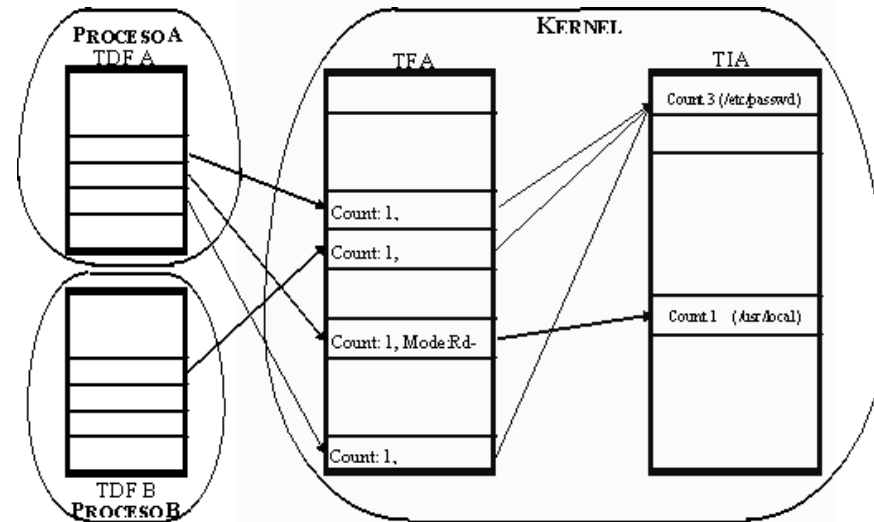


Manejo e Implementación de Archivos

Guatemala 19 de
septiembre de 2023

Ing. David Luna



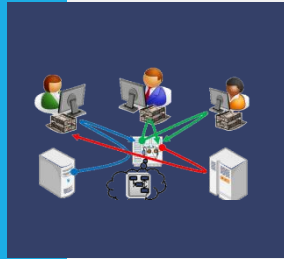
Universidad
Rafael Landívar
Tradición Jesuita en Guatemala

Agenda

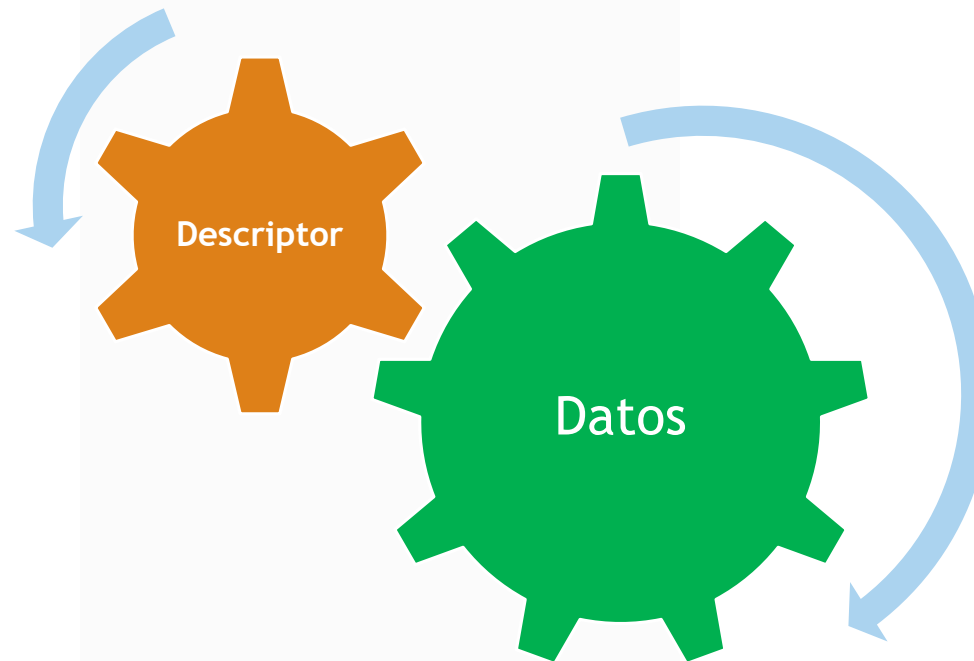


Clasificación de Archivos

Componentes de un archivo



Componentes de un Archivo



Descriptor:

El descriptor de archivos o bloque de control de archivos, son los datos que el sistema necesita para administrar dicho archivo.

- ⦿ Nombre simbólico del archivo
- ⦿ Localización del archivo en el almacenamiento secundario
- ⦿ Organización del archivo (método de organización y acceso)
- ⦿ Datos de control de acceso
- ⦿ Tipo (archivo de datos, programa objeto, programa fuente, etc.)
- ⦿ Disposición (permanente, temporal)
- ⦿ Fecha y hora de creación
- ⦿ Fecha de la última modificación
- ⦿ Usuario que modificó
- ⦿ Estadísticas de acceso (número de lecturas, escrituras, tiempos de acceso, etc.)
- ⦿ Registros activos e inactivos.
- ⦿ Metadatos
- ⦿ Y más...

Contenido del Archivo:

Datos, que pueden estar estructurados en la forma de:

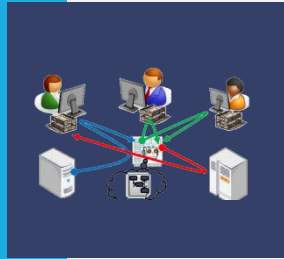
➤ Registros,

- Están compuestos por campos. Estos campos se identifican por lo que se conoce como "metadatos" (información acerca de los datos), en el caso de los campos el metadato es el nombre de dicho campo.

➤ Formación abstracta,

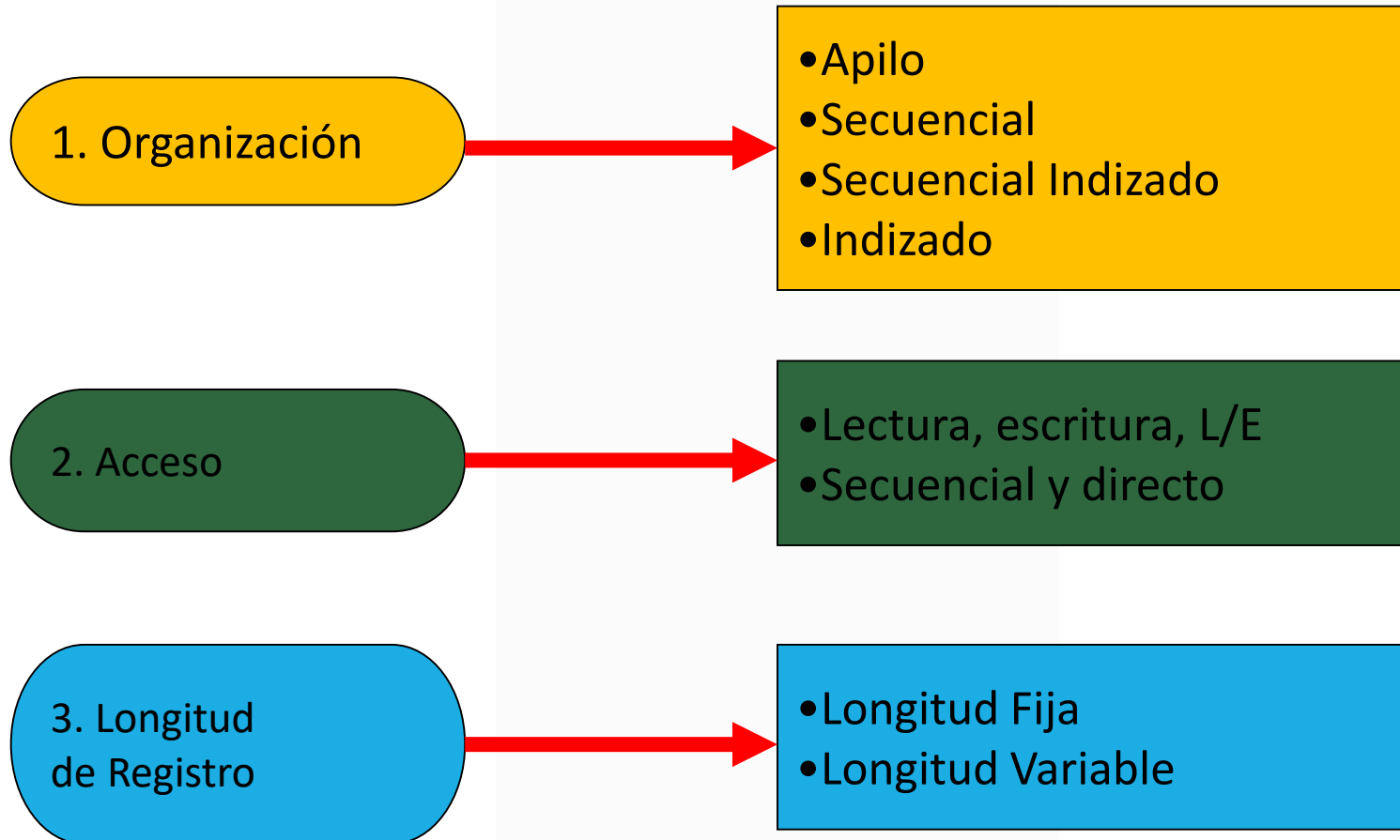
- El archivo tiene al inicio (a veces al final) una serie de bytes que representan los metadatos del archivo, es decir, si por ejemplo el archivo es una imagen al inicio tendremos bytes que nos digan el largo, el ancho, el número de colores, etc, estos bytes se encuentran especificados en el "formato" del archivo. Son formatos estándar de aplicaciones específicas que se emplean en todas las plataformas

Clasificación de archivos



Clasificación de Archivos

Factores de Clasificación



Clasificación de Archivos

1. Organización de Archivos:

Forma en que se disponen los datos para su almacenamiento, recuperación y procesamiento.

Normalmente es una tarea de los sistemas operativos. Para la recuperación eficaz de los datos, estos se organizan en estructuras lógicas (pilas, listas, tablas, árboles, etc).

- 1.1 Apilo

- 1.2 Secuencial

- 1.3 Secuencial Indizado

- 1.4 Indizado



Clasificación de Archivos

2. Acceso:

Forma en que puede llegarse a los datos.

2.1 Lectura, escritura, L/E

2.2 Secuencial y directo



Clasificación de Archivos

3. Longitud de registro:

3.1 Longitud fija:

Independientemente de los datos almacenados, cada registro siempre tiene la misma cantidad de bytes.

3.2 Longitud variable

Los datos almacenados en un registro pueden tener mayor o menor longitud que el resto de registros.

Selección del tipo de Archivo

¿Cómo se elije el tipo de archivo?

- Aplicación
- Tamaño del archivo
- Velocidad búsqueda
- Manejo de operaciones
- Reorganización

Operaciones de un Archivo

¿Qué puede hacerse con un archivo?

- Creación
- Eliminación
- Inserción de datos
- Actualización de datos
- Eliminación de datos
- Reorganización
- Reestructuración



Reorganización de un Archivo

Los registros eliminados de forma lógica, se eliminan permanentemente de forma física. Será necesario realizar varias operaciones para completarla.

Estas operaciones podrían incluir:

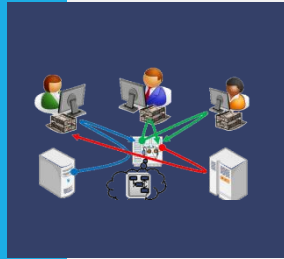
- Crear un archivo nuevo
- Insertar/eliminar registros
- Eliminar el archivo antiguo
- Modificar los descriptores

Reestructuración de un Archivo

Cuando es necesario implementar cambios en la estructura de los datos almacenados en los archivos, se habla de una “reestructuración”. Puede originarse por:

- Cambios drásticos en el volumen de datos
- Cambios en las reglas del negocio
- Cambios en la normativa legal y/o políticas estatales
- Bajo desempeño del sistema de archivos

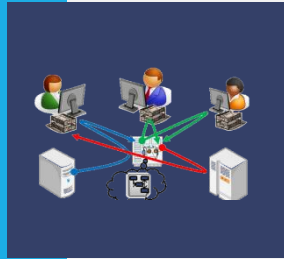
Clasificación de Archivos (Organización)



1. Organización

- Apilo
- Secuencial
- Secuencial Indizado
- Indizado

Archivo Apilo





Archivo Apilo

- Los datos se guardan en el orden en que llegan, sin secuencia lógica.
- **No existe una llave que identifique a los registros.**
- El tiempo para encontrar un registro es largo puesto que hay que buscar todos los registros para localizarlos.
- El tiempo que se emplea para añadir registros es mínimo ya que sólo es necesario agregarlo al final del archivo.

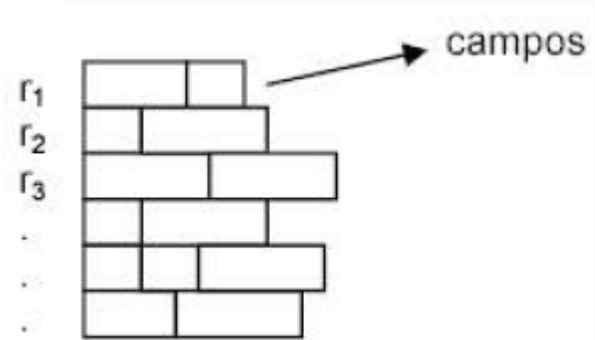


Archivo Apilo

- Registros de longitud variable.
- Cada registro almacena estructura y datos.
- El orden de los campos en un registro no es necesariamente la misma.

Archivo Apilo

Apariencia del archivo



Operaciones en un Apilo

a) Inserción:

- Se realiza de forma cronológica, el último registro se coloca al final del archivo.
- El archivo no tiene orden, cada registro no está identificado por una llave.
- Actualización del descriptor.



Operaciones en un Apilo

b) Búsqueda

- Secuencial
- La búsqueda arranca desde el inicio y deben compararse los campos hasta encontrar el registro o los registros que satisfacen los criterios de búsqueda.

Operaciones en un Apilo

c) Eliminación

- Implica la búsqueda del registro que se eliminará
- Se coloca una marca sobre el registro a eliminar, por lo cual la eliminación es lógica en lugar de física
- Actualización del descriptor.

Operaciones en un Apilo

d) Actualización

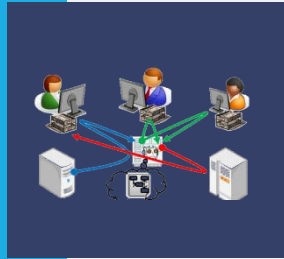
- Implica la búsqueda y eliminación del registro que contiene la información que se desea actualizar. Se realiza eliminación lógica del registro.
- Se inserta un nuevo registro con la información actualizada.
- Genera varios registros marcados como eliminados, debido a las actualizaciones.
- Actualización del descriptor.

Operaciones en un Apilo

e) Reorganización

- Se crea un archivo temporal.
- Se leen cronológicamente los registros del archivo físico y se trasladan al archivo temporal los registros NO marcados como eliminados.
- Se elimina el archivo físico y se renombra el archivo temporal, el cual pasa a ser el archivo físico.
- Deben actualizarse los datos del descriptor.

Archivo Secuencial



Archivo Secuencial

- Los registros de datos se guardan con una secuencia lógica, se guardan ordenados.
- Los registros de datos tienen longitud fija.
- Los registros de datos tienen los mismos atributos y todos en la misma posición; los nombres de los atributos sólo necesitan aparecer una vez en el descriptor y a cada atributo se le asocia un conjunto de datos que se denomina columna.

Archivo Secuencial

Apariencia del archivo

Llave

Estructura de Atributos

Registros de Datos

DPI	Nombres	Apellidos
102083	Arturo	López
300310	Gabriela	Salazar
315555	Mario	López
383721	Dominic	Batres
821810	Graciela	Colindres
827274	Mariano	Gutiérrez
876659	Mario	López

Archivo Secuencial

- Se define una llave para determinar la secuencia lógica de ordenamiento. No hay duplicados.
- Uno o más atributos se toman como la llave de cada registro y se denomina llave primaria.
- Algunas veces se agregan atributos para conformar una llave única de los registros.
- El archivo secuencial tiene dos componentes:
 - un archivo maestro y
 - la bitácora de transacciones que es un archivo tipo apilo.

Archivo Secuencial

Bitácora de
Transacciones(apilo)

DPI	Nombre	Apellido	Estatus
990377	Esteban	Morales	1
138309	Luisa	Keller	0
333377	Juan	Tello	1
212111	Patricia	Acevedo	1
...	

Archivo Maestro

DPI	Nombre	Apellido	Estatus
102083	Arturo	López	1
300310	Gabriela	Salazar	1
315555	Mario	López	0
383721	Dominic	Batres	1
821810	Graciela	Colindres	0
827274	Mariano	Gutiérrez	1
876659	Mario	López	1



Operaciones Archivo Secuencial

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 recopilan en la bitácora de transacciones (apilo), hasta que este archivo crece lo suficiente para realizar una actualización en lote del archivo maestro (una reorganización) y/o se reorganiza bajo demanda.
- Sólo se inserta un registro si se verifica que no existe la llave. Esto se hace para garantizar la unicidad de dicha llave.

Operaciones Archivo Secuencial

b) Búsqueda:

- Deben compararse los campos hasta encontrar el registro o los registros que satisfacen los criterios de búsqueda.
- La búsqueda puede incluir ambos archivos: el maestro y la bitácora de transacciones. La búsqueda en la bitácora puede realizarse antes de revisar el maestro y siguiendo la lógica del apilo (secuencial), o bien, pueden realizarse ambas búsquedas en paralelo.
- Algunos métodos específicos sólo buscan en el archivo maestro, lo cual implica que antes de iniciar la búsqueda debe hacerse previamente la reorganización.

Operaciones Archivo Secuencial

b) Búsqueda (continuación):

● Puede ser secuencial o binaria.

- Secuencial a través del archivo, es decir, registro por registro.
- Binaria comienza con un acceso directo al centro del archivo, luego se segmenta en subsecuentes mitades el bloque donde se considere está la llave de búsqueda.

Operaciones Archivo Secuencial

c) Eliminación:

- Implica la búsqueda del registro que se eliminará.
- Se coloca una marca sobre el registro a eliminar en el archivo maestro o en la bitácora de transacciones, por lo cual la eliminación es lógica en lugar de física.
- Al realizarse la reorganización el registro marcado como inactivo, será físicamente eliminado.

Operaciones Archivo Secuencial

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 o en la bitácora de transacciones.
- Si la actualización SI incluye la llave primaria, se inserta un nuevo registro en la bitácora de transacciones y se elimina de forma lógica el registro antiguo, colocándole una marca, ya sea en el archivo maestro o la bitácora de transacciones.

Operaciones Archivo Secuencial

e) Reorganización:

- Consiste en intercalar ordenadamente el archivo maestro y la bitácora de transacciones, generando un nuevo archivo físico.
- Primero se ordena la bitácora de transacciones de acuerdo a las llaves de ordenamiento del archivo maestro.
- Se crea un archivo temporal.
- Se leen los registros no marcados del archivo maestro y de la bitácora de transacciones y se trasladan ordenadamente al archivo temporal.
- Se eliminan el archivo maestro y la bitácora de transacciones y se renombra el archivo temporal, el cual pasa a ser el archivo maestro. Se crea nuevamente la bitácora de transacciones (ahora vacía).
- Deben actualizarse los datos de los 2 descriptores (el de la bitácora y el del archivo maestro).

Aplicaciones Archivo Secuencial

Procesamiento de datos orientado al manejo por lotes, cuando dicho procesamiento tiene cierta periodicidad.

Ejemplos:

- Asignaciones de un ciclo (propuestas - > Confirmadas).
- Agregar diagnósticos al historial médico de una persona.

Gracias

¿ALGUNA PREGUNTA?