

BD NoSQL

NO SQL
Not Only SQL

BD NoSQL

Procesamiento Batch Tradicional
VS
Procesamiento Transaccional

Procesamiento Batch

Mainframe job

Planificador de jobs corre procesos
secuencialmente

Los Procesos tienen su propia CPU, sus propios
datos. No se comparte nada. No conflictos.

Procesamiento Transaccional

SBD

Sesiones de usuario conectados a SGBD

Las tablas no son archivos y el usuario no se conecta a un archivo. El DCL (data control language) dentro del SGBD decide cuales tablas son accesibles a que usuarios.

Todo compartido. Maxima concurrencia.

Jim Gray 70's → Transaccion → ACID (Atomicity, Consistency, Isolation, Durability) → Sistemas de control de concurrencia (SC)

Procesamiento Transaccional

SC Pesimista → Locks

SC Optimista

Centralizado Vs Distribuido

SD (Sistemas Distribuidos), Eric Brewer 00's,
Teorema CAP (**C**onsistency, **A**vailability, **P**artition
Tolerance): dos de tres, pero no los tres (solo en caso
de falla de la red se elige entre C o A)
C=ACID=todos los nodos ven lo mismo
A=disponibilidad, cada nodo debe responder ante una
peticion del usuario
P=si uno o mas nodos fallan, el sistema debe seguir
funcionando

Centralizado Vs Distribuido

SD ej Google BigTable, Amazon Dynamo, Facebook Cassandra = **BASE** (**B**asically **A**vailable, **S**oft state, **E**ventual consistency)

BA=SD garantiza disponibilidad del dato (x CAP), aunque la respuesta puede ser “error/fallo” si el dato es inconsistente o ha cambiado de estado

S=SD en permanente cambio → estado soft, aunque algunos datos no cambien (ej tablas con datos fijos)

E=estado consistente alcanzado luego de que el SD detenga su input de datos → t de inconsistencia aceptable

Centralizado Vs Distribuido

Conclusiones

SGBD = ACID = Elijen consistencia sobre disponibilidad

SD = BASE = NoSQL (mayoritariamente) = Elijen disponibilidad sobre consistencia

Why Not SQL?

- Los datos no están en una máquina, incluso en una red
- Muchos de esos datos ni siquiera son datos nuestros
- El volumen de datos es tan grande que no podemos ponerlos en un solo lugar
- Los datos no arriban en forma coordinada: ni en tiempo ni espacio
- No siempre son datos estructurados que fácilmente podemos procesar en un SGBDR

Clasificación BD No SQL

- Orientadas a Objetos
- Columnares
- Grafos
- Modelo MapReduce
- Streaming and Complex Events DB
- Clave-Valor
- Textuales / Documentales
- Geograficas
- Big Data and Cloud Computing
- Biometricas

Clasificación BD No SQL

- Analíticas
- Multivaluadas
- Jerárquicas y de Red

Bibliografía

- [1] Joe Celko's, "Complete Guide to NoSQL. What Every SQL Professional Needs to Know About NonRelational Databases", Elsevier Inc., 2014. Introduction, Chapter 1.