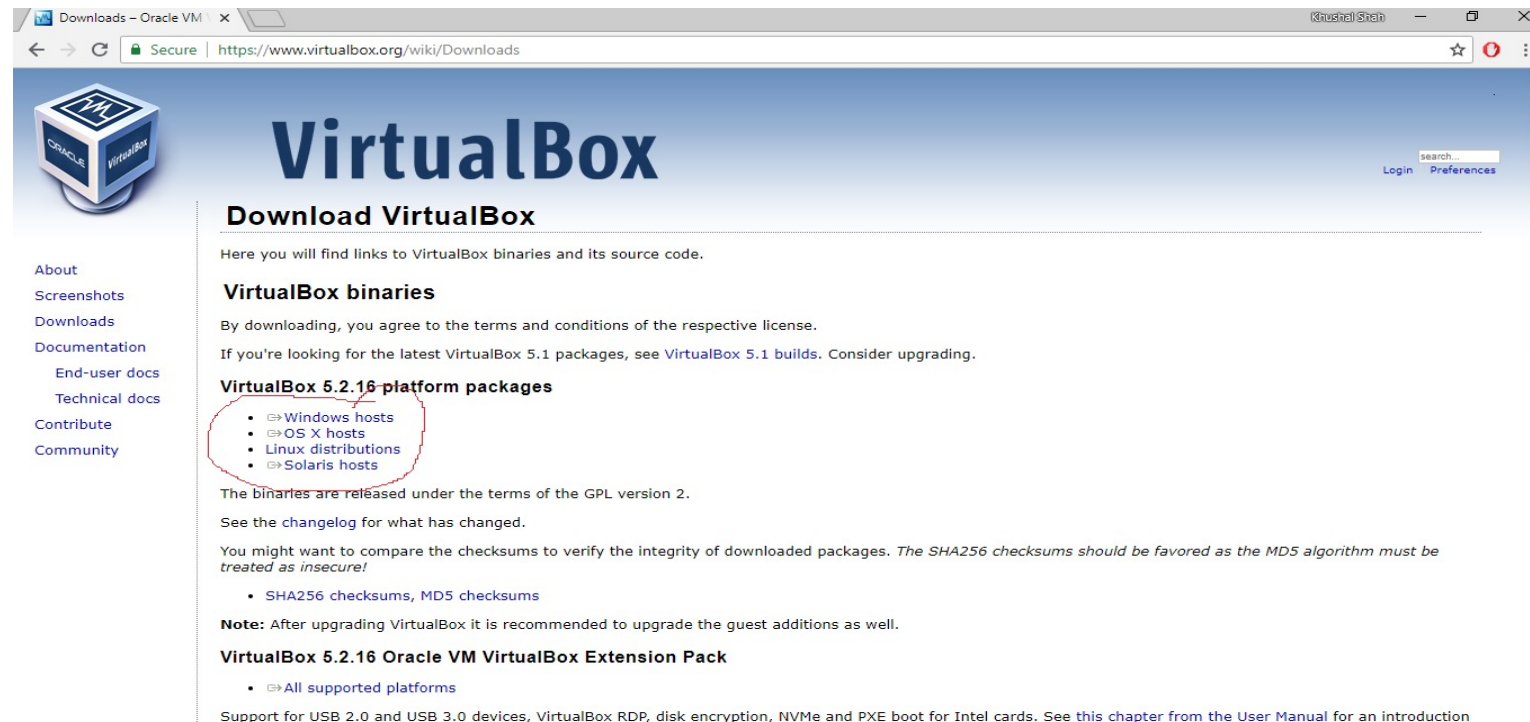# CSCE 5300
In Class Exercise 1

Cloudera Installation

1. Download cloudera operating system given in canvas under the first module.

Download Oracle Virtual Box

Use the following link to download oracle virtual box

https://www.virtualbox.org/wiki/Downloads

Extract cloudera files to a directory
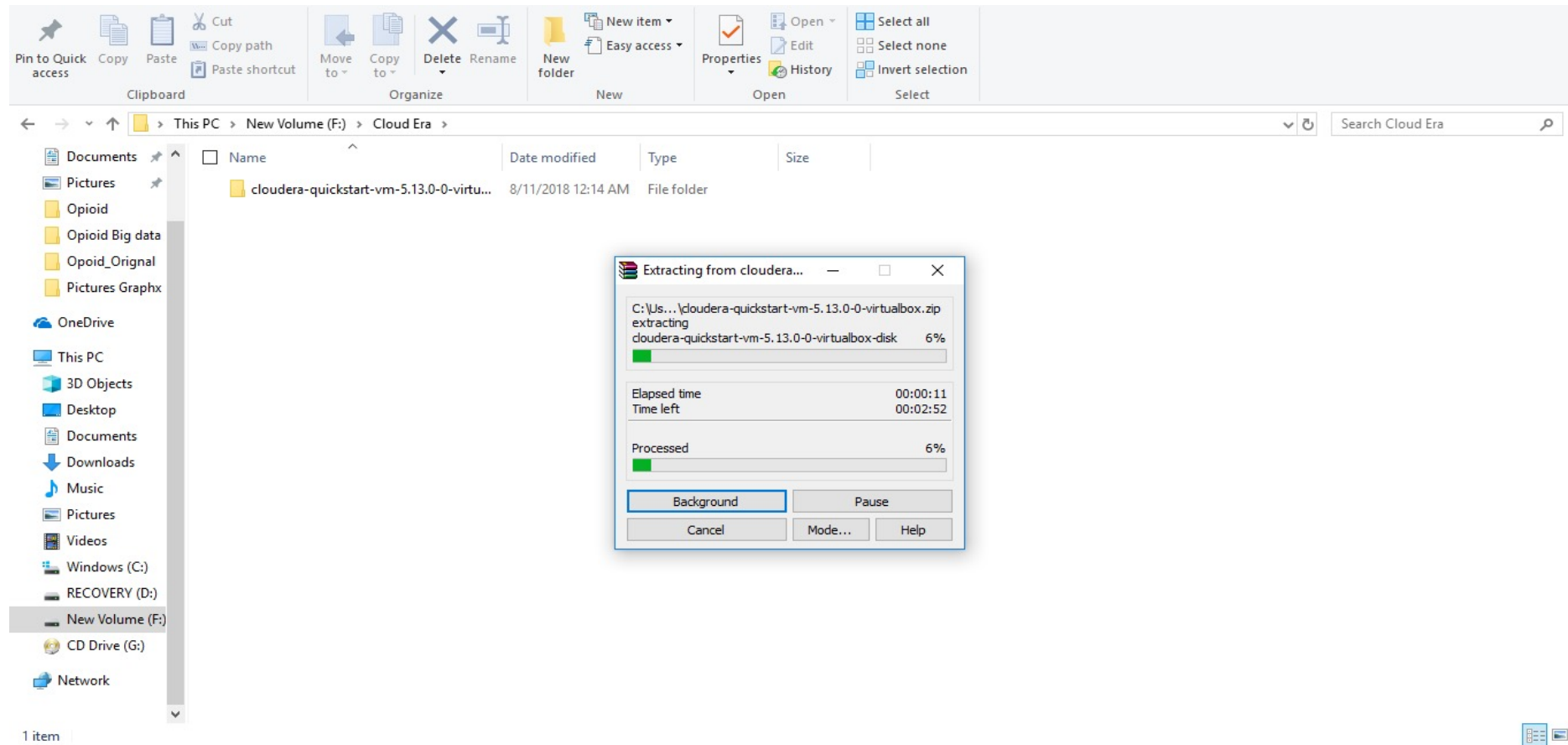
Installation

1. Install Oracle Virtual box
2. Create a Virtual Machines.
3. Go to file and select import Appliance to select the VM image from your hard drive.
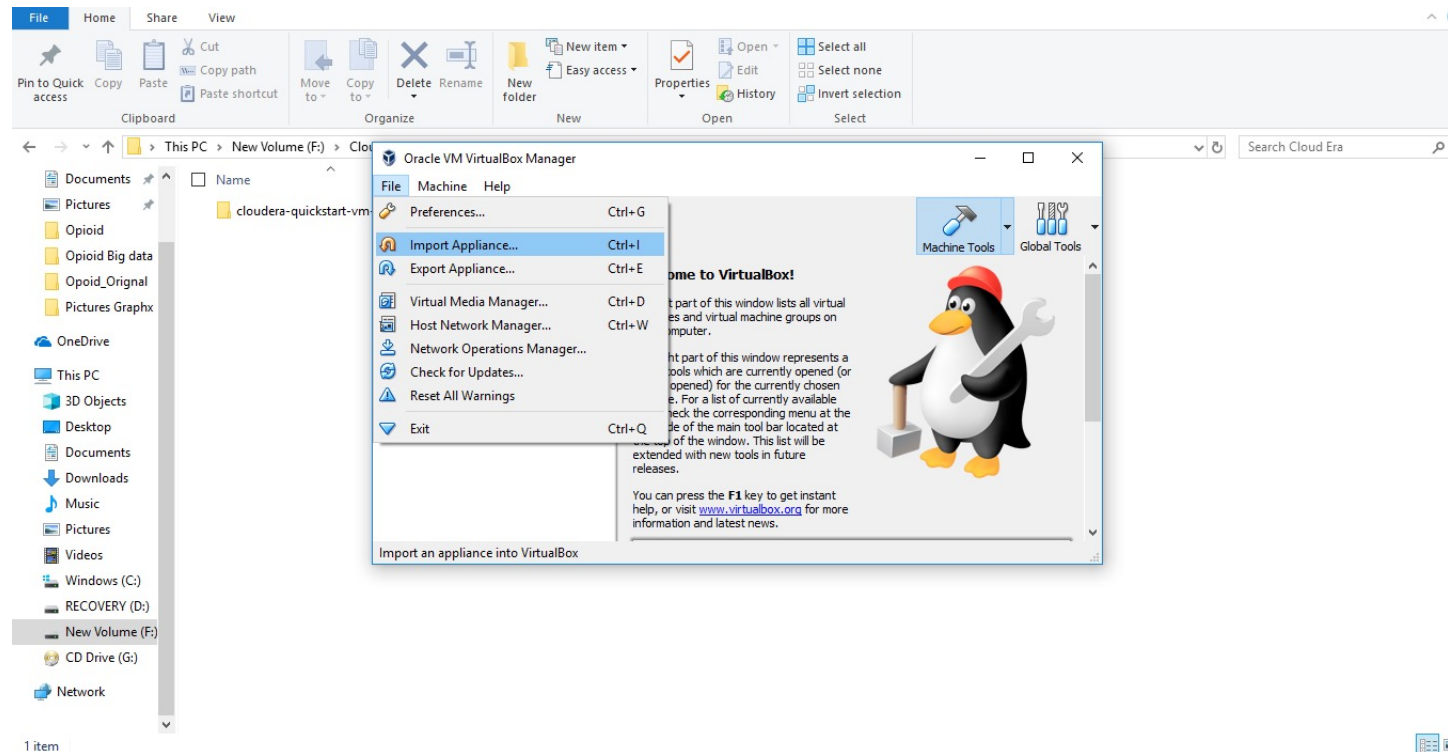
Continued ..

- Select the virtual machine to start loading the VM image in the virtual box. (cloudera-quickstart-vm-5.13.0-0-virtualbox)

- It will take several minutes to start the Cloud era operating system
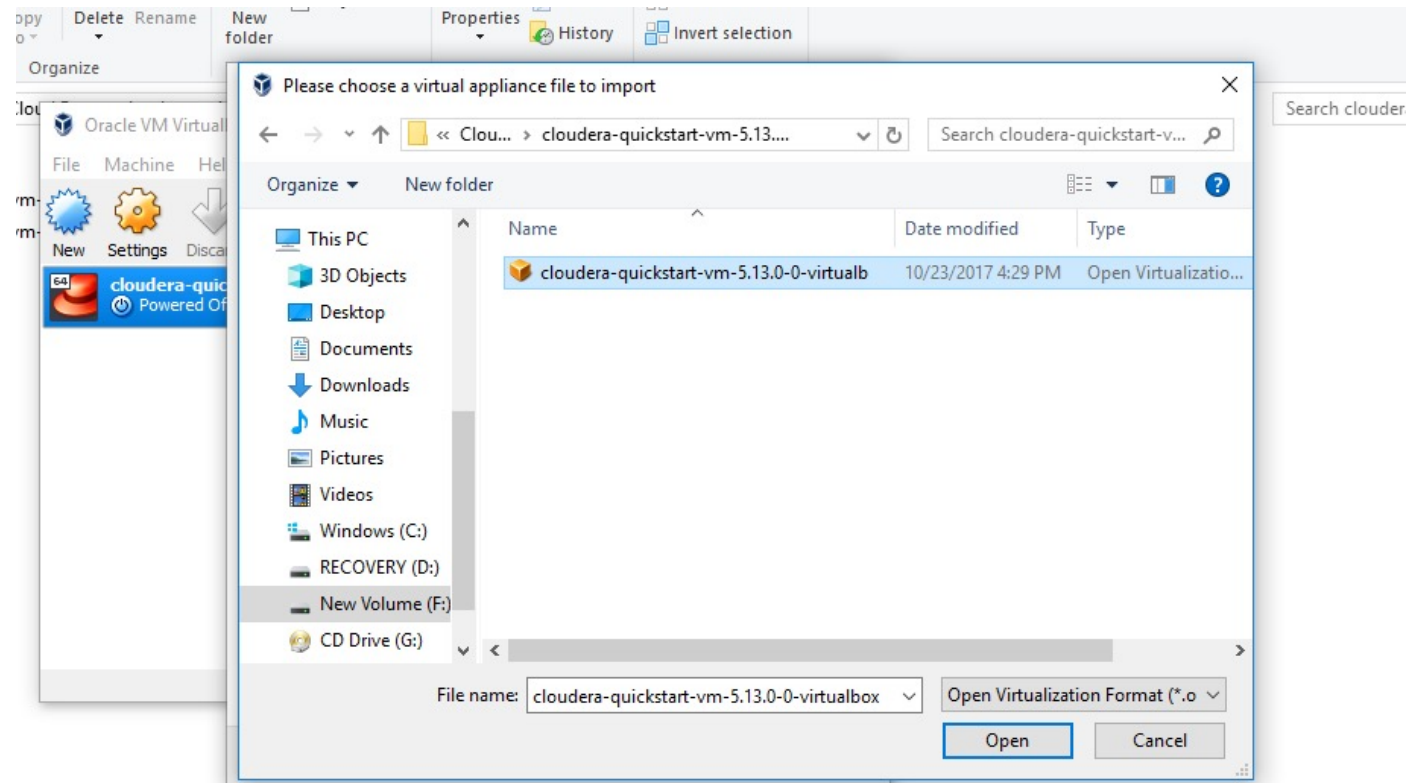
Enter Cloudera (Unix System)

- Cloud era is a Unix system and is ready for Big Data exercises.

Start by opening Terminal

- Start terminal from the upper right corner as shown in the picture.
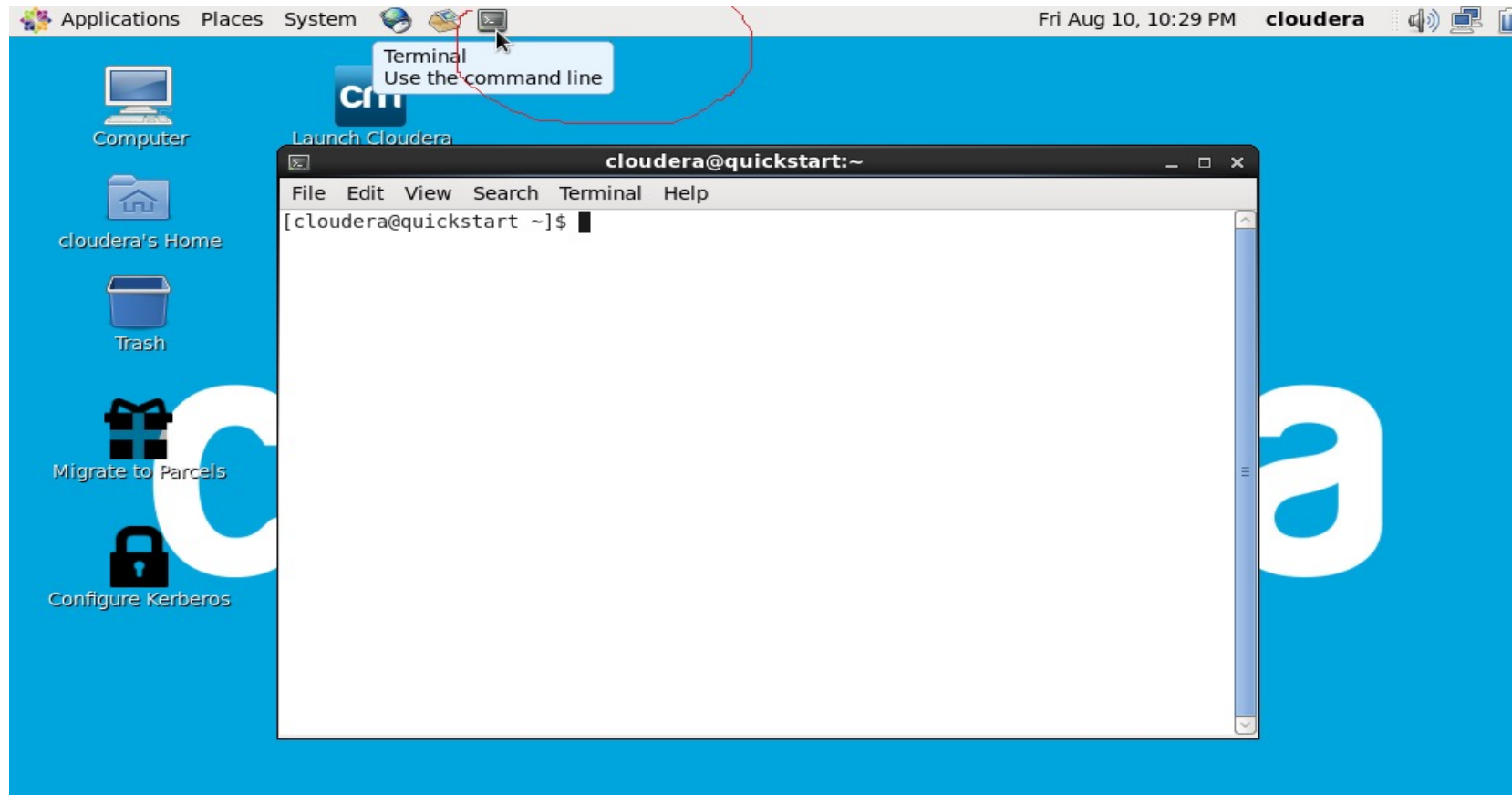
Running Hadoop (HDFS)

- Create a file named as localfile

- Open terminal and type the following command "hdfs dfs -appendToFile localfile /user/hadoop/hadoopfile"

- Check if the command was successful.

- If the command was successful, it means the hadoop server is running successfully on cloudera or any other operating system
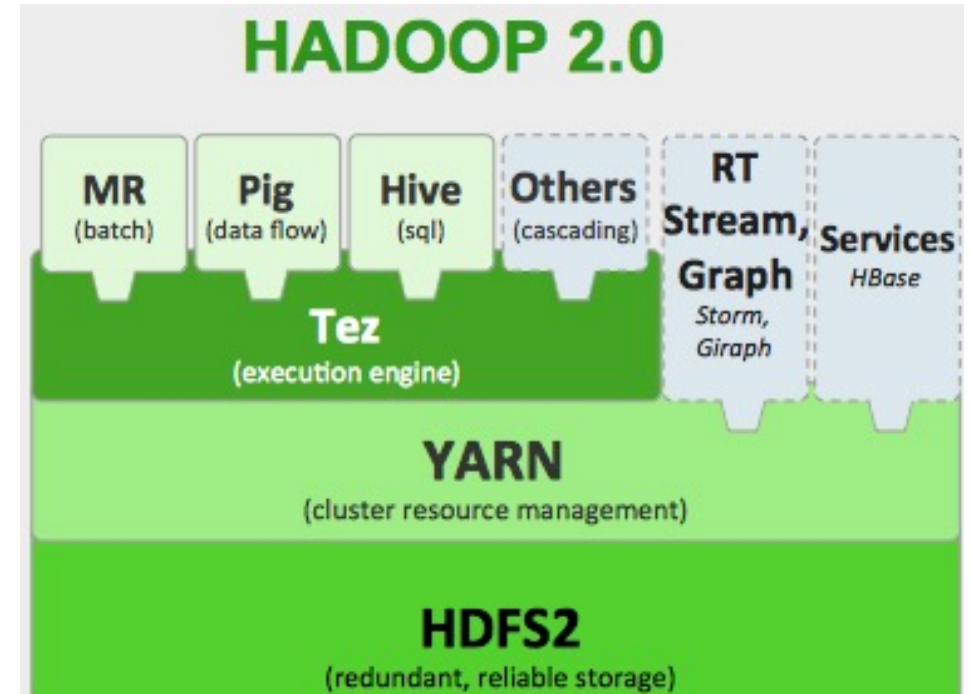
Hue

- **Hue** (Hadoop User Experience) is an open-source Web interface that supports Apache Hadoop and its ecosystem, licensed under the Apache v2 license

- Hue is an open-source Analytics Workbench for browsing, querying and visualizing data.

- All hadoop task can be queried and displayed in Hue interface.
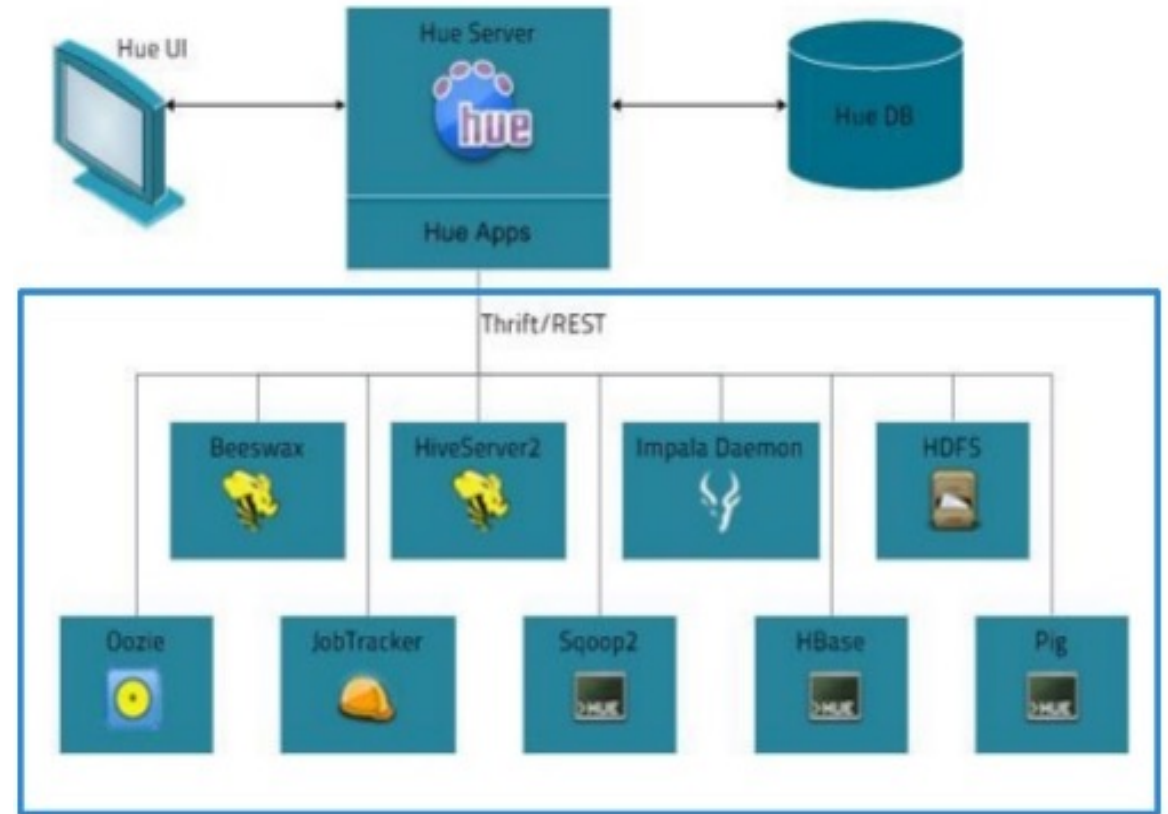
In Class Exercise

You are required to follow the steps below to complete your ICE today

- Use the datasets given in canvas
- Load it in hadoop hdfs
- Use the second file
- Append it to the first file
- Visualize file with Hue
- View the first and last lines (approximately 5) of merged dataset using appropriate hdfs commands
- Create a new text file and load it into hdfs and try to append all three datasets.

Basic Commands

List files in Hadoop

hadoop fs -ls /user/cloudera

Make new Directory:

hadoop fs -mkdir /user/cloudera/Zeenat

Move file to Zeenat

hadoop fs -put /home/cloudera/Desktop/word_list.txt /user/cloudera/zeenat