

1970 SQL

① Volume

bcp de données

② Vitesse

un débit monstrueux

③ variété

des données hétérogènes

2010 NoSQL

clés-valeurs

Colonnes

Documents

Graphes

cf p 150

## Base de données relationnelles & propriété

Les bases de données relationnelles ont été inventées en 1970. Il s'agit de documents tabulaires dans laquelle les données sont définies afin d'être accessibles et de pouvoir être réorganisées de différentes manières.



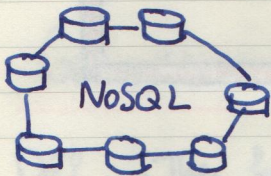
**Atomicité :** Une transaction s'effectue entièrement ou pas du tout

**Cohérence :** Le contenu d'une base doit être cohérent au début et à la fin d'une transaction

**Isolation :**

**Durabilité :**

## Base de données distribuées & propriété



Basically : quelle que soit la charge  
Avalable de la base de données,  
le système garantie un taux  
de disponibilité de la donnée

S oft-state : La base peut changer lors des mises à jour ou lors d'ajout / suppression de serveurs. La base NoSQL n'a pas à être cohérente à tout instant

E ventually consistent : À terme, la base atteindra un état cohérent

## Toutes base de données & Théorème de CAP

Dans toute base de données, vous ne pouvez respecter au plus que 2 propriétés parmi la cohérence, la disponibilité et la distribution.

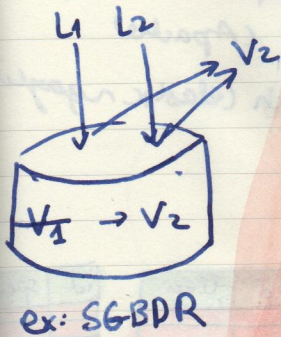
Consistency (cohérence): Une donnée n'a qu'un seul état visible quel que soit le nb de réplias

Availability (Disponibilité): Tant que le système tourne, la donnée doit être disponible.

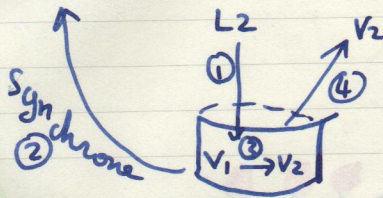
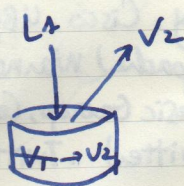
**Partition Tolerance (Distribution):** Quel que soit le nb de serveurs, toute requête doit fournir un résultat correct.



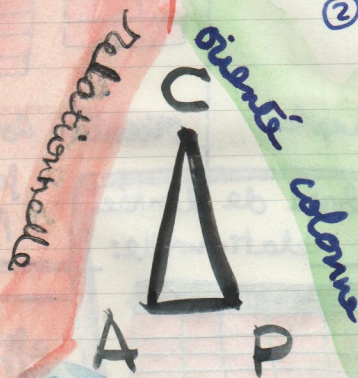
Relationnel Orienté clé-valeur, Orienté colonne  
Orienté Document Datum / Document Orienté graphe



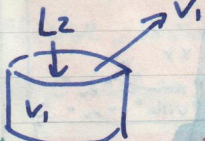
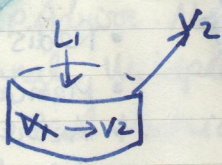
Oracle  
MySQL  
SQL Server  
PostgreSQL  
DB2



Cosmos DB\*  
HBase / BigTable  
MongoDB\*



NoSQL / clé-valeur /  
graphe / orienté document



SimpleDB\*  
Memcached\*  
Redis

Elasticsearch  
Spark

DynamoDB\*  
CouchBase\*  
Cassandra\*

Neo4j  
OrientDB  
FlockDB

\* Possibilité de changer le mode de cohérence à disponibilité ou de disp → coh

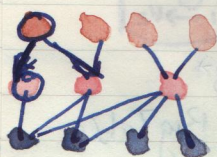


05/07/2020

Datum / Date: NoSQL

- Neo4j eBay, Cisco, UBS, HP
- OrientDB (Apache) Warner Music Group, Comcast
- Flock DB (Twitter) : Twitter

- BigTable (Google)
- HBase (Apache, Hadoop)
- Spark SQL (Apache)
- Elasticsearch (elastic ~ google)



orienté graphe

+ relation

orienté colonne

localiser sur une ou plusieurs colonnes  
ex: count() etc

id	type

id	lieu

id	spec

base de donnée relationnelle:

id	type	lieu	...	spec

Structure complexe

efficacité & simplicité

orienté doc

clé doc

clé	doc
	{ "id": "xx", "type": "xx", "lieu": { "nom": "xx", "ville": "xx" }, : }

clé - valeur

seuls opérations

- Redis (VMWare):
- Memcached (Dang)
- Azure Cosmos DB
- SimpleDB (Amazon)

Create (key, value)

Read (key)

Update (key, value)

Delete (key)

- Mongo DB (Adobe, eBay, Bosch)
- CouchBase (Apache, Hadoop), Disney, PayPal, Ryanair
- DynamoDB (Amazon): BMW, ...
- Cassandra (Facebook → Apache) NYTimes, eBay

clé:

c1	c2	...	CN
type: xx			
lieu: xx			
spec: xx			

valeur