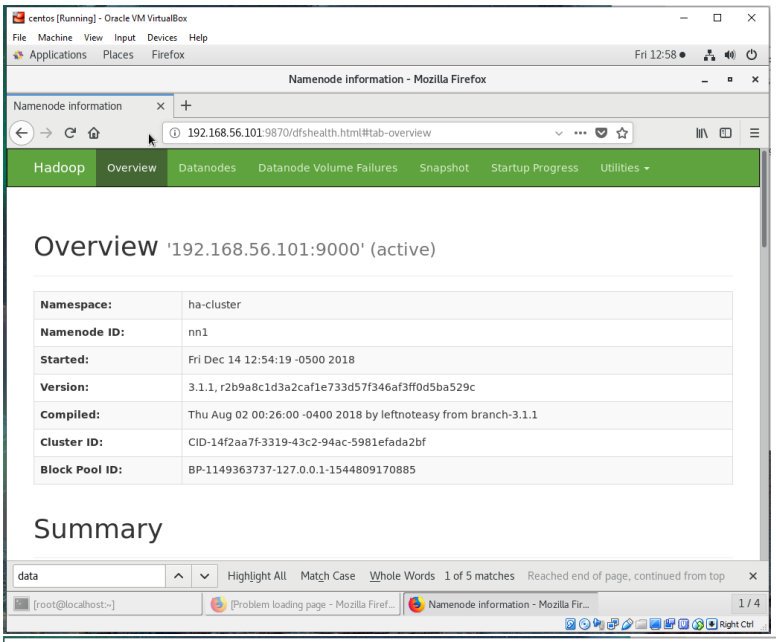
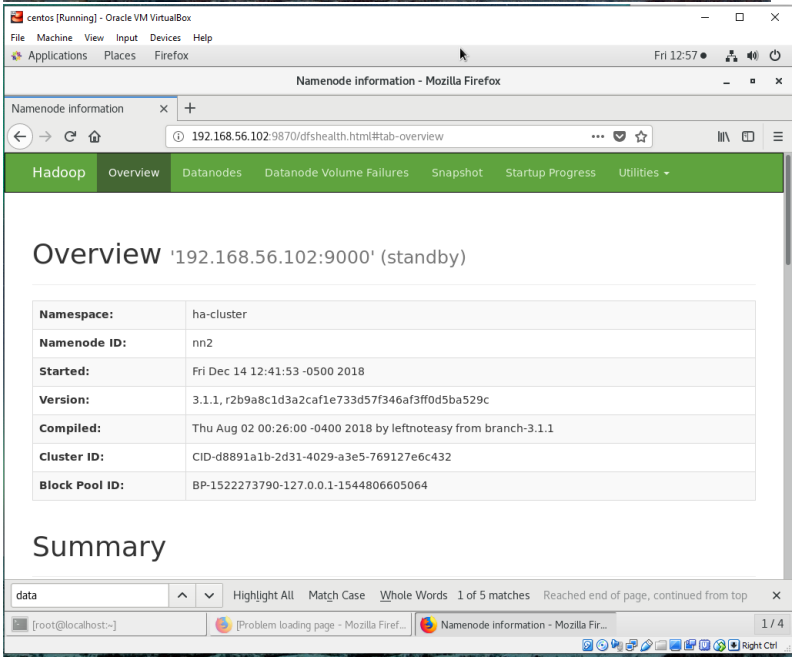
Hadoop advanced.

Task1.

Here are the 2 of my namenodes:

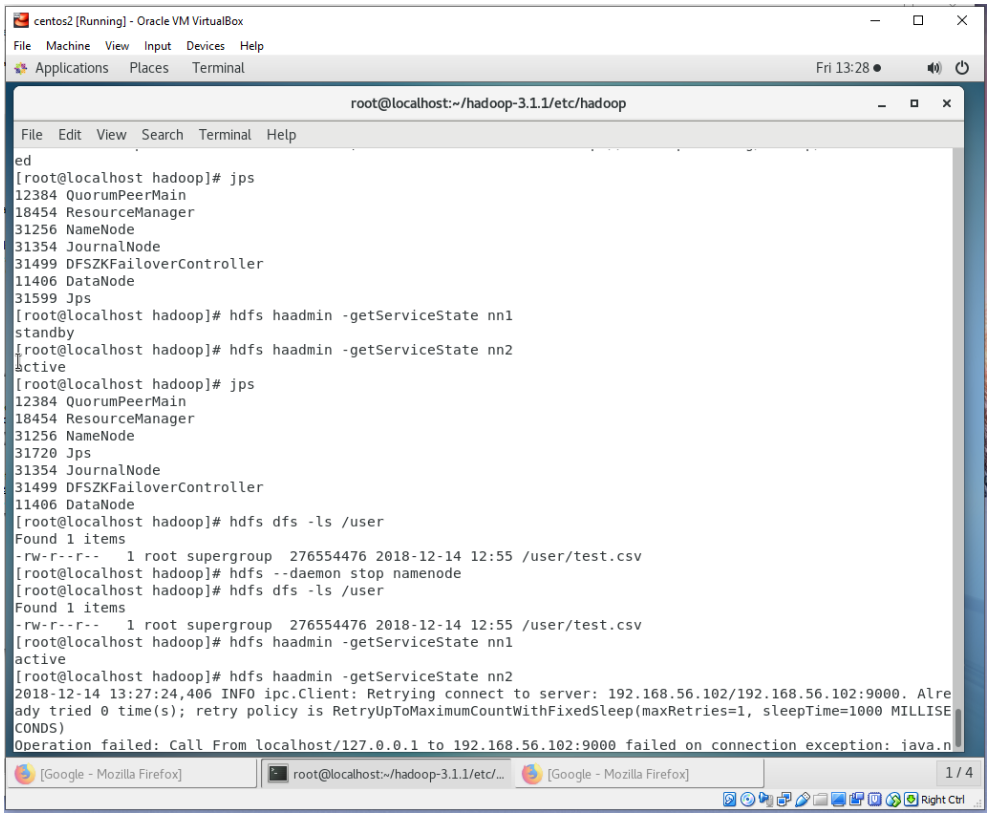




Here I uploaded test.csv into hdfs and executed “ls” to see that it’s there.

```
[root@localhost ~]# hdfs dfs -copyFromLocal test.csv /user
[root@localhost ~]# hdfs dfs -ls /user
Found 1 items
-rw-r--r--   1 root supergroup  276554476 2018-12-14 12:55 /user/test.csv
[root@localhost ~]#
```

Next, in the screenshot below, I first check the running services with “jps”, as you can see NameNode is there. Then I stop it by executing “hdfs –daemon stop namenode” and voila, I can still access hdfs and my standby node become the active one.



Task2.

I followed the tutorial on spring’s website, everything in my code is the same as there, except the Pojo implementation, which I modified to suit task 2.

Here’s how I start my application:

```
[pavel_orekhov@ecsc00a02339 HadoopAdvanced]$ java -jar gs-yarn-basic-dist/target/gs-yarn-basic-dist/gs-yarn-basic-client-0.1.0.jar --my.client.filePath=/user/train.csv
```

Here’s the result for train.csv:

```
[pavel_orekhov@ecsc00a02339 HadoopAdvanced]$ yarn logs -applicationId application_1545290316073_0014 | grep JobPojo
18/12/20 09:27:10 INFO client.RMProxy: Connecting to ResourceManager at ecsc00a022c6.epam.com/10.6.218.24:8032
[2018-12-20 09:26:39.618] boot - 26751  INFO [main] --- JobPojo: Answer: 1277716
[2018-12-20 09:26:39.618] boot - 26751  INFO [main] --- JobPojo: Answer: 275737
[2018-12-20 09:26:39.618] boot - 26751  INFO [main] --- JobPojo: Answer: 141535
[pavel_orekhov@ecsc00a02339 HadoopAdvanced]$
```