

Sistema distribuido

Conjunto de nodos autónomos que trabajando de manera solidaria se comporta como un único sistema coherente para los usuarios.

La idea es compartir servicios y recursos.

Ventajas:

- Económicas: más barato muchos ordenadores que un supercomputador
- Fiabilidad: El sistema funciona aunque algún nodo falle y permiten la capacidad de aplicar cambios en caliente sin que el sistema caiga.
- Se pueden añadir recursos en demanda
- Distribución inherente de aplicaciones

Objetivo: Transparencia

Esconder el hecho de que el sistema está formado por distintos componentes

Tipos de transparencia:

- Transparencia de acceso: Esconder detalles de la representación de los datos y sus mecanismos de acceso.
- Transparencia de localización; Esconder la localización de los recursos mediante la asignación de nombres lógicos.
- Transparencia de migración: El usuario no se entera de si el recurso se mueve de un componente a otro.
- Transparencia de replicación: Se oculta si hay replicas del recurso (requiere transparencia de localización)
- Transparencia de concurrencia: Oculta que componentes del sistema necesitan compartir objetos y cooperar para proporcionar el recurso conservando un estado coherente
- Transparencia de fallos: El usuario no se entera si ciertos componentes del sistema fallan.

Grado de Transparencia

- Compromiso entre transparencia y eficiencia
- Restricciones físicas (latencia)
- No siempre conviene ofrecer un tipo de transparencia

Objetivo: sistema abierto

Los recursos se sirven de una forma estándar, siguiendo una semántica para proporcionar interoperabilidad, portabilidad y extensibilidad.

Esto en la vida real no ocurre.

Objetivo: Escalabilidad

El sistema debe soportar su crecimiento, su distribución geográfica y su administración.

- Sistemas centralizados tienen cuellos de botella
- Sistemas centralizados están lejos de algunos clientes
- Servicios centralizados son más sencillos de proteger y administrar

Escalaabilidad: Datos

→ Distribución de los datos en dominios (DNS)

Escalaabilidad: Algoritmos

- Algoritmo centralizado: los nodos tienen que recolectar todos los datos sobre el estado del sistema para ejecutar el algoritmo. (Puede no ser asumible)
- Algoritmo distribuido:
 - Ningún nodo tiene la información completa
 - Los nodos toman decisiones en base a su estado
 - El fallo de un nodo no arruina el algoritmo
 - No se asume un reloj común exactamente sincronizado.

Escalabilidad : Geografica

- Comunicación asíncrona
- Agrupación de operaciones (batching)
- Protocolos con pocos RTTs (round-trips)
- Preprocesado en el cliente
- Caching → coherencia de la cache

Tipos de sistemas distribuidos

→ Transacciones ACID:

- Atomicity: Operación indivisible
- Consistency: Las operaciones no violan invariantes del sistema.
- Isolation: Las operaciones no interfieren entre sí
- Durability: una vez acabada es permanente

→ Clusters:

- El control lo tiene un nodo llamado Master
- Nodos homogéneos (HW, OS)
- Comunicados por LAN

→ Grids

- Control distribuido por dominios
- Organizaciones que comparten nodos
- Nodos heterogéneos
- Comunicados por internet

