



REPORTE DE PRÁCTICA Biblioteca

ALUMNO: Kevin Badillo Olmos

Dr. Eduardo Cornejo-Velázquez

Las 12 leyes de bases de datos distribuidas

Las 12 leyes de las bases de datos distribuidas, formuladas por C.J. Date, describen los principios que un sistema de gestión de bases de datos distribuidas (DDBMS) ideal debe cumplir. Estas leyes se basan en la idea de que la distribución debe ser transparente para el usuario. Aquí tienes un resumen de cada una de ellas:

Leyes de Transparencia

1. **Transparencia de Localización:** Los usuarios no deben saber dónde están almacenados los datos. Deben poder acceder a la información sin importar su ubicación física en la red.
2. **Transparencia de Fragmentación:** La fragmentación de los datos (la división de una tabla en fragmentos más pequeños) debe ser invisible para el usuario. Los datos, aunque divididos y almacenados en diferentes lugares, deben parecer una sola tabla.
3. **Transparencia de Réplica:** Los usuarios no deben saber si los datos están replicados (copiados) en múltiples sitios. El sistema debe gestionar las réplicas automáticamente para garantizar la consistencia.
4. **Transparencia de Concurrencia:** Los usuarios no deben preocuparse por si otros usuarios están accediendo a los mismos datos al mismo tiempo. El DDBMS debe gestionar los bloqueos y la concurrencia de forma automática para evitar conflictos.
5. **Transparencia de Fallos:** El sistema debe poder recuperarse de fallos en el hardware, software o en la red, de manera que los usuarios no se vean afectados por la caída de uno o más sitios.

Otras Leyes de Diseño

6. **Independencia de la Red:** El DDBMS no debe depender del protocolo de la red ni de las tecnologías de comunicación específicas.
7. **Independencia de la Replicación:** El sistema debe ser capaz de gestionar las réplicas sin requerir la intervención del usuario o de la aplicación.
8. **Procesamiento de Consultas Distribuido:** El sistema debe poder procesar consultas de manera eficiente en toda la red, utilizando algoritmos de optimización para minimizar la transferencia de datos entre sitios.
9. **Gestión de Transacciones Distribuidas:** Las transacciones que involucran múltiples sitios deben ser atómicas, lo que significa que o se completan en todos los sitios o no se completan en ninguno. Esto se logra con protocolos como la confirmación de dos fases (2PC).
10. **Independencia del Hardware:** El sistema debe poder funcionar en diferentes tipos de hardware sin problemas.

11. **Independencia del Sistema Operativo:** El DDBMS debe ser compatible con varios sistemas operativos.
12. **Independencia del SGBD (Sistema de Gestión de Bases de Datos):** Si es posible, un DDBMS debe poder interactuar con diferentes SGBD subyacentes, lo que se conoce como heterogeneidad.

Una base de datos distribuida es un conjunto de bases de datos lógicamente relacionadas que están almacenadas físicamente en múltiples computadoras o servidores interconectados por una red. A pesar de estar fragmentada en varios nodos, funciona como una única base de datos lógica para el usuario, que no necesita saber dónde se guardan los datos. Este tipo de sistema surgió para abordar las limitaciones de las bases de datos centralizadas, como la vulnerabilidad a fallos y los cuellos de botella en el rendimiento.

Ventajas

Las bases de datos distribuidas ofrecen una serie de ventajas clave sobre los sistemas centralizados:

- **Rendimiento y Escalabilidad:** Al distribuir la carga de trabajo entre varios nodos, el sistema puede manejar más usuarios y datos sin ralentizarse, mejorando el rendimiento y la latencia. Es más fácil expandir el sistema agregando nuevos sitios a la red sin afectar las operaciones existentes.
- **Fiabilidad y Tolerancia a Fallos:** Si un nodo falla, el sistema puede seguir funcionando gracias a la **replicación de datos**, donde las copias de los datos se guardan en diferentes nodos. Esto garantiza que la información esté siempre disponible.
- **Acceso más Rápido:** Los datos pueden almacenarse físicamente cerca de los usuarios que los necesitan, lo que reduce la latencia y mejora la velocidad de acceso. Esto es especialmente útil para empresas con operaciones en múltiples ubicaciones geográficas.
- **Autonomía Local:** Cada sitio o nodo puede tener su propia autonomía para gestionar sus datos y operaciones, sin depender de un sitio central.

Desventajas 🙄

A pesar de sus beneficios, las bases de datos distribuidas presentan desafíos y desventajas:

- **Mayor Complejidad:** Son más difíciles de diseñar, implementar y gestionar en comparación con las bases de datos centralizadas. Requieren un software más complejo para coordinar los nodos y manejar la comunicación.
- **Costos:** La instalación y el mantenimiento pueden ser más costosos debido a la necesidad de múltiples servidores e infraestructura de red.
- **Problemas de Sincronización:** Sincronizar los datos entre múltiples sitios para mantener la coherencia puede ser un desafío. La duplicación de datos y las actualizaciones pueden generar conflictos.
- **Seguridad:** Resulta más complicado establecer mecanismos de seguridad y control de acceso en un sistema distribuido, ya que los errores en la red pueden tener consecuencias graves para la privacidad de la información.

Conclusiones

En un mundo cada vez más interconectado, las bases de datos distribuidas han surgido como una solución fundamental para superar las limitaciones de los sistemas centralizados. Aunque presentan desafíos significativos en cuanto a complejidad, seguridad y consistencia, sus ventajas inherentes de escalabilidad, fiabilidad y mayor rendimiento las convierten en una tecnología crucial. Al fragmentar y replicar datos estratégicamente, permiten a las organizaciones acercar la información a los usuarios finales, garantizando una alta disponibilidad incluso en caso de fallos. La implementación de un sistema distribuido ideal, guiado por principios como las 12 leyes de Date, asegura que la complejidad subyacente de la red y los datos sea transparente para el usuario final, permitiendo un acceso fluido y eficiente a la información.

Referencias

□ Platzi. (n.d.). *Bases de datos distribuidas*. Recuperado de <https://platzi.com/blog/bases-de-datos-distribuidas/>