NoSQL

<snappy title here>

Sander van Vliet
@SanderMvanVliet

What I hope you can learn today

- What NoSQL is
- Why your life just got complicated

Something about applications

- What is important?
- So what about the data?

Storage as we used to do it

- Text files
- Binary files
- XML
- SQL Databases

The RDBMS

Pros:

- One-Stop deal
- Common query & definition language
- Back-up and maintenance
- Tool ecosystem

Cons:

- Big!
- Expensive!
- Limited scalability!

The RDBMS and your app

- It gives you ACID
- Strong schemas
- Not a perfect fit

Back to square one

- Goal: Storing data
- Ideally: Same format as in our app
- Blazing performance
- And reliable

The silver bullet called NoSQL

- Some say:
  - NO! SQL
  - Not Only SQL
- A class of database systems

So how does NoSQL help us?

- Storage that matches your entities
- Schemaless
- Quicker development

Data storage: the NoSQL way

- Tailored to your app
  - Key/Value
  - Document
  - Graph
  - Column Family
  - Event store
- Aggregates make life easier

Oh and I also want…

- Durability
- Scalability
- Reliability

You can't have it all

The CAP theorem

- Consistency
- Availability
- Partition tolerance

How to deal with this?

- Distributing data
  - Sharding
- Replication
  - Master/Slave
  - Peer-to-peer

Data consistency

- No ACID means no guarantees
- Do it yourself
  - Write and Read-Write conflicts
  - BASE

How about durability?

- Can your app tolerate failures?
- Replication factor
  - Depends on the type of replication
  - W > N/2
- Data versioning
  - Stamps, clocks and vectors

Where are we now?

- NoSQL gives you great flexibility
- With great power comes great responsibility

Why NoSQL may not be for you

- NoSQL systems are new
- The RDBMS just fits
- Don’t (always) go all out

So, questions?

- I hope so!

Key/Value stores

- MongoDB
- Riak
- Azure Table Storage (Microsoft)
- BigTable (Google)
- BerkelyDB
- Memcached
- DynamoDB (Amazon)
- Cassandra (Apache)
- HyperDex

Document stores

- CouchDB (Apache)
- MongoDB
- Oracle NoSQL
- Cassandra (Apache)
- RavenDB

Graph stores

- Neo4j
- OrientDB
- DB2 (IBM)
- Openlink Virtuoso

Column-Family stores

- Cassandra (Apache)
- HBase
- BigTable (Google)

Further reading

- Google!
- Martin Fowler
- http://nosql.mypopescu.com
- http://nosql-database.org

Data structure versioning

- Update an entity (schema) on update
- No need to shutdown the entire system on deployment
- Drawback: you need to build it

Sharding

Contact details

- https://barad-dur.nl/
- @SanderMvanVliet