

Aspectos Generales de los Sistemas Distribuidos

Diego Alberto Rincón Yáñez MSc darincon@ucatolica.edu.co

Sistemas Distribuidos

Departamento de Ingeniería de Sistemas y Computación

Afiliada a la Federación Internacional de Universidades Católicas (FIUC) www.ucatolicas.edu.co



Contenido

- Definición
- Comparación Sistemas Distribuidos/Centralizado
- Conceptos de Hardware
- Propiedades (Conceptos de Diseño)
- Ventajas
- Desventajas











Comparación

Sistemas distribuidos con respecto a los centralizados/servidores

- Economía: Relación precio/rendimiento.
- Desempeño: Se pueden unir gran cantidad de CPU y tener una capacidad de computo imposible de obtener en un mainframe.
- Distribución inherente: Uso de máquinas separadas en distancia.











Comparación

Sistemas distribuidos con respecto a los centralizados/servidores

- Confiabilidad: Si una máquina se descompone, el sistema puede sobrevivir como un todo.
- Disponibilidad: Distribución de la carga.
- Crecimiento por incrementos: Se puede añadir poder de cómputo en pequeños incrementos.









Comparación



Sistemas distribuidos con respecto a los centralizados/servidores

- Complejidad:
 - Nivel de Abstracción
 - Los algoritmos son totalmente distintos a la forma natural de proponer soluciones.
- Redes: Dependencia directa de la comunicación.
- Seguridad: requerimientos de diseño versus sistemas seguros.









Conceptos de Hardwar UNIVERSIDAD CATÓLICA de Colombia

- Aunque todos los sistemas distribuidos constan de varias CPU, existen diversas formas de organizar el hardware.
- Taxonomía más citada Flynn(1972).
- Basado en los números de flujos de instrucciones y de datos.









Conceptos de Hardware



- SISD(Single Instruction, Single Data): Un flujo de instrucciones y un flujo de datos.
 - Ejemplo: Todas las computadoras tradicionales de un procesador









Conceptos de Hardware



- SIMD(Single Instruction, Multiple Data): Un flujo de instrucciones y varios flujos de datos.
 - Ejemplo: Sumando todos los elementos de 64 vectores independientes.
 - Útil: Cómputos que repiten los mismos cálculos en varios conjuntos de datos









Conceptos de Hardware



- MISD(Multiple Instruction, Simple Data):
 Múltiples flujos de instrucciones y un flujo de datos.
 - Ejemplo: Ninguna de los computadores conocidas se ajusta a este modelo.









Conceptos de Hardwar UNIVERSIDAD CATÓLICA de Colombia

- MIMD (Multiple Instruction, Multiple Data): Varios flujos de instrucciones y varios flujos de datos.
 - Computadoras independientes interconectadas, todos los sistemas distribuidos hacen parte de esta categoría.



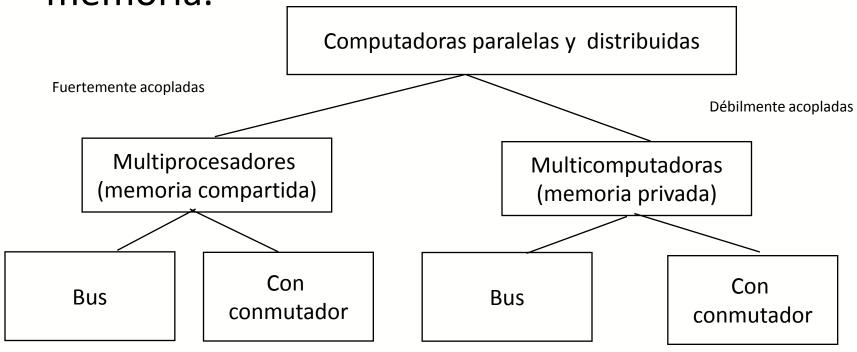






Conceptos de Hardwa Luniversidad CATÓLICA de Colombia

 MIMD se dividen según compartan o no la memoria.











Conceptos de Hardwar UNIVERSIDAD CATÓLICA de Colombia

 Multiprocesadores (memoria compartida):
 Espacio de direcciones virtuales, compartido por todos los CPU

Multicomputadoras (memoria privada):
 Cada máquina tiene su propia memoria









Conceptos de Hardwar UNIVERSIDAD CATÓLICA de Colombia

Multicomputadoras con conmutador

- Se han construido varias redes de interconexión, pero cada CPU tiene acceso a su propia memoria particular.
- Topologías populares:
 - Retícula: Basados en las tarjetas de circuitos impresos.
 - Hipercubo: Cubo n-dimensional









Conceptos de Softwar UNIVERSIDAD CATÓLICA de Colombia

 La forma de pensar de los usuarios de un sistema queda determinada por el software del sistema operativo.

• Tipos:

- Débilmente acoplados: Maquinas y usuarios sean independientes entre sí en lo fundamental, pero interactúen en cierto grado cuando sea necesario.
- Fuertemente acoplados: Multiprocesador dedicado a la ejecución del ajedrez en paralelo. Cada CPU evalúa un tablero de ajedrez y reporta resultados.









Conceptos de Softwar

Sistemas operativos de redes

- Software débilmente acoplado y hardware débilmente acoplado.
- Red de estaciones de trabajo interconectados por una LAN (rlogin, rcp).
- Cada máquina tiene gran autonomía y existen pocos requisitos a lo largo del sistema (sistema operativo de red)









Conceptos de Software de Colombia

Sistemas realmente distribuidos

- Software fuertemente acoplado y hardware débilmente acoplado.
- Crear la ilusión que toda la red de computadores es un sistema de tiempo compartido.
- Características:
 - Mecanismo de comunicación global entre procesos.
 - Administración la misma en todas partes.
 - Sistema de archivos debe tener la misma apariencia en todas partes.









Conceptos de Softwar UNIVERSIDAD CATÓLICA de Colombia

Sistemas de multiprocesador con tiempo compartido

- Software y hardware fuertemente acoplados.
- Ejemplos: Multiprocesadores que operan con un sistema de tiempo compartido (linux)
- Característica:
 - Existencia de una cola de ejecución (Estructura de datos contenida en la memoria compartida)











Transparencia

- Imagen de un sistema.
- Transparencia de localización: Los usuarios no pueden indicar la localización de los recursos.
- Transparencia de migración: Los recursos se pueden mover a voluntad sin cambiar nombres.
- Transparencia de réplica: Los usuarios no pueden compartir recursos de manera automática.









Aspectos de diseño

Transparencia

- Transparencia de concurrencia: Varios usuarios pueden compartir recursos de manera automática.
- Transparencia de paralelismo: Las actividades pueden ocurrir en paralelo sin el conocimiento de los usuarios.









Aspectos de diseño

Flexibilidad

- En cuanto a la estructura de los sistema distribuidos hay dos clases:
 - Cada máquina debe tener un núcleo tradicional que provea la mayoría de los servicios.
 - Núcleo que provea lo menos posible y que el grueso de los servicios del S.O. se obtengan a partir de los servidores al nivel usuario.









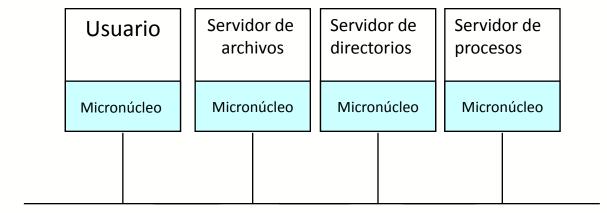


UNIVERSIDAD CATÓLICA de Colombia

Flexibilidad

Núcleo monolítico

Núcleo monolítico



Micronúcleo











Flexibilidad

- El micronúcleo es más flexible ya que casi no hace nada:
- Proporciona:
 - Un mecanismo de comunicación entre procesos.
 - Cierta administración de memoria.
 - Una cantidad limitada de planificación y administración de procesos de baja nivel.
 - Entrada/salida de bajo nivel.









Aspectos de diseño

Confiabilidad

- Disponibilidad: Fracción de tiempo en que se puede utilizar el sistema.
- Seguridad: Los archivos y otros recursos deben ser protegidos contra el uso no autorizado.
- Tolerancia a fallas: Cuando un servidor falla y vuelve a arrancar de manera súbita. ¿Qué ocurre?









Aspectos de diseño

Desempeño

- Cuando se ejecuta una aplicación en un sistema distribuido no debe parecer peor que su ejecución en un procesador.
- Diferentes métricas:
 - Tiempo de respuesta
 - Número de trabajos por hora (rendimiento)
 - Uso del sistema











Escalabilidad

- Los sistemas distribuidos están diseñados para cientos de usuarios pero cuando son millones?
- Cuellos de botella potenciales:
 - Componentes centralizados: Un solo servidor de correo para todos los usuarios.
 - Tablas centralizadas: Un directorio telefónico en línea.
 - Algoritmos centralizados: Realización de un ruteo con base en la información completa.









Arquitecturas de Sistemas Distribuidos

- Prácticamente todos los grandes sistemas informáticos son sistemas distribuidos.
- En un sistema distribuido el procesamiento de información se distribuye sobre varias computadoras en vez de estar confinado en una única máquina.









Arquitecturas de Sistemas Distribuidos

- El desafío es diseñar software y hardware para proporcionar características deseables a los sistemas distribuidos y minimizar los problemas propios de ellos.
- Para eso, debemos comprender las ventajas y desventajas de las diferentes arquitecturas de sistemas distribuidos.











1. **Compartir recursos.** Un sistema distribuido permite compartir recursos hardware y software (discos, impresoras, ficheros y compiladores) que se asocian con computadoras de una red.











2. **Apertura.** Son normalmente sistemas abiertos: se diseñan sobre protocolos estándares que permiten combinar equipamiento y software de diferentes vendedores.











3. **Concurrencia.** Varios procesos pueden operar al mismo tiempo sobre diferentes computadoras de la red. Hasta pueden comunicarse con otros durante su funcionamiento.











4. **Escalabilidad.** Los sistemas distribuidos son escalables mientras la capacidad del sistema pueda incrementarse, añadiendo nuevos recursos para cubrir nuevas demandas sobre el sistema.

En la práctica, si se añaden muchas computadoras nuevas, la capacidad de la red puede saturarse.











- 5. **Tolerancia a defectos.** Contar con varias computadoras y el potencial para reproducir información significa que los sistemas distribuidos pueden ser tolerantes a algunas fallas de funcionamiento del hardware y del software.
- •En la mayoría de los sistemas distribuidos, puede haber un servicio degradado, ante fallas de funcionamiento. Una completa pérdida de servicio sólo ocurre cuando existe una falla de funcionamiento en la red.











- 1. **Complejidad.** Los sistemas distribuidos son más complejos que los sistemas centralizados; lo que hace más difícil comprender sus propiedades emergentes y probar estos sistemas.
- Por ejemplo, en vez de que el rendimiento del sistema dependa de la velocidad de ejecución de un procesador, depende del ancho de banda y de la velocidad de los procesadores de la red.
- Mover los recursos de una parte del sistema a otra puede afectar de forma radical al rendimiento del sistema.











- 2. **Seguridad.** Puede accederse al sistema desde varias computadoras diferentes, y el tráfico en la red puede estar sujeto a escuchas indeseadas.
- Es más difícil mantener la integridad de los datos en el sistema y que los servicios del sistema no se degraden por ataques.









- 3. *Manejabilidad*. Las computadoras en un sistema pueden ser de diferentes tipos y ejecutar versiones diferentes de sistemas operativos.
- Los defectos en una máquina pueden propagarse a otras, con consecuencias inesperadas.
- Esto significa que se requiere más esfuerzo para gestionar y mantener el funcionamiento del sistema.











- 4. *Impredecibilidad*. Los sistemas distribuidos tienen una respuesta impredecible.
- •La respuesta depende de la carga total en el sistema, de su organización y de la carga de la red.
- Como todos ellos pueden cambiar rápidamente, el tiempo requerido para responder a una petición de usuario puede variar drásticamente, de una petición a otra.











Modos de Transmisión

MODOS DE TRANSMISIÓN

Una transmisión de datos tiene que ser controlada por medio del tiempo, para que el equipo receptor conozca en que momento se puede esperar que una transferencia tenga lugar. Hay dos principios de transmisión para hacer esto posible:

Transmisión Sincronía.
 Transmisión Asíncrona.







