Mapreduce Jar Creation

...

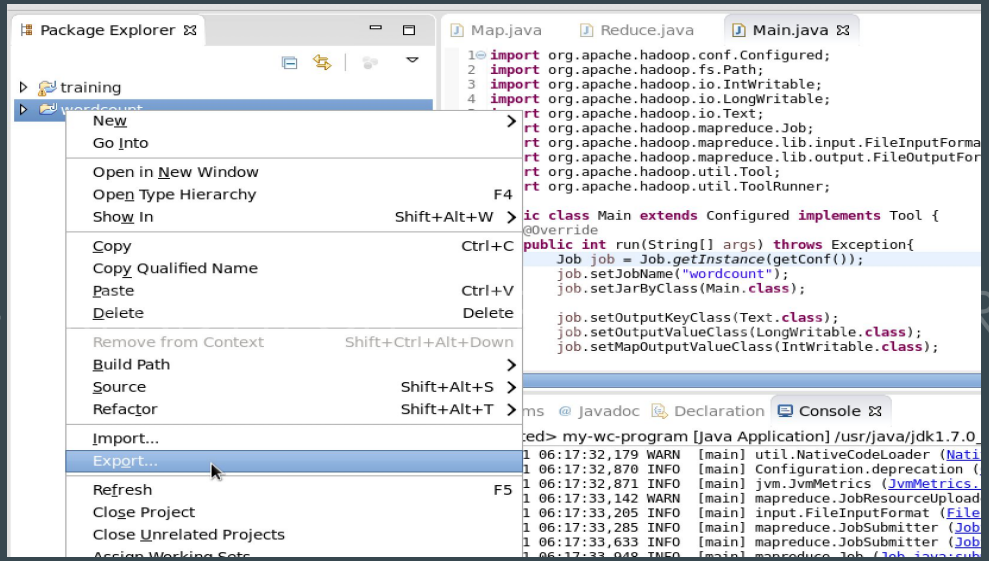
Running MR Job on HDFS File

Moving the input file from local to hdfs

```
[cloudera@quickstart Desktop]$
[cloudera@quickstart Desktop]$ hadoop fs -mkdir /user/cloudera/inputfolder
[cloudera@quickstart Desktop]$ hadoop fs -put inputfolder/inputfile.txt inputfolder
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/cloudera/inputfolder
Found 1 items
-rw-r--r--   1 cloudera cloudera         94 2020-04-11 12:22 /user/cloudera/inputfo
lder/inputfile.txt
[cloudera@quickstart Desktop]$
```

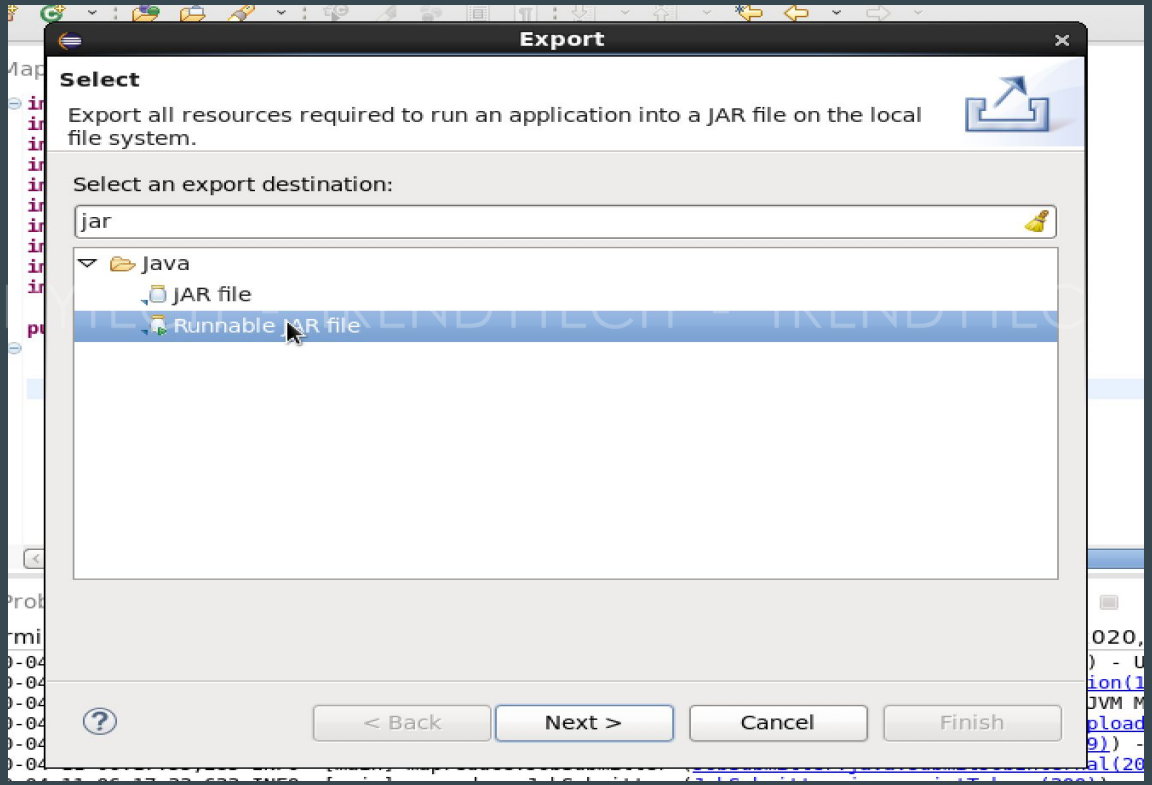
```
[cloudera@quickstart Desktop]$
[cloudera@quickstart Desktop]$ hadoop fs -cat /user/cloudera/inputfolder/*
we are all geared up to learn big data
big data tech stack is fun
yes it is an ablsolute fun

[cloudera@quickstart Desktop]$
```

Exporting our project as jar file

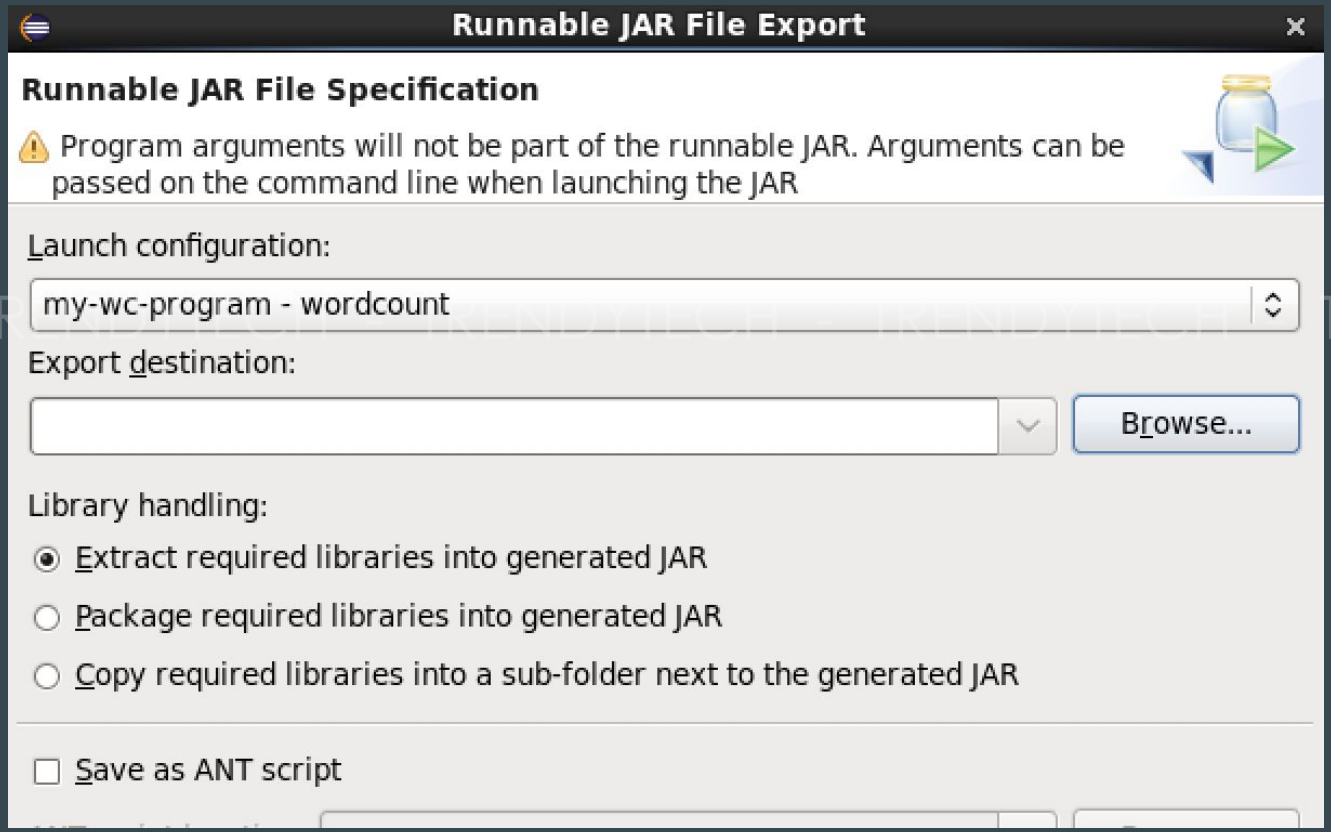


Right click on project -> export -> runnable jar

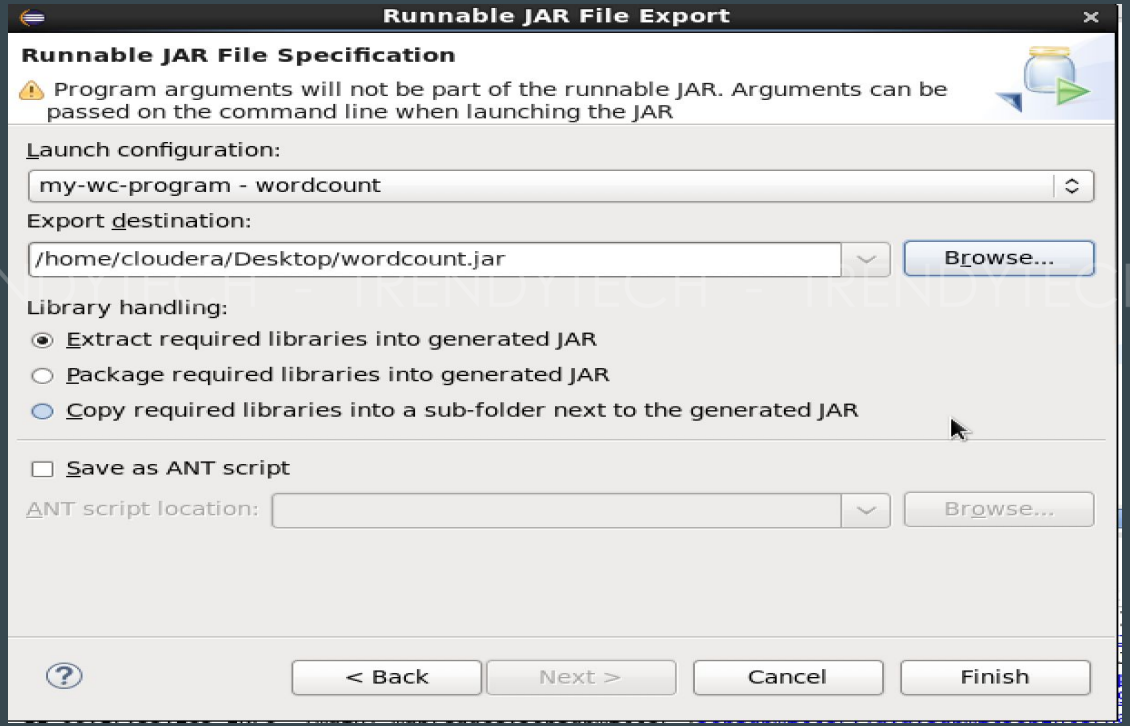
Exporting our project as jar file



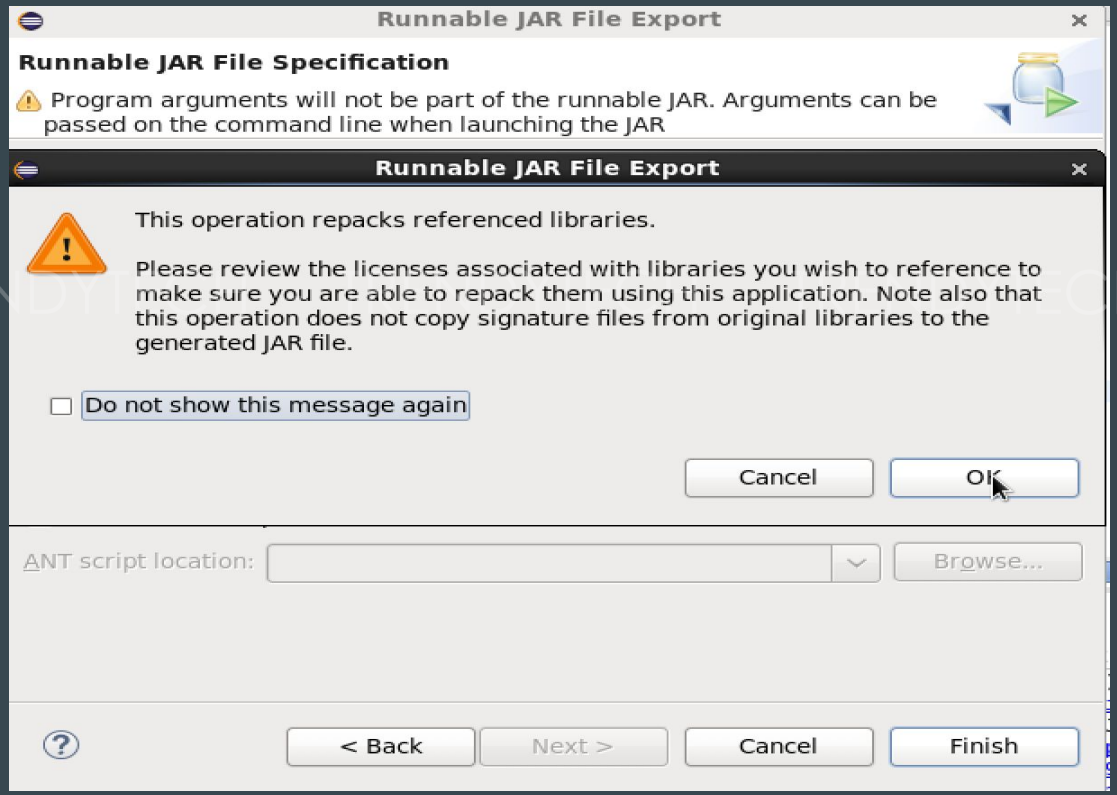
Select the run configuration for your project



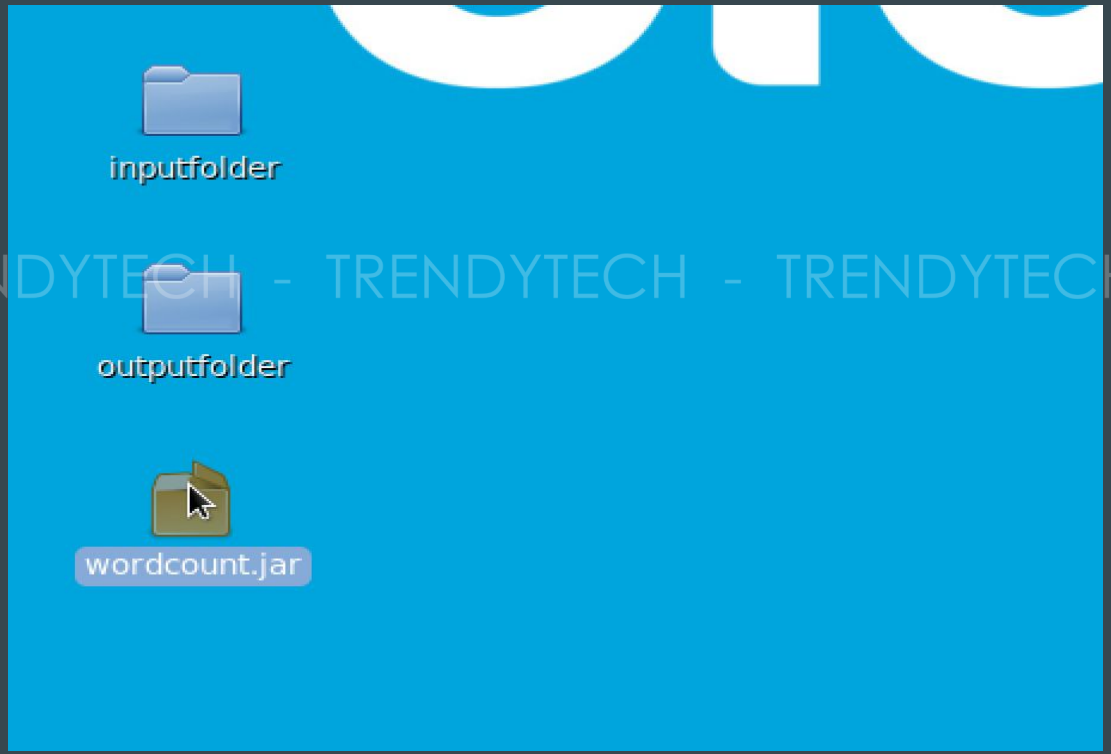
Give the name of the jar and also the location where jar should be created



Ignore the warning while importing



You can see the jar is now present on Desktop



Run the jar now using below command

hadoop jar <jarname> <input path> <output path>

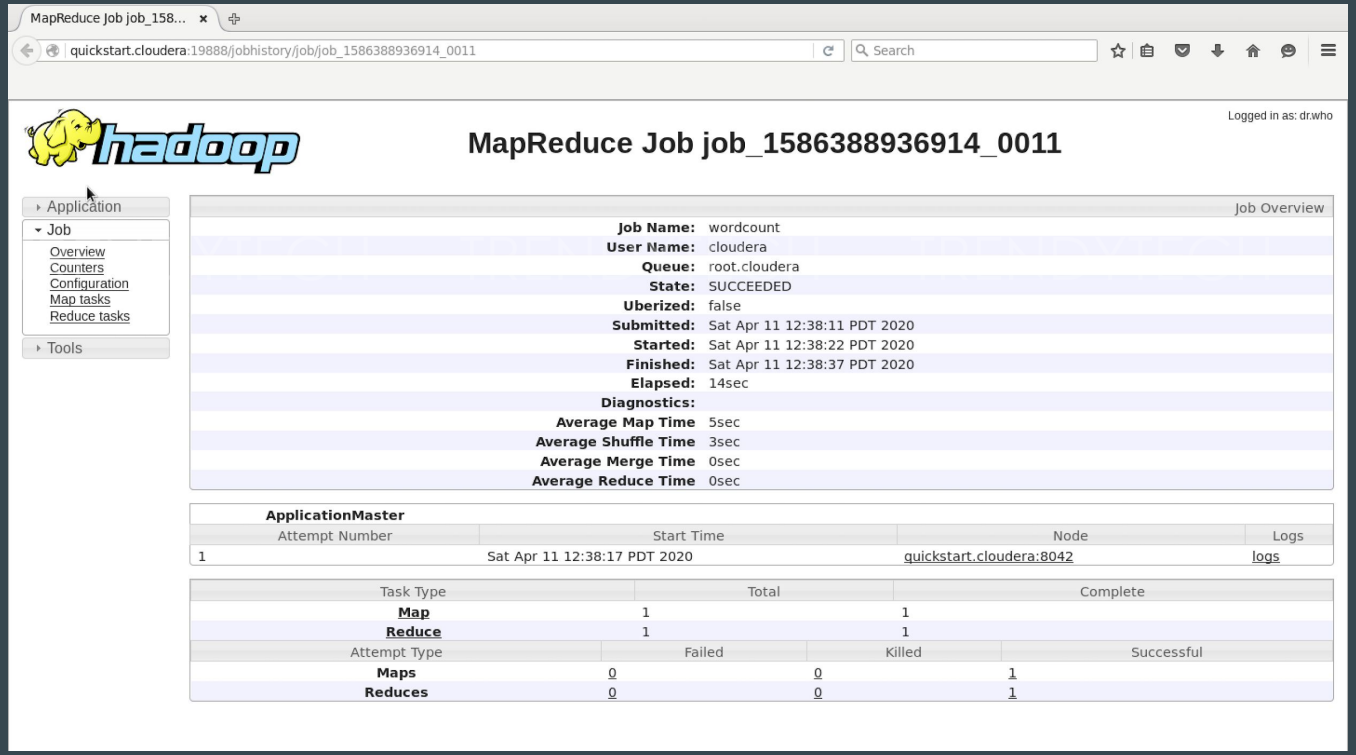

```
[cloudera@quickstart ~]$
[cloudera@quickstart ~]$
[cloudera@quickstart ~]$ hadoop jar /home/cloudera/Desktop/wordcount.jar /user/clou
dera/inputfolder/inputfile.txt /user/cloudera/outputfolder
```

Job Tracking URL will be displayed when Job Runs (in the logs)

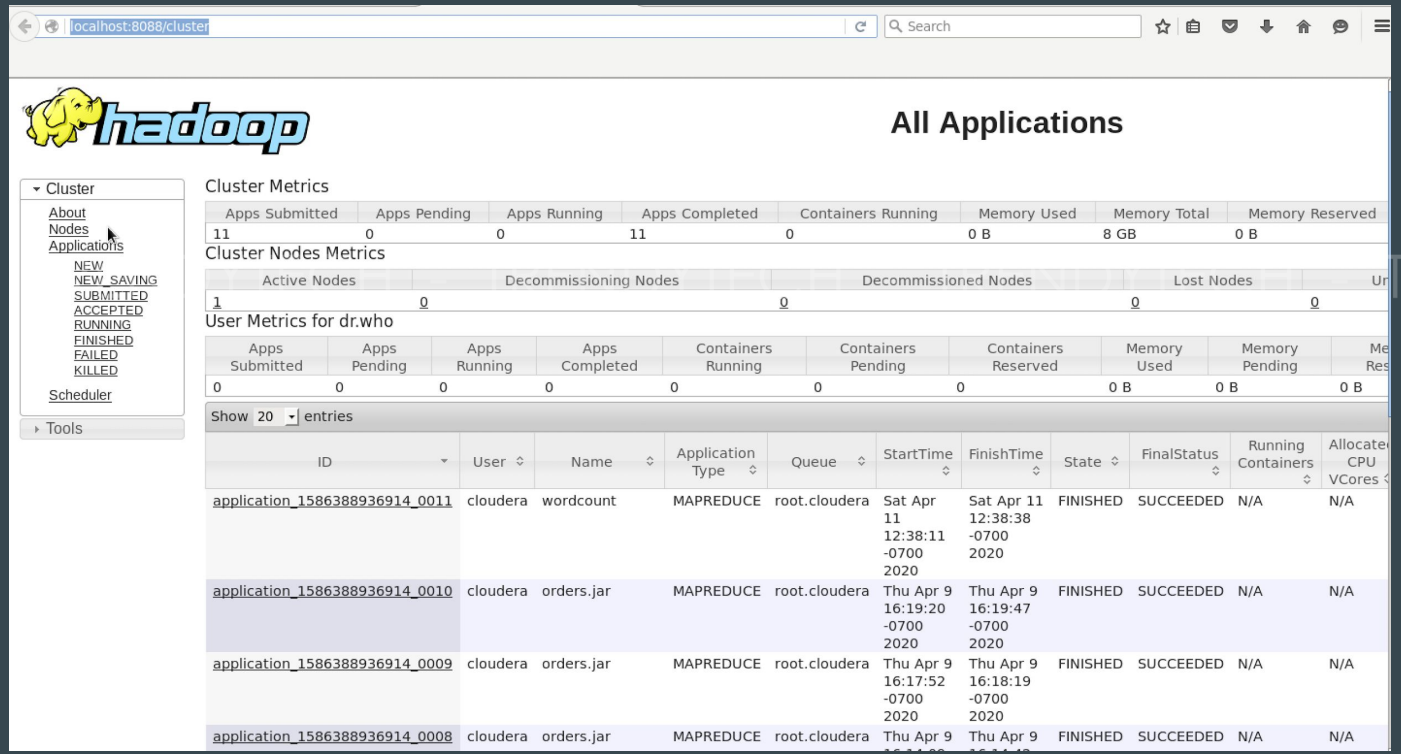
```
20/04/11 12:38:11 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_15863
88936914_0011
20/04/11 12:38:12 INFO impl.YarnClientImpl: Submitted application application_15863
88936914_0011
20/04/11 12:38:12 INFO mapreduce.Job: The url to track the job: http://quickstart.c
loudera:8088/proxy/application_1586388936914_0011/
20/04/11 12:38:12 INFO mapreduce.Job: Running job: job_1586388936914_0011
20/04/11 12:38:24 INFO mapreduce.Job: Job job_1586388936914_0011 running in uber mo
de : false
20/04/11 12:38:24 INFO mapreduce.Job:  map 0% reduce 0%
```

We can try opening the url in our cloudera browser.

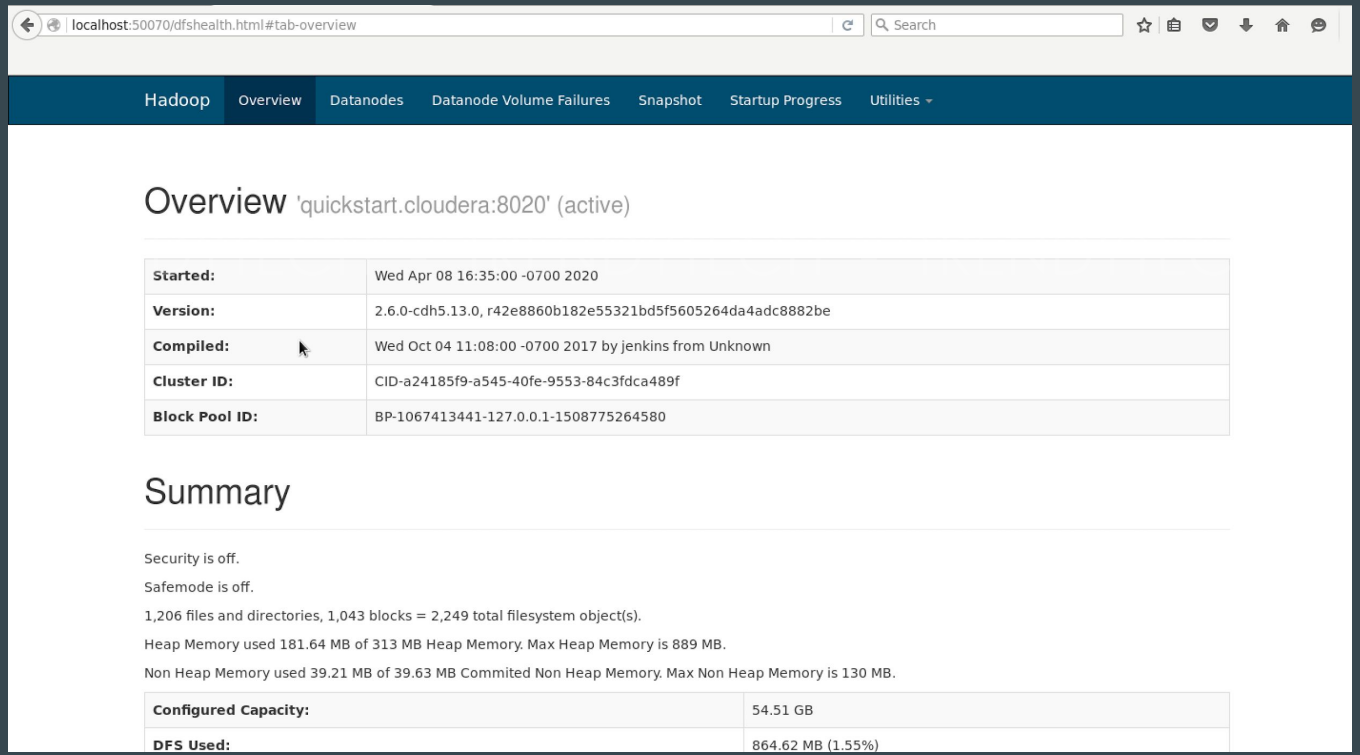
The job tracking URL will show details related to the job



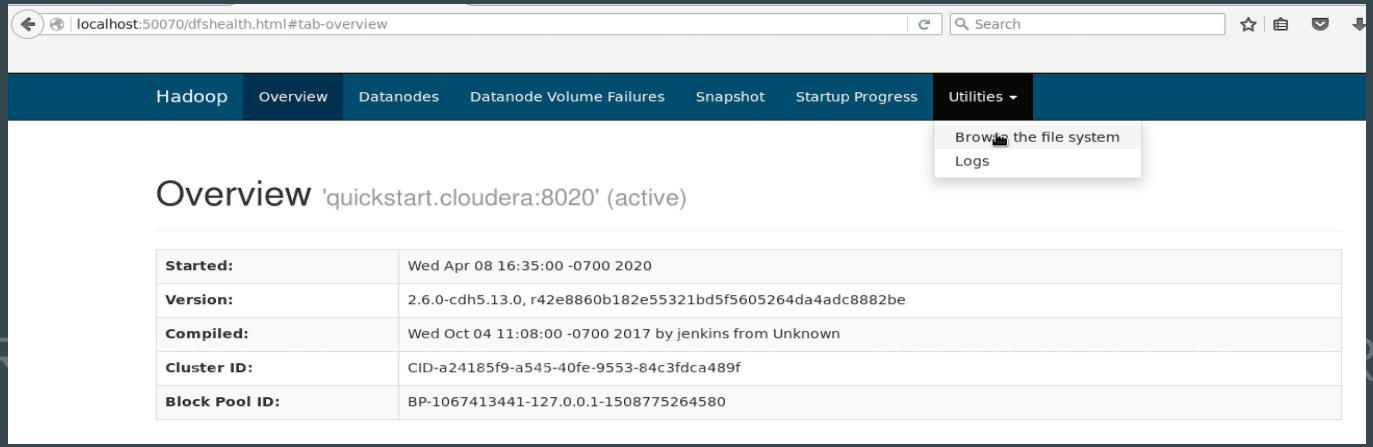
See list of all jobs on port 8088 (localhost:8088)

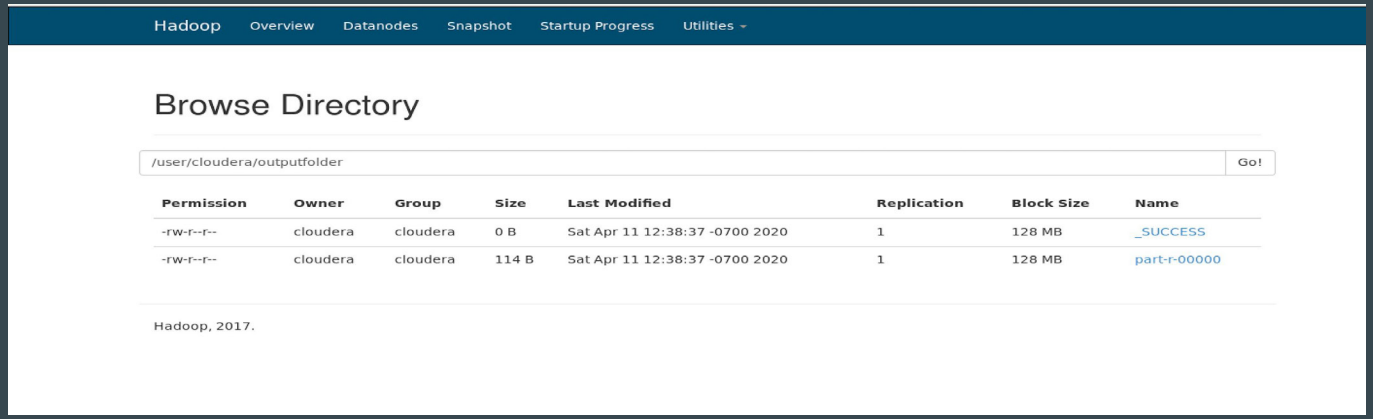


See HDFS details through web UI on Port 50070 (localhost:50070)



Browse HDFS filesystem





We have successfully Created & Executed the Jar

Happy Learning!!!

Google REVIEWS

5 Star Google Rated
Big Data Course

LEARN FROM THE EXPERT

9108179578

Call for more details

Follow US

Trainer	Mr. Sumit Mittal
Phone	9108179578
Email	trendytech.sumit@gmail.com
Website	https://trendytech.in/courses/big-data-online-training/
LinkedIn	https://www.linkedin.com/in/bigdatabysumit/
Twitter	@BigdataBySumit
Instagram	bigdatabysumit
Facebook	https://www.facebook.com/trendytech.in/
Youtube	TrendyTech