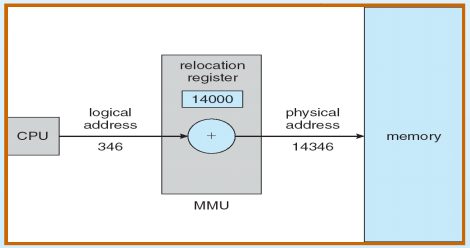
Binding of Instructions and Data to Memory

- Address binding of instructions and data to memory addresses can happen at three different stages
  - **Compile time**: If memory location known a priori, **absolute code** can be generated; must recompile code if starting location changes
  - **Load time**: Must generate **relocatable code** if memory location is not known at compile time: reload only user code when location is known
  - **Execution time**: Binding delayed until run time if the process can be moved during its execution from one memory segment to another. Special hardware support must be available (e.g., base and limit registers)



Memory-Management Unit (MMU)

- Hardware device that maps virtual to physical address
- In MMU scheme, the value in the relocation register is added to every address generated by a user process at the time it is sent to memory
- The user program deals with *logical* addresses; it never sees the *real* physical addresses
  - Could set (a=346) / inc (a++) / compute / compare…
  - But when use as an address *a → relocated relative to the base
  - Final location not known until the reference is made (exec-time binding)



Dynamic Loading

- Up to now, entire program and all data must be in physical memory
  - Size of a process <= size of physical memory
- Dynamic loading:
  - Routine is not loaded until it is called
  - Better memory-space utilization: unused routine is never loaded
  - Useful when large amounts of code are needed to handle infrequently occurring cases
  - No special support from the operating system is required; implemented through program design

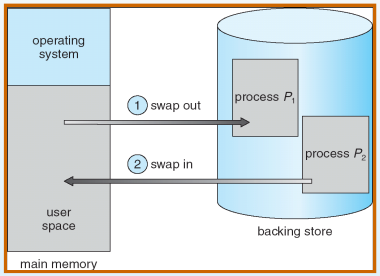
Dynamic Linking

- Linking postponed until execution time
- Small piece of code, *stub*, used to locate the appropriate memory-resident library routine
- Stub replaces itself with the address of the routine, and executes the routine
- Operating system needed to check if routine is in processes' memory address
- Dynamic linking is particularly useful for libraries
- System also known as **shared libraries**

2. 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: 10 MB at 40 MB per second → around ¼ second
- Modified versions of swapping are found on many systems (i.e., UNIX, Linux, and Windows)
- System maintains a **ready queue** of ready-to-run processes that have memory images on disk

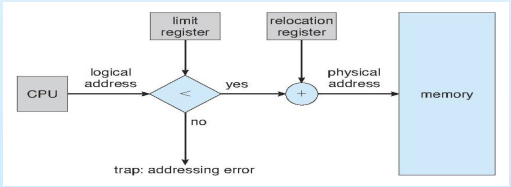
Schematic View of Swapping



- Context switch time problem
- Process must be completely idle
  - Problem with pending I/Os
  - → never swap process with pending I/0
  - Or execute I/O only into OS buffers

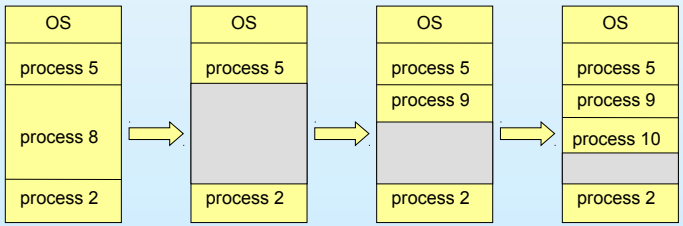
3. Contiguous Allocation

- 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*



Contiguous Allocation (Cont.)

- 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)



Dynamic Storage-Allocation Problem

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

Bit table or linked list to keep track of holes

Fragmentation

- **External Fragmentation** – total memory space exists to satisfy a request, but it is not contiguous – 50-percent rule: 1/3 of memory unusable! (statistical analysis)
- **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
    - Last job in memory while it is involved in I/O
    - Do I/O only into OS buffers

4. 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 16 MB)
- 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

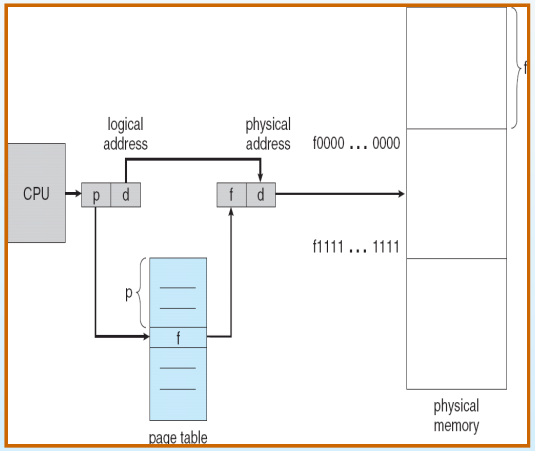
Address Translation Scheme

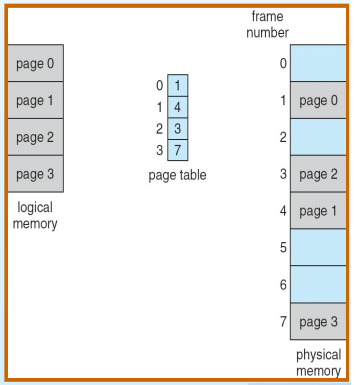
- Address generated by CPU is divided into:
  - **Page number (*p*)** – used as an index into a *page table* which contains base address of each page in physical memory
  - **Page offset (d)** – combined with base address to define the physical memory address that is sent to the memory unit

page number	page offset
p	*d*
m - n	*n*

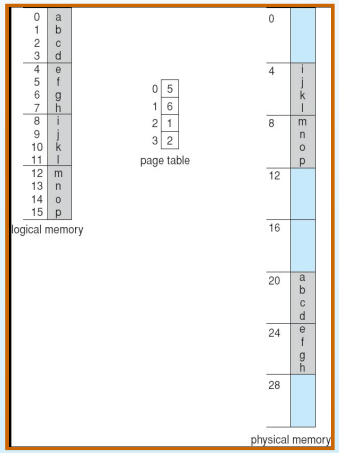
for given logical address space of size 2^m and page size 2^n

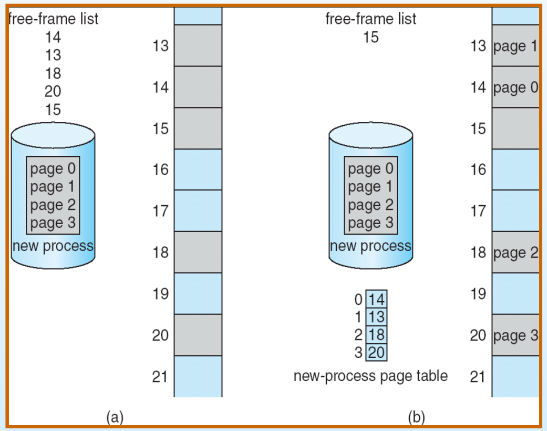
Paging Hardware, Logical and Physical Memory





Paging Examples





Before allocation

After allocation

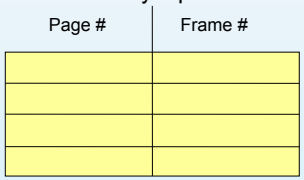
Example with 32-byte memory and 4-byte pages

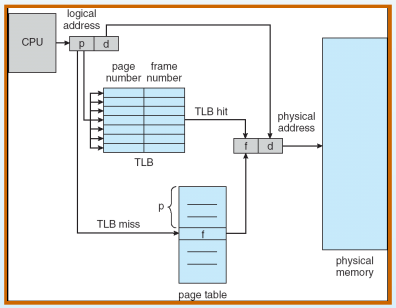
Implementation of Page Table

- Page table is kept in main memory
- **Page-table base register (PTBR)** points to the page table
- **Page-table length register (PTLR)** indicates size of the page table
- In this scheme every data/instruction access requires two memory accesses. One for the page table and one for the data/instruction.
- The two memory access problem can be solved by the use of a special fast-lookup hardware cache called **associative memory** or **translation look-aside buffers (TLBs)**
- Some TLBs store **address-space identifiers (ASIDs)** in each TLB entry – uniquely identifies each process to provide address-space protection for that process

Associative Memory

- TLBs typically small (64 to 1024 entries)
- Associative memory – parallel search





Address translation (p, d)

- If p is in associative register, get frame # out
- Otherwise get frame # from page table in memory
- Replacement policies must be considered, some entries *wired down*

Effective Access Time

- Associative lookup = ε time unit, can be < 10% of memory access time
- Assume memory access time is 1 microsecond
- Hit ratio – percentage of times that a page number is found in the associative registers; ratio related to number of associative registers
- Hit ratio = α
- **Effective Access Time** (EAT)

  EAT = $(1 + \varepsilon)\,\alpha + (2 + \varepsilon)(1 - \alpha)$

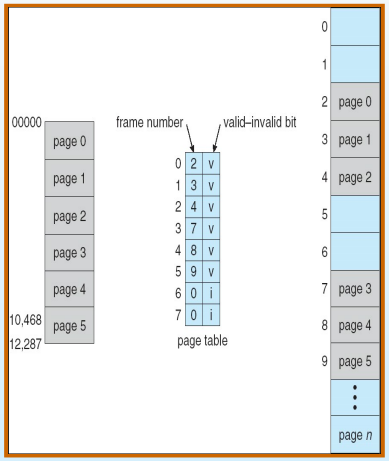
  $= 2 + \varepsilon - \alpha$

  Hit ratio 80%, and ε = 0.2 → 1,4 microseconds

  Hit ratio 98%, and ε = 0.2 → 1,22 microseconds

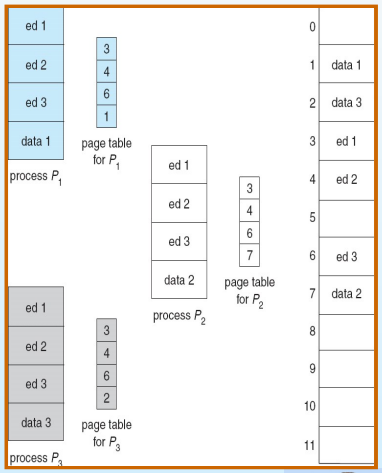
Memory Protection

- Memory protection implemented by associating protection bit with each frame: read-write vs read-only
- **Valid-invalid** bit attached to each entry in the page table:
  - "valid" indicates that the associated page is in the process's logical address space, and is thus a legal page
  - "invalid" indicates that the page is not in the process's logical address space
  - Problem with invalid addresses due to internal fragmentation
  - Page Table Length Register (PTLR)



Shared Pages

- **Shared code**
  - One copy of read-only (reentrant) code shared among processes (i.e., text editors, compilers, window systems)
  - Similar to multiple threads sharing the same process space
  - Also useful for interprocess communication if sharing of read-write pages is allowed
- **Private code and data**
  - Each process keeps a separate copy of the code and data
  - The pages for the private code and data can appear anywhere in the logical address space
- **Shared memory**
  - Could be implemented with shared pages

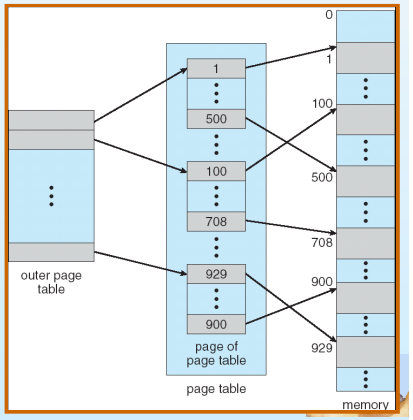


5. Structure of the Page Table

- Hierarchical Paging
- Hashed Page Tables
- Inverted Page Tables

5a. Hierarchical Page Tables

- Large logical address space 2^32 to 2^64
  - → large page table (page = 4KB → 1million entries (2^32/2^12))
  - →4MB for the table
- Break up the logical address space into multiple page tables
- A simple technique is a two-level page table
  - The page table is itself paged!

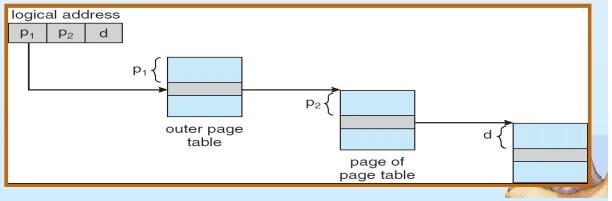


Two-Level Paging Example

- A logical address (on 32-bit machine with 1KB page size) is divided into:
  - a page number consisting of 22 bits
  - a page offset consisting of 10 bits
- Since the page table is paged, the page number is further divided into:
  - a 12-bit page number
  - a 10-bit page offset
- Thus, a logical address is as follows:

page number		page offset
p_1	p_2	d
12	10	10

where p_1 is an index into the outer page table, and p_2 is the displacement within the page of the outer page table



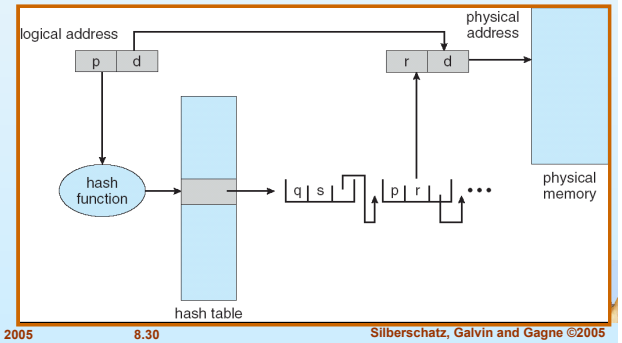
Three-level Paging Scheme

outer page	inner page	offset
p_1	p_2	d
42	10	12

2nd outer page	outer page	inner page	offset
p_1	p_2	p_3	d
32	10	10	12

5b. Hashed Page Tables

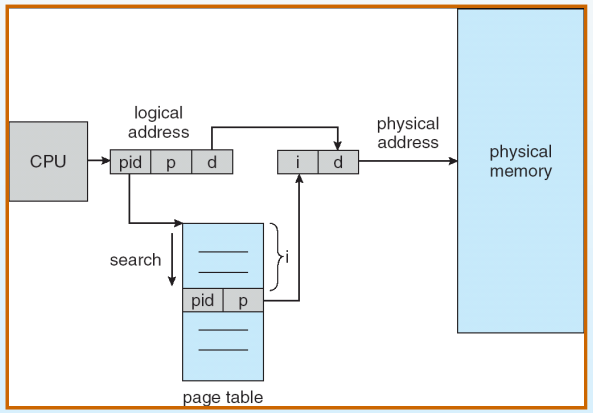
- Common in address spaces > 32 bits
- The virtual page number is hashed into a page table. This page table contains a chain of elements hashing to the same location.
- Virtual page numbers are compared in this chain searching for a match. If a match is found, the corresponding physical frame is extracted.



5c. Inverted Page Table

- One associated page table per process → large amount
- Solution: one entry for each **real** page (frame) 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

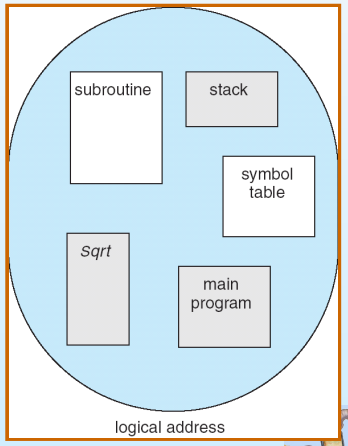
Inverted Page Table Architecture



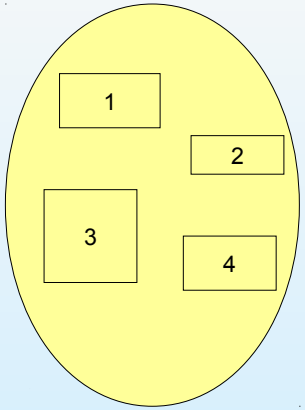
What about shared memory?

6. Segmentation

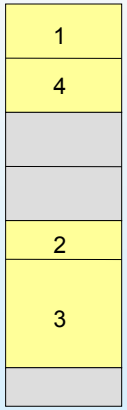
- Memory-management scheme that supports user view of memory
- A program is a collection of segments
- A segment is a logical unit such as:
  - main program
  - procedure
  - function
  - method
  - object
  - local variables, global variables
  - common block
  - stack
  - symbol table
  - arrays



Logical View of Segmentation



user space



physical memory space

Segmentation Architecture

- Logical address consists of a two tuple:

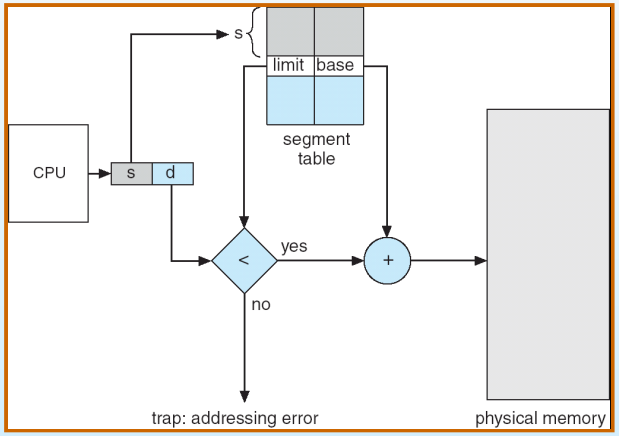
  <segment-number, offset>
- **Segment table** – maps two-dimensional physical addresses; each table entry has:
  - **base** – contains the starting physical address where the segments reside in memory
  - **limit** – specifies the length of the segment
- **Segment-table base register (STBR)** points to 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**

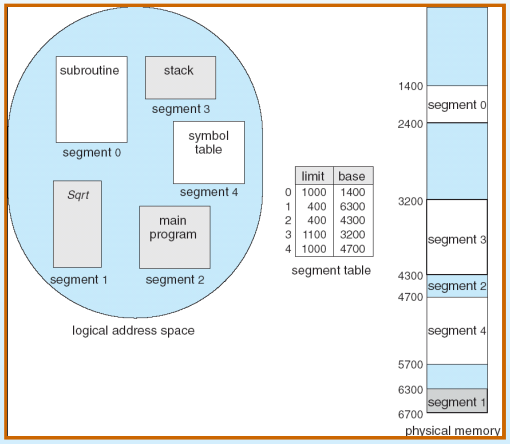
Segmentation Architecture (Cont.)

- Protection
  - With each entry in segment table associate:
    - validation bit = 0 → illegal segment
    - read/write/execute privileges
- Protection bits associated with segments; code sharing occurs at segment level
- Since segments vary in length, memory allocation is a dynamic storage-allocation problem
- A segmentation example is shown in the following diagram
- Map two-dimensional user-defined address into one-dimensional physical address == segment table

Segmentation Hardware



Example of Segmentation

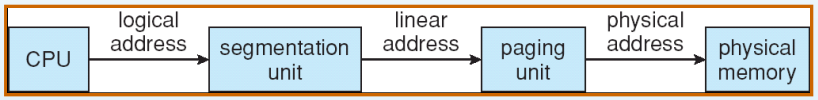


7. Example: The Intel Pentium

- Both paging and segmentation have advantages and disadvantages
- Supports both segmentation and segmentation with paging
- CPU generates logical address
  - Given to segmentation unit
    - which produces linear addresses
  - Linear address given to paging unit
    - which generates physical address in main memory

Segmentation and paging units form equivalent of MMU

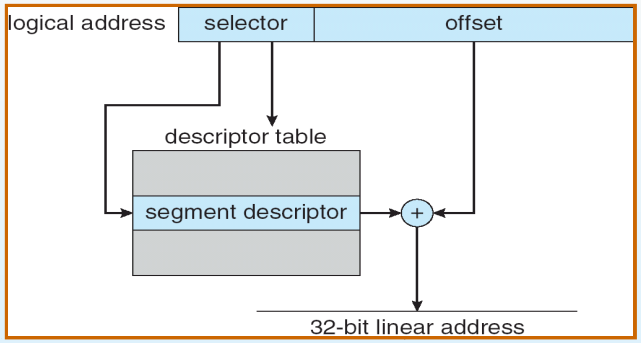
Logical to Physical Address Translation in Pentium



page number		page offset
p_1	p_2	d
10	10	12

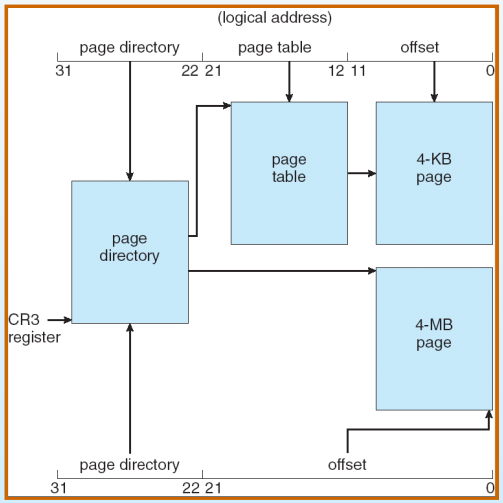
Page size: 4KB or 4MB – Two-level paging for 4-KB pages

Intel Pentium Segmentation



Segments as large as 4GB
Maximum number of segments per process: 16K
Logical space of a process divided into 2 partitions:
up to 8K segments private (LDT Local Descriptor Table)
+ up to 8K segments shared (GDT Global Descriptor Table)
Selector = s(13) g(1 – LDT/GDT) p(2 – protection)

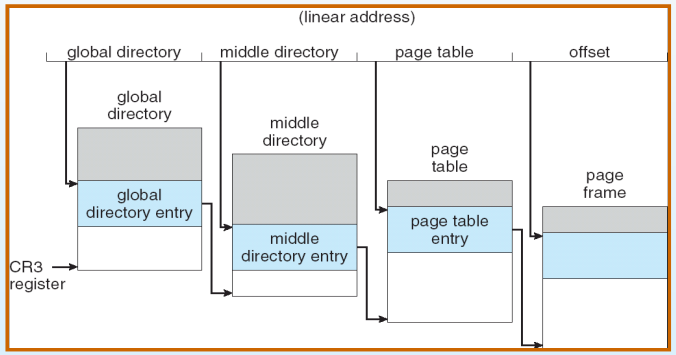
Pentium Paging Architecture



Linear Address in Linux

Broken into four parts:

Three-level Paging in Linux



On the pentium: middle directory of size 0 bits
CR3 register points to global directory for task currently executing

Résumé

- Algorithmes de gestion de la mémoire: contraints par le hardware
  - Allocation contiguë (C), pagination (P), segmentation (S)
  - Support hardware: registres de base et limite suffisant pour C, mais besoin de tables en plus pour P et S
  - Performance: temps requis pour mapper une adresse logique à une adresse physique – Utilisation de TLBs pour P et S
  - Fragmentation: 1-partition et P: fragmentation interne
    multi-partitions et S: fragmentation externe
  - Swapping: copier dans et hors de la mémoire
  - Partage: pages ou segments à partager
  - Protection: protéger le code et les données (read only, ...)
- Processus entier en mémoire avant exécution
- Utilisation de mémoire virtuelle pour ne pas avoir cette limitation!

Exercices

1) Comparer la pagination et la segmentation en terme de mémoire requise pour pouvoir traduire les adresses virtuelles en adresses physiques.

2) Un programme est généralement structuré ainsi: le code à partir d'une adresse virtuelle 0, suivi des données (variables du programme). La pile part de l'autre bout de l'espace d'adressage du processus.

   Comment se comporte l'algorithme de gestion de la mémoire avec (a) l'allocation contiguë, (b) la segmentation, et (c) la pagination?