

Sistemas Distribuidos

Félix García Carballera

Objetivos del curso

- Presentar una visión global del estado del arte y los aspectos más novedosos del diseño y construcción de sistemas distribuidos.
- Desarrollar ejemplos prácticos que permitan consolidar los conceptos teóricos.

Contenido

- **Introducción**
 - Características de los sistemas distribuidos. Modelos
 - Objetivos y problemas de diseño
 - Ejemplos de sistemas distribuidos
- **Redes e interconexión**
 - Conceptos y propiedades de los sistemas de comunicación
 - Tipos de dispositivos de comunicación y tendencias futuras
 - Protocolos de comunicación. Rendimiento y fiabilidad
- **Comunicación en sistemas distribuidos**
 - Mecanismos básicos de comunicación entre procesos
 - Modelo cliente/servidor y comunicación en grupos
 - Colas de Mensajes POSIX. Sockets en Unix y Java
 - Llamadas a procedimientos remotos (RPC)
 - Entornos orientados a objetos. CORBA, RMI
 - Comunicación en aplicaciones paralelas. MPI

Contenido

- **Sistemas operativos distribuidos**
 - Sistemas operativos en red y distribuidos
 - Objetivos y aspectos de diseño
 - Asignación de procesadores y algoritmos de reparto de carga
 - Middlewares
 - Ejemplos
- **Sistemas de ficheros distribuidos**
 - Conceptos básicos y estructura
 - Servicio de directorio
 - Servicio de ficheros
 - Implementación. Semántica de coutilización, métodos de acceso, cache, coherencia de cache
 - Incremento de prestaciones. Paralelismo
 - Ejemplos

Contenido

- **Memoria global distribuida**
 - Conceptos básicos
 - Modelos de consistencia
 - Aspectos de implementación
 - Ejemplos
- **Sincronización y coordinación distribuida**
 - Mecanismos de sincronización entre procesos
 - Modelos de sistemas distribuidos
 - Relojes lógicos, relojes vectoriales
 - Entrega causal. Estados globales consistentes
 - Sincronización de relojes físicos
 - Exclusión mutua distribuida y algoritmos de elección

Contenido

- **Fiabilidad en sistemas distribuidos**
 - Distintos aspectos de la fiabilidad
 - Replicación
 - Protocolos de consenso
 - Comunicación en grupos
 - Compromiso distribuido
 - Transacciones
- **Seguridad**
 - Requisitos de seguridad
 - Métodos para conseguir seguridad
 - Cifrado. Firmas digitales
 - Autenticación
 - Kerberos
 - Cortafuegos

Contenido

- **Diseño de aplicaciones distribuidas**
 - Elementos de una aplicación distribuida
 - Diseño de aplicaciones cliente/servidor
 - Uso de WWW en el diseño de aplicaciones distribuidas
 - Herramientas para construir aplicaciones distribuidas
 - Web, Java, CORBA, JDBC, ...
- **Trabajos prácticos**
 - Desarrollo y evaluación de pequeñas aplicaciones distribuidas utilizando diferentes esquemas
 - Estudio de aspectos relacionados con los sistemas distribuidos

Bibliografía

- *Distributed Systems. Concepts and Design*. 2º edición
G. Coulouris, J. Dollimore, T. Kindberg.
Addison-Wesley, 1994
<http://www.dcs.qmw.ac.uk/research/distrib/book.html>
- *Building Secure and Reliable Network Applications*
K. P. Birman
Manning Publications Co., 1996
- *Distributed Operating Systems*
A. S. Tanenbaum
Prentice-Hall, 1995

Bibliografía

- *Distributed Systems*. 2º edición
S. Mullender (Editor)
Addison-Wesley, 1993
- *Client/Server Computing for Technical Professionals. Concepts and Solutions*.
J. M. Hart, B. Rosenberg
Addison-Wesley, 1995
- *Client/Server Programming with Java and Corba (2º ed.)*
R. Orfali, D. Harkey
Wiley Computer Publishing, 1998
- Cursos sobre sistemas distribuidos y paralelos
 - <http://joda.cis.temple.edu/courses-para.html>

Revistas

- IEEE Concurrency, Parallel, Distributed and Mobile Computig
- Journal of Parallel and Distributed Computing
- Distributed Computing
- Communications of the ACM
- IEEE Computer
- ACM Computing Surveys

Introducción

Contenido

- Introducción y conceptos previos
- Ejemplos de aplicaciones distribuidas
- Características de los sistemas distribuidos
- Desventajas
- Sistemas paralelos
- Principales aspectos de diseño

Evolución de la informática

- En los 70
 - Mainframes centrales
 - Sistemas de tiempo compartido
 - Recursos centralizados
 - Interfaces de usuario poco amigables
 - Aparecen las primeras redes
- En los 80:
 - PCs y estaciones de trabajo
 - Interfaces amigables
 - Redes de área local
 - Aparecen los primeros sistemas operativos distribuidos
 - Mach, Sprite, Chorus, ...

Sistemas Distribuidos

12

Félix García Carballera (1999)

Evolución de la informática

- En los 90:
 - Despegue de las aplicaciones cliente/servidor
 - Mas descentralización
 - Enorme difusión de internet gracias al Web
 - Nuevas necesidades y aplicaciones basadas en el Web
 - Comercio electrónico
 - Multimedia
 - Sistemas de control
 - Aplicaciones médicas

Sistemas Distribuidos

13

Félix García Carballera (1999)

Conceptos previos

- Un **programa** es un conjunto de instrucciones.
- Un **proceso** es un programa en ejecución.
- Una **red de computadores** es un conjunto de computadores conectados por una red de interconexión.
- Un **sistema distribuido (SD)**
 - Modelo físico: conjunto de nodos (procesadores sin memoria ni reloj común) conectados por una red.
 - Modelo lógico: conjunto de procesos que ejecutan *concurrentemente* en uno o más computadores que colaboran y comunican intercambiando *mensajes*.
- Un **protocolo** es un conjunto de reglas e instrucciones que gobiernan la comunicación en un sistema distribuido, es decir, el intercambio de mensajes.

Sistemas Distribuidos

14

Félix García Carballera (1999)

Otras definiciones

- “Un sistema distribuido es aquél en el que no puedes trabajar con tu máquina por el fallo de otra máquina que ni siquiera sabías que existía”
-Leslie Lamport

Sistemas Distribuidos

15

Félix García Carballera (1999)

Modelos de sistemas distribuidos

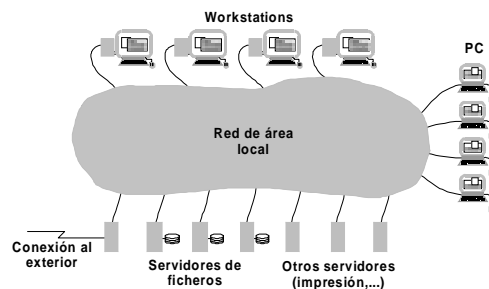
- **Sistema distribuido asíncrono:**
 - Tiempo de entrega de un mensaje no está acotado.
 - Relojes no sincronizados.
- **Sistema distribuido síncrono:**
 - Tiempo de entrega de un mensaje está acotado.
 - Desviación de los relojes acotada.
 - Ventaja: se pueden utilizar el esquema de *timeout* para detectar fallos o pérdidas de mensajes.

Sistemas Distribuidos

16

Félix García Carballera (1999)

Ejemplos: red de área local



Sistemas Distribuidos

17

Félix García Carballera (1999)

Ejemplos de aplicaciones distribuidas

- Correo electrónico, transferencia de ficheros
- Servicios de *News*
- *World Wide Web*
- Sistemas de control de tráfico aéreo
- Aplicaciones bancarias
- Comercio electrónico
- Aplicaciones multimedia (vídeoconferencias, video bajo demanda, etc.)
 - El ancho de banda en estas aplicaciones es un orden de magnitud mayor que en otras
- Aplicaciones médicas (transferencia de imágenes)

Sistemas Distribuidos

18

Félix García Carballeira (1999)

Características de los SD

- Compartir recursos (HW, SW, datos).
 - Acceso a recursos remotos.
 - Modelo cliente-servidor
 - Modelo basado en objetos
- Ofrecen una buena relación coste/rendimiento
- Capacidad de crecimiento
- Tolerancia a fallos, disponibilidad
 - Replicación
- Concurrencia
- Velocidad
 - Paralelismo

Sistemas Distribuidos

19

Félix García Carballeira (1999)

Características de los SD

- Sistemas abiertos
 - Se caracterizan por emplear unas interfaces públicas estandarizadas.
 - Pueden construirse sobre HW y SW heterogéneo de diferentes vendedores.
 - Independiente de los vendedores.
 - Puede extenderse fácilmente:
 - Hardware: nuevos computadores
 - Software: nuevos servicios
- Transparencia
 - Un SD se percibe como un sistema único, no como un conjunto de componentes independientes

Sistemas Distribuidos

20

Félix García Carballeira (1999)

Transparencia

- **Acceso:** acceso a recursos remotos y locales de igual forma
- **Posición:** acceso a los recursos sin necesidad de conocer su situación
- **Concurrencia:** acceso concurrente a recursos compartidos sin interferencias
- **Replicación:** Acceso a recursos replicados sin conocimiento de que lo son
- **Fallos:** mantenimiento del servicio en presencia de fallos.
- **Migración:** permite que los recursos y objetos se muevan sin afectar a la operación de los programas.
- **Capacidad de crecimiento:** facilidad para crecer sin afectar a la estructura del sistema

Sistemas Distribuidos

21

Félix García Carballeira (1999)

Desventajas de los sistemas distribuidos

- Interconexión
 - Coste
 - Fiabilidad, pérdida de mensajes
 - Saturación
- Comunicaciones inseguras
- Software más complejo
- Potencia de cada nodo no adecuada

Sistemas Distribuidos

22

Félix García Carballeira (1999)

Sistemas distribuidos y paralelos

- **Sistemas distribuidos**
 - Objetivo: compartir recursos y colaborar.
 - Redes de computadores
- **Sistemas paralelos**
 - Objetivo: ejecutar un programa muy rápido (*speedup*).
 - Máquinas paralelas (arquitecturas dedicadas)
 - Multiprocesadores
 - Multicomputadores
 - Redes de estaciones de trabajo trabajando como un multicomputador (*cluster*)

Sistemas Distribuidos

23

Félix García Carballeira (1999)

Objetivos de diseño

- Rendimiento
- Capacidad de crecimiento
- Transparencia
- Consistencia
- Fiabilidad:
 - Tolerancia a fallos
 - Disponibilidad
 - Recuperación
 - Consistencia
 - Seguridad
 - Confidencialidad
 - Rendimiento predecible, calidad de servicio

Sistemas Distribuidos

24

Félix García Carballera (1999)

Problemas de diseño

- Nombrado
- Comunicación y sincronización entre procesos
- Estructura software
- Reparto de la carga
- Coherencia
- Calidad de servicio
 - Rendimiento
 - Fiabilidad

Sistemas Distribuidos

25

Félix García Carballera (1999)

Nombrado

- Los usuarios designan a los objetos mediante un *nombre*.
- Los programas designan a los objetos mediante un *identificador*.
- *Resolver* un nombre implica obtener el identificador a partir del nombre.
- **Objetivo importante:** nombres independientes de la posición
- Consideraciones de diseño a tener en cuenta:
 - El espacio de nombres (tamaño, estructura, jerarquía, ...)
 - El servicio de nombres que realiza la resolución

Sistemas Distribuidos

26

Félix García Carballera (1999)

Comunicación y sincronización (C y S)

- Forma básica de C y S: *paso de mensajes*.
 - Mecanismos síncronos
 - Mecanismos asíncronos
- Comunicación cliente-servidor
 - Primitivas de comunicación básicas (*send, receive*)
 - Llamadas a procedimientos remotos
 - Invocación de objetos remotos
- Comunicación en grupos
 - *Multicast, broadcast*
 - Útil para localizar un objeto, tolerancia a fallos, mejorar el rendimiento (replicación), asegurar consistencia

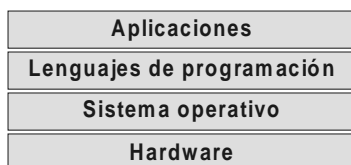
Sistemas Distribuidos

27

Félix García Carballera (1999)

Estructura software

- Estructura software típica de un sistema centralizado



- El sistema operativo
 - Gestiona los recursos
 - Ofrece servicios

Sistemas Distribuidos

28

Félix García Carballera (1999)

Estructura software

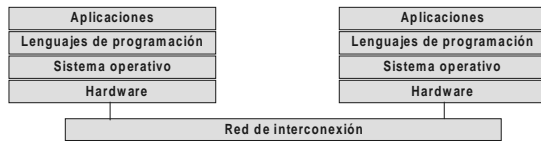
- Existen tres posibilidades para estructurar el software de un sistema distribuido
 - Emplear sistemas operativos en red
 - Utilizar un sistema operativo distribuido
 - Utilizar *middlewares* o entornos distribuidos
- Lo importante es ofrecer un soporte para la programación de aplicaciones distribuidas fácil y transparente

Sistemas Distribuidos

29

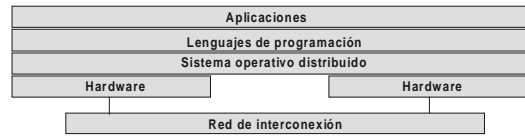
Félix García Carballera (1999)

Sistema operativo en red (SOR)



- El usuario ve un conjunto de máquinas independientes
 - No hay transparencia
- Se debe acceder de forma explícita a los recursos de otras máquinas
- Difíciles de utilizar para desarrollar aplicaciones distribuidas

Sistema operativo distribuido (SOD)



- Se comporta como un SO único
 - Distribución. Transparencia
- Se construyen normalmente como micronúcleos que ofrecen servicios básicos de comunicación
 - Mach, Amoeba, Chorus.
- Todos los computadores deben ejecutar el mismo SOD

Middleware y entornos distribuidos



- Servicios y protocolos estandarizados: Sistemas abiertos
- Ofrecen servicios no incluidos en el SO (servicios de ficheros distribuidos, servicios de nombres, ...)
- Facilitan el desarrollo de aplicaciones distribuidas
- Independientes del HW y del SO subyacente.
- DCE, CORBA, DCOM, Legion, Globe, Globus

Reparto de la carga

- ¿Cómo se asigna memoria y procesador a las aplicaciones?
- Modelos:
 - Estaciones de trabajo: modelo típico
 - Modelo del pool de procesadores (Amoeba)
 - Permiten paralelismo a lo usuarios
 - Uso de una estación de trabajo libre

Asegurar la coherencia

- El problema de la coherencia surge cuando varios procesos acceden y actualizan datos de forma concurrente
 - Coherencia de las actualizaciones
 - Coherencia de la replicación
 - Coherencia de caches
 - Coherencia ante fallos
 - Relojes consistentes

Calidad de servicio

- Rendimiento
 - Tiempo de respuesta adecuado
 - El rendimiento viene determinado por
 - La red de comunicación
 - Los servicios de comunicación empleados
 - El sistema operativo
 - El soporte para la programación de sistemas distribuidos
- Fiabilidad
 - Disponibilidad
 - Consistencia
 - Seguridad
 - Confidencialidad