

Tarea 6

11.1 Cite ejemplos de recursos reutilizables y consumibles

R: Recursos consumibles:

- Interrupciones
- Señales
- Mensajes
- Información en los buffers de E/S

Recursos Reutilizables:

- Procesadores
- Canales de E/S
- Memoria primaria y secundaria
- Ficheros
- Base de datos
- Semáforos

11.2 Enumere y defina brevemente las 3 técnicas para realizar E/S.

R: Técnicas de realización de E/S:

- E/S Programada: el procesador emite una orden de E/S de parte de un proceso a un modulo de E/S; el proceso espera entonces a que termine la operación, antes de seguir.
- E/S dirigida por Interrupciones: el procesador emite una orden de E/S de parte de un proceso, continua la ejecución de las instrucciones siguientes y el modulo de E/S lo interrumpe cuando completa su trabajo. Las instrucciones siguientes pueden ser del mismo proceso, si no es necesario para este esperar la terminación de la E/S. En otro caso, el proceso se suspende a la espera de la interrupción, mientras se realiza otro trabajo.
- Acceso Directo a la Memoria (DMA): un modulo de DMA controla el intercambio de datos entre la memoria principal y un modulo de E/S. El procesador envía una petición de transferencia de un bloque de datos al modulo de DMA y se interrumpe solo cuando se ha transferido el bloque entero.

11.3 ¿Cuál es la diferencia entre E/S lógica y E/S a dispositivo?

R: Diferencias:

- La E/S lógica considera al dispositivo como un recurso lógico, y no al dispositivo en si. Se ocupa de gestionar las funciones generales con órdenes como abrir, cerrar, leer y escribir.
- La E/S salida a dispositivo si considera al dispositivo y se encarga de utilizar secuencias adecuadas de instrucciones de E/S y ordenes para el canal y el controlador del dispositivo, ya que esta si se ocupa del control del dispositivo.

11.4 ¿Cuál es la diferencia entre dispositivos orientados a bloques y dispositivos orientados a flujos de caracteres? Proponga un ejemplo de cada tipo.

R: Diferencias:

- Los dispositivos orientados a bloques almacenan la información en bloques, por lo general de tamaño fijo, haciendo las transferencias de un bloque a la vez.
- Los dispositivos orientados a flujo transfieren los datos como una serie de bytes. Por lo general, los dispositivos de almacenamiento secundario son orientados a bloque, y el resto son orientados a flujo. Las cintas son un ejemplo de dispositivo orientado a bloques y las impresoras son un ejemplo de dispositivo orientado a flujo.

11.5 ¿Por qué podría esperar una mejora del rendimiento utilizando para la E/S unbuffer doble en lugar de un único buffer?

R: La memoria intermedia (buffer) doble mejora el rendimiento ya que permite tener un almacenamiento en uso, mientras el sistema operativo llena o vacía el otro. No hay necesidad de esperar a la siguiente E/S, al menos que se de el caso de que el dispositivo se adelante a la memoria intermedia doble.

11.6 ¿Qué elementos de retardo están involucrados una lectura o escritura de disco?

R: Retardos en una lectura o escritura de disco:

- Tiempo de búsqueda
- Retardo de giro
- Tiempo de transferencia
- Espera a que el dispositivo este libre
- Espera por el canal de E/S
- RPS (si el sistema cuenta con esta técnica)

11.7 Defina brevemente las políticas de planificación de disco mostradas en la figura 11.7

R: Políticas de planificación:

- Primero en Entrar, primero en salir (FIFO): el primero en entrar es el primero en salir, por lo tanto los elementos de la cola se procesan en un orden secuencial.
- Prioridad (PRI): se le da prioridad a los trabajos por lotes que sean cortos y a los trabajos interactivos, más que a los trabajos más grandes que consisten en grandes operaciones.
- Ultimo en entrar, primero en salir (LIFO): le da prioridad al último elemento de la cola, esto es de gran utilidad ya que reduce el tiempo de búsqueda por la cercanía del brazo y reduce las colas. Si la cola de trabajo es muy larga, se produce inanición.
- Primero el tiempo de servicio más corto (SSTF): se atiende la solicitud de E/S a disco que requiere el menor movimiento posible del brazo del disco. Esta política reduce el tiempo de búsqueda, pero no garantiza que el tiempo medio de búsqueda sea también el mínimo en una serie de movimientos del brazo.
- SCAN: el brazo solo se puede mover en un sentido hasta que resuelve todas las peticiones en esa ruta, evitando así dejar alguna petición sin resolver (LOOK). Cuando termino con todas las peticiones, entonces cambia de dirección y hace nuevamente el recorrido.
- C-SCAN (Circular SCAN): restringe el rastreo a una sola dirección. Cuando se haya visitado la ultima pista en un sentido, el brazo vuelve al extremo opuesto del disco y comienza a recorrerlo de nuevo, esto reduce el retardo máximo sufrido por las nuevas solicitudes.
- SCAN de N pasos: divide la cola de solicitudes del disco en sub-colas de longitud N. Las sub-colas procesan una a una mediante SCAN. Cuando se este procesando una cola, las nuevas solicitudes se añaden a las otras.
- FSCAN: emplea dos sub-colas, cuando comienza el rastreo todas las solicitudes están en una de las colas y la otra permanece vacía. Durante el recorrido, todas las solicitudes nuevas se colocan en la cola que estaba vacía.

11.8 Defina brevemente los siete niveles del RAID

R: Niveles del RAID:

- RAID 0: Por no incluir redundancia para mejorar el rendimiento, no es realmente miembro de la familia RAID. El usuario y los sistemas están distribuidos a lo largo de todo el vector de discos. Si se dan las condiciones, se pueden realizar en paralelo solicitudes, reduciendo el tiempo en la cola de

E/S. Los datos están distribuidos entre los discos disponibles. Todos los usuarios y los datos se ven como si estuviesen almacenados en un disco lógico. Este disco esta dividido en bandas, estas bandas pueden ser bloques físicos, sectores o alguna otra unidad. Las bandas se asignan mediante turnos rotatorios a miembros consecutivos del vector. Un conjunto de bandas consecutivas lógicamente que se corresponden exactamente con una banda de cada miembro del vector, se denomina franja.

- **RAID 1:** La redundancia se consigue duplicando los datos. Cada banda lógica corresponde con dos discos físicos independientes, por lo que cada disco del vector tiene un disco espejo con los mismos datos. De aquí su principal desventaja, el coste, ya que requiere el doble de espacio de disco del disco lógico que soporta. RAID 1 puede mejorar el rendimiento de RAID 0 en lo que a lectura se refiere, pero si la mayoría de las solicitudes son de escritura puede que esto ya no sea una ventaja. RAID 1 ofrece una copia de respaldo en tiempo real de todos los datos, por lo tanto debería limitarse su uso para información crítica.
- **RAID 2:** Usa una técnica de acceso paralelo, en donde todos los discos participan en la ejecución de cada solicitud de E/S. Normalmente, el eje de las unidades individuales esta sincronizado, por lo que cada cabeza de disco esta en la misma posición de cada disco en un instante dado. Se calcula un código de corrección de errores a lo largo de los bits correspondientes sobre cada disco de datos, y los bits del código se almacenan en las respectivas posiciones de bit sobre discos de paridad múltiple. Normalmente se utiliza el código Hamming, que es capaz de corregir errores de 1 solo bit y detectar errores de 2 bit. Se mantiene la característica del alto coste, es efectivo en entornos donde se producen muchos errores de disco.
- **RAID 3:** Se parece a RAID 2, pero solo requiere un disco redundante, no importa el tamaño del vector de discos. Utiliza acceso paralelo, con los datos distribuidos en pequeñas bandas. Se calcula un solo bit de paridad para el conjunto de bits en la misma posición de todos los discos de datos. En el caso de un fallo de disco, se accede a la unidad de paridad y se reconstruyen los datos desde los dispositivos restantes. Una vez que se reemplaza la unidad fallida, se pueden restaurar sobre la nueva unidad los datos perdidos y continúa el funcionamiento. Puede conseguir una tasa de transferencia de datos alta. Para grandes transferencias el rendimiento es bueno, no así en un entorno orientado a transacciones, ya que solo se puede realizar una solicitud por instante.
- **RAID 4:** Usa una técnica de acceso independiente, por lo que se pueden satisfacer en paralelo solicitudes de E/S individuales. También utiliza división de datos y las bandas son relativamente grandes. Se calcula una banda de la paridad bit a bit a lo largo de las bandas correspondientes de cada disco de datos, y los bits de paridad se almacenan en la banda correspondiente del disco de paridad. Cada operación de escritura debe implicar al disco de paridad, lo que puede llegar a ser un cuello de botella.
- **RAID 5:** Esta organizado de forma similar a RAID 4. La diferencia es que este distribuye las bandas de paridad a través de todos los discos. Una asignación normal es un esquema de turno rotatorio. Para un vector de n discos, la banda de prioridad esta sobre un disco distinto para las n primeras bandas, y después, se repite el patrón. La distribución de las bandas de paridad a través de todos los discos evita el potencial cuello de botella en la E/S del sistema RAID 4, en donde solo había un disco de paridad.
- **RAID 6:** Realiza dos cálculos distintos de paridad y se almacenan en bloques independientes de diferentes discos. De este modo un vector en RAID 6 cuyos datos requiere N discos esta formado por N + 2 discos. La principal ventaja es que ofrece una gran disponibilidad de datos, deben fallar 3 discos en el intervalo MTTR (Mean Time To Repair, Tiempo Medio de Reparación) para que los datos no estén disponibles. Por otro lado, tenemos que también sufre de una penalización de escritura, porque cada escritura afecta a 2 bloques de paridad.

11.8 ¿Cual es el tamaño normal de un sector de disco?

R: Por lo general se usa un sector fijo de 512 bytes.