NoSQL - MongoDB

DIS Exercise Course

Relational Dominance

1980 1990 2000

Persistent Data

Application Integration

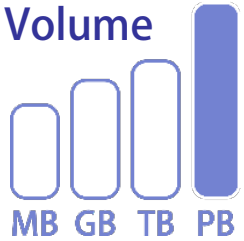
Mostly Standard

Concurrency Control

Big Data

2000 2010 2020

Volume



20 hours of video uploaded on every min YouTube

Global Online Population

2012
2+ Billion

2000
360 Million

Velocity

400 Million a Day in 2012

reached 1 billion users in 2012

Variety

Global Data Storage in Exabytes

2012
2,837

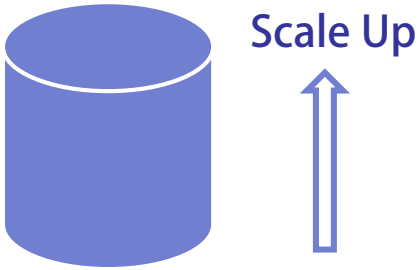
2010
1,227

2005
130

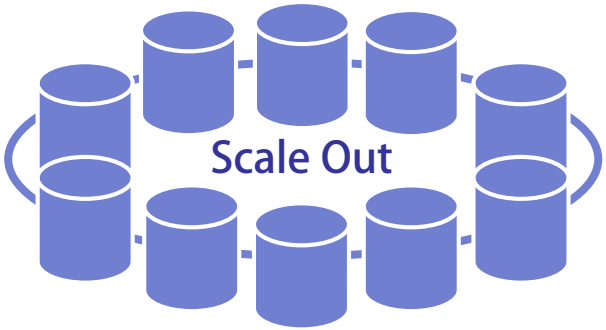
What's the problem with RDBs?

RDBs designed to run on a single machine.

You need to buy a bigger machine to scale vertically.



But it's cheaper and more effective to scale horizontally by buying lots of machines.



Google and Amazon were both early adopters of large clusters, and both eschewed RDBs.

Amazon DynamoDB

Google Bigtable

So Now We Have NoSQL Databases

Non-Relational

Cluster-Friendly

Schema-Less

Open-Source

21st Century Web

NoSQL Data Models

Key-Value

Column-Family

Document

Graph

But ...

The relational model is still relevant !

ACID Transactions

Tools

Familiarity

(from hu"mongo"us)

A scalable, high-performance, schema-free, open-source document-oriented database.

- JSON-Style Documents
- Auto-Sharding
- Replication & High Availability
- Full Index Support
- Fast In-Place Updates
- Rich, Document-Based Queries
- Map/Reduce

What's a Document Anyway ?

```
{
        _id : "2016711",
        title : "Star Trek: Nemesis",
        year : "2002",
                                ...
        genre : [
                "Action",
                "Sci-Fi",
                "Thriller"
        ],
        releases : [
                {
                        country : "USA",
                        date : ISODate("2002-12-08T23:00:00Z")
                },
                {
                        country : "Germany",
                        date : ISODate("2003-01-15T23:00:00Z")
                }
        ]
}
```

Learn More ...

Master Project NoSQL
WiSe 2013/2014 & SoSe 2014