

Evaluación del rendimiento

La selección de una determinada política de planificación de procesos se basa en un conjunto de parámetros de rendimiento cuya importancia relativa depende de algunas características particulares del sistema (por ejemplo, interactivo o batch, existencia de procesos de tiempo real...), lo que determina los compromisos que hay que establecer en la selección de las política y mecanismos de la gestión de procesos.

En general, todo mecanismo introduce una sobrecoste en tiempo de ejecución que contrarresta parte de los beneficios obtenidos por la introducción del mecanismo.

Los parámetros de rendimiento que utilizaremos como criterios para analizar la calidad de una política de planificación, particularizados para su aplicación al recurso procesador, son:

Eficiencia – Uso efectivo de CPU

Se refiere a la eficiencia temporal. Se expresa como el porcentaje de tiempo en que la CPU se mantiene ocupada haciendo trabajo útil. Por trabajo útil se entiende la ejecución de código de los programas (y de los servicios solicitados por éstos). Cabe esperar que un sistema multiprogramado sea mucho más eficiente que uno monoprogramado, ya que en éstos la CPU está ociosa cuando en programa espera por una operación de E/S pudiendo haber programas esperando a ejecutarse, por lo que ese tiempo contará como tiempo perdido. Sin embargo, un sistema multiprogramado tiene que ejecutar las tareas propias de la gestión de procesos (scheduler y del dispatcher), y éstas deben computarse también como tiempo perdido.

$$Ef_t = \frac{t_{\text{útil}}}{T} = \frac{t_{\text{útil}}}{t_{\text{útil}} + t_{\text{perdido}}}$$

A veces resultará interesante referirse a la pérdida de eficiencia como:

$$1 - Ef_t = \frac{t_{\text{perdido}}}{T}$$

Productividad (throughput)

En lo que respecta a la gestión de procesos, mide el número de programas que se ejecutan por unidad de tiempo. Incluye otras muchas características que afectan el rendimiento del sistema, como por ejemplo la velocidad del procesador, que habrá que compensar si se comparan máquinas con distinto hardware.

Tiempo de retorno o finalización

Considera el rendimiento del sistema desde el punto de vista del programa que se ejecuta. Globalmente, se puede expresar como el tiempo desde que se solicita la ejecución de un programa hasta que ésta finaliza. Es una medida válida para sistemas batch.

Lo denotaremos t_f

Tiempo de espera

Mide exclusivamente los tiempos totales de espera de un proceso en la cola de preparados, lo denominamos por t_w , eliminando la dependencia de la duración del propio programa. Depende en cierta medida, sin embargo, del número de veces que éste se bloquea.

Debe tenerse en cuenta la siguiente relación entre tiempos: si t_{CPU} es el tiempo que pasa el proceso en la CPU y t_{bloq} es el tiempo total que el proceso está bloqueado, entonces

$$t_f = t_w + t_{CPU} + t_{bloq}$$

El tiempo de espera no cuenta la sobrecarga del sistema para los cambios de contexto/tarea. Por tanto, contando con este tiempo perdido o de sobrecarga t_{sob} , el tiempo de retorno o finalización de un proceso sería finalmente.

$$t_f = t_w + t_{CPU} + t_{bloq} + t_{sob}$$

Tasa de CPU

La relación entre el tiempo de CPU del programa y su tiempo de espera expresa la tasa de CPU, que indica el grado de disfrute del procesador que ha tenido el proceso:

$$TasaCPU = \frac{t_{CPU}}{t_{CPU} + t_w}$$

Latencia o Tiempo de respuesta

Mide el tiempo desde que un proceso entra en el estado de preparado (porque se crea o porque se desbloquea) hasta que entra en ejecución. Sólo tiene sentido en sistemas interactivos. Lo denotaremos t_r .