

This is the html version of the file <https://moodle.polymtl.ca/mod/resource/view.php?id=255463>.
Google automatically generates html versions of documents as we crawl the web.

LOG8430: Architecture logicielle et conception avancée

Choosing Persistence Architecture Automne 2017

Fabio Petrillo
Chargé de Cours

Page 2

Example of architecture for a data system

2

Page 3

Persistence

3

Page 4

Persistence

- Persistence simply means that we would like our application's data to **outlive the applications process**.
- The best-known persistence data model today is probably that of SQL, based on the relational model proposed by Edgar Codd in 1970: data is organized

into relations (called tables in SQL), where each relation is an unordered collection of tuples (rows in SQL).

- In the 2010s, NoSQL is the latest attempt to overthrow the relational model's dominance, addressing several issues.

Object-Relational Impedance Mismatch Problem

- After 1990's due to popularity of **HTTP**, the cost of posting and exchanging information became **cheaper** which led to the flooding of information on Internet.
- It was realized that **traditional techniques** of data storage will soon become **stale and inefficient** to handle such vast amount of **unstructured and semi-structured** data.
- Not all the information generated on Web is structured, rather interactive Web has produced more semi-structured or unstructured data.
- All the available rich information cannot be **forcefully** made to fit in the tabular format of **relational databases**.

Page 6

Page 7

NoSQL

7

Page 10

Page 11

Slides: slideshare.net/felixgessert
Article: medium.com/baqend-blog

12

Page 13

Page 14

Page 15

NoSQL

- The major problem is that **legacy tools** require data **schema** to be **defined *a priori*** to the data creation. In today's world deciding on **rigid pre-defined schema is unrealistic**.
- NoSQL solutions should not be thought of as replacement for RDBMS, instead as a complementary product for handling issues of **scalability and complexity**.
- These NoSQL databases may require **additional storage**, since data is **denormalized**, but results in overall improvements in **performance, flexibility and scalability**.
- These systems typically **sacrifice some of these dimensions**, e.g. database-wide **transaction consistency**, in order to achieve others, e.g. higher availability and scalability.

15

Page 16

NoSQL systems key features

- the ability to horizontally scale “simple operation” throughput over many servers
- the ability to replicate and to distribute (partition) data over many servers

- a simple call level interface or protocol (in contrast to a SQL binding)
- a **weaker** concurrency model than the **ACID transactions** of most relational (SQL) database systems
- efficient use of distributed indexes and RAM for data storage
- the ability to dynamically add new attributes to data records
- A key feature of NoSQL systems is “shared nothing” horizontal scaling – replicating and partitioning data over many servers. This allows them to support a large number of simple read/write operations per second.

Rick Cattell. 2011. *Scalable SQL and NoSQL data stores*. SIGMOD Rec. 39, 4 (May 2011), 12-27

16

Page 22

Page 23

Page 24

Page 27

27

Page 28

Page 29

BASE Transactions

- **BA**sically Available
- Soft state
- **E**ventually consistent
- The idea is that by giving up ACID constraints, one can achieve much higher performance and scalability.

Rick Cattell. 2011. *Scalable SQL and NoSQL data stores*. SIGMOD Rec. 39, 4 (May 2011), 12-27

29

Page 31

Page 32

Page 34

Page 35

Page 36

Page 39

Page 40

Page 43

Page 44

Page 47

Page 48

Choosing Persistence Architecture

48

Page 49

Choosing the right NoSQL database

Summary quality attributes for popular databases

João Ricardo Lourenço, Bruno Cabral, Paulo Carreiro, Marco Vieira and Jorge Bernardino. *Choosing the right NoSQL database for the job: a quality attribute evaluation*. Journal of Big Data 2015 Springer.

49

Page 50

Techniques vs functional/non-functional properties

Gessert, Felix and Wingerath, Wolfram and Friedrich, Steffen and Ritter, Norbert. "NoSQL database systems: a survey and decision guidance. 2017. Computer Science - Research and Development. Volume 32 number 3.

50

Page 41

Functional/non-functional requirements and techniques - MongoDB, Redis, HBase, Riak, Cassandra and MySQL

Gessert, Felix and Wingerath, Wolfram and Friedrich, Steffen and Ritter, Norbert. "NoSQL database systems: a survey and decision guidance. 2017. Computer Science - Research and Development. Volume 32 number 3.

51

The storage pyramid and its role in NoSQL systems

Gessert, Felix and Wingerath, Wolfram and Friedrich, Steffen and Ritter, Norbert. "NoSQL database systems: a survey and decision guidance. 2017. Computer Science - Research and Development. Volume 32 number 3.

52

Page 54

Decision tree for requirements to NoSQL database

Gessert, Felix and Wingerath, Wolfram and Friedrich, Steffen and Ritter, Norbert. "NoSQL database systems: a survey and decision guidance. 2017. Computer Science - Research and Development. Volume 32 number 3.

54

Page 55

Polyglot persistence

55

RDBMS for every aspect for an application

Page 58

Using services instead of talking to databases

Polyglot approach - Persistence as a Service

