Database NoSQL

Docente: Matteo Mendula,
PhD student - DISI - University of Bologna

matteo.mendula@unibo.it

Course program:

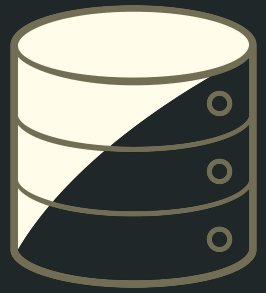
- **Lecture 01**: Introduction to NoSQL databases
- **Lecture 02**: Cassandra
- **Lecture 03**: Test + Cassandra
- **Lecture 04**: Cassandra + Neo4J
- **Lecture 05**: Neo4J
- **Lecture 06**: Neo4J + MongoDB Find
- **Lecture 07**: Test Cassandra + MongoDB Find
- **Lecture 08**: Test Neo4J + MongoDB Find
- **Lecture 09**: MongoDB Find / Aggregate
- **Lecture 10**: Test MongoDB + Redis + Real datasets

https://github.com/MatteoMendula/FitStic_noSQL

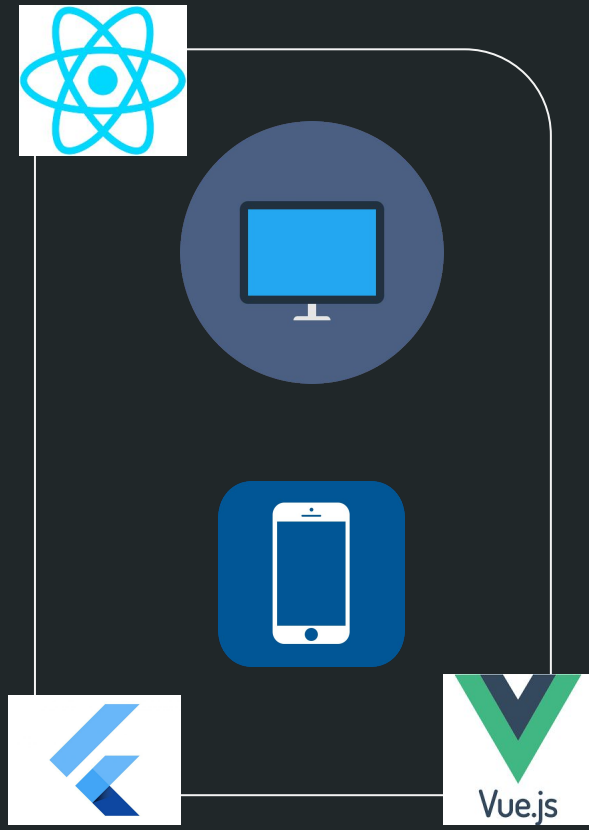
Small poll:

Preferred programming language?

3 layer architecture



3 layer architecture

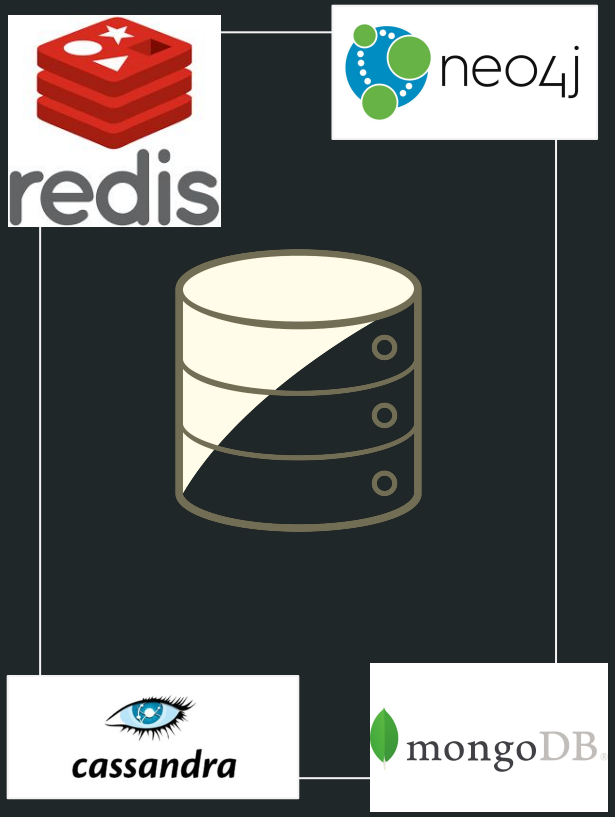


3 layer architecture

Spring Boot

node JS

3 layer architecture

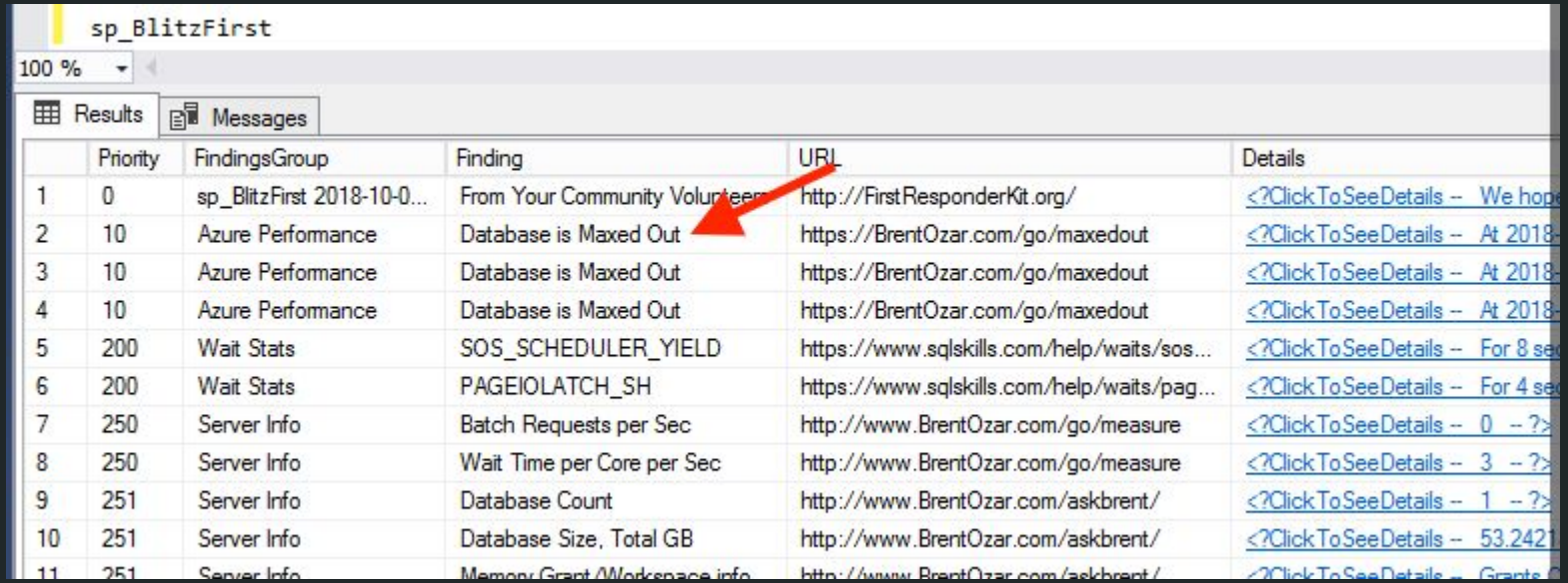


SQL

Structured Query Language (SQL) is a standardized programming language that is used to manage **relational databases** and perform various operations on the data in them.

It is particularly useful in handling **structured** data.

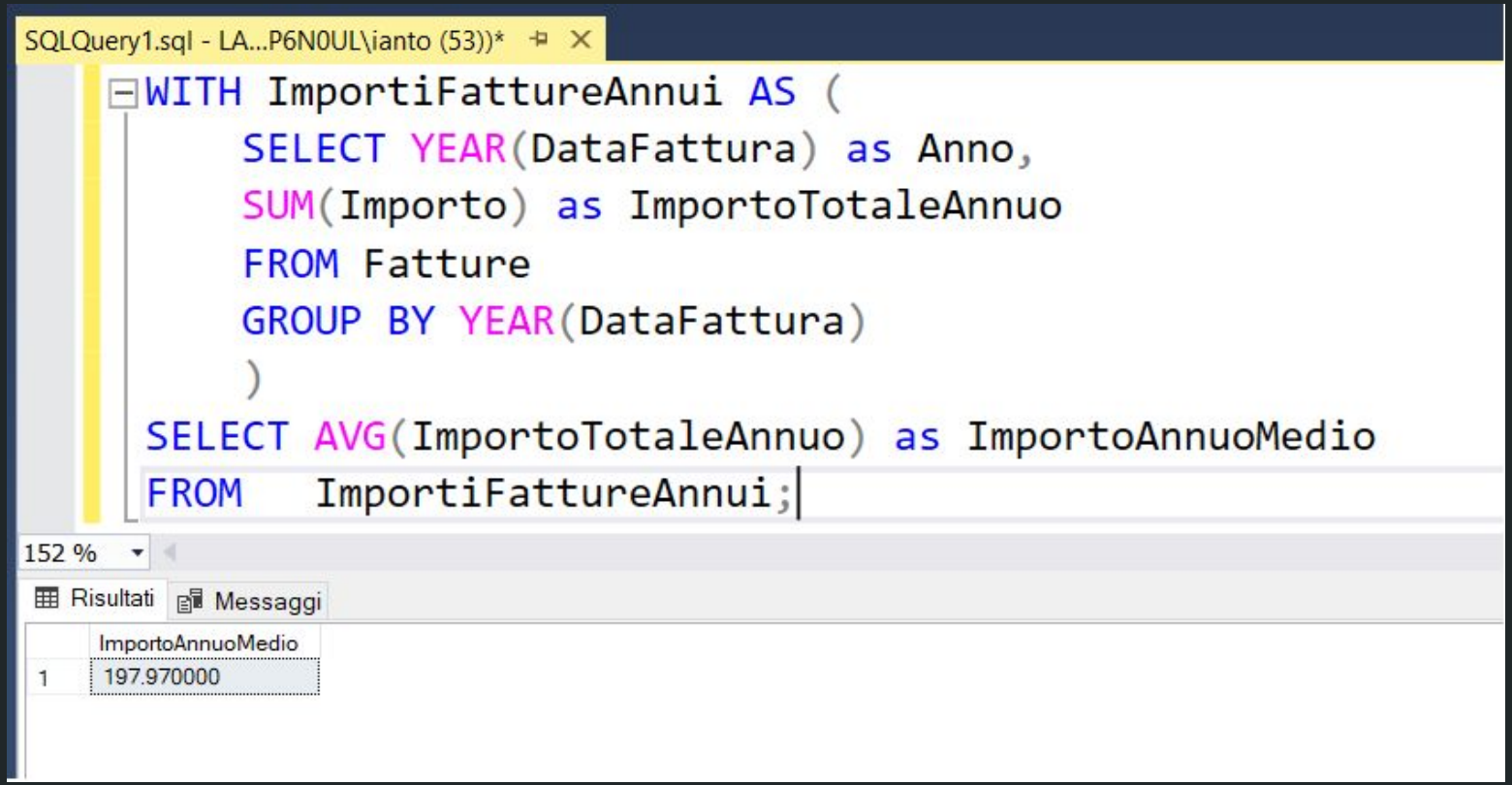
SQL



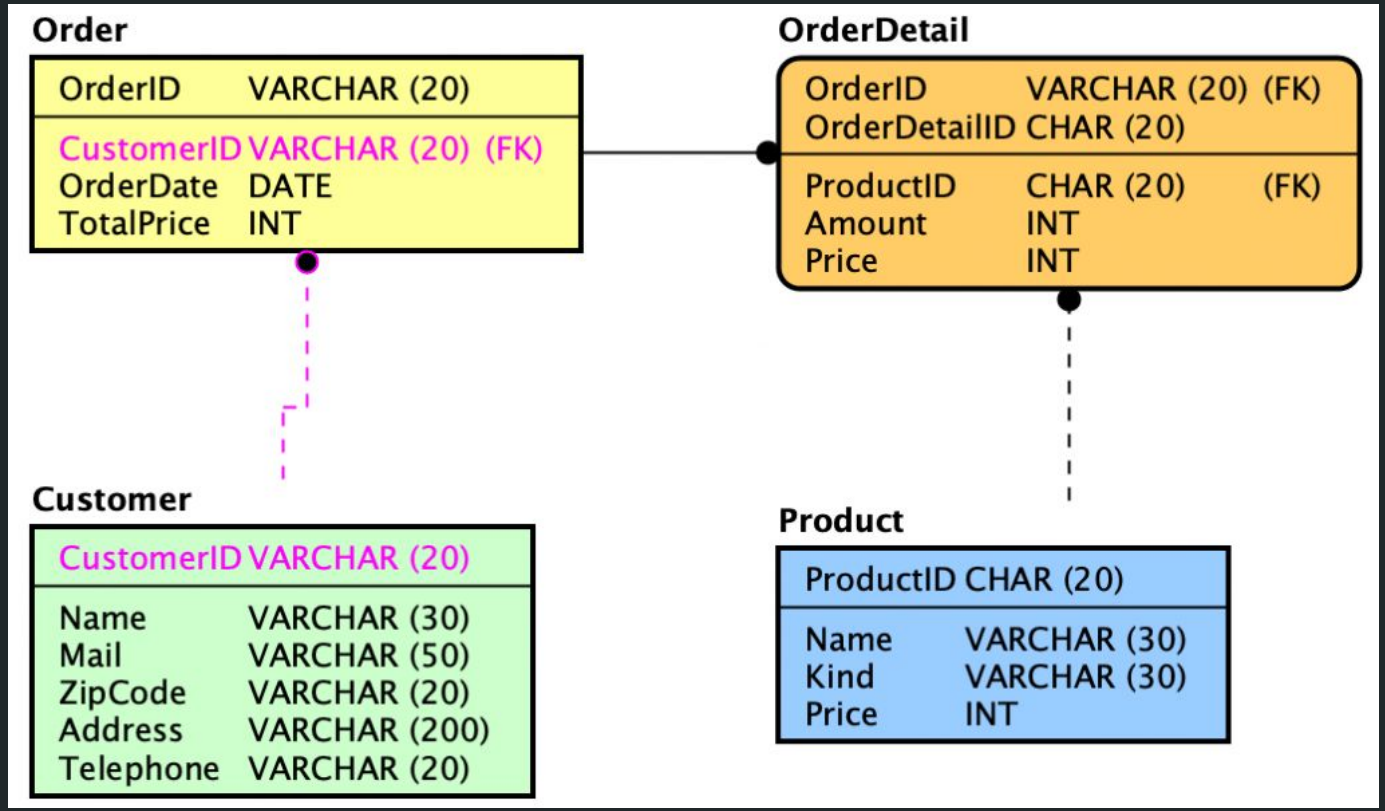
	Priority	FindingsGroup	Finding	URL	Details
1	0	sp_BlitzFirst 2018-10-0...	From Your Community Volunteers	http://FirstResponderKit.org/	<?ClickToSeeDetails -- We hope
2	10	Azure Performance	Database is Maxed Out	https://BrentOzar.com/go/maxedout	<?ClickToSeeDetails -- At 2018-
3	10	Azure Performance	Database is Maxed Out	https://BrentOzar.com/go/maxedout	<?ClickToSeeDetails -- At 2018-
4	10	Azure Performance	Database is Maxed Out	https://BrentOzar.com/go/maxedout	<?ClickToSeeDetails -- At 2018-
5	200	Wait Stats	SOS_SCHEDULER_YIELD	https://www.sqlskills.com/help/waits/sos...	<?ClickToSeeDetails -- For 8 sec
6	200	Wait Stats	PAGEIOLATCH_SH	https://www.sqlskills.com/help/waits/pag...	<?ClickToSeeDetails -- For 4 sec
7	250	Server Info	Batch Requests per Sec	http://www.BrentOzar.com/go/measure	<?ClickToSeeDetails -- 0 -- ?>
8	250	Server Info	Wait Time per Core per Sec	http://www.BrentOzar.com/go/measure	<?ClickToSeeDetails -- 3 -- ?>
9	251	Server Info	Database Count	http://www.BrentOzar.com/askbrent/	<?ClickToSeeDetails -- 1 -- ?>
10	251	Server Info	Database Size, Total GB	http://www.BrentOzar.com/askbrent/	<?ClickToSeeDetails -- 53.2421
11	251	Server Info	Memory Grant Workspace info	http://www.BrentOzar.com/askbrent/	<?ClickToSeeDetails -- Grants C

Database (RDBMS) = N * TABELLA
TABELLA = M * TUPLE

SQL


SQLQuery1.sql - LA...P6N0UL\ianto (53))*
WITH ImportiFattureAnnui AS (
SELECT YEAR(DataFattura) as Anno,
SUM(Importo) as ImportoTotaleAnnuo
FROM Fatture
GROUP BY YEAR(DataFattura)
)
SELECT AVG(ImportoTotaleAnnuo) as ImportoAnnuoMedio
FROM ImportiFattureAnnui;
152 %
Risultati
Messaggi
ImportoAnnuoMedio
1
197.970000

ER (entity relation) diagram:



Is SQL just enough?

Object–relational impedance mismatch*: it is a set of conceptual and technical difficulties that are often encountered when a relational database management system (RDBMS) is being served by an application program (or multiple application programs) written in an* ***object-oriented programming language*** *or style, particularly because objects or class definitions must be mapped to database tables defined by a relational schema.*

Person:

- Name
- Surname
- Age
- Sex

Customer:

- ?

Is SQL just enough?

Object–relational impedance mismatch*: it is a set of conceptual and technical difficulties that are often encountered when a relational database management system (RDBMS) is being served by an application program (or multiple application programs) written in an* ***object-oriented programming language*** *or style, particularly because objects or class definitions must be mapped to database tables defined by a relational schema.*

Person:

- Name
- Surname
- Age
- Sex

Customer:

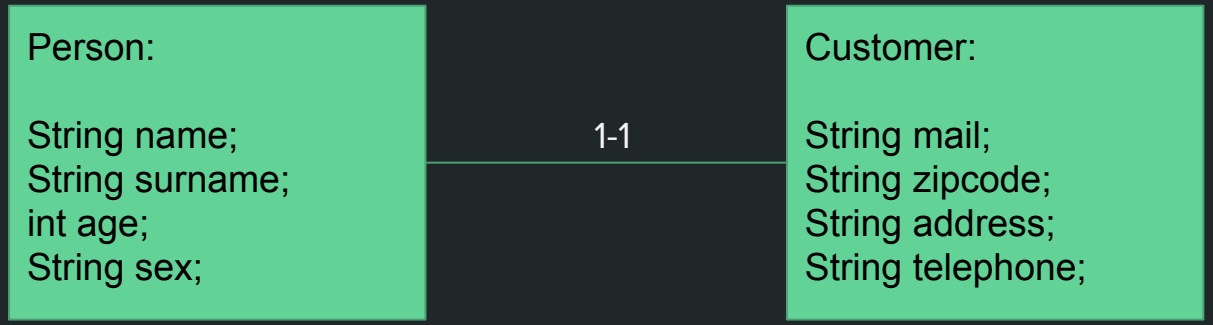
- Mail
- Zipcode
- Address
- Telephone

Person - Customer - Java implementation OOP

```java
public class Person {
  String name;
  String surname;
  int age;
  String sex;
}
```

```java
public class Customer
extends Person {
  String mail;
  String zipcode;
  String address;
  String telephone;
}
```

Person - Customer - Database mapping



Person - Customer - Database mapping

PersonID: MatteMendu9516	
Name	Matteo
Surname	Mendula
Age	26
Sex	Male

CustomerID: CustomerMatteMendu9516	
PersonID	MatteMendu9516
Mail	mattemendu@gmail.com
Zipcode	12345
Address	Via delle vie 17b

Open questions to you: new attribute?

CustomerID	PersonID	Mail	Zipcode	Address
CustomerMatteMendu9516	MatteMendu9516	mattemendu@gmail.com	12345	Via delle vie 17b
CustomerFilippoFavero1256	FilippoFavero1256	filippopippo@gmail.com	67891	Via di Filippo 2
…	…	…	…	…
CustomerPaoloBassi1246	PaoloBassi1246	paolo.b@gmail.com	23456	Via Paolo the best 12

Open questions to you: new attribute?

CustomerID	PersonID	Mail	Zipcode	Address	Premium
CustomerMatteMendu9516	MatteMendu9516	mattemendu@gmail.com	12345	Via delle vie 17b	
CustomerFilippoFavero1256	FilippoFavero1256	filippopippo@gmail.com	67891	Via di Filippo 2	
…	…	…	…	…	
CustomerPaoloBassi1246	PaoloBassi1246	paolo.b@gmail.com	23456	Via Paolo the best 12	TRUE

Open questions to you: new attribute?

CustomerID	PersonID	Mail	Zipcode	Address	Premium
CustomerMatteMendu9516	MatteMendu9516	mattemendu@gmail.com	12345	Via delle vie 17b	?
CustomerFilippoFavero1256	FilippoFavero1256	filippopippo@gmail.com	67891	Via di Filippo 2	?
…	…	…	…	…	?
CustomerPaoloBassi1246	PaoloBassi1246	paolo.b@gmail.com	23456	Via Paolo the best 12	TRUE

Open questions to you: new attribute?

CustomerID	PersonID	Mail	Zipcode	Address	Premium
CustomerMatteMendu9516	MatteMendu9516	mattemendu@gmail.com	12345	Via delle vie 17b	NULL
CustomerFilippoFavero1256	FilippoFavero1256	filippopippo@gmail.com	67891	Via di Filippo 2	NULL
…	…	…	…	…	NULL
CustomerPaoloBassi1246	PaoloBassi1246	paolo.b@gmail.com	23456	Via Paolo the best 12	TRUE

Open questions to you: new attribute as array?

CustomerID	PersonID	Mail	Zipcode	Address	WishList
CustomerMatteMendu9516	MatteMendu9516	mattemendu@gmail.com	12345	Via delle vie 17b	
CustomerFilippoFavero1256	FilippoFavero1256	filippopippo@gmail.com	67891	Via di Filippo 2	
…	…	…	…	…	
CustomerPaoloBassi1246	PaoloBassi1246	paolo.b@gmail.com	23456	Via Paolo the best 12	[order_12, order_13, …]

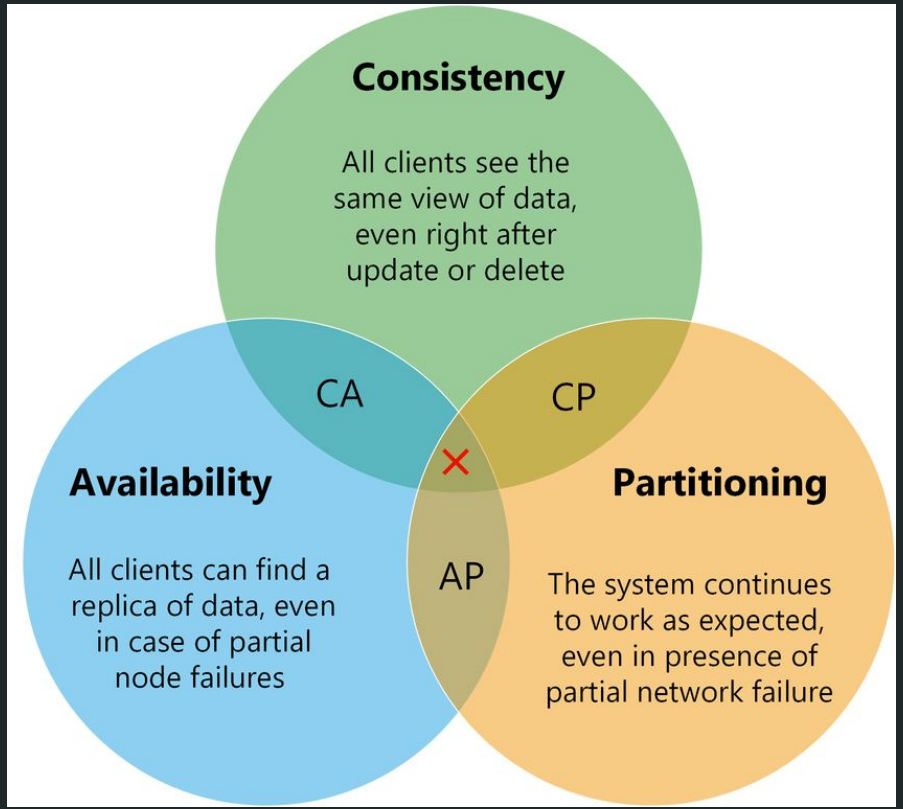
Open questions to you: new attribute as array?

OrderID	**CustomerID**	**…**
Order_12	CustomerMatteMendu9516	…
…	…	…
Order_99	CustomerPaoloBassi1246	…

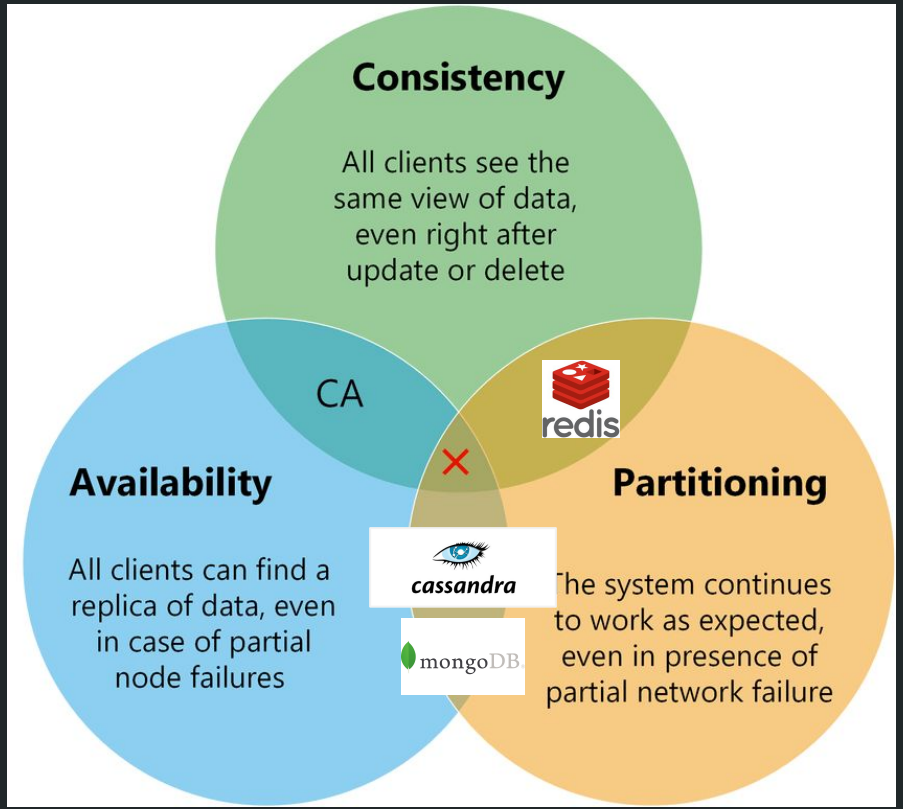
Object–relational impedance mismatch

Objects (instances) reference one another and therefore form a **graph in the mathematical** sense. Relational schemas are, in contrast, **tabular and based on relational algebra**, which defines linked heterogeneous tuples (groupings of data fields into a "row" with different types for each field).

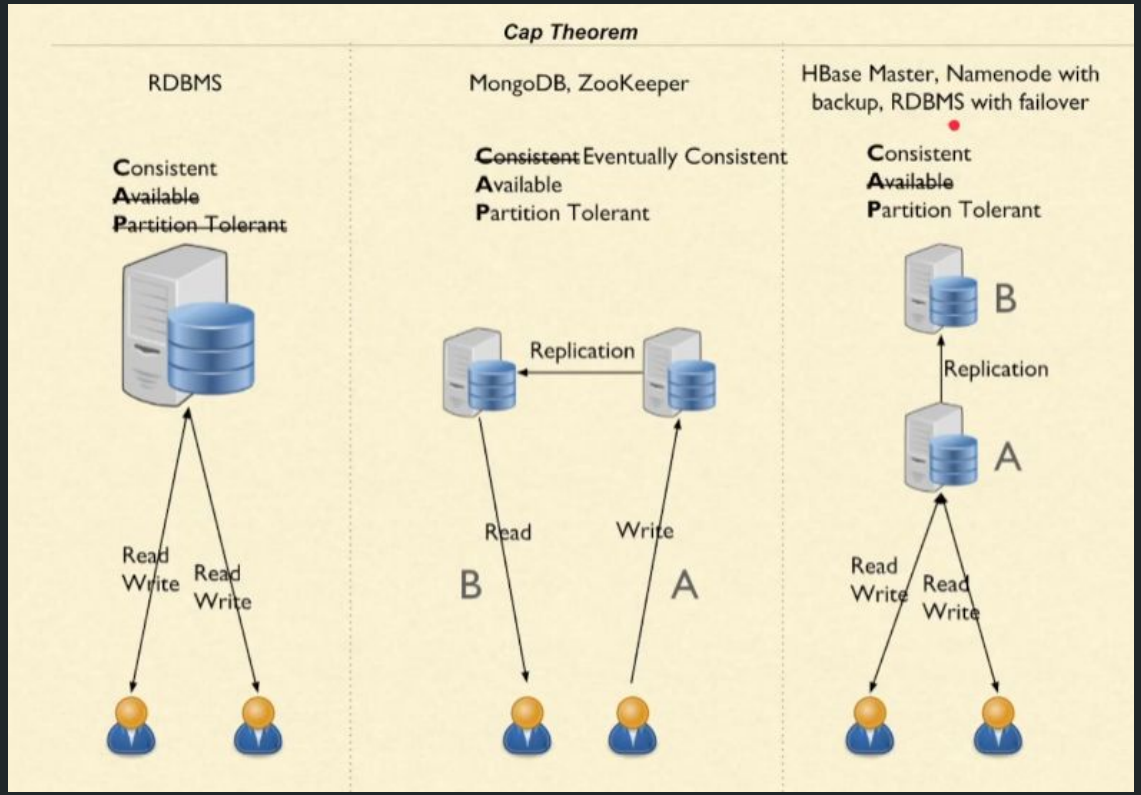
CAP theorem


Consistency
All clients see the same view of data, even right after update or delete
CA
CP
Availability
All clients can find a replica of data, even in case of partial node failures
AP
Partitioning
The system continues to work as expected, even in presence of partial network failure

CAP theorem


Consistency
All clients see the same view of data, even right after update or delete
CA
redis
Availability
All clients can find a replica of data, even in case of partial node failures
cassandra
mongoDB.
Partitioning
The system continues to work as expected, even in presence of partial network failure

Eventual consistency



Not only SQL

A NoSQL (originally referring to "non-SQL" or "non-relational") database provides a mechanism for storage and retrieval of data that is modeled in means other than the tabular relations used in relational databases. Such databases have existed since the late 1960s, but the name "NoSQL" was only coined in the early 21st century. Triggered by the needs of Web 2.0 companies. NoSQL databases are increasingly used in big data and real-time web applications. NoSQL systems are also sometimes called **Not only SQL** to emphasize that they may support SQL-like query languages or sit alongside SQL databases.

Type	Notable examples of this type
Wide Column Store	Azure Cosmos DB, Amazon DynamoDB, Bigtable, Cassandra, Google Cloud Datastore, HBase, Hypertable, ScyllaDB
Graph database	Azure Cosmos DB, AllegroGraph, ArangoDB, InfiniteGraph, Apache Giraph, MarkLogic, Neo4J, OrientDB, Virtuoso
Document store	Azure Cosmos DB, ArangoDB, BaseX, Clusterpoint, Couchbase, CouchDB, DocumentDB, eXist-db, IBM Domino, MarkLogic, MongoDB, Qizx, RethinkDB, Elasticsearch, OrientDB
Key–value cache	Apache Ignite, Couchbase, Coherence, eXtreme Scale, Hazelcast, Infinispan, Memcached, Redis, Velocity

Data model

Post instagram

FB profile and friendship network

Cassandra?

https://phoenixnap.com/kb/install-cassandra-on-windows

https://phoenixnap.com/kb/create-drop-alter-and-truncate-tables-in-cassandra

https://www.youtube.com/watch?v=8AoWhAhfuYk