

TEMA 2

GESTIÓN DE PROCESOS

2.1 INTRODUCCIÓN A LOS PROCESOS

Sistema operativo: Conjunto de extensiones software del hardware original, que culminan en una máquina virtual que sirve como un entorno de programación de alto nivel que gestiona el flujo de trabajo de una red de computadores. El S.O. proporciona una serie de servicios y una interfaz a dichos servicios.

Multitarea: Capacidad que tienen los S.O. de ejecutar de forma simultánea varios procesos.

Multiprogramación: Se intenta mantener del modo más ocupado posible a todos los recursos del sistema con el entrelazado y el solapamiento de la ejecución de más de un programa.

Solamente cambiando la atención de un único procesador entre varios programas y dejando que cada uno se ejecute unas milésimas de segundo antes de ir a otro, permite crear la ilusión de la ejecución concurrente o al unísono de los programas.

Además de la multitarea incluye la capacidad de gestión de la memoria y de los ficheros.

Sistemas de tiempo compartido: Sistemas con multiprogramación interactiva en los que se asigna a cada programa un mismo intervalo de tiempo de ejecución.

2.1.2 DISTINCIÓN ENTRE PROGRAMAS Y PROCESOS

Cada invocación a un programa genera un proceso distinto e independiente. Los procesos compiten por el uso de los recursos del sistema, de manera que se necesitan herramientas que permitan la sincronización y la comunicación entre los procesos. En cualquier instante de tiempo, sólo un proceso estará en ejecución mientras que los otros estarán suspendidos, preparados o esperando a que ocurra algún evento.

2.2 RELACIÓN ENTRE LOS PROCESOS

El sistema operativo debe proporcionar los servicios necesarios que permitan :

- Ejecución concurrente de los procesos.
- Sincronización entre los procesos.
- Comunicación entre procesos.

Además, debe de disponer de algoritmos de gestión y planificación de procesos que se encarguen de :

- Decidir que procesos se ejecutará o tomará el procesador.
- Llevar cuenta de los estados de los procesos, sus prioridades y toda la información relevante acerca de ellos.

Dependiendo de la interacción entre los procesos, éstos se clasifican en:

- Independientes:** No se comunican o sincronizan entre ellos. En un sistema con un solo procesador, los procesos independientes en sentido estricto no existen.
- Cooperativos:** Se comunican y sincronizan sus actividades para realizar una labor común.
- Competitivos:** Todos compiten por los recursos del computador. El acceso ordenado a los recursos a veces necesita de la sincronización y a veces, también de la comunicación entre los procesos.

2.3 ESPECIFICACIÓN DE LOS PROCESOS

Implícitos: Procesos generados y eliminados por el S.O.

Explícitos: Definidos por el usuario realizando llamadas al S.O.

Proceso Padre: Que crea otro proceso. Puede tener varios hijos.

Proceso Hijo: Proceso creado por otro. Sólo puede tener un padre

El S.O. hace que los argumentos y variables locales del procedimiento o función que da lugar al proceso se asocien a éste y no al código en el que aparecen. El sistema crea para cada proceso un conjunto independiente de variables locales, parámetros formales y llamadas a procedimientos y a las funciones.

TEMA 2

GESTIÓN DE PROCESOS

2.4 ESTADOS DE LOS PROCESOS

- a) **Activo:** Proceso que está ejecutándose en un instante dado.
- b) **Preparado:** Listo para ejecutarse, esperando a que el procesador quede libre.
- c) **Bloqueado o Suspendido:** Esperando a que se cumpla alguna condición. (E/S, señal, etc.)
- d) **Muerto:** Ha terminado su ejecución o el S.O. ha detectado un error fatal.
- e) **Nonato:** Existe, pero todavía no es conocido por el S.O.

Estado global del sistema: Conjunto de recursos y procesos existentes con sus estados correspondientes en un instante determinado. El S.O. lo cambia en respuesta a eventos externos o internos.

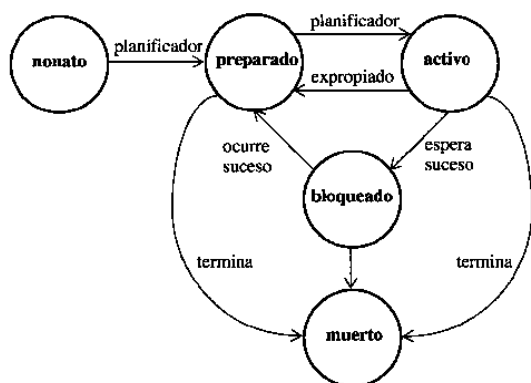
El S.O. dispone de algoritmos de planificación y de gestión de los procesos y de los recursos de modo que, considerando como entradas los eventos y el estado global del sistema, genera un nuevo estado global.

2.4.1 TRANSICIONES ENTRE LOS ESTADOS

Distribuidor (dispatcher): Módulo del S.O. que se encarga de activar los procesos que están en el estado preparado.

Interrupción: Hace que la tarea que está activa deje de ejecutarse a favor del S.O. que decidirá cuál de los procesos preparados pasa a activo.

Planificador (scheduler) : Hace que el proceso activo vuelva a preparado sin que medie una llamada al sistema (por fin de tiempo asignado o proceso preparado con mayor prioridad).



De	A	Motivos
nonato	preparado activo	Al crearse
preparado	activo	Activado por el distribuidor o por el planificador
activo	preparado	Interrupción
activo	bloqueado	Solicita un recurso necesario para continuar (ej.: E/S)
nonato		Necesita que tenga lugar un evento
preparado	bloqueado	Acción determinada, como una llamada al sistema del proceso activo
bloqueado	preparado	Tiene lugar un evento que estaba esperando
activo	muerto	Finaliza
preparado		Eliminado por otro proceso
bloqueado		

2.5 EL BLOQUE DE CONTROL DE PROCESOS

Definición. Estructura de datos en la que el sistema mantiene toda la información sobre un proceso. Contiene:

- a) Identificador único de proceso (pid).
- b) Estado del proceso (activo, preparado, bloqueado, etc.)
- c) Prioridad
- d) Estado hardware (contador de programa, códigos de condición, punteros de pila, etc.)
- e) Información para gestionar la memoria (punteros, tablas, registros)
- f) Información de contabilidad y planificación.

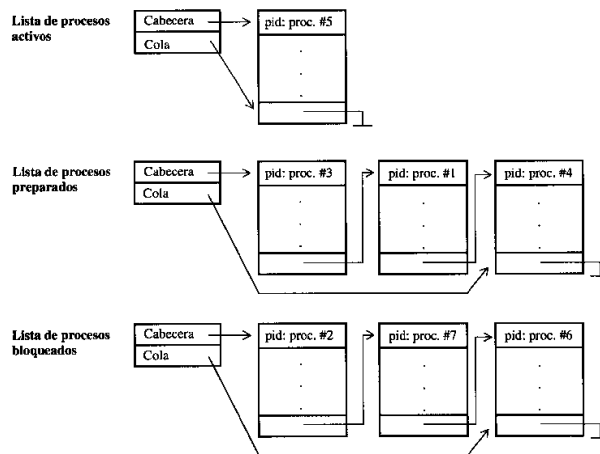
pid del proceso
estado del proceso
prioridad
contador de programa
registros
...

TEMA 2

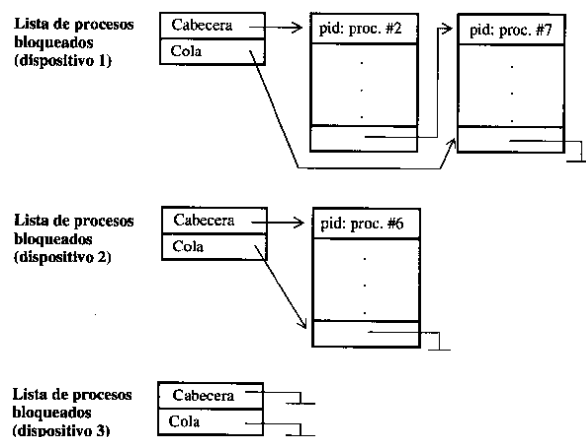
GESTIÓN DE PROCESOS

2.5.1 LISTAS DE PROCESOS

El S.O. mantiene listas de BCP para cada uno de los estados del sistema (preparado, activo, etc.). Cada proceso pertenece a una única lista.



El planificador se encarga de gestionar el paso de los procesos de una lista a la otra. Puede haber una lista de estados suspendidos ligada en exclusiva a un dispositivo o evento.



2.6 PROCESOS Y HEBRAS

Podemos concebir el proceso como una entidad formada por una o más unidades de ejecución denominadas hebras (threads) o procesos primitivos y un conjunto de recursos asociados.

Cada hebra posee su propio vector de estado. No pueden existir fuera de un proceso y sólo pueden pertenecer a uno de ellos.

Los recursos no están asociados a la hebra sino al proceso.

TEMA 2

GESTIÓN DE PROCESOS

2.7 EL PLANIFICADOR DE PROCESOS

Definición. Software del S.O. que se encarga de asignar los recursos de un sistema entre los procesos que los solicitan.

PLANIFICADOR A LARGO PLAZO (PLP)

También llamado planificador de trabajos. Determina que procesos se admiten en el sistema para su procesamiento y son cargados en la memoria disponible. Debe realizar una mezcla adecuada de trabajos destinados al procesador y trabajos destinados al sistema de E/S. Lo más usual es definir una función de prioridad y asignarle a cada programa un valor para la misma que se actualiza en instantes determinados por el PLP.

Existen sistemas que no tienen algoritmo PLP o este es mínimo, como algunos sistemas de tiempo compartido, que se limitan a poner en el estado de preparado cualquier proceso nuevo. En general, el PLP sólo existe en sistemas que admiten procesamiento por lotes.

PLANIFICADOR A CORTO PLAZO (PCP)

También llamado planificador del procesador. Gestiona los procesos entre dos instantes de gestión del PLP. Selecciona al proceso que pasará al estado activo de entre todos los preparados. Su elevada frecuencia de uso hace que el algoritmo suela ser sencillo. Es invocado siempre que se origina un cambio de estado del sistema por un evento:

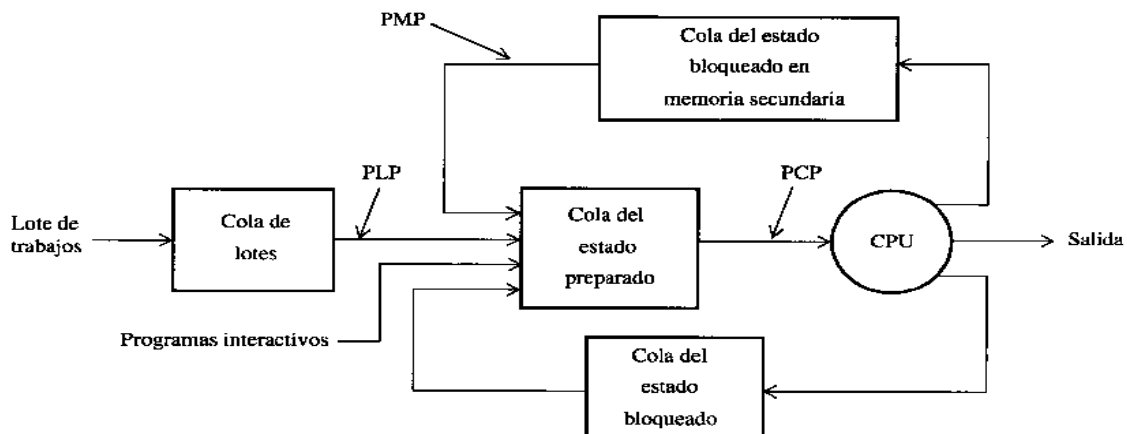
- Señales de reloj del sistema.
- Interrupciones.
- Finalización de operaciones de E/S.
- Llamadas al S.O.
- Envío y recepción de señales.
- Activación de programas interactivos.



PLANIFICADOR A MEDIO PLAZO (PMP)

En algunos sistemas. Se encarga del **intercambio** (swapping) o proceso de salvar una tarea suspendida en la memoria secundaria para :

- Liberar memoria principal.
- Separar los procesos destinados al procesador de los destinados al sistema de E/S.



TEMA 2

GESTIÓN DE PROCESOS

2.7.2 CRITERIOS PARA LA PLANIFICACIÓN

- a) **Eficacia:** Porcentaje del tiempo medio de utilización.
- b) **Rendimiento** (throughput) : Número de procesos completados por unidad de tiempo.
- c) **Tiempo de Regreso** (turnaround) : Intervalo de tiempo que transcurre desde que un proceso se crea o presenta hasta que se completa por el sistema.
- d) **Tiempo de Espera** : Tiempo que el proceso espera hasta que se le concede el procesador. Puede resultar una medida más adecuada de la eficiencia del sistema.
- e) **Tiempo de Respuesta a un Evento** : Intervalo de tiempo que transcurre desde que se señala un evento hasta que se ejecuta la primera instrucción.

2.7.3 PLANIFICACIÓN POR EXPROPIACIÓN

Los procesos con mayor prioridad reemplazan en el estado de ejecución a cualquier tarea con menor prioridad. Adecuada en aquellos sistemas en los que los procesos con una prioridad alta requieren una atención rápida. Cada vez que se produce un evento que origina un cambio en el estado global del sistema, se activa el PCP.

En los algoritmos con estrategia de no expropiación, el proceso que está activo permanece con el procesador hasta que él mismo devuelve el control al S.O. Hace un uso menos frecuente del PCP, lo que implica una mayor dedicación a los procesos de los usuarios frente a los del sistema.

2.7.4 PLANIFICACIÓN POR PRIORIDADES

Cada proceso tiene asignada una prioridad y el de mayor prioridad en el estado preparado es el que toma el procesador. El valor inicial puede ser asignado por el usuario o por el sistema. La asignación puede ser:

- a) Estática.
- b) Dinámica. Puede ser modificada por el usuario o por el sistema.

La modificación se suele realizar en función de ciertos parámetros como la cantidad de memoria que utiliza, el número de acciones de E/S que lleva realizado, el tiempo medio de utilización del procesador hasta ese momento, el número de ficheros abiertos, etc.

Los algoritmos pueden ser expropiativos o no expropiativos. Los algoritmos con expropiación y con prioridades se dicen que están guiados por eventos (event-driven) .

Se suele ir aumentando la prioridad de aquellos procesos que llevan un tiempo de espera muy elevado, prioridad por envejecimiento.

2.7.5 PLANIFICACIÓN FCFS: primero en llegar primero en ser servido.

First Come First Served.

Modo más sencillo de planificación en el que los procesos en el estado preparado acceden al procesador en el orden con que llegan dicho estado. No es expropiativo. Suele emplearse dentro de otros esquemas.

2.7.6 PLANIFICACIÓN SJF: primera tarea más corta.

Shortest Job First

Se selecciona el trabajo con el menor valor de tiempo restante de ejecución. Si dos trabajos tienen el mismo tiempo, se seleccionan según la estrategia FCFS.

Minimiza el tiempo medio de espera de los trabajos. Su dificultad estriba en conocer el tiempo que le queda a cada trabajo para terminar de ejecutarse.

En entornos de producción donde los trabajos se ejecuten regularmente es posible tener una estima aceptable.

2.7.7 PLANIFICACIÓN SRT: tiempo que queda más corto.

Shortest Remaining Time.

Versión expropiativa del método SJF. Se elige a continuación aquel proceso al que le queda menos tiempo para terminar su ejecución.

Mayor frecuencia de invocación del planificador y mayor carga en las labores que tiene que realizar.

TEMA 2

GESTIÓN DE PROCESOS

2.7.8 PLANIFICACIÓN RR: prioridad circular.

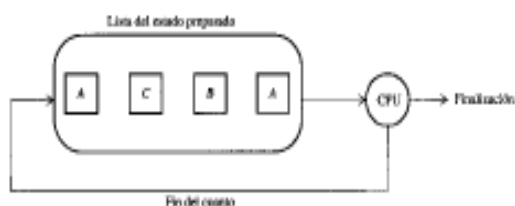
Round Robin.

A todos los procesos en estado preparado se les asigna un tiempo de ejecución denominado **cuanto**.

Muy apropiado en sistemas de tiempo compartido

Necesita de un temporizador que controle el cuanto asignado a los procesos y genere una interrupción siempre que se produzca su finalización. La interrupción hace que se llame al PCP, que se encarga de guardar el contexto del proceso en ejecución, lo lleva al final de la lista de procesos preparados para ejecutarse y pasa a ejecución al proceso cabecera de la lista.

La elección del cuanto supone llegar a un compromiso entre la eficacia del procesador (que exigiría un valor no demasiado pequeño) y el tiempo de respuesta (que si lo exigiría).



El proceso A pasa ejecución, una vez transcurrido un cuanto vuelve a la cola del estado preparado.

2.7.9 PLANIFICACIÓN BASADA EN EL RELOJ DE TIEMPO REAL (RTR)

La mayoría de los sistemas disponen de un reloj de tiempo real que se encarga de generar una señal de forma periódica utilizada para producir una interrupción a un intervalo de tiempo fijo, previamente establecido.

La atención a estas interrupciones se debe diseñar para que se realice rápidamente. Los procesos suelen dar la mayor prioridad a las interrupciones del reloj.

En muchos casos puede ser conveniente ralentizar el reloj para, a pesar de perder precisión, lograr una funcionalidad más rica. A la velocidad real de ejecución efectiva del código de tratamiento se le conoce como velocidad de tic.

El S.O. utiliza el RTR para :

- Limitar el intervalo de tiempo que un proceso puede estar en ejecución.
- Proporcionar servicios a los usuarios, como iniciar tareas cíclicas o la función de retardo.

El sistema mantiene una lista de eventos ordenados por el instante de tiempo en el que tienen que ocurrir. Cuando se produce una interrupción del RTR, se examina la lista y se inician aquellos para los que ya ha transcurrido un ciclo o el retardo.

La mayoría de los sistemas disponen además del RTR un reloj con la hora del día. Es posible tanto para el usuario como para el sistema conocer la fecha y la hora y utilizar estos datos para desencadenar eventos.

2.7.10 PLANIFICACIÓN MLQ: colas multinivel

Multi Level Queues.

Se clasifican las tareas en diferentes grupos a los que se aplican distintas estrategias de planificación. Se crean colas de tareas separadas que se gestionan por criterios diferentes. Cada tarea se asigna a una sola cola de acuerdo con alguna propiedad de la tarea. Además, debe existir un criterio de planificación entre las colas. Normalmente éste suele ser el de prioridad fija con expropiación. No puede ejecutarse ninguna tarea de una cola si las que le preceden en prioridad tienen alguna tarea pendiente.

Otra posibilidad es la de compartir el tiempo entre las colas, por ejemplo, de modo proporcional a las prioridades de las cola.

TEMA 2

GESTIÓN DE PROCESOS

2.7.11 PLANIFICACIÓN MLFQ: colas multinivel con retroalimentación.

Se permite la movilidad entre las colas. Las colas se organizan según las características de uso del procesador. Las tareas que emplean poco el procesador se mantienen en las colas de mayor prioridad y las que lo utilizan mucho se sitúan en las colas de menor prioridad.

El movimiento de las tareas entre las colas se realiza según su comportamiento dinámico a lo largo de su tiempo de vida efectiva.

Los algoritmos posibles pueden definirse por los siguientes parámetros :

- Número de colas.
- Algoritmo de planificación de cada cola.
- Métodos que determinan el movimiento de las tareas entre las colas.
- Método que indica la cola en la que entra una tarea cuando necesita un servicio.

