JSON Stores Lab (MongoDB)

Ioana Manolescu
INRIA Saclay
ioana.manolescu@inria.fr
http://pages.saclay.inria.fr/ioana.manolescu/

M2 Data and Knowledge
Université de Paris Saclay

MongoDB: setting up

- Install MongoDB.
- Create a directory for the server and launch it. Write down the command for launching it; on which port it runs?
- Import the document moviepeople-10.json into the server
- Launch a client (mongo shell), retrieve all the documents by asking a query in the client.

→1. Keep track of: the commands you used

MongoDB: import, query

- Import the documents moviepeople-3000.json and cities.json into the server
- In the mongo shell client, write queries for finding:
  2. The person named Anabela Teixeira
  3. The birthplace of Steven Spielberg
  4. The number of people born in Lisbon
  5. The people taller than 170 cm
  6. The names of people whose information contains "Opera"

  → Keep track of: the queries and answers

MongoDB: more querying

7. For each movie person whose birth place is known, find the lattitude, longitude and population of that city (if that information exists in the city document)
   – You may use functions, several commands etc.

→ Keep track of: your commands, functions, queries etc.

MongoDB replication

- Create working directories for **3** MongoDB servers
- Create a replication set for a collection named **small-movie**
- Launch 3 MongoDB servers (in different shells). The server program is **mongod**. Leave those shells alone.
- Connect a client (**mongo**) to one server. Through the client, initialize the replication: add the other replica server, and the arbiter.
- Identify the master from the outputs in the servers' shell and by requesting replica set information from the servers.

→ 8. Keep track of: your commands; server output

MongoDB replication, sharding

Replication (continued)

Import moviepeople-10.json through the master; note the output of the two other servers.

When the synchronization has finished, stop (ctrl-c) the master; note the output of the two other servers.

→ 9. Keep track of: output of the other two servers

Sharding

Start two shard servers; shard the cities by the country.

→ 10. Keep track of: your commands; output from the servers.