

Almacenamiento Secundario y Terciario

Unidad 4.

Almacenamiento y Estructuras de Archivos

Medios de Almacenamiento

- Los medios de almacenamiento son memoria externa
- El acceso a estos dispositivos siempre es más lento que el acceso a memoria
- Medio de Almacenamiento Secundario
 - Almacenamiento a largo plazo
- Medio de Almacenamiento Terciario
 - Enfocados en el almacenamiento industrial

Discos Magnéticos

Discos Magnéticos

- Mejor conocidos como discos duros (*Hard Drive*)
- Son un ensamble de discos de aluminio recubiertos por una capa de óxido ferroso

Almacenamiento en un Disco

- La información se almacena en n superficies que giran alrededor de un eje
- Cada superficie cuenta con una cabeza de lectura/escritura unida a un peine de cabezas
- Cada superficie cuenta con n pistas
- Cada pista se divide en n sectores

Conceptos Fundamentales

- Pista. Superficie de una cara de un disco que queda debajo de una cabeza de L/E
- Sector. Cantidad mínima de información que es leída o escrita en una operación de E/S. Actualmente un valor común para la cantidad de sectores es 36 sectores con una capacidad de 512 bytes cada uno

Conceptos Fundamentales

- Cilindro. Conjunto de pistas que quedan debajo una de otra en diferentes superficies. Todas las pistas en un cilindro dado pueden ser accedidas sin tener que mover el brazo de acceso

Cálculo de la Capacidad de un Disco

- Se parte de la capacidad en bytes de un sector (C_{sec})
- Capacidad de pistas (C_{pis}) se calcula como $C_{\text{sec}} \times N_{\text{sec}}$ (número de sectores)
- Capacidad de superficie (C_{sup}) se calcula como $C_{\text{pis}} \times N_{\text{pis}}$ (número de pistas)
- Capacidad de disco (C_{dis}) se calcula como $C_{\text{sup}} \times N_{\text{sup}}$ (número de superficies)

Lectura

- Cuando una instrucción de *lectura* busca un byte en particular de un archivo en un disco, el sistema operativo encuentra la superficie correcta, pista y sector, posteriormente se lee el sector entero en un área especial de memoria llamada *buffer* y encuentra el byte solicitado en ese buffer

Bloque

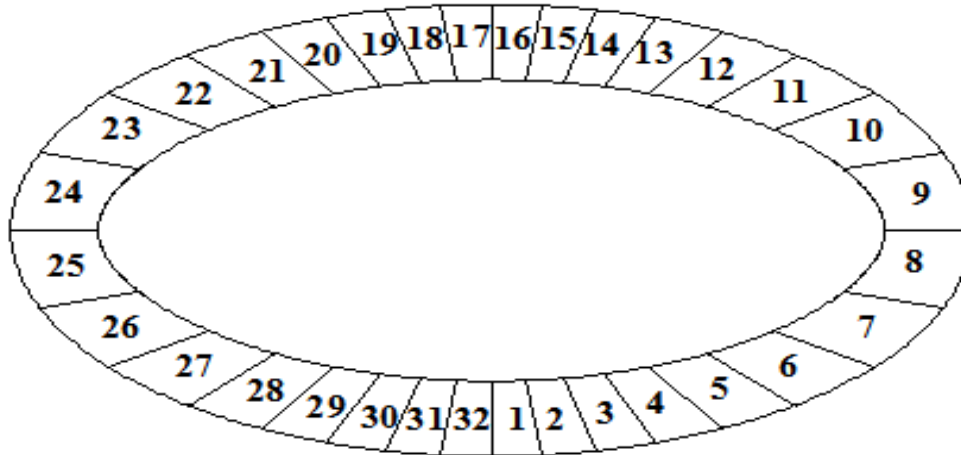
- Es la unidad de la organización de datos correspondiente a la cantidad de datos transferidos en un solo acceso.
- Un bloque es llamado ocasionalmente un registro físico; un sector es ocasionalmente llamado un bloque

Cluster

- Cantidad de espacio que consiste en un grupo de sectores contiguos
- Una vez que se localiza un cluster (o grupo de sectores contiguos) se pueden leer todos los sectores sin tener que realizar una búsqueda

Manejo Lógico de los Sectores

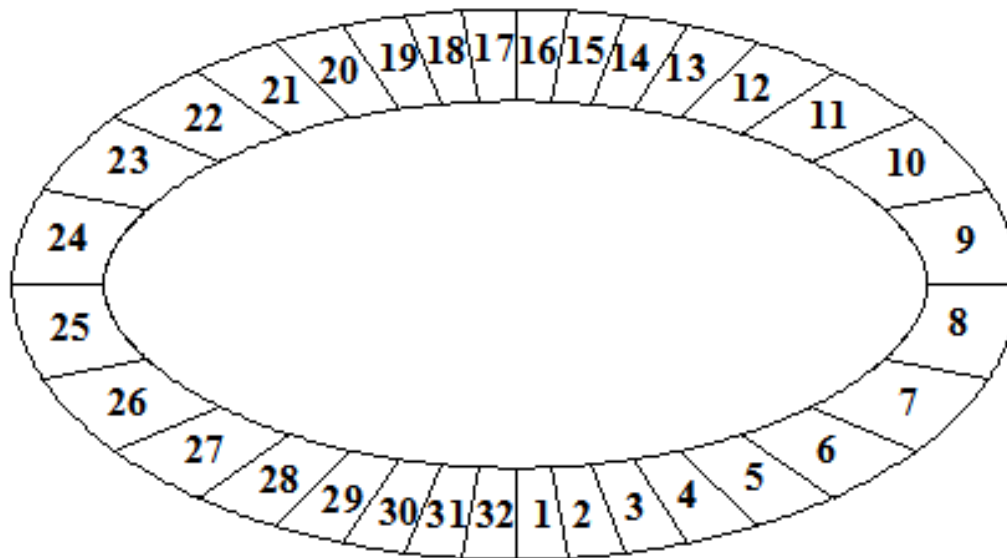
- Una opción es numerar a los sectores como están acomodados de manera física



- La principal desventaja es que para leer sectores contiguos podría tener que dar una vuelta completa

Manejo con Intervalos

- Una manera más conveniente de manejarlo es con un intervalo de sectores físicos entre sectores lógicos adyacentes



Manejo con Intervalos

- Para poder ver al archivo como una serie de clusters y mantener el esquema de sectores, se ata un sector lógico a su correspondiente sector físico usando una tabla FAT (*file allocation table*)
- Esta tabla contiene una lista de los clusters en un archivo ordenados de acuerdo al orden lógico de los sectores que contienen

Costos de Acceso a Disco

Tiempos de Acceso

- Un acceso a disco se divide en tres distintas operaciones físicas, cada una de las cuáles tiene su propio costo:
 - Tiempo de búsqueda
 - Retraso por rotación
 - Tiempo de transferencia

Tiempo de Búsqueda

- Tiempo requerido para mover el brazo de acceso al cilindro correcto
- La cantidad de tiempo gastado en una búsqueda durante un acceso a disco depende de que tan lejos tenga el brazo que moverse
- Siempre se busca minimizar el costo de ésta operación

Tiempos de Búsqueda Típicos

- El valor mínimo es normalmente el necesario para viajar de una pista a la pista contigua
- El máximo es el necesario para recorrer desde la primera a la última pista.
 - tiempo de búsqueda mínimo(tb_{min}) 2 ms
 - tiempo de búsqueda promedio(tb_{pro}) 12 ms
 - tiempo de búsqueda máximo(tb_{max}) 20 ms

Retraso por Rotación

- Se refiere al tiempo que toma al disco rotar para que el sector que se desea quede debajo de la cabeza de lectura/escritura
- En promedio el retraso rotacional es la mitad del tiempo de una rotación.
- El tiempo de rotación se calcula dividiendo 60 (60 seg tiene un minuto) entre la velocidad de rotación
- Los discos duros actuales rotan entre 3600 y 7200 rpm, lo que significa una rotación entre 16mseg y 8mseg en promedio.
- De esta manera el retraso rotacional es la mitad del tiempo de una rotación, 8mseg o 4mseg

Tiempo de Transferencia

- Cuando el dato al que se desea se encuentra debajo de la cabeza de lectura/escritura, éste debe ser transferido
- El tiempo de transferencia esta dado por la siguiente fórmula:
- Tiempo de transferencia =

$$\frac{\text{número de bytes transferidos}}{\text{número de bytes en una pista}} \times \text{tiempo de rotación}$$

Medios de Almacenamiento Óptico

Actividad

- Realizar la actividad relacionada con los medios de almacenamiento óptico

Cintas Magnéticas

Introducción

- Dispositivos que no proporcionan un método de acceso directo
- Compactas, resistentes a las condiciones ambientales y fáciles de almacenar y transportar
- Actualmente son principalmente utilizadas como resguardo de información

Composición

- Una cinta puede verse como si estuviera dividida en pistas
- Cada pista es una sucesión de bits
- Un marco se refiere a los bits que se encuentran en la misma posición de todas las pistas
- El último bit de un marco (columna) se utiliza para verificar la información

Composición

- Esta verificación se utiliza mediante la técnica de paridad par o impar de la cantidad de unos en un marco
- Los marcos se agrupan en bloques
- Cada bloque se separa por un espacio sin datos (*Inter Record Gap*, IRG)
- El espacio tiene como objetivo sincronizar el flujo de datos

Composición

Pista 1							
Pista 2	M						
Pista 3	A			I			
Pista 4	C						
Pista 5	R			R			
Pista 6	O						
Pista 7				G			
Pista 8	Bit Paridad	Bit Paridad	Bit Paridad		Bit Paridad	Bit Paridad	Bit Paridad

Propiedades

- Las cintas vienen en distintas formas, tamaños y velocidades, el desempeño de ellas puede ser medida en base a tres elementos:
 - Densidad de la cinta. Comúnmente de 800 a 30,000 bits por pulgada (bpi) por pista
 - Velocidad de la cinta. Comúnmente de 30 a 200 pulgadas por segundo (ips)
 - Tamaño de los intervalos. Comúnmente entre 0.3 y 0.75 pulgadas

Longitud Requerida

- Calcular la cantidad de cinta requerida para almacenar cierta cantidad de información
- Existen dos contribuidores principales:
 - Intervalo de brecha
 - Bloques de datos

Información Necesaria

- La información involucrada en el cálculo es:
 - $b = \text{longitud física de un bloque de datos}$
 - $g = \text{longitud del intervalo de brecha}$
 - $n = \text{el número de bloques de datos}$

Cálculo de b

- b es un valor que depende de la cantidad de bytes por bloque que pueda contener

$$b = \frac{\textit{bits por bloque}}{\textit{bits por pulgada}}$$

Cálculo de n

- Este valor indica la cantidad registros que contendrá cada bloque
- El número de registros almacenados en un bloque físico es llamado el *factor de bloque*

Cálculo de n

- En general el cálculo de n es el siguiente:
- $n = \text{nRegistros} / \text{fBloque}$
- Donde:
 - n es el número de bloques de datos
 - nRegistros es el número de registros del archivo
 - fBloque es el factor de bloque

Longitud de la Cinta

- Finalmente la longitud de la cinta está dada por la siguiente fórmula:

$$s = n \times (b + g)$$

Ejemplo

- Se desea almacenar una copia de resguardo de un archivo con un millón de registros de 100 bytes

Datos Necesarios

- La cinta tiene las siguientes propiedades:
 - Densidad de 6,250 bpi
 - Intervalo de brecha de 0.3 pulgadas
 - Se utilizará un bloque para cada registro
 - El factor de bloque es 1

Ejemplo

- Cálculo de b

$$b = \frac{\text{bytes por bloque}}{\text{bytes por pulgada}} = \frac{800 \text{ bits}}{6250 \text{ bpi}} = 0.128 \text{ pulgadas}$$

- Cantidad de cinta

$$\begin{aligned} s &= 1,000,000 \times (0.128 + 0.3) \text{ pulgadas} \\ &= 1,000,000 \times 0.428 \text{ pulgadas} \\ &= 428,000 \text{ pulgadas} \\ &= 35,666 \text{ ft} \\ &= 35666 \times 0.3 = 10,700 \text{ m} \end{aligned}$$

Problemas

- El tamaño más común de las cintas es 2400ft
- El espacio se puede reducir incrementando el factor de bloque
- Una medida efectiva del efecto de la elección de diferentes tamaños de bloque es la *densidad efectiva de almacenamiento*

Densidad Efectiva de Almacenamiento

- Refleja la cantidad de datos que pueden ser almacenados por pulgada en una cinta, se define:

$$\frac{\text{número de bytes por bloque}}{\text{longitud física de un bloque de datos} + \text{longitud del intervalo de brecha}}$$

- En el ejemplo, se utiliza un bloque por registro
- Se tienen 100 bytes por bloque
- La densidad efectiva es:

$$\frac{800 \text{ bits}}{0.428 \text{ pulgadas}} = 1869 \text{ bpi}$$

Tiempos de Transmisión

- Dos factores que afectan el rango de transmisión de datos de o hacia la cinta son: la tasa nominal de transferencia y la tasa efectiva de transferencia

Tasa nominal de transferencia

- Indica la velocidad máxima a la que se pueden transmitir datos de y hacia la cinta
- Se calcula con:

$$\text{tasa} = d\text{Cinta} * v\text{Cinta}$$

- Donde:
 - $d\text{Cinta}$ es la densidad de la cinta
 - $v\text{Cinta}$ es la velocidad de la cinta

Tasa real de transferencia

- Es la velocidad real a la que se pueden transmitir datos de y hacia la cinta
- Se calcula con:

$$\frac{\text{bytes por bloque}}{b + s} \times v$$

Del ejemplo

- Si la cinta tiene una velocidad de 200ips, la tasa nominal de transferencia es:

$$6250 \times 200 = 1,250,000 \text{ bits/sec}$$

- La tasa real de transferencia es:

$$1869 \times 200 = 373,800 \text{ bits/sec}$$

Medios de Estado Sólido

- Realizar una investigación sobre los medios de almacenamiento de estado sólido:
 - Historia de su aparición
 - ¿Qué utilizan en lugar de elementos físicos o magnéticos?
 - Capacidad máxima de almacenamiento
 - Ventajas de estos medios sobre los discos duros tradicionales
 - Desventajas a comparación de los discos duros
 - Tipo de memoria que utilizan