

**"2015. Año del Bicentenario Luctuoso de José María Morelos y Pavón"**

# **Centro Universitario UAEM Zumpango**

## **Ingeniería en Computación**

### **Unidad de Aprendizaje: Organización de Archivos**

#### **Unidad de Competencia I:**

Adquirir la capacidad para identificar los distintos medios de almacenamiento secundario y manejar los fundamentos de archivos.

**MTE-MI. Rosa Erendira Reyes Luna**

Agosto 2015





# Propósito de la Unidad de Aprendizaje

Que el docente adquiriera la habilidad para aplicar las estructuras en memoria secundaria y la organización de archivos como herramienta para la implementación de los principales algoritmos aplicables a la creación de programas de cómputo y software que de solución a problemas de almacenamiento, procesamiento y acceso de información.



# Introducción

Las estructuras de datos revisadas en unidades de aprendizaje anteriores se almacenaban en la memoria principal. Este tipo de almacenamiento se le conoce como almacenamiento primario, tiene la ventaja de su pequeño tiempo de acceso, el cuál es el mismo para acceder a cualquier parte del dispositivo.

Sin embargo no siempre es confiable guardar en la memoria principal debido a que este dispositivo tiene las siguientes limitaciones:

- La cantidad de datos que puede manipular un programa no puede ser muy grande debido a la capacidad de la memoria central de la computadora (aunque hoy en día esto ya casi no se cumple).
- La existencia de los datos en la memoria principal está supeditada al tiempo que la computadora está encendida y el programa ejecutándose (tiempo efímero).



# Introducción

Estas limitaciones dificultan:

- La manipulación de gran número de datos.
- Que la transmisión de salida de resultados de un programa pueda ser tratada como entrada a otro programa.

Para superar estas dificultades se requieren dispositivos de almacenamiento secundario, donde se almacenará la información o datos que podrán ser recuperados para su tratamiento posterior. Las estructuras de datos aplicadas a colección de datos en almacenamiento secundario se llaman organización de archivos.





**UAEM** | Universidad Autónoma  
del Estado de México.

Adquirir la capacidad para identificar los distintos medios de almacenamiento secundario y manejar los fundamentos de archivos.

## **UNIDAD DE COMPETENCIA I**



[www.uaemex.mx](http://www.uaemex.mx)



# Contenido

1. Clasificación de Archivos
2. Sistema de Archivos
3. Soportes Secuenciales y Direccionables
4. Organización de Archivos
5. Operaciones sobre Archivos
6. Archivos vistos desde C++
7. Gestión de Archivos
8. Conclusiones
9. Referencias





**UAEM** | Universidad Autónoma  
del Estado de México.

# CLASIFICACIÓN DE ARCHIVOS



[www.uaemex.mx](http://www.uaemex.mx)

## Clasificación de Archivos por función

1. Archivo Maestro
2. Archivo de Transacciones
3. Archivo de Reporte
4. Archivo de Trabajo
5. Archivo de programa
6. Archivo de Texto



## Clasificación de Archivos por función

- Archivo Maestro: representa una visión estática de algún aspecto de los negocios de una organización en un momento dado.

- Archivo de Transacciones: recolecta los cambios que han sido aplicados a los archivos maestros.

Una transacción puede contener datos para agregar un nuevo registro o para borrar o modificar un registro existente en un archivo maestro.

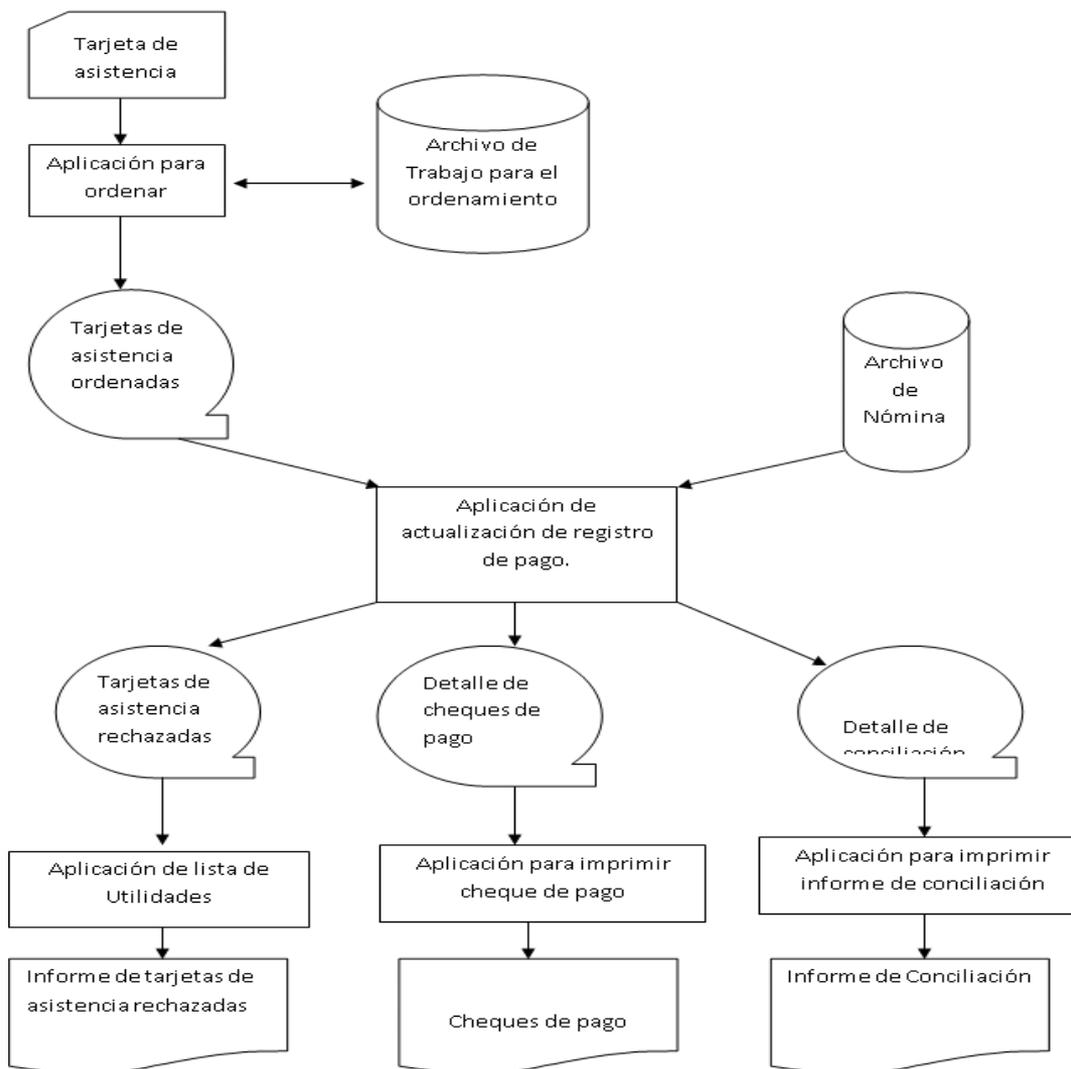
- Archivo de Reporte: contiene datos que son formateados para su presentación al usuario.



## Clasificación de Archivos por función

- Archivo de trabajo: es un archivo temporal en el sistema. No tiene las características de Entrada/Salida de un archivo de transacciones o de reporte ni las de largo plazo del archivo maestro.
- Archivo de programa: contiene instrucciones para procesar datos, los cuales pueden almacenarse en otros archivos o residir en la memoria principal.
- Archivo de texto: contiene datos alfanuméricos y gráficos ingresados a través de un programa editor de textos.





Actividad Llena la tabla de acuerdo a lo representado por el diagrama de flujo.

Archivo	Función
Tarjeta de Asistencia	
Aplicación para ordenar	
Archivo de trabajo para el ordenamiento	
Tarjetas de asistencia ordenadas	
Archivo de nómina	
Aplicación de actualización de registros de pagos	
Tarjetas de Asistencia rechazadas	
Detalle de cheques de pago	
Detalle de Conciliación	
Aplicación de lista de utilidades	
Informe de rechazos	
Aplicación para imprimir cheques de pago	
Cheques de pago	
Aplicación para imprimir reporte de conciliación	
Informe de Conciliación	





**UAEM** | Universidad Autónoma  
del Estado de México.

# SISTEMA DE ARCHIVOS



CENTRO UNIVERSITARIO UAEM  
ZUMPAANGO



[www.uaemex.mx](http://www.uaemex.mx)



## Sistema de Archivos

El acceso de datos (lectura y escritura) en archivos requiere de mucha actividad que es transparente al programador de la aplicación. Los lenguajes de programación permiten a los programadores definir técnicas de organización de archivos bastante complejas con instrucciones bastantes simples.

**Definición:** Es aquel que proporciona el apoyo que permite al programador acceder archivos sin preocuparse de los detalles sobre las características de almacenamiento y tiempos de acceso.

El sistema de archivo convierte a las instrucciones de acceso de archivos sencillos al programador en instrucciones de bajo nivel de Entrada/Salida.



## Responsabilidades de un Archivo

1. Mantener un directorio de identificación de archivos y localización de información.
2. Establecer rutas de flujo de datos entre la memoria principal y los dispositivos de almacenamiento secundario.
3. Coordinar la comunicación entre la unidad central de procesamiento (CPU) y los dispositivos de almacenamiento secundario y viceversa, lo cual incluye:
  - a. Manipular la desproporción en la rapidez de la CPU de la computadora y de los dispositivos de almacenamiento, de tal forma que la CPU no desperdicie una excesiva cantidad de tiempo, esperando ociosamente que las Entradas/Salidas sean completadas.
  - b. Manipular datos, de tal forma que éstos puedan ser conservados si el emisor (CPU o dispositivo de almacenamiento secundario) y el receptor (dispositivo de medio secundario o CPU) no están al mismo tiempo.
4. Preparar archivos como entrada o salida.
5. Manipular los archivos cuando su uso de entrada o salida haya terminado.



## Directorios

- Un directorio por si solo es un archivo que le pertenece al SO
- Provee mapeo entre los nombres de archivos y los mismos archivos
- Lista de entradas, una por cada archivo
- Archivo secuencial con el nombre del archivo como llave
- No provee ayuda para organizar los archivos
- Forzar al usuario a ser cuidadoso en no usar el mismo nombre para dos archivos
- Un directorio para cada usuario y un directorio maestro
- El directorio maestro contiene una entrada para cada usuario
  - Incluye una dirección e información de control de acceso
- Cada directorio de usuario es una simple lista de archivos del usuario
- No ofrece a los usuarios ayuda alguna para estructurar sus colecciones de archivos



## Soportes Secuenciales y Direccionables

El soporte es el medio físico donde se almacenan los datos. Los tipos de soportes utilizados en la gestión de archivos son:

- Soportes secuenciales.- son aquellos en los que los registros están escritos unos a continuación de otros y para acceder a un determinado registro  $n$  se necesita pasar por los registros  $n-1$  registros anteriores.

La secuencia puede corresponder al orden físico de los registros en el archivo o bien al orden de claves (ascendente o descendente) de los registros.

- Soportes Direccionables.- se estructuran de modo que las informaciones registradas se pueden localizar directamente por su dirección y no se requiere pasar por los registros precedentes.



## Organización de Archivos

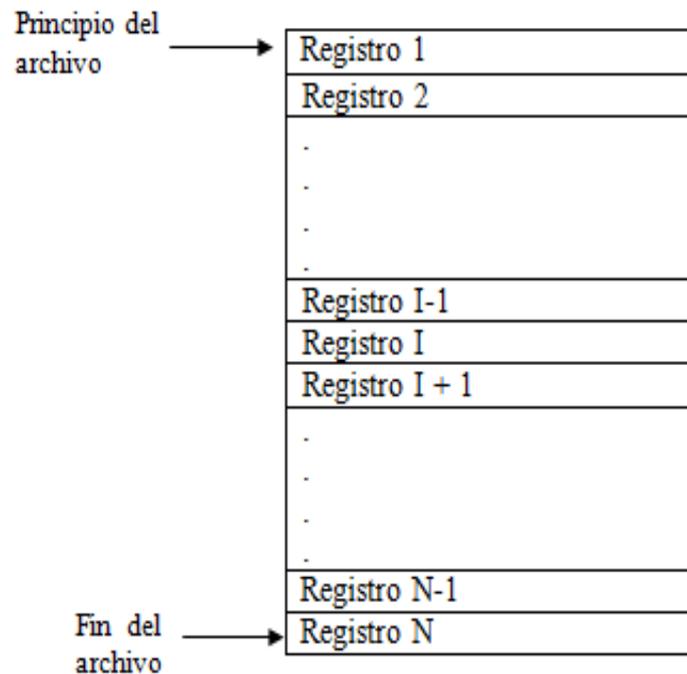
La organización de un archivo define la forma en que los registros se disponen sobre el soporte de almacenamiento, o también se define la organización como la forma en que se estructuran los datos en un archivo. En general se consideran tres organizaciones fundamentales:

- Organización secuencial
- Organización directa o aleatoria (random)
- Organización secuencial indexada (indexed)



## Organización secuencial

- Un archivo con esta organización es una sucesión de registros almacenados consecutivamente sobre el soporte externo, de tal modo que para acceder a un registro  $n$  dado es obligatorio pasar por todos los  $n-1$  artículos que le preceden.
- Los registros se graban consecutivamente cuando el archivo se crea y se debe acceder consecutivamente cuando se leen dichos registros.



# Organización Secuencial

- El orden físico en que fueron grabados (escritos) los registros es el orden de lectura de los mismos.
- Todos los tipos de dispositivos de memoria auxiliar soportan la organización secuencial.
- Los archivos organizados secuencialmente contienen un registro particular –el último- que contiene una marca fin de archivo detectable mediante las funciones
  - EOF.- End of File
  - FDA.- Final de Archivo
  - FDF.- Final de Fichero
  - FF.- Fin Fichero
  - FA.- Fin Archivo

Nota: Esta marca fin de archivo suele ser un carácter especial, como '\*'.





## Organización Directa

- Un archivo está organizado en modo directo cuando el orden físico no se corresponde con el orden lógico. Los datos se sitúan en el archivo y se accede a ellos directa-aleatoriamente-mediante su posición, es decir, el lugar relativo que ocupan.
- Ventaja.- esta organización permite que se puedan leer y escribir los registros en cualquier orden y posición; además son de muy rápido acceso a la información que contienen.
- Desventaja.- Esta organización requiere programar la relación existente entre el contenido de un registro y la posición que ocupa; además el acceso a los registros en modo directo implica la posible existencia de huecos libres dentro del soporte y, por consecuencia, pueden existir huecos libres entre los registros.



## Organización Directa

- Las condiciones para que un archivo sea de organización directa son:
  - Almacenamiento en un soporte direccionable.
  - Los registros deben contener un campo específico denominado clave que identifica a cada registro como único, es decir, dos registros no pueden tener el mismo valor clave.
  - Existencia de una correspondencia entre los posibles valores de la clave y las direcciones posibles sobre el soporte.
- En la práctica el programador no gestiona directamente relaciones absolutas, sino direcciones relativas respecto al principio del archivo. La manipulación de dirección relativa permite diseñar el programa con independencia de la posición absoluta del archivo en el soporte.



## Organización Secuencial Indexada

- El diccionario es un ejemplo típico de archivo secuencial indexado con dos niveles de índices, el nivel superior para las letras iniciales y el nivel menor para las cabeceras de página.
- En una organización de computadora las letras y las cabeceras de páginas se guardarán en un archivo de índice independiente de las entradas del diccionario (archivo de datos). Por lo que cada archivo secuencial indexado consta de un archivo índice y un archivo de datos.



## Organización Secuencial Indexada

- Un archivo está organizado en forma secuencial indexada si:
  - El tipo de sus registros contiene un campo clave identificador.
  - Los registros están situados en un soporte direccionable por el orden de los valores indicados por la clave.
  - Existe un índice con cada una de las posiciones direccionables, que almacena la dirección de la posición y el valor de la clave; en esencia, el índice contiene la clave del último registro y la dirección de acceso al primer registro del bloque.





**UAEM** | Universidad Autónoma  
del Estado de México.

# OPERACIONES SOBRE ARCHIVOS



CENTRO UNIVERSITARIO UAEM  
ZUMPAANGO



[www.uaemex.mx](http://www.uaemex.mx)



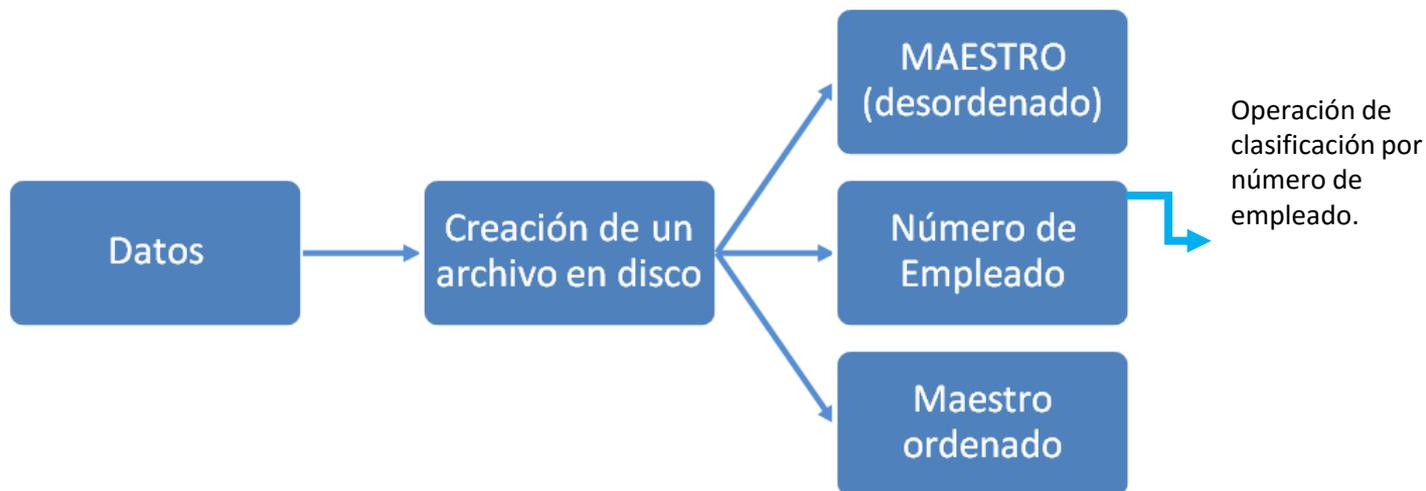
## Operaciones sobre Archivos

- Las distintas operaciones que se pueden realizar a los archivos son las siguientes:
  - Creación
  - Consulta
  - Actualización (altas, bajas, modificación, consulta)
  - Clasificación
  - Reorganización
  - Destrucción (borrado)
  - Reunión/Fusión
  - Rotura/Estallido



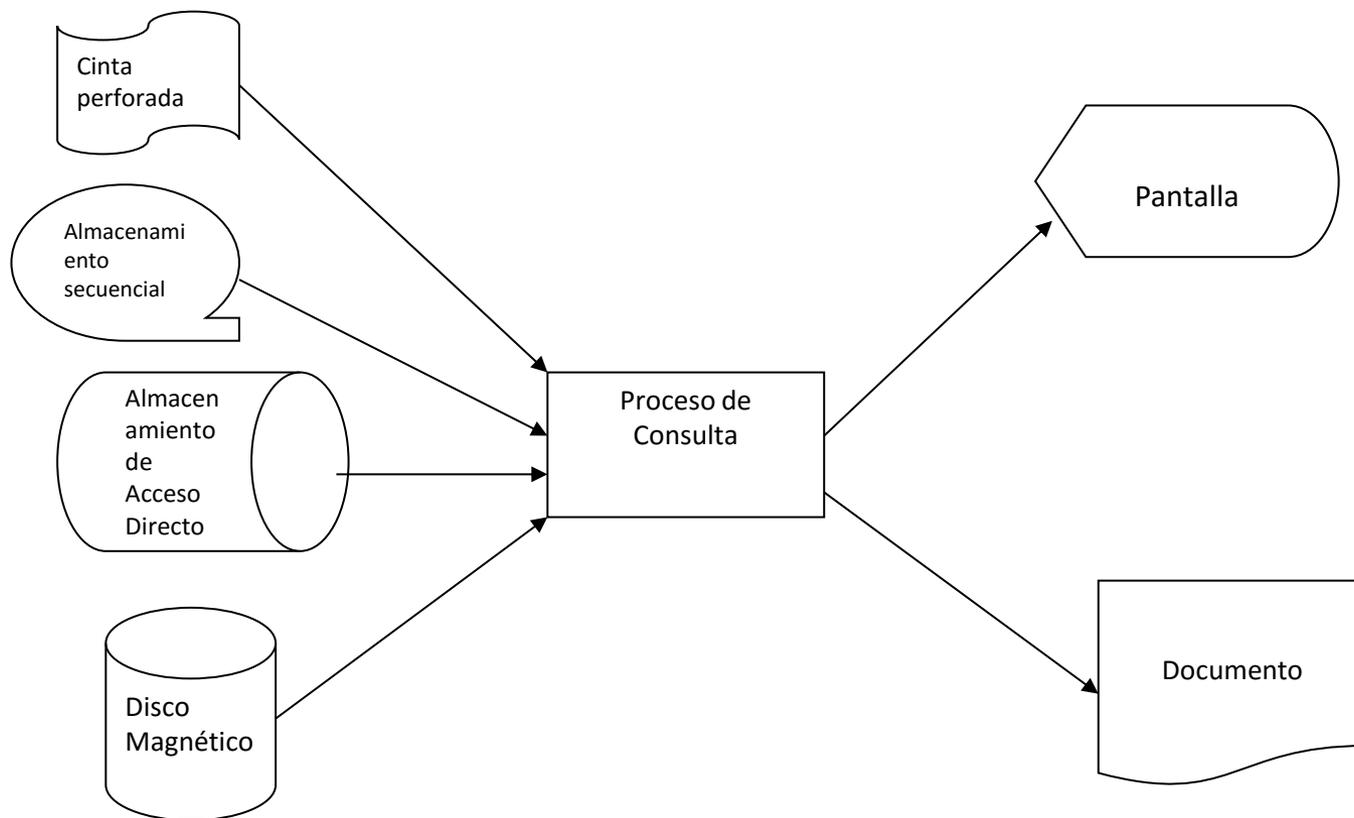


# Creación de un archivo



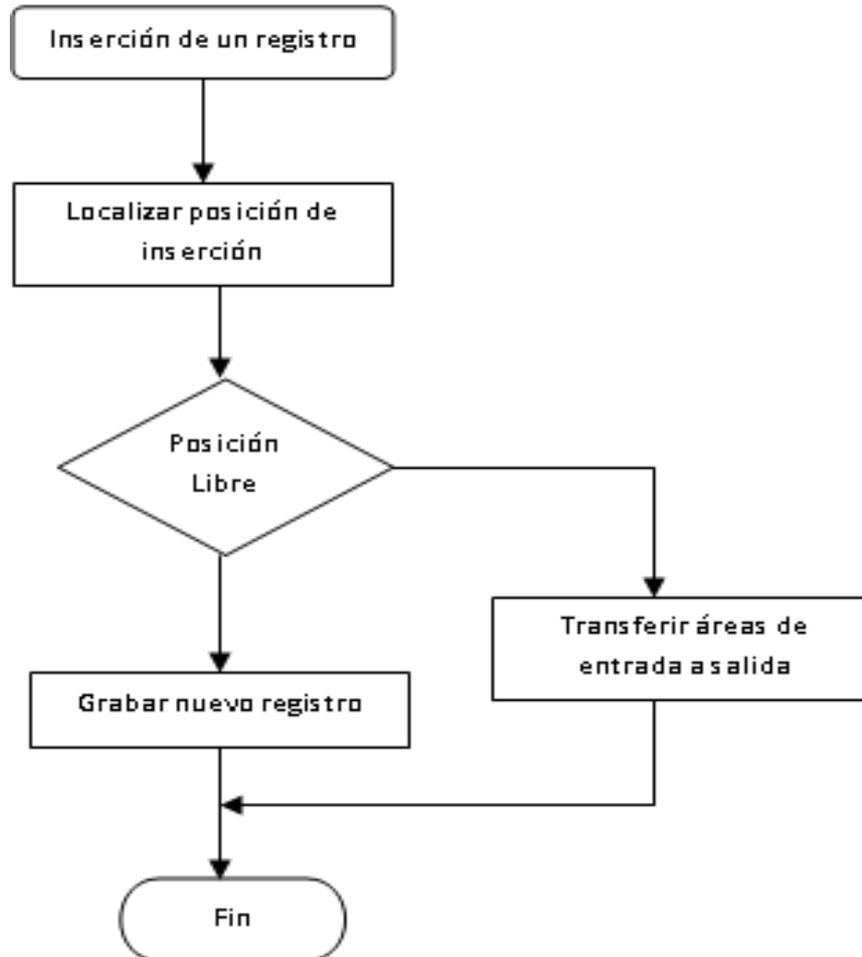


# Consulta de un Archivo



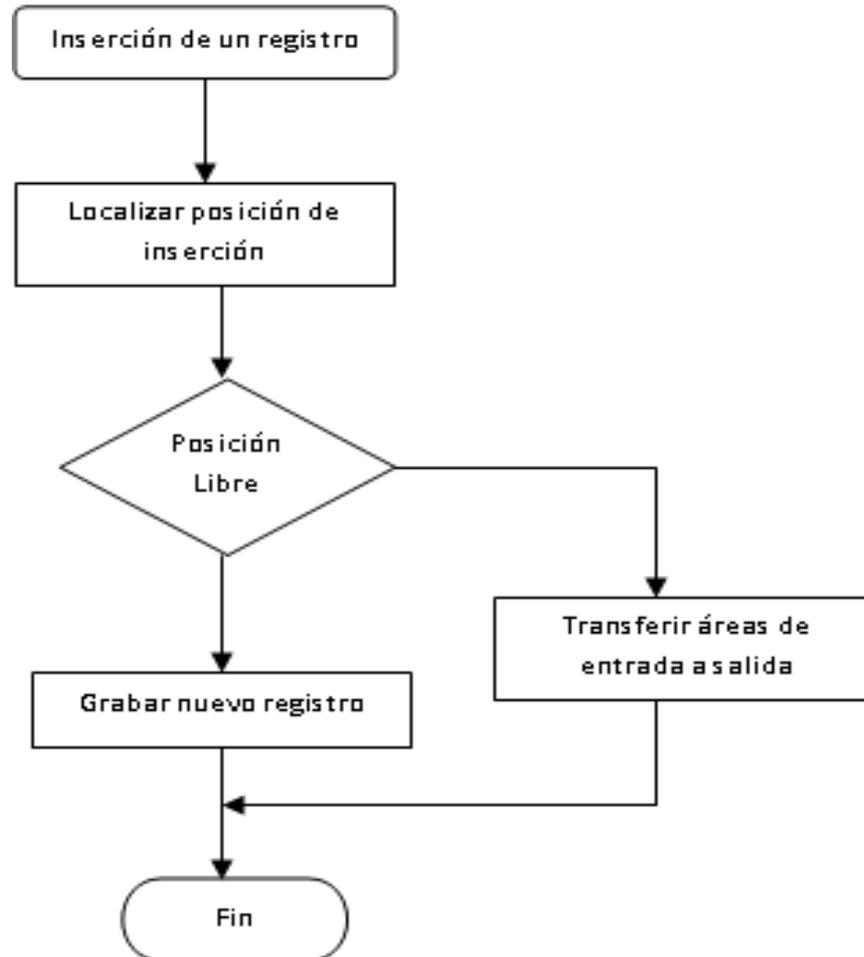


# Actualización de un Archivo





# Actualización de un Archivo



## Clasificación de un Archivo

Una operación muy importante en un archivo es la clasificación u ordenación (sort, en inglés).

Esta clasificación se realizará de acuerdo con el valor de un campo específico, pudiendo ser ascendente (creciente) o descendente (decreciente): alfabética o numérica.





## Reorganización de un Archivo

- Las operaciones sobre archivos modifican la estructura inicial o la óptima de un archivo. Los índices, enlaces (punteros), zonas de sinónimos, zonas de desbordamiento, etc. se modifican con el paso del tiempo, lo que hace a la operación del acceso al registro sea cada vez más lenta.
- La reorganización suele consistir en la copia de un nuevo archivo a partir del archivo modificado, a fin de obtener una nueva estructura lo más óptima posible.





## Destrucción de un Archivo

Es la operación inversa a la creación de un archivo (kill, en inglés).

Cuando se destruye (anula o borra) un archivo, éste ya no puede ser utilizado y, por consiguiente, no se podrá acceder a ninguno de sus registros.



## **Destrucción de un Archivo**

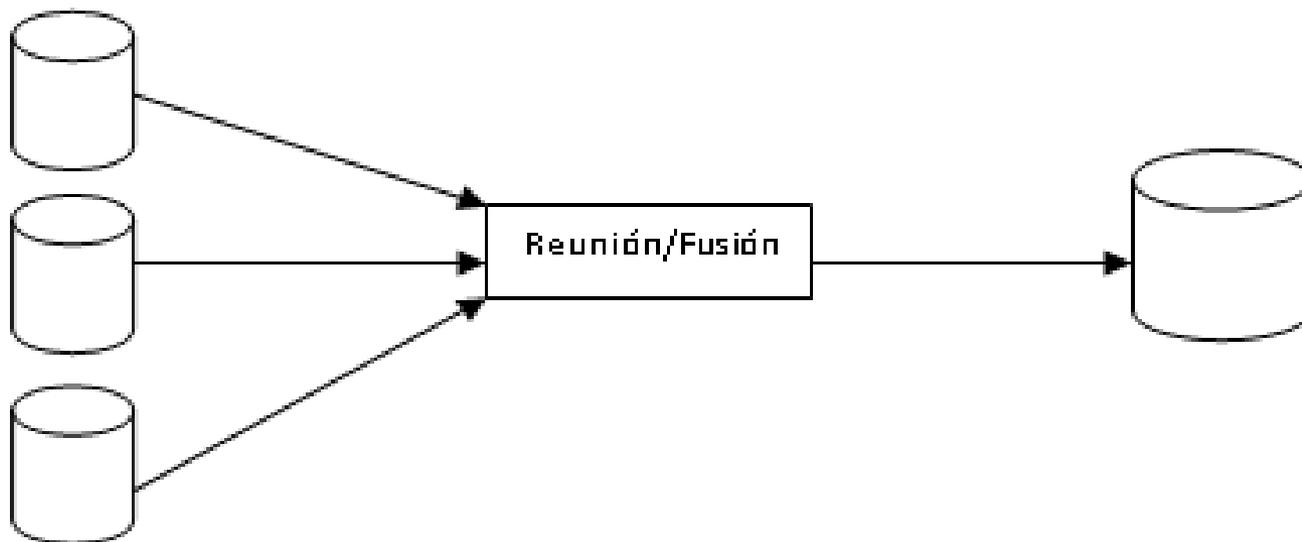
Es la operación inversa a la creación de un archivo (kill, en inglés).

Cuando se destruye (anula o borra) un archivo, éste ya no puede ser utilizado y, por consiguiente, no se podrá acceder a ninguno de sus registros.



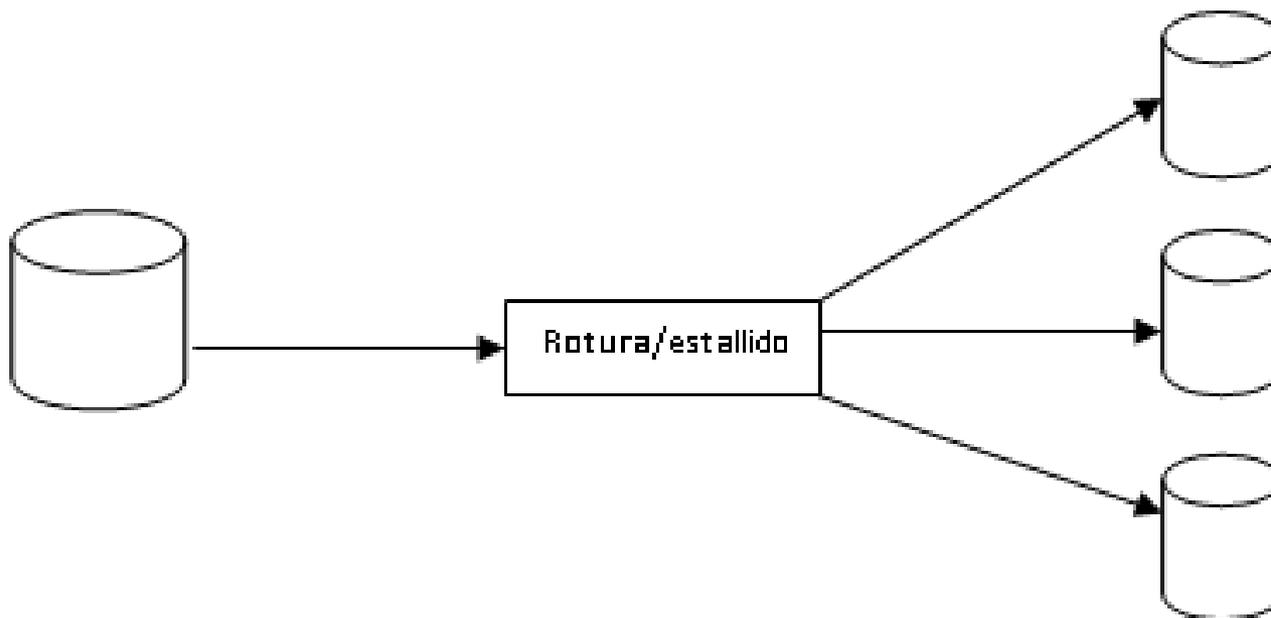


# Reunión/Fusión de un Archivo



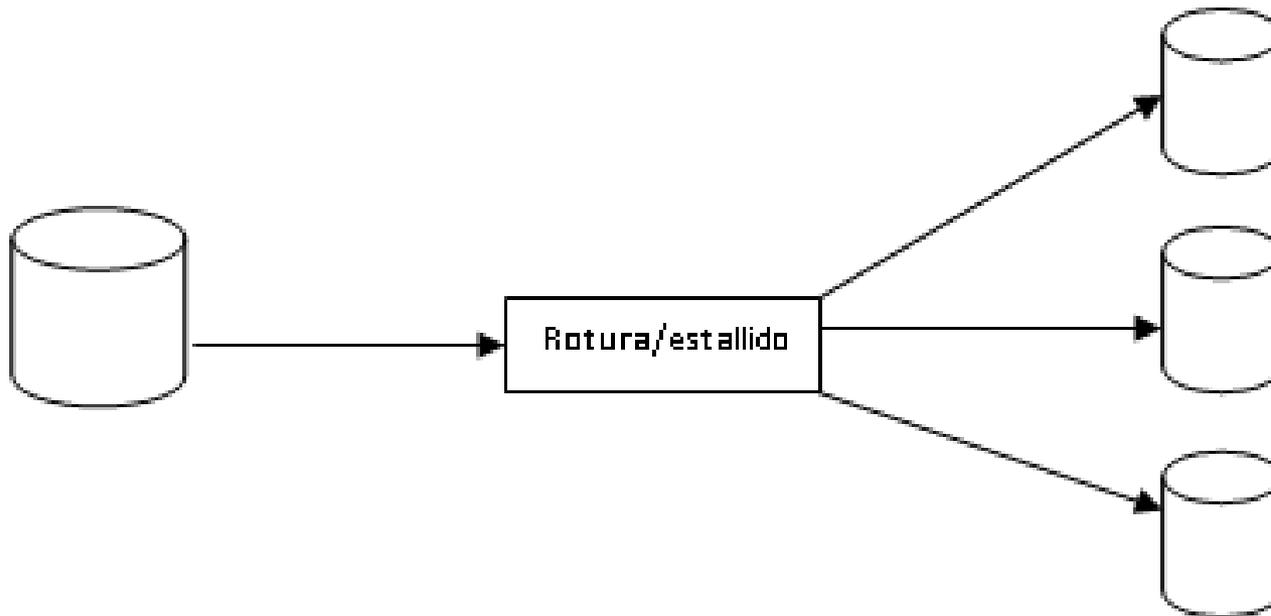


# Reunión/Fusión de un Archivo



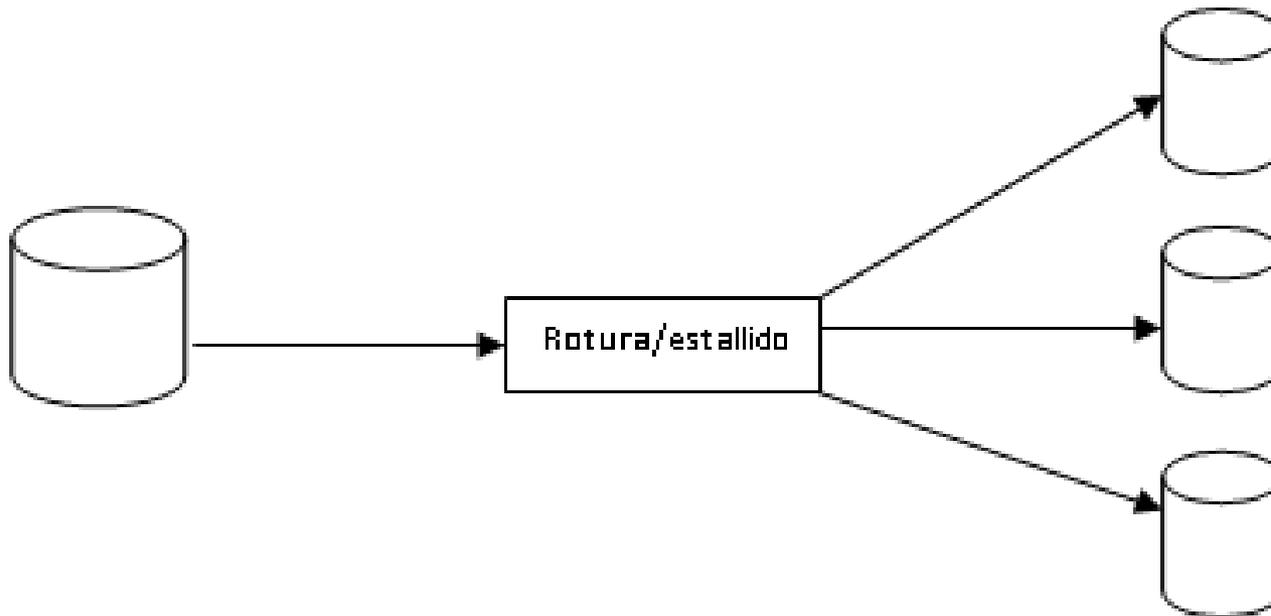


# Rotura/estallido de un Archivo





# Rotura/estallido de un Archivo





**UAEM** | Universidad Autónoma  
del Estado de México.

# ARCHIVOS VISTOS DESDE C++



CENTRO UNIVERSITARIO UAEM  
ZUMPAANGO



[www.uaemex.mx](http://www.uaemex.mx)

C++ ve a cada archivo como una secuencia de bytes, tal como lo muestra la imagen:



Cada archivo termina con un marcador de fin de archivo o en un número byte específico, registrado en una estructura de flujo con ese objeto.

Para llevar a cabo el procesamiento de archivo en C++, deben incluirse los archivos de encabezado `<iostream.h>` y `<fstream.h>`.



# Modos de apertura de archivos en C++

Modo	Descripción
<code>ios::app</code>	Anexa toda la salida al final del archivo
<code>ios::ate</code>	Abre un archivo en modo de salida y se desplaza al final de éste (por lo general se utiliza para anexar datos a un archivo). Los datos pueden describirse en cualquier parte del archivo.
<code>ios::in</code>	Abre un archivo en modo de entrada
<code>ios::out</code>	Abre un archivo en modo de salida
<code>ios::trunc</code>	Descarta el contenido del archivo si es que existe
<code>ios:: binary</code>	Abre un archivo en modo binario de entrada o salida





## Gestión de Archivos

Las operaciones más usuales en los registros son:

- Consulta: lectura del contenido de un registro
- Modificación: alterar la información contenida en un registro
- Inserción: añadir un nuevo registro al archivo
- Borrado: suprimir un registro del archivo



# Gestión de Archivos

Las operaciones sobre archivos se realizan mediante programas, de modo tal que los archivos se identifican por un nombre externo al que están asociados. Pueden existir programas que procesen el mismo archivo de datos. La mayoría de los programas ejecutarían las siguientes clases de funciones:

- Crear archivos (create)
- Abrir o arrancar (open) un archivo que fue creado con anterioridad a la ejecución de este programa.
- Cerrar el archivo después que el programa ha terminado de utilizarlo (close)
- Borrar (delete) un archivo ya existente
- Transferir datos desde (leer) o a (escribir) el dispositivo utilizado por el programa



# Gestión de Archivos

Las operaciones sobre archivos se realizan mediante programas, de modo tal que los archivos se identifican por un nombre externo al que están asociados. Pueden existir programas que procesen el mismo archivo de datos. La mayoría de los programas ejecutarían las siguientes clases de funciones:

- Crear archivos (create)
- Abrir o arrancar (open) un archivo que fue creado con anterioridad a la ejecución de este programa.
- Cerrar el archivo después que el programa ha terminado de utilizarlo (close)
- Borrar (delete) un archivo ya existente
- Transferir datos desde (leer) o a (escribir) el dispositivo utilizado por el programa



## Conclusiones

- Un archivo está siempre almacenado en un soporte externo a la memoria central.
- Existe independencia de las informaciones respecto de los programas.
- Todo programa de tratamiento intercambia información con el archivo y la unidad básica de entrada/salida es el registro.
- La información almacenada es permanente.
- En un momento dado, los datos extraídos por el archivo son los de un registro y no los de un archivo completo.



## Referencias

- Joyanes, Luis; Zahonero, Ignacio. 2004. Estructuras de datos, Algoritmos, Abstracción y Objetos. McGraw Hill. España.
- Joyanes, Luis; Fernández, Matilde; Sánchez, Lucas; Zahonero, Ignacio. 2005. Estructuras de datos en C. McGraw Hill. España.
- Hopcroft, John E.; Ullman, D. Jeffrey. 2002. Estructura de datos y Algoritmos. Pearson Addison Wesley Longman. México.





**UAEM** | Universidad Autónoma  
del Estado de México.

## GRACIAS

Organización de Archivos  
Unidad de Competencia I  
MTE-MI. Rosa Erendira Reyes Luna  
rereyesl@uaemex.mx



[www.uaemex.mx](http://www.uaemex.mx)



**UAEM** | Universidad Autónoma  
del Estado de México.

# GUÍA EMPLEO DE MATERIAL

Organización de Archivos  
Unidad de Competencia I



[www.uaemex.mx](http://www.uaemex.mx)

# GUÍA EMPLEO DE MATERIAL

Las primeras diapositivas muestran el propósito, justificación y objetivos de la unidad de aprendizaje. Se presentan para que el alumno identifique dichos elementos.

El contenido, conforme a la unidad de aprendizaje, maneja los temas de un menor a mayor grado de dificultad.

La diapositiva 12 tiene como objetivo corroborar que el estudiante haya comprendido la clasificación de archivos conforme a su función dentro de un sistema informático.



# GUÍA EMPLEO DE MATERIAL

De las dispositivas 26 a la 38 explican las operaciones que se llevan a cabo con los archivos, esquematizando el procedimiento y resultado al trabajar con los registros contenidos en el archivo.

Las diapositivas 40 y 41 muestran la forma en que el lenguaje de programación C++ reconoce los archivos.

