

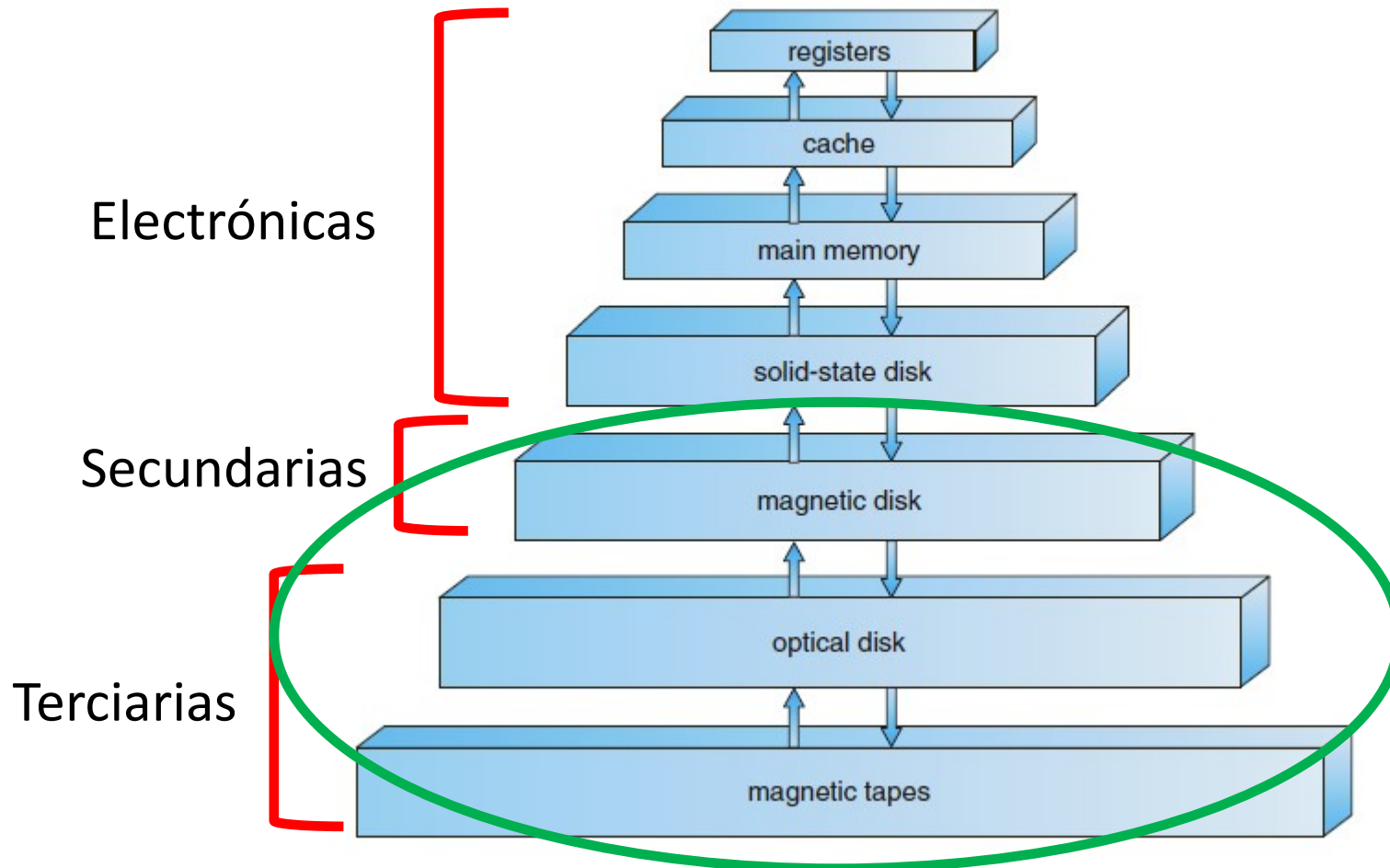
Sistemas Operativos

Cursada 2022

Comisión S21 y S22

Sistema de Almacenamiento

Jerarquía de Memorias (en general)



Sistema de Almacenamiento

- **El periférico fundamental de una computadora, serian sus dispositivos de memoria periférica.**
- **Tienen como característica principal la no volatilidad de los datos guardados en el.**
- **Necesitamos todo un modulo del S.O. dedicado a la administración de la memoria periférica.**
- **El tema principal es como guardamos la información en estos dispositivos y como la organizamos.**

Sistema de Almacenamiento

Estos tipos de memorias son las que se conectan a través del canal

Memorias Secundarias: Unidades Discos Solidos, Rígidos. Las conectadas permanentemente al equipo

Memorias Terciarias: Pendrive, incluimos CD/DVD, tarjetas, etc. Dispositivos removibles.

Sistema de Almacenamiento

Vamos a usar el concepto de “**Archivo**” en el concepto mas restringido.

Que es un Archivo?

- Es un conjunto de registros (llamado fichero)
- Un conjunto de datos organizados
- Un contenedor donde guardo algo, que?
Si o si contiene datos?
- **Entonces un archivo es un conjunto de bit identificables.**
- Hay otras definiciones, pero esas tienen que ver con los tipos de archivos

Sistema de Almacenamiento

Para manejar los archivos que nuestra unidad básica de almacenamiento requiere, vamos a necesitar un software específico para administrar los **Archivos**, esto se llama

Sistema de Archivos (File System)

Unix: unix file system (UFS)

Linux: ext2, ext3, ext4

Microsoft: FAT (2, 16, 32), NTFS

Cada uno maneja su propio File System

Hoy en día es tan así?

Sistema de Almacenamiento

Características de un Archivo o Atributos

➤ **Identificador (Id)** Atributo principal

➤ **Nombre**

➤ **Ubicación**

➤ **Tipo de Archivo**

➤ **Tamaño**

➤ **Protección**

➤ **Owner**

➤ **Fecha**

- Creación
- Modificación
- Ultimo Acceso

Todas estos Atributos los
los llamamos

Metadatos

Sistema de Almacenamiento

Operaciones sobre un *Archivo*

Creación (debe llamar a una SC al SO)

Escribir

Leer

Eliminar

Reposicionamiento

Split

Merge

Sistema de Almacenamiento

Hay muchas operaciones que se pueden hacer sobre un **archivo**, esto depende del tipo de **File System**

➤ **Hay algunos complejos**

Active Directory de Microsoft

➤ **Otros Sencillos**

CD de Música

Tipos de Archivos

Que define el tipo de archivo?

- **Que voy a guardar en el**
- **Como deben estar organizados ese conjunto de bits**
- **Definir una estructura interna.**
- **Cada contenedor va hacer de un tipo distinto de acuerdo a lo que guarde**

Tipos de Archivos

Texto

- En el sentido básico
- Texto puro
- No de un procesador de texto
- Ej. Archivo de script

Tienen algunas características particulares

Lo que llamamos formato del archivo texto
Este tiene delimitadores (por ej fin de línea)
ASCII 10 (LF) 13 (CR)

Tipos de Archivos

Binarios

- A veces llamados ejecutables
- Permiso de Ejecucion en UNIX/Linux
- Ejecutable no es un tipo
- Ej. Archivo de script
- Si hablamos de extensión tenemos
los
 - obj, exe, lib msi, com
- Estos lo que tienen es código hexa

Tipos de Archivos

Registros

- Uno de los mas importantes
- Base en mainframe de como guardarlos datos
- Un archivo es un conjunto de registros
- Un registro es un conjunto de campos
- Pueden ser numéricos o alfabético

• L	Nro Alu	Ny Ap	Sexo	Edad
-----	---------	-------	------	------

Tipos de Archivos

Registros

- Todo el Sistema de mainframe basado en archivos de cinta (decada 70/80), los ficheros están basados en **registros**
- En base de datos lo llamamos **fila**
- Una **Tabla en BD** seria un conjunto de registros
- Elemento básico para guardar datos en informática. Muy Usado

Sistema de Almacenamiento

Tipos de Archivos

Después de la aparición de la PC empezó haber una cantidad considerable de tipos de archivos

Ej, archivos de sonido o imágenes, y de estos hay de varios tipos, pueden ser libres o propietarios

El **Sistema Operativo** debe manejar todos los tipos de archivos?

Sistema de Almacenamiento

Tipos de Archivos

Lo que se hizo fue generar una tabla donde cada tipo de archivo esta asociado a una aplicación que maneja el mismo

El **S.O.** hará nacer un proceso para que ese archivo pueda ser tratado.

Esto de tener una extensión viene de microsoft, esta extensión se divide por un **punto**

Algunos File System guardan en la estructura del archivo el tipo de aplicación que lo debe tratar

Sistema de Almacenamiento

Estructura de un Archivos

Las estructuras van a depender del tipo de archivo, por ejemplos cuyos delimitadores de párrafo van a ser caracteres especiales

Un binario no va a tener delimitadores, sino un conjunto de byte generado por el compilador, va a tener una **cabecera** indicando que es de tipo **binario** y como **metadato** el atributo de ejecutable.

Esto porque le puedo dar el atributo de ejecutable a un tipo texto (script de bash)

Sistema de Almacenamiento

Estructura de un Archivos

Entonces los archivos van a tener una estructura dividida en 2 partes (salvo texto puro)

Header (Cabecera del archivo)

Datos sobre la estructura interna del archivo

Cuerpo (El resto del archivo)

Cada norma va a definir esto.

Sistema de Almacenamiento

Métodos de acceso

Directo

- Nos permite leer un registro directamente sin tener que leer otros antes
- Este método se usa sobre discos rígidos

Ej: leemos un Numero de Alumno determinado, o modificamos el registro que corresponde a ese Alumno

Hay dos formas de tratarlos

- ***Metodo Relativo***
- ***Metodo Hash***

Métodos de acceso

Secuencial

- El mas básico y viene de las cintas magnéticas

- Tiene una cabecera con el Id del archivo

- Tipo de registros y cantidad en algunos casos

- Un delimitador de registros

- Un EOF

Ejemplos

- Un CD de música es muy simple se lee en forma secuencial

- Una copia de backup

- Los archivos de logs

Sistema de Almacenamiento

Métodos de acceso **Relativo** (Mas Usado)

- Trabaja con una variable de tipo **record**
- Llamamos **Registro lógico**
- Sabemos cuanto byte ocupa cada variable
- Esta en nuestro programa ese dato
- Por lo tanto sabemos el tamaño cuando vamos a grabar o leer al Disco
- La unidad básica será el Registro Físico

Sistema de Almacenamiento

Métodos de acceso *Relativo*

Como sabemos donde ir a leer ?

- En el directorio esta el dato donde comienza el archivo.
- El S.O. nos da soporte de como resolverlo haciendo cuentas matemáticas.
- EL S.O. ve el disco como un arreglo de sectores
- El drivers lo mapeara en pista y sector
- Para el programador es transparente, puedo trabajar con el numero que mas que convenga

Sistema de Almacenamiento

Métodos de acceso *Hash*

Evidentemente habrá repetidos si paso de un dominio al otro

Esto se llama **Sinónimo** en la técnica de hash

Dos números del dominio que dan el mismo numero del codominio

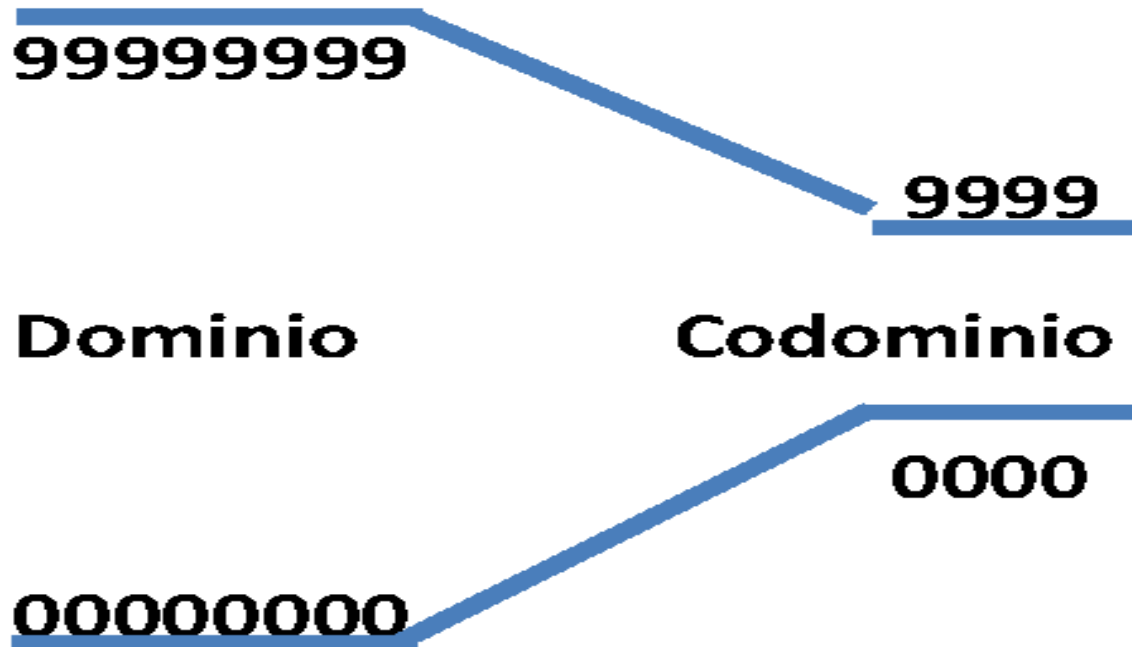
Desde el punto de vista matemático la función es compleja.

Sistema de Almacenamiento

Métodos de acceso *Hash*

Como el DNI tiene 8 dígitos, y si lo uso como Nro_cliente, es raro tener un volumen tan grande.

Usamos la **técnica de hash** para realizar la búsqueda
Ej.



Sistema de Almacenamiento

Métodos de acceso *Hash*

Se utiliza dentro de la función hash los números randón.

Que pasa cuando tengo sinónimos, hay al menos 2 políticas

Guárdalo en el lugar que sigue, para esto debo definir un codominio generoso y fijo, cosa de que siempre encuentre un lugar libre

Sistema de Almacenamiento

Métodos de acceso *Hash*

Cuando hago la búsqueda voy comparando hasta que coincide lo que busco

Busco **20346775**

20346771
20346772
20346773
20346775

Sistema de Almacenamiento

Métodos de acceso *Hash*

La otra técnica es generar un **segundo archivo** donde guardo los sinónimos

Crear una **lista encadenada** para los sinónimo, seria otra técnica

Los motores de base de datos usan estas técnicas.

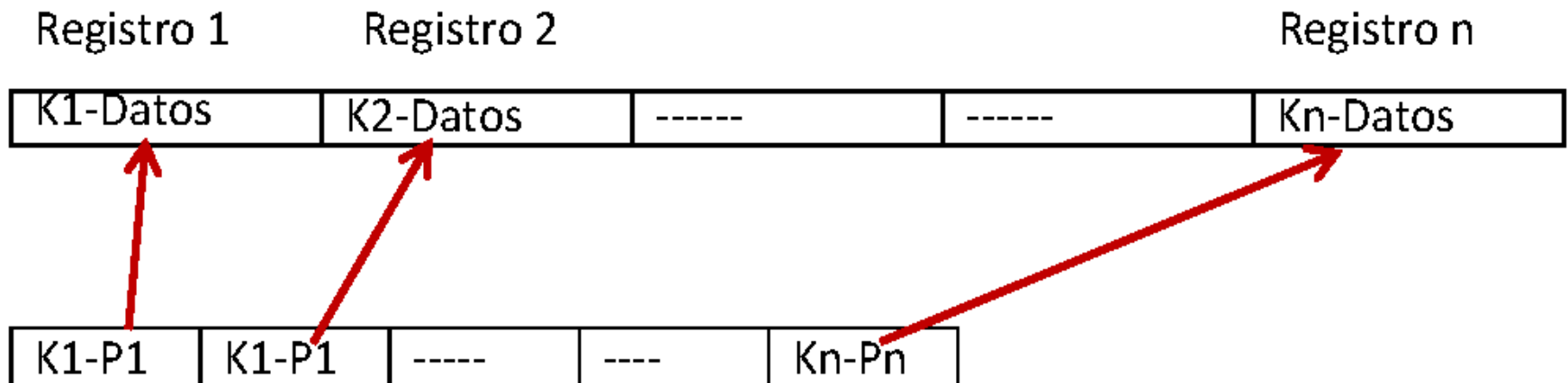
Para nosotros es transparente, le damos la clave y el motor la busca

Sistema de Almacenamiento

Archivos ISAM (Índex Secuencial Access Method)

Método desarrollado por IBM en sus mainframe para archivos grandes

Archivo de Registros



Archivo de Índices (ordenado)

Suponemos que K1, K2, Kn es una clave que puede ser NroAlumno

Sistema de Almacenamiento

Archivos ISAM (Índex Secuencial Access Method)

A medida que doy altas voy agregando al final, y también debo ir actualizando nuestro **Archivo de Índice**

Que debo hacer con el A.I. si lo actualice?

Ordenar (Utilitario Sort)

Por lo tanto la búsqueda que voy hacer sobre el A.I. será **binaria**

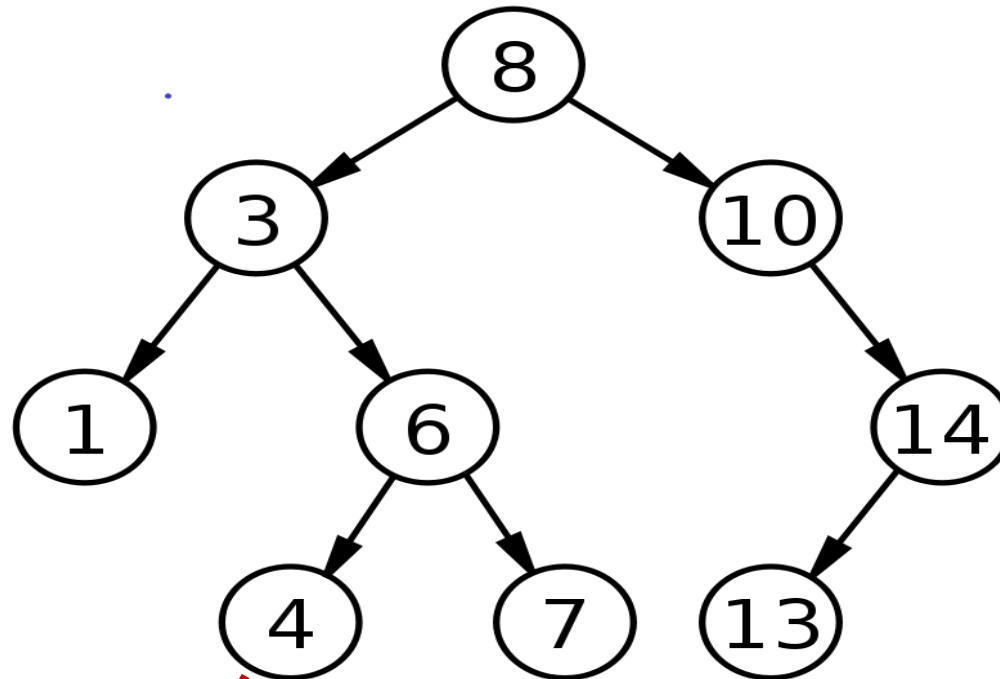
Sistema de Almacenamiento

Archivos ISAM (Índex Secuencial Access Method)

- Se uso mucho en los mainframe en archivos de millones de registros.
- Esto porque me permite hacer pocas lecturas
- Si tengo 64000reg a lo sumo hare unas 10 lecturas mas o menos
- Este método de llama **ISAM Puro**
- Hay otras metodologías usadas, para la búsqueda, que es la **Teoría de Grafo y Arboles**

Sistema de Almacenamiento

No necesariamente tiene que ser un **árbol binario**
Puede ser arboles **tipo B** (Cada nodo tiene varios hijos)
Árbol **Rojo-Negro** (Variante binario – complejo)



Sistema de Almacenamiento

- Me permite hacer búsquedas de archivos de millones de registro.
- Todas estas teorías son usadas en los motores de Base de Datos
- En definitiva todos estos métodos de acceso que hemos mencionado se aíslan por debajo de una capa de software que es el propio motor de **Base de Datos**
- El usuario le tira la clave de búsqueda y este se encarga de buscarla
- EL tema de usar una **BD** significa que es mas lenta la búsqueda
- Usar un archivo de **acceso directo** seria mas rápido

Los dejo a su criterio

Fin del Tema