CS 237 Project presentation

Zhenyuan Zhang
Pansheng Fang
Dehao Li

Motivation & Goals

———

- Nowadays, recommendation kind of systems become more and more important in lots of fields. Distributed system is considered to be one of the most popular ways to realize the recommendation systems and middleware is truly one of the core parts of the system.
- Due to this reason, we decide to design a basic middleware and build a simplified version of recommendation system by using that middleware.

Related work

- A web application based on java and tomcat.

- Java based middleware combined server.

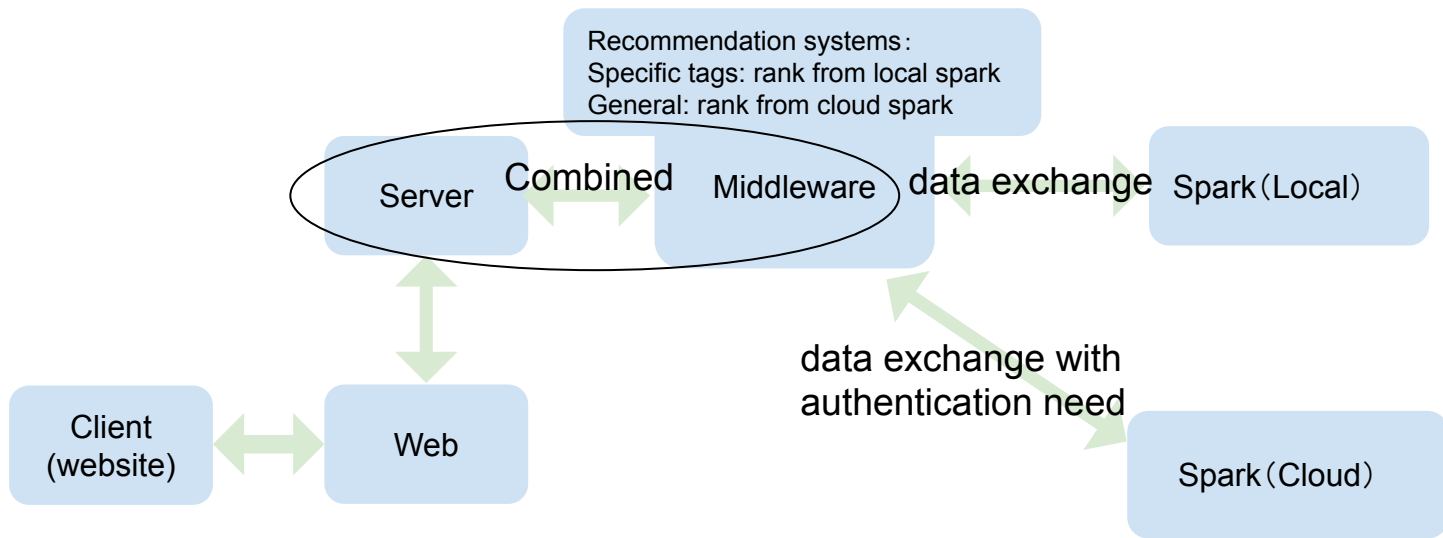- Spark based distributed system with cloud and local nodes.

Frequency count & Spark on AWS cluster

———

- A spark job that can get a huge amount of data as input.
- The data should be in the form of text such as .txt file.
- The system can read through the files one by one and count the frequency of each item that is shown in the file.
- Spark environment has been implemented on the server.
- Three nodes.
- One of the three nodes acts as the master node.

System

– – –

Recommendation systems：
Specific tags: rank from local spark
General: rank from cloud spark

Server   Combined   Middleware   data exchange   Spark（Local）

Client
(website)   Web

data exchange with
authentication need

Spark（Cloud）

Testing

———

1. Prestore movie related data in cloud spark.
2. Directly run the middleware combined server.
3. (Cloud push and pull data would run periodically.
4. Using browser to open the web site offered by server.
5. Show the rank of the movies(should have rank from cloud).
6. Watch specific type of the movie.
7. Show the rank of the movies(should have rank from local).