



Instituto Politécnico Nacional
Escuela Superior de Cómputo



TAREA 4

Sistemas Operativos

Integrantes:

Mora Ayala José Antonio
Ramírez Cotonieto Luis Fernando
Torres Carrillo Josehf Miguel Ángel
Tovar Jacuinde Rodrigo

Profesor:

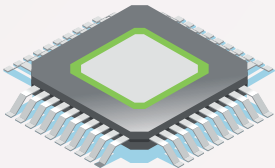
Cortés Galicia Jorge
2CM17

PLANIFICACIÓN DE CPU

La planificación (scheduling) es la base para lograr la multiprogramación. Un sistema multiprogramado tendrá varios procesos que requerirán el recurso procesador a la vez. Esto sucede cuando los procesos están en estado ready (pronto). Si existe un procesador disponible y existen procesos en estado ready, se debe elegir el que será asignado al recurso para ejecutar. El componente del sistema operativo que realiza la elección del proceso es llamada planificador (scheduler).

ÚNICO PROCESADOR

Sólo puede ejecutarse un proceso cada vez; cualquier otro proceso tendrá que esperar hasta que la CPU quede libre y pueda volver a planificarse.



Utilidad

Los esquemas de planificación son útiles según el ambiente donde sean aplicados:

– *Sistemas por lotes*

Como no existe interacción con usuarios, los planificadores no expropiativos son ideales.

– *Sistemas interactivos*

Debido a que existen procesos de usuarios ejecutando a la vez, los planificadores expropiativos son ideales para mantener un buen tiempo de respuesta para los usuarios.

– *Sistemas de tiempo real*

No es necesario un planificador expropiativo ya que los procesos puede que no ejecuten por un buen tiempo, pero cuando lo hacen es por un período muy corto

UTILIZACIÓN DE LA CPU

Deseamos mantener la CPU tan ocupada como sea posible. Conceptualmente, la utilización de la CPU se define en el rango comprendido entre el 0 y el 100 por cien.

TASA DE PROCESAMIENTO

Si la CPU está ocupada ejecutando procesos, entonces se estará lle vando a cabo algún tipo de trabajo. Una medida de esa cantidad de trabajo es el número de procesos que se completan por unidad de tiempo, y dicha medida se denomina tasa de procesamiento.

TIEMPO DE RESPUESTA

Tiempo que el proceso tarda en empezar a responder, no el tiempo que tarda en enviar a la salida toda la información de respuesta. Generalmente, el tiempo de respuesta está limitado por la velocidad del dispositivo de salida.

Corto plazo

Determina qué proceso es ejecutado en el procesador
Se ejecuta frecuentemente y debe ser eficiente
Evento típicos que desencadena al despachador: interrupciones del reloj (quantum), interrupciones de I/O, llamados al sistemas, señales, etc.

Mediano plazo

Determina si agregar más programas a los que ya están parcialmente o totalmente en memoria principal

Largo plazo

Determina qué programas son admitidos al sistema para ejecución
Controla el grado de multiprogramación.
Mientras más procesos son admitidos, cada uno tendrá un porcentaje menor de uso del procesador

MULTIPROGRAMACIÓN

El objetivo es tener continuamente varios procesos en ejecución, con el fin de maximizar el uso de la CPU.

Se intenta usar ese tiempo de forma productiva. En este caso, se mantienen varios procesos en memoria a la vez. Cuando un proceso tiene que esperar, el sistema operativo retira el uso de la CPU a ese proceso y se lo cede a otro proceso.

TIEMPO DE EJECUCIÓN

Ese tiempo de ejecución es la suma de los períodos que el proceso invierte en esperar para cargarse en memoria, esperar en la cola de procesos preparados, ejecutarse en la CPU y realizar las operaciones de E/S.

TIEMPO DE ESPERA

El tiempo de espera es la suma de los períodos invertidos en esperar en la cola de procesos preparados.

ALGORITMOS DE PLANIFICACIÓN

PLANIFICACIÓN FCFS

La implementación es fácil a través de una cola FIFO.

Los procesos son ejecutados en el orden que llegan a la cola de procesos listos.

Es adecuado para sistemas por lotes (batch).

PLANIFICACIÓN SJF

Cuando el procesador queda disponible se le asigna al proceso que tenga el menor CPU-burst.

El algoritmo asocia a los procesos el largo de su próximo CPU-burst.

Si dos procesos tiene el mismo CPU-burst se desempata de alguna forma.

PLANIFICACIÓN MEDIANTE COLAS MULTINIVEL REALIMENTADAS

• El número de colas.

• El algoritmo de planificación de cada cola.

PLANIFICACIÓN POR PRIORIDADES

Es adecuado para sistemas interactivos.

A cada proceso se le asigna un número entero que representa su prioridad.

El planificador asigna el procesador al proceso con la más alta prioridad.

PLANIFICACIÓN DE SISTEMAS N MULTIPROCESADOR

En un sistema simétrico cualquier procesador podrá ejecutar procesos de usuario.

PLANIFICACIÓN POR TURNOS

A cada proceso se le brinda un intervalo de tiempo para el uso del procesador (time quantum).

Al finalizar el tiempo, el procesador le es expropiado y vuelve al estado pronto (ready) al final de la cola.

Es ideal para sistemas de tiempo compartido.

PLANIFICACIÓN MEDIANTE COLAS MULTINIVEL

1. Procesos del Sistema.
2. Procesos interactivos.
3. Procesos de edición interactivos.
4. Procesos por lotes.
5. Procesos de estudiantes.

PLANIFICACIÓN DE CPU

Los mecanismos de planificación de la CPU son la base de los sistemas operativos multiprogramados. Mediante la conmutación de la CPU entre distintos procesos, el sistema operativo puede hacer que la computadora sea más productiva.

PROCESO

Un proceso se ejecuta hasta que tenga que esperar, normalmente porque es necesario completar alguna solicitud de E/S. En un sistema informático simple, la CPU permanece entonces inactiva y todo el tiempo de espera se desperdicia; no se realiza ningún trabajo útil.

CPU INACTIVA

Cuando la CPU queda inactiva, el sistema operativo debe seleccionar uno de los procesos que se encuentran en la cola de procesos preparados para ejecución.

ÚNICO PROCESADOR

Sólo puede ejecutarse un proceso cada vez; cualquier otro proceso tendrá que esperar hasta que la CPU quede libre y pueda volver a planificarse.

MULTIPROGRAMACIÓN

El objetivo es tener continuamente varios procesos en ejecución, con el fin de maximizar el uso de la CPU.

PLANIFICADOR A CORTO PLAZO

El planificador a corto plazo (o planificador de la CPC) lleva a cabo esa selección del proceso. El planificador elige uno de los procesos que están en memoria preparados para ejecutarse y asigna la CPU a dicho proceso.

CICLO DE RAFAGAS DE CPU Y E/S

La ejecución del proceso comienza con una ráfaga de CPU. Esta va seguida de una ráfaga de E/S, a la cual sigue otra ráfaga de CPU, luego otra ráfaga de E/S, etc. Finalmente, la ráfaga final de CPU concluye con una solicitud al sistema para terminar la ejecución.

TIEMPO

Generalmente, la curva es de tipo exponencial o hiperexponencial, con un gran número de ráfagas de CPU cortas y un número menor de ráfagas de CPU largas.

PLANIFICACIÓN APROPIATIVA

SITUACIONES

1. Cuando un proceso cambia del estado de ejecución al estado de espera (por ejemplo, como resultado de una solicitud de E/S o de una invocación de wait para esperar a que termine uno de los procesos hijo).
2. Cuando un proceso cambia del estado de ejecución al estado preparado (por ejemplo, cuando se produce una interrupción).
3. Cuando un proceso cambia del estado de espera al estado preparado (por ejemplo, al completarse una operación de E/S).
4. Cuando un proceso termina.

PRESENTA

No hay más que una opción en términos de planificación: debe seleccionarse un nuevo proceso para su ejecución (si hay algún proceso en la cola de procesos preparados).

Necesidad de crear un nuevo mecanismo para coordinar el acceso a los datos compartidos.

COSTE

Limitaciones con el acceso de datos compartidos.

PERJUDICIAL AL KERNEL

La técnica de desalojo también afecta al diseño del kernel del sistema operativo. Durante el procesamiento de una llamada al sistema, el kernel puede estar ocupado realizando una actividad en nombre de un proceso. Tales actividades pueden implicar cambios importantes en los datos del kernel (por ejemplo, en las colas de E/S).

DESPACHADOR

El despachador es el módulo que proporciona el control de la CPU a los procesos seleccionados por el planificador a corto plazo.

- Cambio de contexto.
- Cambio al modo usuario.
- Salto a la posición correcta dentro del programa de usuario para reiniciar dicho programa.

DEBE SER

Lo más rápido posible, ya que se invoca en cada conmutación de proceso.

LATENCIA DE DESPACHO

El tiempo que tarda el despachador en detener un proceso e iniciar la ejecución de otro.

CRITERIOS DE PLANIFICACIÓN

Los diferentes algoritmos de planificación de la CPU tienen distintas propiedades, y la elección de un algoritmo en particular puede favorecer una clase de procesos sobre otros. A la hora de decidir qué algoritmo utilizar en una situación particular, debemos considerar las propiedades de los diversos algoritmos.

OBJETIVO

Maximizar la utilización de la CPU y la tasa de procesamiento, y minimizar el tiempo de ejecución, el tiempo de espera y el tiempo de respuesta. En la mayoría de los casos, lo que se hace es optimizar algún tipo de valor promedio.

UTILIZACIÓN DE LA CPU

Deseamos mantener la CPU tan ocupada como sea posible. Conceptualmente, la utilización de la CPU se define en el rango comprendido entre el 0 y el 100 por cien.

TASA DE PROCESAMIENTO

Si la CPU está ocupada ejecutando procesos, entonces se estará llevando a cabo algún tipo de trabajo. Una medida de esa cantidad de trabajo es el número de procesos que se completan por unidad de tiempo, y dicha medida se denomina tasa de procesamiento. Para procesos de larga duración, este valor puede ser de un proceso por hora; para transacciones cortas, puede ser de 10 procesos por segundo.

TIEMPO DE EJECUCIÓN

El intervalo que va desde el instante en que se ordena la ejecución de un proceso hasta el instante en que se completa es el tiempo de ejecución.

Ese tiempo de ejecución es la suma de los periodos que el proceso invierte en esperar para cargarse en memoria, esperar en la cola de procesos preparados, ejecutarse en la CPU y realizar las operaciones de E/S.

TIEMPO DE ESPERA

El tiempo de espera es la suma de los periodos invertidos en esperar en la cola de procesos preparados.

TIEMPO DE RESPUESTA

Tiempo que el proceso tarda en empezar a responder, no el tiempo que tarda en enviar a la salida toda la información de respuesta. Generalmente, el tiempo de respuesta está limitado por la velocidad del dispositivo de salida.

ALGORITMOS DE PLANIFICACIÓN

PLANIFICACIÓN FCFS

First Come First Served (FCFS)

Los procesos son ejecutados en el orden que llegan a la cola de procesos listos.

La implementación es fácil a través de una cola FIFO.

Es adecuado para sistemas por lotes (batch).

Es un algoritmo no expropiativo: una vez que el procesador le es asignado a un proceso este lo mantiene hasta que termina o se bloquea (por ejemplo al generar un pedido de E/S).

El tiempo de espera promedio por lo general es alto.

PLANIFICACIÓN SJF

Shortest Job First (SJF)

El algoritmo asocia a los procesos el largo de su próximo CPU-burst.

Cuando el procesador queda disponible se le asigna al proceso que tenga el menor CPU-burst.

Si dos procesos tienen el mismo CPU-burst se desempata de alguna forma.

Es adecuado para sistemas por lotes (batch).

Su funcionamiento depende de conocer los tiempos de ejecución lo cual en la mayoría de los casos no sucede.

PLANIFICACIÓN POR PRIORIDADES

Es adecuado para sistemas interactivos.

A cada proceso se le asigna un número entero que representa su prioridad.

El planificador asigna el procesador al proceso con la más alta prioridad.

Se utiliza en general un esquema expropiativo ya que si un proceso con mayor prioridad que el que está ejecutando arriba a la lista de procesos listos (ready queue), será asignado al procesador.

SJF se puede ver como un algoritmo de prioridad donde la prioridad está dada por el próximo CPU-burst.

DESVENTAJAS

Sufre de posposición indefinida ya que un proceso de baja prioridad quizás no pueda ejecutar nunca.

La solución es utilizar prioridades dinámicas de envejecimiento: incrementar la prioridad según pasa el tiempo sin ejecutar.

La prioridad de un proceso para el uso del recurso procesador deberá ser inversamente proporcional al uso que el proceso haga del mismo.

Por lo tanto un proceso tipo I/O-bound deberá tener, en general, mayor prioridad que uno tipo CPU-bound.

PLANIFICACIÓN POR TURNOS

RR ROUND ROBIN

Especialmente para los sistemas de tiempo compartido. Es similar a la planificación FCFS, pero se añade la técnica de desalojo para conmutar entre procesos.

A cada proceso se le brinda un intervalo de tiempo para el uso del procesador (time quantum).

Al finalizar el tiempo, el procesador le es expropiado y vuelve al estado pronto (ready) al final de la cola.

Es fácil de implementar ya que solamente es necesario una cola de procesos listos. Cuando un proceso consume su quantum es puesto al final de la cola.

El quantum debe ser bastante mayor a lo que lleva a realizar un cambio de contexto, sino se tendrá mucho overhead. A su vez, el tiempo de quantum incide en los tiempos de retorno.

Es ideal para sistemas de tiempo compartido.

No hay posposición indefinida.

PLANIFICACIÓN MEDIANTE COLAS MULTINIVEL

SECUENCIA DE COLAS MULTINIVEL

1. Procesos del Sistema.
2. Procesos interactivos.
3. Procesos de edición interactivos.
4. Procesos por lotes.
5. Procesos de estudiantes.

Si los procesos se pueden clasificar según sus cualidades, es posible dividir la lista de procesos listos (ready queue) en varias colas (una para cada clasificación).

Los procesos son asignados permanentemente a una de las colas.

Cada cola tendrá su propio algoritmo de planificación propio.

Además, se debe tener una estrategia de planificación entre las diferentes colas. Por ejemplo, una cola tendrá prioridad sobre otra.

SE DEFINE MEDIANTE:

- El número de colas.
- El algoritmo de planificación de cada cola.
- El método usado para determinar cuándo pasar un proceso a una cola de prioridad más alta.
- El método usado para determinar cuándo pasar un proceso a una cola de prioridad más baja.
- El método usado para determinar en qué cola se introducirá un proceso cuando haya que darle servicio.

Se basa en categorizar los procesos según el uso de CPU (CPU-burst) que tengan.

La cola de mayor prioridad será la de los procesos I/O-bound y la de menor la de procesos con alto CPU-bound.

Los procesos, según el consumo de CPU que hagan, serán promovidos a una cola de mayor prioridad o rebajados a una de menor prioridad.

De esta forma, se garantiza que los procesos con poco uso de procesador tengan mayor prioridad, y los que consumen mucho procesador tendrán baja prioridad.

En un sistema simétrico cualquier procesador podrá ejecutar procesos de usuario.

Una posibilidad es asignar una cola de procesos listos para cada procesador y de esa forma mantener los procesos asignados a un procesador (afinidad de procesador).

Esto es conveniente para aprovechar los datos que están frescos en la memoria cache del procesador, ya que al ejecutar un proceso en un procesador se nutre su cache con datos del proceso.

De esta forma, se logra mantener un mayor índice de cache hit y, por lo tanto, un mayor rendimiento en el sistema.

Un problema que puede surgir es un desbalance en la cantidad de trabajo por procesador. En estos casos se migrarán procesos de cola para lograr balancear nuevamente la carga.

MÉTODOS DE PLANIFICACIÓN EN LOS SISTEMAS MULTIPROCESADOR

PLANIFICACIÓN DE SISTEMAS N MULTIPROCESADOR

El método usado para determinar en qué cola se introducirá un proceso cuando haya que darle servicio.