

Implementing Billing Services in Distributed Systems with Hadoop


CS230

Julius Aguma
Mowhebat Bazargani
Elaine Thai



Project Introduction (Motivation + Goals)

- In our project, we are trying to simulate the billing services in a distributed system.
- Our goal is to split the workload between different machines in the Hadoop cluster to improve the performance and availability of the billing services when we have a high load.



Project Design and Implementation Specifics

Server:

- Runs on multi node cluster
- Uses sockets and ssh for communication
- Has a local master, local slave and aws cloud node
- Uses opencsv library for manipulating csv files

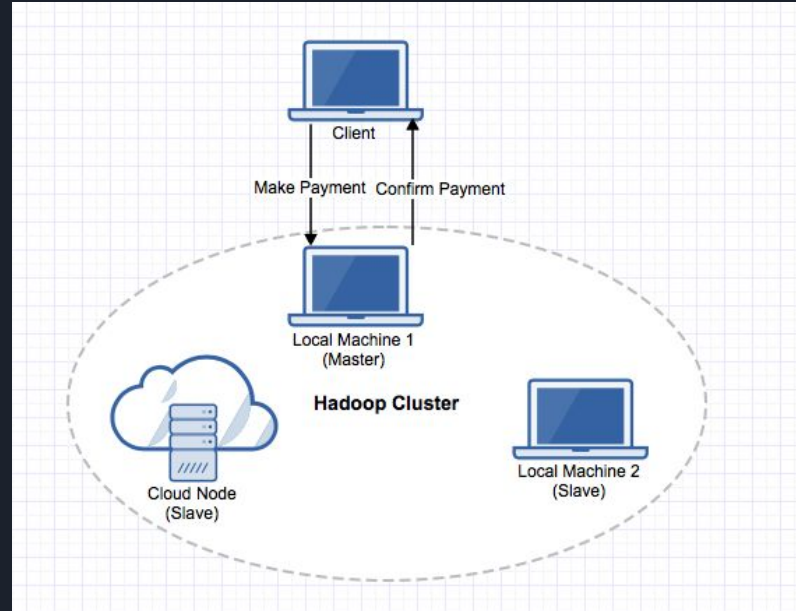
Client:

- Runs on a local machine
- Randomly selects test student
- Uses sockets to communicate with server

Project Architecture

We will have one Hadoop cluster with one master node and two slave nodes. One local machine will act as the master while another and a cloud node will act as slaves.

The database is stored in the master node of the Hadoop cluster, including students' name and the amount of money they owe. When requests are received from the client, the Hadoop cluster splits the works between the nodes.



Project Testing & Evaluation

Running
client code:

```
the8hire@hadoop-master: ~/Desktop/hadoop project
File Edit View Search Terminal Tabs Help

hadoop@hadoop-master: ~/finalRun x the8hire@hadoop-master: ~/Desкто... x

inet6 fe80::4455:821f:bf55:f256 prefixlen 64 scopeid 0x20<link>
ether f8:16:54:7e:f0:0e txqueuelen 1000 (Ethernet)
RX packets 9773 bytes 8455581 (8.4 MB)
RX errors 0 dropped 0 overruns 0 frame 0
TX packets 6788 bytes 1331164 (1.3 MB)
TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

the8hire@hadoop-master:~$ cd Desktop
the8hire@hadoop-master:~/Desktop$ cd 'hadoop project'
the8hire@hadoop-master:~/Desktop/hadoop project$ javac -cp .:opencsv-5.1.jar:commons-lang3-3.1.jar HadoopClient.java
the8hire@hadoop-master:~/Desktop/hadoop project$ java -cp .:opencsv-5.1.jar:commons-lang3-3.1.jar HadoopClient
Connected to Server
Sending name:Terra Dobney
payment:161
Server says: Found and updated
the8hire@hadoop-master:~/Desktop/hadoop project$ java -cp .:opencsv-5.1.jar:commons-lang3-3.1.jar HadoopClient
Connected to Server
Sending name:Bernard Baldini
payment:207
Server says: Found and updated
the8hire@hadoop-master:~/Desktop/hadoop projects$
```

```
hadoop@hadoop-master: ~/finalRun
File Edit View Search Terminal Tabs Help

hadoop@hadoop-master: ~/finalRun x the8hire@hadoop-master: ~/Deskto... x

ShutDown HSB: Shutting down namenode at hadoop-master/10.120.1.10.201
*****/
hadoop@hadoop-master:~/finalRun$ sudo /usr/local/hadoop/sbin/start-all.sh
Starting namenodes on [hadoop-master]
Starting datanodes
Starting secondary namenodes [hadoop-master]
Starting resourcemanager
Starting nodemanagers
hadoop@hadoop-master:~/finalRun$ sudo make
javac -classpath .:opencsv-5.1.jar:commons-lang3-3.1.jar:'/usr/local/hadoop/bin/
hadoop classpath' HadoopServer.java
jar -cvf HadoopServer.jar HadoopServer.class
added manifest
adding: HadoopServer.class(in = 3149) (out= 1791)(deflated 43%)
hadoop@hadoop-master:~/finalRun$ sudo make test
/usr/local/hadoop/bin/hadoop jar HadoopServer.jar HadoopServer
Socket open and listenin...
client info: Terra Dobney,161
old payment:161
Found and updated
client info: Bernard Baldini,207
old payment:207
Found and updated
```

Running server code



Contribution

Elaine Thai:

Elaine implemented early cloud cluster instances, wrote up the project proposal, worked on the project slides and wrote up the first draft of the project report. We intended to run the client on Elaine's local machine but firewall permissions prevented that.

Mowhebat Bazargani:

Mowhebat implemented one of the local slave nodes, worked on the project slides and contributed to the project proposal and project report. We intended to run the hadoop slave 1 on Mowhebat's local machine but firewalls and network filter blocks prevented that.

Julius Ceasar Aguma:

Cesar was responsible for setting up the final cloud instance, wrote the server code and final draft of the project report. We intended to use Cesar's laptop as the hadoop master but ended up having to run both server and client on it due to obstacles with multi-machine connection using ip addresses.



Conclusion

- Limitations:

Firewalls, and IP blocking

Long configuration process