Relational VS Non-Relational Databases (Distinct CAP/Brewer's Theorem Trade-offs)

Relational (CP)

Data follows a predefined schema with strict rules of structure fulfilled by validating keys, columns and data types.

Minimizes repetition and ensures data integrity

Key concepts:

The keys **Data integrity validators**

- Primary key (**PK**): A column (or set) that uniquely identifies each row in a table. It must be unique and not null;
- Foreign key (**FK**): A column in one table that references the primary key in another table, establishing a relationship between the two.

Indexes **Boost query speed**

- Only used on huge data sets;
- Only benefit read operations;
- Relate to a single table.

Views **Reusable logic**

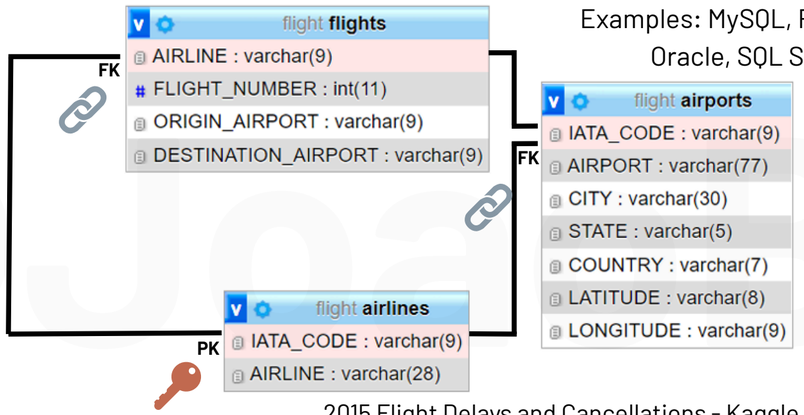
- Allows CRUD operations directly;
- Only make sense when they cross an acceptable number of tables

ACID Properties

Atomicity: All operations succeed or none.
Consistency: Valid data state transitions.
Isolation: No interference.
Durability: Data is permanently saved.

Scaling Vertical, by replacing a resource with another with higher or lower capacity

Examples: MySQL, PostgreSQL, Oracle, SQL Server.



2015 Flight Delays and Cancellations - Kaggle

Non-Relational (AP)

Data is stored in flexible formats without a fixed schema.

Prioritizes availability and flexibility

Key concepts:

It's a collection - a set of documents/variables


```
{
    "_id" : ObjectId("61c5bcf244128795cb4022fd"),
    "id" : "2",
    "firstname" : "Hendrik A.",
    "surname" : "Lorentz",
    "born" : "1853-07-18",
    "died" : "1928-02-04",
    "bornCountry" : "the Netherlands",
    "bornCountryCode" : "NL",
    "bornCity" : "Arnhem",
    "diedCountry" : "the Netherlands",
    "diedCountryCode" : "NL",
    "gender" : "male",
    "prizes" : [
        {
            "year" : "1902",
            "category" : "physics",
            "share" : "2",
            "motivation" : "\"in recognition of the extraordinary service they
rendered by their researches into the influence of magnetism upon radiation
phenomena\"",
            "affiliations" : [
                {
                    "name" : "Leiden University",
                    "city" : "Leiden",
                    "country" : "the Netherlands"
                }
            ]
        }
```


Example document in a NoSQL collection (MongoDB-style)

Structure **Schema-less and flexible**

- Data stored as JSON-like documents, key-value pairs, or wide-column formats.
- No need to define a fixed schema in advance.

Collections

- Group of documents, similar to a table but without a rigid structure.

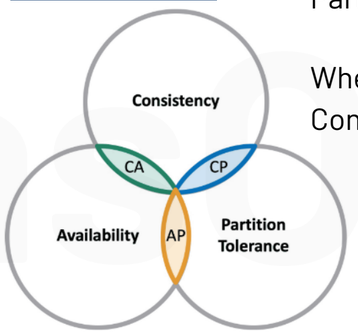
Scaling

Horizontal, by adding more servers (nodes).

Examples: MongoDB, Couchbase, Cassandra, DynamoDB.

CAP Theorem

A distributed system can't guarantee Consistency, Availability, and Partition Tolerance simultaneously.



When partitions occur, systems are forced to trade off either Consistency or Availability, depending on the specific requirements.

CP - System remains consistent, but may be unavailable.

AP - System remains responsive, but may return stale/inconsistent data .