

4. La localidad tiende a ser pequeña.
5. La fragmentación interna se reduce con páginas pequeñas.

Los diseñadores sugieren el uso de páginas pequeñas.

Muchos experimentos han examinado el comportamiento de los sistemas de computo con paginación y los resultados han sido interesantes:

1. Cuando un proceso comienza a ejecutarse, por lo regular hace referencia rápidamente aun gran porcentaje de sus páginas.
2. El número de fallas experimentadas por un proceso en ejecución tiende a crecer a medida que aumenta el tamaño de la página suponiendo que permanece constante el almacenamiento primario asignado al proceso.
3. El tiempo entre fallos experimentado por un proceso en ejecución crece a medida que aumenta la cantidad de marcos de página asignados al proceso, una vez asignado al proceso un número suficiente de marcos de páginas para contener su conjunto de trabajo disminuye la tasa de incremento.
4. El número de instrucciones ejecutadas en una página antes de que el control salga de la página tiende a ser pequeño.

## Unidad V

### SISTEMAS DE ARCHIVO

**Objetivo de esta unidad.-** Durante el desarrollo de esta unidad, se analizará como almacenar una gran cantidad de información, que no se pierda hasta cuando el proceso que la generó termina su ejecución y que dos o más procesos tengan acceso concurrente a la información.

El alumno deberá comprender la forma de uso de los archivos y las propiedades que estos poseen; la utilización de los directorios, su organización, sus propiedades y las operaciones que se lleven a cabo con ellos; la implantación de un sistema de archivo, la seguridad de la información y mecanismos de protección debe de tener un sistema de archivos.

**Archivo.-** Es un conjunto de datos al cual se le asigna un nombre, reside, por lo regular, en el almacenamiento secundario ya sea en cinta ó en disco, se puede manejar como una unidad mediante operaciones como:

- abrir
- cerrar
- destruir
- copiar
- renombrar
- listar

A los elementos individuales de información dentro del archivo se le pueden aplicar las siguientes operaciones:

- leer
- escribir
- modificar
- insertar
- eliminar

Los archivos pueden caracterizarse por su:

**Volatilidad.-** Se refiere a la frecuencia con la que se añade información a un archivo y se borra de él.

**Actividad.-** Se refiere al porcentaje de los registros de un archivo al cual se tuvo acceso durante un periodo dado.

**Tamaño.-** Se refiere a la cantidad de información almacenada en el archivo

#### Sistemas de Archivos

Los sistemas de archivo contienen:

**Método de acceso.-** Se ocupan de la forma en que se obtiene acceso a los datos almacenados en archivos.

**Administración de archivos.-** Se ocupa de ofrecer mecanismos para almacenar, compartir y asegurar archivos, y hacer referencia a ellos.

**Administración de almacenamiento secundario.-** Se ocupa de asignar espacio a los archivos en dispositivos de almacenamiento secundario.

**Mecanismos de integridad de los archivos.-** Se ocupan de garantizar que no se corrompa la información en un archivo. Cuando se asegura la integridad de los archivos, cualquier información que deba estar en un archivo estará ahí.

*El sistema de archivos se ocupa primordialmente de administrar el espacio de almacenamiento secundario, sobre todo el espacio en disco.*

#### **Funciones del Sistemas de Archivos**

Además de que los usuarios deben poder crear, modificar y eliminar archivos las funciones normalmente atribuidas a los sistemas de archivos son:

**Compartir información.-** Los usuarios deben ser capaces de compartir sus archivos entre sí en forma cuidadosamente controlada para aprovechar y continuar el trabajo de los demás.

**Mantener privada la información, cifrarla o descifrarla.-** En ambientes delicados en los cuales la información debe mantenerse segura y privada, como los sistemas de transferencia de fondos, sistemas de expedientes criminales, sistemas de expedientes médicos, etc, es conveniente ofrecer estas funciones, lo cual hace que la información solo resulte útil a quienes esta destinada.

**Obtener acceso a la información.-** El mecanismo para compartir debe ofrecer varios tipos de acceso controlado, como acceso para lectura, acceso para escritura, acceso para ejecución o las diversas combinaciones de estos.

**Respaldo y Recuperación de la información.-** Deben ofrecerse recursos de respaldo y recuperación para evitar la pérdida accidental o la destrucción mal intencionada de información.

**Independencia con respecto a los dispositivos.-** Los usuarios deben poder hacer referencia a sus archivos mediante nombres simbólicos en vez de tener que usar nombres de dispositivos físicos

Lo más importante de todo es que el sistema de archivos debe ofrecer una interfaz amable con el usuario. Debe proporcionar a los usuarios una vista lógica de sus datos y las funciones que puede realizar con ellos, más que una vista física. El usuario no debe preocuparse de los dispositivos específicos en los cuales esta almacenada la información, la

forma que adoptan los datos en esos dispositivos o los mecanismos físicos para transferir datos de esos dispositivos o hacia ellos.

#### **Jerarquía de los Datos**

La estructura de datos esta formada por entidades de complejidad creciente:

bits → bytes → campos → registros → archivos → bases de datos

#### **División en Bloques y Empleo de Buffer**

**Registro físico.-** Es la unidad de información que se lee realmente de un dispositivo o se graba en él.  
**o Bloque físico**

**Registro Lógico.-** Es un conjunto de datos manejado como una unidad desde el punto de vista del usuario.  
**o Bloque lógico**

**Registro sin Bloques.-** Cuando cada registro físico contiene sólo un registro lógico se dice que el archivo esta formado por *registros sin bloques*.

**Registro en Bloques.-** Cuando cada registro físico puede contener varios registros lógicos se dice que el archivo esta formado por *registros en bloques*.

En un archivo con *registros de longitud fija*, todos los registros tienen el mismo tamaño; el tamaño de bloque casi siempre es un múltiplo entero de la longitud del registro.

En un archivo con *registros de longitud variable*, el tamaño del registro puede variar, sin rebasar el tamaño del bloque.

El *empleo de Buffer* (áreas de almacenamiento temporal) hace posible que el cómputo se efectúe de forma simultáneo con la entrada/salida. Se reservan espacios en almacenamiento primario para guardar varios bloques físicos de un archivo a la vez, cada uno de estos espacios se llama buffer. El método más común son los buffers dobles y funciona como siguen (para la salida):

1. Existen dos buffers
2. Los registros generados por un proceso se depositan en el primer buffer hasta que se llena.
3. Se inicia la transferencia del bloque en el primer buffer al almacenamiento secundario.
4. Mientras se realiza esta transferencia, el proceso continúa generando registros que se depositan en el segundo buffer
5. Cuando se llena el segundo buffer y una vez finalizada la transferencia del primer buffer, se inicia la transferencia del segundo buffer.

El proceso sigue generando registros, los cuales se depositan ahora en el primer buffer. Esta alternación de buffers permite que se lleve a cabo la entrada/salida en paralelo con los cálculos de un proceso.

### Organización de los Archivos

La organización de los archivos se refiere a la forma como se acomodan los registros de un archivo en almacenamiento secundario.

Sistemas de Organización de Archivos más comunes:

**Secuencial.-** Los registros se colocan en orden físico. Es una organización natural para archivos grabados en cinta magnética, medio de almacenamiento que por su naturaleza es secuencial. Los archivos en disco también se pueden organizar secuencialmente, aunque por diversas razones, los registros de un archivo secuencial de disco no se almacenan en forma contigua por fuerza.

**Directo.-** Se obtiene acceso directo (aleatorio) a los registros por su dirección física en un dispositivo de almacenamiento de acceso directo (*DASD, direct access storage device*)

**Secuencial indexado.-** Los registros se acomodan en secuencia lógica de acuerdo con una clave contenida en cada registro. El sistema mantiene un índice con las direcciones físicas de ciertos registros principales. El acceso a los registros secuencial indexado puede obtenerse secuencialmente por orden de clave o de manera directa, mediante una búsqueda dentro del índice creado por el sistema. Los archivos secuencial indexado casi siempre se almacenan en discos.

**De partición.-** Este es en esencia un archivo de subarchivos secuenciales. Cada subarchivo secuencial se llama miembro. La dirección inicial de cada miembro se almacena en el directorio del archivo.

El término *volumen* se usa para referirse al medio de grabación de cada dispositivo auxiliar de almacenamiento en particular. El volumen empleado en una unidad de cinta es el carrete de cinta y el volumen empleado en una unidad de disco es el disco.

### Métodos de acceso

El acceso a los archivos se logra mediante funciones de los sistemas operativos llamados *métodos de acceso*, estos se agrupan en dos categorías:

**Método de Acceso por Colas.-** Se usan con archivos de organización secuencial; realizan *transferencias anticipadas a buffers* y programan la E/S además de ofrecer agrupación en bloques y separación de bloques en forma automática.

**Métodos de Acceso Básicos.-** Se utilizan por lo regular cuando no es posible anticipar el orden en el que se van a procesar los registros, sobre todo en el caso del acceso directo. En los métodos básicos, el método leer y escribe bloques físicos; si se requiere en la aplicación, el usuario se encarga de la formación y división de bloques.

### Asignación y Liberación de Espacio

Con la asignación contigua cada archivo se asigna a un área única del almacenamiento secundario. La asignación contigua facilita el acceso rápido pero adolece de serios problemas de fragmentación.

Si la asignación es no contigua, el archivo puede estar disperso en varias áreas de almacenamiento secundario. La asignación no contigua es más flexible pero puede requerir búsquedas frecuentes.

### Descriptor de Archivos

Llamado también Bloque de Control de Archivo. Es un bloque de control con información que el sistema necesita para administrar un archivo. Un descriptor de archivo representativo incluye lo siguiente:

1. Nombre simbólico del archivo.
2. Localización del archivo en el almacenamiento secundario.
3. Organización del archivo.
4. Tipo de dispositivo
5. Datos para el control de acceso.
6. Tipo (Si es archivo de datos, programa objeto, o programa fuente).
7. Tratamiento (Temporal o permanente).
8. Fecha y hora de creación.
9. Fecha de destrucción.
10. Fecha y hora de la última modificación.
11. Conteo De la actividad de acceso.

Los descriptores de archivo se mantienen en el almacenamiento secundario y se transfieren al almacenamiento primario cuando se abre un archivo. El descriptor de archivo es controlado por el sistema operativo, el usuario no puede hacer referencia directa a él.

### Tipos de Control de Acceso

El control de acceso a los archivos puede manejarse mediante una matriz para control de acceso que indica cuales usuarios tienen que tipos de acceso a cuales archivos. Lo más frecuente es que el control de acceso se maneje por clases de usuarios donde por ejemplo se puede conceder acceso al propietario ó aun usuario específico, a un miembro de un grupo ó a un miembro del público en general.

### Respaldo y Recuperación

Son funciones muy importantes de cualquier sistema de archivos. La técnica más común es el respaldo periódico. Otra técnica es vaciado por incrementos.

**Vaciado por Incrementos.-** Los archivos modificados por un usuario durante una sesión determinada se respaldan cuando ese usuario sale del sistema.

**Bitácora de Transacciones.-** Todas la líneas tecleadas por un usuario se copian en un archivo de bitácora. La recuperación implicaría literalmente una nueva aplicación de todas las transacciones realizadas después del último respaldo periódico principal.

### Bases de Datos

Es un conjunto integrado de datos controlados centralmente. Unas de la ventajas principales son:

1. Se puede reducir la redundancia.
2. Evitar la inconsistencia.
3. Compartir datos.
4. Imponer Normas.
5. Se puede aplicar restricciones de seguridad.
6. Mantener la integridad.
7. Se puede equilibrar requerimientos en conflicto.

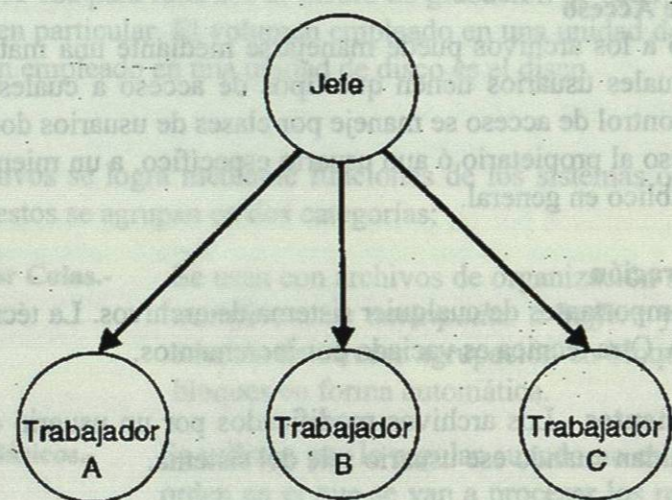
**Independencia de los datos.-** Hace posible modificar una aplicación y desarrollar nuevas aplicaciones sin tener que alterar la estructura del almacenamiento de los datos y la estrategia de acceso.

**Bases de Datos Distribuida.-** Está distribuida o dispersa en todos los sistemas de cómputo mediante una red.

### Tipos de Organización de Bases de Datos

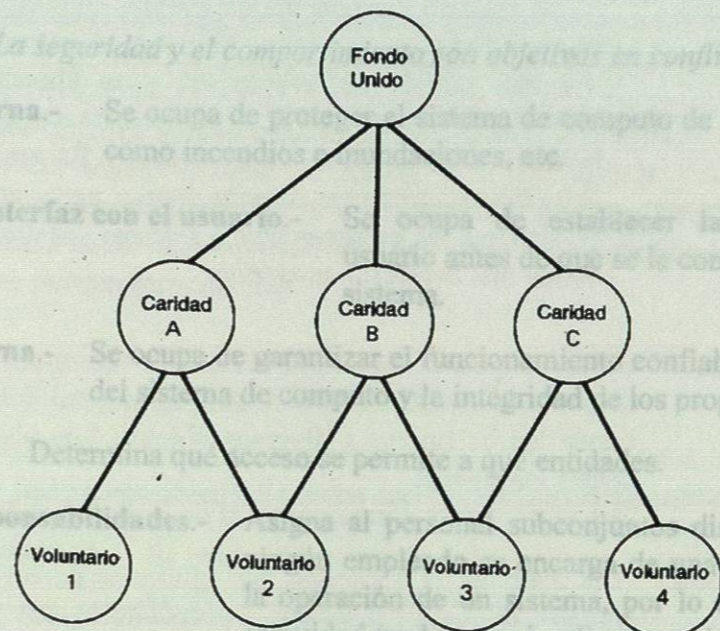
1. Jerárquica
2. Red
3. Relacional

En el enfoque *jerárquico* los datos se organizan según interrelaciones padre-hijo, cada hijo tiene un solo padre y cada padre puede tener muchos hijos. Búsquedas y mantenimientos fáciles, pero es limitada la flexibilidad del usuario para diseñar interdependencias complejas de los datos.



En el enfoque de *red* se pueden expresar en forma conveniente interdependencias muy generales, pero las estructuras resultantes pueden ser difíciles de entender, modificar o reconstruir en caso de falla (Red).

### SEGURIDAD EN LOS SISTEMAS OPERATIVOS



El enfoque *relacional* ofrece muchas ventajas con respecto a las anteriores, la representación tabular es más fácil de comprender y llevar a la práctica. Otros esquemas se pueden convertir con mucha facilidad a la organización relacional.

Relación: EMPLEADO

Número	Nombre	Departamento	Salario	Localización
23603	RAMOS, A.	413	1100	SEGOVIA
24568	CORTÉS, R.	413	2000	SEGOVIA
34589	LÓPEZ, P.	642	1800	ZARAGOZA
35761	MENDOZA, B.	611	1400	MADRID
47132	NOGUERA, C.	413	9000	SEGOVIA
78321	SALAS, T.	611	8500	MADRID

Una tupla

Clave primaria      Un dominio

Operaciones útiles como la proyección y la reunión facilitan la creación de nuevas relaciones. Los datos delicados pueden asegurarse colocándolos en relaciones separadas. Las búsquedas son más directas y rápidas, modificaciones directas y mejora la claridad y visibilidad de los datos.