Persistence

Code 401 (Class 09)

Persistence

Our http and chat servers to this point have been operating solely in memory.

Every time we stop the server, we lose everything

- Chat History
- Login History
- Our Notes List

Web applications become really interesting when they are able to manage and maintain states of data, even in a stateless environment.

Data Storage - What and Why

- Login and User Information
- Order History
- Todo List
- Forum Posts
- Chat History -- this is what differentiates Slack and HIpChat from IRC
- Website Metadata (Google)

Data Storage - How

At the end of the day, data is ultimately stored in a file.

- “comma”,”separated”,”values”
- Fixed Length Records
- Berkley DB (the first key/value data store)
- SQL Databases (Oracle, SQL Server, mySQL, Postgres)
  - Also known as “Relational” databases.
  - Great at dealing with interrelated data and indexing
- NOSQL Databases
  - More advanced key/value store
  - Deals with actual objects and documents

Data Storage - Big Data

You'll hear this term a lot, and everyone loves to think they are "BIG DATA"

Enormous data stores …

- Google's index of every website in the world
- Amazon's customer, order, and product information datasets
- NYSE transaction history

Wickedly Fast

- Billions of results searched and indexed in microseconds
- Smart, distributed caches
- Striped and

Secure and Redundant

- Failovers on failovers
- Encryption and locking

Read @ AWS Database types: https://aws.amazon.com/

Data Storage - Modeling

The nugget of all of this is in how we model our data to most effectively be stored and retrieved by our services

For our notes database, we are going to use a simple flat (text) file to store all of the notes, and will update our RESTful APIs to use the fs module to read and write to it.

Decisions to be made:

- What should the shape of each note be (do we need to revisit our note model?)
- Should each note be in a separate file or one?
- If one, is it an array of notes?
- ...or is it a keyed map?
- If we wanted to search the database, how would that change our answers?
- Do we need to manage an index?