



Universidad
Nacional de
General
Sarmiento



Sistemas Operativos y redes



Lic. Mariano Vargas

¿Que es HPC?



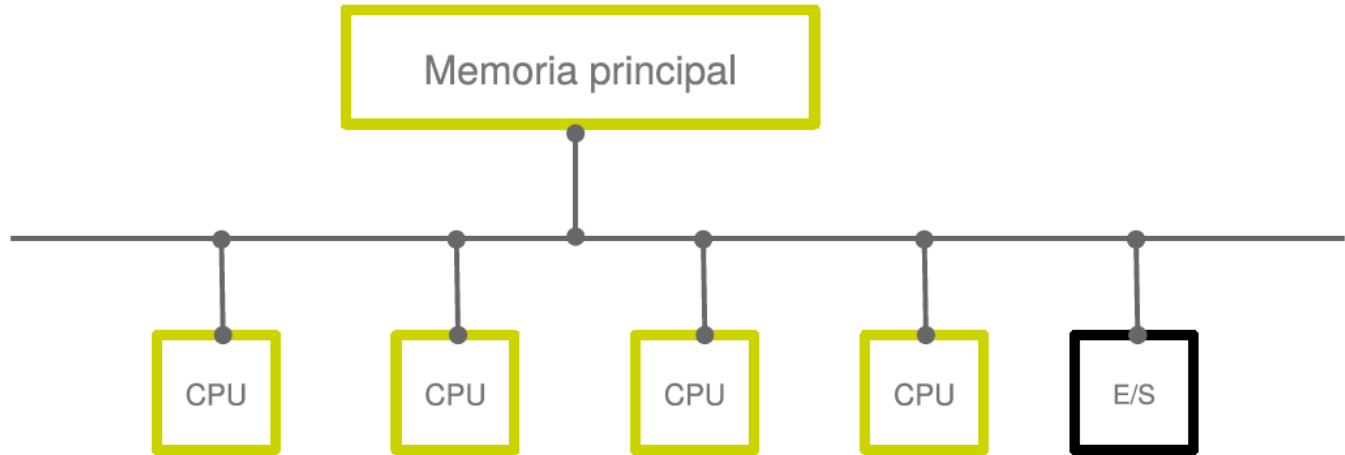
- HPC: High Performance Computing
- Consiste en aumentar poder de cómputo con el fin de lograr resultados mucho más rápido de los que se logra con una computadora.
- Generalmente involucra grandes problemas en ciencia, ingeniería o negocios
- Se puede dividir las estrategias de procesamiento en 3 tipos generales: **SMP, Heterogéneo y pasaje de mensajes.**

Multiprocesamiento simétrico (SMP)



- Varias unidades de procesamiento idénticas comparten un único espacio de memoria.
- Todas las unidades tienen acceso a los dispositivos de entrada y salida (E/S).
- Un sistema operativo común.
- Caso típico: procesadores multi-núcleo.

Multiprocesamiento simétrico (SMP)

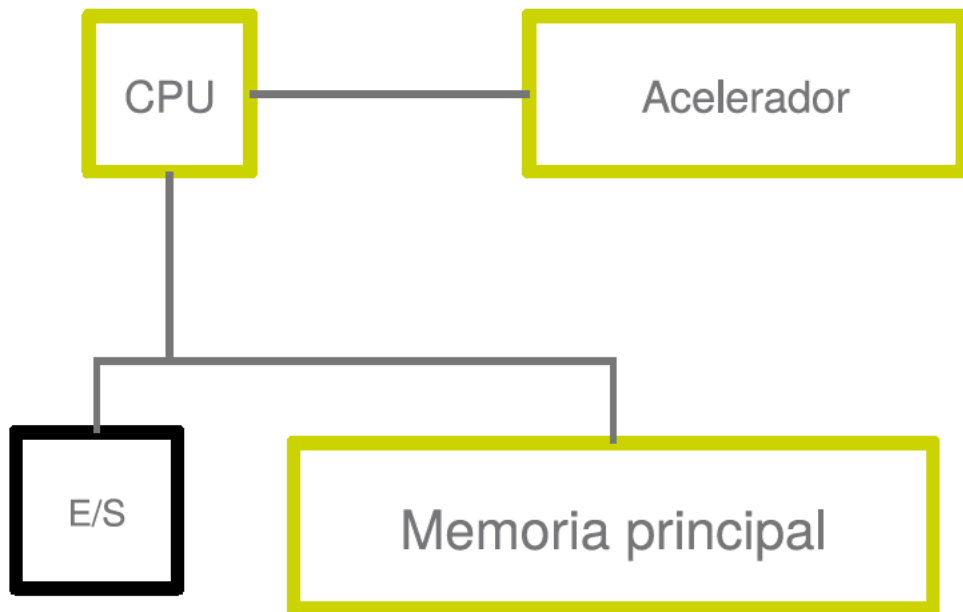


Heterogéneo



- Unidades de propósito general manejan la memoria y la E/S.
- Otras unidades especializadas computan tareas específicas.
- Espacio de memoria compartido y sistema operativo común.
- Caso típico: PC con GPU.

Heterogéneo

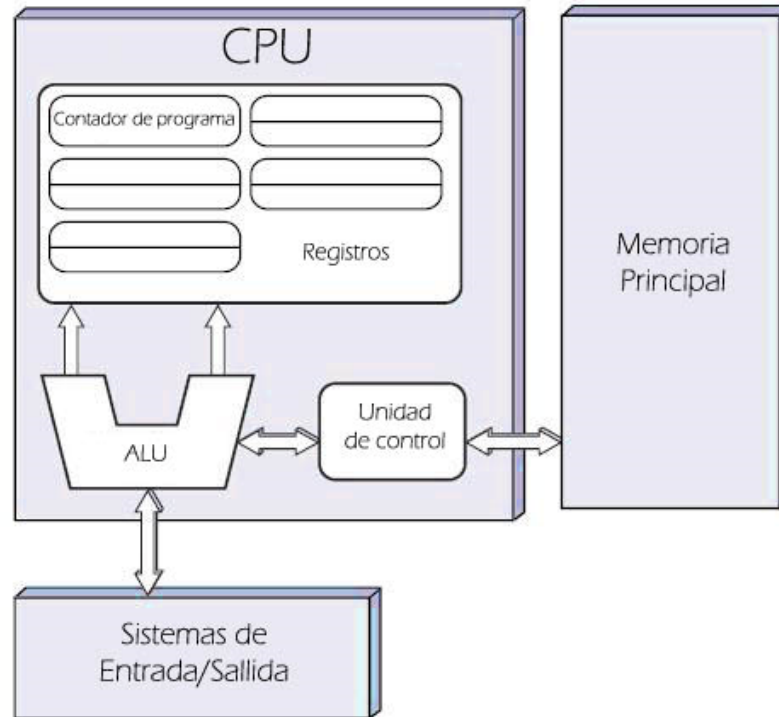


Pasaje de mensajes



- Múltiples nodos se comunican a través de una red por pasaje
- de mensajes.
- Los nodos no comparten memoria.
- No es necesario compartir el sistema operativo.
- Caso típico: Cluster.

Arquitectura de Von Neumann



Arquitectura de Von Neumann

Cuatro componentes principales

- Memoria
- Unidad de Control
- Unidad aritmética lógica
- Entrada/Salida

- I/O: memoria de acceso aleatorio para almacenar datos e instrucciones.
- La unidad de control trae instrucciones y datos de la memoria, decodifica las instrucciones y secuencialmente coordina las operaciones para llevar a cabo la tarea programada.
- La Unidad aritmética lógica realiza las operaciones aritméticas/lógicas básicas
- La entrada/salida es la interfaz con el humano
- Cache:



Elementos mínimos de una computadora paralela

- Hardware

- Múltiples procesadores.
- Múltiples memorias.
- Interconexión.

- Software de Sistema

- Sistema Operativo Paralelo
- Herramientas

- Aplicaciones



Organización



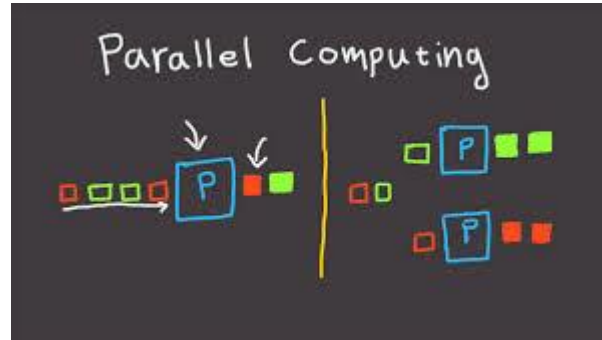
- Organización Lógica
 - Lo que ve el usuario de la maquina a traves del software
- Organización Física
 - El hardware de la maquina
- La arquitectura física es independiente de la arquitectura lógica.
- Es necesario conocer el hardware si queremos que el software corra de forma eficiente.
- La computadora ideal es aquella que nos oculta los detalles físicos.

Taxonomía de Flynn (1966)



- Se separan las instrucciones y los datos en dos dimensiones
- Pueden ser simples o múltiples
- SISD: Single Instruction, Single Data
- SIMD: Single Instruction, Multiple Data
- MISD: Multiple Instruction, Single Data
- MIMD: Multiple Instruction, Multiple Data

Paralelismo



Paralelismo

- Ver mapa conceptual

¿Qué es y para qué sirve un clúster?





¿Qué es y para qué sirve un clúster?

- Clúster: conjunto de computadoras conectadas por un medio de comunicación que cooperan entre sí para resolver un problema en un tiempo mas corto (se puede trabajar en forma concurrente).
- Problemas:
- Simulaciones
- Manejo de imágenes
- Cálculos de la ciencia y la Tecnología y...



¿Qué es y para qué sirve un clúster?

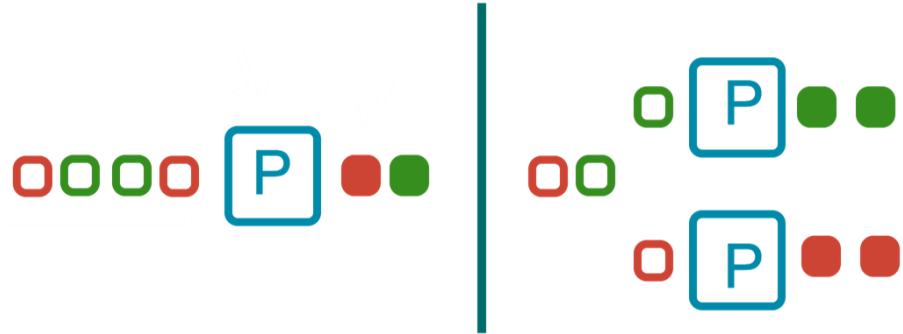
- En un clúster, cada computadora, llamada nodo, ejecuta su propio sistema operativo y software, y se comunica con los otros nodos a través de una red de alta velocidad para coordinar sus actividades.
- El objetivo principal de un clúster es proporcionar una mayor potencia de procesamiento y capacidad de almacenamiento que una sola computadora podría ofrecer, lo que permite procesar grandes cantidades de datos y realizar cálculos complejos en **un tiempo más corto**.

Aprovecha la concurrencia



- Concurrencia:
- Pipeline
- Paralelismo
- Simultaneidad

Parallel computing



Diferencias entre cluster y supercomputadora



- **Arquitectura:** Las supercomputadoras están diseñadas para realizar un gran número de cálculos complejos y para trabajar en aplicaciones que requieren un alto rendimiento de cálculo y/o una gran cantidad de memoria.
- Por lo general, tienen una arquitectura única y personalizada que se adapta a los requisitos específicos de las aplicaciones que ejecutan. Los clusters de computadoras, por otro lado, son sistemas más comunes que están contruidos a partir de computadoras estándar, que se conectan mediante una red para trabajar juntas como un solo sistema.

Diferencias entre cluster y supercomputador



- Escalabilidad, Disponibilidad y Costo.
- Los clusters son sistemas más comunes y escalables contruidos a partir de nodos individuales, mientras que las supercomputadoras son sistemas personalizados con arquitecturas específicas para ejecutar aplicaciones de alta demanda de recursos.

Tipos de clúster de supercomputadoras



- High Availability Clusters (HA)
- Load Balancing Clusters
- High Performance Computing Cluster (HPC)

High performance computing cluster



- Se utiliza para procesamiento intensivo de datos y cálculos complejos. Está diseñado para mejorar el rendimiento y reducir el tiempo de procesamiento, utilizando varios nodos que trabajan juntos en paralelo para realizar una tarea.

Arquitectura de clúster



- Nodos, nodos maestros y almacenamiento compartido
- Comunicación y coordinación entre nodos
- Protocolos de comunicación y gestión



Ventajas de los clústers

- **Mayor capacidad de procesamiento:** Los clústers de computadoras combinan múltiples nodos para proporcionar una capacidad de procesamiento mucho mayor que la de una sola computadora.
- **Mayor disponibilidad:** En un clúster de computadoras, si un nodo falla, otros nodos pueden tomar su lugar para continuar el procesamiento. Esto significa que los servicios pueden seguir funcionando incluso si hay problemas con uno de los nodos.
- **Escalabilidad:** Un clúster de computadoras puede escalar fácilmente añadiendo más nodos, lo que significa que puede manejar mayores cargas de trabajo.
- **Mayor tolerancia a fallos:** Con un clúster de computadoras, si un nodo falla, otros nodos pueden tomar su lugar para garantizar que el servicio siga funcionando. Esto significa que un clúster de computadoras es más tolerante a fallos que una sola computadora.

Desventajas



- **Costo:** La configuración de un clúster de computadoras puede ser costosa debido a la necesidad de múltiples nodos y servidores.
- **Configuración y mantenimiento:** La configuración y el mantenimiento de un clúster de computadoras pueden ser complicados debido a la necesidad de configurar y mantener múltiples nodos y servidores.
- **Requerimientos de red:** Un clúster de computadoras requiere una buena infraestructura de red para funcionar de manera eficiente.
- **Complejidad de programación:** La programación en un clúster de computadoras puede ser más complicada que la programación en una sola computadora, debido a la necesidad de coordinar el procesamiento entre múltiples nodos.

En resumen



En general, los clústers de computadoras son una herramienta poderosa para proporcionar una mayor capacidad de procesamiento y una mayor disponibilidad de servicios, pero requieren una inversión en recursos (humanos también) y esfuerzos de configuración y mantenimiento.



Ejemplos

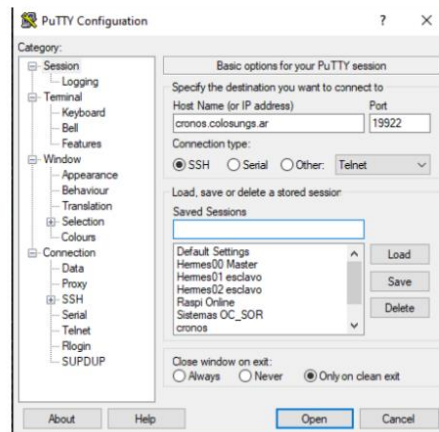
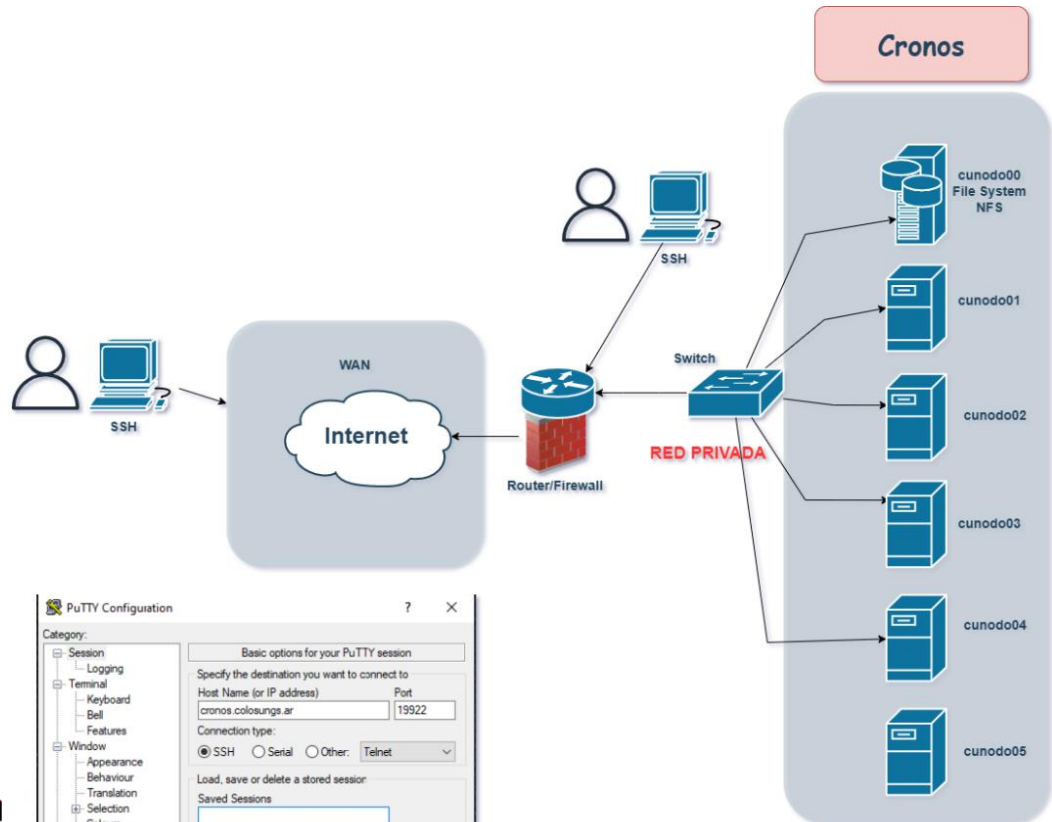


Google: Google utiliza clusters de computadoras para su motor de búsqueda y otros servicios, como Google Drive. Utilizan miles de servidores que trabajan juntos para procesar consultas y proporcionar resultados de búsqueda en tiempo real.

Clúster de Netflix: Netflix utiliza clusters de computadoras para almacenar y procesar grandes cantidades de datos de los usuarios, como las selecciones de películas y programas de televisión. Estos clusters están diseñados para proporcionar una experiencia de usuario personalizada y eficiente.



Arquitectura de Cronos



Arquitectura de Cronos



**Universidad
Nacional de
General
Sarmiento**



Hardware cronos: Single Board Computer



Raspberry Pi 4

Software

- Sistema Operativo Raspbian Lite 64 bits
- Munge
- Slurm
- OpenMP



iFIN!

