

GESTIÓN DE E/S

Capítulo 7

Contenidos

- Introducción
- Software de E/S
 - ▣ Objetivos
 - ▣ Niveles
- Discos
 - ▣ Hw de disco
 - ▣ Sw de disco
 - Traducción de direcciones lógicas/físicas
 - Algoritmos de planificación del brazo

Conceptos básicos

Gestión de Entrada/Salida

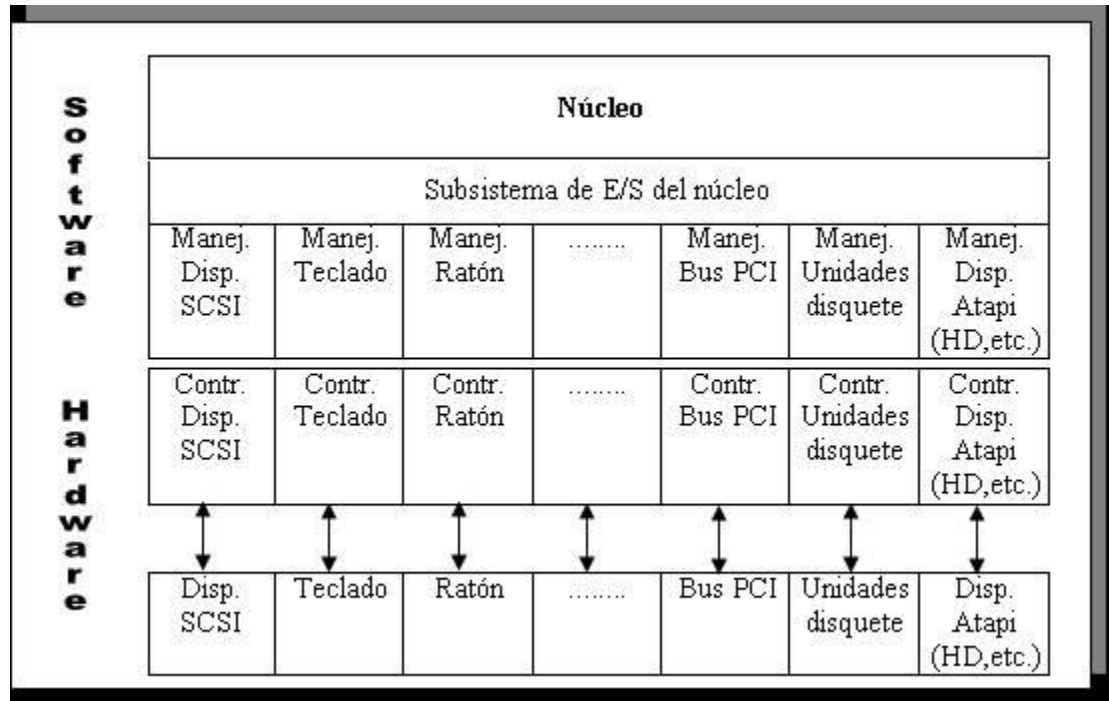
- Parte del SO encargada de la gestión de los dispositivos de E/S que actúa como interfaz entre los usuarios del sistema y los dispositivos de E/S, de forma que en las capas superiores los dispositivos se tratan de una manera uniforme, sencilla, segura y eficaz (**Independencia de Dispositivo**).

Conceptos básicos

Sistema de Entrada/Salida

Se divide en:

- Hardware de E/S. Formado por el dispositivo físico de E/S que es la parte electromecánica y el controlador del dispositivo que es la parte electrónica.
- Software de E/S. Formado por el software que permite la comunicación con el controlador del dispositivo.



Software de entrada/salida

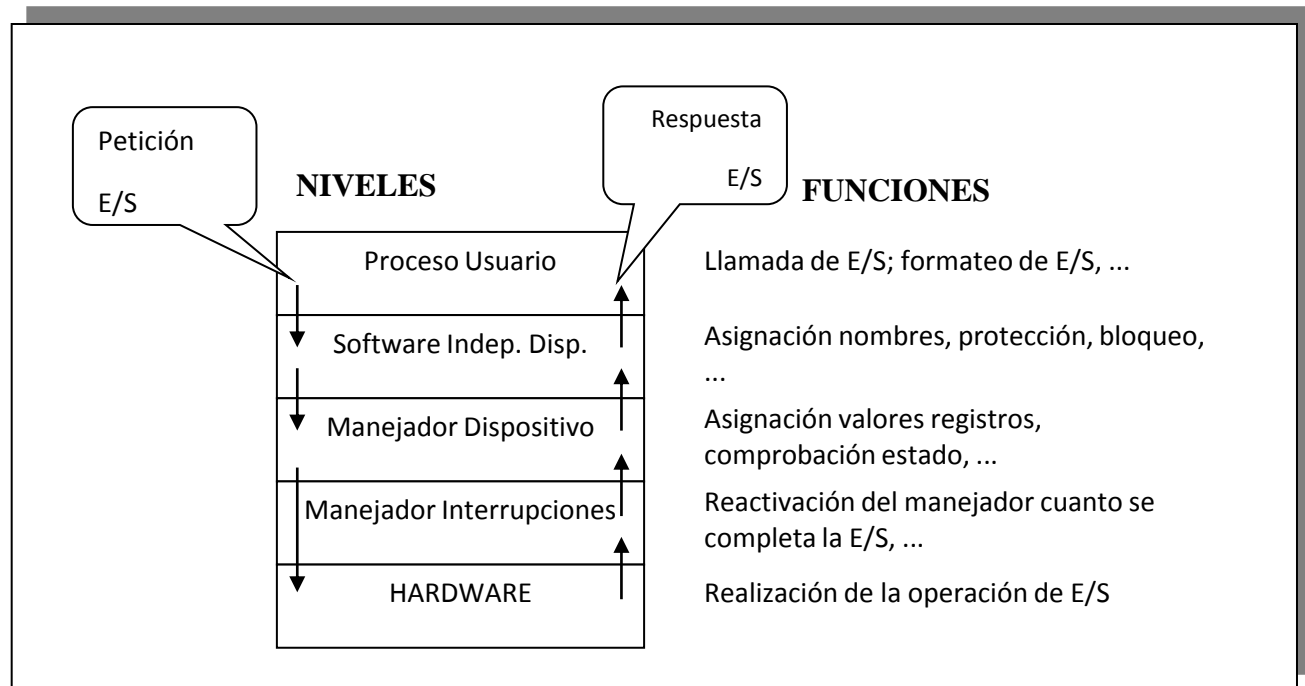
objetivos

- Proporcionar independencia de dispositivo
- Nombres uniformes (del archivo o dispositivo y no depender del dispositivo)
- Manejo de errores (se debe manejar lo más cerca del Hardware)
- Independencia del tipo de transferencia (síncrona mediante bloqueo o asíncrona mediante interrupciones)
- Dispositivos compartidos o exclusivos

Software de entrada/salida

niveles

- En una operación de Entrada/Salida intervienen 4 tipos o niveles de programas



Software de entrada/salida

niveles

Lectura de un bloque de archivo

- Se invoca al SO para llevar a cabo la llamada.
- El Software independiente del dispositivo busca, por ejemplo, en la caché.
- Se llama al manejador de dispositivos, si el bloque no estuviera en la caché.
- El manejador pasa la solicitud al Hardware.
- El proceso se bloquea hasta que termina la operación.
- Una vez terminada la operación, el Hardware genera una interrupción.
- Se ejecuta el manejador de interrupciones para descubrir lo ocurrido.
- Extrae el estado del dispositivo y lo despierta para enviarle los datos y que continúe.

Software de entrada/salida

niveles

Manejador de interrupciones

- Es el encargado de ocultar las interrupciones a los niveles superiores y que se gestionen en los niveles inferiores.
- Es el encargado de bloquear y despertar al proceso que solicita la entrada/salida.

Software de entrada/salida

niveles

Manejador de dispositivos

- Funciones:
 - ▣ Recibir peticiones abstractas de niveles superiores
 - ▣ Enviar las órdenes al controlador y comprobar que se realizan adecuadamente
 - ▣ Devolver datos, información de estado y error
- Es el software que contiene todo el código dependiente de dispositivo. Es el único programa que tiene acceso a los registros del controlador.
- Suele manejar solo una clase de dispositivos.
- Después de enviar las órdenes al controlador, el manejador puede:
 - ▣ Bloquearse hasta recibir interrupción si la operación tarda mucho
 - ▣ No bloquearse si la operación ejecuta rápidamente

Software de entrada/salida

niveles

Programas del SO independientes del dispositivo

Funciones:

- Proporcionar operaciones de E/S comunes a todos los dispositivos
- Proporcionar interfaz uniforme a los programas de usuario
- Asignar nombre a los dispositivos
- Proteger los dispositivos
- Proporcionar tamaño de bloque independiente de los dispositivos
- Almacenar temporalmente los datos en las transferencias E/S (buffering)
- Gestionar la asignación de espacio en los dispositivos de bloques
- Reservar y liberar los dispositivos
- Informar de los errores y actuar en consecuencia si el tratamiento es independiente del dispositivo

Software de entrada/salida

niveles

Nivel de usuario

- La mayoría del software de E/S están dentro del SO, pero una pequeña parte está fuera de él:
 - ▣ Bibliotecas que se enlazan con los programas de usuario. Llamadas al sistema, etc.
 - ▣ Sistemas de spooling. Consisten en un directorio para almacenar los ficheros a transferir y un proceso especializado (**daemon**) que es el único que puede utilizar el manejador del dispositivo para que se envíen los ficheros. Ej.: impresora.

Discos

intro

- Presentan tres ventajas respecto al uso de la Memoria Principal:
 - ▣ La capacidad del espacio de almacenamiento disponible es mucho más grande
 - ▣ El precio por bit es más barato
 - ▣ No son volátiles

Discos

Hw de disco

- En el hardware de disco pueden distinguirse dos partes:
 - ▣ La unidad de disco que es el dispositivo electromecánico. Ej: FDD: Floppy Disk Drive.
 - ▣ El controlador de la unidad de disco. Ej: FDC: Floppy Disk Controller.

Discos

Hw de disco

El disco

- Un **disco magnético** es un cilindro recubierto en las superficies planas de material magnético sobre el que se efectúa la grabación de datos.
- El **cabezal del disco** es un transductor que permite leer/escribir datos en el material magnético.

Tipos de discos:

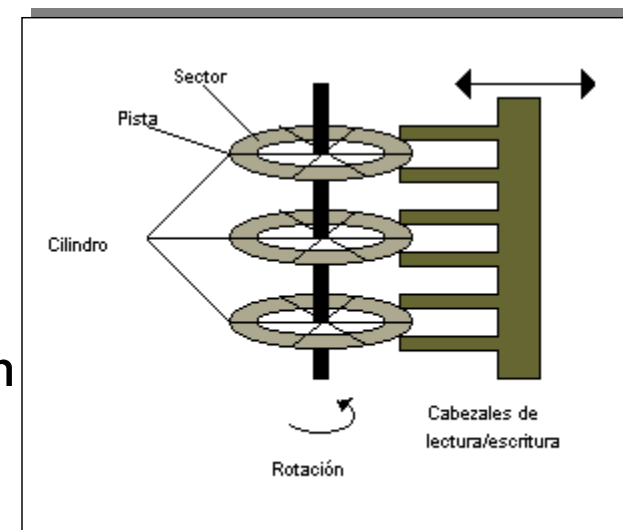
- Desplazamiento cabezas: fija y móvil.
- Transportabilidad de disco: fijo (rígido) y extraíble (flex.)
- Superficies: única y doble.
- Número disco: único y múltiples.
- Mecanismo de la cabeza: contacto, fija y aerodinámica

Discos

Hw de disco

Organización del disco

- ❑ Cilindro: Conjunto de pistas accesibles en una posición del brazo. Se supone que hay una cabeza por cada cara.
- ❑ Pista: Círculo determinado por una cabeza. Está formada por sectores.
- ❑ Sectores: Unidad básica de trasiego de información.
- ❑ La organización en cilindros, pistas y sectores se efectúa mediante la operación dar **formato**.



Discos

Hw de disco

Organización-Ejemplo:

Unidad de Disco	Número Cilindros	Pistas por Cilindro	Sectores por Pista	Sectores por Disco	Bytes por sector	Bytes por disco
5 1/4" Doble Densidad	40	2	9	720	512	368640
5 1/4" Alta Densidad	80	2	15	2400	512	1228800
3 1/2" Doble Densidad	80	2	9	1440	512	737280
3 1/2" Alta Densidad	80	2	18	2880	512	1474560
3 1/2" Extra Densidad	80	2	36	5760	512	2949120

Discos

Hw de disco

Direccionamiento

- A un determinado sector físico se puede acceder mediante su dirección lógica. Por ejemplo, para el mismo disco flexible

Dirección física sector			Dirección lógica sector
Cilindr	Pista	Sectores	
0			
0	0	1..9	0..8
0	1	1..9	9..17
1	0	1..9	18..26
....
39	1	1..9	711..719

- El controlador numera los sectores dentro de una pista desde 1 a 9, y no desde 0 a 8. Cilindro y cabezales sí son numerados a partir del 0.

El dispositivo

- El dispositivo electromecánico está constituido por motores, cabezas de lectura y escritura, carcasa, etc. y la lógica relacionada. Típicamente consta de:
 - ▣ Motores: para el giro del disco y para el posicionamiento del brazo.
 - ▣ Detectores de posición: de inicio de pista, para posición de referencia de las cabezas (pista 00) y de protección contra escritura.
 - ▣ Circuitos para efectuar las lecturas y las escrituras.

El dispositivo

- Ante una operación de lectura o escritura, la unidad, en un esquema simplificado, opera de la siguiente manera:
 - ▣ Enciende el motor.
 - ▣ Posiciona las cabezas y espera un tiempo para que estas apoyen sobre la superficie.
 - ▣ Localiza el sector y efectúa la operación.
 - ▣ Para el motor.

Discos

Hw de disco

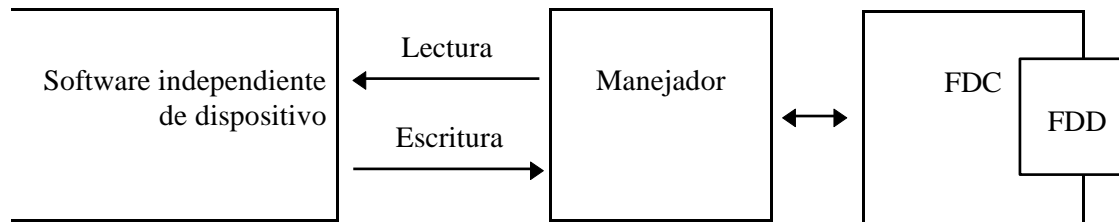
- Los tiempos que intervienen en las operaciones de entrada/salida a disco son:
 - ▣ **Tiempo de búsqueda** del cilindro o tiempo de posicionamiento de brazo. Es el tiempo necesario para mover el brazo del disco hasta la pista solicitada. Este tiempo incluye el tiempo de arranque y el tiempo de estabilización de velocidad.
 - ▣ **Tiempo de latencia** o retardo de rotación. Tiempo que transcurre desde que la cabeza está situada en el cilindro hasta que la cabeza se sitúa en el sector solicitado. El tiempo de acceso es igual al tiempo de búsqueda más el tiempo de latencia.
 - ▣ **Tiempo de transferencia.** Tiempo que se necesita para realizar la transferencia de los datos solicitados.

El controlador

- Es la parte electrónica que establece la interacción entre el procesador y el disco recibiendo peticiones del primero, gestionando el trasvase y notificando al procesador el fin de operación.
- Realiza las siguientes operaciones:
 - ▣ Control de la unidad de disco. Movimiento del brazo para localizar el cilindro, selección de la cabeza correspondiente a la pista, operación sobre el sector interesado, recalibrado, etc.
 - ▣ Traspase de información. Comprobación de la corrección de los datos, interpretación y ejecución de las órdenes relacionadas con el disco.

Esquema general

- Función principal del manejador es proporcionar un interfaz independiente del dispositivo.



- El esquema del manejador es un bucle infinito que atiende peticiones, encarga su realización e informa del resultado

Discos

Sw de disco

Manejador

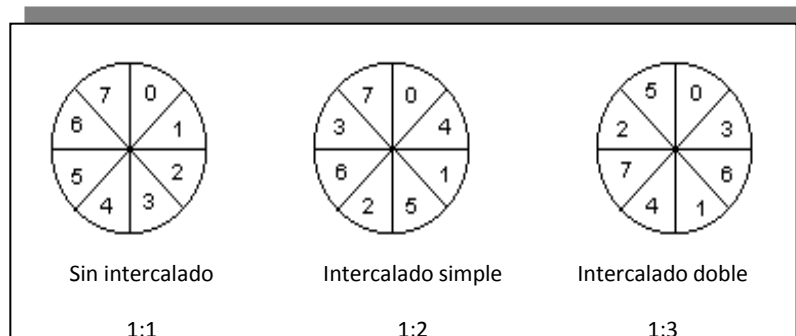
```
mensaje mens;                                     /*buffer del mensaje */
manejador_disco ()
{
    int r, peticionario;

    iniciar();
    while (TRUE)
    {
        recibir_petición (&mens);                /*esperar petición */
        peticionario = mens.peticionario; /*proceso que envió mensaje */
        switch(mens.m_tipo)
        {
            /*tratamiento tipo peticion */
            case READ:                             r = leer(); break;
            case WRITE:                            r = escribir(); break;
            default:                                r = ERROR;

        }
        mens.tipo_op = RESPUESTA;
        mens.STATUS = r;      /* código de resultado */
        enviar(peticionario, &mens);               /* mensaje respuesta remitente*/
    }
}
```

Conversión de direcciones lógicas a direcciones físicas

- Para la conversión es necesario conocer la numeración de los cilindros, pistas y sectores. Generalmente:
 - ▣ Numeración de cilindros comenzando desde el más externo (cilindro físico 0).
 - ▣ Numeración de pistas comenzando desde la superior (cabeza 0)
 - ▣ Numeración de sectores a partir de la marca de inicio de pista y considerando el factor de intercalación.
- El **factor de intercalación** indica la posición que ocupa un sector lógico respecto de su predecesor en una pista.
- Este mecanismo permite reducir el tiempo que necesita la unidad de disco para localizar un sector en una pista:



Conversión de direcciones lógicas a direcciones físicas

- La función que convierte la dirección lógica que recibe el manejador al formato de cilindro-pista-sector para acceso al disco será:

```
int convertir_dirección(long dir_disco)    /*dirección en palabra, no en bloque */
{
    ...
    bloque = (int) (dir_disco / tam_sector);
    cilindro = (int) (bloque / (n_pist * n_sect));
    pista = (int) ((bloque % ((n_pist * n_sect)) / n_sect);
    sector_lógico = bloque - cilindro * n_pist * n_sect - pista * n_sect;
    sector_físico = (sector_lógico * intercalación) % n_sect;
    ...
}
```

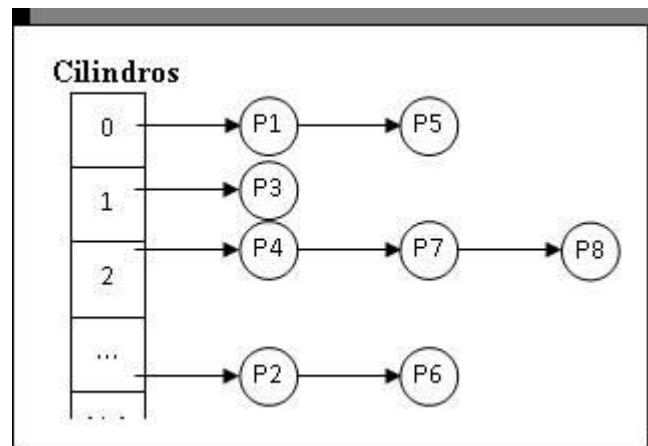
Conversión de direcciones lógicas a direcciones físicas

- Por ejemplo, en un disco de 2 pistas, 9 sectores por pista y 512 bytes por sector (se suponen los bloques del mismo tamaño que los sectores), las posiciones en el disco para varias direcciones se muestran en la tabla:

intercalac.	dir disco	bloque	cil.	pista	s. lóg.	sect. fís.
1	34309	67	3	1	4	4
2						8
1	57212	111	6	0	3	3
2						6
1	40450	79	4	0	7	7
2						5

Planificación del Movimiento del Brazo

- El Sistema Operativo mantiene una cola de peticiones de acceso para cada dispositivo de Entrada/Salida. Dependiendo de cómo se atiendan esas peticiones el tiempo de posicionamiento será mayor o menor. Por ejemplo, para un determinado dispositivo



Discos

Sw de disco

Planificación del Movimiento del Brazo

FCFS: Primero en llegar, primero en ser servido

- Las peticiones son atendidas en el orden de llegada.
- Ejemplo: Sea un disco con 200 cilindros y se supone que las peticiones en un instante determinado son en orden de llegada: 55, 58, 39, 18, 90, 160, 150, 38 y 184. Suponer que inicialmente la cabeza está situada en la pista 100 y se mueve en direcciones crecientes de cilindro. Indicar en qué orden se atenderán.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
1								
	2							
		3						
			4					
				5				
					6			
						7		
							8	
								9

- Es el algoritmo más fácil de implementar.
- Movimientos bruscos.
- Recorridos totales elevados.

Discos

Sw de disco

Planificación del Movimiento del Brazo

SSTF (SSF) : Primero más corto

- Se atiende primero las peticiones más próximas a la posición actual de la cabeza. Si dos peticiones son equidistantes, se puede usar un algoritmo aleatorio para decidir o bien se da prioridad al sentido de movimiento actual de la cabeza.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
				1				
	2							
3								
		4						
							5	
			6					
						7		
					8			
								9

- Se minimiza el tiempo medio de posicionamiento. Mejor que FCFS.
- Problema de inanición
- El brazo tiende a permanecer en la mitad del disco.

Planificación del Movimiento del Brazo

SCAN: Algoritmo del ascensor

- El brazo sólo se puede mover en un sentido, resolviendo todas las peticiones pendientes en su ruta hasta que alcance el último cilindro. A continuación, la cabeza se mueve en el otro sentido atendiendo las peticiones que se encuentre.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
						1		
					2			
								3
				4				
	5							
6								
		7						
							8	
			9					

- Presenta la ventaja de que atiende cualquier petición en un tiempo máximo.
- Normalmente, es peor que el algoritmo SSTF.

Discos

Sw de disco

Planificación del Movimiento del Brazo

CSCAN: Algoritmo del ascensor circular

- El brazo sólo se puede mover en un sentido, resolviendo todas las peticiones pendientes en su ruta. Cuando llega al último cilindro, se desplaza al primero y empieza a resolver las nuevas peticiones.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
						1		
					2			
								3
			4					
							5	
		6						
7								
	8							
				9				

- Si el movimiento por defecto es descendente la vuelta circular sería hacia el último cilindro.

Discos

Sw de disco

Planificación del Movimiento del Brazo

LOOK

- Es una variante del algoritmo SCAN. Cambia de sentido cuando se ha atendido la última petición y no tiene que llegar hasta el último cilindro.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
						1		
					2			
								3
				4				
	5							
6								
		7						
							8	
			9					

- Los recorridos totales son inferiores al algoritmo SCAN
- Más complejo que el algoritmo SCAN.

Planificación del Movimiento del Brazo

CLOOK

- Es una variante del algoritmo CSCAN. Volviendo al primer cilindro cuando se ha atendido la última petición y no tiene que llegar hasta el último cilindro.

Cilindros Solicitados								
055	058	039	018	090	160	150	038	184
						1		
					2			
								3
			4					
							5	
		6						
7								
	8							
				9				

- Los recorridos totales son inferiores al algoritmo CSCAN pero más elevados que LOOK.
- Más complejo que el algoritmo CSCAN.
- Si el movimiento por defecto es descendente la vuelta circular sería hacia el último cilindro.

Transferencia

- La transferencia es una operación sencilla a nivel del software porque el trabajo se realiza entre el controlador y el DMA. Todo lo que se requiere es indicar:
 - ▣ el tipo de operación (lectura o escritura),
 - ▣ la unidad de disco,
 - ▣ la dirección física (cilindro, pista, sector),
 - ▣ el número de sectores por pista,
 - ▣ el espacio entre sectores y
 - ▣ la longitud de datos.
- Seguidamente, el manejador debe bloquearse en espera de la interrupción del controlador indicando el final de operación. Ocurrida la interrupción, se obtiene el estado del controlador y se comprueba el resultado de la transferencia.

Tratamiento de Errores

- Se pueden distinguir dos tipos de errores:
 - ▣ Sistemático, porque algún elemento haya perdido operatividad o alguna de las actividades manuales se haya realizado incorrectamente (ausencia de disco, protección contra escritura en su caso, etc.). Se emite un mensaje de error y no se reintentan.
 - ▣ Transitorio, desde una mota de polvo sobre la superficie hasta un error en el controlador. Se resuelven reintentando la operación hasta un número limitado de veces. Caso de no subsanarse, se entenderá que se trata de un error sistemático.
- Cuando la causa del error es de deterioro de un sector, se toman soluciones como:
- Crear un archivo de sectores defectuosos cuyo contenido son los sectores que no pueden ser utilizados. Este fichero es transparente al usuario.
- Reemplazar las pistas dañadas por otras al final del disco reservadas a tal efecto.

Tratamiento de Errores

Los más comunes:

- De programación (algún parámetro incorrecto). Ej.: solicitar sector inexistente.
 - ▣ algunos son detectados por el controlador
 - ▣ en general, el manejador deberá devolver error
- Transitorios de transferencia. Ej.: provocado por suciedad.
 - ▣ se detectan mediante sumas de control de los datos trasvasados
 - ▣ se reintenta la operación varias veces
- Permanentes de transferencia. Ej.: sectores dañados.
 - ▣ los bloques inservibles deben estar caracterizados
 - ▣ se pueden utilizar ficheros de bloques inservibles (ojo a operaciones que no operan sobre ficheros)

Tratamiento de Errores

Los más comunes:

- De búsqueda (posicionamiento incorrecto del brazo)
 - ▣ algunos controladores los corrigen automáticamente
 - ▣ otros controladores ponen a 1 un bit de error y declinan la responsabilidad
 - ▣ en el manejador (suele enviar comando RECALIBRATE)
- Del controlador. Ej.: el controlador no acepta los comandos.
 - ▣ caso de malfunción, es típico poner a 1 un bit de error para forzar la
 - ▣ reinicialización del propio controlador

Discos

Sw de disco

Otras Operaciones

Iniciación del dispositivo

- El manejador informa al controlador del número de unidades de disco, calibrado del brazo, unidad de disco por defecto, parámetros de estabilización de disco, etc.
- Comprobar que las operaciones se han realizado satisfactoriamente.

Carga de DMA

- Sentido de la transferencia (Controlador → Memoria y Controlador ← Memoria).
- Dirección de inicio en memoria principal.
- Cantidad de palabras a transferir.

Discos

Sw de disco

Otras Operaciones

Parada del motor

- No se para el motor después de cada transferencia sino que se realiza una **parada retardada** con un temporizador de guardia.

Encendido del motor

- Para encender el motor hay que activar un bit y esperar un tiempo de estabilización.

```
int arrancar_motor()
{
    ...
    if (motor_girando)
    {
        inhabilitar_parada();
        return;
    }
    escribir_en_puerto ( );
    esperar (tiempo_estabilización);           /* temporizador de guardia */
}
```

Mejoras

Mejoras de Eficiencia

Se suelen utilizar otras técnicas para mejorar las prestaciones globales.

- Solapamiento de búsquedas. Solapadas entre sí y solapadas con operación lectura o escritura.
- Otras planificaciones del brazo. Ascensor combinado con datos de uso más frecuente en cilindros centrales.
- Cache interna del controlador. Cuando se da orden de leer un sector, se lee toda la pista y se almacena en la cache.
- Defragmentar.
- Intercalación entre pistas.
- Intercalación de discos. Acceso a varios discos en paralelo.

Mejoras

Mejoras de Fiabilidad

- Empleo de grupos redundantes de discos de bajo coste **Sistemas RAID** (redundant arrays of inexpensive disks).
- Esta redundancia puede organizarse con diverso rendimiento y rentabilidad. Existen diferentes niveles o configuraciones RAID dependiendo del grado de redundancia.