

# ***Bases de datos para la Web***

---

Marco Donoso

**Centro de Estudios para la Comunidad  
Escuela Politécnica Nacional**

# Conceptos generales

---

- ◆ Con el origen de las computadoras, se comenzó a utilizarlas para almacenar datos y recuperar información a partir de esos datos.
- ◆ El mecanismo físico de almacenamiento y acceso a esos datos fueron inicialmente archivos de texto.

# RDBMS

## Bases de Datos Relacionales

---

- ◆ Conforme creció el volumen de los datos, fue necesario desarrollar software de propósito específico para manejarlos correctamente.
- ◆ **Base de datos Relacional:**
  - ◆ Almacena gran cantidad de información de forma organizada.
  - ◆ El usuario puede hacer consultas (queries) sobre el contenido de esta información de manera rápida y flexible

# La Torre de Babel

---

- ◆ Cada fabricante de Bases de Datos desarrolló su producto con las características y utilizando los formatos que creía necesarios.
- ◆ Como consecuencia de esto, existía una gran cantidad de:
  - ◆ Plataformas
  - ◆ Formatos de almacenamiento de datos
  - ◆ Formatos de acceso a los datos.

# La Torre de Babel (2)

---

- ◆ Por tanto, era muy complicado integrar en un mismo sistema varios tipos de bases de datos.
- ◆ Era necesario desarrollar sistemas específicos para cada aplicación.
- ◆ Nunca una base de datos cubría los requerimientos específicos de cada cliente. El trabajo de adaptación e implementación era muy complejo.

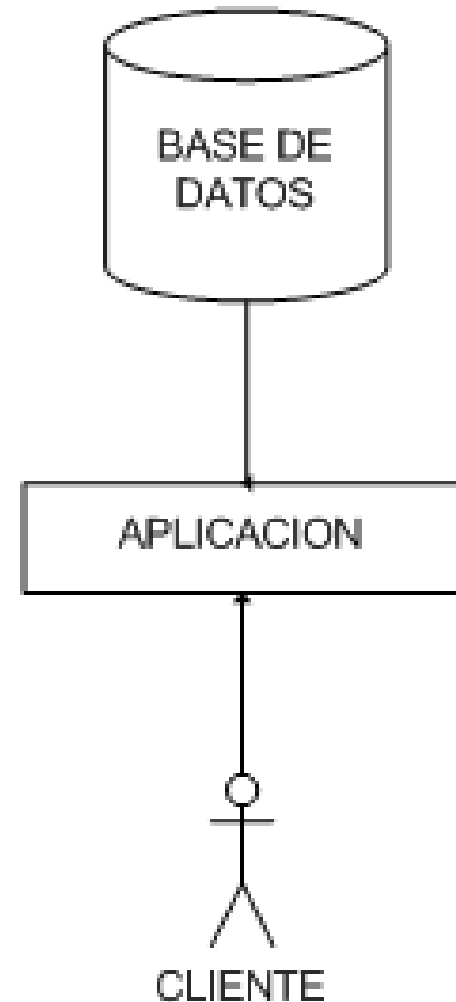
# Estándares: SQL

---

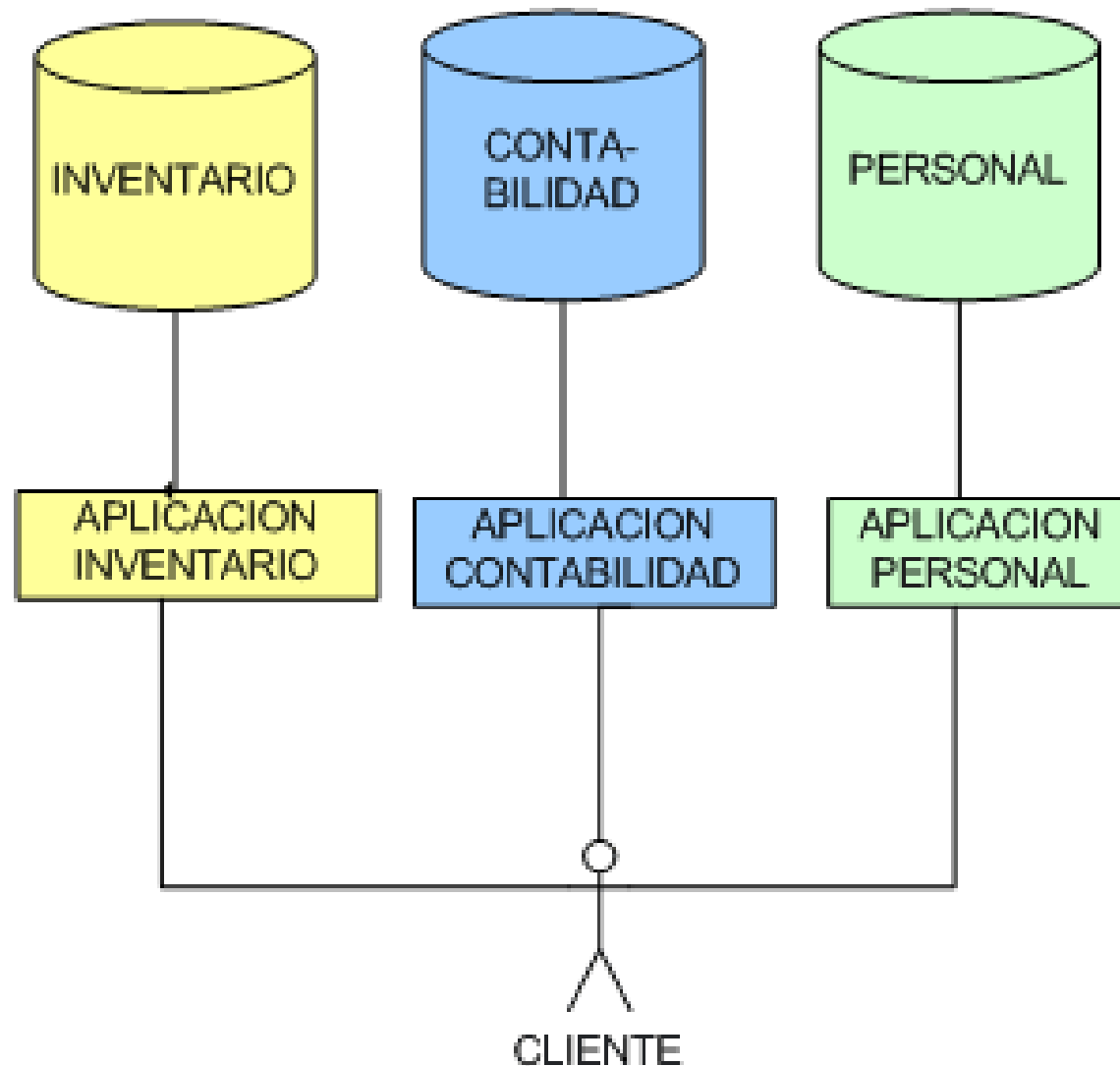
- ◆ ANSI SQL-92
- ◆ Formato estándar de sintaxis para acceder a la información de una Base de Datos Relacional.
- ◆ En la realidad, cada fabricante implementa un conjunto de 'sentencias' SQL estándar y un conjunto de sentencias SQL no estándar.

# Estructura de una aplicación con Bases de Datos

- ◆ Base de Datos  
(servidor de datos)
- ◆ Aplicación  
(servidor de aplicaciones)
- ◆ Cliente



# Estructura de una aplicación con Bases de Datos (2)



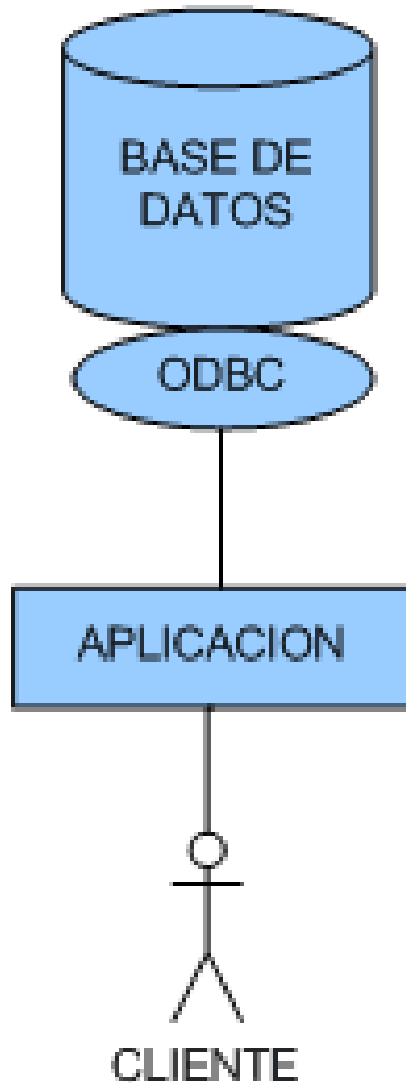
# Más Estándares: ODBC

---

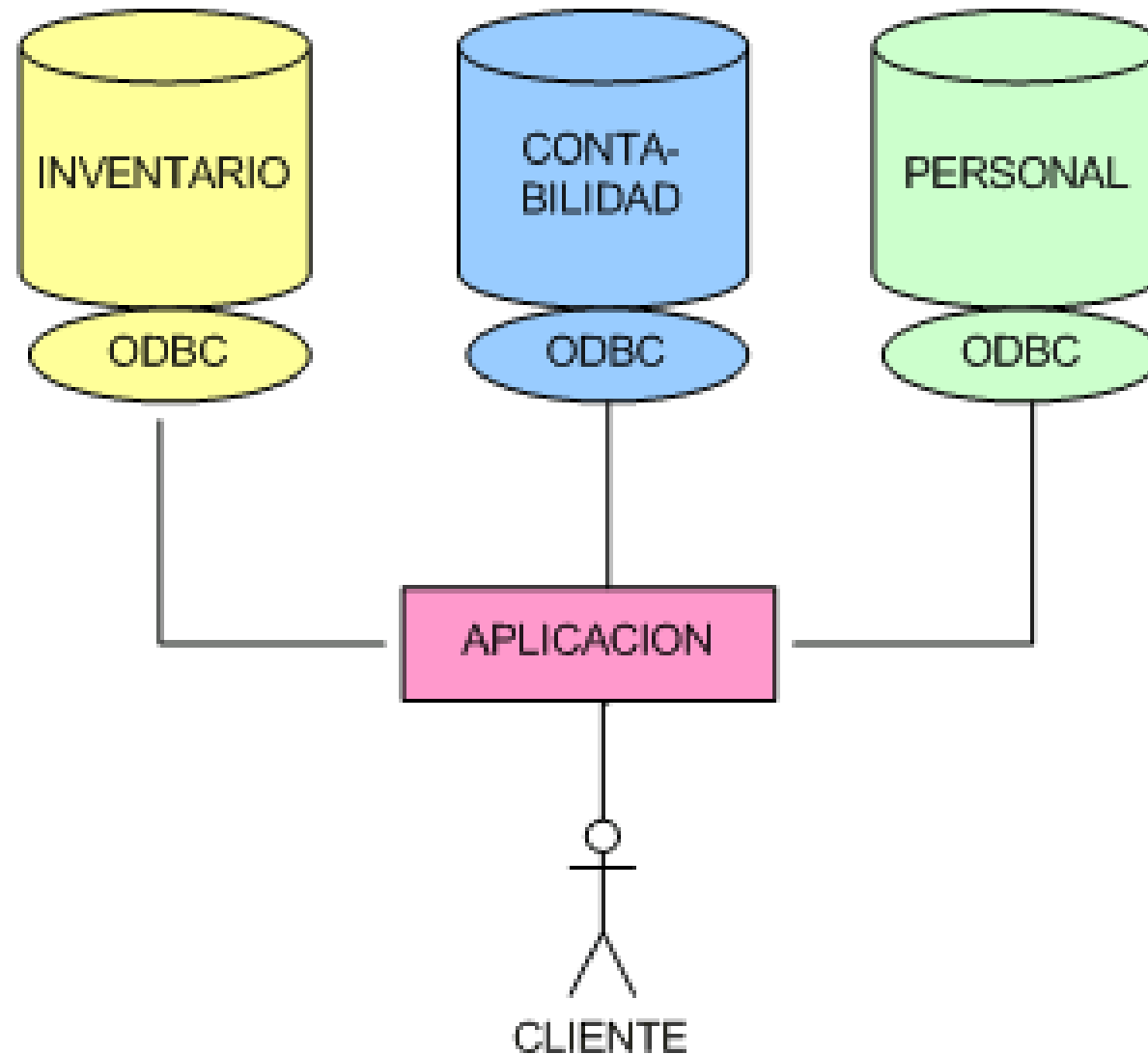
- ◆ ODBC (Open Database Connectivity)
- ◆ Estándar que permite a cualquier aplicación o cliente obtener información de cualquier Base de Datos relacional.
- ◆ Prácticamente todos los fabricantes de bases de datos garantizan que su producto maneja ODBC.

# Más estándares: ODBC

---



# Estructura de una aplicación con bases de datos



# Algunas Bases de Datos

---

## ◆ Comerciales

- ◆ Oracle
- ◆ Sybase
- ◆ Microsoft SQL
- ◆ Access

## ◆ GNU

- ◆ Postgres SQL
- ◆ MySQL
- ◆ Firefox

# Parámetros a tomar en cuenta para decidir sobre una Base de Datos

---

- ◆ Costos (TCO: Costo total de Operación)
  - ◆ Costo inicial
  - ◆ Costo de mantenimiento
- ◆ Desempeño
  - ◆ Capacidad
  - ◆ Robustez

# Costo inicial

---

- ◆ Costo inicial del licenciamiento de la Base de Datos.
  - ◆ Depende del esquema de licenciamiento del fabricante (por cliente, por microprocesador)
- ◆ Costo de la infraestructura física
  - ◆ Hardware
  - ◆ Conectividad (redes)
- ◆ Costo de la implementación
  - ◆ Desarrollo o migración de aplicaciones

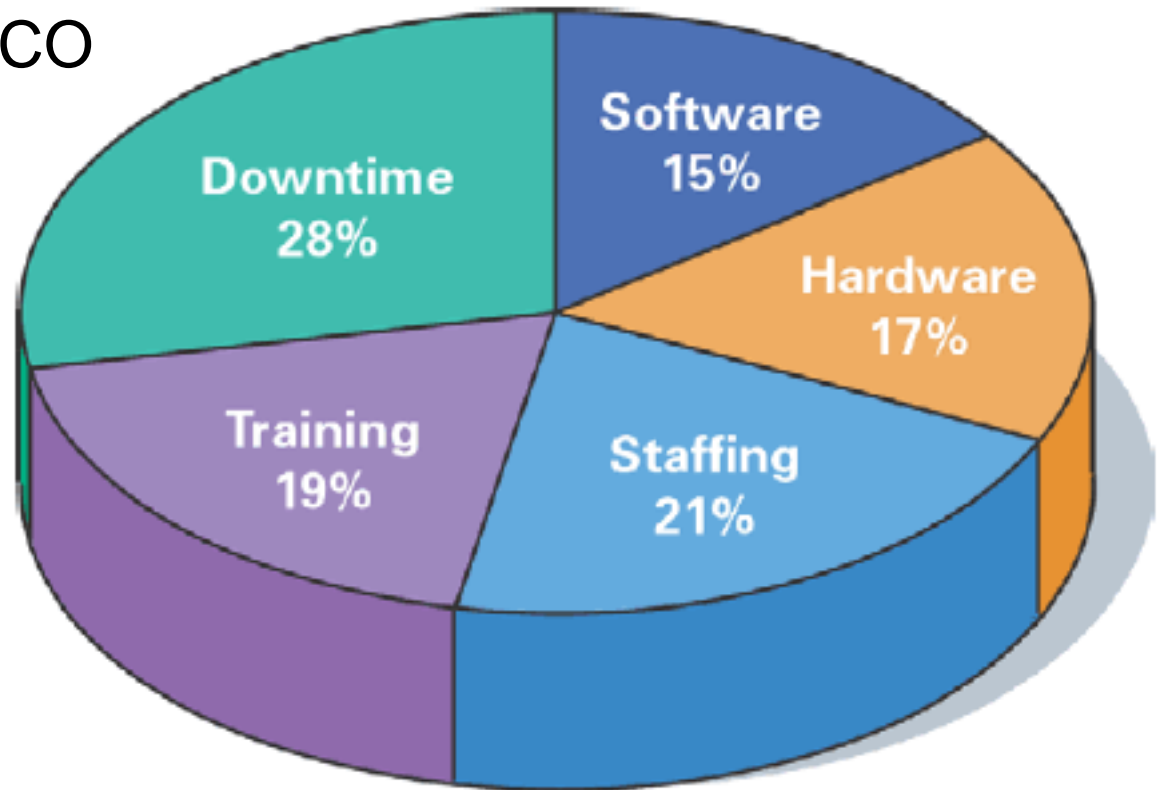
# Costo de mantenimiento

---

- ◆ Costo de capacitación
  - ◆ Personal técnico y administrativo
  - ◆ Cliente o usuario final
- ◆ Costo de soporte técnico del fabricante
- ◆ Costo de paro por fallas
  - ◆ Recursos humanos
  - ◆ Material y producción
  - ◆ Pérdida de competitividad

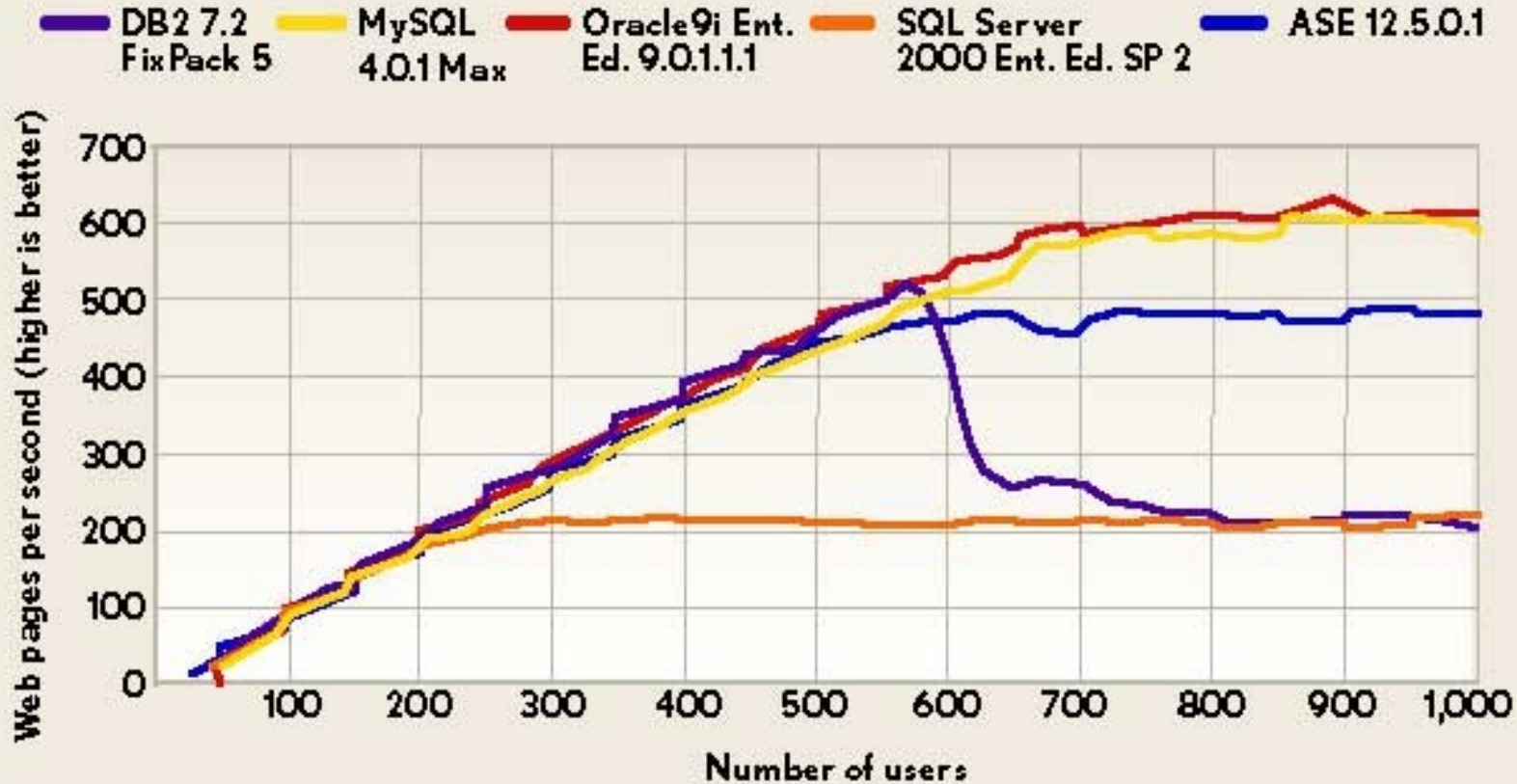
# Costo Total de Operación TCO

- ◆ Comparación de reducción de costos por rubros en la implementación de una base de datos MySQL contra una base de datos Comercial
- ◆ En general, se habla de una reducción del TCO desde un 70% hasta un 85%



# Desempeño

## Oracle9i and MySQL top throughput



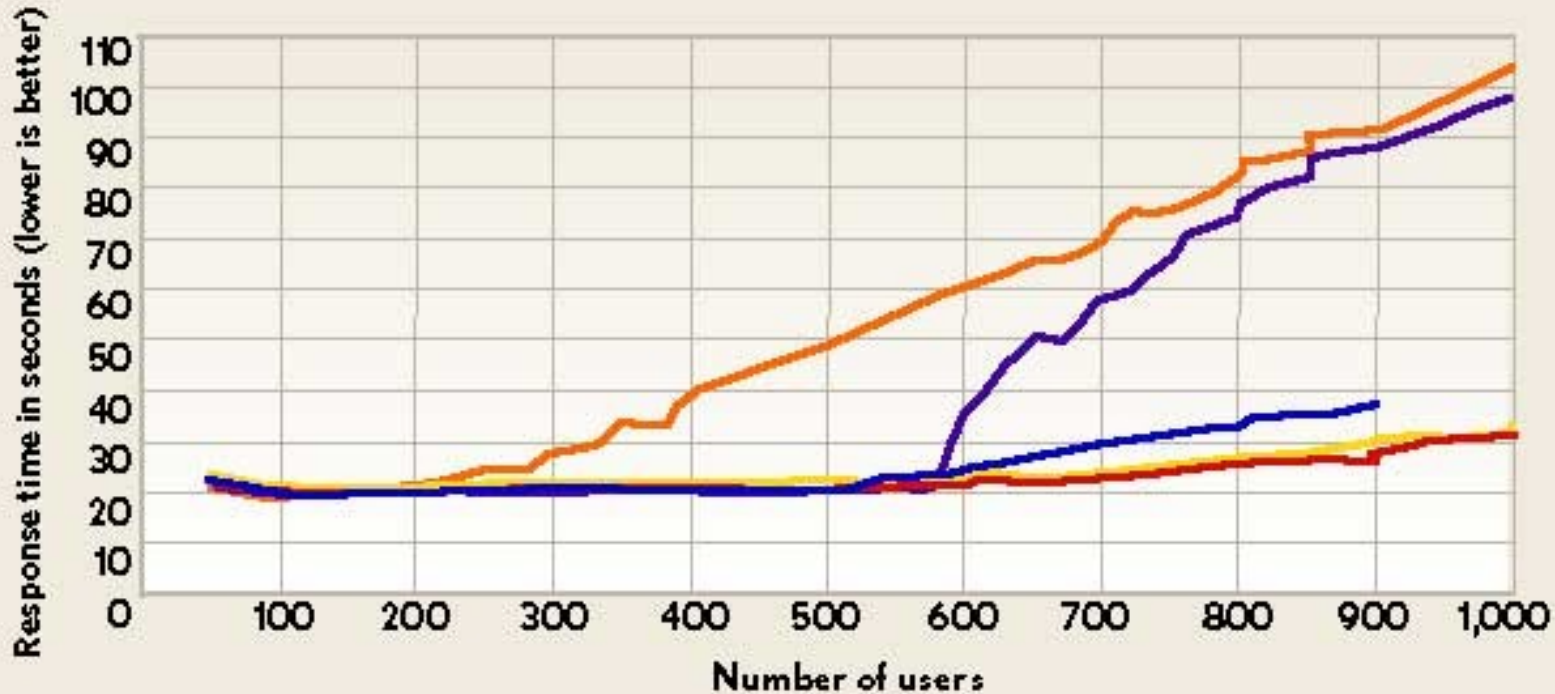
Throughput is in returned Web pages per second from the application server. Number of users is number of concurrent Web clients driving the load. Response time is the time to complete the six bookstore user action sequences, weighted by frequency of each sequence in the mix. All tests were conducted on an HP NetServer LT 6000r with four 700MHz Xeon CPUs, 2GB of RAM, a Gigabit Ethernet Intel Corp. Pro/1000 F Server Adapter and 24 9.1GB Ultra3 SCSI hard drives used for database storage.

Gráfico comparativo de la 'capacidad de respuesta' de varias bases de datos

# Desempeño

## Oracle9i and MySQL offered the fastest response times

DB2 7.2 Fix Pack 5    MySQL 4.0.1 Max    Oracle9i Ent. Ed. 9.0.1.1.1    SQL Server 2000 Ent. Ed. SP 2    ASE 12.5.O.1



Throughput is in returned Web pages per second from the application server. Number of users is number of concurrent Web clients driving the load. Response time is the time to complete the six bookstore user action sequences, weighted by frequency of each sequence in the mix. All tests were conducted on an HP NetServer LT 6000r with four 700MHz Xeon CPUs, 2 GB of RAM, a Gigabit Ethernet Intel Corp. Pro/1000 F Server Adapter and 24 9.1GB Ultra3 SCSI hard drives used for database storage.

Gráfico comparativo de la 'velocidad de respuesta' de varias bases de datos

# Casos de Exito

---

- ◆ Yahoo!
- ◆ Google
- ◆ Cox Communications
- ◆ DaimlerChrysler
- ◆ The Associated Press
- ◆ Earthlink
- ◆ Alcatel SEL
- ◆ Service
- ◆ US Census Bureau
- ◆ Ericsson
- ◆ Hoover's Online
- ◆ Lufthansa
- ◆ NASA
- ◆ Caterpillar
- ◆ Omaha Steaks
- ◆ Powell's Bookstore
- ◆ Sabre Holdings
- ◆ Texas Instruments