
TEMA 1

Conceptos básicos de Sistemas Distribuidos



Lección 1

Introducción a los Sistemas Distribuidos



Definiciones

“Un sistema distribuido es una colección de computadores independientes que aparecen ante el usuario del sistema como un único computador”. Tanenbaum

Un sistema distribuido para nosotros:

- Máquinas independientes,**
- Una red de comunicaciones que las une**
- Comparten algunos recursos, pero sin que el usuario deje de percibir que el sistema está formado por máquinas independientes.**



Ventajas S. D. vs. Mainframe

- **Economía.**- ¿Poder computacional de una CPU proporcional al precio?. Antes sí, ahora más precio → más velocidad procesador, pero mismo chip.
- **Potencia.**- Existe un límite en la capacidad de un único procesador.
- **Escalabilidad.**- Si un *mainframe* llega al límite de su capacidad es necesario adquirir otro. En un S. D. la incorporación de nuevas máquinas es menos traumático.
- **Naturaleza de las organizaciones y del trabajo.**- Muchas aplicaciones son intrínsecamente distribuidas: sucursales, oficinas, departamentos, grupos de trabajo, etc.
- **Fiabilidad.**- Si falla el *mainframe*, falla el sistema. Si falla una máquina, el S.D. puede continuar.



Desventajas S. D. vs. Mainframe

- **Software.**- No hay mucho software para sistemas distribuidos y no hay todo tipo de software para todos los problemas.
- **Comunicaciones.**- Se depende de algún tipo de red para el funcionamiento de los S.D. Deben ser fiables. Si se satura la red y es necesario un cambio de tecnología → altos costes.
- **Accesibilidad.**- La alta accesibilidad de los datos puede ser un problema. Problemas de seguridad y vulnerabilidad ante ataques tanto internos como externos. ¿Mejor en *mainframes*?
- **Administración.**- Los sistemas centralizados son mucho más fáciles de administrar: usuarios, programas, seguridad, copias de seguridad, etc.



Sistemas Distribuidos. Retos.

Un sistema distribuido debería presentar las siguientes características:

- Transparencia**
- Heterogeneidad**
- Fiabilidad**
- Escalabilidad**
- Sincronizado**
- Seguridad**



Transparencia

Mide la aproximación de un sistema distribuido a un sistema centralizado. Si un S.D. es indistinguible de un sistema centralizado, es transparente.

Tipos de transparencia:

- **Acceso.**- Al ejecutar un programa, no se sabe que involucra diversas máquinas.
- **Localización.**- No se sabe dónde están los recursos que se están usando.
- **Replicación.**- El sistema hace copias de los recursos sin conocimiento del usuario.
- **Concurrencia.**- No se tienen que percibir otros usuarios en el sistema.
- **Paralelismo.**- El reparto de la carga se hace automáticamente entre los procesadores disponibles.
- **Migración.**- Se pueden mover los recursos de un sitio a otro sin afectar a las aplicaciones.
- **Fallos.**- No se perciben los posibles fallos del sistema.



Heterogeneidad

La heterogeneidad de un sistema distribuido viene dada por la diversidad de elementos que lo componen.

Fuentes de heterogeneidad:

- **Redes.**- Cable, inalámbrica, fibra óptica, teléfono, etc. *Diferentes protocolos de transporte.*
- **Hardware computadores.**- PCs, Mac, Workstations, PDA, teléfonos móviles, etc. *Los datos no se representan de la misma manera.*
- **Sistemas Operativos.**- Windows, Linux, Un*x, PalmOS, etc. *Diferentes APIs de programación.*
- **Lenguajes de programación.**- C, C++, C#, Java, Python, etc. *Se necesita middleware.*
- **Programadores.**- *Son necesarios protocolos comunes para que los programas cooperen.*



Fiabilidad

Una de las ventajas de los S.D. es la **fiabilidad**: si tenemos n máquinas y una falla, las $n-1$ máquinas restantes seguirán con el trabajo.

Medida de Fiabilidad:

Disponibilidad.- Fracción de tiempo que el sistema es usable.

¿Cómo incrementar la disponibilidad?

- **Diseño**. No usar múltiples máquinas para completar un trabajo.
- **Utilizar la redundancia**. Si se replica parte del hardware o del software, si se produce un fallo, las copias se encargan de continuar el trabajo.

Problema: Cuando se incrementa la redundancia aparecen problemas de **inconsistencia**. Hay que garantizar que las copias que circulan por el sistema son correctas y están actualizadas.



Escalabilidad

La **escalabilidad** es la capacidad de crecimiento del sistema. Un sistema distribuido debería ser capaz de crecer *sin límites*, ya que se pueden suministrar recursos ilimitados.

La escalabilidad se puede conseguir:

- **Evitando (previando) limitaciones software.**- P. ej., usando cantidades de un determinado número de bits, como las direcciones IP.
- **Evitando recursos centralizados.**- Ej. un sistema de nombres centralizado.
- **Evitando algoritmos no escalables.**- P. ej., búsquedas.
- **Evitar algoritmos centralizados.**- P. ej., planificación de rutas. No se tiene que necesitar información completa sobre el estado del sistema.
- **Ignorando la existencia de un reloj común.**

