Introduction to Distributed Systems

RSS Feeds Classification using MapReduce

Group 6
Sohaila Baset
Ernesto Serrano
Mohammad Parhizkar
Sezin Yildiz

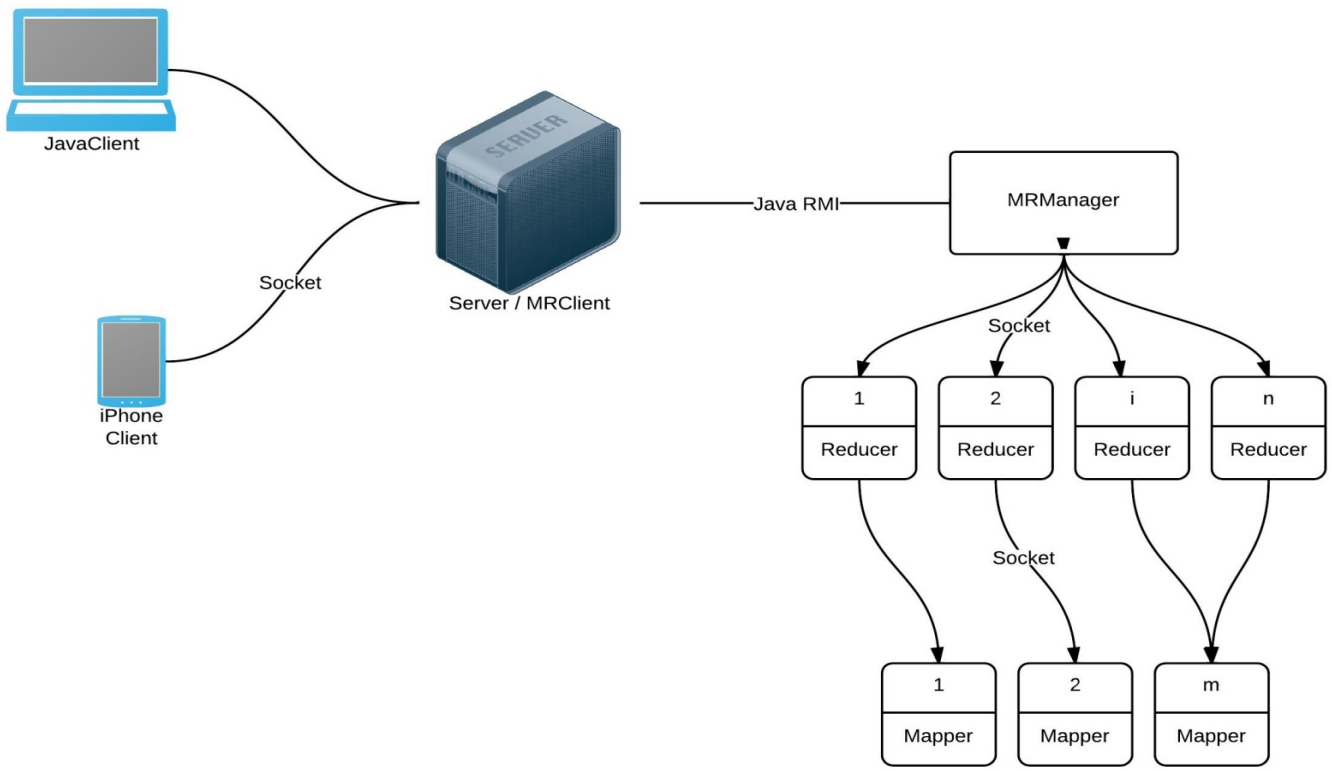
- RSS feeds are growing more and more in size and broadness.

- Information we receive from RSS can be irrelevant or at least not very easy to sort.

Demo

….!!!

Architecture Client-Server



IPhone/Java Clients

- Add RSS URLs to be searched for
- Add keyword(s) to the search queries
- Show the result of the queries to the user

Server/MRClient

- Receive the queries from the clients
- Parse the RSS URLs and return the results back
- Connect to the MRManager through Java RMI
- Request for filtering RSS feeds with the keywords (RSSJob)
- Send the result of the filtered RSS feeds to clients

Server/MRManager

- Receive the filter requests from MRClient with Java RMI Connection
- Connect to Reducers through Sockets and divide the job into tasks (RSSTask)
- Distribute the tasks to Mappers through Sockets
- Receive the result of the tasks from Reducers and return the final result to MRClient (TaskResult)

Communication

- Sockets
  - IPhone Client – MRClient
  - MRManager – Reducers
  - Reducers – Mappers
- Java RMI
  - MRClient – MRManager

Work Distribution

- Ernesto – Java Client
- Mohammad – Iphone Application
- Sezin – MapReduce Client
- Sohaila – Server / MapReduce Manager

Thank you!

QUESTIONS..?