PROF 741

LAB 1: Installation Hortonworks Sandbox with VirtualBox

Virtual Machine Overview

The Hortonworks Sandbox is delivered as a virtual appliance. The virtual appliance (indicated by an .ovf or .ova extension in the filename) runs in the context of a virtual machine (VM), a piece of software that appears to be an application to the underlying (host) operating system (OS), but that looks like a bare machine, including CPU, storage, network adapters, and so forth, to the operating system and applications that run on it.

To use the Hortonworks Sandbox, one of the supported virtual machine applications needs to install on your host machine:

- **VirtualBox (Recommended)**
- VMWare Fusion
- Hyper-V

This document describes importing the Hortonworks Sandbox virtual appliance to VirtualBox.

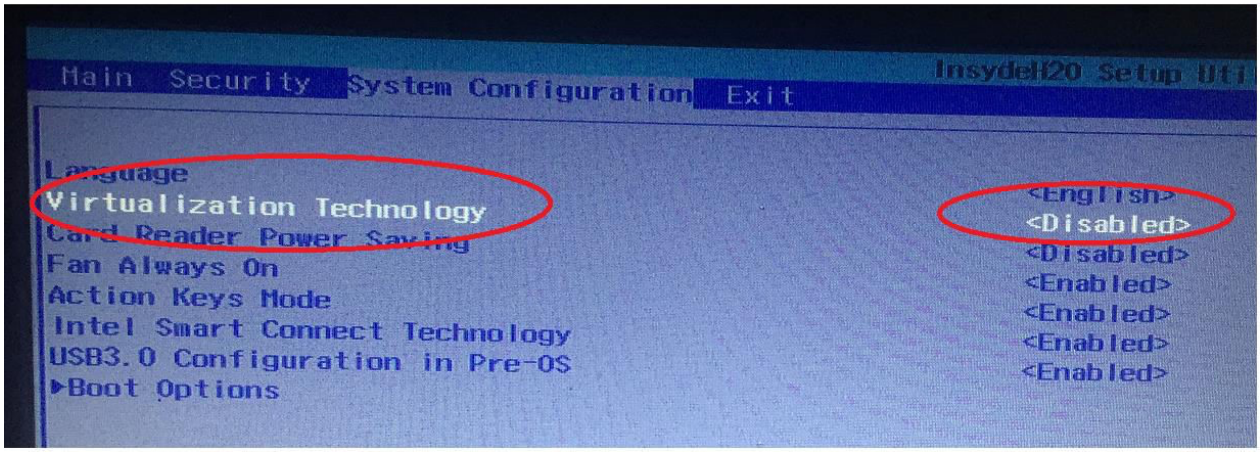
Prerequisites

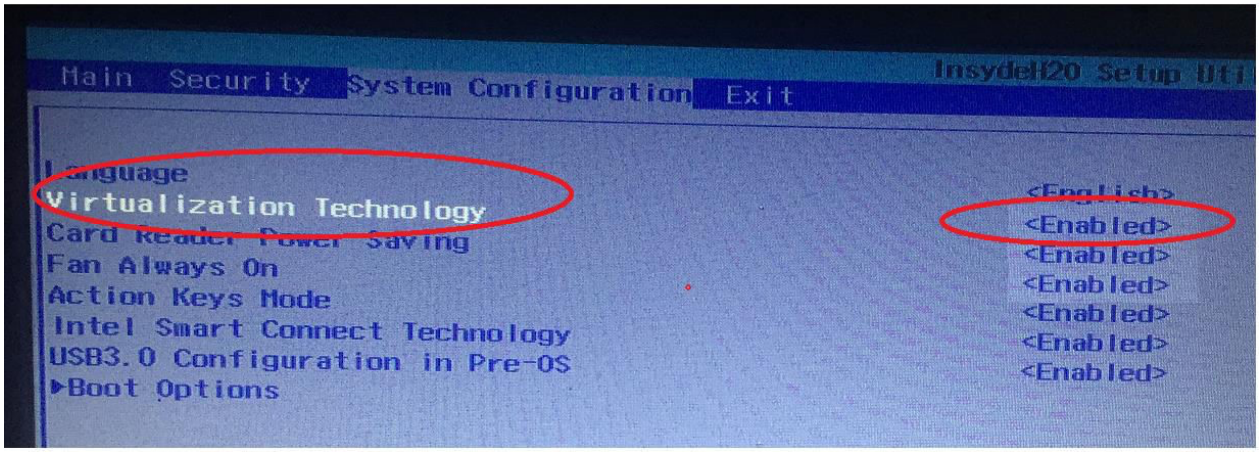
To use the Hortonworks Sandbox with VirtualBox the following requirements need to be met:

- **Oracle VirtualBox installed**
  Version 4.2 or later
- **Host Operating Systems:**
  Host operating system refers to the operating system of your computer.
- **Hardware** (The newer the hardware the better):
  - A 64-bit machine with a multi-core CPU that supports virtualization.
  - BIOS that has been enabled for virtualization support. Your Operating system (WINDOWS) must be able to support Hardware Virtualization. (In next step, you will get to know how to enable Virtualization enabled. It might already be enabled on your machine.

(OPTIONAL) Step 1: Go to your windows machine BIOS setting and enable the virtualization as shown in below. (Your machine could have little different way to do this)

To enter in BIOS, while starting your machine use F10 key ( )





- At least 4 GB of RAM

If you wish to enable services such as Ambari, HBase, Storm, Kafka,or Spark please ensure you have at least 10 Gb of physical RAM in order to run the VM using 8 GB.

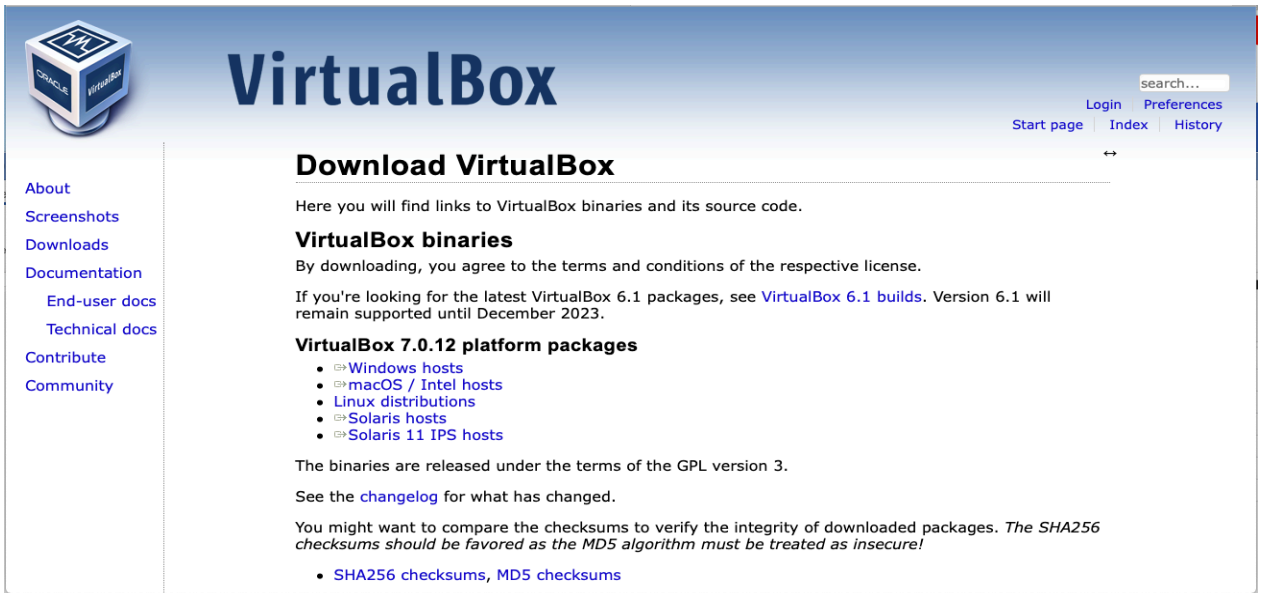
- **Browsers Supported**
- Chrome 25+,
- IE 9+ (Sandbox will not run on IE 10)
- Safari 6+

INSTALLATION ORACLE VIRTUALBOX

Procedure

1. Download the Oracle Virtual Box using the following URL:

URL: https://www.virtualbox.org/wiki/Downloads
Select the latest version available:



Step 2: After downloaded click on .exe or dmg to proceed for the installation. Once installation is finished you are done with installation of VirtualBox, and we need to install Hortonworks sandbox in the virtual box.

INSTALLING HDP Sandbox

Hortonworks Data Platform (HDP) is a straightforward, pre-configured, learning environment that contains the latest developments from Apache Hadoop. The Sandboxcomes packaged in a virtual environment that can run in the cloud or on your personal machine. The Sandbox allows you to learn and explore HDP on your own.

DOWNLOADING HDP Sandbox

You can find the download links of the Sandbox in ***.ova*** format with respect to the chosen hypervisor.

RECOMMENDED FOR THE LAB: If you are using VirtualBox, then download from here:
URL: https://archive.cloudera.com/hwx-sandbox/hdp/hdp-2.6.5/HDP_2.6.5_virtualbox_180626.ova

For VMware users, the download link is here:
URL: https://archive.cloudera.com/hwx-sandbox/hdp/hdp-2.6.5/HDP_2.6.5_vmware_180622.ova
You can also download them from the official website
URL: https://www.cloudera.com/downloads/hortonworks-sandbox/hdp.html
but it needs an account on Cloudera website.

INSTALLING the SANDBOX

The steps provided describe how to import the Hortonworks Sandbox virtual appliance into VirtualBox. The screenshots displayed are taken from Mac OS X and Windows machine running the VirtualBox software.

1. Open the **Oracle VM VirtualBox Manager**
   You can do so by double clicking the icon:

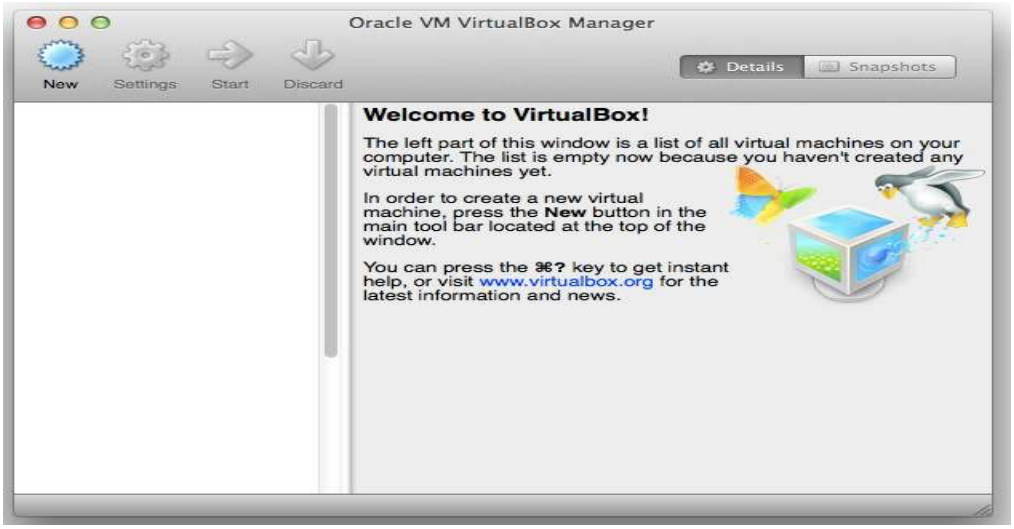
   

   On a Mac OS X machine, by default the application is installed within theApplications folder, hence the icon to launch the application can be foundthere.
   On a Windows machine, the executable (.exe) file is typically found in `c:\Program Files\Oracle\VirtualBox\VBoxManage.exe`

NOTE: For M1 and M2 based MacOS users, please install CentOs 7.5 or Ubuntu 18.04. Then using Docker where you need to install Docker on the virtual machine and run cluster containers.
Use the following image to download CentOs:

http://mirrors.coreix.net/centos-altarch/7.9.2009/isos/aarch64/CentOS-7-aarch64-Minimal-2009.iso

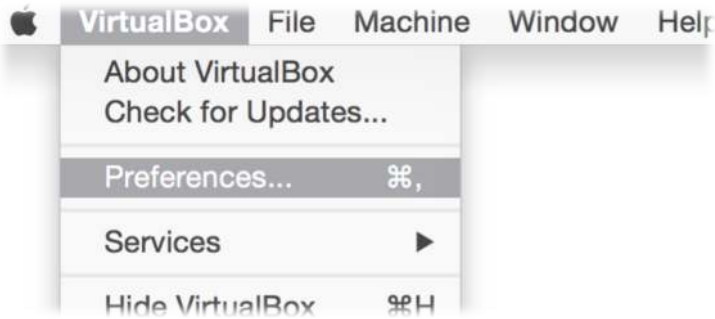
2. The **Oracle VM VirtualBox Manager** window opens.
   Mac OS X:



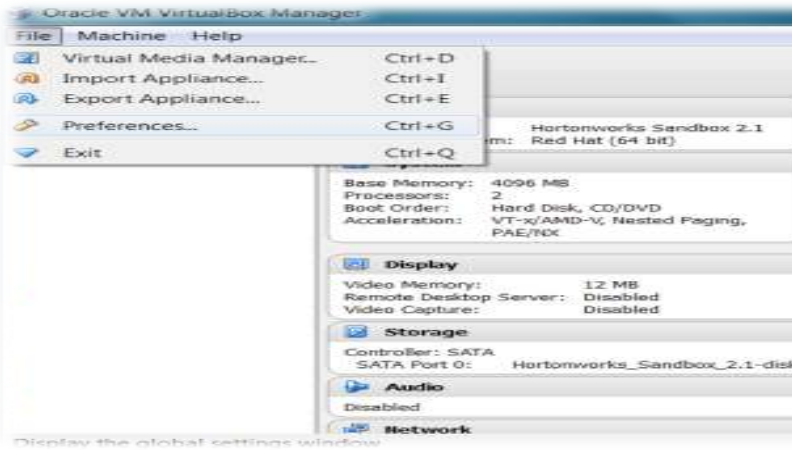
   Windows:



3. Open the preferences dialog window.
   Mac OS X: Select **VirtualBox->Preferences…** from the menu bar while the Oracle VM VirtualBox Manager application is in focus.

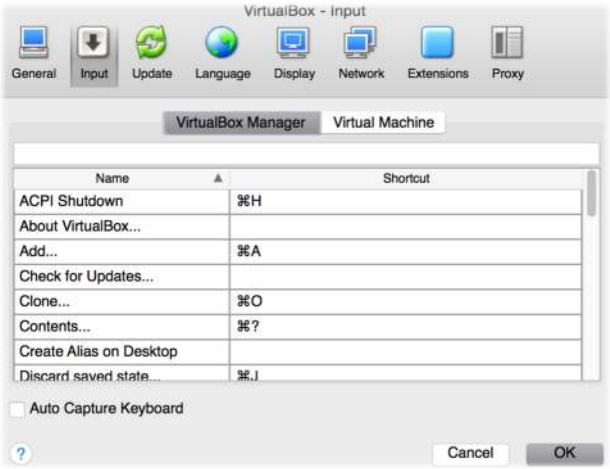


Windows:
Select **File->Preferences…** within the Oracle VM VirtualBox Manager



4. Uncheck **Auto-Capture Keyboard** within the Preferences dialog window.

   Mac OS X: Select the **Input** icon button from the top of the window first to getto the window containing the **Auto-Capture Keyboard** checkbox.



   Windows: Select the **Input** icon button from the left hand pane of the window first to get to the following window.



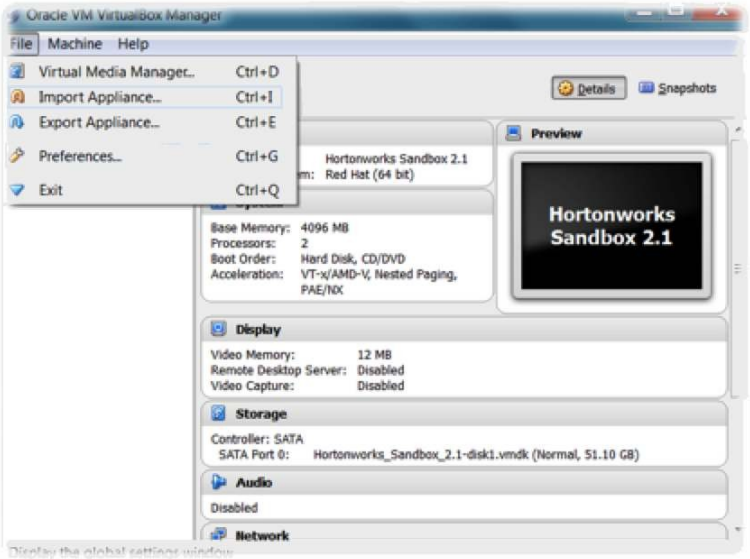
Click the **OK** button once done. This will close the Preferences window.

5. Open the Import Appliance window.

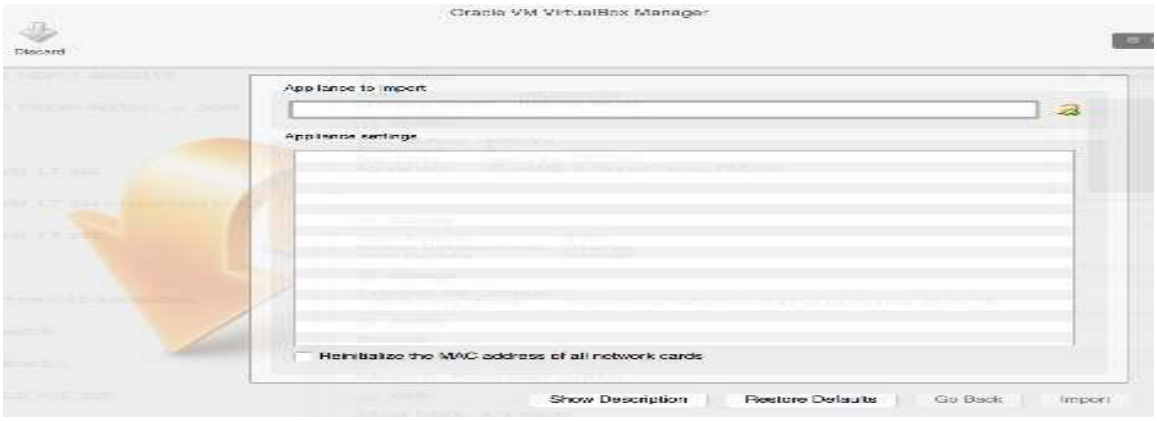
Mac OS X: Select **File->Import Appliance…** from the menu bar while the Oracle VM VirtualBox Manager application is in focus.



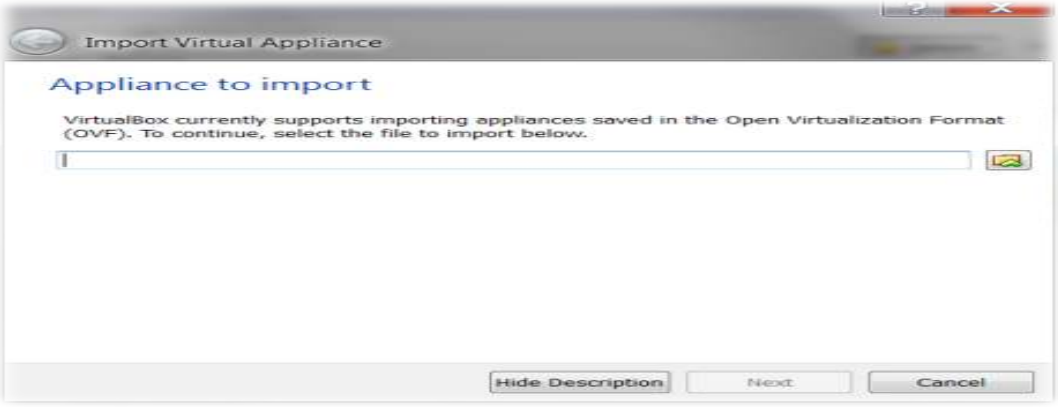
Windows:
Select **File->Import Appliance…** within the Oracle VM VirtualBox Manager



6. **Appliance to Import** dialog window displays.
Mac OS X: The drop-down screen displays

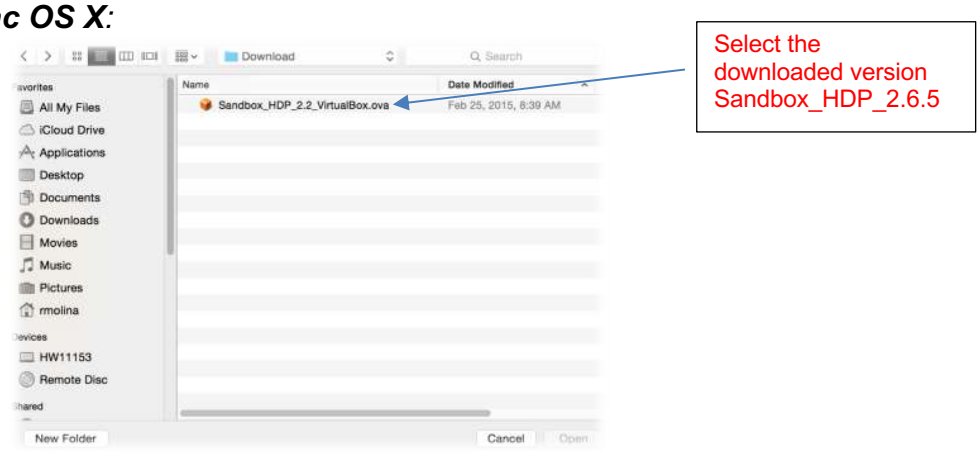


Windows: A separate dialog window is put in front of the VM VirtualBoxManager window:

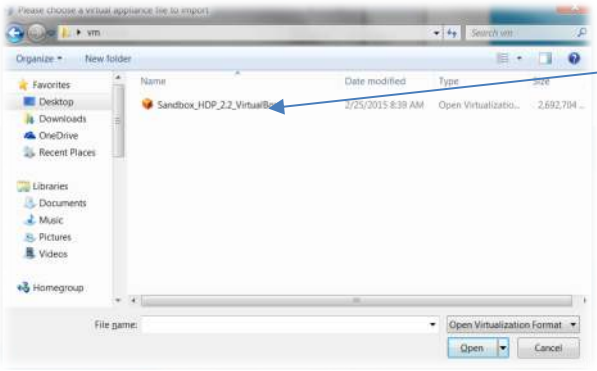


7. Click on the **folder** icon that will open a file dialog window to allow selecting the virtual appliance file that you downloaded as a prerequisite.After selecting the file click the **Open** button.

Mac OS X:



Windows:

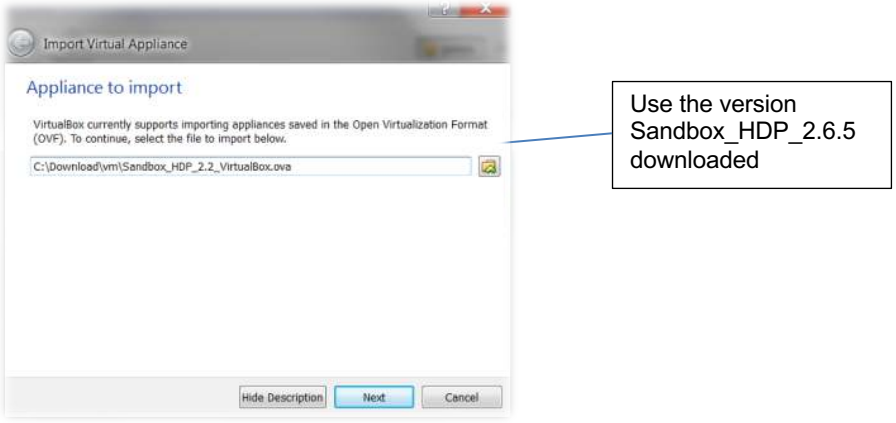


***NOTE**: The name of the file (Sandbox_HDP_2.6.5) you have downloaded depends on the version of the HortonworksSandbox you have chosen to download.*

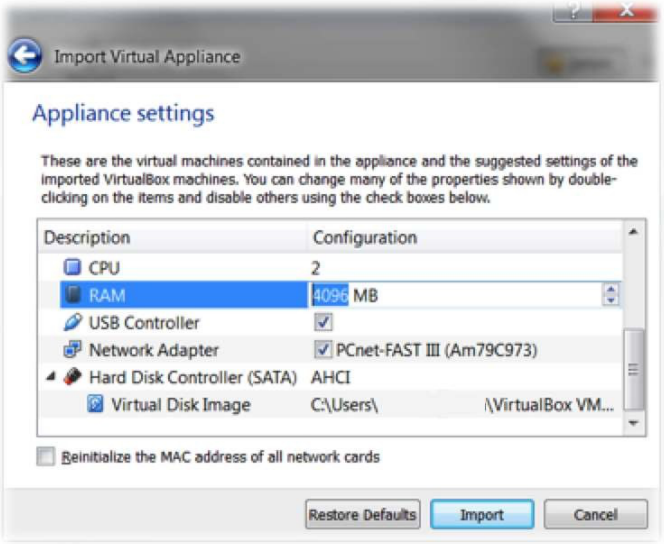
8. Application settings are now displayed.
   Mac OS X:



Windows:
On Windows after you select the virtual appliance file, you are brought backto this window.

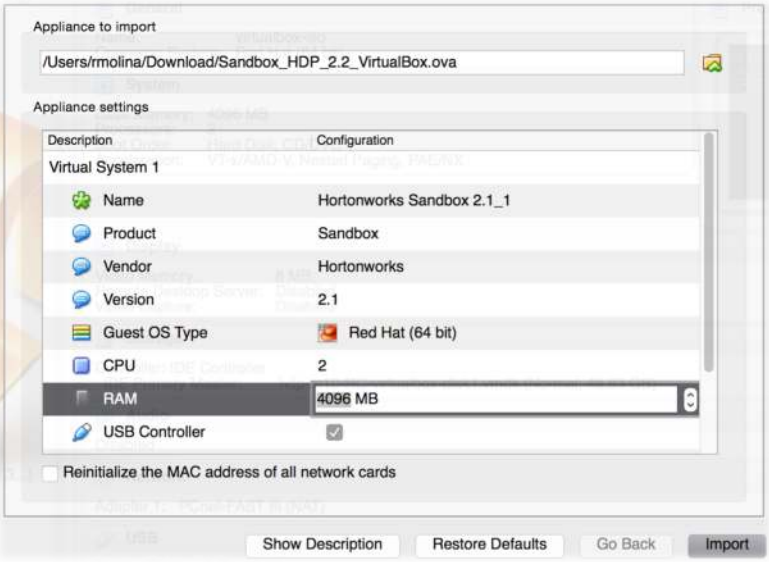


After clicking on **Next**, the Appliance Settings are displayed.

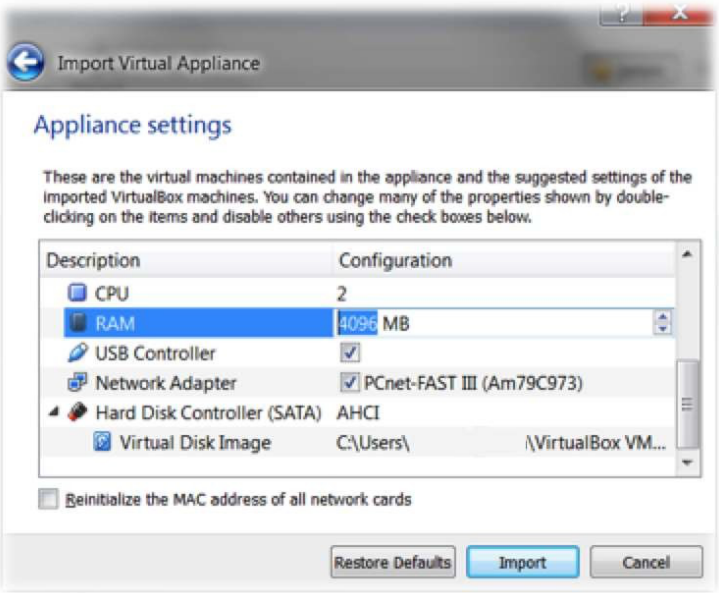


9. **Modify Appliance Settings as needed**.
Within the Appliance Settings section you may wish to allocate more RAM tothe virtual appliance. Setting 8GB of RAM to the Hortonworks Sandbox virtual appliance will improve the performance. Make sure you have enough physical RAM on the host machine to make this change. To make the change, click on the specific value to modify and make your edits. Once finished configuring, click **Import**.

Mac OS X:

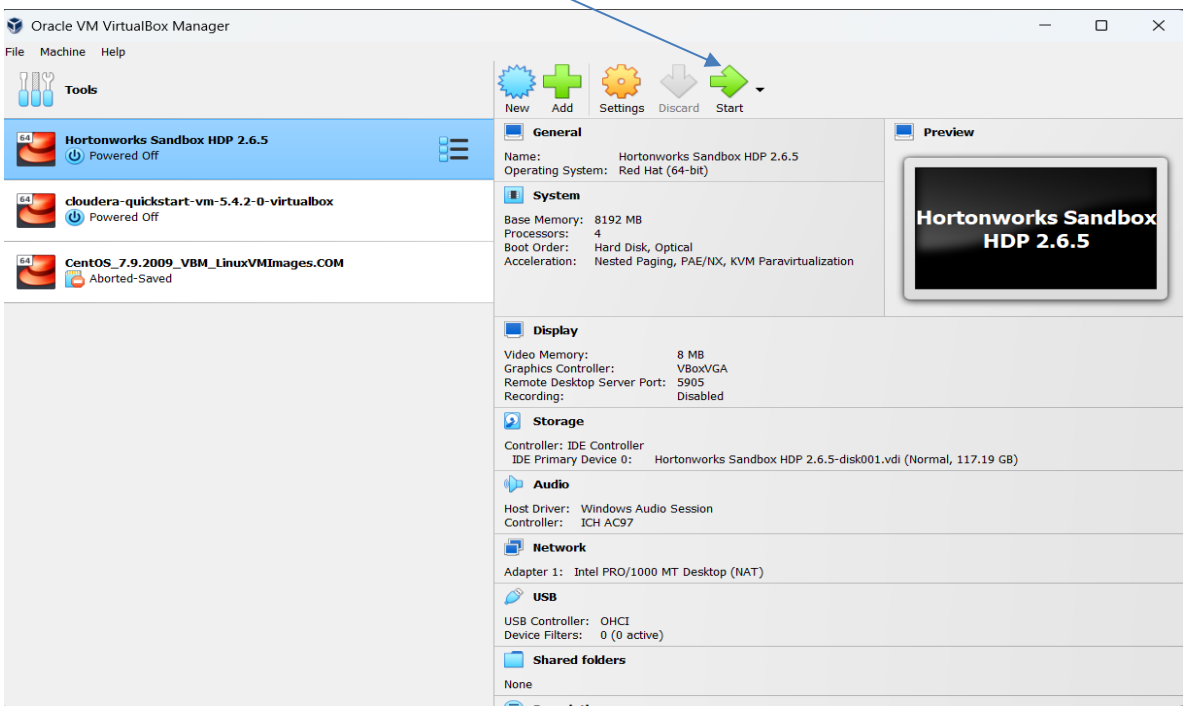


Windows:



10. Once the import finishes, you are brought to the main Oracle VM VirtualBox Manager screen. From the left-hand pane, select the appliance you just imported and click the green **Start** arrow.

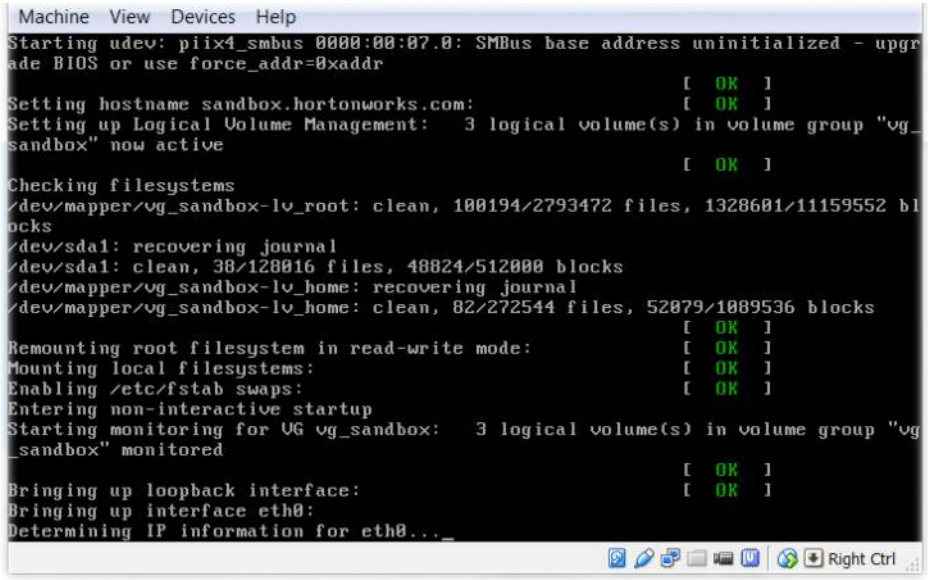
Mac OS X / Windows



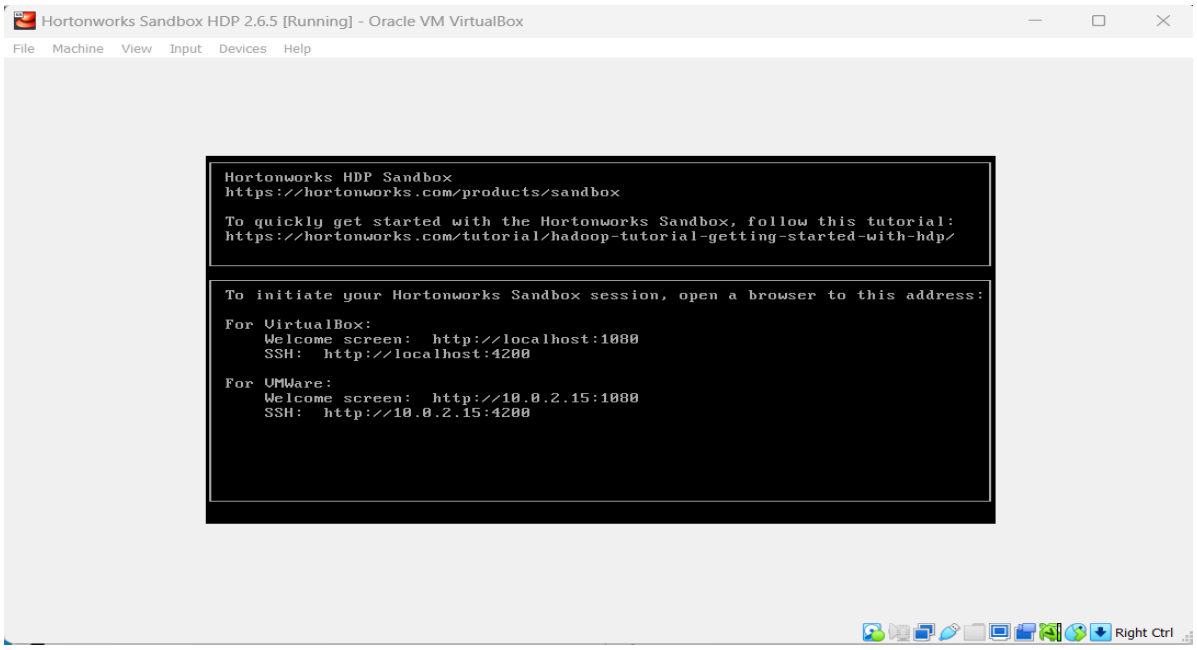
Running the VM

The first boot of HDP Sandbox takes a lot of time, please take a break and wait till it finishes.

11. A console window opens and displaying the boot up information.

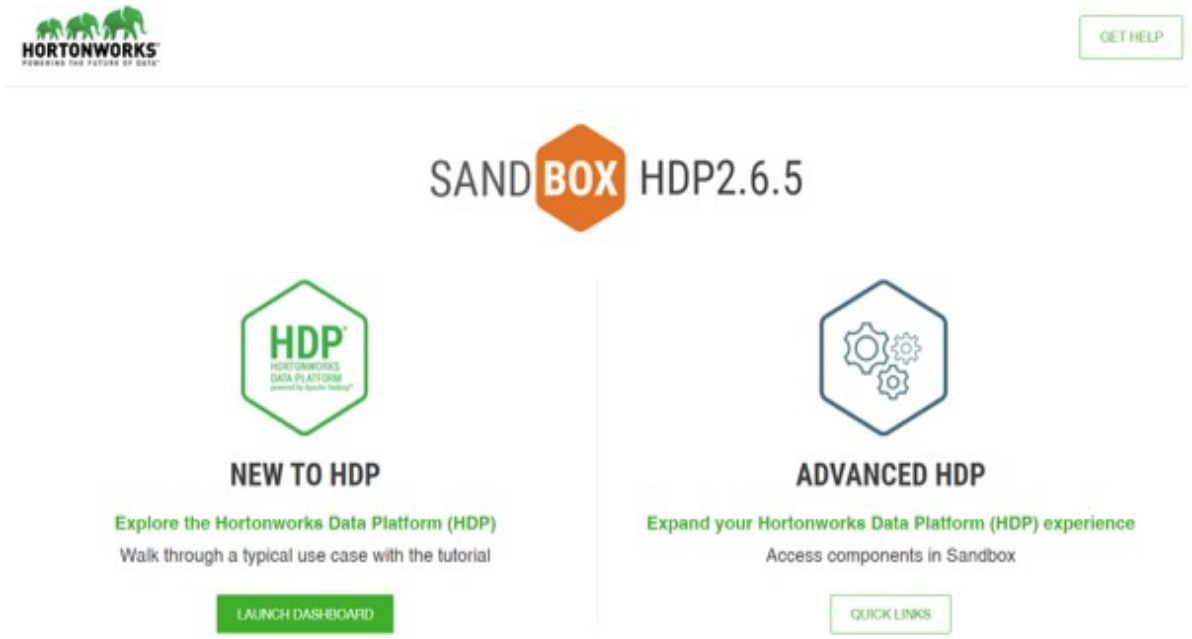


12. Once the virtual machine fully boots up, the console displays the logininstructions.



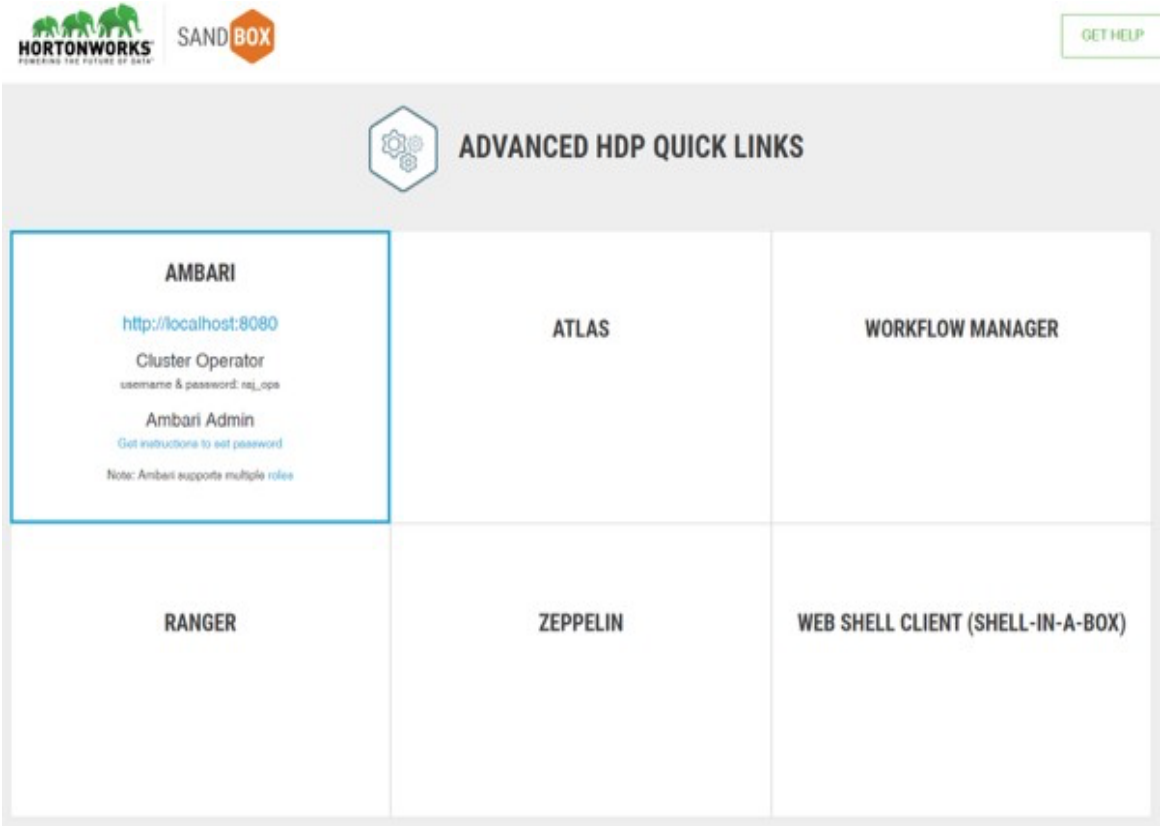
Access HDP Sandbox Cluster

After the cluster installation you would end up here. The installed HDP Sandbox cluster is a single node implementation of HDP. It is packaged as a virtual machine to make evaluation and experimentation withHDP fast and easy. You can access the webpage of the cluster via http://localhost:1080 for HDP 2.6.5



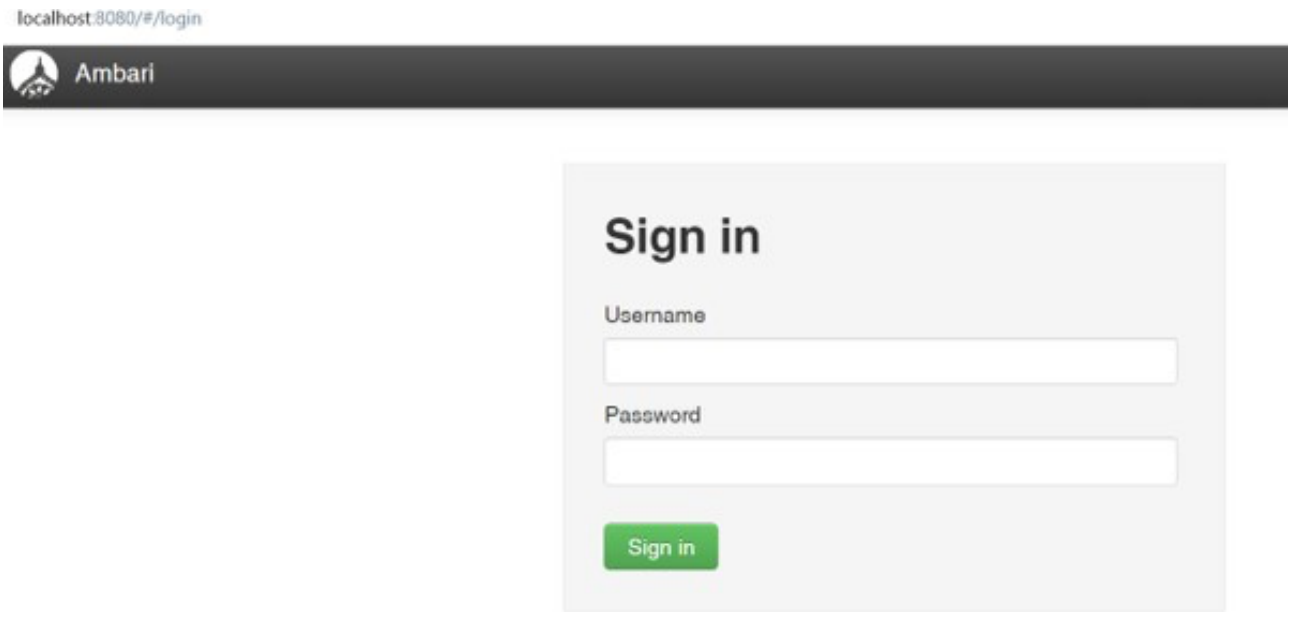
HDP Sandbox webpage

The button *Launch Dashboard* will transfer to portal.
The button *Quick Links* will transfer you to the page of links where you can access someservices of the cluster.



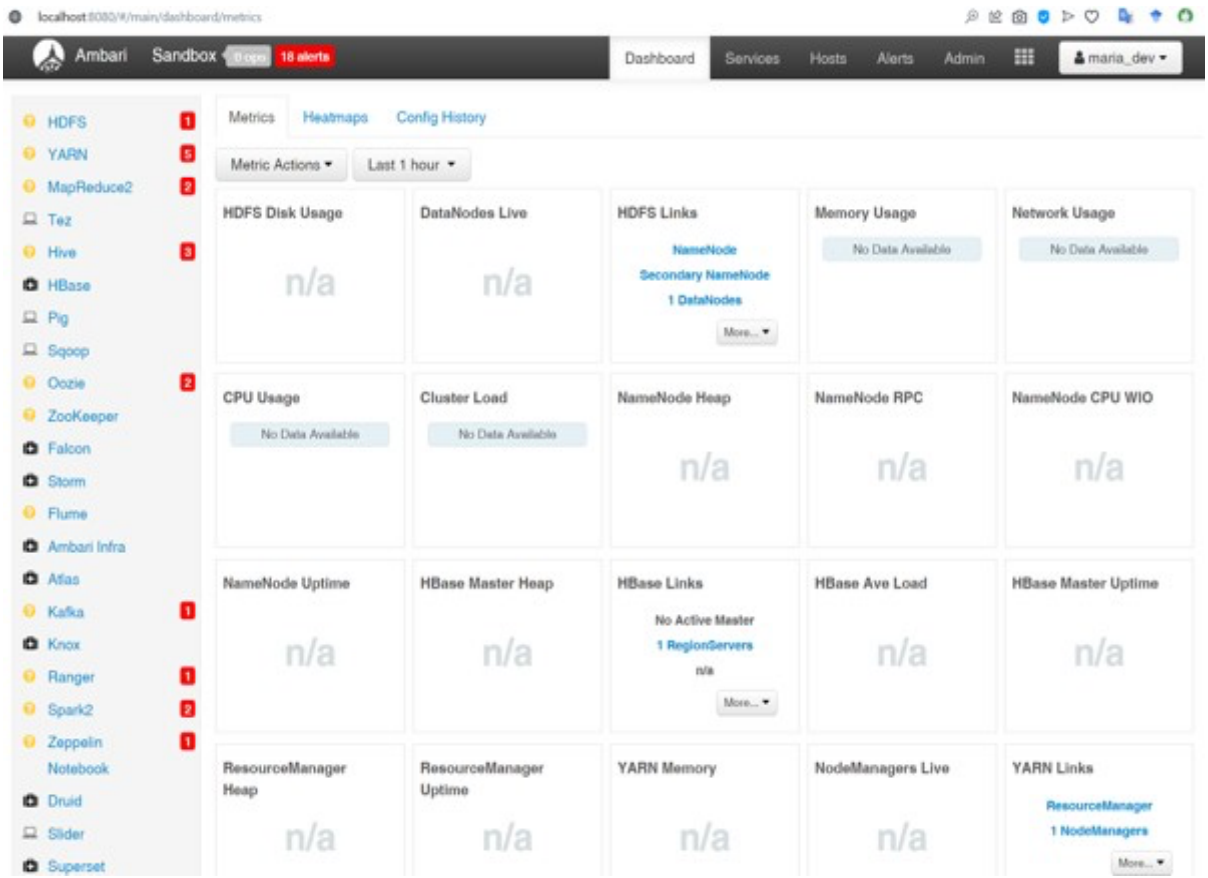
HDP Sandbox quick links webpage

In order to see all services of the cluster, you need to access *Ambari* service at http://localhost:8080 where you can monitor and manage all services.



Ambari login page

You need to log in in order to access this service. You can use the credentials of the user `maria_dev/maria_dev` as (username/password). HDP Sandbox comes with 4 default users with different roles in the cluster and there is also Ambari Admin who can manage the other users in the cluster.



Ambari homepage

As you can see in Ambari homepage, most services are showing alerts since they did not start so far or due to some problems. You need to wait till services start then you can access them. If you have set less resources than required then it is probably that most services cannot be run, so you can stop the services which are not needed to let the required services run.

Note: You can reset the password of Ambari Admin by running the command `ambari-admin-password-reset` via `ssh` as follows:

```
[root@sandbox-hdp  ~]#  ambari-admin-password-reset
```

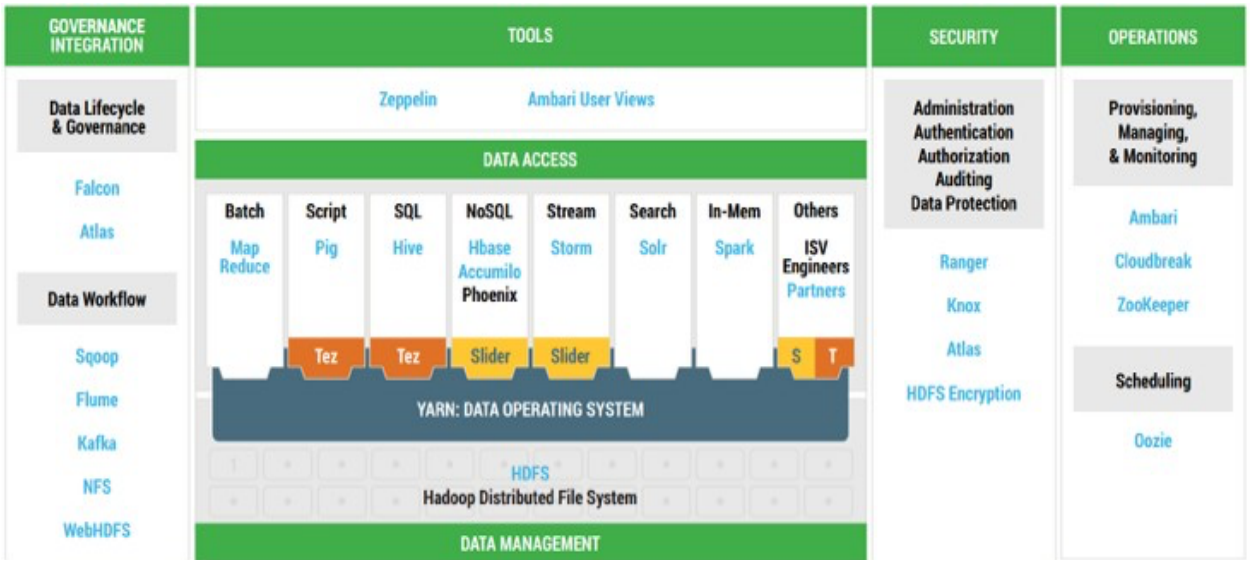
```
localhost:4200
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
Last login: Sat Jan 21 13:35:36 2023 from 172.18.0.2
[root@sandbox-hdp ~]# ambari-admin-password-reset
Please set the password for admin:
Please retype the password for admin:

The admin password has been set.
Restarting ambari-server to make the password change effective...

Using python  /usr/bin/python
Restarting ambari-server
Ambari Server is not running
Ambari Server running with administrator privileges.
Organizing resource files at /var/lib/ambari-server/resources...
Ambari database consistency check started...
Server PID at: /var/run/ambari-server/ambari-server.pid
Server out at: /var/log/ambari-server/ambari-server.out
Server log at: /var/log/ambari-server/ambari-server.log
Waiting for server start.....................................
Server started listening on 8080

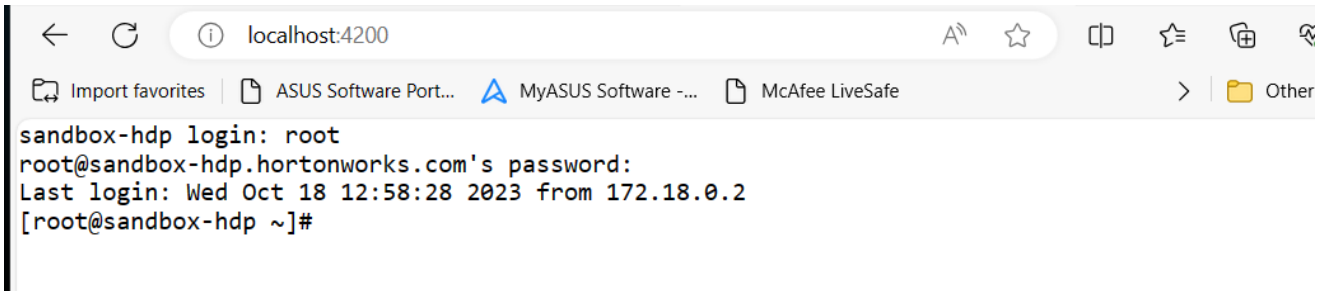
DB configs consistency check: no errors and warnings were found.
[root@sandbox-hdp ~]#
```

Resetting Ambari Admin credentials



You can access the cluster via shell web client by following the address http://localhost:4200 in your browser.

For the very first time, the default credentials are `root/hadoop` and you will be asked to reset the password. You need to set a strong password to pass the password-reset step.



Web Shell client for HDP Sandbox