Universidad Católica de Santiago del Estero Departamento Académico Rafaela – DAR

Cátedra: Sistemas Operativos

Objetivos

Las políticas de planificación intentan cubrir los siguientes objetivos:

- Justicia. La política debe ser lo más justa posible con todo tipo de procesos, sin favorecer a unos y perjudicar a otros.
- *Máxima capacidad de ejecución*. Debe dar un servicio aceptable para que todos los trabajos se realicen lo más rápidamente posible. Esto se logra disminuyendo el número de cambios de proceso.
- *Máximo número de usuarios interactivos*. En los sistemas de tiempo compartido se tratará de que puedan estar trabajando el mayor número de usuarios simultáneamente.
- **Predecibilidad**. La política de planificación debe concebirse de tal forma que en todo momento pueda saberse cómo será su ejecución.
- Minimización de la sobrecarga. La computadora debe tener poca sobrecarga ya que ésta incide directamente sobre el rendimiento final del sistema: a menor sobrecarga, mayor velocidad de proceso. Por ello, los cambios de contexto deben disminuirse.
- **Equilibrio en el uso de recursos**. Para obtener un buen rendimiento en el uso de los recursos y que éstos estén ocupados equitativamente el mayor tiempo posible.
- **Seguridad de las prioridades**. Si un proceso tiene mayor prioridad que otro, éste debe ejecutarse más rápidamente.

Criterios

Criterios a considerar a la hora de elegir o diseñar un buen algoritmo de planificación:

1- **Equidad.**

Garantizar que cada proceso obtiene su proporción justa de la CPU.

2- Eficacia, eficiencia, utilización de la CPU.

Mantener ocupada la CPU el 100% del tiempo.

El procesador debe estar el mayor tiempo posible ocupado para lograr así un gran rendimiento.

En un sistema real, deberá variar entre el 40% (en un sistema con poca carga) y el 90% (en un sistema muy cargado).

3- Tiempo de respuesta.

Velocidad con que el ordenador da respuesta a una petición. Depende mucho de la velocidad de los dispositivos de entrada/salida.

Universidad Católica de Santiago del Estero Departamento Académico Rafaela – DAR

Cátedra: Sistemas Operativos

Tiempo que transcurre entre la presentación de una solicitud y la producción de la primera respuesta. Tiempo que el proceso tarda en comenzar a responder, pero no incluye el tiempo que toma exhibir la respuesta.

Minimizar el tiempo de respuesta para los usuarios interactivos.

4- Tiempo de regreso o de servicio.

Es el tiempo que tarda en ejecutarse un proceso, donde se incluye el tiempo de carga del programa en memoria, el tiempo de espera en la cola de procesos preparados, el tiempo de ejecución en el procesador y el tiempo consumido en operaciones de entrada/salida.

El intervalo entre el momento de presentación de un proceso y el momento en que se termina. Es la suma de los períodos durante los cuales espera entrar en la memoria, espera en la cola de procesos listos, se ejecuta en la CPU y realiza E/S.

En un sistema interactivo, el tiempo de retorno tal vez no sea el mejor criterio.

Minimizar el tiempo que deben esperar los usuarios por lotes para obtener sus resultados.

5- Tiempo de ejecución.

Es idéntico al tiempo de servicio menos el tiempo de espera en la cola de procesos preparados; es decir, es el tiempo teórico que necesitaría el proceso para ser ejecutado si fuera el único presente en el sistema.

6- Tiempo de procesador.

Es el tiempo que un proceso está utilizando el procesador sin contar el tiempo que se encuentra bloqueado por operaciones de entrada/salida.

7- Tiempo de espera.

Es el tiempo en que los procesos están activos pero sin ser ejecutados, es decir, los tiempos de espera en las distintas colas.

Es la suma de los períodos que el proceso para esperando en dicha cola. El algoritmo de planificación de la CPU afecta el tiempo que un proceso pasa esperando en la cola de procesos listos.

8- Rendimiento.

Es el número de trabajos o procesos realizados por unidad de tiempo, que debe ser lo mayor posible.

Es el número de procesos que se completan por unidad de tiempo. En el caso de procesos largos, esta taza podría ser de un proceso por hora; en caso de transacciones cortas, podría ser 10 procesos por segundo.

Maximizar el número de tareas procesadas por hora.