

Gestión y administración de la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

Eloy Anguiano

Rosa M^a Carro

Ana González

Escuela Politécnica Superior
Universidad Autónoma de Madrid

Introducción

¿Qué es la memoria?

- La memoria es una **matriz de palabras o bytes direccionables** (accesibles mediante una dirección única) por la CPU para la carga de instrucciones o datos (ej: operandos) y para el almacenamiento de datos (ej: resultados).
- Las direcciones de un proceso pueden ser representadas de modo diferente en las sucesivas etapas del ciclo de un programa de usuario (compilación, carga y ejecución).
- Las direcciones que entiende el controlador de la memoria son direcciones absolutas. La conversión entre las distintas representaciones de las direcciones y las direcciones absolutas se denomina **vinculación**, y puede llevarse a cabo en cualquiera de las fases del programa.

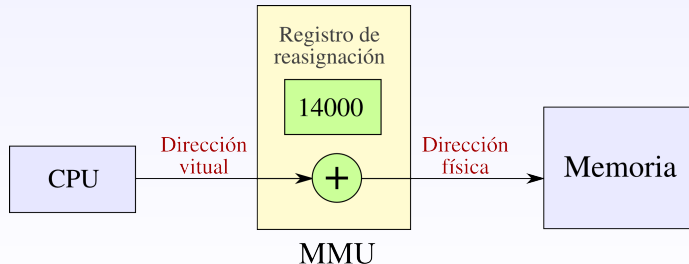
Contiene

- El sistema operativo: núcleo.
- Una zona de usuarios: imágenes de los procesos en ejecución.

Introducción

Direccionamiento

- **Dirección lógica**: Dirección generada por la CPU.
- **Dirección física**: Dirección real percibida por la unidad de memoria.
- Si la **vinculación** de direcciones (lógica-física) se realiza en **tiempo de compilación o de carga**, entonces la **dirección física = dirección lógica**.
- Si la **vinculación** de direcciones se realiza en **tiempo de ejecución** (por ser el código reasignable), las direcciones **lógicas y físicas difieren** y llamamos a las direcciones lógicas **direcciones virtuales**.





Escuela
Politécnica
Superior

Gestión y
administración de
la memoria

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

Introducción Gestión

La necesidad de la gestión de la memoria se debe a que tenemos que:

- Ubicar (cargar) los procesos en la memoria.
- Subdividir la memoria para hacer sitio a varios procesos.
- Repartir eficientemente la memoria para introducir tantos procesos como sea posible.

Introducción

Requisitos de la gestión

Gestión y
administración de
la memoria

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

1 Reubicación:

- El programador no conoce qué otros programas residirán en la memoria en el momento de la ejecución.
- Mientras que se está ejecutando el programa, puede que se descargue al disco y que vuelva después de nuevo a la memoria principal, pero en una ubicación distinta a la anterior (reubicación).
- Se deben traducir las referencias a la memoria encontradas en el código del programa a las direcciones físicas reales.

Introducción

Requisitos de la gestión

Gestión y
administración de
la memoria

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

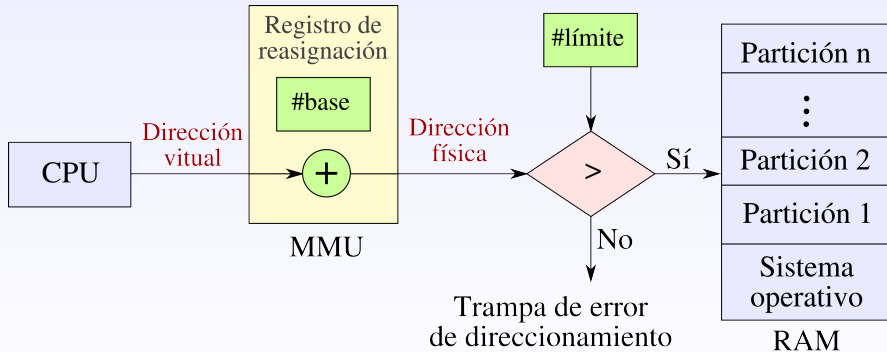
2 Protección:

- El código de un proceso no puede hacer referencia a posiciones de memoria de otros procesos sin permiso.
- Es imposible comprobar las direcciones absolutas de los programas, puesto que se desconoce la ubicación de un programa en la memoria principal.
- Debe comprobarse durante la ejecución:
 - El sistema operativo no puede anticiparse a todas las referencias a la memoria que hará un programa.

Introducción

Requisitos de la gestión

- La protección de memoria se puede realizar utilizando dos registros (base=reasignación y límite) que controlan el acceso a la memoria física.
- Sólo el SO puede modificar los registros de base y límite.**



Introducción

Requisitos de la gestión

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

3 Compartición:

- Permite el acceso de varios procesos a la misma zona de la memoria principal.
 - Para cooperación es necesario tener acceso compartido a estructuras de datos.
 - Es mejor permitir a cada proceso (persona) que acceda a la misma copia del programa, en lugar de tener cada uno su propia copia aparte.

Introducción

Requisitos de la gestión

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

4 Organización lógica:

- La memoria está organizada como un espacio de direcciones unidimensional
- La mayoría de los programas se organizan en módulos.
 - Los módulos pueden escribirse y compilarse independientemente.
 - Pueden otorgarse un grado distinto de protección a cada uno de los módulos (sólo lectura, sólo ejecución).
 - Se pueden compartir módulos.

Introducción

Requisitos de la gestión

Gestión y
administración de
la memoria

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

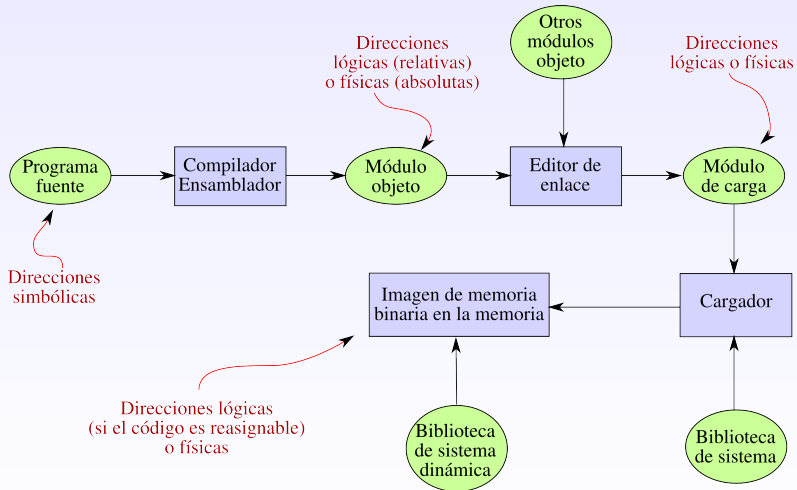
Particiones
dinámicas

Paginación

Segmentación

Intercambio

Proceso en etapas de un programa de usuario



Introducción

Requisitos de la gestión

Introducción

¿Qué es la memoria?

Direccionamiento

Gestión

Requisitos de la
gestión

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación

Intercambio

5 Organización física:

- **Memoria principal:** rápida, costosa, pequeña, volátil
- **Memoria secundaria:** más lenta, barata, gran capacidad, no volátil, ...
- El flujo entre ambas es responsabilidad del S.O. (antes lo realizaba el programador):
 - La memoria disponible para un programa y sus datos puede ser insuficiente: la superposición es una técnica que permite que varios módulos sean asignados a la misma región de memoria.
 - El programador en principio no conoce el espacio disponible.

Técnicas de gestión de memoria

La tarea principal del sistema de gestión de memoria es **la carga de programas en memoria principal para su ejecución en el procesador** (ubicar las imágenes de los procesos en memoria principal)

Técnicas simples

- Particionamiento (utilizada con distintas variantes en antiguos SO)
- Paginación simple
- Segmentación simple

Memoria Virtual (próximo tema)

- Sistemas multiprogramados modernos
- Basada en segmentación y paginación

Particiones estáticas

Condiciones

- El Sistema Operativo ocupa una parte fija de la memoria
- El resto está disponible para ser usada por los procesos:
 - Límite en el número de procesos listos, bloqueados o en ejecución, determinado por el número de particiones existentes.
- Tamaños de las particiones y algoritmos de ubicación:
 - Particiones de igual tamaño
 - Particiones de distinto tamaño

Particiones estáticas

De igual tamaño

Condiciones

- Cualquier proceso cuyo tamaño sea menor o igual que el tamaño de la partición puede cargarse en cualquier partición libre.
- Si todas las particiones están ocupadas, el sistema operativo puede sacar un proceso de una partición (suspensión).

Problemas

- Un programa puede que no se ajuste a una partición (porque sea mayor). El programador debe diseñar el programa mediante superposiciones.
- El uso de la memoria principal es ineficiente. Cualquier programa, aunque sea pequeño, ocupará una partición completa. Se produce el fenómeno de la **fragmentación interna**: existen espacios de memoria no utilizados, dentro de una partición, que no están disponibles para poderse utilizar/asignar.

Sistema operativo 8 MB
8 MB
8 MB
8 MB
8 MB
8 MB

Particiones estáticas

De distinto tamaño

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

De igual tamaño
De distinto tamaño
En general

Particiones
dinámicas

Paginación

Segmentación

Intercambio

Reducen los problemas

- Se pueden alojar programas más grandes sin superposición.
- Se reduce el desperdicio de memoria producido por programas pequeños.

Problemas

- Caben menos programas grandes sin superposición.
- Si hay muchos programas pequeños la fragmentación interna es muy grande en las particiones grandes.

Dos formas de gestionarlas:

- Una cola por partición.
- Una cola única.

Sistema
operativo
8 MB

2 MB

4 MB

6 MB

8 MB

8 MB

12 MB

Particiones estáticas

De distinto tamaño

Una cola por partición

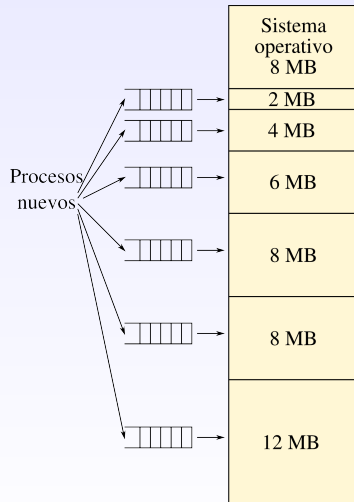
Método: Ubicar cada proceso en la partición de tamaño **menor** en la que quepa. Si la partición no está libre, espera en la cola correspondiente. Hace falta una cola para cada partición.

Ventaja

Se minimiza la memoria desaprovechada dentro de cada partición (fragmentación interna)

Desventaja

Pueden existir procesos suspendidos esperando en colas correspondientes a las particiones que mejor se ajustan a su tamaño, habiendo otras particiones vacías.



Particiones estáticas

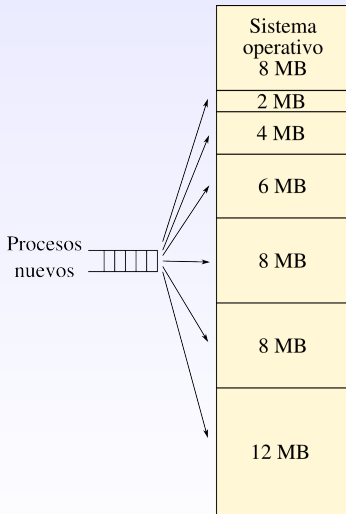
De distinto tamaño

Una cola única

Método: Cuando se va a cargar un proceso, se selecciona la partición **más pequeña disponible** donde quepa.

Si todas las particiones están ocupadas, los criterios para suspender otro proceso se relacionan con:

- Usar la partición más ajustada (minimiza la fragmentación interna).
- Considerar las prioridades de procesos ya cargados.
- Suspender a los procesos bloqueados antes que a los procesos listos.



Particiones estáticas

En general

Desventajas

- El número y el tamaño de las particiones especificadas al principio limitan el número y el tamaño máximo de procesos activos (no suspendidos)
- Si hay muchos procesos pequeños no se aprovecha bien la memoria (fragmentación interna)
- Si hay muchos procesos grandes, las particiones pequeñas pueden resultar infrutilizadas.

Actualmente no se usan particiones estáticas

Particiones dinámicas

Definición

Introducción

Técnicas de gestión de memoria

Particiones estáticas

Particiones dinámicas

Definición

Fragmentación
externa

Fragmentación
interna

Ubicación con mapa
de bits

Ubicación con listas
enlazadas

Ubicación por ajuste

Sistema de colegas

Paginación

Segmentación

- Las particiones son variables en número y longitud.
- Al proceso se le asigna exactamente tanta memoria como necesite.

Particiones dinámicas

Fragmentación externa

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

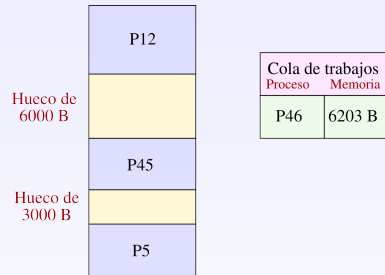
Particiones
dinámicas

Definición
Fragmentación
externa
Fragmentación
interna
Ubicación con mapa
de bits
Ubicación con listas
enlazadas
Ubicación por ajuste
Sistema de colegas

Paginación

Segmentación

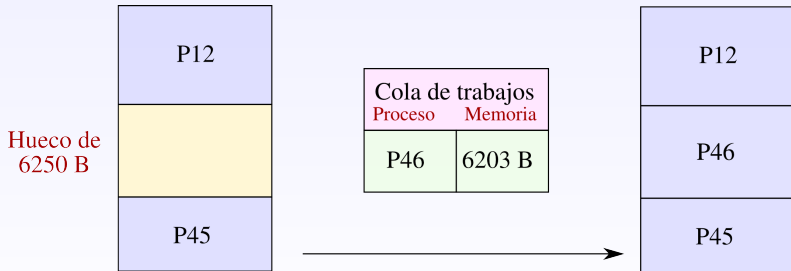
- En un sistema de particiones dinámicas, tras la reubicación de procesos, se van generando huecos en memoria entre las distintas particiones que, unidos, serían útiles para albergar procesos, pero, separados, no sirven. Este fenómeno se denomina **fragmentación externa**.
- Se debe usar la **compactación** para desplazar los procesos de modo que queden situados en particiones contiguas (normalmente en el comienzo de la memoria) y toda la memoria libre quede junta en un único bloque grande (normalmente al final de la memoria).



Particiones dinámicas

Fragmentación interna

- Es el desperdicio de memoria que se produce al asignar a un proceso más memoria de la que solicita (por conveniencia o simplicidad en la asignación).
- No es frecuente encontrar fragmentación externa en sistemas de gestión de memoria basados en particiones dinámicas, pues normalmente se asigna a cada proceso justo el espacio que necesita.



Particiones dinámicas

Ubicación con mapa de bits

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Definición

Fragmentación
externa

Fragmentación
interna

Ubicación con mapa
de bits

Ubicación con listas
enlazadas

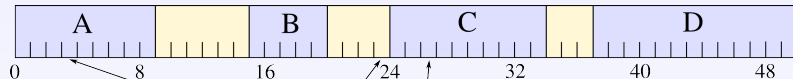
Ubicación por ajuste

Sistema de colegas

Paginación

Segmentación

- La memoria principal está dividida en unidades de asignación (bloques).
- Para conocer qué bloques están libres/ocupados, se utiliza un mapa de bits.
- Cada bloque se representa mediante un bit en el mapa de bits
 - Compromiso entre el tamaño del mapa y la optimización de la gestión.
- La asignación de espacio a un proceso de tamaño k unidades consiste en una búsqueda de k ceros contiguos en el mapa de bits: LENTO.



11111111
10000001
11110000
11111111
11000111
11 ...

Unidad de
asignación

Mapa de Bits

Particiones dinámicas

Ubicación con listas enlazadas

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

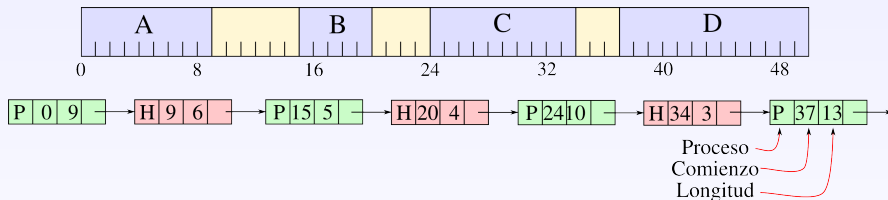
Particiones
dinámicas

Definición
Fragmentación
externa
Fragmentación
interna
Ubicación con mapa
de bits
**Ubicación con listas
enlazadas**
Ubicación por ajuste
Sistema de colegas

Paginación

Segmentación

- La memoria está representada por una lista ligada de zonas de memoria ocupadas (P) y libres (H).



Particiones dinámicas

Ubicación con listas enlazadas

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Definición

Fragmentación
externa

Fragmentación
interna

Ubicación con mapa
de bits

Ubicación con listas
enlazadas

Ubicación por ajuste

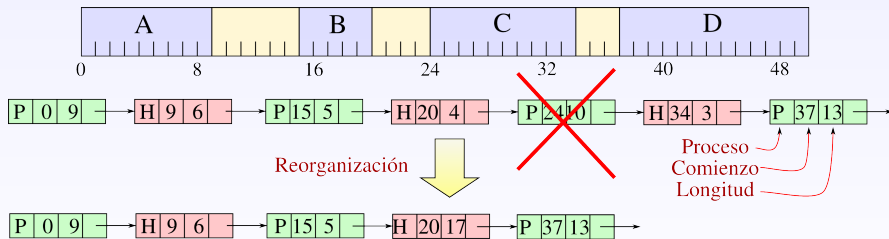
Sistema de colegas

Paginación

Segmentación

Ventajas

- La búsqueda de huecos es más rápida, puesto que cada elemento de la lista contiene el tamaño del mismo.
- Es fácil reorganizar la lista al terminar un proceso o al introducir un proceso nuevo.



Particiones dinámicas

Ubicación por ajuste

Best-fit

- Elige el bloque de tamaño más próximo al solicitado.
- Proporciona en general los peores resultados.
- Como este algoritmo busca el hueco más pequeño para el proceso, garantiza que el fragmento que se deja es lo más pequeño posible y, por lo tanto, suele ser insuficiente en el futuro y se debe compactar más frecuentemente.

First-fit

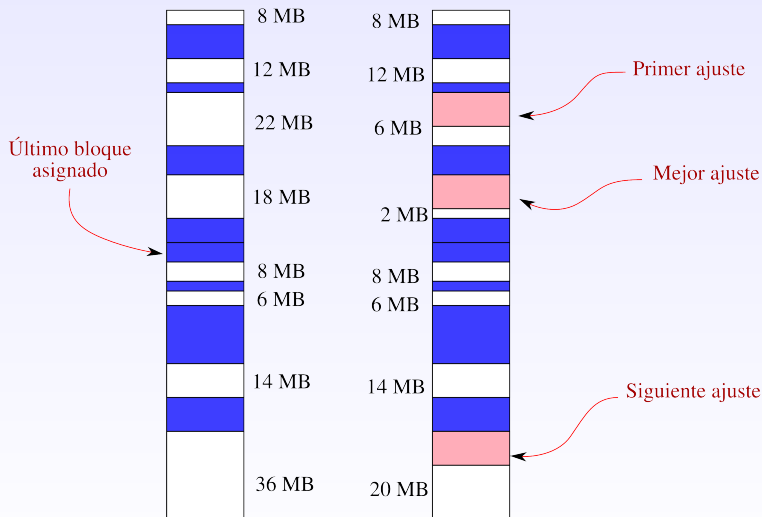
- Es más rápido.
- Puede tener varios procesos cargados en el extremo inicial de la memoria que es necesario recorrer cuando se intente encontrar un bloque libre.

Next-fit

- Lleva frecuentemente a la asignación de un bloque de memoria de la última ubicación, donde se encuentra el bloque más grande.
- El bloque de memoria más grande se divide en fragmentos pequeños.
- Hará falta la compactación para obtener un bloque de memoria grande al final del espacio de memoria.

Particiones dinámicas

Ubicación por ajuste



Particiones dinámicas

Sistema de colegas

Introducción

Técnicas de gestión de memoria

Particiones estáticas

Particiones dinámicas

Definición

Fragmentación
externa

Fragmentación
interna

Ubicación con mapa
de bits

Ubicación con listas
enlazadas

Ubicación por ajuste

Sistema de colegas

Paginación

Segmentación

- El espacio entero disponible para la asignación se trata como un solo bloque de tamaño 2^U .
- Si se hace una solicitud de tamaño s tal que $2^{U-1} < s \leq 2^U$, entonces el bloque entero se asigna:
 - En otro caso, el bloque se divide en dos colegas de igual tamaño.
 - Este proceso continúa mientras que el bloque más pequeño sea mayor o igual que s .

Particiones dinámicas

Sistema de colegas

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Definición

Fragmentación
externa

Fragmentación
interna

Ubicación con mapa
de bits

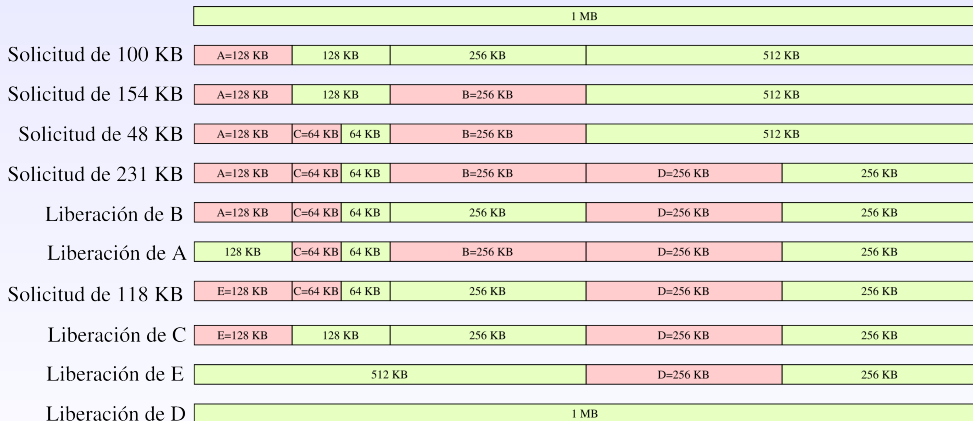
Ubicación con listas
enlazadas

Ubicación por ajuste

Sistema de colegas

Paginación

Segmentación



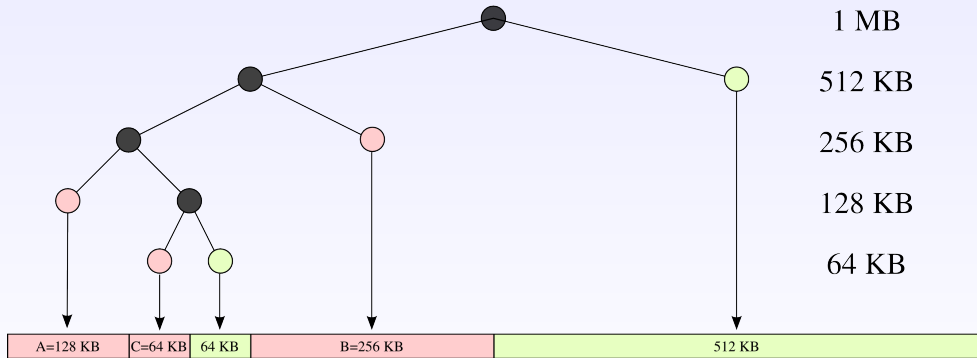
(Errata: Color del hueco que deja B al liberarse: verde desde que se libera)

Particiones dinámicas

Sistema de colegas

Gestión y
administración de
la memoria

Representación en árbol del sistema de colegas



Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Definición
Fragmentación
externa
Fragmentación
interna
Ubicación con mapa
de bits
Ubicación con listas
enlazadas
Ubicación por ajuste
Sistema de colegas

Paginación

Segmentación

Paginación

Definición

- La memoria principal se encuentra dividida en pequeñas partes iguales de tamaño fijo. Cada una de esas partes se denomina **marco**.
- La imagen de cada proceso se divide (desde el punto de vista lógico) en pequeñas partes iguales de tamaño fijo e igual al de los marcos. Cada una de esas partes se denomina **página**.
- Los marcos de memoria principal albergan ('envuelven') las páginas.
- El sistema operativo mantiene una **tabla de páginas** para cada proceso:
 - Cada entrada en esa tabla se corresponde con una página del proceso e indica el número del marco donde se encuentra dicha página en memoria principal.
- Una **dirección lógica** consta de un número de página y de un desplazamiento dentro de la página.
- Una **dirección física** consta de un número de marco y de un desplazamiento dentro de ese marco.



Tiempo 

Quince marcos libres	Carga proceso A	Carga proceso B	Carga proceso C	Descarga proceso B	Carga proceso D
0	A.0	A.0	A.0	A.0	A.0
1	A.1	A.1	A.1	A.1	A.1
2	A.2	A.2	A.2	A.2	A.2
3	A.3	A.3	A.3	A.3	A.3
4		B.0	B.0		D.0
5		B.1	B.1		D.1
6		B.2	B.2		D.2
7			C.0	C.0	C.0
8			C.1	C.1	C.1
9			C.2	C.2	C.2
10			C.3	C.3	C.3
11					D.3
12					D.4
13					
14					

Intercambio

Paginación

Asignación de páginas a marcos libres

Gestión y
administración de
la memoria

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Definición

Asignación de
páginas a marcos
libres

Segmentación

Intercambio

0	A.0
1	A.1
2	A.2
3	A.3
4	D.0
5	D.1
6	D.2
7	C.0
8	C.1
9	C.2
10	C.3
11	D.3
12	D.4
13	
14	

Estructura
de datos
correspondiente

0	0	0	-	0	7	0	4
1	1	1	-	1	8	1	5
2	2	2	-	2	9	2	6
3	3			3	10	3	11
						4	12

Segmentación

Definición

- La imagen de un proceso está formada por múltiples segmentos.
- No es necesario que todos los segmentos de todos los programas tengan la misma longitud.
- Cada segmento se almacena entero de forma contigua en memoria principal.
- Existe una longitud máxima de segmento.
- En un sistema con gestión de memoria basada en segmentación, una **dirección lógica** consta de dos partes: un número de segmento y un desplazamiento.
- Se crea una tabla de segmentos para cada proceso.
 - La tabla tiene una entrada por cada segmento del proceso.
 - Cada entrada de la tabla indica, para el segmento correspondiente, la dirección base donde comienza en memoria principal y su tamaño.

Segmentación

Definición

Como consecuencia del empleo de segmentos de distinto tamaño, la segmentación resulta similar a la partición dinámica.

- Un proceso puede ocupar más de un segmento.
- En memoria no tienen porque estar contiguos los distintos segmentos del proceso.
- No se produce fragmentación interna
- Sí se puede producir fragmentación externa, aunque menor que con las particiones dinámicas.

Segmentación

Definición

Introducción

Técnicas de
gestión de
memoria

Particiones
estáticas

Particiones
dinámicas

Paginación

Segmentación
Definición

Intercambio

Proceso de carga

- Se divide el programa en segmentos
- Se cargan todos sus segmentos en bloques libres de memoria
- Se actualiza la lista de bloques libres

Intercambio

Definición

- En determinadas circunstancias, un proceso que está en memoria puede ser almacenado temporalmente (intercambiado) en un almacenamiento auxiliar (normalmente en disco magnético), de modo que el planificador pueda asignar su espacio de CPU a otro proceso situado en el almacenamiento auxiliar.
- Normalmente, un proceso intercambiado a disco se intercambia de regreso a la mismo espacio de memoria que ocupó anteriormente. Esto es **obligatorio si la vinculación de direcciones se realizó en tiempo de compilación o carga**, no así si la vinculación de direcciones se realiza en **tiempo de ejecución**.
- El sistema de almacenamiento debe tener espacio suficiente para alojar las imágenes de memoria de todos los procesos y proporcionar un acceso rápido a las mismas.

Intercambio

Definición

- Los procesos listos que están en disco pueden estar intercalados con los procesos listos en memoria en la lista de procesos listos del planificador o estar en una cola de segundo nivel para ser intercambiados a disco en grupos cada cierto tiempo.
- No deben realizarse intercambios de procesos que tengan pendientes operaciones de E/S que requieran el uso de buffers en el espacio de memoria propio (OK si utilizan buffers del SO).
- La duración del cambio de contexto desde disco depende principalmente del tamaño de la zona de memoria a intercambiar (la zona de intercambio es un área independiente del sistema de archivos que permite un acceso rápido a la información). En cualquier caso es superior al cambio de contexto entre procesos que están en memoria.
- Los algoritmos de asignación de espacio en disco para intercambios son los mismo que los explicados para la asignación de memoria.