

CONCEPTOS BÁSICOS DE ADMINISTRACIÓN DE ARCHIVOS

Bruno López Takeyas
Instituto Tecnológico de Nuevo Laredo
Reforma Sur 2007, C.P. 88250, Nuevo Laredo, Tamps. México
<http://www.itnuevolaredo.edu.mx/takeyas>
E-mail: takeyas@itnuevolaredo.edu.mx

1. Conceptos Básicos

Archivo: En términos computacionales es una colección de datos que tiene un nombre y se almacena regularmente en un disco o en una cinta.

Datos: Los elementos individuales de los archivos se llaman datos o campos. Por ejemplo un cheque de un banco tiene los siguientes campos: Cuenta habiente, Numero de cheque, Fecha, Persona a la que se le paga, Monto numérico, Monto con letra, Nota, Identificación del banco, Número de cuenta y Firma.

Registro: Es el conjunto completo de datos relacionados pertenecientes a una entrada, como un cheque de banco.

Cada campo tiene su longitud y tipo.

Tipos de archivo por contenido {
* Texto
* Binarios

Archivos de texto: Son aquellos que pueden contener cualquier clase de datos y de tal manera que son “entendibles” por la gente. Los datos en un archivo de texto se almacenan usando el código ASCII, en el cual cada carácter es representado por un simple byte. Debido a que los archivos de texto utilizan el código ASCII, se pueden desplegar o imprimir.

Archivos Binarios: Son aquellos que almacenan los datos numéricos con su representación binaria. Pueden ser archivos que contienen instrucciones en lenguaje máquina listos para ser ejecutados. Por ejemplo, cuando escribimos un programa en un lenguaje en particular (como C++, Pascal, Fortran, etc), tenemos las instrucciones almacenadas en un archivo de texto llamado programa fuente, pero una vez que lo sometemos a un proceso de compilación y ejecución nuestro programa lo trasladamos a un programa ejecutable (en lenguaje máquina), que es directamente entendido por la computadora y se crea un archivo binario.

Tipos de archivos por función {
* Maestro
* Trans.
* Reporte
* Trabajo

Archivo Maestro: Es un conjunto de registros relacionados con un aspecto importante de las actividades de una organización. Por ejemplo, una organización de manufactura puede tener un archivo maestro de nómina, de clientes, personal, inventario, etc. Los archivos maestros son útiles solo mientras se mantengan exactos y actualizados ya que reflejan el estado actual de eventos específicos o indicadores de la empresa y se logra por medio de transacciones.

Archivo de Transacciones: Es un archivo con dos propósitos acumular datos de los eventos al momento que ocurran y actualizar los archivos maestros para reflejar los resultados de las transacciones actuales. El término transacción se refiere a cualquier evento que afecte la organización y sobre el cual se calculan los datos. Por ejemplo, compras, pagos, contratar personal, pagar a empleados y registrar ventas. Los archivos maestros son permanentes y duran mientras exista el sistema. Sin embargo, los contenidos de los archivos cambian como resultado del procesamiento y la actualización. Por otro lado, los archivos de transacciones son temporales, ya que en algún momento ya no son necesarios y se borran o se destruyen.

Archivos de Reportes: Contiene datos que son formateados para su presentación al usuario. Pueden imprimirse o desplegarse en la pantalla. Regularmente son archivos temporales que se utilizan cuando el tiempo de impresión no está disponible para todos los reportes producidos, en este caso, la computadora escribe el reporte a un archivo de disco donde permanece hasta que pueda imprimirse.

Archivos de Trabajo: Es un archivo temporal. No tiene las características de Entrada/Salida de un archivo de transacciones o de reporte, ni las de largo plazo del archivo maestro. Se utiliza regularmente para pasar datos creados de un programa a otro.

Otro tipo de Archivos

- ? Programas: Contiene instrucciones para procesar datos, ya sea código fuente, ejecutables o compilados.
- ? Respaldos: Son copias de seguridad.

Organización de Archivos

La técnica utilizada para representar y almacenar registros en archivos es llamada organización de archivos. Las cuatro técnicas fundamentales son:

1. Secuencial
2. Secuencial indexado
3. Relativa (acceso directo)
4. Multi-llave

Hay aspectos básicos en que difieren las de organización de archivos.

1. La secuenciación de registros: Es decir la ordenación física de los registros almacenados.
2. La organización del archivo determina el conjunto de operaciones necesarias para encontrar registros.

Las operaciones básicas que se ejecutan en archivos son:

1. Creación
2. Actualización, incluyendo:

- ? Inserción de registros
- ? Modificación
- ? Supresión
- 3. Recuperación
 - ? Consulta
 - ? Reportes
- 4. Mantenimiento
 - ? Estructuración
 - ? Reorganización

2.- Organización de Archivos Secuenciales

En un archivo organizado secuencialmente, los registros quedan grabados consecutivamente cuando el archivo se crea y se deben accederse consecutivamente.

Ventajas y Desventajas.

La ventaja mas importante es la capacidad de acceder al "siguiente" registro rápidamente, además de que son muy sencillos de usar y de aplicar.

Si la secuencia de acceso a registros en un archivo secuencial es conforme al ordenamiento físico de los mismos, entonces los tiempos de acceso serán muy buenos, sin embargo, si el acceso no esta basado en el orden físico de los registros, entonces la eficiencia del programa puede ser terrible dando lugar a tiempos de acceso muy altos, provocando una desventaja.

Selección de la llave.

La llave de un archivo secuencial determina el orden en el cual son accedidos los registros, en otras palabras, es un campo del registro que sirve para identificarlo, diferenciarlo de otros registros y determinar el orden de acceso.

3.- Archivo Secuencial-Indexado.

Se utiliza este tipo de organización de archivo cuando existe la necesidad tanto de acceder los registros secuencialmente, por algún valor de llave, como de accederlos individualmente. Un archivo secuencial indexado proporciona la combinación de tipos de acceso que manejan un archivo secuencial y un archivo relativo o de acceso directo.

Estructura lógica de Archivo Secuencial-Indexado

En este tipo de organización de archivos se dispone de una tabla en que aparecen ordenados secuencialmente los números de la clave del archivo y asociados a cada uno de ellos se da la dirección del registro correspondiente.

4.- Organización de Archivos Relativos o de Acceso Directo

Se utiliza la organización relativa cuando existe la necesidad de acceder registros individuales directamente.

En un archivo relativo existe una relación predecible entre la llave usada para identificar un registro y su localización dentro del archivo. Sin embargo es importante comprender que el ordenamiento lógico de los registros no necesita tener ninguna relación con su secuencia física. Los registros no necesariamente aparecen físicamente ordenados de acuerdo al valor de sus llaves.

A cada archivo relativo debe definírsele una relación que será utilizada

para obtener una dirección física (o lógica) a partir de un valor llave. Esta relación R es una función de mapeo y se obtiene mediante métodos de conversión clave-dirección o técnicas hashing.

R(llave) _____ Dirección

Operaciones

Cuando se desea grabar en un archivo relativo, la función de mapeo R se usa para traducir el valor de la llave del registro a una dirección, la cual indica donde deberá almacenarse el registro.

Cuando es necesario recuperar el registro con un valor de llave particular, la función R es aplicada a ese valor de llave, traduciéndolo a una dirección donde se encuentra el registro.

Ventajas y desventajas

La ventaja principal de un archivo relativo es la habilidad de acceder registros individuales directamente.

Su desventaja radica (dependiendo de la función de mapeo) en que se puede presentar una misma dirección para diferentes valores de llave, es decir, se presentan colisiones.

$$R(K_1) = R(K_2) : \text{Colision}$$

Donde:

R = Función de mapeo

K_1, K_2 = Valores de clave o llave y

$K_1 \neq K_2$

Otra desventaja es el desperdicio de espacio debido a la distribución al azar de los registros.

5.- Bibliografía

- ? [García Badell, J. Javier. "Turbo C. Programación en manejo de archivos". Macrobit.](#)
- ? [Joyanes Aguilar, Luis. "Problemas de Metodología de la Programación". McGraw Hill. 1990.](#)
- ? [Loomis, Mary E.S. "Estructura de Datos y Organización de Archivos". Prentice Hall. México. 1991.](#)
- ? [Martin James. "Organización de las bases de datos". Prentice Hall. 1993.](#)
- ? [Rose, Cesar E.. "Archivos. Organización y Procedimientos". Computec. 1993.](#)
- ? [Sedgewick, Robert. "Algorithms". Second edition. Addison Wesley. USA. 1988.](#)
- ? [Sedgewick, Robert. "Algoritmos en C++". Addison Wesley. USA. 1995.](#)
- ? [Tsai, Alice Y. H. "Sistemas de bases de datos. Administración y uso". Prentice Hall. 1988.](#)