

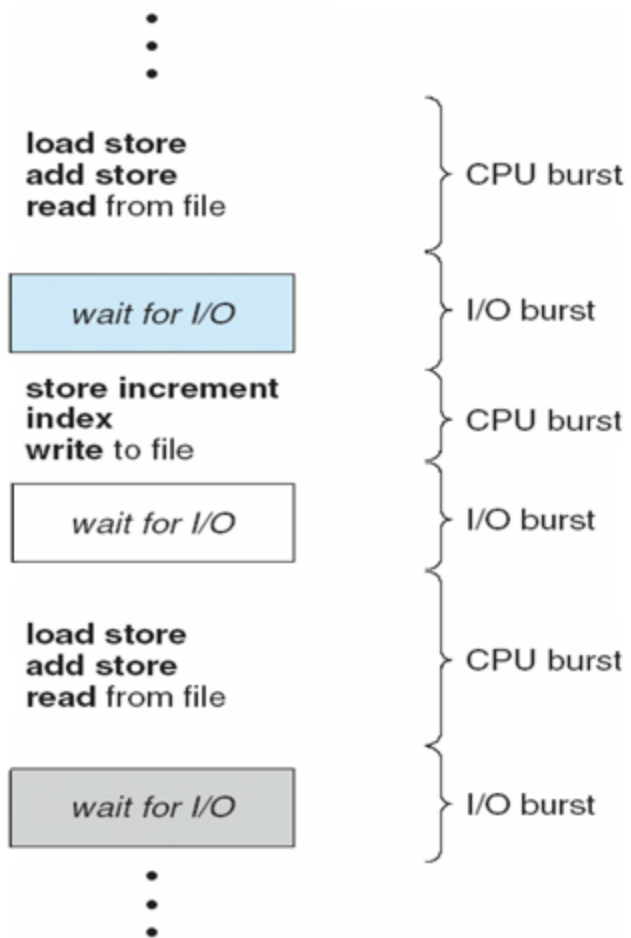
Sistemas Operativos I

Planificación

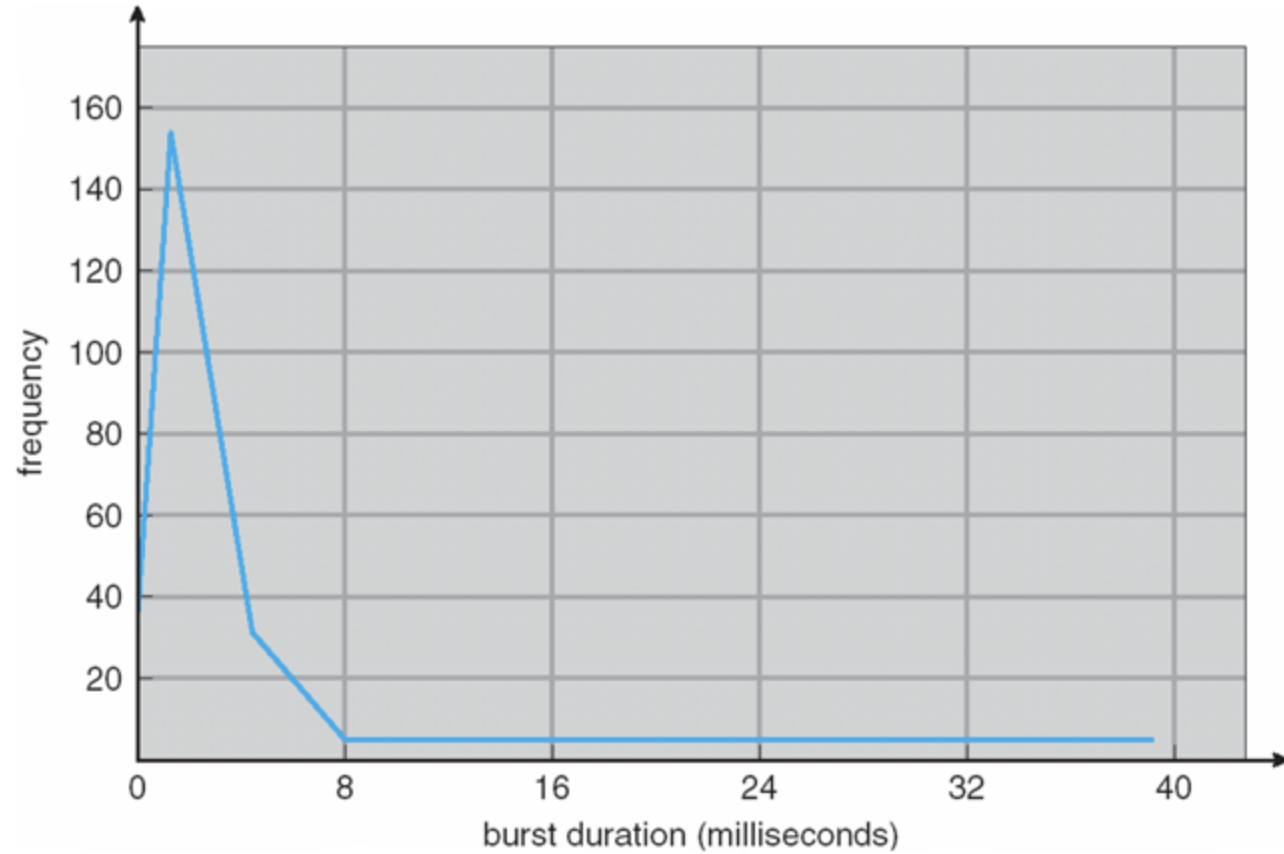
Conceptos Básicos

- La utilización máxima de CPU se obtiene por medio de la multiprogramación
- Ciclo de ráfagas de CPU–E/S – Los proceso en ejecución consisten de ciclos de ejecución en CPU u espera por E/S
- Siempre empieza con una ráfaga (primera instrucción) y termina con una ráfaga (pedido de terminación del proceso – `exit`)

Distribución de Ráfagas



Frecuencia de Ráfagas de CPU



Planificador de Corto Plazo (CPU)

- Selecciona entre los procesos en memoria que están listos para ejecutarse, y le asigna la CPU a uno
- La planificación de CPU puede suceder cuando un proceso:
 - Cambia de ejecutando a esperando (p. ej., esperando por un hijo o E/S)
 - Cambia de ejecutando a listo (p. ej., una interrupción)
 - Cambia de esperando a listo (p. ej., se completa la E/S)
 - Termina

Planificador de CPU

- La planificación en los casos 1 y 4 es no-expropiativa
 - Una vez que la CPU se asigna a un proceso, el proceso la mantiene hasta que la libera o bien terminando o cambiando a esperando
- El resto de los casos la planificación es expropiativa
- Tiene un costo asociado con el acceso a información compartida

Dispatcher

- El despachador (dispatcher) es el módulo del SO que da control de la CPU al proceso seleccionado por el planificador de corto plazo; esto implica:
 - Cambiar de contexto
 - Cambiar a modo usuario
 - Ir a la dirección apropiada en el proceso del usuario y reiniciarlo
- Latencia de despacho: es el tiempo que toma al despachador para el proceso actual y arrancar el proceso seleccionado

Criterios de Planificación

- Utilización CPU: mantener la CPU lo más ocupada posible (p. ej. 40-90%)
- Rendimiento: # de procesos que terminan por unidad de tiempo
- Tiempo de retorno: tiempo insumido en ejecutar un proceso en particular, incluye:
- Tiempo de espera para ingresar en memoria
- Tiempo de espera en la cola de listos
- Tiempo de ejecución de CPU
- Tiempo de E/S

Criterios de Planificación

- Tiempo de espera – cantidad de tiempo que un proceso queda esperando en la cola de listos a ser ejecutado (este tiempo sólo es afectado por el algoritmo de planificación)
- Tiempo de respuesta – cantidad de tiempo que lleva un proceso hacer un pedido hasta la primer salida (para sistemas interactivos o tiempo compartido)

Criterios de Planificación

- Máxima utilización de CPU
- Máximo rendimiento
- Mínimo tiempo de retorno
- Mínimo tiempo de espera
- Mínimo tiempo de respuesta

Para sistemas interactivos – minimizar la varianza del tiempo de retorno

Planificación FCFS

Proceso	Ráfaga
P1	24
P2	3
P3	3



- Efecto Convoy – procesos cortos detrás de procesos largos → menor utilización de CPU y dispositivos
- FCFS es no-expropiativo

Shortest Job First (SJF)

- Asocia con cada proceso la longitud de la siguiente ráfaga de CPU. Usa esa información para planificar el proceso con la ráfaga más corta
- 2 formas:
 - **no-expropiativo** – una vez que se asignó la CPU a un proceso, no puede expropiarla hasta que el proceso termine su ráfaga
 - **expropiativo** – si llega un proceso con una ráfaga de CPU más corta que el remanente del proceso en ejecución, se expropia. Este esquema se conoce como Shortest-Remaining-Time-First (SRTF)

Proceso	Llegada	Ráfaga
P1	0.0	10
P2	2.0	8
P3	4.0	2
P4	5.0	8



Reflexiones sobre SJF

- En realidad es “la ráfaga más corta primero”...
- SJF da el menor tiempo de espera promedio para cualquier conjunto de procesos (moviendo el más corto antes de uno largo disminuye el tiempo de espera más que lo que aumenta la espera del proceso largo)
- Lo difícil es conocer la longitud de la siguiente ráfaga de CPU

Ideal para planificadores de largo plazo si se pide al usuario que estime el tiempo de procesamiento:

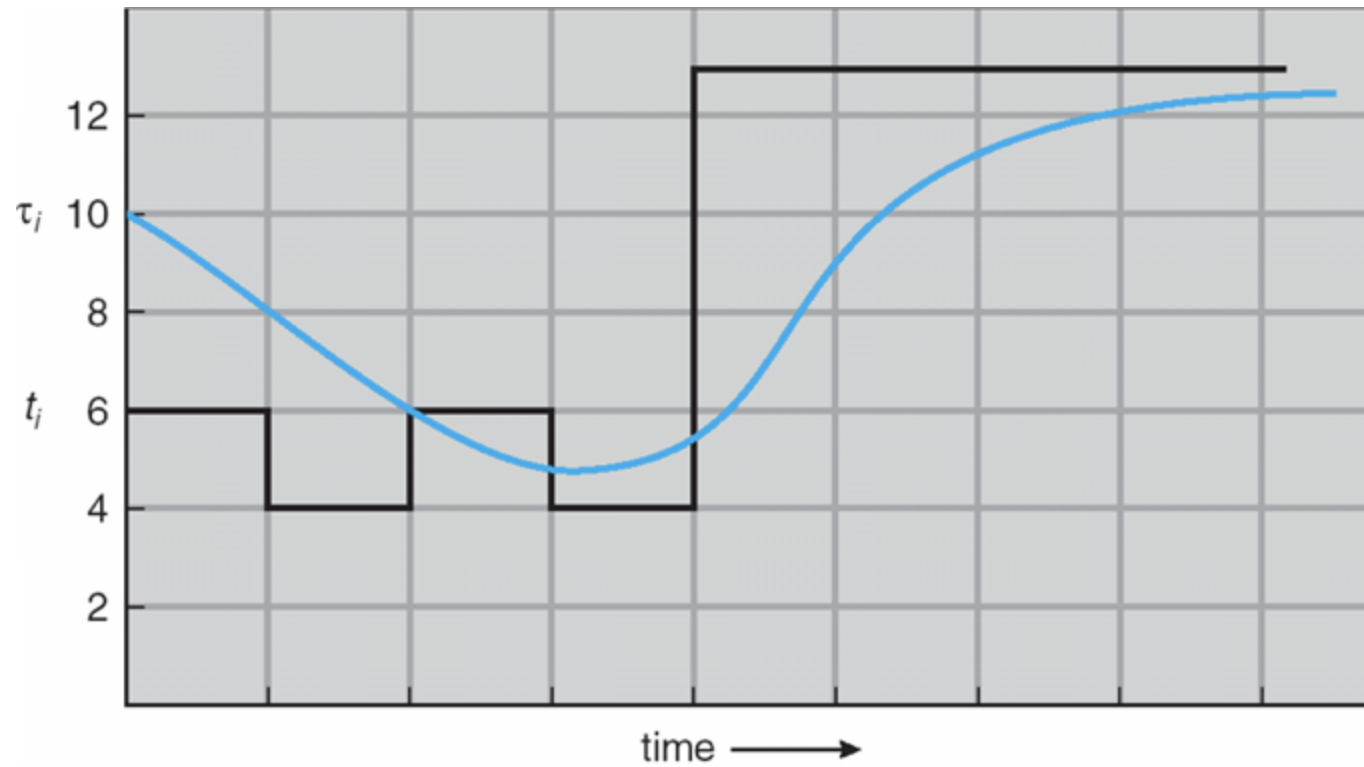
- **Estimación baja** – enviar nuevamente
- **Estimación alta** – la tarea va a ser postergada

Estimación de la Siguiente Ráfaga

- Sólo podemos estimar la longitud de la ráfaga
- Puede hacerse utilizando las longitudes de las ráfagas previas usando un promedio exponencial

1. t_n = longitud de la ráfaga de CPU n
2. τ_{n+1} = valor estimado de la siguiente ráfaga
3. $\alpha, 0 \leq \alpha \leq 1$
4. Se define: $\tau_{n+1} = \alpha t_n + (1 - \alpha)\tau_n$

α determina el peso del comportamiento pasado



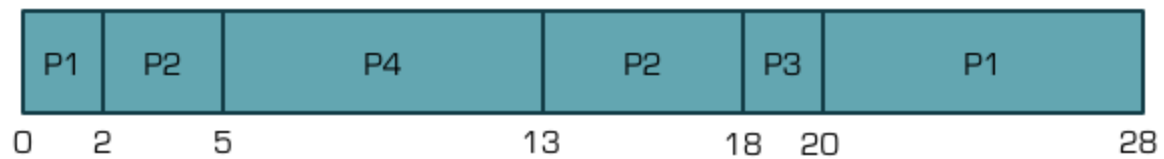
CPU burst (t_i)		6	4	6	4	13	13	13	...
"guess" (τ_i)	10	8	6	6	5	9	11	12	...

¿Cuál es el valor de α utilizado?

Planificación Por Prioridad

- Se asocia a cada proceso un número de prioridad (entero)
- Se asigna la CPU al proceso con mayor prioridad: expropiativo y no-expropiativo
- SJF es un caso particular de prioridad donde la prioridad es la longitud de la siguiente ráfaga
- **Problema Inanición** – los procesos de baja prioridad podrían no ejecutarse nunca

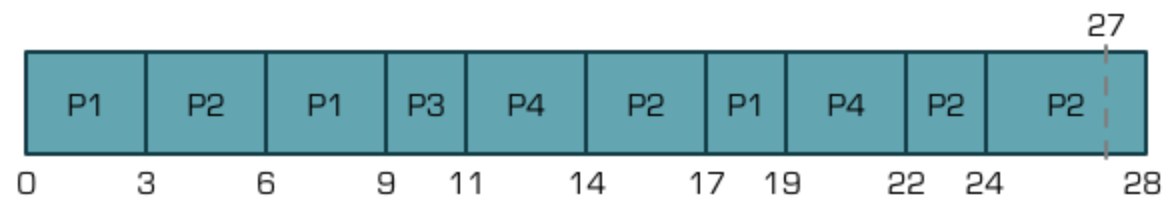
Proceso	Llegada	Ráfaga	Prioridad
P1	0.0	10	3
P2	2.0	8	2
P3	4.0	2	2
P4	5.0	8	1



Round-Robin

- Diseñado especialmente para sistemas interactivos
- Cada proceso obtiene una pequeña unidad de tiempo de CPU (time quantum), generalmente de **10-100ms**. Una vez finalizado, el proceso es expropiado y agregado a la cola de listos
- Si hay n procesos listos y un quantum de q , entonces cada proceso obtiene $1/n$ del tiempo de CPU en porciones de a lo sumo q unidades de tiempo
- Rendimiento
 - q grande \Rightarrow FIFO
 - q chico \Rightarrow q debe ser grande respecto al cambio de contexto, de otra forma el overhead es alto

Proceso	Llegada	Ráfaga
P1	0.0	10
P2	2.0	8
P3	4.0	2
P4	5.0	8



Tiempo de CPU = 10



quantum

12

cambios
de contexto

0



6

1



1

9