

8. Complements

- ✓ 1. SGBD
- ✓ 2. Escalabilitat
- ☞ 3. NO SQL
- 4. Altres "models"

8.3 NoSQL

- ✓ a. Conceptes i principis
 - Big Data
 - NoSQL
 - ACID vs BASE
 - Teorema CAP
- ☞ b. Tipus
- c. Pros/Contres

8.3 NoSQL

b. Tipus

Es poden trobar múltiples classificacions, basades habitualment en la forma/tipus de dades que gestionen. Per exemple:

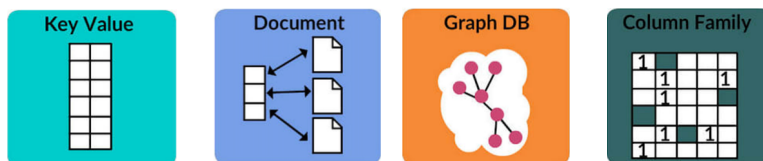
- Clau-valor
- Documents
- Grafs
- Famílies de columnes
- Multimodel
- Objectes
- Cloud
- XML
- Multivalor
- ...

3

8.3 NoSQL

b. Tipus

Però, a dia d'avui, la classificació més habitual és:

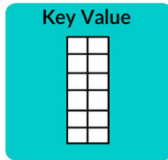


... i, més recentment, multimodel!

4

8.3 NoSQL

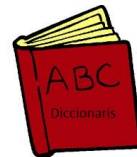
b. Tipus



Son els models més simples. Cada ítem de la bd està emmagatzemat en forma de parella:

Key: el nom de l'ítem
Value: el seu valor

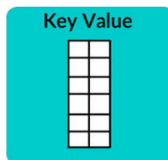
- Escalable, permet manejar grans volums de dades i d'usuaris
- Queries simples, variades i molt ràpides
- DML també molt senzill: get(key), put(key,value), delete(key)
- Permet consistència
- Transaccional



5

8.3 NoSQL

b. Tipus



Limitacions

- Relacions molt limitades
- Les transaccions son multi-operació: afecten a grups de claus
- Les cerques (sempre sobre claus) basades en informació de la part valor
- Les operacions sempre son sobre una única clau (*tipus* de clau). No es pot operar sobre varies claus simultàniament

Son molt adequades per control de perfils: *userId*, *username*



Amazon DynamoDB

Airbnb Netflix
Nike ...



Uber



Zendesk Amazon

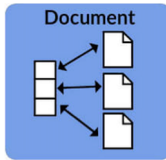


Twitter Pinterest
Coinbase

6

8.3 NoSQL

b. Tipus



Gestionen parells clau-document. Un document es una estructura complexa formada per parells clau-valor o clau-array o clau-document.

Un document es tracta com una unitat que es pot descompondre en les seves parts.

- Escalable, permet manejar objectes complexos
- Poden ser ACID
- Poden implementar característiques dels relacionals

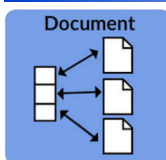
Limitacions

- Les pròpies del document
- Informació repetida entre documents
- Disseny (modelització) complex

7

8.3 NoSQL

b. Tipus



```
<order>
  <number> 1 </Number>
  <date> 11/20/2002 </date>
  <customercode> C7854 </customercode>
  <orderline number="1">
    <productcode> CDR1080 </productcode>
    <description> CD </description>
    <quantity> 10 </quantity>
  </orderline>
</order>
```



```
{
  "orders": [
    {
      "orderno": "748745375",
      "date": "June 30, 2088 1:54:23 AM",
      "trackingno": "TN0039291",
      "custid": "11045",
      "customer": {
        "custid": "11045",
        "fname": "Sue",
        "lname": "Hatfield",
        "address": "1409 Silver Street",
        "city": "Ashland",
        "state": "NE",
        "zip": "68003"
      }
    }
  ]
}
```

Els documents
poden ser XML o
JSON

{ JSON }

JavaScript Object Notation

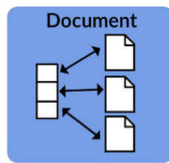
JSON

mongoDB

8

8.3 NoSQL

b. Tipus



Exemples:

XML

eXistdb

MarkLogic™

Grans bancs,
Chevron



mongoDB.



CouchDB

JSON



OrientDB
An SAP Company



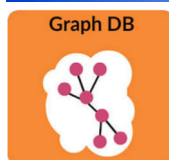
Amazon DocumentDB

SEGA, Cisco,
Weather Channel
Alibaba ...

9

8.3 NoSQL

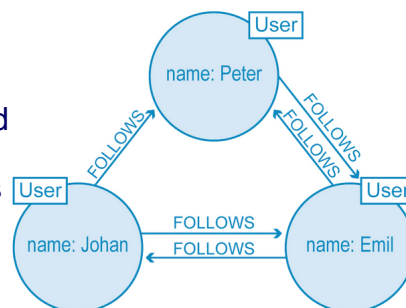
b. Tipus



Emmagatzemen i gestionen estructures graf: col·leccions de nodes i arestes emprades per guardar informació sobre xarxes: neuronals, comunicacions, xarxes socials, etc.

Parts d'un graf:

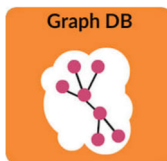
- Node: representen entitats
- Propietats: informació dels nod
 - Es poden indexar
- Arestes: relacions entre nodes
 - Es poden indexar
 - Direccionals/bidireccionals



10

8.3 NoSQL

b. Tipus



- Escalen “per complexitat” del graf
- Focalitzades en les inter-conexions

Limitacions

- Limitades quan a *performance* de la concurrència
- Un graf no sol ser particionable: problemes de concurrència
- No hi ha cap tipus d'estàndard quan als llenguatges
- Disseny (modelització) complex

Aplicacions: aeronàutica, control de xarxes de transport (carreteres, trens ...), logística, històries mèdiques, detecció de frau, etc.



Walmart
Amazon
Cisco

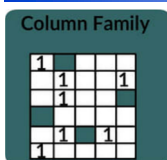


Yahoo
Facebook
Twitter

11

8.3 NoSQL

b. Tipus



Column Family / Wide Column / Columnar

Les dades s'organitzen per **columnes**.

Aquestes es poden agrupar per **famílies**.

Podem pensar una família com una fila del relacional.

Les famílies s'agrupen en **keyspaces**

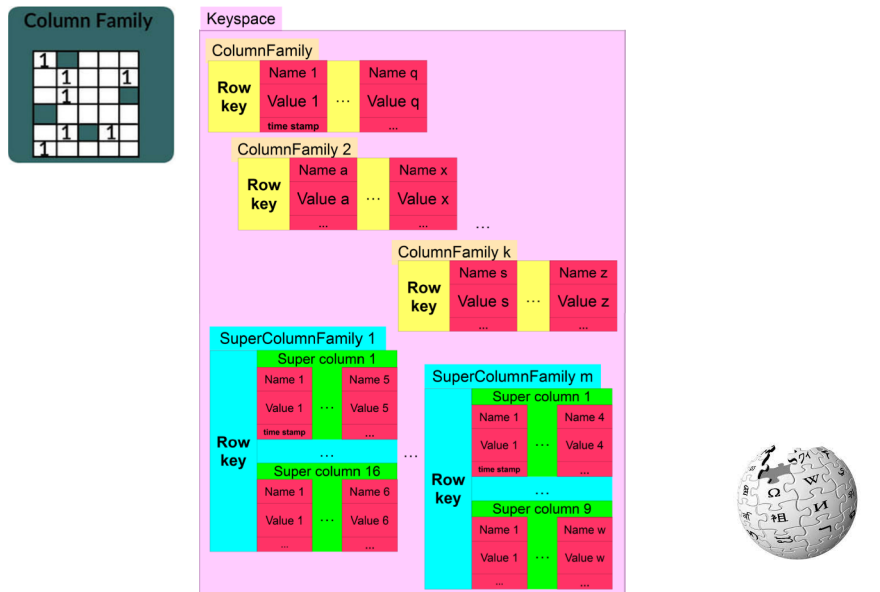
*A **column family** is a database object that contains columns of related data*

*A **keyspace** is an object that holds together all column families of a design*

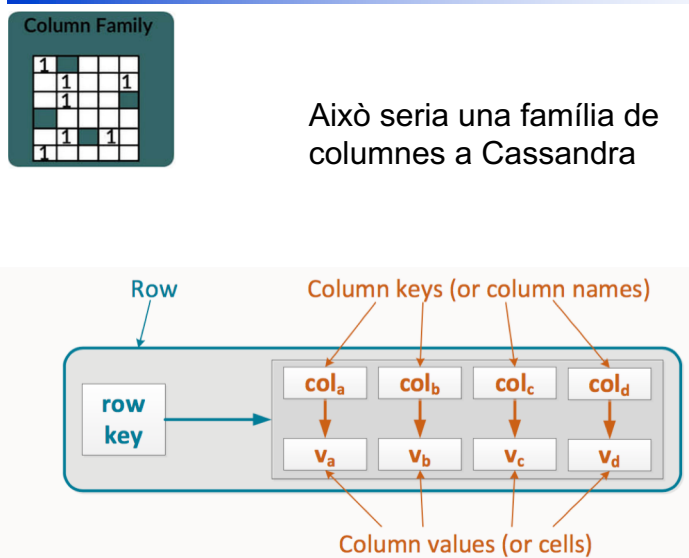


12

b. Tipus



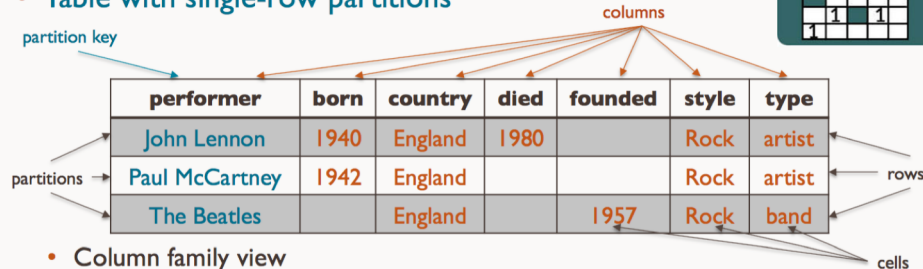
b. Tipus



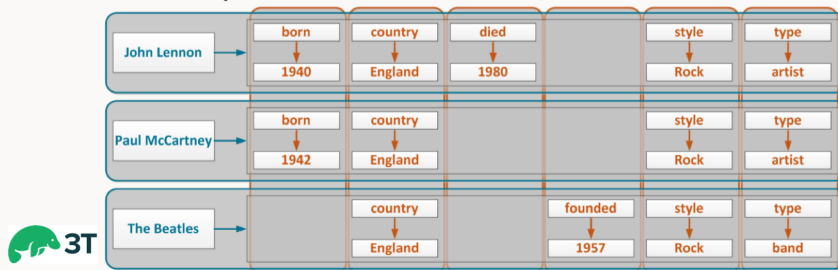
8.3 NoSQL

b. Tipus

- Table with single-row partitions



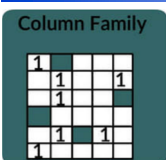
- Column family view



15

8.3 NoSQL

b. Tipus



- Son molt escalables
- *Very high performance*: es carregen molt aviat
→ queries molt ràpides
- Permeten molt bona compressió de dades
- Donen molts bons resultats per consultes amb agregacions

Limitacions

- Escripcions lentes
- Treballen amb grups grans de dades, les operacions sobre ítems concrets son lentes.



cassandra
Spotify, Outbrain



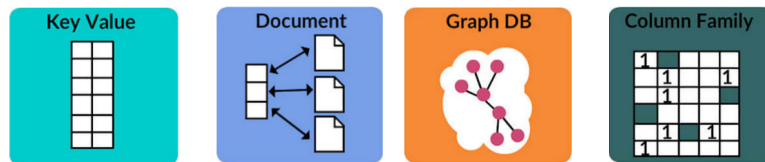
Facebook



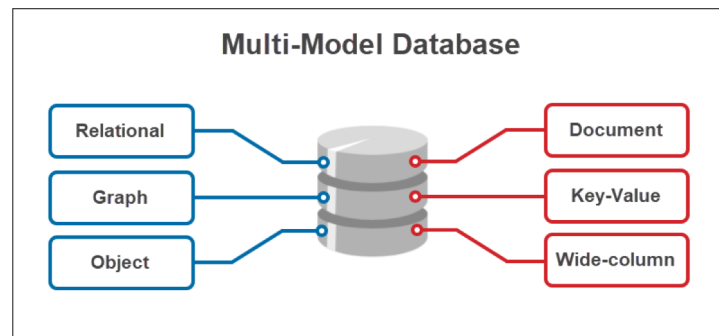
16

8.3 NoSQL

b. Tipus



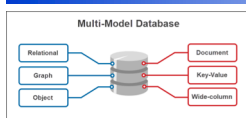
... i, més recentment ...



17

8.3 NoSQL

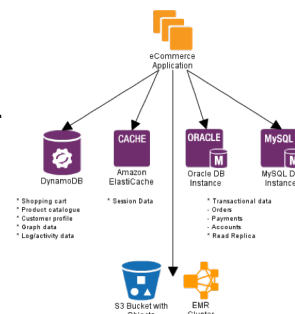
b. Tipus



No confondre amb:



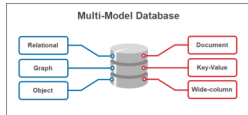
Polyglot persistence is a term that refers to using multiple data storage technologies for varying data storage needs across an application or within smaller components of an application.



18

8.3 NoSQL

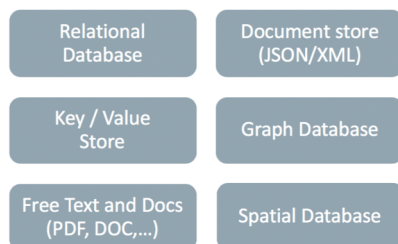
b. Tipus



[ORACLE Multimodel Database](#), maig 2019

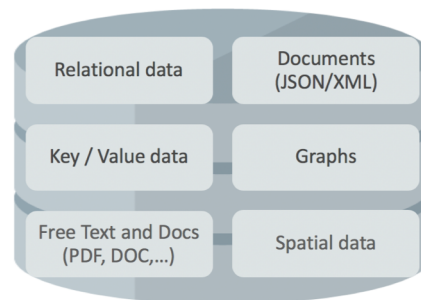
Polyglot

Multiple Databases with Single-model



Multi-model

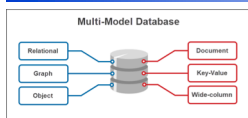
Single Database with Multiple Models



19

8.3 NoSQL

b. Tipus



[ORACLE Multimodel Database](#), maig 2019

Multimodel Database

- Oracle Database supports multiple models
 - Relational, In-memory, Sharded
 - Document Store
 - JSON
 - XML
 - Text
 - Spatial Database
 - Graph Database and Triple Store
- Oracle Database support multiple languages and access protocols



20

8.3 NoSQL

b. Tipus. Multimodel



Database	SQL	Document	Graph	Object	License	Transactions
ApacheDB	No	Yes	Yes	No	Apache 2 License	Full ACID, pessimistic locking, configurable durability
Azure Cosmos DB	Yes	Yes	Yes	Yes	Proprietary	Full ACID within a partition, multiple consistency models
CrateDB	Yes	Yes	No	Yes	Apache 2 License	Eventual consistency , Optimistic concurrency control
EnterpriseDB	Yes	Yes	No	Yes	Proprietary	Full ACID
MarkLogic	Yes	Yes	Yes	No	Proprietary	Full ACID
OrientDB	Yes	Yes	Yes	Yes	Apache 2 License	Full ACID, even distributed
SAP HANA	Yes	Yes	Yes	No	Proprietary	Full ACID
Virtuoso	Yes	Yes	Yes	Yes	Proprietary or GNU GPL v2	Full ACID

21

8.3 NoSQL

- ✓ a. Conceptes i principis
 - Big Data
 - NoSQL
 - ACID vs BASE
 - Teorema CAP
- ✓ b. Tipus
- ☞ c. Pros/Contres

22

8.3 NoSQL

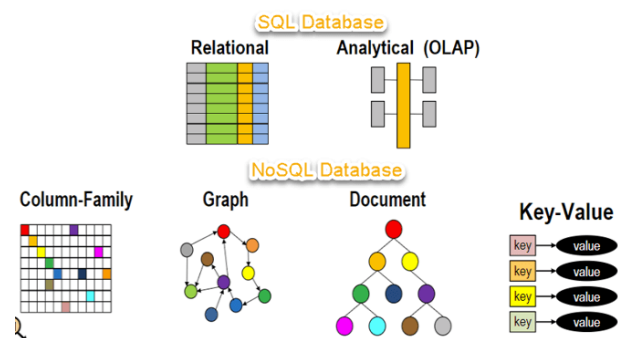
c. Pros/Contres

Feature	NoSQL Databases	Relational Databases
Performance	High	Low
Reliability	Poor	Good
Availability	Good	Good
Consistency	Poor	Good
Data Storage	Optimized for huge data	Medium sized to large
Scalability	High	High (but more expensive)

23

8.3 NoSQL

c. Pros/Contres



24

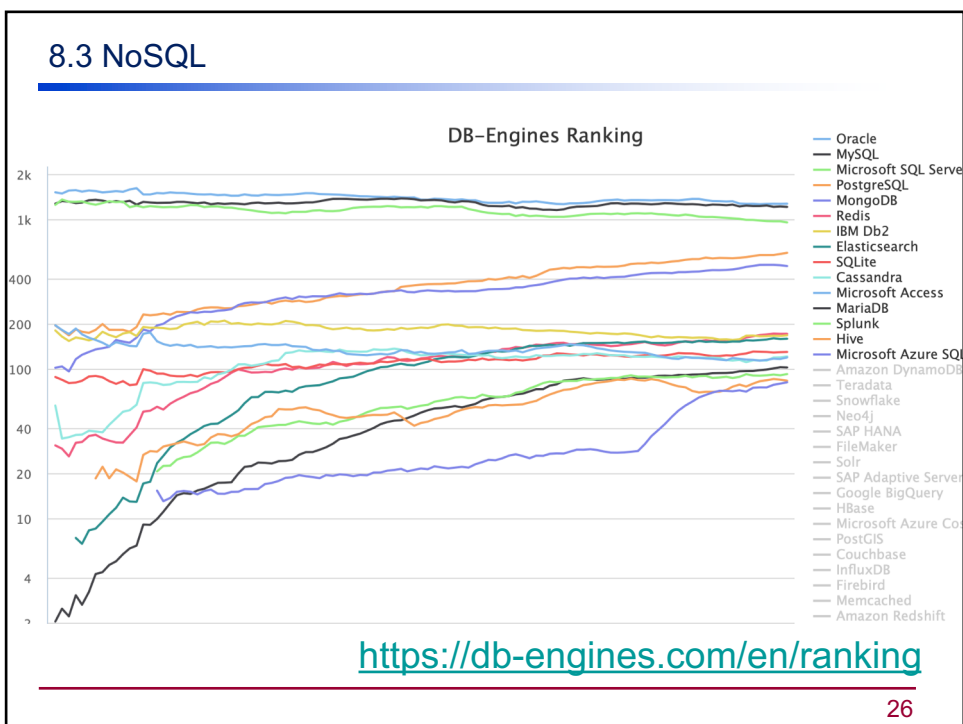
8.3 NoSQL

Nov 2021	Oct 2021	Nov 2020	DBMS	Database Model	Nov 2021
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1272.73
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1211.52
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	954.29
4.	4.	4.	PostgreSQL + ⓘ	Relational, Multi-model ⓘ	597.27
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	487.35
6.	6.	7. ↑	Redis +	Key-value, Multi-model ⓘ	171.50
7.	7.	6. ↓	IBM Db2	Relational, Multi-model ⓘ	167.52
8.	8.	8.	Elasticsearch	Search engine, Multi-model ⓘ	159.09
9.	9.	9.	SQLite +	Relational	129.80
10.	10.	10.	Cassandra +	Wide column	120.88

<https://db-engines.com/en/ranking>

25

8.3 NoSQL



26