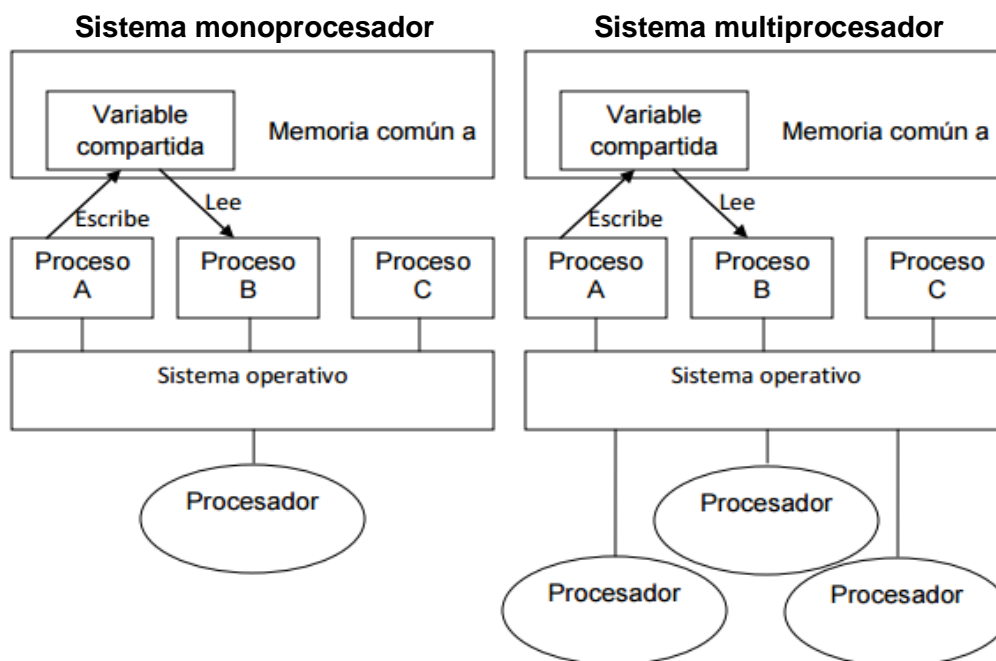


Semana # 10 – Paralelismo, Segmentación Encauzada & Sistemas Multiprocesadores

PARALELISMO EN SISTEMAS MONOPROCESADORES

Cuando un equipo de cómputo cuenta con un único procesador para ejecutar todos los procesos o tareas, se define entonces como un **sistema monoprocesador**. En la buena teoría, en este tipo de sistemas no existe el paralelismo.

La siguiente imagen muestra de forma resumida la diferencia entre un sistema monoprocesador y uno multiprocesador:



Aun así, el paralelismo podría ser posible en sistemas monoprocesadores o mejor dicho de alguna manera sucede. Por ejemplo, la Unidad Aritmético Lógica (UAL) realiza operaciones matemáticas y lógicas, por lo que podemos decir que esta parte del CPU se subdivide en dos o más unidades funcionales especializadas. Cada unidad lleva a cabo un ciclo de instrucción desarrollando tareas en paralelo.

Un segundo mecanismo es el llamado **segmentación del cauce** que consiste en describir las fases de ejecución de una cola de instrucción (**pipeline**). Dichas fases consisten es:

- Extracción de la instrucción.
- Decodificación.
- Extracción de operandos.
- Ejecución de operación.
- Almacenamiento.

Un tercer mecanismo consiste en ejecutar simultáneamente las operaciones de entrada/salida por medio de cálculos realizados por el CPU. Un problema que se puede presentar con esto, es que el proceso se puede volver muy lento. Para alivianar un poco la carga de trabajo, la **unidad de acceso directo a la memoria (DMA)** permite que la comunicación entre los diferentes dispositivos sea directa, sin necesidad de hacer uso del CPU.

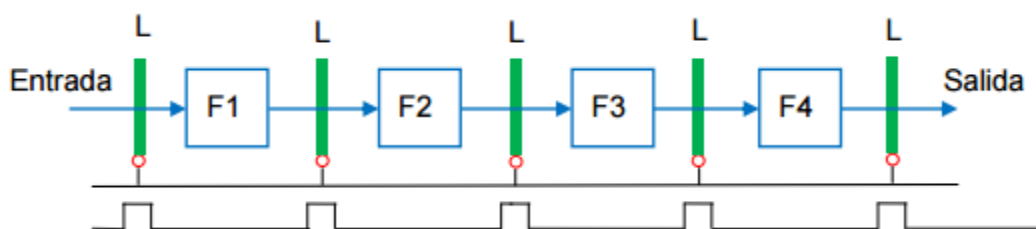
El último de los criterios que veremos en este tema es la **multiprogramación** y el **tiempo compartido**. Con la multiprogramación se busca la interacción de unos programas con otros ya que esto favorece una mejor gestión de los recursos mediante el uso de las operaciones de entrada/salida. El sistema de tiempo compartido es una extensión de la multiprogramación, debido a esto podemos decir que aplica el mismo principio. La principal ventaja del uso de tiempo compartido es que **evita que un proceso sea monopolizado por el procesador**. Este proceso es gestionado por el Sistema Operativo.

Una computadora paralela la podemos definir como un dispositivo que cuenta con dos o más procesadores interconectados, los cuales son capaces de intercambiar información y ejecutar instrucciones de forma simultánea. Este tipo de computadoras están divididas en tres tipos: **de segmentación encauzada, matriciales y sistemas multiprocesadores**, cada una de ellas explota y obtiene el paralelismo de forma diferente.

SEGMENTACIÓN ENCAUZADA

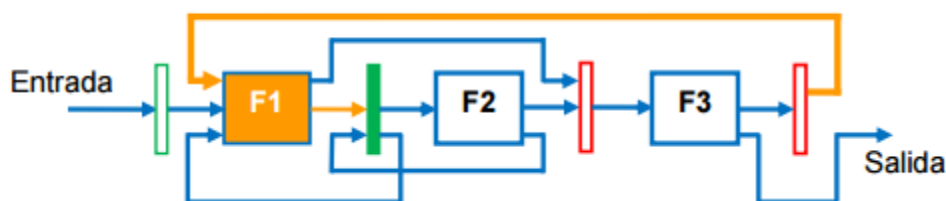
Segmentación encauzada lineal

En este tipo de segmentación, todas las etapas de las unidades funcionales son ejecutadas de manera secuencial, con la característica de que una tarea no puede dar inicio hasta que la tarea que la precede haya finalizado. En este tipo de segmentación no pueden haber lazos hacia atrás, algo que sí sucede con la **segmentación encauzada no lineal**. Veamos la siguiente imagen:



Segmentación encauzada lineal

Máquina X

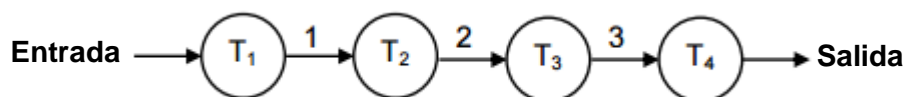


Segmentación encauzada no lineal

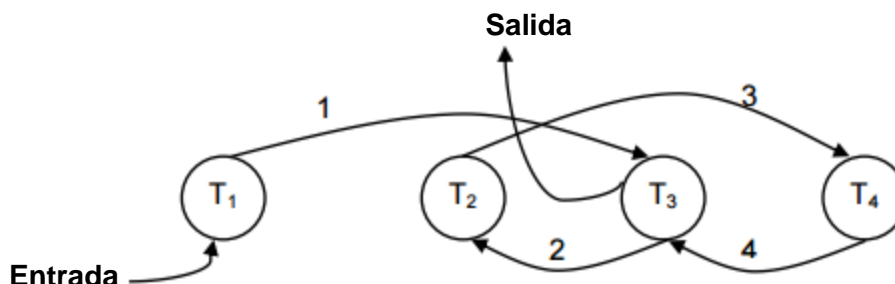
En el caso de la segmentación encauzada no lineal, hay posibilidad de tener lazos hacia atrás y se puede pasar más de una vez por cada etapa en cada iniciación de datos.

Veamos otro ejemplo, haciendo uso de grafos de procedencia.

En el caso de la segmentación lineal, las tareas pasan por todas y cada una de las etapas, en orden y una única vez, con un solo punto de entrada y otro de salida. Se observa también que en la segmentación no lineal, cada tarea puede no pasar por todas las etapas y tener múltiples puntos de entrada y salida.



Grafo de precedencia lineal



Grafo de precedencia no lineal

Clasificación de los procesadores encauzados

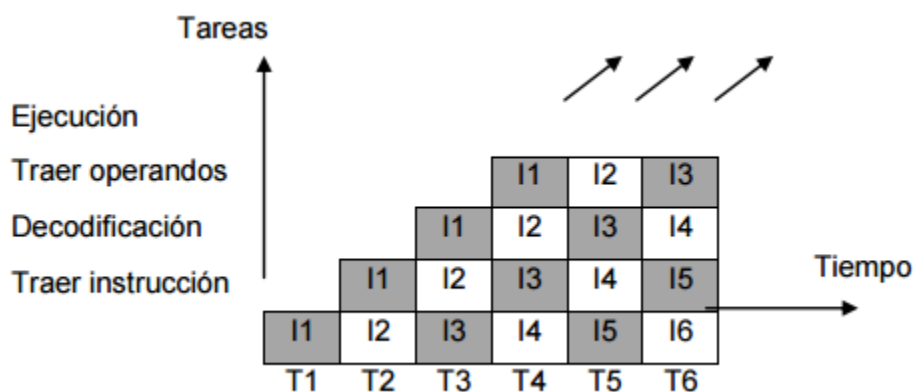
Existen varios criterios:

Frecuencia de cambio de función	Capacidad	Posibilidad o no de retroalimentación	Aplicación
Estáticos: Asumen una configuración funcional única.	Unifunción: Puede realizar una única función sobre los datos, por ejemplo el sumador de coma flotante.	Lineales: La ejecución de la tarea es secuencial.	Operadores aritméticos segmentados: Realizan varias operaciones encauzadas de ALU en diferentes formatos de datos.
Dinámicos: Asumen varias configuraciones funcionales simultáneamente y por esto pueden ser de rendimiento muy bajo.	Multifunción: Dependiendo de la señal de control, puede ejecutar varias funciones sobre los datos.	No lineales: Permiten tener varios puntos de entrada y salida.	Cauces del flujo de instrucción: Por lo general cauces lineales, algunas de sus fases pueden sub-segmentarse y permitir solapar la ejecución de la instrucción con la decodificación y búsqueda de los operandos de las instrucciones subsiguientes.

Cómo resumen podemos decir sobre la **segmentación encauzada**:

Se mejora el rendimiento incrementando la productividad de las instrucciones en lugar del tiempo de ejecución en cada una de ellas, es decir que no reduce el tiempo que tarda una instrucción en ejecutarse, sino que incrementa el número de instrucciones para ser ejecutadas simultáneamente.

Veamos la siguiente figura:



Representación gráfica de tiempo/tareas de una instrucción

La imagen que arriba lo que muestra es que una instrucción requiere de cuatro etapas:

- Traer la instrucción
- Decodificación
- Traer operandos
- Ejecutar la instrucción

En este caso, la segmentación encauzada mejora la velocidad porque permite que las cuatro etapas sean independientes y trabajen de manera paralela en el mismo ciclo de reloj.

La duración de ciclo de reloj es importante en la segmentación encauzada, ya que es la encargada de sincronizar el total de las etapas que constituyen la segmentación.

SISTEMAS MULTIPROCESADORES

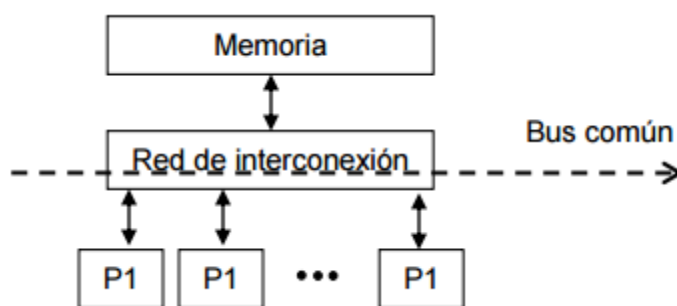
En la actualidad, los sistemas que se emplean para brindar servicios en una red se les conoce con el nombre de **servidores**. Podemos afirmar que las nuevas arquitecturas vienen con al menos dos sockets, para instalar uno o dos procesadores. Es también muy común encontrar servidores con múltiples sockets y esto se debe a la carga de trabajo que el sistema puede llevar a cabo. Imaginemos un servidor de correo electrónico para una empresa con 6000 empleados, esto sucede por ejemplo en HP Costa Rica.

En este caso, el servidor de correo o los servidores de correo no podrían sobrevivir con un único CPU y es por esto que la posibilidad de tener más de un CPU es hoy en día toda una realidad. A esta configuración se le conoce con el nombre de **sistemas multiprocesadores**.

Son arquitecturas construidas con dos o más procesadores, esto nos permite ejecutar de manera simultánea varios procesos, su ventaja reside en la operación llamada cambio de contexto que consiste en quitar un proceso de la CPU, ejecutar otro y volver a colocar el primero sin que se entere de nada. Los dos objetivos primordiales de los multiprocesadores son acelerar la ejecución de aplicaciones orientadas a la supercomputación y ejecutar más aplicaciones por unidad de tiempo.

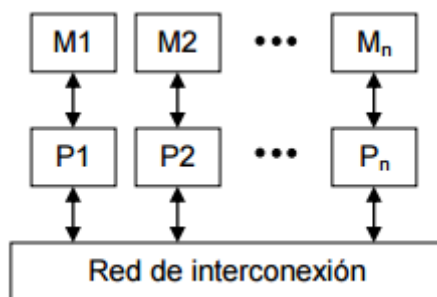
Los dos modelos de organización son los siguientes:

- **Memoria compartida:** Por medio de una red de interconexión, cada procesador puede acceder a cualquier posición de la memoria. El problema, es que dicha red de interconexión se puede convertir en un cuello de botella, por lo tanto este modelo **no es escalable**.



Multiprocesadores con memoria compartida

- **Memoria distribuida:** La red de interconexión permite que cualquier procesador se comuniquen con cualquier memoria del sistema. Este modelo presenta una gran potencia de cálculo.



Multiprocesadores con memoria distribuida