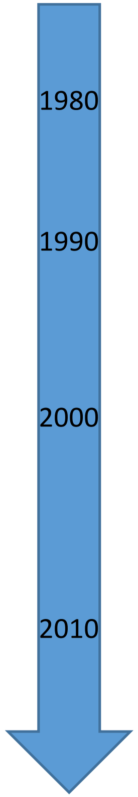
Introduction to MongoDB

Requirements of Modern Application

- Big Data
    - Lots of storage
    - Lots of access
- Ever Changing Features
    - Simpler data models
- Flexible deployments
    - Cloud-enabled

Relational Databases



Persistence

Reporting

Integration

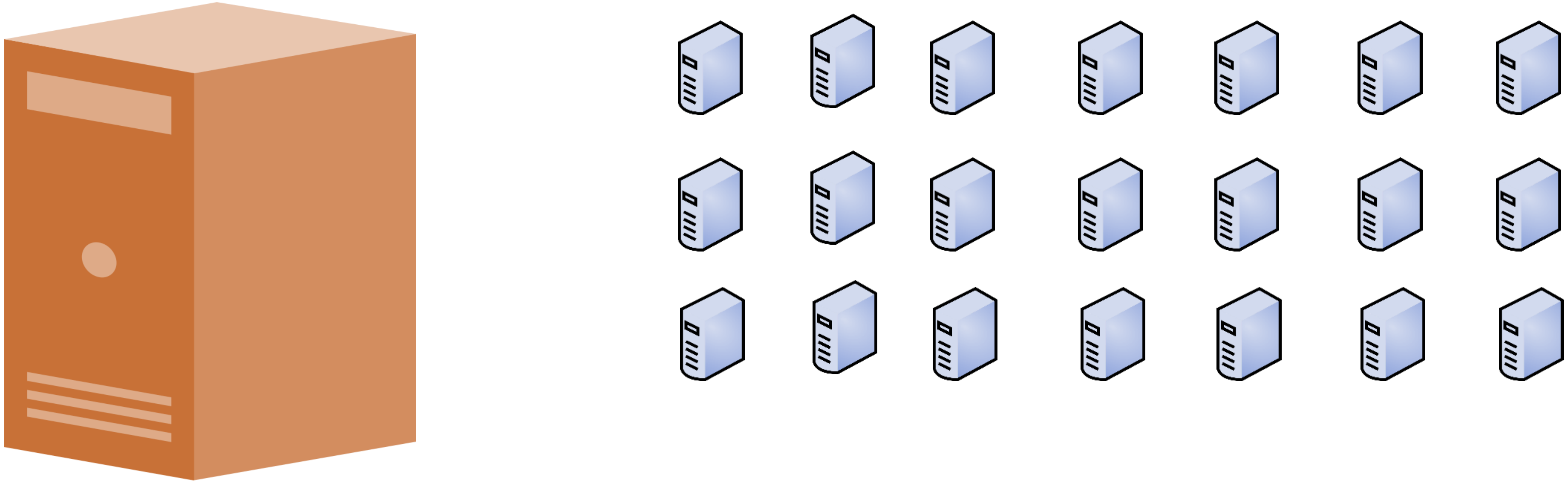
Transactions

SQL

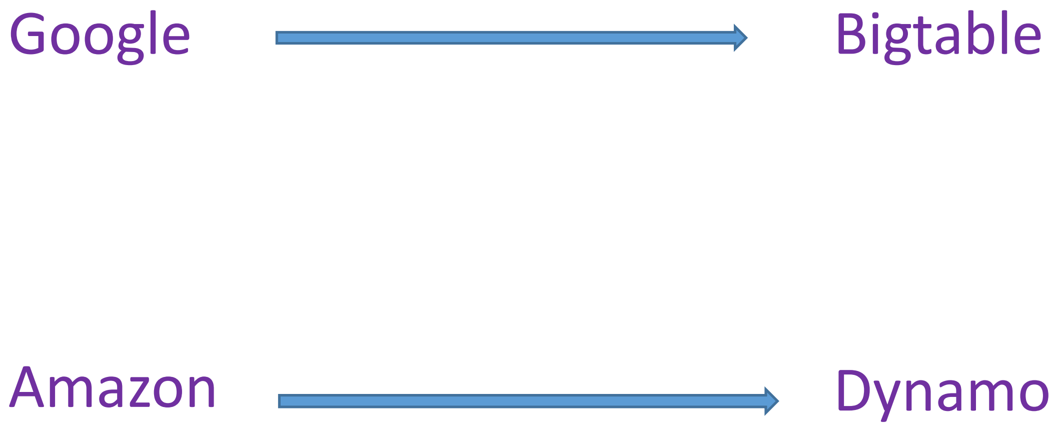
Relational Database

- Impedance Mismatch

The age of Big Data



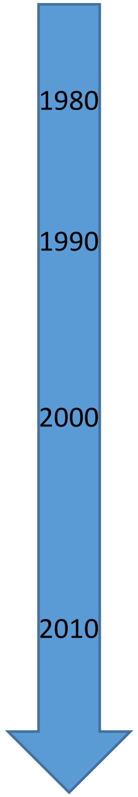
Not a Relational Database



Google → Bigtable

Amazon → Dynamo

NoSQL



Non-relational

Open-source

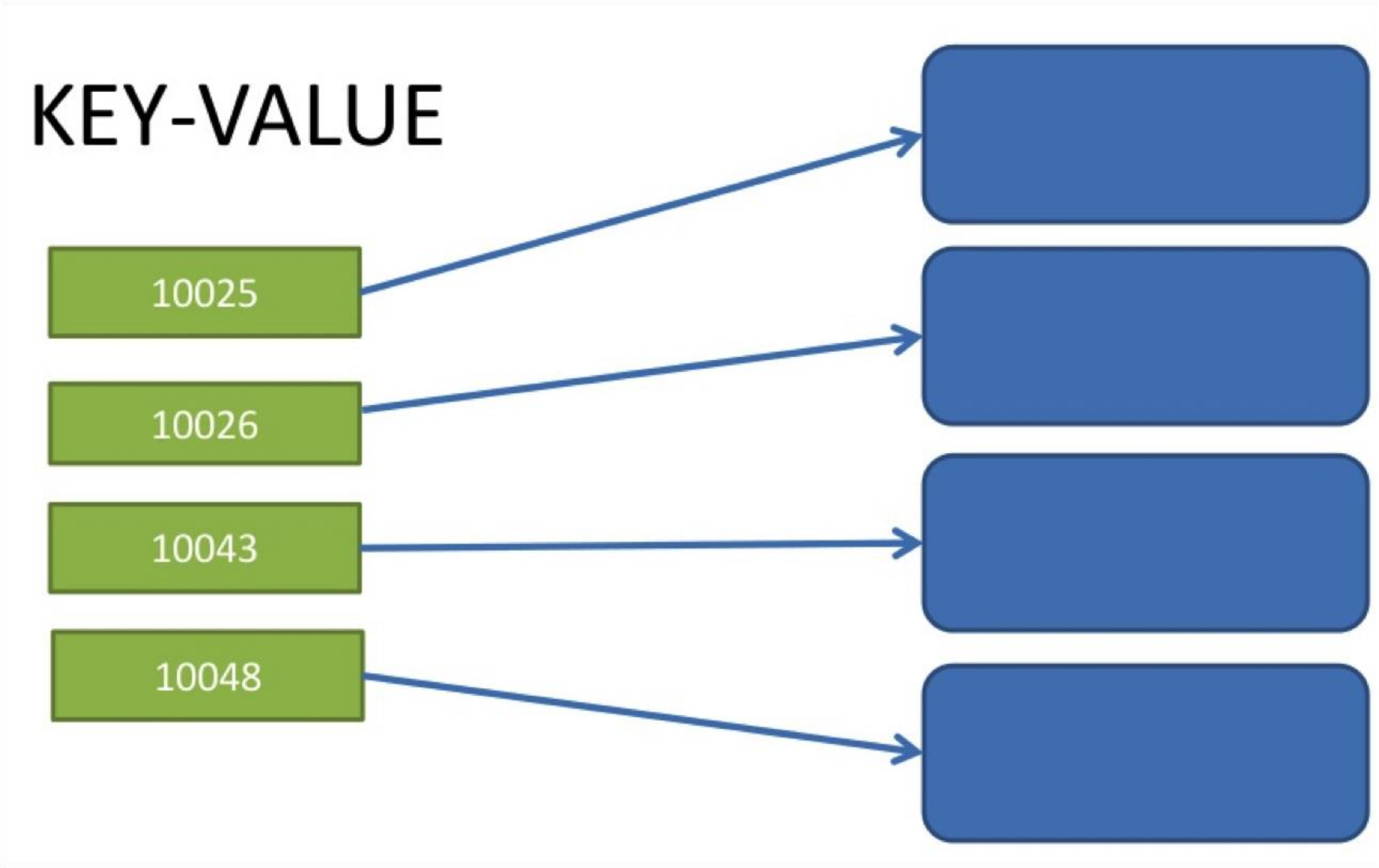
Cluster friendly

Schema-less

Data Model

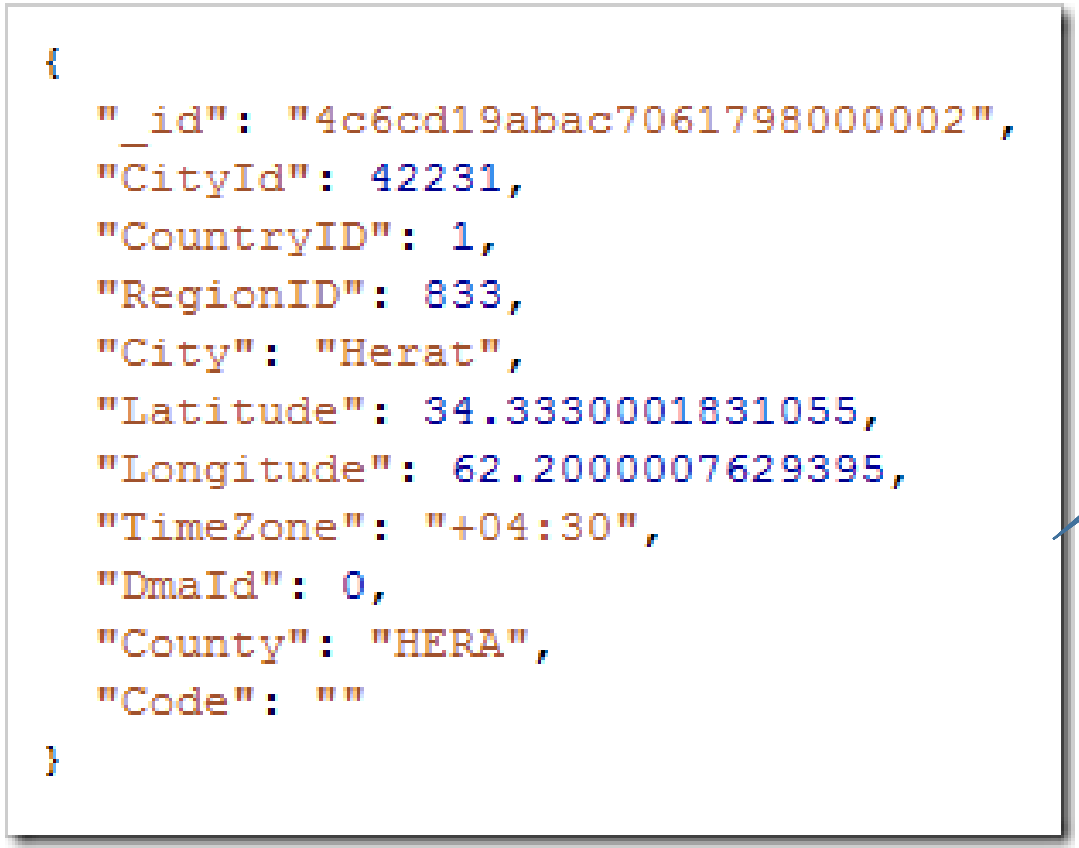
- Key-value
- Document
- Graph
- Column-family

Key-value



Source: http://bigdata-blog.com/wp-content/uploads/2015/05/keyvalue.png

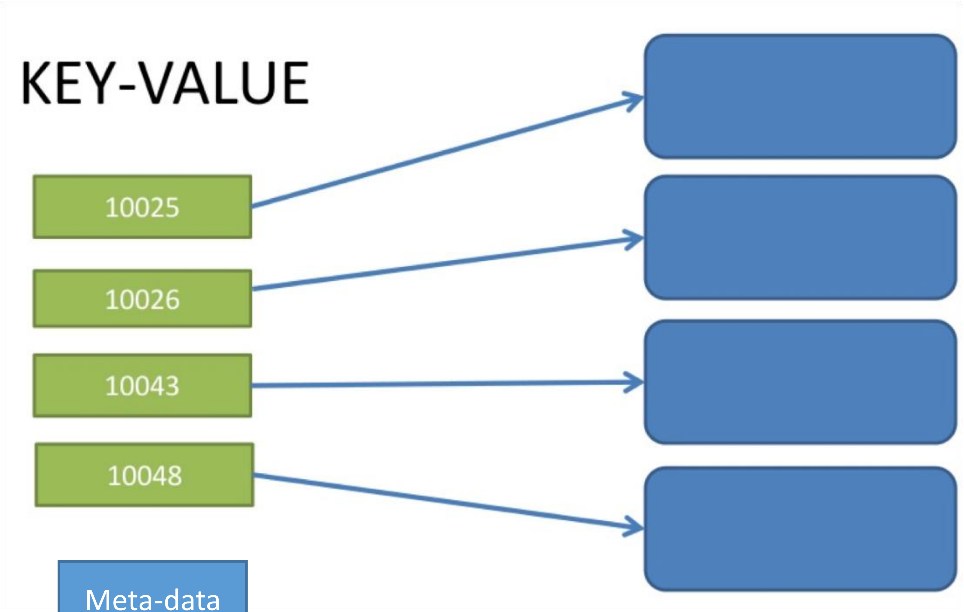
Document



```
{
  "_id": "4c6cd19abac7061798000002",
  "CityId": 42231,
  "CountryID": 1,
  "RegionID": 833,
  "City": "Herat",
  "Latitude": 34.3330001831055,
  "Longitude": 62.2000007629395,
  "TimeZone": "+04:30",
  "DmaId": 0,
  "County": "HERA",
  "Code": ""
}
```

No Schema

Source: http://new.mongovue.com/wp-content/uploads/2010/11/image2.png

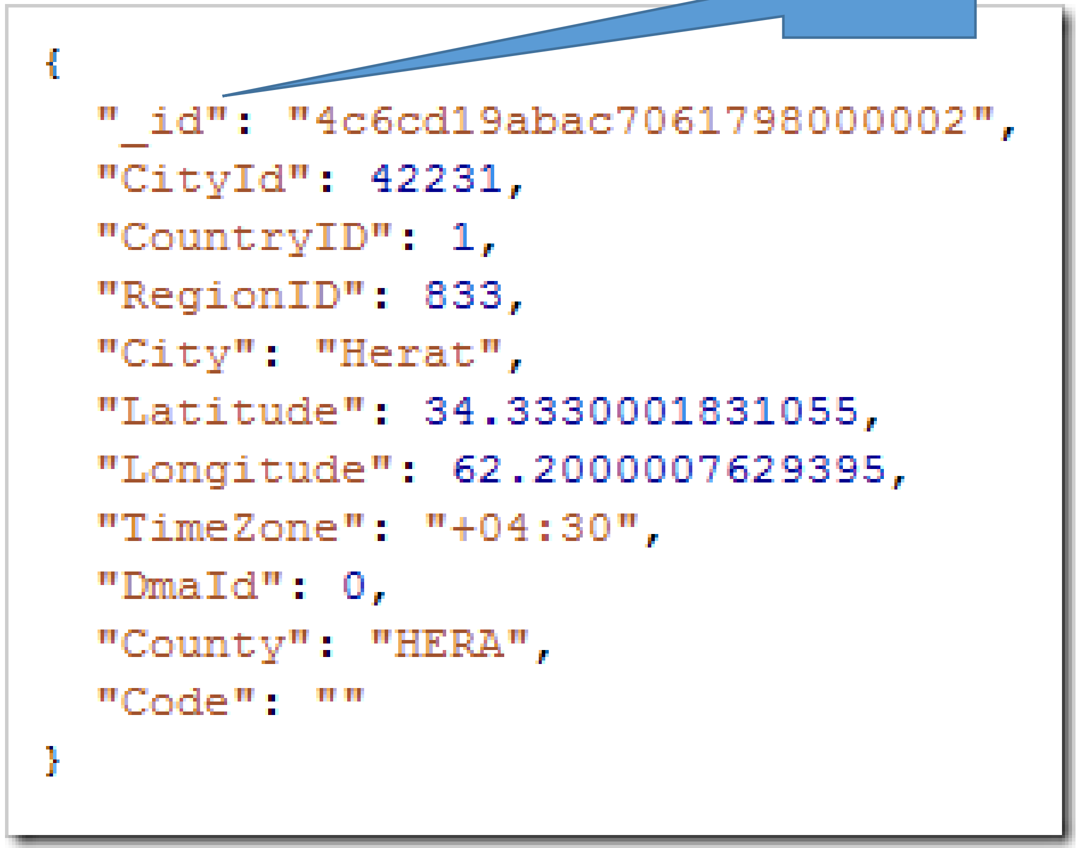


Meta-data

Customer_id = 123132

Document

Key



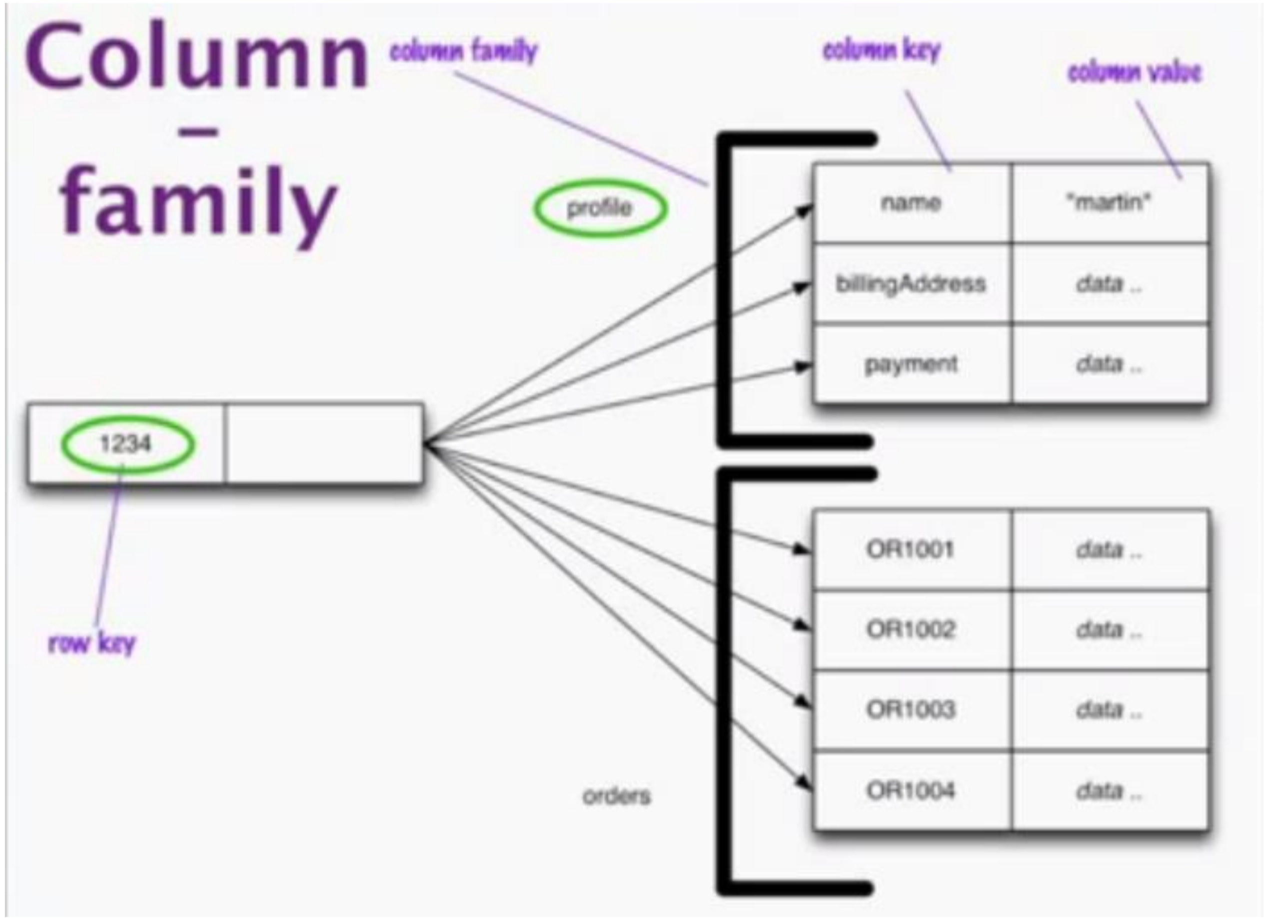
```
{
    "_id": "4c6cd19abac7061798000002",
    "CityId": 42231,
    "CountryID": 1,
    "RegionID": 833,
    "City": "Herat",
    "Latitude": 34.3330001831055,
    "Longitude": 62.2000007629395,
    "TimeZone": "+04:30",
    "DmaId": 0,
    "County": "HERA",
    "Code": ""
}
```

Source: http://new.mongovue.com/wp-content/uploads/2010/11/image2.png

Aggregate

- Aggregate-oriented databse
    - Value
    - Document
- Aggregate is the single unit that we access

Column-family



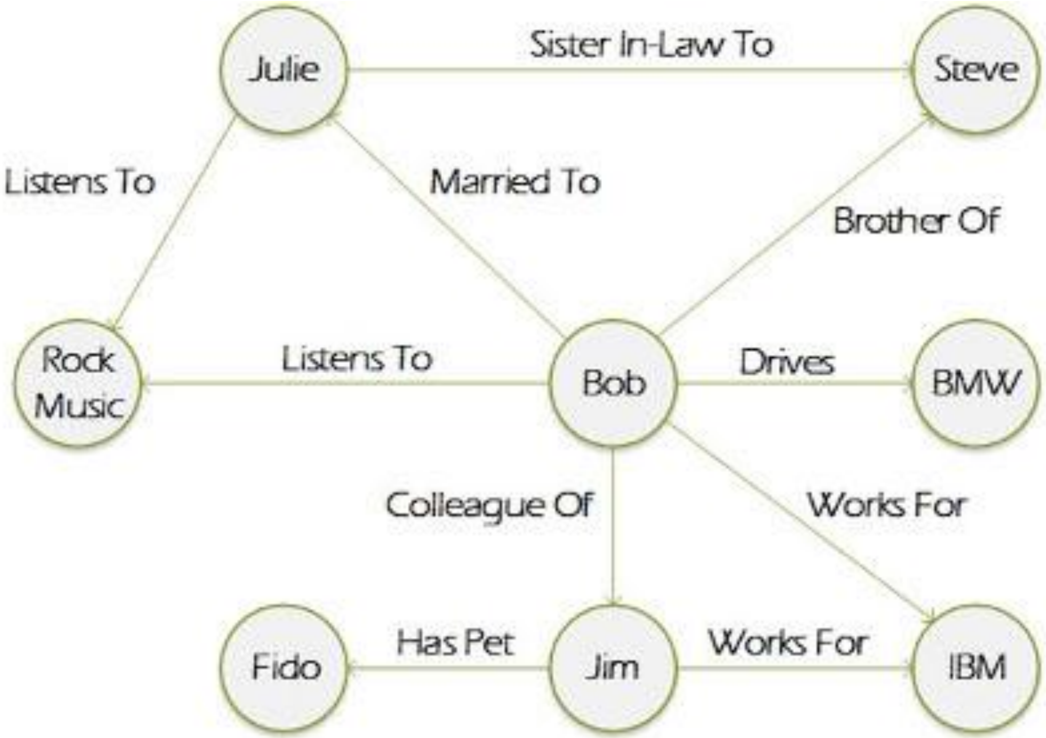
name	"martin"
billingAddress	data ..
payment	data ..

OR1001	data ..
OR1002	data ..
OR1003	data ..
OR1004	data ..

Source: https://csetech.files.wordpress.com/2013/06/colfamily.png

Aggregate model on Clusters

Graph



Source: http://itknowledgeexchange.techtarget.com/overheard/files/2014/01/Graph-database-sketch.jpg

NoSQL and Consistency

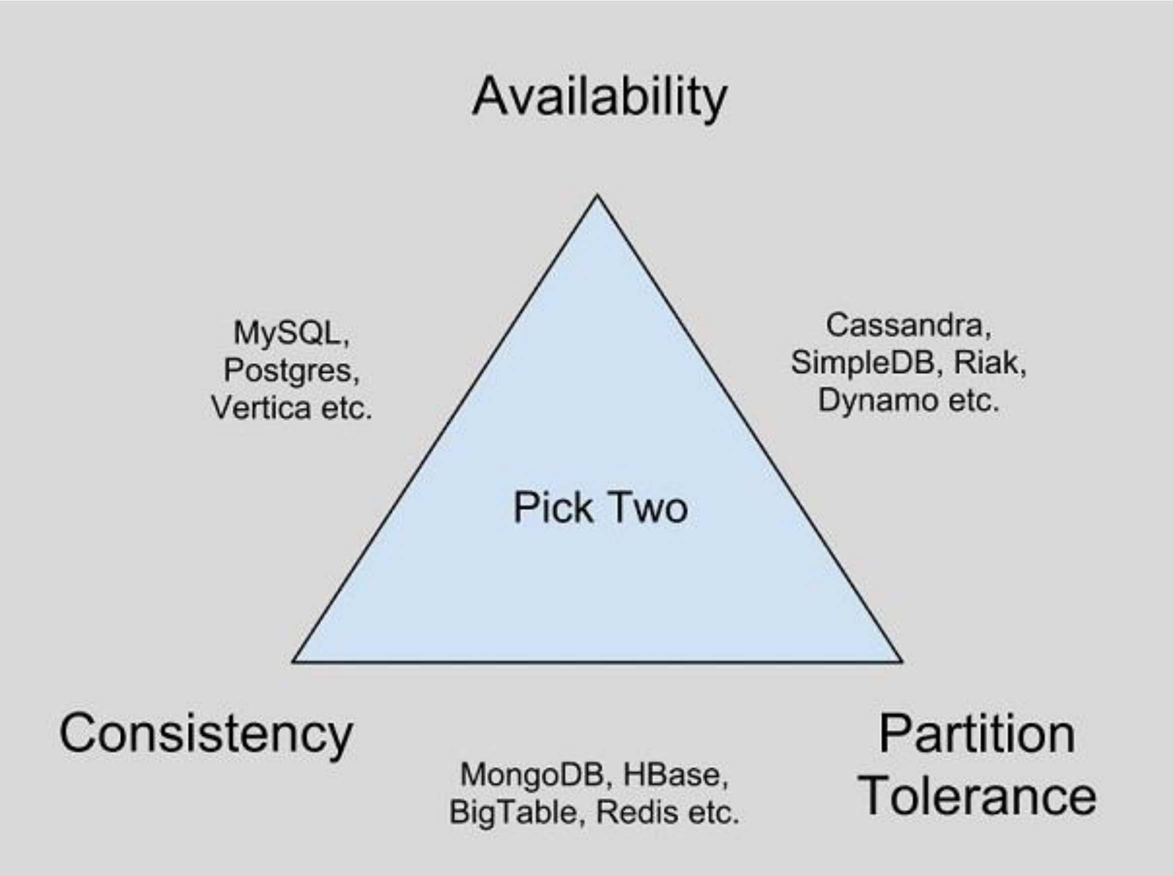
- ACID?

NoSQL and Consistency

- Graph
- Aggregation-oriented
    - Aggregate is the boundary
- Version stamps

NoSQL and Consistency

- Logical
- Replication


Availability
MySQL, Postgres, Vertica etc.
Cassandra, SimpleDB, Riak, Dynamo etc.
Pick Two
Consistency
MongoDB, HBase, BigTable, Redis etc.
Partition Tolerance

Source: http://wp1.cdn.brainsins.com/en/wp-content/uploads/2013/10/cap-theorem.png

Trade-offs

- Consistency vs. Response-time

When and Why to use a NoSQL database

- Large scale data
  - Mobile Cloud Applications
- Easier Development
  - Natural Aggregate
- Web Services/Applications enabling access to data