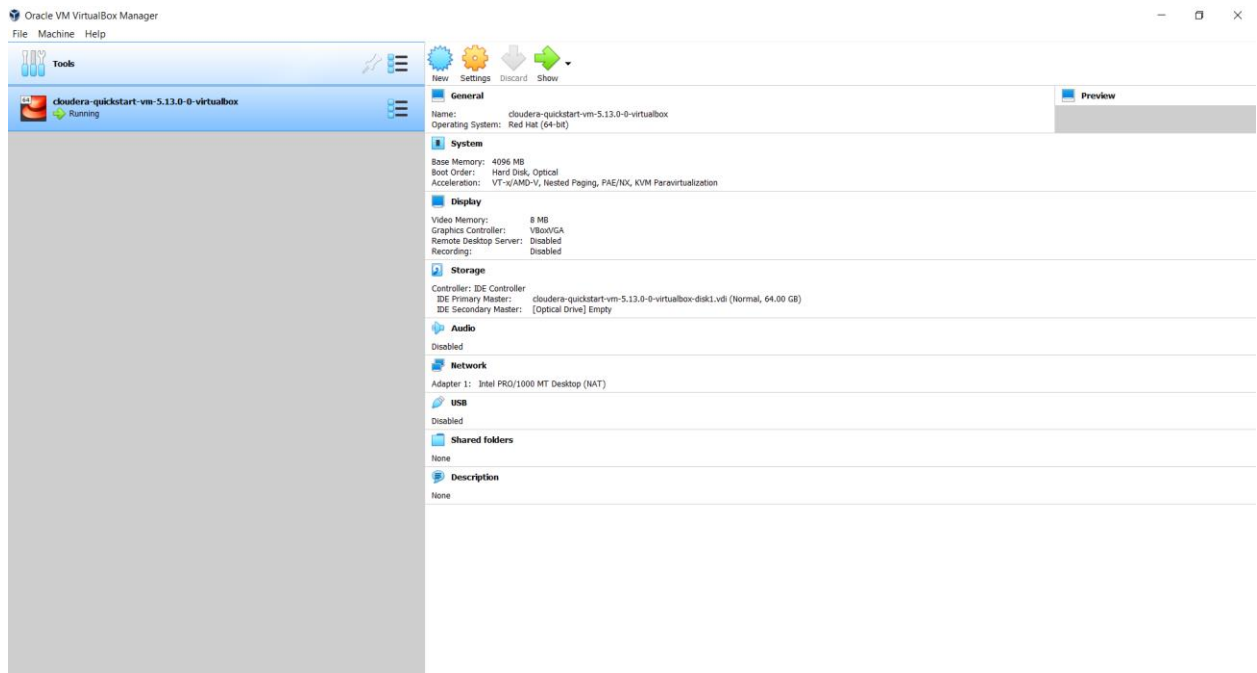
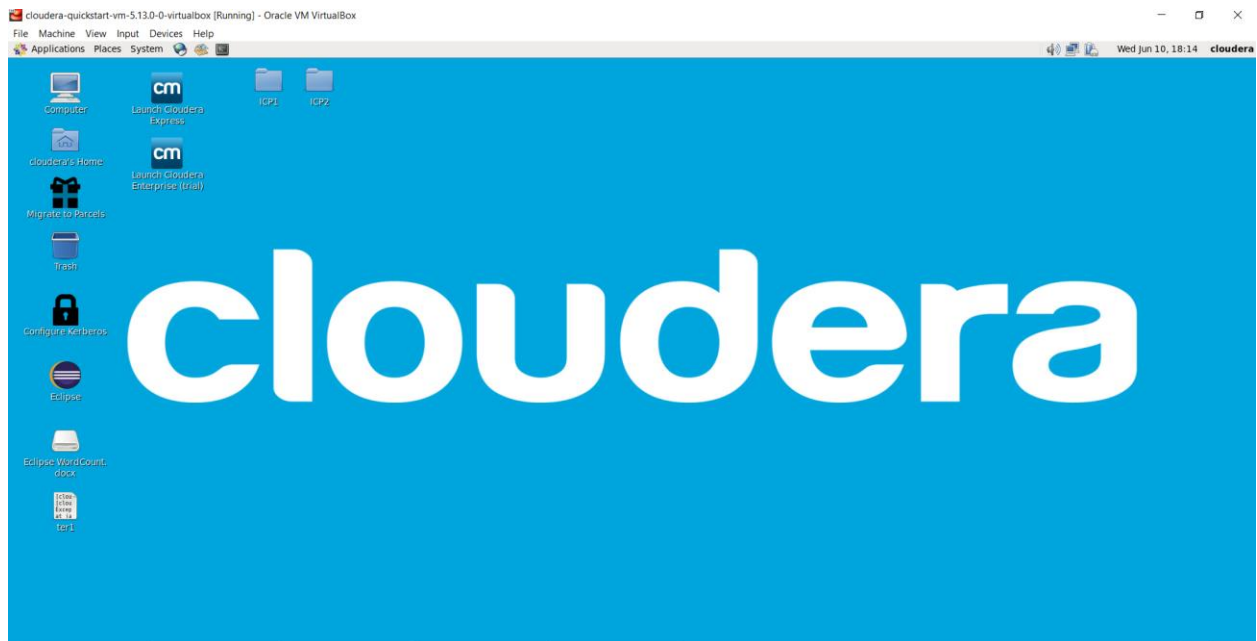


1. Download Virtual Box and Cloudera



2. Import Cloudera to Virtual Box and run it

3. Run Cloudera

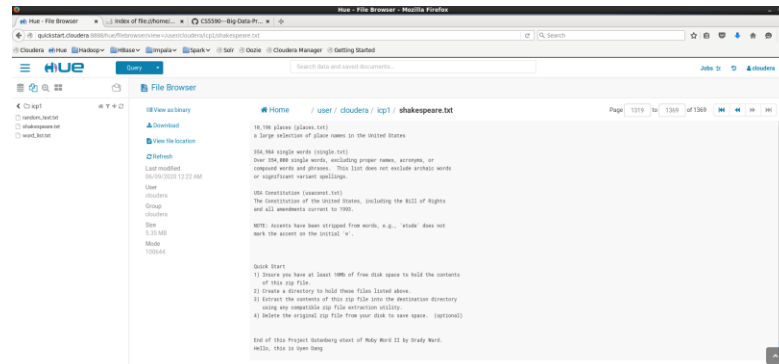


- Create file in hdfs:

`hadoop fs -mkdir /user/cloudera/icp1`

- Push shakespeare.txt to hdfs

`hadoop fs -put shakespeare.txt /user/cloudera/icp1`



- Append word_file.txt

`hadoop fs -appendToFile word_list.txt /user/cloudera/icp1/shakespeare.txt`

- Query the first and last five lines

`hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | head -n 5`

```
[cloudera@quickstart ~]$ hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | head -n 5
The Project Gutenberg EBook of The Complete Works of William Shakespeare, by
William Shakespeare
```

This eBook is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it, give it away or

`hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | tail -n 5`

```
[cloudera@quickstart ~]$ hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | tail -n 5
4) Delete the original zip file from your disk to save space. (optional)
```

```
End of this Project Gutenberg etext of Moby Word II by Grady Ward.
Hello, this is Uyen Dang
[cloudera@quickstart ~]$
```

- Create a new text file, push it to hdfs and append it to merged dataset.

```
hadoop fs -put random_text.txt /user/cloudera/icp1 hadoop fs -appendToFile  
random_text.txt /user/cloudera/icp1/shakespeare.txt
```