Motivación

Las BBDD son la piedra angular de la gestión de datos. Son ofrecidos a otras aplicaciones para su uso o compartición

La capa de gestión de datos	tendremos procesamiento de datos
la capa de explotación	es un conjunto de aplicaciones que utilizan los datos ofrecidos por la capa de gestión de datos

🚀Las aplicaciones de explotación típicamente poseen requisitos de calidad relacionados con el tiempo

🚀Si un sistema va demasiado lento, buscará otro que requiera menos tiempo.

EFICIENCIA

🚀 La eficiencia es el grado en el que el software, tal como una base de datos, emplea óptimamente los recursos del sistema, según el comportamiento del tiempo y de los recursos

¿Cómo hacemos una BD más eficiente?

MÄS RÁPIDO	más memoria caché o prefetch
	discos duros más rápidos
	mejores CPUs
	cambiar la compresión de los datos
	escalar la base de datos
MENOS RECURSOS consumidos	Haciendo mejor uso de los métodos de acceso a la BD. Método de acceso es la manera en la que una BD encuentra y devuelve un dato concreto, o un conjunto de datos.
	Típicamente utilizan estructuras de datos complementarias

Métodos de acceso a los datos

El método que utilizan las BD para acceder a los datos es el escaneo secuencial. Va fila por fila.

Hay métodos de acceso más eficientes: los SGBD normalmente utilizan índices. Estos índices son metadatos: datos sobre los datos

Tipos de índices: el más común es el B+ (es un árbol). Nos permite acceder a los datos en tiempo de acceso logarítmico

CUIDADO: Si creamos un índice que no es útil nos puede empeorar el rendimiento, tiene un coste.

Métodos de acceso a los datos (cont)

INDICE COMPUESTO: conseguimos encontrar elementos que cumplan dos condiciones a la vez

```
CREATE INDEX titletype_index;
ON title_basics (titletype);

DROP INDEX titletype_index;
```

EL TEOREMA CAP

las RDBMS, dentro del mundo de los DBMS, tienen un grave problema o, mejor dicho, una limitación teórico-práctica

Todas las bases de datos tienen un problema, respecto al teorema CAP.

No se puede elegir no tener ningún problema, pero si elegir cuál problema quieres tener.

En el mundo de los RDBMS, típicamente, ese problema es fijo: No se puede elegir

➢ El teorema CAP describe las limitaciones de las bases de datos distribuidas en una red: Un clúster de bases de datos, o cualquier base de datos escalada horizontalmente.

Es imposible garantizar simultáneamente la consistencia, la disponibilidad y la tolerancia a particiones en una base de datos distribuida.	La consistencia se refiere a la propiedad de una base de datos de garantizar que todas las copias de la base de datos en diferentes nodos de la red contengan los mismos datos.
	La disponibilidad se refiere a la propiedad de una base de datos de estar siempre disponible para realizar lecturas y escrituras.
	La tolerancia a particiones se refiere a la capacidad de una base de datos para seguir funcionando correctamente incluso si hay una interrupción en la comunicación entre algunos de los nodos de la red.

Los diseñadores de bases de datos distribuidas deben elegir dos de estas propiedades para garantizar y sacrificar la tercera.

Los sistemas de bases de datos relacionales son típicamente CA, y típicamente sirven para entidades bancarias. ▪ No son tolerantes a particiones.

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com

Recuperación

Una BD debe guardar los datos de forma persistente:	si se corrompe --> poder recuperarla
	si falla --> No perder cambios de transacciones confirmadas, ni mantener cambios de transacciones abortadas
PELIGROS:	cancelación de transacciones
	caida del sistema (cancelar todas las transacciones activas)
	destruccion de BD
RECUPERACION	restauracion: ROLLBACK. restauracion hacia atras; RESTAURACION hacia delante. rehacer cambios de transaccion confirmada. EL LOG guarda información sobre los cambios en la BD
	reconstrucción: se realiza mediante dos fuentes diferentes. BACKUP: completo o incremental. LOG: a partir del momento que se hizo el backup.
	De manera externa mediante RAIDs o mediante otra base de datos (distribuida)
ESCALADO	Proceso de aumentar la capacidad de la BD para mayor cantidad de datos o para mas solicitudes
	- añadiendo hardware, o distribuyendo BD
	objetivo siempre mejorar el rendimiento de la BD

ESCALADO

Escalado de una BD es el proceso de aumentar la capacidad de la BD para manejar una mayor cantidad de datos o un mayor número de solicitudes de consulta

- Escalada vertical: se refiere a la adición de más recursos de HW a una sola máquina para mejorar el rendimiento de la BD. TIENE UN LIMITE. Es bastante simple, no hay que cambiar nada.

- Escalado horizontal: distribución de la BD en varias máquinas para mejorar el rendimiento o la tolerancia a fallos de la BD. Es más complejo, pues implica que varias máquinas tienen que comunicarse entre sí de manera continua, distribuir las tareas y datos entre ellas, etc. VER REPLICACIÖN y PARTICIONAMIENTO.

- Escalado diagonal: mezclar el horizontal y el vertical

REPLICACION: consiste en crear copias de la BD en varias máquinas para distribuir la carga de trabajo entre varias máquinas y mejorar el rendimiento.

PARTICIONAMIENTO: dividir grandes tablas en fragmentos más pequeños y distribuirlos en varias máquinas. Esto mejora el rendimiento y capacidad de manejo de los datos. Requiere un SHARD KEY

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com