



Educación para una Vida Exitosa

UNIVERSIDAD INTERNACIONAL DEL ECUADOR

ESCUELA DE INFORMÁTICA Y MULTIMEDIA

UIDE-BITS

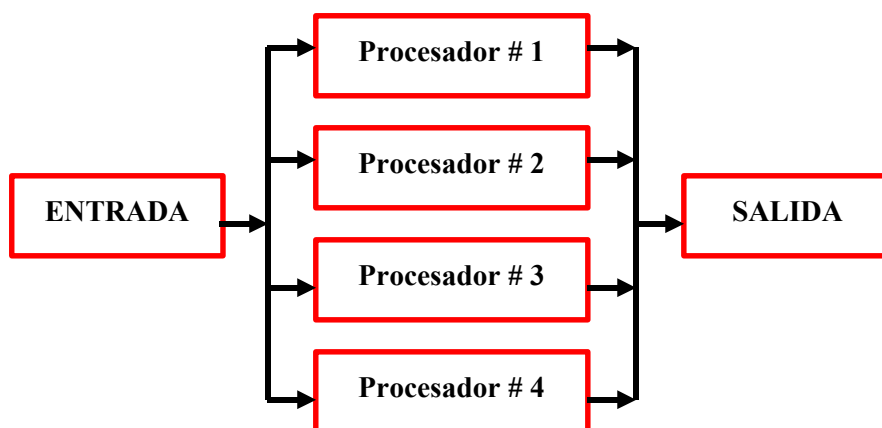
Boletín 5

PROCESAMIENTO PARALELO

La idea básica detrás del procesamiento paralelo es que varios dispositivos (procesadores), ejecutando simultánea y coordinadamente las tareas, pueden rendir más que un único dispositivo. El problema fundamental son las innovaciones tecnológicas que se requieren para obtener ese rendimiento mejorado.

REFLEXIÓN

Las Tecnologías de Información y comunicación no son lo más importante del mundo, pero sí pueden ayudarnos a realizar cosas importantes.



Si bien el procesamiento paralelo ofrece una ventaja definitiva en cuanto a costos, su principal beneficio, la **escalabilidad** (capacidad de crecimiento), puede ser difícil de alcanzar. Esto se debe a que conforme se añaden procesadores, las disputas por los recursos compartidos se intensifican.

Algunas alternativas de arquitecturas de procesamiento paralelo enfrentan este problema fundamental, con diferentes resultados, entre las que se puede mencionar:

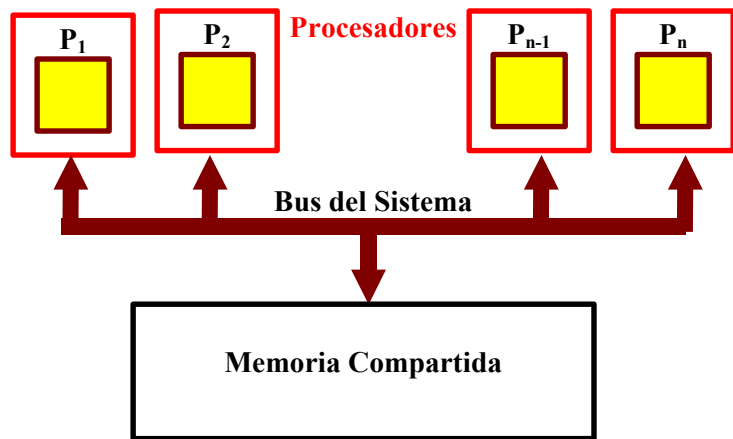
- Multiprocesamiento simétrico
- Procesamiento masivamente paralelo
- Procesamiento paralelo escalable

Cada diseño tiene sus propias ventajas y desventajas.

MULTIPROCESAMIENTO SIMÉTRICO:

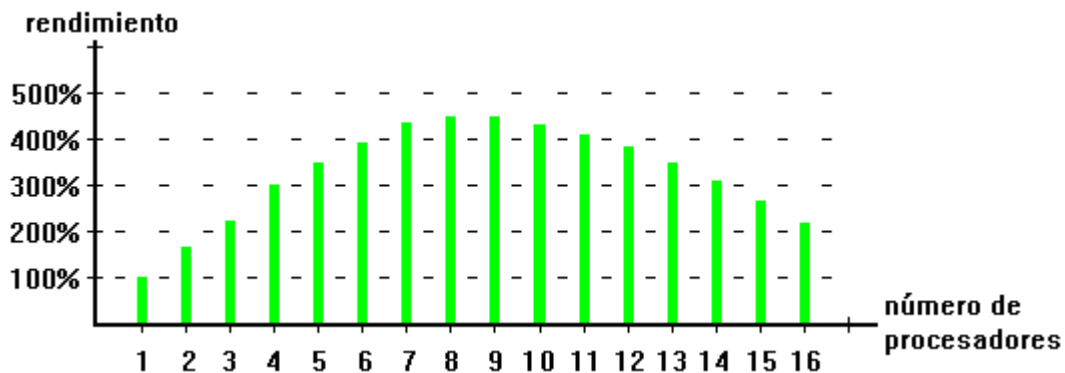
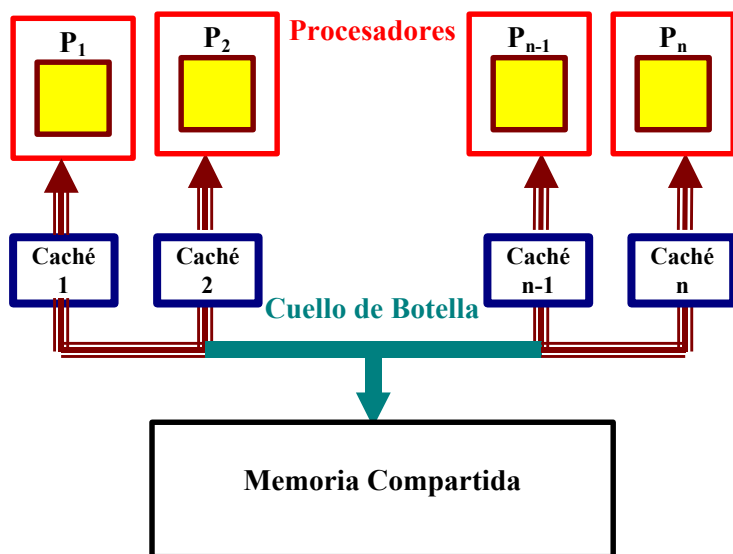
El **Multiprocesamiento Simétrico** (Symmetric Multiprocessing / SMP) tiene un diseño simple, efectivo y económico. En **SMP**, muchos procesadores comparten la misma memoria RAM y el bus del sistema.

La presencia de un solo espacio de memoria simplifica tanto el diseño del sistema físico (hardware) como la programación de las aplicaciones (software). Esa memoria compartida permite que un **Sistema Operativo con Multiconexión** distribuya las tareas entre varios procesadores, o que una aplicación obtenga toda la memoria que necesita para una simulación compleja. La memoria globalmente compartida también vuelve fácil la sincronización de los datos.



SMP es uno de los diseños de procesamiento paralelo más maduros. Sin embargo, la memoria global contribuye al problema más grande de SMP: conforme se añaden procesadores, el tráfico en el bus de memoria se satura. Al añadir memoria caché a cada procesador se puede reducir algo del tráfico en el bus.

Al manejarse ocho o más procesadores, el cuello de botella se vuelve crítico, inclusive para los mejores diseños, por lo que SMP es considerada una tecnología poco escalable.

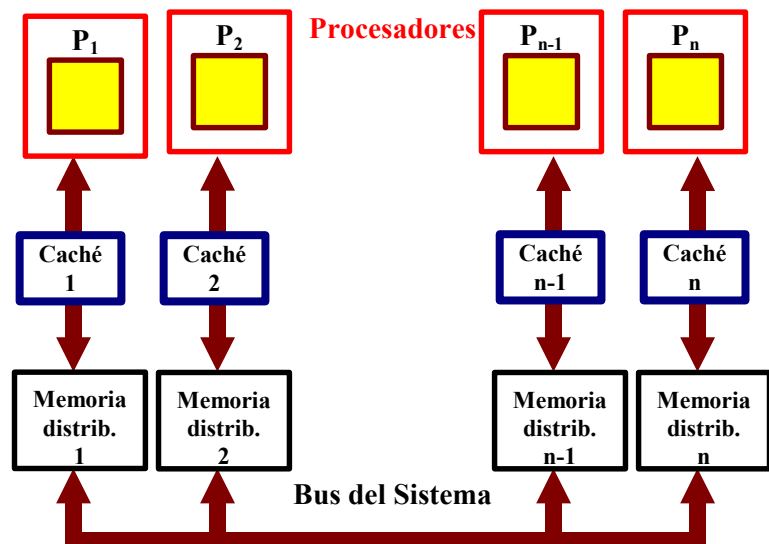


Si bien los primeros componentes utilizados con la tecnología SMP fueron procesadores RISC, en la actualidad, debido a su bajo costo, los procesadores CISC avanzados como Pentium III y Pentium IV son empleados con mayor frecuencia.

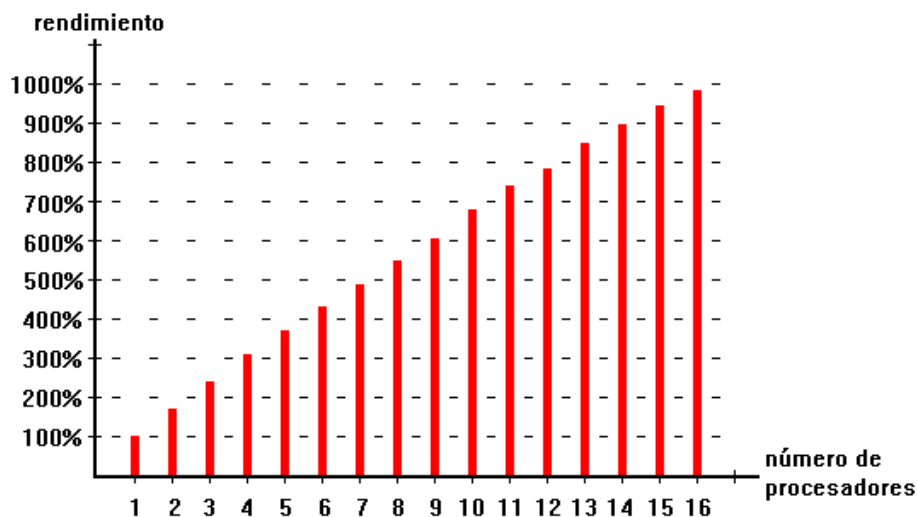
En el mercado se encuentran fácilmente sistemas Pentium III y Pentium IV de dos, cuatro y ocho vías (two-way, four-way, eight-way). Dos, cuatro y ocho vías significan dos, cuatro y ocho procesadores conectados en paralelo, respectivamente.

PROCESAMIENTO MASIVAMENTE PARALELO:

El **Procesamiento Masivamente Paralelo** (Massively Parallel Processing / MPP) es una arquitectura computacional de alto rendimiento. Para evitar los cuellos de botella en el bus de memoria, MPP no utiliza memoria compartida; en su lugar, distribuye equitativamente la memoria RAM entre los procesadores de modo que se asemeja a una red (cada procesador con su memoria distribuida asociada es similar a un computador dentro de una red de procesamiento distribuido).



Para tener acceso a las áreas de memoria fuera de su propia RAM (memoria libre no empleada por los otros procesadores), los procesadores utilizan un esquema de **paso de mensajes** análogo a los **paquetes de datos** en redes. Este sistema reduce el tráfico del bus, debido a que cada sección de memoria interactúa únicamente con aquellos accesos que le están destinados, en lugar de interactuar con todos los accesos a memoria, como ocurre en un sistema SMP. Esto permite la construcción de sistemas MPP de gran tamaño, con cientos y aún miles de procesadores, por lo que MPP es una tecnología altamente escalable.

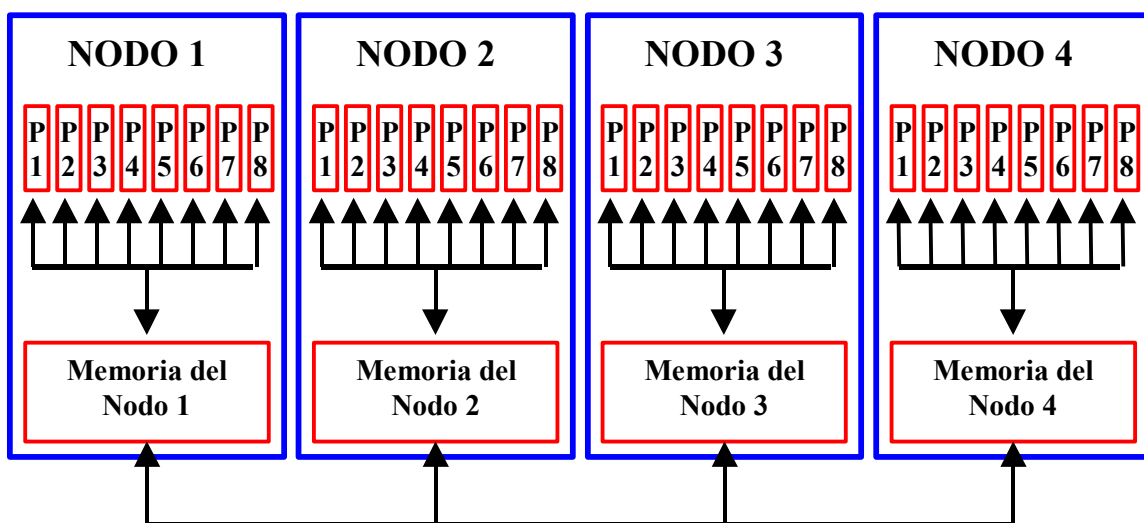


La parte negativa de MPP —desde el punto de vista tecnológico— es que la programación se vuelve difícil, debido a que la memoria se rompe en pequeños espacios separados. Sin la existencia de un espacio de memoria globalmente compartido, ejecutar una aplicación que requiere una gran cantidad de RAM (comparada con la memoria local), puede ser difícil. La sincronización de datos entre tareas ampliamente distribuidas también se complica, particularmente si un mensaje debe pasar por muchos componentes de hardware hasta alcanzar la memoria del procesador destino.

Escribir una aplicación MPP también requiere estar al tanto de la organización de la memoria manejada por el programa. Donde sea necesario, se deben insertar comandos de paso de mensajes dentro del código del programa. Además de complicar el diseño del software, tales comandos pueden crear dependencias de hardware en las aplicaciones (sólo funcionaría en equipos de cierta marca y modelo). Sin embargo, la mayor parte de vendedores de computadores han salvaguardado la portabilidad de las aplicaciones adoptando, sea un mecanismo de dominio público para paso de mensajes conocido como **Máquina Virtual Paralela** (Parallel Virtual Machine / PVM), o un estándar llamado **Interfaz de Paso de Mensajes** (Message Passing Interface / MPI).

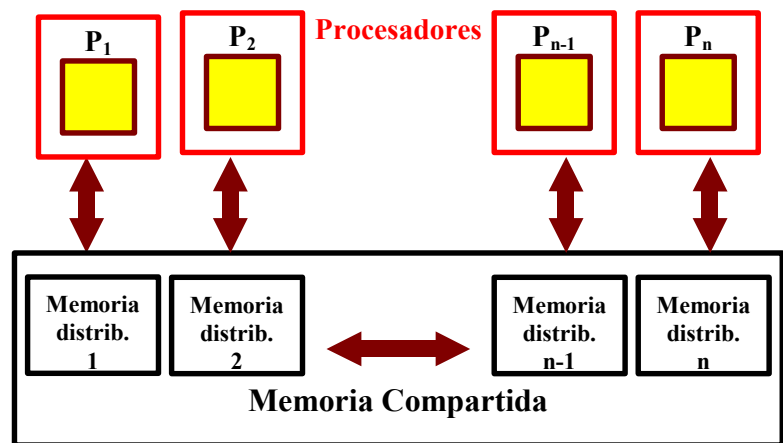
El costo de las soluciones basadas en MPP es mucho más alto que el costo por procesador de las soluciones SMP, por lo que su uso sólo se justifica cuando la necesidad de procesamiento es muy alta.

El **RS/6000 Scalable Powerparallel System** de IBM (SP2) es un ejemplo de sistema MPP. El enfoque dado por IBM para manejar un gran número de procesadores es una pequeña variación —muy ingeniosa— de la tecnología MPP estándar. Los componentes son conectados en grupos de 8 procesadores, llamados nodos. Cada nodo tiene un bloque de memoria compartida (tecnología SMP), pero todos los nodos están conectados utilizando tecnología MPP convencional (cada nodo tiene una memoria distribuida independiente). Lo importante de esta arquitectura es que aprovechan los bajos costos de la tecnología SMP.



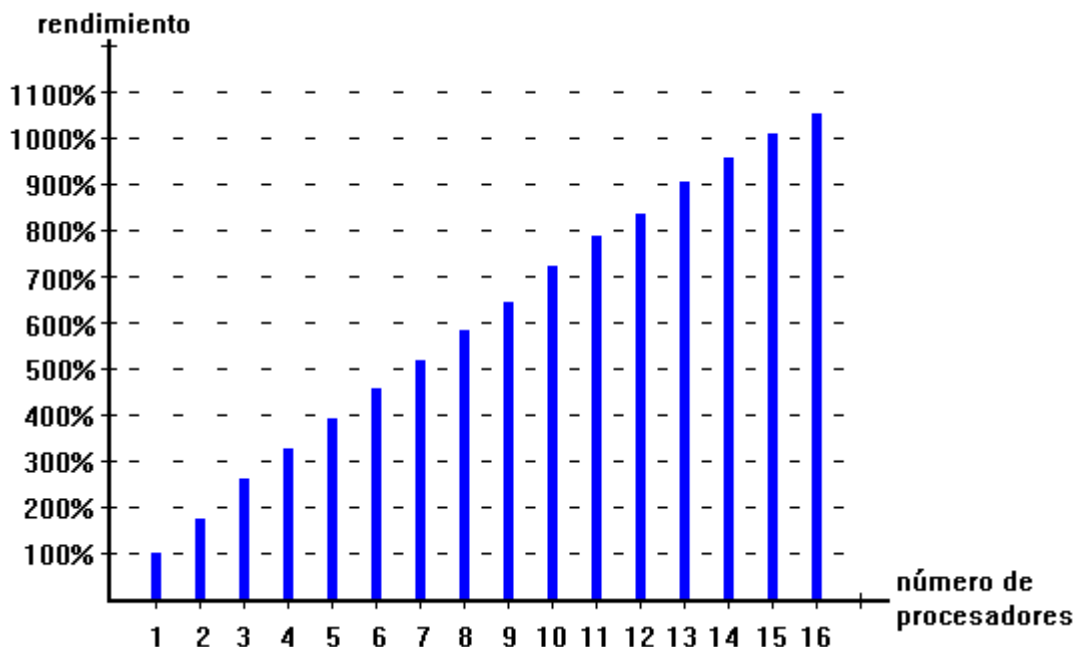
PROCESAMIENTO PARALELO ESCALABLE:

La tercera arquitectura paralela, el **Procesamiento Paralelo Escalable** (Scalable Parallel Processing / SPP), es un híbrido de SMP y MPP, que utiliza una memoria jerárquica de dos niveles para alcanzar la escalabilidad. La primera capa consiste de componentes de memoria distribuida que son esencialmente parte de sistemas MPP completos, con múltiples nodos (nodo = procesador + memoria distribuida), y el segundo nivel de memoria está globalmente compartido al estilo SMP.



Se construyen sistemas SPP grandes interconectando dos o más nodos a través de la segunda capa de memoria, de modo que esta capa aparece, lógicamente, como una extensión de la memoria individual de cada nodo.

La memoria de dos niveles reduce el tráfico de bus debido a que solamente ocurren actualizaciones para mantener coherencia de memoria. Por tanto, SPP ofrece la facilidad de programación del modelo SMP, a la vez que provee una escalabilidad similar a la de un diseño MPP.



Marcelo Romo
8 de Junio de 2003