

# Practical 1: To Install Cloudera QuickStart VM on VMware

Cloudera is software that provides a platform for data analytics, data warehousing, and machine learning. Initially, Cloudera started as an open-source Apache Hadoop distribution project, commonly known as Cloudera Distribution for Hadoop or CDH. It contains Apache Hadoop and other related projects where all the components are 100% open-source under Apache License.

Cloudera provides virtual machine images of complete Apache Hadoop clusters, making it easy to get started with Cloudera CDH. This assignment covers the following topics related to Cloudera QuickStart VM Installation:

1. What is Cloudera QuickStart VM?
2. Prerequisites for Cloudera QuickStart VM Installation
3. Downloading the Cloudera QuickStart VM
4. Cloudera QuickStart VM Installation on Windows

What is Cloudera QuickStart VM?

Cloudera QuickStart VM includes everything you need for using CDH, Impala, Cloudera Search, and Cloudera Manager. The Cloudera QuickStart VM uses a package-based install that allows you to work with or without the Cloudera Manager. It provides a sample of Cloudera's platform for "Big Data."

Prerequisites for Cloudera QuickStart VM Installation



- A virtual machine such as Oracle VirtualBox or VMware
- RAM of 12+ GB, that is 4+ GB for the operating system and 8+ GB for Cloudera
- 80 GB hard disk
- Oracle VirtualBox downloaded from <https://www.virtualbox.org/wiki/Downloads> and installed on your system























Downloading the Cloudera QuickStart VM

- Cloudera QuickStart VMs are available as Zip archives in VirtualBox, VMware, and KVM formats. To download the VM, go to <https://www.cloudera.com/downloads.html> and select the appropriate version of CDH that you require.
- Click on the "GET IT NOW" button, and it will prompt you to fill in your details.

- Once the file is downloaded, go to the download folder and unzip the files. They can then be used to set up a single-node Cloudera cluster.

The two virtual images of Cloudera QuickStart VM are shown below:

Name	Date modified	Type	Size
 cloudera-quickstart-vm-5.13.0-0-vmware	20-03-2023 04:56 PM	File folder	
 cloudera-quickstart-vm-5.13.0-0-vmware.zip	17-03-2023 12:13 AM	ZIP File	56,22,959 KB

Name	Date modified	Type	Size
 cloudera-quickstart-vm-5.13.0-0-vmware.nvram	23-10-2017 05:40 PM	VMware Virtual M...	9 KB
 cloudera-quickstart-vm-5.13.0-0-vmware.vmdk	23-10-2017 05:33 PM	VMware virtual dis...	2 KB
 cloudera-quickstart-vm-5.13.0-0-vmware.vmsd	23-10-2017 04:51 PM	VMware snapshot ...	0 KB
 cloudera-quickstart-vm-5.13.0-0-vmware.vmx	23-10-2017 05:40 PM	VMware virtual m...	3 KB
 cloudera-quickstart-vm-5.13.0-0-vmware.vmxs	23-10-2017 04:51 PM	VMware Team Me...	1 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s001.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	7,01,184 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s002.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	25,45,280 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s003.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	17,31,136 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s004.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	94,016 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s005.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	6,66,240 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s006.vmdk	23-10-2017 05:42 PM	VMware virtual dis...	15,64,352 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s007.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	1,56,288 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s008.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	41,408 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s009.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	704 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s010.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	1,22,112 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s011.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	40,128 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s012.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	1,024 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s013.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	9,55,520 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s014.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	4,08,320 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s015.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	960 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s016.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	512 KB
 cloudera-quickstart-vm-5.13.0-0-vmware-s017.vmdk	23-10-2017 05:43 PM	VMware virtual dis...	128 KB

- Now that the downloading process is done with, let's move forward with this Cloudera QuickStart VM Installation guide and see the actual process.

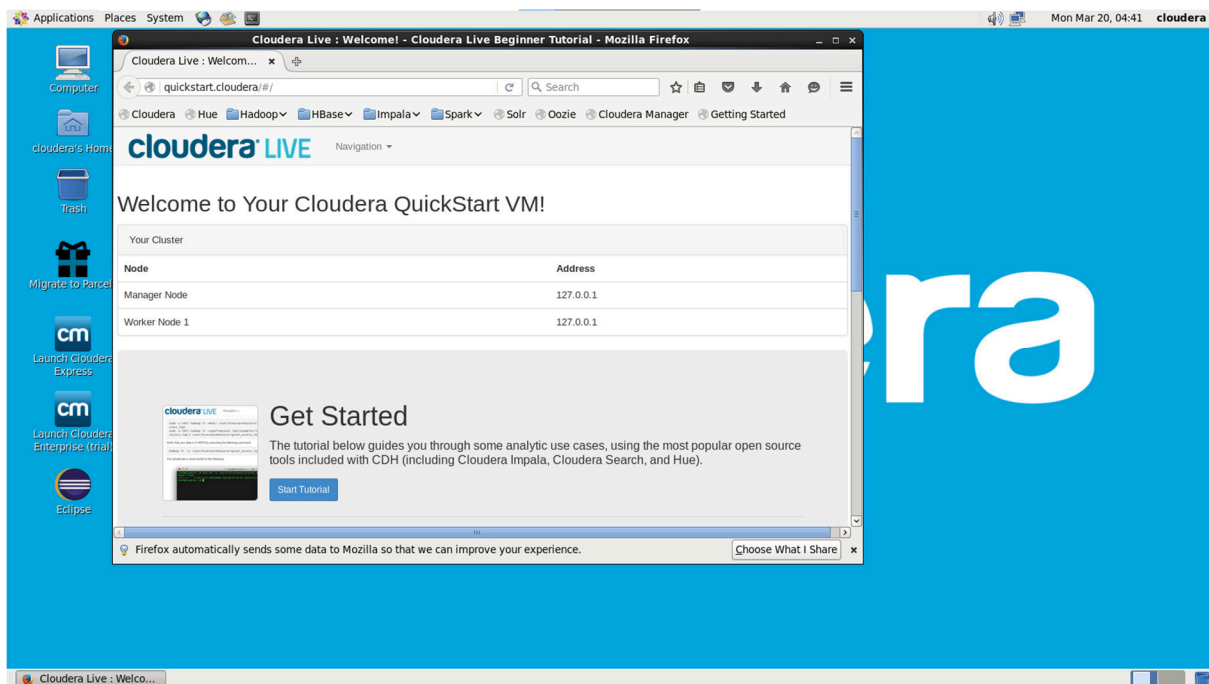
#### Cloudera QuickStart VM Installation

- Before setting up the Cloudera Virtual Machine, you would need to have a virtual machine such as VMware or Oracle VirtualBox on your system.
- In this case, we are using Oracle VirtualBox to set up the Cloudera QuickStart VM.
- In order to download and install the Oracle VirtualBox on your operating system, click on the following link: Oracle VirtualBox(<https://www.virtualbox.org/wiki/Downloads>).

- To set up the Cloudera QuickStart VM in your Oracle VirtualBox Manager, click on file with extension as “.VMX “ and then it will automatically open in VMware Workstation.
- Once complete you can see the Cloudera QuickStart VM on the left side panel.



- The next step will be going ahead and starting the machine by clicking the ‘Start’ symbol on top.
- Once your machine comes on, it will look like this:



- Next, we have to follow a few steps to gain admin console access. You need to click on the terminal present on top of the desktop screen, and type in the following:

1. `hostname #` This shows the hostname which will be `quickstart.cloudera`
  2. `hdfs dfs -ls / #` Checks if you have access and if your cluster is working. It displays what exists on your HDFS location by default
  3. `service cloudera-scm-server status #` Tells what command you have to type to use cloudera express free
  4. `su - #` Login as root
  5. `service cloudera-scm-server status #` The password for root is cloudera
- Once you see that your HDFS access is working fine, you can close the terminal. Then, you have to click on the following icon that says 'Launch Cloudera Express'.



- You are required to copy the command, and run it on a separate terminal. Hence, open a new terminal, and use the below command to close the Cloudera based services. It will restart the services, after which you can access your admin console.

```
File Machine View Input Devices Help
Applications Places System
Browse and run installed applications cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ sudo /home/cloudera/cloudera-manager --force --express
[QuickStart] Shutting down CDH services via init scripts...
kafka-server: unrecognized service
[QuickStart] Disabling CDH services on boot...
error reading information on service kafka-server: No such file or directory
[QuickStart] Starting Cloudera Manager server...
[QuickStart] Waiting for Cloudera Manager API...
[QuickStart] Starting Cloudera Manager agent...
[QuickStart] Configuring deployment...
[QuickStart] Deploying client configuration...
[QuickStart] Starting Cloudera Management Service...
[QuickStart] Enabling Cloudera Manager daemons on boot...

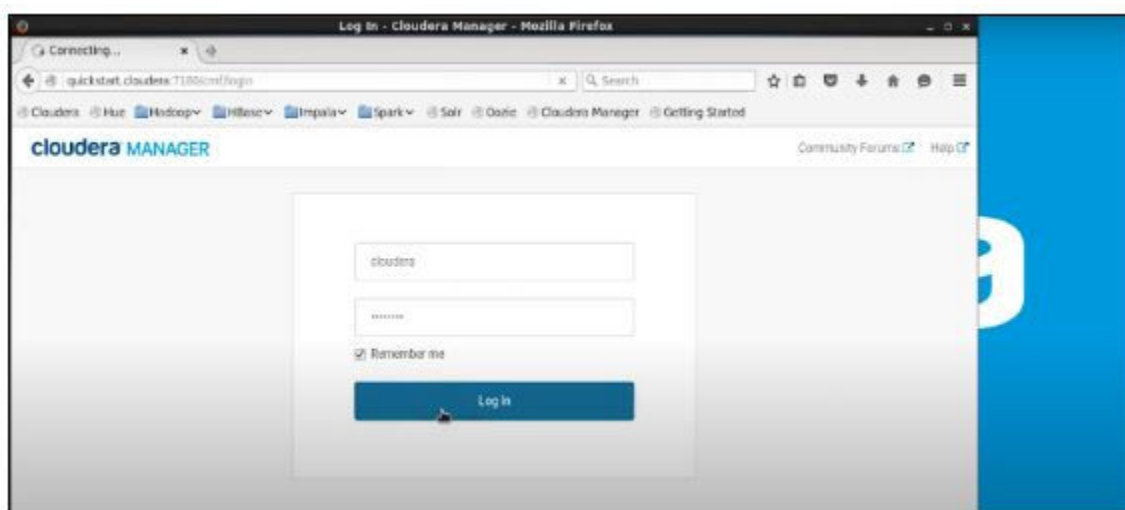
Success! You can now log into Cloudera Manager from the QuickStart VM's browser:

http://quickstart.cloudera:7180

Username: cloudera
Password: cloudera

[cloudera@quickstart ~]$
```

- Now that our deployment has been configured, client configurations have also been deployed. Additionally, it has restarted the Cloudera Management Service, which gives access to the Cloudera QuickStart admin console with the help of a username and password.
- Go on and open up the browser and change the port number to 7180.
- You can log in to the Cloudera Manager by providing your username and password.



- You can go ahead and restart the services now. It will ensure that the cluster becomes accessible either by Hue as a web interface or Cloudera QuickStart Terminal, where you can write your commands.