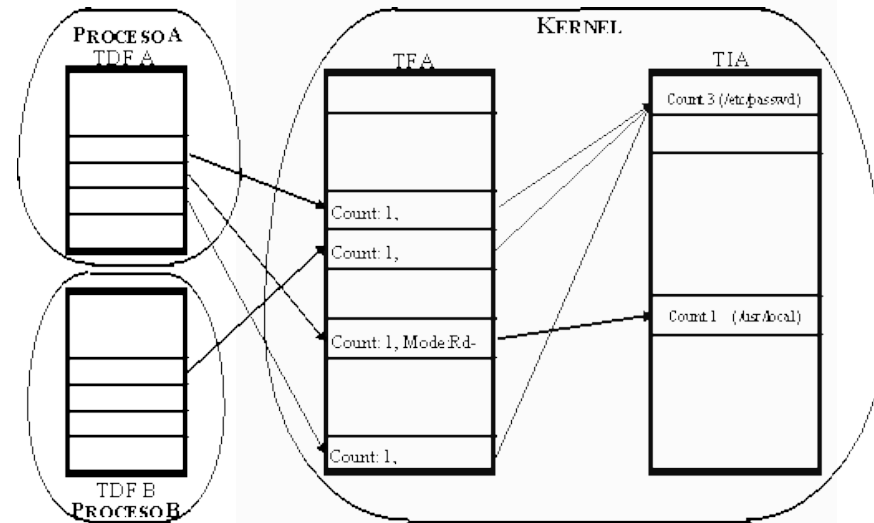


Manejo e Implementación de Archivos

Guatemala 17 de octubre de 2023

Ing. David Luna



Universidad
Rafael Landívar
Tradición Jesuita en Guatemala

Agenda



Secuencial

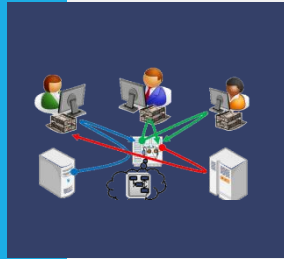


Secuencial indizado



Índices

Índices



Índices

¿Qué es?

- ☐ Estructura de datos que optimiza el acceso a los registros del archivo maestro.
- ☐ Debido a que almacena ciertos datos, suele caber en memoria principal y por ende su acceso es más rápido que acceder al archivo maestro.
- ☐ Debe almacenar, como mínimo, un campo para identificar al dato y la dirección física para localizarlo.
- ☐ Es responsabilidad del diseñador del sistema de archivos, definir de manera óptima los índices, su organización y su contenido.

Índices

Características generales:

- Los índices siempre deben estar clasificados (ordenados), según su atributo llave.
- Un archivo de datos puede tener asociado más de un índice.
- Un índice puede estar asociado a uno o más campos de un registro.

Índices

Características generales:

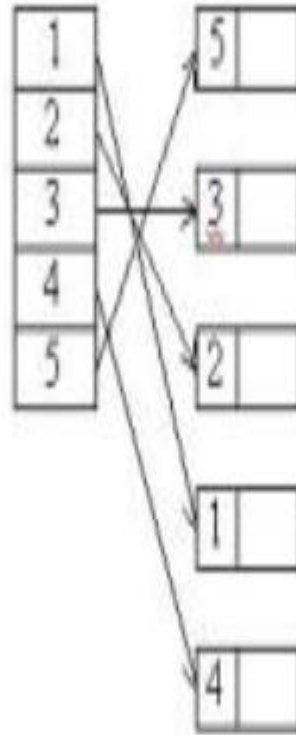
- Unicidad: es la propiedad que exige que la llave de búsqueda sea única.
- Duplicidad: es la propiedad que permite tener valores repetidos para la llave de búsqueda.



Clasificación

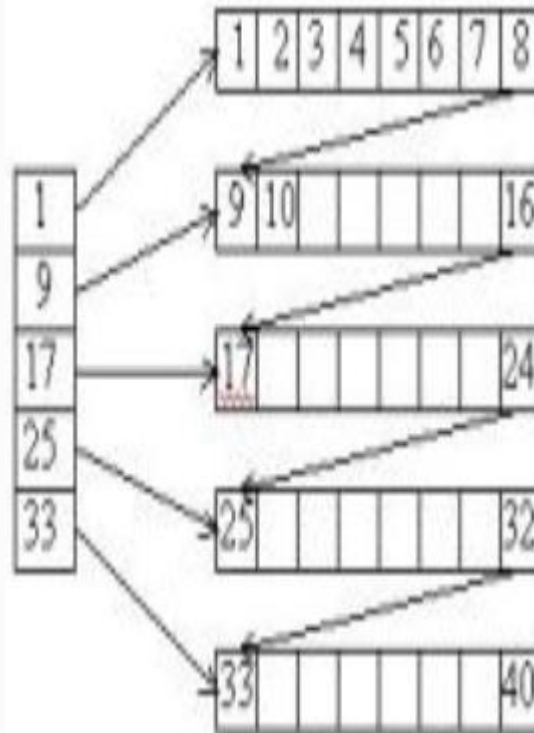
Clasificación por densidad:

● Un índice **denso** o **exhaustivo** es el que asocia una entrada en el índice, por cada registro de datos.



Clasificación por densidad:

● Un índice **disperso**, **escaso** o **selectivo** es el que no necesariamente asocia una entrada en el índice, por cada registro de datos. Pueden haber registros en el archivo de datos, que no tengan una correspondiente llave en el índice.



Clasificación por frecuencia de actualización:

- Un índice en línea es el que se actualiza en sincronización con cada operación en el archivo de datos.
- Un índice en batch se usa cuando se almacena una bitácora de transacciones y en un momento determinado se lee dicha bitácora para actualizar el índice con todas las operaciones, es decir que la actualización del archivo de datos y el índice no se hace al mismo tiempo.

Clasificación por tipo de llave:

- Un índice simple o primario

Es aquel que sólo registra la llave primaria, la cual identifica de forma única a cada registro.

El índice primario no permite la búsqueda de datos por campos que no sean la llave primaria.

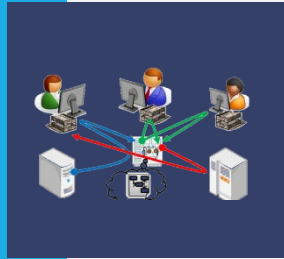
Clasificación por tipo de llave:

- Un índice secundario

Incluye más campos que la llave primaria y permite llaves duplicadas, lo cual es ventajoso para las búsquedas, pero puede reducir el desempeño durante las otras operaciones

.

Secuencial Indizado

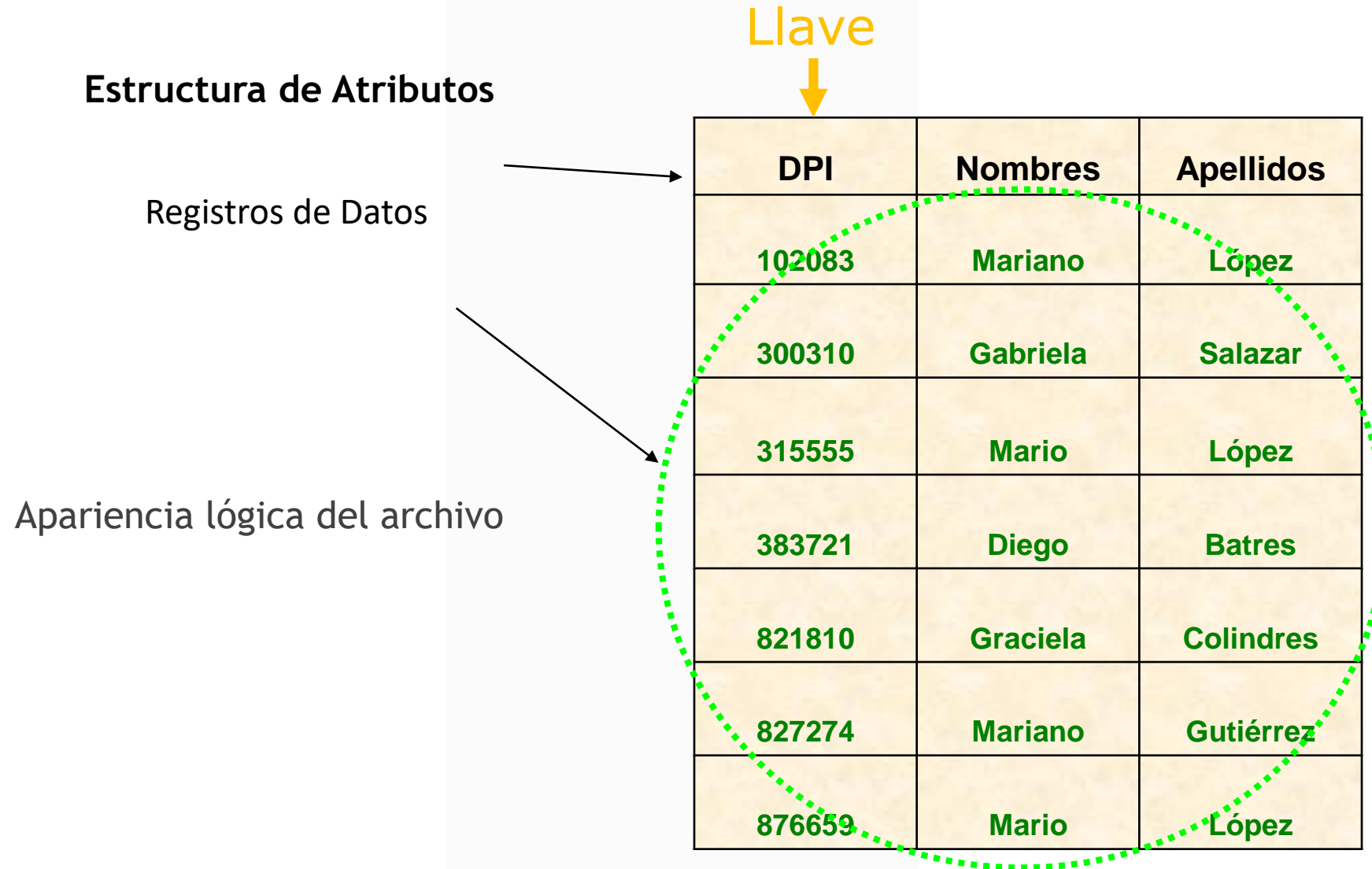


Archivo Secuencial Indizado

Características de los Registros:

- Almacenados con una secuencia lógica.
- Longitud fija.
- Atributos en la misma posición, sólo necesitan aparecer una vez en el descriptor y a cada uno se le asocia un conjunto de datos que se denomina columna.

Archivo Secuencial Indizado



Archivo Secuencial Indizado

- Se define una llave para determinar la secuencia lógica de ordenamiento.
- Uno o más atributos se toman como la llave de cada registro y se denomina llave primaria.
- Si una llave primaria no es única, entonces pueden especificarse atributos llave secundarios, terciarios y así sucesivamente.

Archivo Secuencial Indizado

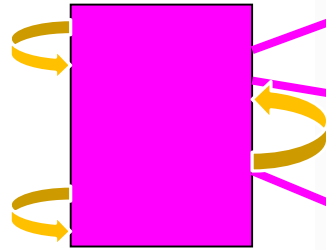
Está compuesto por:



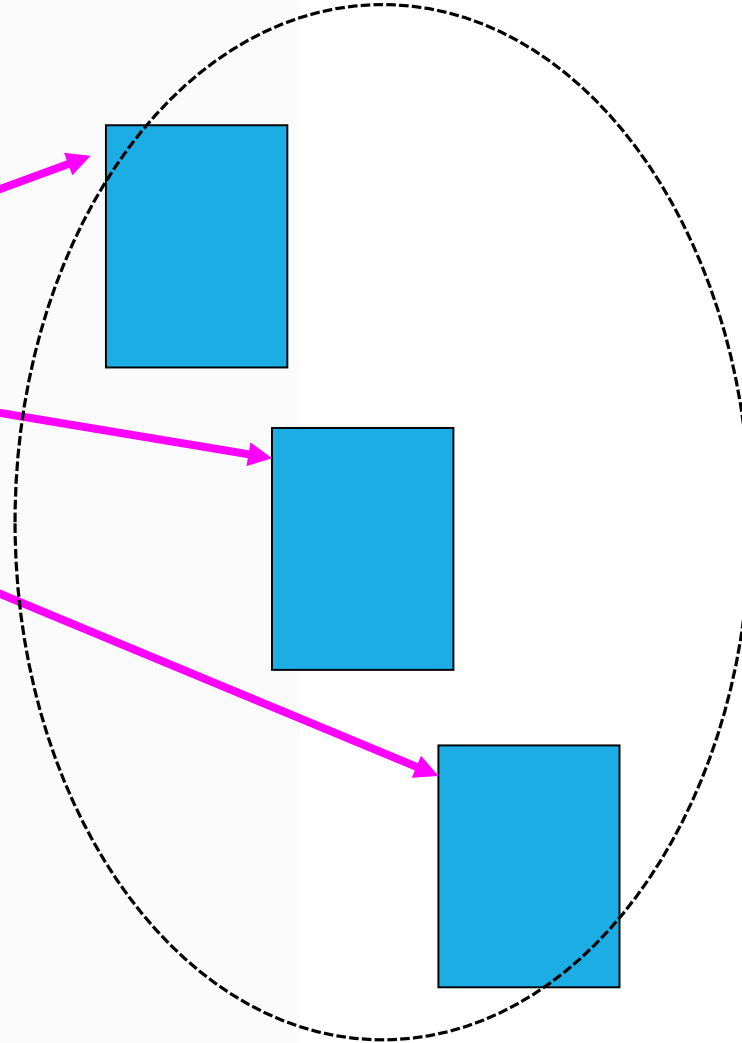
- **Objetivo:** Proveer acceso más rápido a los datos.
- Búsquedas a través de la llave.
- Puede estar compuesto por “n” bloques de longitud fija.
- Al crearse el archivo, se crea el primer bloque.
- Al llenarse un bloque, se crea uno nuevo.

Archivo Secuencial Indizado

Archivo Índice
(lista enlazada)

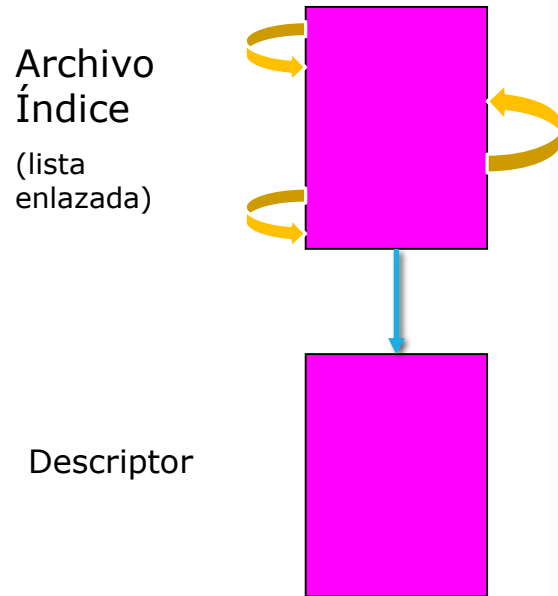


Archivo Maestro



Archivo Índice

- El archivo índice está formado por registros que simulan una lista enlazada; se define un descriptor para el archivo índice.



- Los registros del índice están enlazados.
- Los registros del índice contienen al menos: Registro, posición, llave, siguiente de la lista, estatus.
- La posición incluye: número de registro y número de bloque.

Descriptor
(índice)
Reg.Inicial: 2
No.Bloques:1

inicio →

Índice

Registro	Posición	Llave	Siguiente	Estatus
1	1.1	Alvarado Zapata	...	activo
2	1.2	Alvarado Mack	1	activo
3

Maestro, bloque 1

Descriptor
(bloque 1,
maestro)
No.Registros: 2
Max.Reg: 20

1

2

Apellidos	Nombres	Dirección	Teléfono	Estatus
Alvarado Zapata	Mónica María	Av. Bolívar 5-60	98737361	activo
Alvarado Mack	José Guillermo	1 calle 30-20	55543383	activo
...

Llave

Operaciones

Archivo Secuencial Indizado

a) Inserción:

- Los registros están limitados a un conjunto predeterminado de atributos.
- Debe verificarse que los datos a ingresar pertenezcan al dominio asociado a cada atributo.
- No pueden almacenarse registros de longitud mayor a la definida originalmente.
- **Las inserciones se hacen si no existe la llave.**
- Las inserciones se hacen en el archivo maestro, en el bloque actual (el último creado) . Se utiliza un nuevo bloque si ya se llegó al máximo de registros por bloque.
- Se agrega la llave en el índice y se actualiza el enlace de los registros (se enlazan generando el ordenamiento).
- Con cada inserción se revisa el primero de la lista.
- Se actualizan descriptores del índice y el maestro.

Ejemplo

Se tendrá archivo secuencial indizado, que tiene las siguientes características:

Cada registro debe tener los siguientes campos:

- No_Carnet
- Nombre_completo
- Correo_Electrónico
- Dirección
- Teléfono
- Rol
- Contraseña
- Fecha_Nacimiento
- Estatus

La llave primaria es No_Carnet



Los bloques no pueden tener mas de 4 registros

Pasos

Identificar bloques



Identificar max para
reorganizar



Insertar en el bloque
que corresponde



Identificar inicio del
índice



Insertar en índice



Actualizar punteros



Actualizar descriptores



Operaciones Archivo Secuencial Indizado

b) Búsqueda:

Si el criterio de búsqueda es la llave primaria, se utiliza el índice para ubicar la posición del registro y la búsqueda es directa.

Si el criterio de búsqueda NO incluye la llave primaria, se utiliza búsqueda secuencial en el archivo maestro.



Operaciones Archivo Secuencial Indizado

c) Eliminación:

Implica la búsqueda del registro que se eliminará.

Se cambia el estatus del registro a eliminar en ambos archivos y **se reordenan los enlaces en el archivo índice**; la eliminación es lógica en lugar de física.

Al realizarse la reorganización, los registros marcados serán físicamente eliminados.



Operaciones Archivo Secuencial Indizado

d) Actualización:

Implica la búsqueda del registro que contiene la información que se desea actualizar.

Si la actualización NO incluye la llave primaria, entonces sólo se actualizan datos, en el archivo maestro.

Si la actualización SI incluye la llave primaria, se inserta un nuevo registro y se elimina de forma lógica el registro antiguo, cambiándole el estatus.



Operaciones Archivo Secuencial Indizado

e) Reorganización:

Se crean archivos temporales: índice y maestro.

Se trasladan los registros no marcados de cada archivo y se trasladan al archivo temporal correspondiente: índice y maestro.

Se eliminan los archivos originales y se renombran los temporales.

Deben actualizarse los datos del descriptor.

Asignación / Liberación de espacio



Se deben tener presentes problemas tales como la “fragmentación” creciente del espacio en disco:

- Ocasiona problemas de rendimiento al hacer que los archivos se desperdigen a través de bloques muy dispersos.
- Técnicas para aliviar el problema de la “fragmentación” consisten en realizar periódicamente:
 - “Condensación”: se pueden “reorganizar” los archivos expresamente o automáticamente según algún criterio predefinido.

http://upload.wikimedia.org/wikipedia/commons/4/40/Defragmenting_disk.gif

- “Recolección de basura o residuos” (garbage collector): se puede hacer fuera de línea o en línea, con el sistema activo, según la implementación.



Aplicaciones Archivo Secuencial Indizado

Cuando es crucial la localización de la información.

Ejemplo

Se tendrá archivo secuencial indizado, que tiene las siguientes características:

Cada registro debe tener los siguientes campos:

- No_Carnet
- Nombre_completo
- Correo_Electrónico
- Dirección
- Teléfono
- Rol
- Contraseña
- Fecha_Nacimiento
- Estatus

La llave primaria es No_Carnet



Los bloques no pueden tener mas de 5 registros

Gracias

¿ALGUNA PREGUNTA?