

# Hortonworks Sandbox (HDP 2.2.4) Installation Guide

## Overview

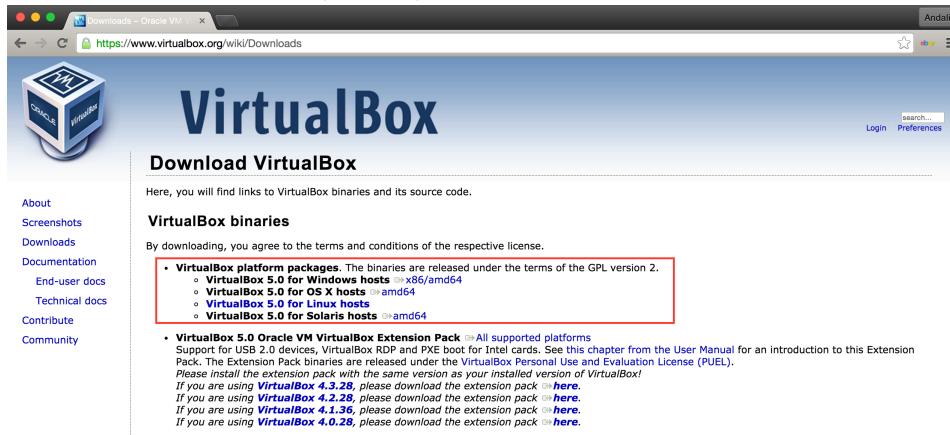
There are various ways to get started with Hadoop. For a beginner, I found Hortonworks Sandbox is an easiest way to get started with Hadoop Ecosystem. The Hortonworks Sandbox is a piece of software that runs on virtual machine. It has everything preconfigured (Hadoop and its component) that you need to get started with Hadoop.

## Prerequisites:

- RAM: At least 4 GB (the more, the better)
- CPU: 64 bit
- OS: Mac/Windows
- Browser: Chrome (version 25+)

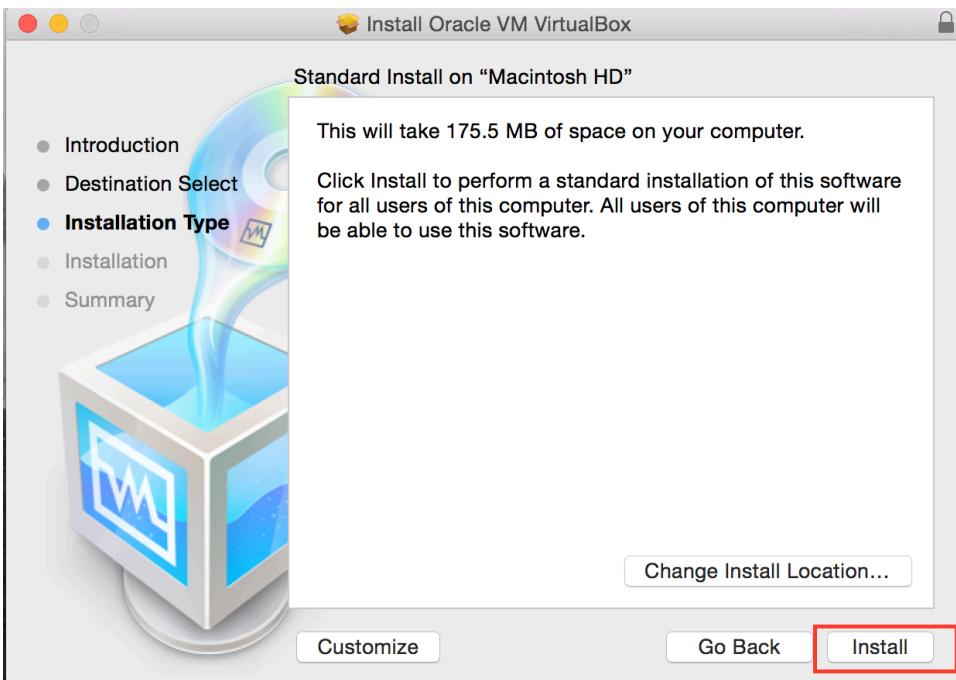
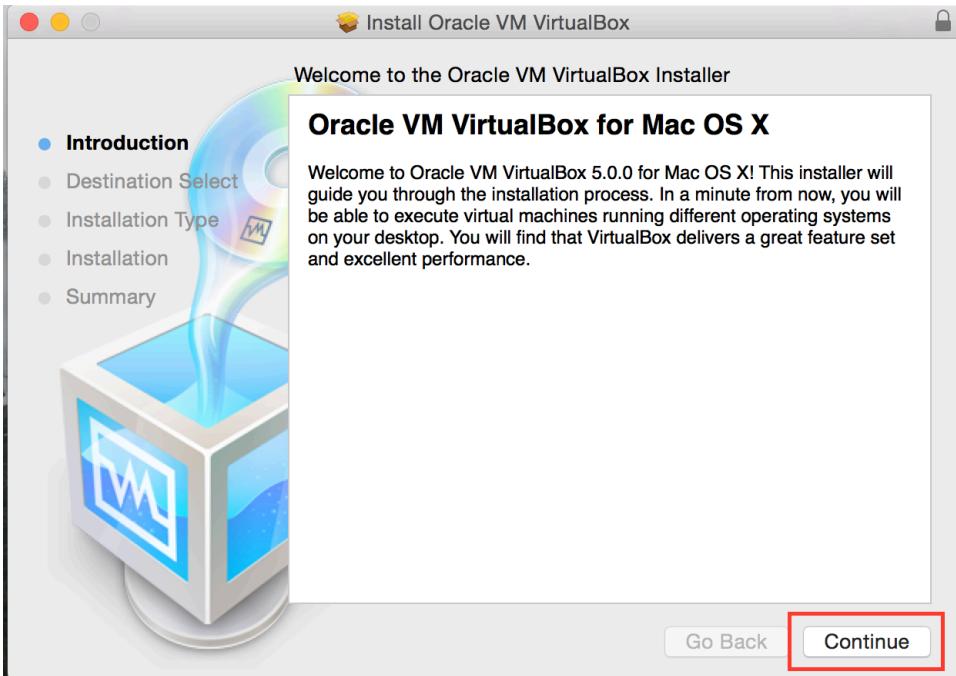
## Installation Steps:

1. Download Oracle Virtualbox (version 5.0) from <https://www.virtualbox.org/wiki/Downloads>.  
Download the file with respect to your OS.

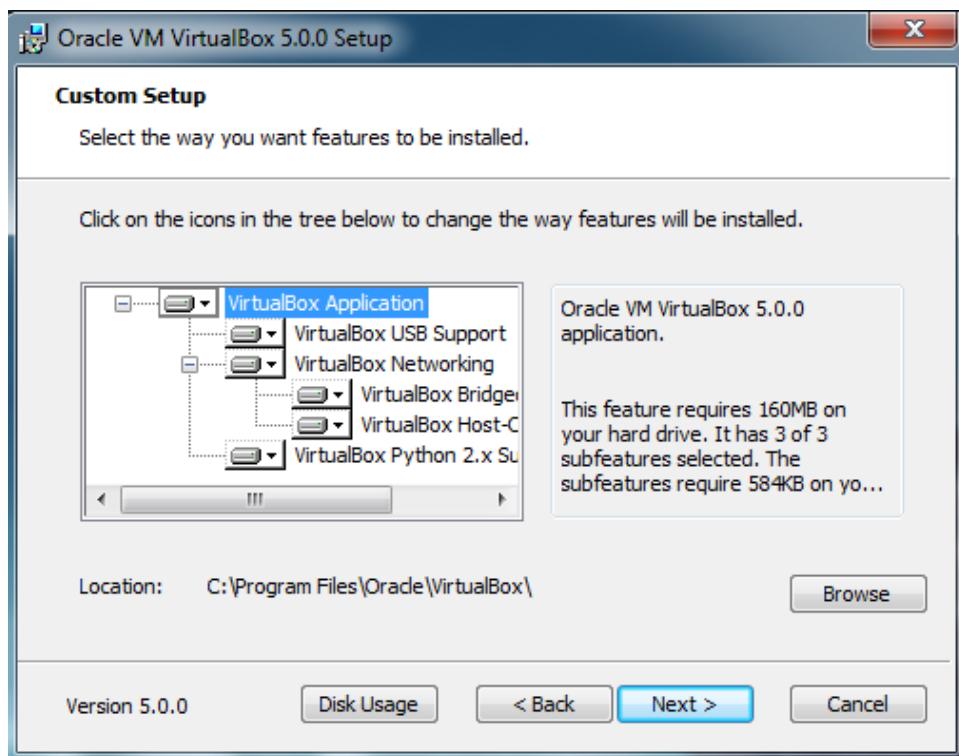


2. Install Virtualbox. Follow the screen instructions

For Mac:



For Windows: You just need to click next button few times

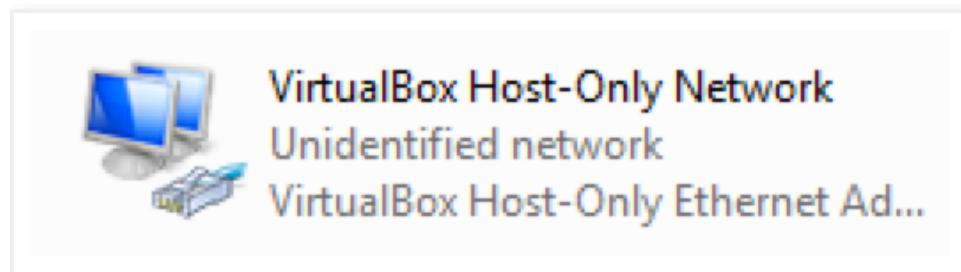


If you get any warning issues, just ignore it. Click yes.



**Note:** In worst case, if you face any network issue after installing Virtualbox, then please follow below instructions:

Start Control Panel->Network and Internet->Network and Sharing Center->Change Adapter Settings. You should see a network connection as below. Right click and disable it. You may have to reconnect your internet once again.



3. Download Hortonworks Sandbox (HDP 2.2.4) from  
<http://hortonworks.com/downloads/#sandbox>

Click on Expand:

The screenshot shows a web browser with the URL [hortonworks.com/downloads/#sandbox](http://hortonworks.com/downloads/#sandbox). The page has a navigation bar with tabs: SANDBOX (which is active), DATAFLOW, DATA PLATFORM, and TECH PREVIEW. Below the tabs, there are three expandable sections for HDP 2.4:

- Tutorials | Release Notes | Import on Virtual Box | MD5 : d42a9bd11f29775cc5b804ce82a72efd** **DOWNLOAD FOR VIRTUALBOX**
- Tutorials | Release Notes | Import on VMware | MD5 : f1d45e93ab9f2a655db559be5b2f2f43** **DOWNLOAD FOR VMWARE**
- Tutorials | Release Notes | Import on Docker | MD5 : c613fab7ed21e15886ab23d7a28aec8a** **DOWNLOAD FOR DOCKER**

### Hortonworks Sandbox in the Cloud

HDP 2.4 on Azure with Hortonworks Sandbox

[Tutorial: Sandbox on Azure](#) | Try it one month for free

**ONE MONTH TRIAL**

Hortonworks Sandbox Archive

**EXPAND ▾**

### Hortonworks DataFlow

☰ MENU

SHARE [in](#) [f](#) [t](#) NEWSLETTER [✉](#)

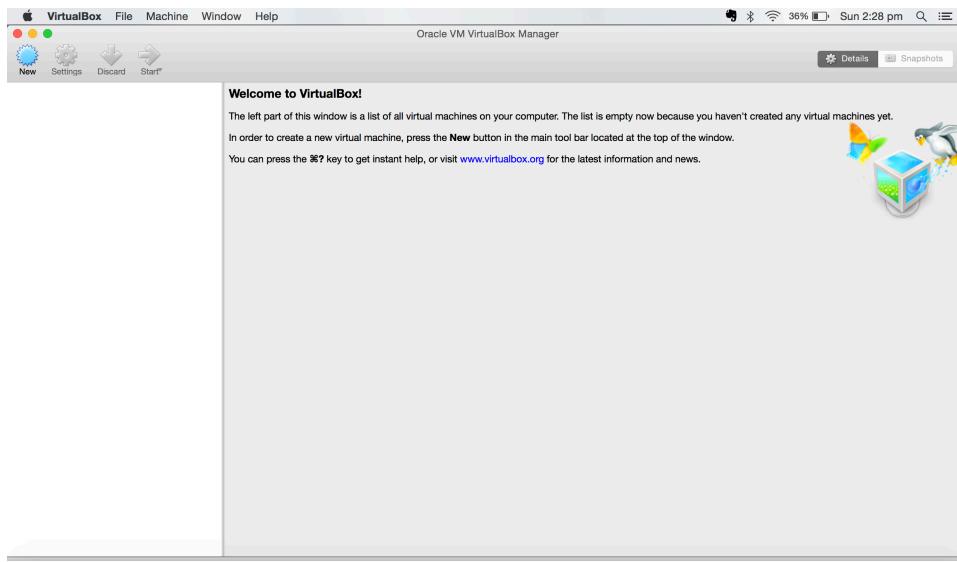
Contact Sales? [✉](#)

Scroll to HDP 2.2.4 on Hortonworks Sandbox, then click on Download for Virtualbox:

The screenshot shows the same web browser interface as the previous one, but with different sections expanded:

- HDP 2.3.2 on Sandbox**
  - [VirtualBox Install Guide](#) **DOWNLOAD FOR VIRTUALBOX**
  - [VMWare Install Guide](#) **DOWNLOAD FOR VMWARE**
- HDP 2.2.4 on Hortonworks Sandbox**
  - [VirtualBox Install Guide](#) **DOWNLOAD FOR VIRTUALBOX**
  - [VMWare Install Guide](#) **DOWNLOAD FOR VMWARE**
  - [HyperV Install Guide](#) **DOWNLOAD FOR HYPERV**
- HDP 2.2 on Hortonworks Sandbox**
  - [VirtualBox Install Guide \(Mac\)](#) | [VirtualBox Install Guide \(Windows\)](#) **DOWNLOAD FOR VIRTUALBOX**

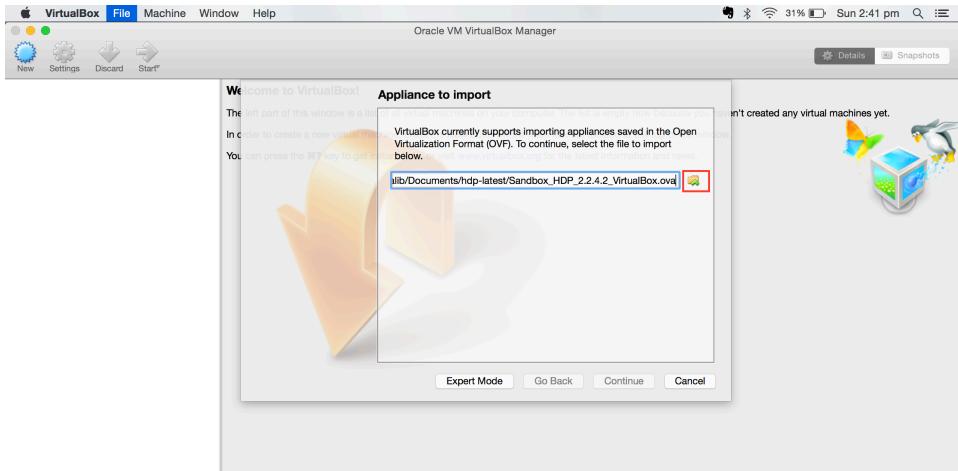
- Once virtualbox installation is finished, start your virtualbox.



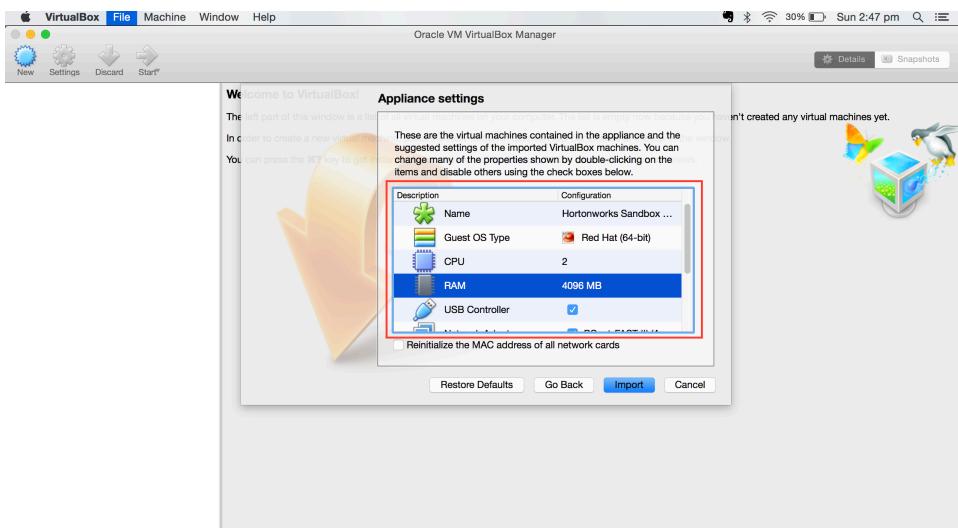
- Import Hortonworks Sandbox file (Sandbox\_HDP\_2.2.4.2\_VirtualBox.ova which you have downloaded from Hortonworks website) into Virtualbox:

**For Mac:**



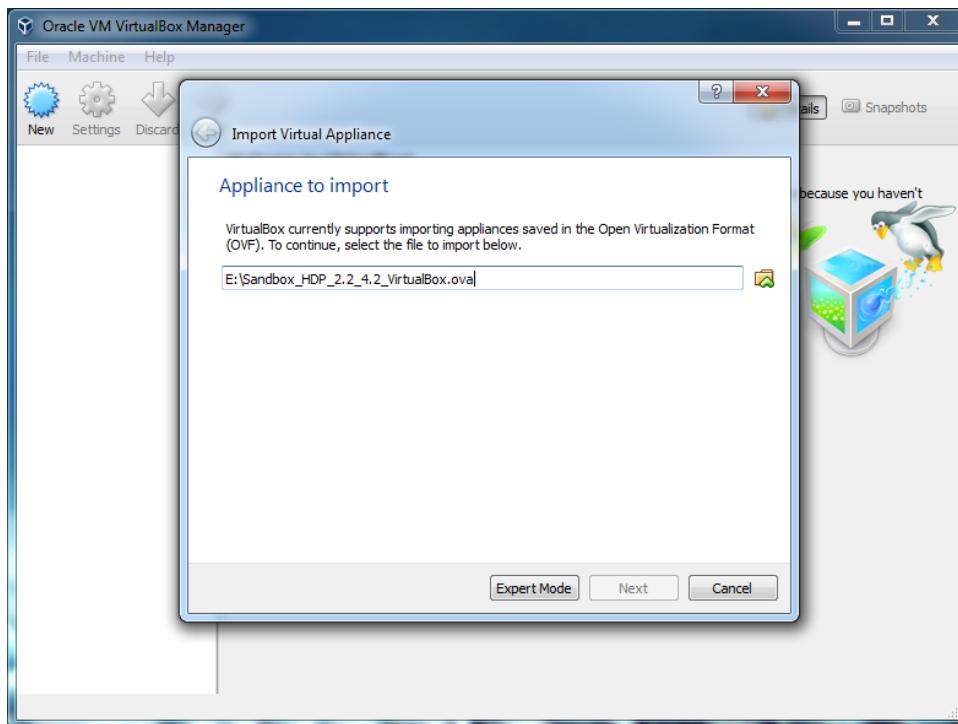
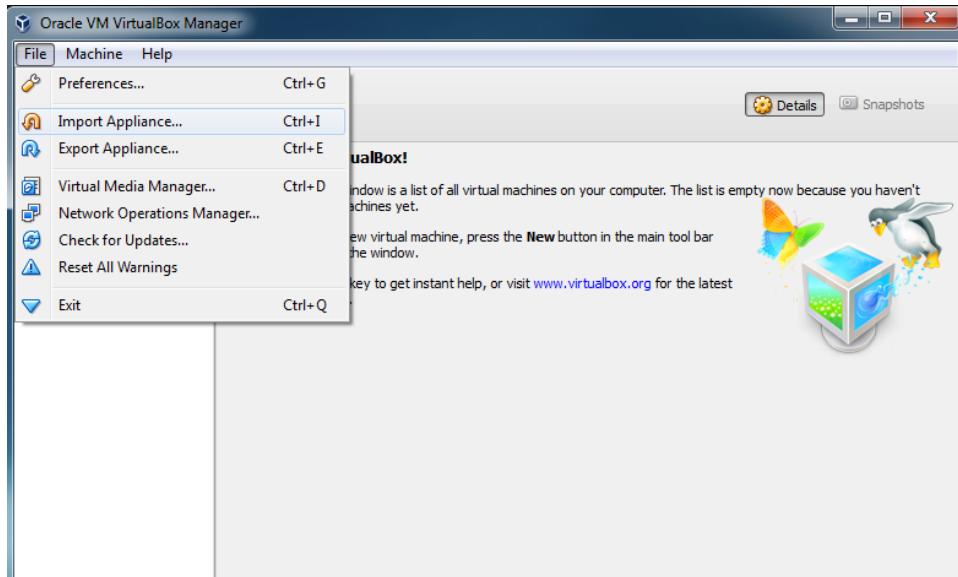


Change your virtualbox setting according to your hardware (like RAM size, number of CPUs etc). If you have 4 GB RAM on your machine, then assign 2.5 GB RAM. But my suggestion would be to have at least 8 GB RAM on your machine and assign 4-5 GB of RAM to your virtualbox to avoid any machine hanging issue or slow tasks.

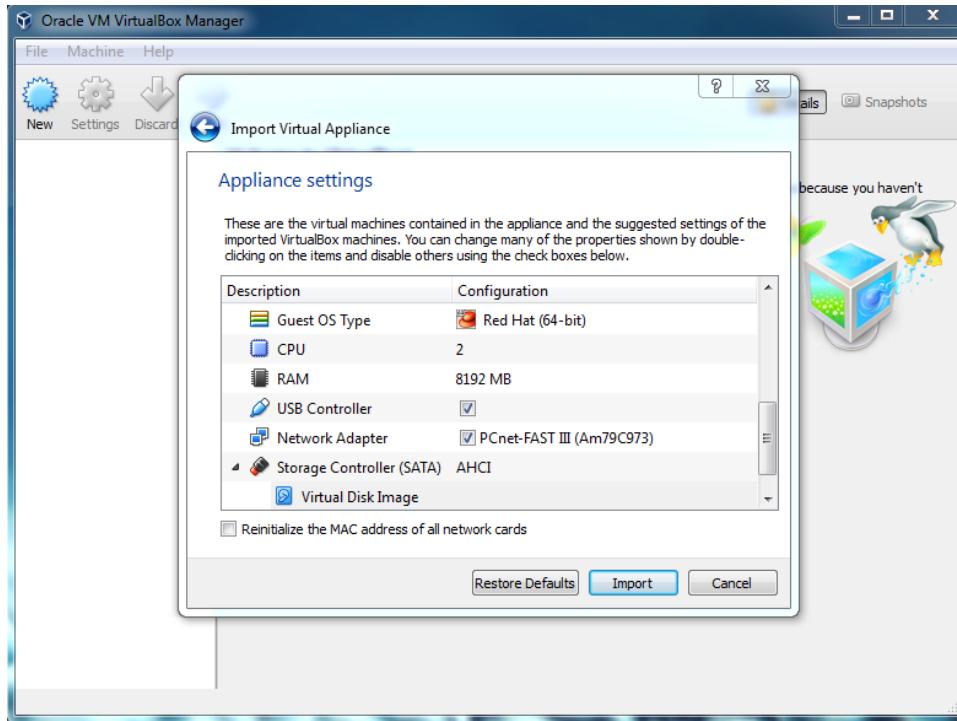


### For Windows:

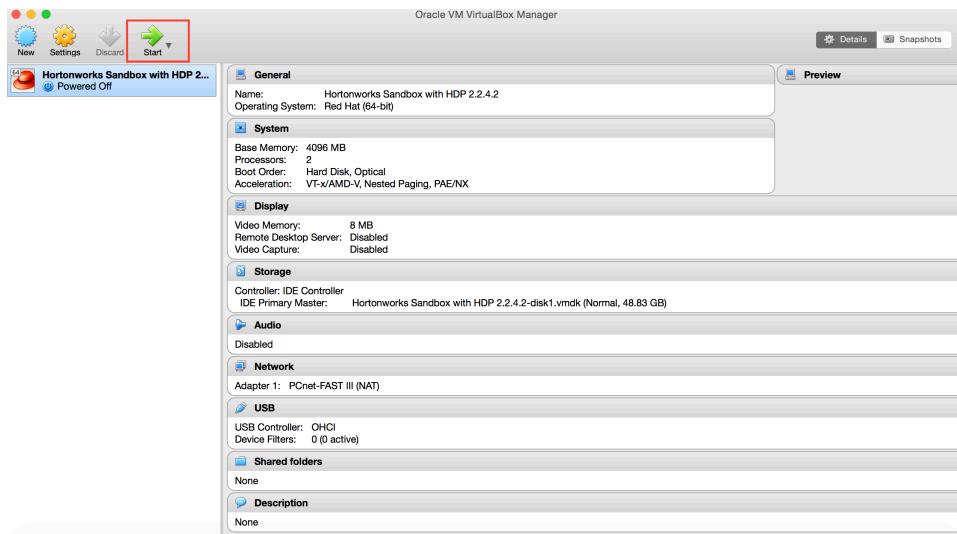
click on Import Appliance



Change your virtualbox setting according to your hardware (like RAM size, number of CPUs etc). If you have 4 GB RAM on your machine, then assign 2.5 GB RAM. But my suggestion would be to have at least 8 GB RAM on your machine and assign 4-5 GB of RAM to your virtualbox to avoid any machine hanging issue or slow tasks.



- Once Import process is completed, start your Sandbox (For both Mac or Windows, the screen will look like below).



After clicking the start button, it will open up a new window like below:

```
Hortonworks Sandbox with HDP 2.2 [Running]
Starting portreserve: [ OK ]
Starting system logger: [ OK ]
Starting tutorials... [ OK ]
Starting system message bus: [ OK ]
Mounting filesystems: [ OK ]
Retrigger failed udev events [ OK ]
Adding udev persistent rules [ OK ]
Starting the VirtualBox Guest Additions [ OK ]
Starting VirtualBox Guest Addition service [ OK ]
Starting sshd: [ OK ]
Starting postfix: [ OK ]
Starting httpd: [ OK ]
Apache Ranger Admin has started
Starting startup_script...
Updating IP...
Starting HDP ...
Starting mysql [ OK ]
Starting Knox ldap [ OK ]
Starting Postgre SQL [ OK ]
Starting name node [ OK ]
Starting secondary name node [ OK ]
Starting data node [ OK ]
Starting Oozie [ OK ]
```

Once your Sandbox is started, the screen should look like below:

```
Hortonworks Sandbox with HDP 2.2 [Running]
Hortonworks Sandbox with HDP 2.2
http://hortonworks.com

To initiate your Hortonworks Sandbox session,
please open a browser and enter this address
in the browser's address field:
http://127.0.0.1:8888/

You can access SSH by $ ssh root@127.0.0.1 -p 2222

Log in to this virtual machine: Linux/Windows <Alt+F5>, Mac OS X <Fn+Alt+F5>
```

Now your portable Hadoop Environment (i.e. Hortonworks Sandbox) is started. There are two ways to log in into Sandbox (i.e. Hadoop):

- You can log in into the same window (above shown window) by pressing 'Alt+F5' (for Windows) or 'Fn+Alt+F5' (for Mac). Username- **root** and Password- **hadoop**

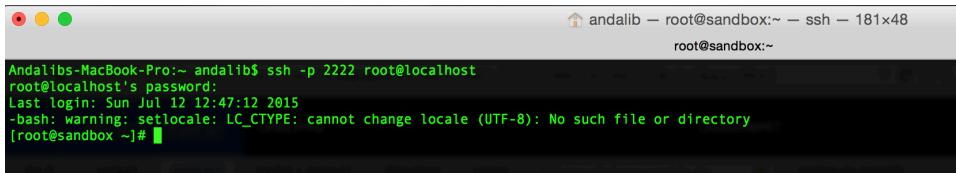
b) Or you can access Hadoop by doing ssh from your local terminal.

**ssh root@127.0.0.1 -p 2222**

or

**ssh -p 2222 root@localhost**

Password- **hadoop**



A screenshot of a macOS terminal window titled "andalib — root@sandbox:~ — ssh — 181x48". The window shows the command "ssh -p 2222 root@localhost" being run, followed by the password prompt "root@localhost's password:". The terminal also displays the last login information and a warning message about locale settings. The prompt "[root@sandbox ~]#" is visible at the bottom.

**Note:** For Windows Users, you need to install a ssh client (download it from here: <http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html>)

**Known Issues during ssh connection:**

If you face any ssh connection issue like "Offending RSA key in /Users/andalib/.ssh/known\_hosts:1" or "Host key verification failed" then follow below step:

- cd /Users/user\_name/.ssh e.g. cd /Users/andalib/.ssh
- rm known\_hosts

or

if you are doing ssh connections to many servers and your offending key is at some line number (e.g. Offending key in /Users/username/.ssh/known\_hosts:6), then you should probably follow below steps:

- vi /Users/username/.ssh/known\_hosts:6
- :6
- dd
- press Esc key and type :wq

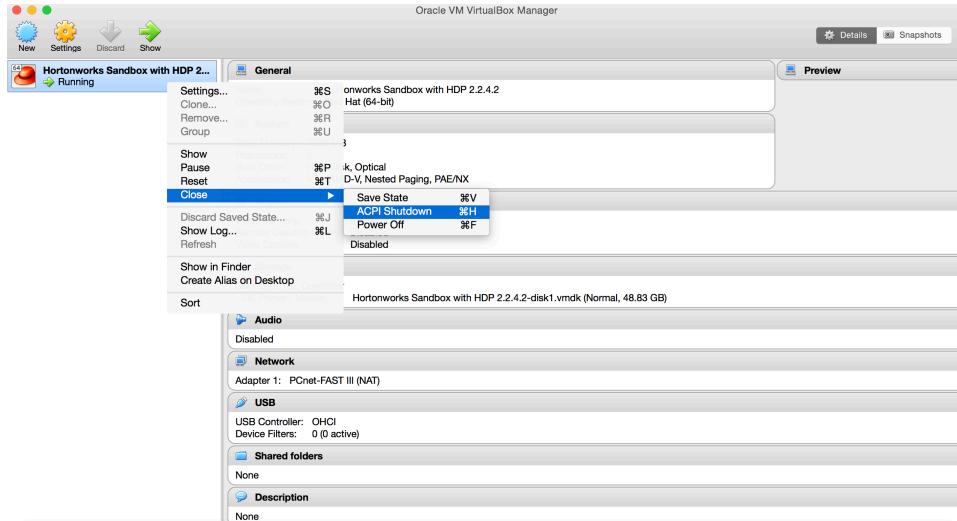
7. Hortonworks Sandbox provides a very nice web interface to interact with Hadoop and its component. Open your Chrome browser and hit <http://127.0.0.1:8000>. It will open up a registration form which will ask your basic details. Once you submit the form, you will be redirected to Sandbox web interface page. We will see this web interface in action in the coming lectures.

The screenshot shows the Hortonworks Sandbox with HDP 2.2 web interface. At the top, there's a navigation bar with links for Configuration, Check for misconfiguration, and Server Logs. Below the navigation bar, the title "Hortonworks Sandbox with HDP 2.2" is displayed. To the left, there's a logo with three elephants and the text "Leave Feedback". On the right, a table lists component names and their versions. The table includes the following rows:

Component	Version
Hue	2.6.1-2041
HDP	2.2.0
Hadoop	2.6.0
Pig	0.14.0
Hive-Hcatalog	0.14.0
Oozie	4.1.0
Ambari	1.7-169
	<input type="button" value="Enable"/>
HBase	0.98.4
Knox	0.5.0
...	...

**Note:** Kindly do not enable Ambari as it consumes a lot of RAM.

## 8. To Shutdown Sandbox, see below screen



**Congratulations, now you have a working Hadoop Ecosystem on your machine..!! :)**



