

Hadoop Installation Steps

VMware player: <https://www.vmware.com/tryvmware/?p=player>

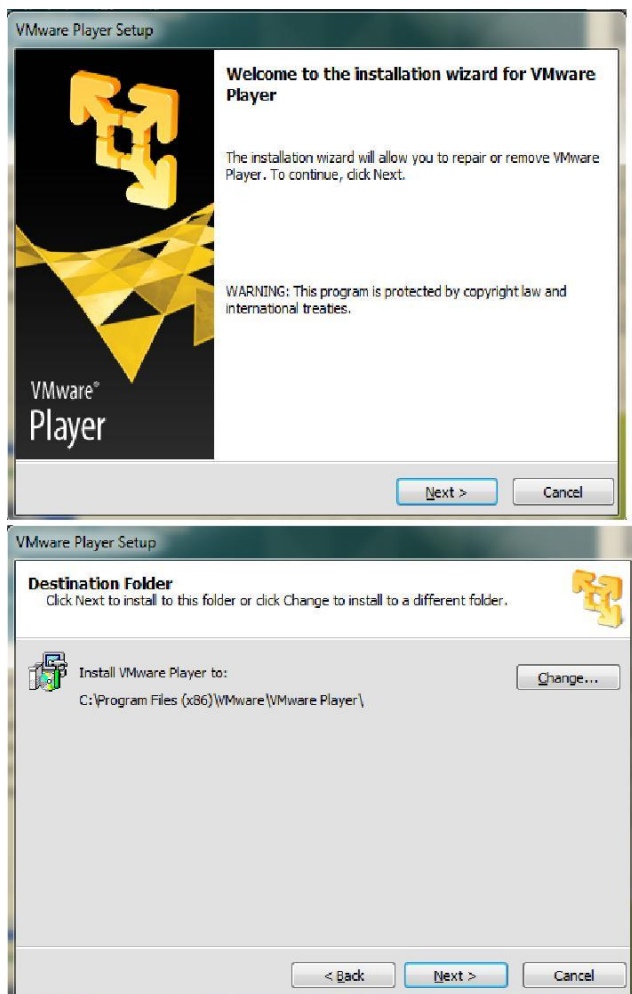
Cloudera Setup files:

Download the Cloudera VM file from **any** of the below link.

Cloudera Official Link: <https://ccp.cloudera.com/display/SUPPORT/Cloudera+QuickStart+VM>

- <https://drive.google.com/file/d/0B2-rlCGKD40NNXQwNm1LeUJpX1U/edit?usp=sharing>
- <https://drive.google.com/file/d/0B2-rlCGKD40NRDlwY19FYWNNQ0U/edit?usp=sharing>
- <https://drive.google.com/file/d/0B2-rlCGKD40NRU1NcklzdDU1azA/edit?usp=sharing>
- <https://drive.google.com/file/d/0B2-rlCGKD40NTDI2ODVaTDB3SE0/edit?usp=sharing>

Step 1: Download the “VMware player” from the link shown above and install it.



Step 2: Download the “Cloudera Setup File” from any of the above links and extract that zipped file on your hard drive.

Overview

Downloads

- Cloudera Manager Downloads
- CDH Downloads
- Cloudera QuickStart VM
- Learn Hadoop
- Get Support

Log In

- Log In
- Sign Up

Customer Portal Feedback

Have comments about this page?

Let us know!

Customer Portal [Log In](#) | [Register](#)

[Downloads](#) > Cloudera QuickStart VM

Cloudera QuickStart VM

This VM contains a sample of Cloudera's platform for "Big Data". Although the true power of Hadoop comes when it can be distributed across hundreds, even thousands of nodes, this VM makes it easy for you to learn about industry-leading tools without having to set up a full cluster.

Last Updated: May 2013

The VM from Cloudera are available in VMware, VirtualBox and KVM flavors, and all require a 64 bit host OS. This VM runs CentOS 6.2 and includes CDH4.2, Cloudera Manager 4.5.2, and Cloudera Impala 1.0.

VMware

(SHA1: 523731a7c2899ec8650847045b28654468f191e)

VMware [Download](#)

Product Documentation
[Administrative Information](#)

KVM

(SHA1: dd8d8b615c4588ad8c91353461de9c15e4d0a57c)

KVM [Download](#)

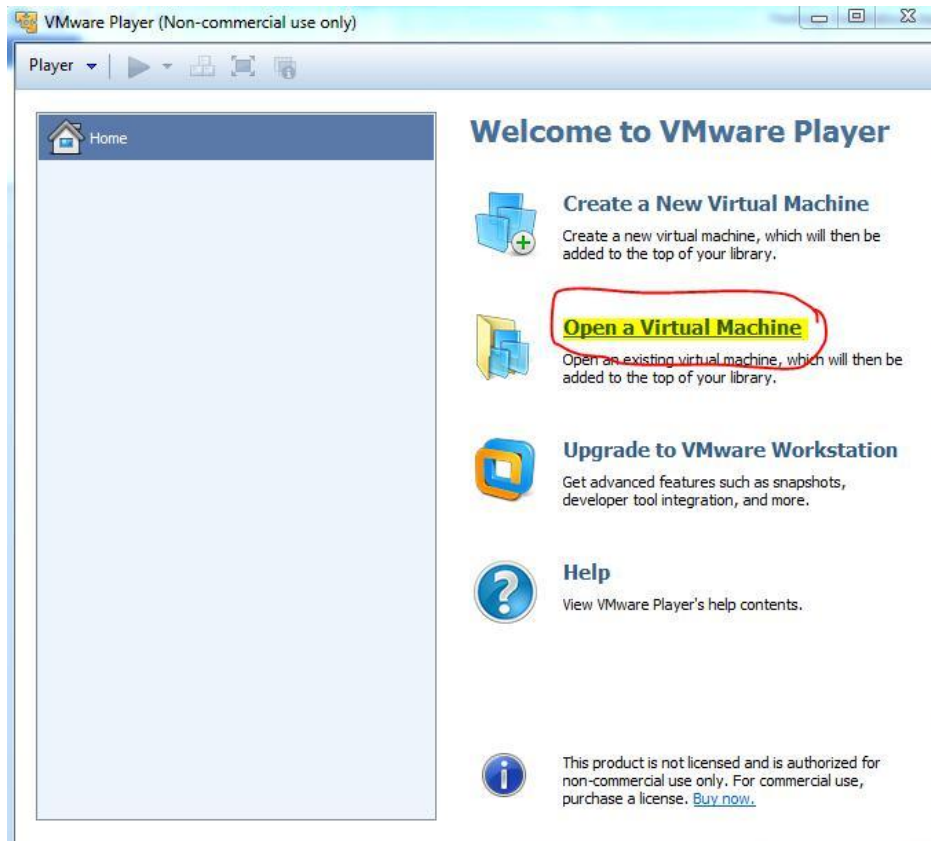
Scroll down and select Accept

Notwithstanding any terms to the contrary in this Agreement, Cloudera may choose to electronically deliver all communications with Customer, which may include email to Customer's email address indicated in Customer's communications with Cloudera. Cloudera's electronic communications to Customer may transmit or convey information about action taken on Customer's request, portions of Customer's request that may be incomplete or require additional explanation, any notices required under applicable law and any other notices. Customer agrees to do business electronically with Cloudera, and to receive electronically all current and future notices, disclosures, communications and information, and that the aforementioned provided electronically satisfies any legal requirement that such communications be in writing. An electronic notice will be deemed to have been received the day of receipt as evidenced by such email.

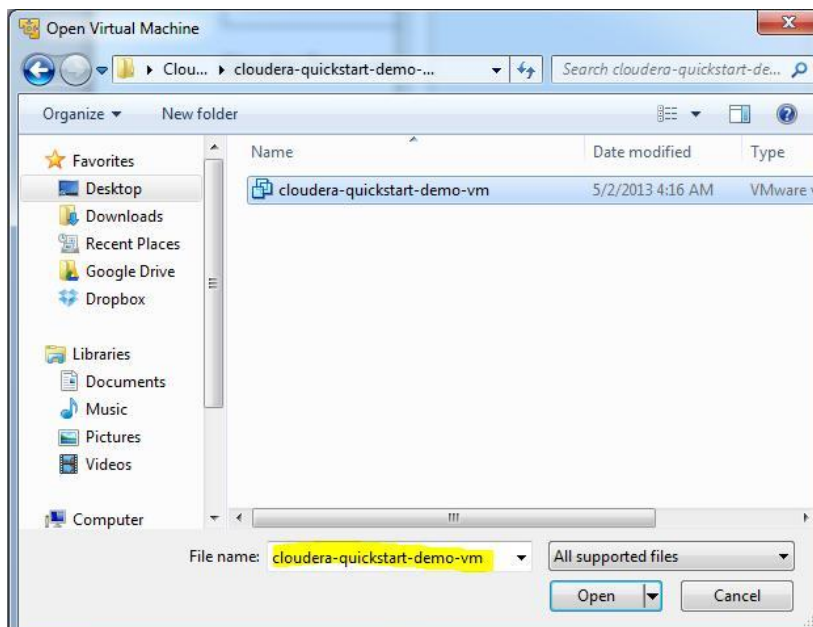
[Accept](#)

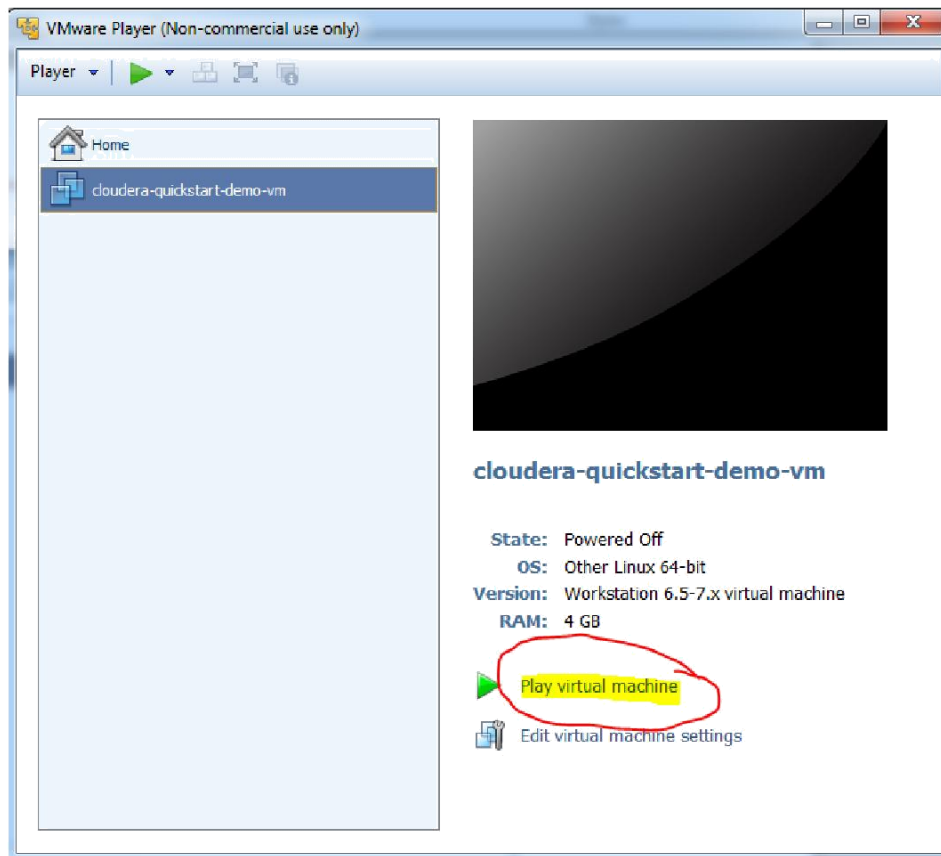
[Reject](#)

Step 3: Start VMPlayer and click open a Virtual Machine.



Browse the extracted folder.





It will start in a couple of minutes.

Login credentials:

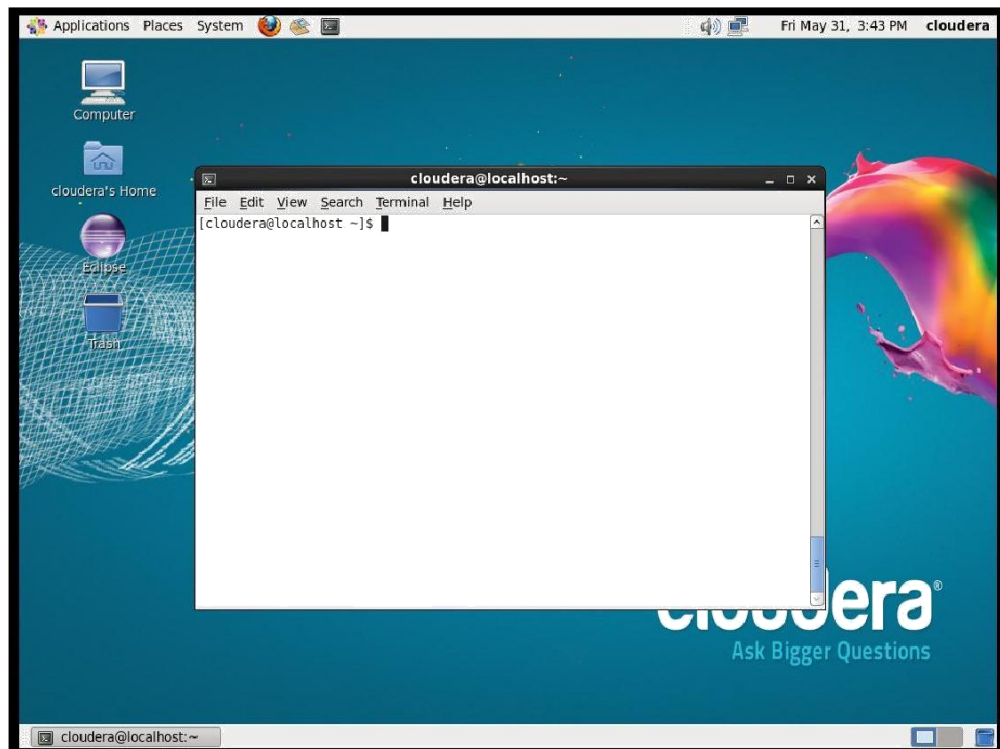
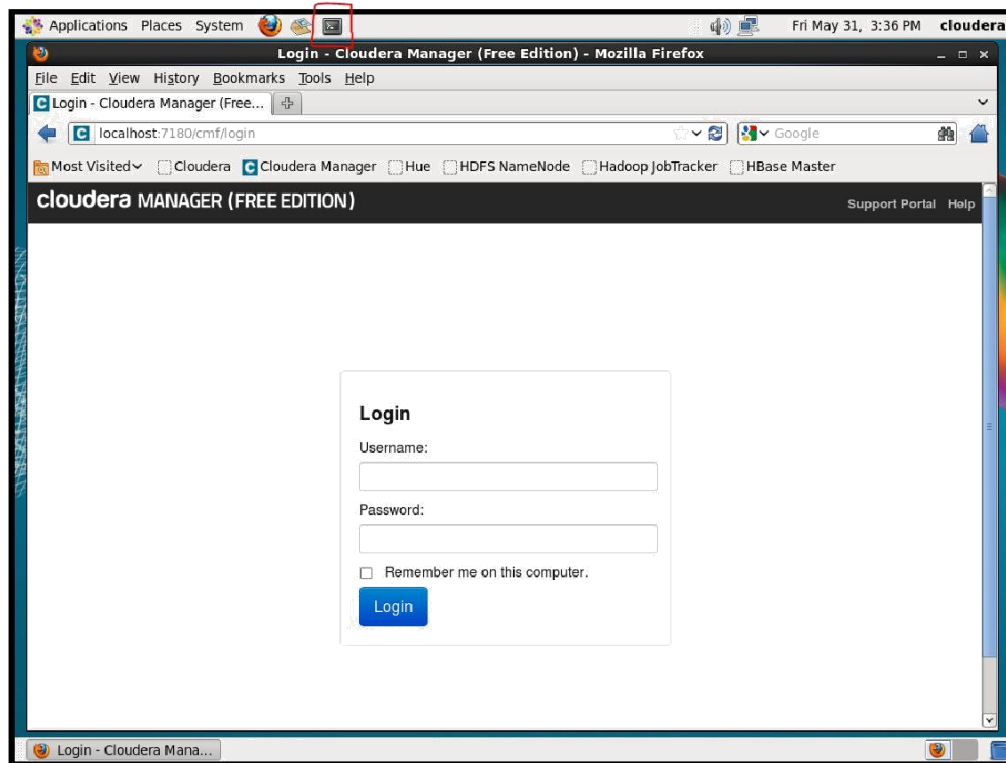
Machine Login credentials are:

- a. Username - **cloudera**
- b. Password - **cloudera**

Cloudera Manager Credentials are:

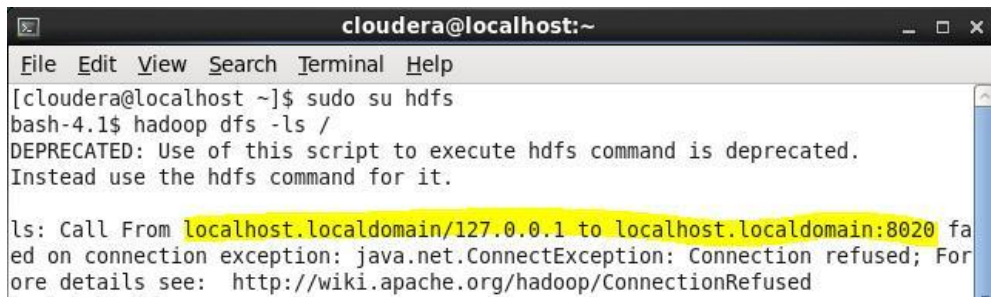
- a. Username - **admin**
- b. Password - **admin**

Click on the black box shown below in the image to start terminal.



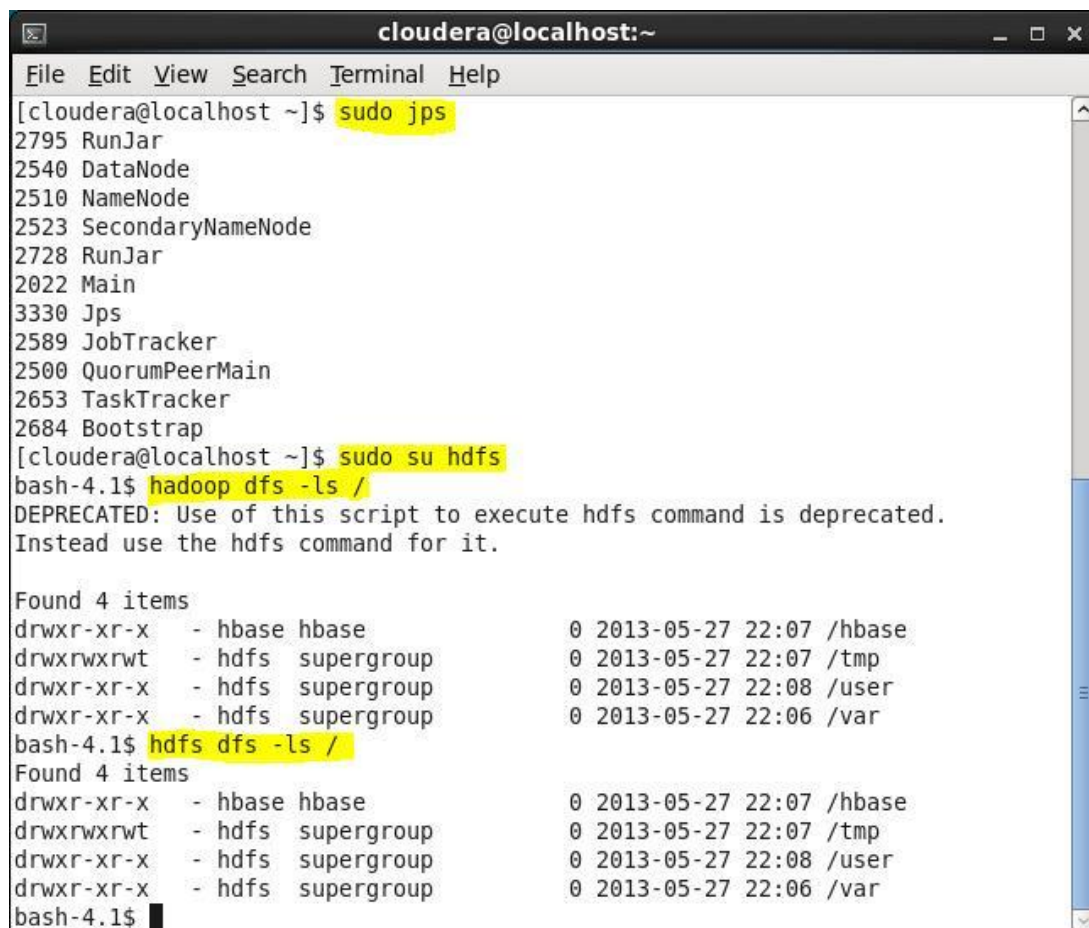
Step 4: Checking your Hadoop Cluster

- Type: `sudo jps` to see if all nodes are running (if you see an error like below, wait for some time and then try again, your threads are not started yet)



```
cloudera@localhost:~  
File Edit View Search Terminal Help  
[cloudera@localhost ~]$ sudo su hdfs  
bash-4.1$ hadoop dfs -ls /  
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.  
  
ls: Call From localhost.localdomain/127.0.0.1 to localhost.localdomain:8020 fa  
ed on connection exception: java.net.ConnectException: Connection refused; For  
ore details see: http://wiki.apache.org/hadoop/ConnectionRefused
```

- Type: `sudo su hdfs`
- Execute your command ie – `hadoop dfs -ls /`



```
cloudera@localhost:~  
File Edit View Search Terminal Help  
[cloudera@localhost ~]$ sudo jps  
2795 RunJar  
2540 DataNode  
2510 NameNode  
2523 SecondaryNameNode  
2728 RunJar  
2022 Main  
3330 Jps  
2589 JobTracker  
2500 QuorumPeerMain  
2653 TaskTracker  
2684 Bootstrap  
[cloudera@localhost ~]$ sudo su hdfs  
bash-4.1$ hadoop dfs -ls /  
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.  
  
Found 4 items  
drwxr-xr-x - hbase hbase 0 2013-05-27 22:07 /hbase  
drwxrwxrwt - hdfs supergroup 0 2013-05-27 22:07 /tmp  
drwxr-xr-x - hdfs supergroup 0 2013-05-27 22:08 /user  
drwxr-xr-x - hdfs supergroup 0 2013-05-27 22:06 /var  
bash-4.1$ hdfs dfs -ls /  
Found 4 items  
drwxr-xr-x - hbase hbase 0 2013-05-27 22:07 /hbase  
drwxrwxrwt - hdfs supergroup 0 2013-05-27 22:07 /tmp  
drwxr-xr-x - hdfs supergroup 0 2013-05-27 22:08 /user  
drwxr-xr-x - hdfs supergroup 0 2013-05-27 22:06 /var  
bash-4.1$
```

Step 5: Download the list of Hadoop commands for reference from the following link: http://hadoop.apache.org/docs/r1.0.4/commands_manual.pdf

Don't be afraid of this file, all important Hadoop commands will be explained in the class.