A short MongoDB Tutorial

Franz Papst

28.7.2021

What is MongoDB?

- NoSQL Database
  - document-oriented database
- Name derived from the word *humongous*
- Rather drastic shift from relational databases
  - no schemas, joins, foreign keys, …
- A good way to store large amounts of data

Why MongoDB?

- Simple
  - setup (at least for single instances)
  - usage
- Flexible
  - requires no fixed schema
- Fast
  - writing compared to relational databases
- Feature rich
  - e.g. data aggregation
- Horizontal scalable
  - creating a cluster
- Free and Open Source
  - Community Version
  - MongoDB Enterprise is really expensive
  - MongoDB Atlas is a hosted and affordable enterprise version

NoSQL

- Weaker emphasis on data consistency compared to relational databases
  - for uses-cases where consistent data is not that important
  - but faster because of this
- Less rigid than relational databases
  - does not need a strict schema
  - flexible and easily adaptable to new data
- Gained traction in the late 2000s
  - around the same time Social Media got big
- Different flavours
  - document-oriented database
  - wide-column store
  - graph databases (?)

What is MongoDB good at?

- Writing values into the database
  - quite fast doing so
  - not as fast as Cassandra though
- Dealing with large amounts of (unstructured) data
  - think of MongoDB as a huge JSON file
- Being flexible
  - documents in a collection do not need to have the same data model
  - data structure can change over time ➡ better equipped for the future

What is MongoDB not so good at?

- Finding data
  - especially searching by values
  - indices help though
- Updating data
  - because you have to find data first
- Data consistency
  - keep track of it yourself ➡ prone to errors

MongoDB for Time Series Data

- MongoDB is well suited for time series data
- Bucketing
  - splitting data into different buckets
- Native aggregation features
- MongoDB 5.0 added native time series support

Hands-on

Show, don’t tell: https://mongoplayground.net/p/o7W_XoBqZzi

Check repo for test data and sample queries
https://github.com/complexity-science-hub/skillup/tree/master/mongodb