

Tema 4: Planificación

Sistemas Operativos

Javier Pérez Rodríguez
javier.perez[at]uco.es

Departamento de Informática y Análisis Numérico
Universidad de Córdoba

12 de diciembre de 2024



1. Definición y tipos de planificación
2. Objetivos y criterios de planificación.
3. Modos de decisión.
4. Tipos de procesos.
5. Algoritmos de planificación.

Qué es la planificación

- La **planificación** en SO es el proceso mediante el cual el sistema decide qué proceso debe ejecutarse en la CPU en un momento dado.
- **Objetivo:** gestionar de manera eficiente los recursos del sistema, garantizando que los procesos se ejecuten de forma ordenada y justa.
- Existen diferentes algoritmos de planificación, cada uno con sus propias estrategias para optimizar el rendimiento y cumplir con diversos criterios.



Tipos de planificación

El objetivo de la planificación de procesos es asignar procesos para ser ejecutados por el procesador de manera que se cumplan los objetivos del sistema, como el tiempo de respuesta y la productividad. En muchos sistemas, esta actividad de planificación se divide en tres funciones independientes: planificación a largo, medio y corto plazo.



Planificación a largo plazo

La planificación a largo plazo se realiza cuando se crea un nuevo proceso y se decide si se añade o no al conjunto de procesos activos. El planificador a largo plazo:

- Determina qué programas se admiten en el sistema para su procesamiento.
- Controla el grado de multiprogramación.
- Limita el número de procesos cargados en RAM para evitar que demasiados procesos compitan por la CPU.

Cuando el procesador está ocioso, se puede invocar al planificador a largo plazo para cargar nuevos procesos.



Planificación a medio plazo

La planificación a medio plazo se ocupa de decidir si se añade un proceso a la memoria principal o se mueve a la memoria secundaria (*swapping*). Este tipo de planificación es parte del manejo de memoria virtual y tiene como objetivo mejorar el rendimiento:

- En situaciones de falta de memoria, se mueven procesos de la memoria principal a disco (swap).
- En sistemas con poca memoria RAM, el planificador a medio plazo decide qué procesos deben ser cambiados a memoria secundaria.

El uso de *swapping* ha disminuido en los sistemas actuales debido a la mayor cantidad de memoria RAM disponible.



Planificación a corto plazo

La planificación a corto plazo decide qué proceso de la lista de Listos se ejecutará a continuación en la CPU. Este planificador se invoca con mucha frecuencia y toma decisiones de grano fino. Se utiliza cada vez que ocurre un evento que puede implicar la retirada del proceso actual, como por ejemplo:

- El proceso actual finaliza su ejecución o se bloquea.
- Un nuevo proceso de mayor prioridad llega a la lista de Listos.

El planificador a corto plazo es el que más afecta al rendimiento inmediato del sistema.



Frecuencia de ejecución de los planificadores

- **Planificador a largo plazo:** Ejecuta con baja frecuencia y toma decisiones de grano grueso, como la admisión de nuevos procesos.
- **Planificador a medio plazo:** Se ejecuta de manera poco frecuente, solo cuando se necesitan decisiones de intercambio debido a limitaciones de memoria.
- **Planificador a corto plazo:** Ejecuta con alta frecuencia y toma decisiones de grano fino sobre qué proceso ejecutar a continuación.

La frecuencia de ejecución de cada planificador varía según la función que desempeñan en el sistema.



Objetivos de la planificación

- La planificación debe realizarse en base a distintos objetivos, algunos de los cuales son contradictorios entre sí, por lo que se deben buscar soluciones de compromiso.
- La naturaleza de cada sistema determinará cuáles son los objetivos más importantes.



Objetivos más relevantes

- **Justicia:** El algoritmo de planificación debe asignar una porción de tiempo de CPU justa a todos los procesos. Ningún proceso debe acaparar la CPU.
- **Aplazamiento indefinido:** Se debe evitar la inanición de los procesos, cuando un proceso nunca recibe servicio por parte del procesador o se posterga por mucho tiempo.
- **Productividad:** Se busca maximizar la cantidad de trabajo finalizado por unidad de tiempo. La productividad depende de la "longitud" de los procesos.



Objetivos más relevantes

- **Costo extra:** El tiempo utilizado en la planificación debe ser mínimo, ya que es tiempo innecesario para el sistema.
- **Tiempo de respuesta aceptable:** El tiempo que transcurre desde que se crea un proceso hasta que se obtiene la primera interacción (proceso ejecutándose).
- **Degradación aceptable:** Ante una subida de carga, el sistema debe procurar una degradación del rendimiento suave, evitando caídas globales.



Criterios para la planificación

Los administradores de centros de cómputo utilizan una serie de métricas para verificar el desempeño de los sistemas y comparar el rendimiento de los algoritmos de planificación. Algunos de estos criterios son:



Tiempo de retorno (Turnaround Time - tt)

El **turnaround time** es el tiempo que transcurre desde que se crea un proceso (estado Nuevo) hasta que se completa (estado Saliente). Incluye:

- Tiempo de ejecución efectiva (uso de CPU)
- Tiempo de espera
- Tiempos de entrada/salida (E/S)



Tiempo de espera (Waiting Time - wt)

- El **tiempo de espera** es la suma de los tiempos que un proceso permanece en estados distintos al de Ejecución.
- Este tiempo no depende del algoritmo de planificación, pero sí de interrupciones y operaciones de E/S.
- El objetivo es minimizar el tiempo medio de espera.



Utilización de la CPU (CPU utilization)

- La **utilización de la CPU**, desde el punto de vista de un proceso, es el tiempo que está en la CPU o el tiempo que necesitaría para completarse si no hubiera otros procesos en el sistema.



Modos de decisión en la planificación

Los algoritmos de planificación se dividen en dos categorías según cómo manejan los instantes de tiempo en los que se selecciona un proceso para ejecutar en la CPU.



Algoritmos No apropiativos o sin expulsión

En los algoritmos no apropiativos, el planificador selecciona un proceso para ejecutarlo y lo deja en la CPU hasta que:

- El proceso se bloquea (por una llamada bloqueante o de E/S).
- El proceso termina (de manera normal o abrupta).
- Se produce una interrupción que no tiene que ver con la planificación (por ejemplo, interrupción de una operación de E/S de otro proceso).

En este modelo no hay rodajas de tiempo.



Algoritmos Apropiativos o con expulsión

En los algoritmos apropiativos, el planificador selecciona un proceso y puede expulsarlo de la CPU si ocurre alguno de los siguientes eventos:

- El proceso se ejecuta durante un tiempo fijo (quantum o rodaja de tiempo). Si termina el quantum, el proceso pasa a la lista de Listos y el planificador selecciona otro proceso.
- Llega un nuevo proceso con mayor prioridad.

Este enfoque previene la inanición de procesos, pero tiene mayor sobrecarga debido a los frecuentes cambios de contexto.



Eventos que provocan la expulsión de un proceso

Independientemente del modo de decisión, algunos eventos comunes provocan la expulsión de un proceso de la CPU:

- Se produce una interrupción (por parte del controlador de E/S) proveniente de una operación de E/S anterior, lo que interrumpe la ejecución del proceso actual.
- El proceso actual pasa a estado Bloqueado, por ejemplo:
 - Llamadas al sistema que provocan E/S (e.g., `open()`, `read()`).
 - Llamadas bloqueantes como `waitpid()` o `pthread_join()`, o mecanismos de exclusión mutua como semáforos (`semWait()`).



Tipos de procesos

La decisión de cuál es el siguiente trabajo a admitir para la CPU se basa en encontrar un compromiso entre procesos limitados por el procesador y procesos limitados por la E/S:

- **Proceso limitado por el procesador:** Realiza mucho trabajo computacional y utiliza los dispositivos de E/S ocasionalmente o nunca.
- **Proceso limitado por la E/S:** Pasa más tiempo utilizando los dispositivos de E/S que el procesador.

El compromiso entre estos tipos de procesos depende de los recursos de E/S disponibles, buscando equilibrar su uso.

