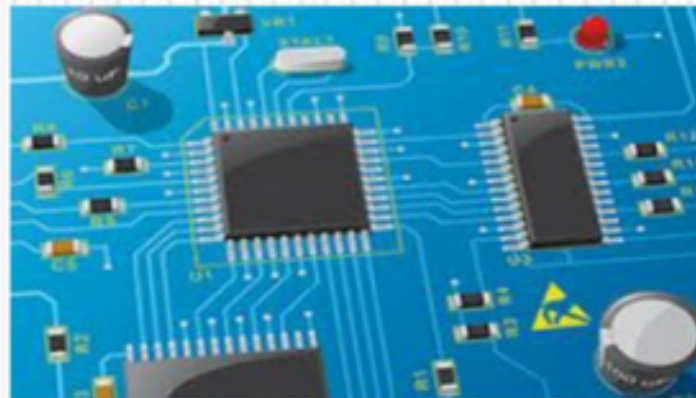




T3. Sistema de Memoria:

3.1 Introducción a la Gestión de Memoria

FUNDAMENTOS DE ARQUITECTURA DE COMPUTADORES



Contenido del capítulo

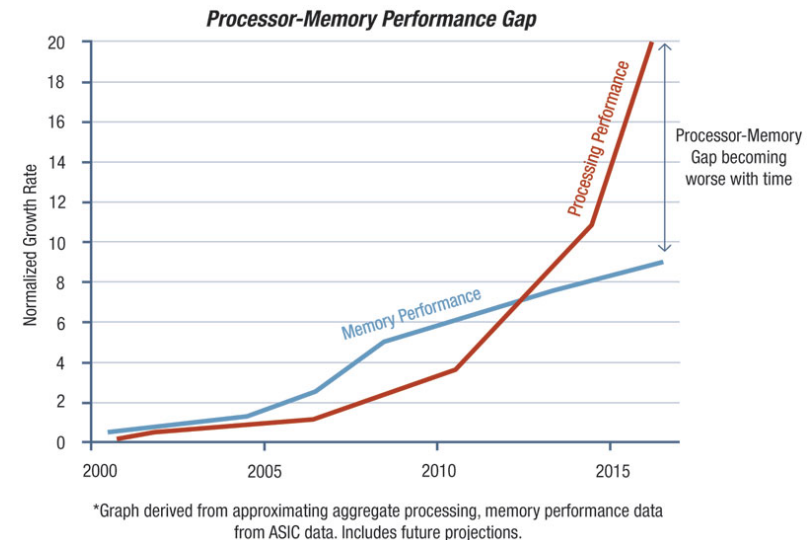
- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso.
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

Introducción

- Los puntos principales que marcan el rendimiento del computador son memoria y microprocesador.
- Mientras que la velocidad del procesador se duplica cada dos años, la velocidad de las memorias apenas crece un 10% anual.
- La arquitectura de computadores establece una **jerarquía de sistemas de memoria** tanto internos como externos para ofrecer al micro los datos cuanto antes.
- En este tema:
 - Estudiaremos conceptos básicos relativos a la gestión de memoria.
 - Analizaremos principales componentes de almacenamiento.



Unidades Básicas



- **Palabra:** Es la unidad “natural” de organización de la memoria.
- **Unidades direccionables:**
 - En muchos sistemas la unidad direccionable es la palabra.
 - Algunos de ellos permiten direccionar a nivel de bytes, dobles o cuádruples palabras.
 - La relación entre el numero N de unidades direccionables, y la longitud A de una dirección, es $N = 2^A$.
- **Unidad de transferencia:** *NO* tiene que coincidir con unidad direccionable.
 - Para la *memoria más interna*, es el numero de bits que se leen o escriben desde o hacia la CPU, usualmente una palabra.
 - Para la *memoria externa*, los datos se transfieren normalmente en unidades más grandes que la palabra, denominadas bloques.

Antes de nada... Cuestiones prácticas!

- Considere un hipotético computador con 100 registros que direcciona la memoria por bytes.
- En este computador se pueden direccionar como mucho 64KiB de memoria.
- Asumiendo que el tamaño de la palabra de este computador coincide con el número de bits empleados para las direcciones de memoria, responde a las siguientes preguntas:
 - ¿Cuántos bits se emplean para las direcciones de memoria?
 - ¿Cuál es el tamaño de los registros?
 - ¿Cuántos bits se almacenan en cada posición de memoria?
 - Si el computador fuera de 32 bits, ¿cuál sería el tamaño de la máxima memoria direccionable (en MiB)?
 - ¿Cuántos bits se necesitan para identificar a los registros?

Contenido del capítulo

- Introducción: Conceptos generales y unidades básicas
- **Métodos de acceso**
- Rendimiento y temporización
- Jerarquía de memoria
 - Introducción
 - Taxonomía Complementaria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria

Métodos de Acceso

- **Acceso secuencial**

- El tiempo necesario para acceder a un registro dado es muy variable.
- Unidades de *cinta*

- **Acceso directo**

- Acceso directo a una vecindad
→ búsqueda secuencial.
- Tiempo de acceso variable.
- Discos *duros*

- **Acceso aleatorio**

- El tiempo de acceso es constante e independiente de la secuencia previa.
- Memoria *principal*

- **Memoria asociativa**

- Realiza una comparación de ciertas posiciones de bits dentro de una palabra buscando que coincidan con unos valores dados (**clave**).
- Memoria *caché*

Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

Rendimiento y Temporización

- **Latencia:**

- Es el tiempo de acceso a la memoria.
- Normalmente mide el **tiempo de repuesta** desde que se le envía una dirección hasta que devuelve el dato que hay en ella.
- Se mide como el tiempo de ciclo o tiempo que transcurre entre que la memoria acepta una petición y queda lista para aceptar la siguiente.

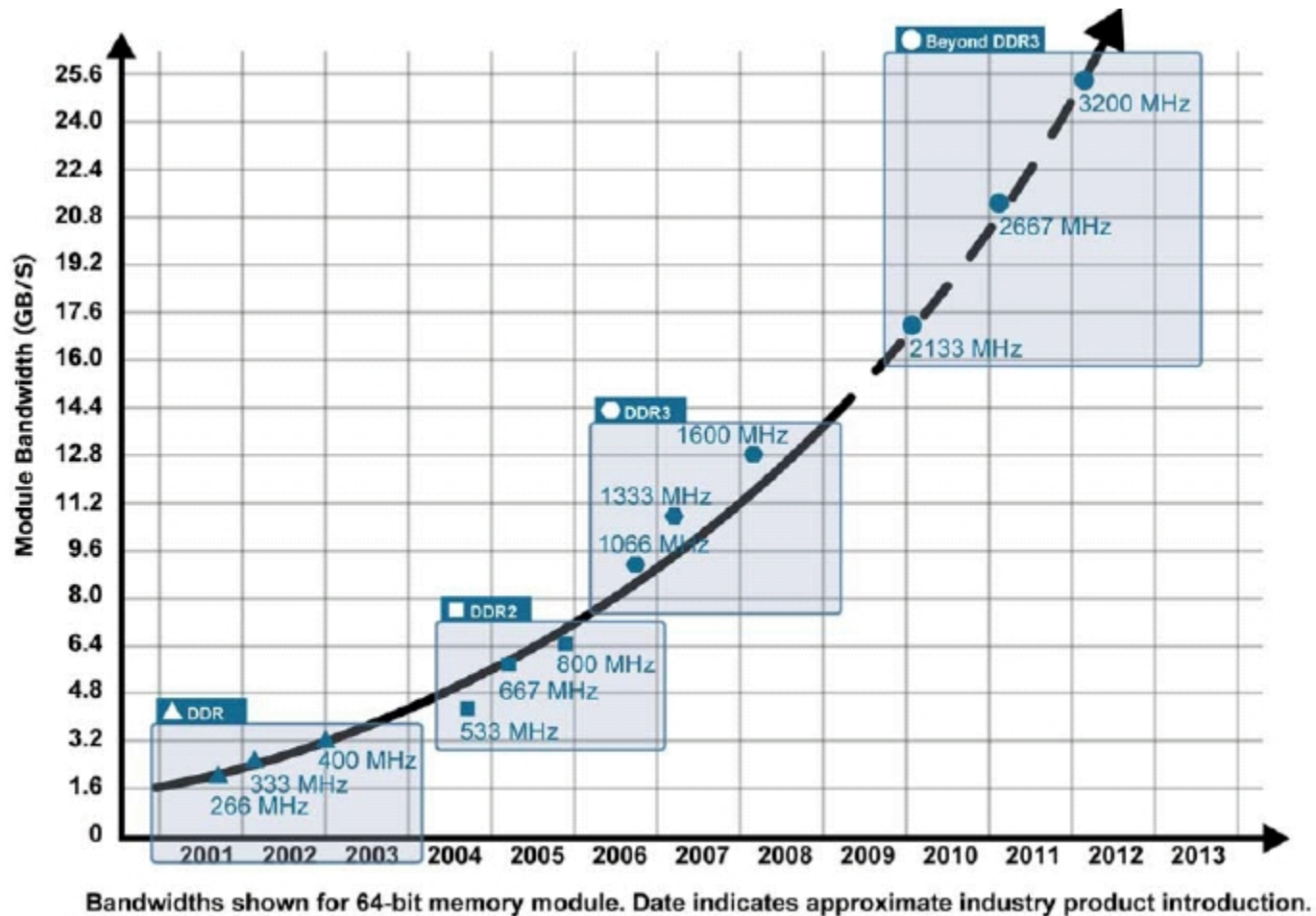
- **Velocidad del bus:**

- Frecuencia de envío (MHz).
- Número de veces que un grupo de bits son enviados por segundo.

- **Ancho de banda:**

- Velocidad de transporte de datos entre memoria y dispositivos.
- Depende de elementos externos y suele medirse en Mbytes/s.

Rendimiento y temporización

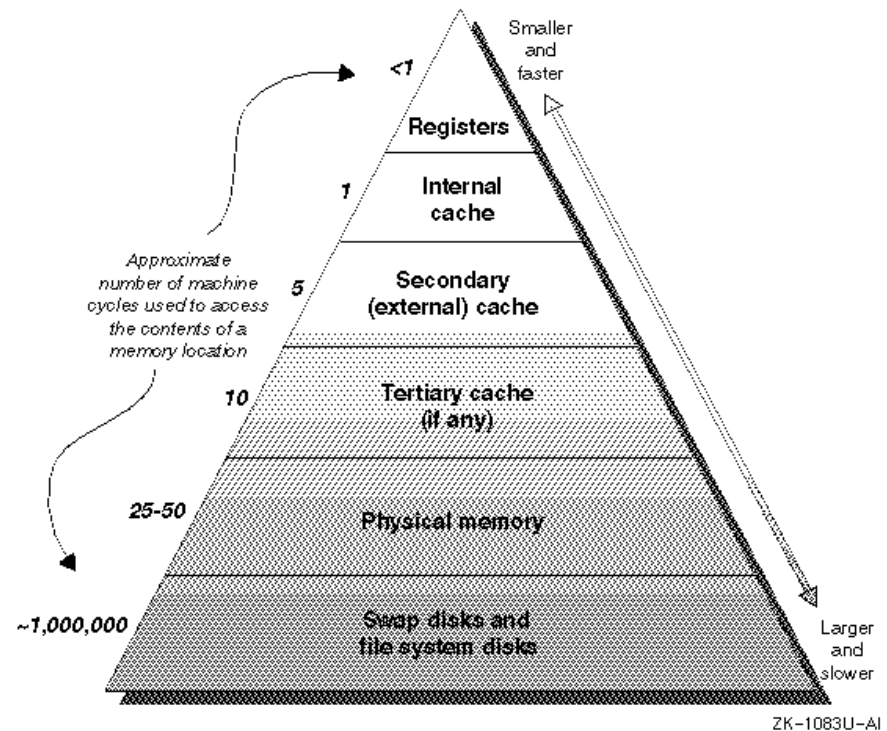


Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

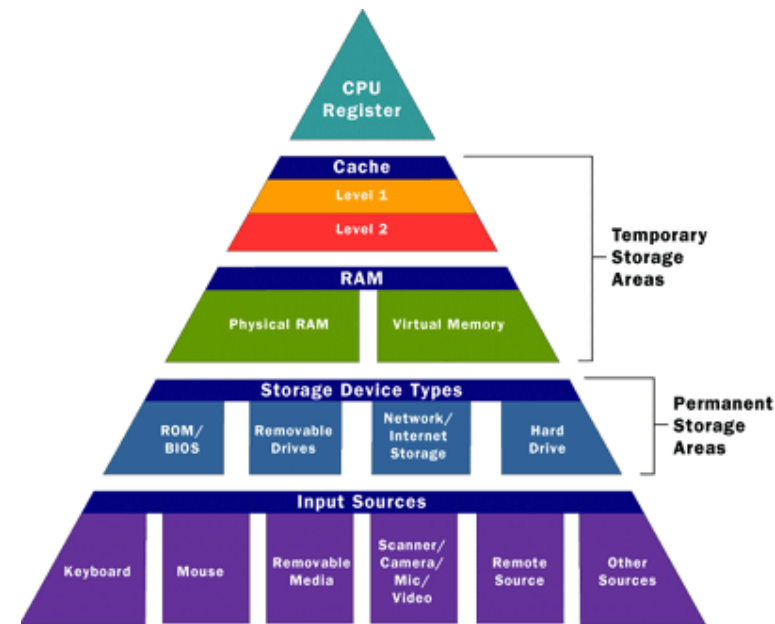
Jerarquía de memoria

- La memoria se organiza en una jerarquía, según **velocidades**.
- A rapidez creciente, la cantidad de información por unidad de volumen **decrece**, y el precio **aumenta**.
- Generalmente, los datos contenidos en cada nivel son un **subconjunto** de los datos en los niveles superiores.



Jerarquía memoria (2): Taxonomía Comp.

- **Almacenamiento Primario:**
 - Está directamente conectada a la CPU de la computadora.
 - Debe estar presente para que la CPU funcione correctamente.
- **Secundario**-“almacenamiento masivo”:
 - Requiere que el uso de canales de E/S para acceder a la información.
 - Se utiliza para almacenamiento a largo plazo de información persistente.
 - SSOO: área de intercambio (Swap) y Memoria Virtual.



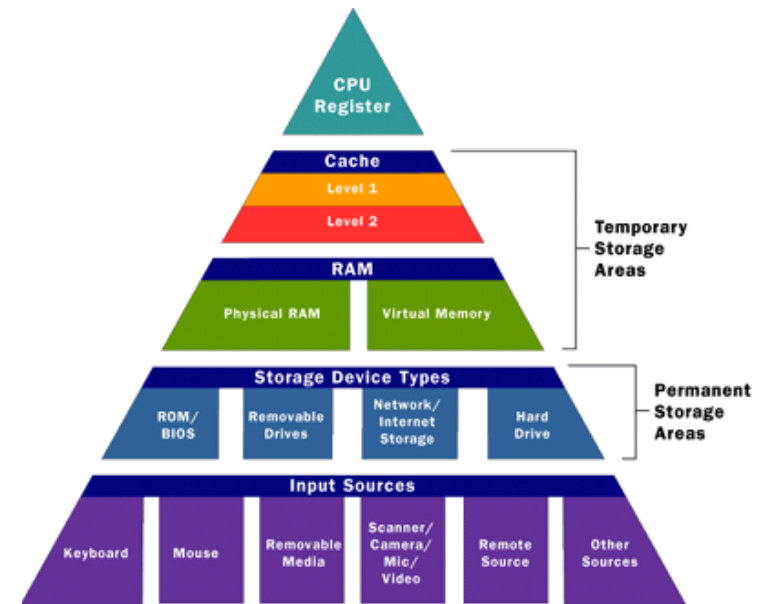
Jerarquía memoria (3): Taxonomía Comp

- **Fuera de línea:**

- El medio de almacenamiento puede ser extraído fácilmente del dispositivo.
- Transporte y archivo de datos.
- Discos ópticos y memorias flash.

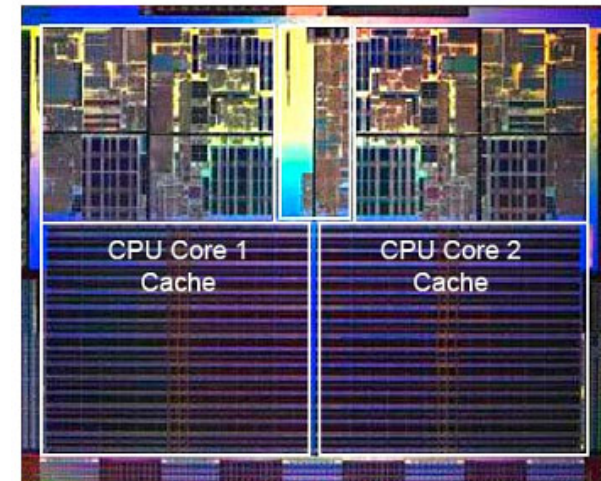
- **De red:**

- Acceso a la información a través de una red informática.
- Permite centralizar el control de información y reducir la duplicidad de la información.



Almacenamiento Primario

- **Registros del procesador:**
 - Contienen información que las unidades de la ALU necesitan llevar a la instrucción en ejecución
- **Memoria caché:**
 - Es un tipo especial de memoria usada en muchas CPU para mejorar su eficiencia o rendimiento.
 - Dividida en “niveles”. Se integra en el propio circuito del procesador (niveles L1 y “L2”).
- **Memoria principal:**
 - Contiene programas en ejecución y datos con los que operan.
 - La ALU puede transferir información muy rápidamente entre un registro del microprocesador y "direcciones de memoria".



Almacenamient. Secund. y Fuera de Línea

- **Discos Magnéticos:**

- Disco duro, disquete, cintas magnéticas...

- **Discos Ópticos:**

- CD, DVD, Blu Ray, etc.
- Algunos dispositivos combinan ambas tecnologías, es decir, son dispositivos de almacenamiento híbridos, por ej., discos Zip.

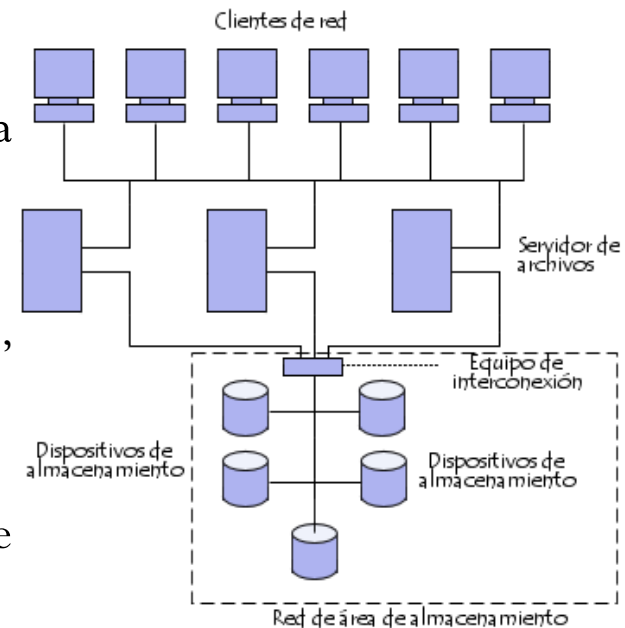
- **Memoria Flash:**

- Tarjetas de Memorias Flash y Unidades de Estado sólido SSD.



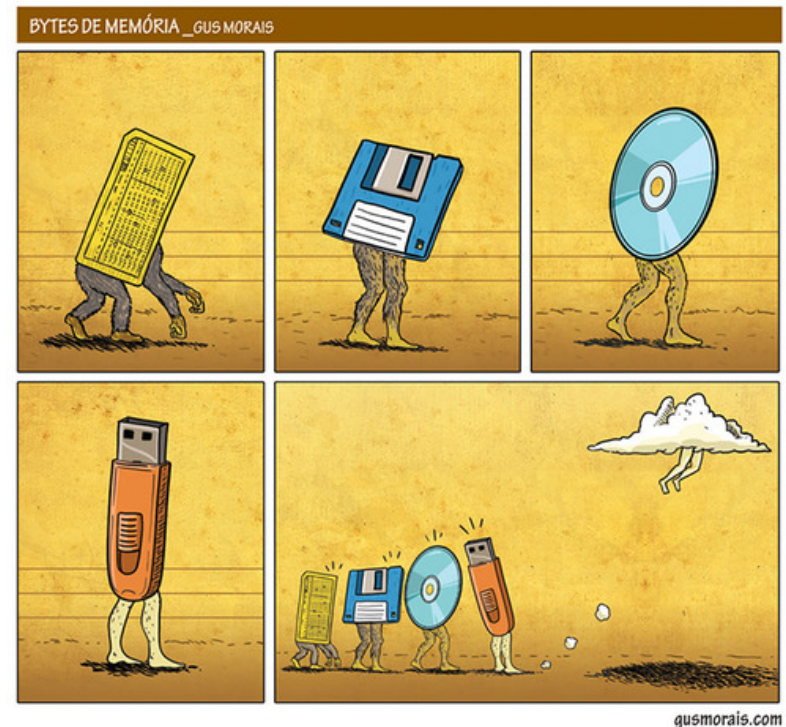
Almacenamiento de Red

- Una "SAN" (Storage Area Network) es una red de almacenamiento integral.
- Se trata de una arquitectura completa que agrupa los siguientes elementos:
 - Red de alta velocidad: canal de fibra o SCSI
 - Equipo de interconexión dedicado (conmutadores, puentes, etc.)
 - Elementos de almacenamiento (discos duros)
- Rendimiento de la SAN está directamente relacionado con el tipo de red.
- Su capacidad se puede extender de manera casi ilimitada y puede alcanzar cientos y hasta miles de terabytes.



Cloud Computing

- Nuevo paradigma de computación
- Los recursos informáticos se encuentran en máquinas remotas. Su uso “se alquila”
- Se puede tener acceso a almacenamiento o tiempo de máquina.
- Cloud público: se evitan las labores de compra, actualización y mantenim. del servidor
- Mucho más práctico que una red SAN



Cloud Computing (2)

- Tres claros niveles:
 - IaaS: Infraestructura
 - PaaS: Plataforma
 - SaaS: Software
- Ejemplos Almacenam:
 - Dropbox,
 - Google Drive,
 - SkyDrive
 - Box...



Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- **Modificación y volatilidad de la información**
- Ejemplo del uso de la memoria
- Bibliografía

Modificación de la información

- Memoria de lectura/escritura:
 - Permiten que la información se reescriba en cualquier momento
 - Memoria RAM ó discos duros
- Memoria de sólo lectura:
 - Retienen la información almacenada en el momento de fabricarse
 - Memoria ROM (PROM, EEPROM, CD-ROM...)
- Memoria de escritura lenta/lectura rápida:
 - La información se reescribe múltiples veces. Velocidad de escritura mucho menor que la de lectura.
 - Discos ópticos re-escribibles: CD-RW

Volatilidad de la Información

- Memoria volátil:
 - Requiere energía constante para mantener la información almacenada.
 - Se suele usar sólo en memorias primarias.
 - La memoria RAM es una memoria volátil, ya que pierde información en la falta de energía eléctrica.
 - Existen dos tipos de memoria volátil:
 - Memoria dinámica (DRAM): periódicamente se refresca la información almacenada.
 - Memoria estática (SRAM): mantiene los datos si no se interrumpe corriente eléctrica.
- Memoria no volátil.
 - Retendrá la información almacenada incluso si no recibe corriente eléctrica constantemente, como es el caso de la memoria ROM.
 - Se usa para almacenamientos a largo plazo y, por tanto, se usa en memorias secundarias y fuera de línea.

Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

Ej. de Funcionamiento de la Memoria

- El ordenador carga datos desde una **unidad ROM** y realiza un “power-on self-test” (**POST**)
 - Chequear que los componentes principales funcionan correctamente.
 - El controlador de memoria comprueba mediante operaciones de R/W todas las direcciones de memoria en busca de errores en los chips.
- Carga la **BIOS** (basic input/output system) desde la **ROM**.
 - La BIOS proporciona la información más básica sobre dispositivos de almacenamiento, secuencia de arranque, seguridad, etc.
- El ordenador carga el S.O. desde el **HDD** a la **RAM** del sistema.
- Los apartados críticos del SO se mantienen en la RAM mientras que el ordenador esté encendido.

Ej. Funcionamiento de la Memoria (2)

- Toda aplicación que se ejecute, se carga en la RAM.
- Para conservar espacio en RAM, muchas aplicaciones sólo cargan las partes esenciales del programa inicialmente, dejando el resto para cuando se necesiten.
- Cualquier fichero que se necesite dicha aplicación se deberá cargar en la RAM.
- Al guardar un fichero, el fichero se escribe en el dispositivo de almacenamiento especificado, tardando un tiempo.
- Al cerrar la aplicación, la memoria que haya sido reservada se puede “eliminar” de la RAM.
- Existe una memoria intermedia (llamada caché) que guarda la información más usada de la aplicación para acceder de forma mucho más rápida y que la ejecución sea más eficiente.

Contenido del capítulo

- Introducción: Conceptos generales, unidades básicas y métodos de acceso
- Rendimiento y temporización
- Jerarquía de memoria
- Modificación y volatilidad de la información
- Ejemplo del uso de la memoria
- Bibliografía

Bibliografía

- Patterson y Hennessy: Estructura y Diseño de Computadores. Capítulo 5.
- Murdocca y Heuring: Principios de Arquitectura de Computadoras: Capítulo 7.