



Lecture Exercise Booklet

BIG DATA HORTONWORKS SANDBOX EDITION

Contents

- Task 1 – Sign in to AWS..... 2
- Task 2 – Create a Hortonworks Sandbox Instance 2
- Task 3 – Interact with your Hortonworks Sandbox 5
- Task 4 – Updating Configuration12
- Task 5 – Interacting with HDFS13
- Task 6 – Creating Your Own.....16

Task 1 – Sign in to AWS

You will be provided with log in details to AWS for the duration of this course. Use these details to log in to AWS.

Task 2 – Create a Hortonworks Sandbox Instance

Create your Hortonworks sandbox instance.

Once you have signed up, you should see a dashboard of options. Click on 'Launch a virtual machine with EC2', then 'advanced EC2 launch wizard'.

Click on 'community AMIs then in the search bar look for 'Hortonworks'.

Step 1: Choose an Amazon Machine Image (AMI) Cancel and Exit

Continuing in the AWS Marketplace, or you can select one of your own AMIs.

Quick Start

My AMIs

AWS Marketplace

Community AMIs

Operating system

☐ Amazon Linux

☐ Cent OS

☐ Debian

☐ Fedora

☐ Gentoo

☐ OpenSUSE


☐ Other Linux

☐ Red Hat

Search: Hortonworks

6 results for "Hortonworks" on AWS Marketplace


Partner software pre-configured to run on AWS



Hortonworks HDPCAdministrator_2.3 PracticeExam_v3 - ami-11829421

Hortonworks HDPCAdministrator_2.3 PracticeExam_v3


Root device type: ebs Virtualization type: hvm



Hortonworks HDPCD Java PracticeExam_v1 - ami-415d5671

Hortonworks HDPCD Java PracticeExam_v1

Root device type: ebs Virtualization type: hvm




Hortonworks HDPCDeveloper_2.2 PracticeExam_v7 - ami-617b6700

Hortonworks HDPCDeveloper_2.2 PracticeExam_v7

Root device type: ebs Virtualization type: hvm

Scroll down until you find the AMI below (the version may change over time). Click Select.



hortonworks-sandbox-2.5 - ami-d278c1b2

[Copied ami-278a7931 from us-east-1] hortonworks-sandbox-2.5

Root device type: ebs Virtualization type: hvm

Select

64-bit

You should then be asked to choose an instance type. Select the t2.medium as highlighted below.

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Add Tags

6. Configure Security Group

7. Review

Step 2: Choose an Instance Type

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance	IPv6 Support
<input type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.micro Free tier eligible	1	1	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes
<input checked="" type="checkbox"/>	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes

Click through configuring the instance and adding storage, but do not change anything. When you move on to 'Add Tags', add a value of the word 'Academy', followed by your first name and last name as shown below.

Big Data Hortonworks Sandbox Booklet

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 5: Add Tags

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver. [Learn more](#) about tagging your Amazon EC2 resources.

Key (127 characters maximum)	Value (255 characters maximum)
Name	AcademyFName.Name

[Add another tag](#) (Up to 50 tags maximum)

Next, configure the security group. Select 'an existing security group', and select the 'Big-Data-Academy' one.

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☐ Create a new security group
☒ Select an existing security group

Security Group ID	Name	Description	Actions
<input type="checkbox"/> sg-11cc0069	AlexMulkerrinDocker_SG	security group for accessing server in docker containers	Copy to new
<input type="checkbox"/> sg-atb6f9c9	BarcelonaR10KDemo	BarcelonaR10KDemo created 2016-07-24T19:55:32.929+01:00	Copy to new
<input checked="" type="checkbox"/> sg-938f97eb	Big-Data-Academy	launch-wizard-14 created 2017-03-23T09:44:26.118+00:00	Copy to new
<input type="checkbox"/> sg-6c7467a1	Default	Default security group for Amazon EC2 instances	Copy to new

Inbound rules for sg-938f97eb (Selected security groups: sg-938f97eb)

Type	Protocol	Port Range	Source
All TCP	TCP	0 - 65535	0.0.0.0/0

If you see the following, then select the recommended fix.

Boot from General Purpose (SSD)

General Purpose (SSD) volumes provide the ability to burst to 3000 IOPS per volume, independent of volume size, to meet the performance needs of most applications and also deliver a consistent baseline of 3 IOPS/GiB.

- ☒ Make General Purpose (SSD) the default boot volume for all instance launches from the console going forward (recommended).
- ☐ Make General Purpose (SSD) the boot volume for this instance.
- ☐ Continue with Magnetic as the boot volume for this instance.

Free tier eligible customers can get up to 30GB of General Purpose (SSD) storage.

Next

You can now launch the instance. Click Launch, then Select existing Key Pair and use the BigDataKeyPair. Accept the terms and launch!

Select an existing key pair or create a new key pair



A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair ▼

Select a key pair




BigDataKeyPair ▼

☒ I acknowledge that I have access to the selected private key file (BigDataKeyPair.pem), and that without this file, I won't be able to log into my instance.

Cancel

Launch Instances

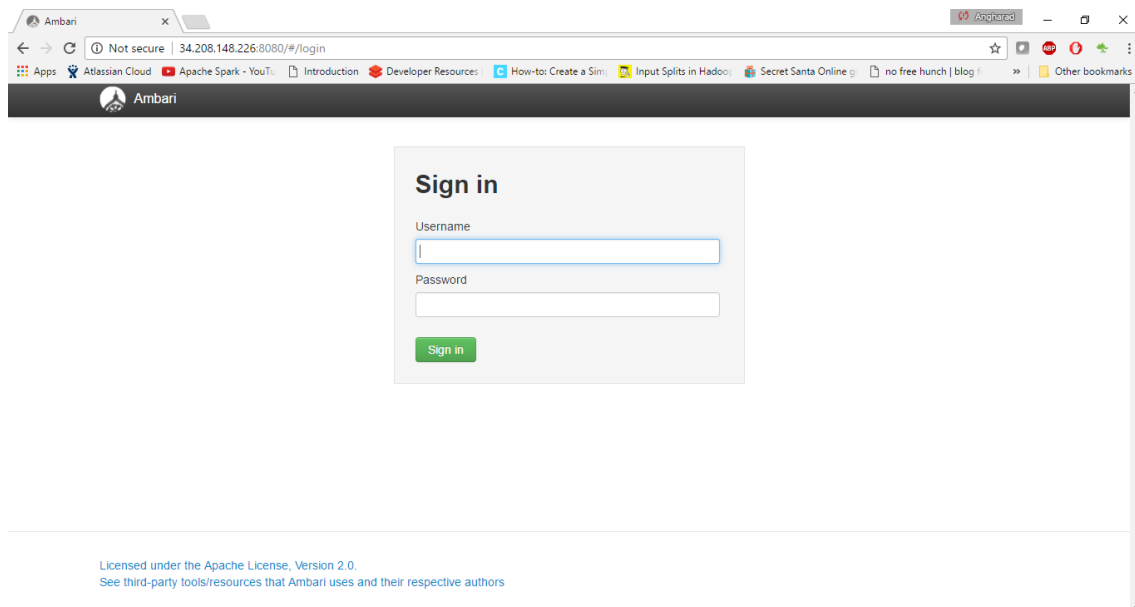
Go to services and click on EC2, then 'Running Instances' and find your instance. Scroll along to find the IP address of your instance.

 running	 Initializing	None	 ec2-34-208-148-226.us-...	<u>34.208.148.226</u>	-	BigDataKeyPair
---	--	------	---	-----------------------	---	----------------

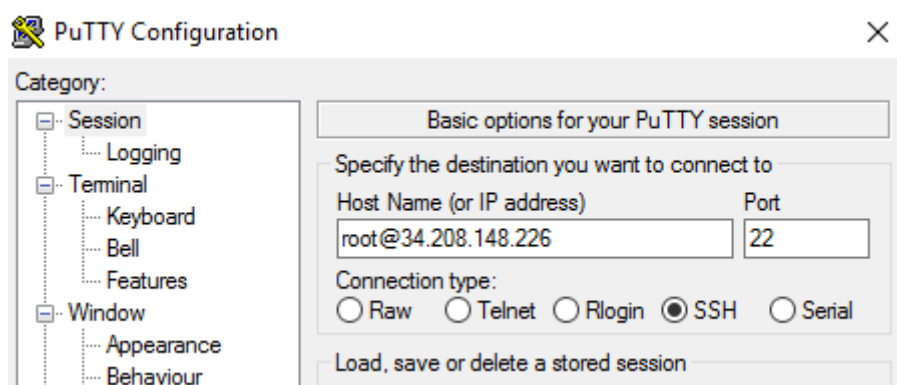
Once it has finished initializing you should be able to connect to it...

Task 3 – Interact with your Hortonworks Sandbox

Enter the IP address you found earlier followed by :8080 in your desktop's web browser to get to the following page.



You can also use putty to access the command line. Find and install putty from the LocalInstall folder on your machines and run it. Access your instance by giving putty the host of root:<your_ip_address>.



You will be prompted for a password, it is 'hadoop'.

The Hortonworks 2.5 sandbox operates as a number of containers running different processes. In order to change some of the Ambari settings you will need to ssh into the Ambari container, you will be prompted for another password, which is also 'hadoop', then change it to something you will remember.

```
[root@sandbox ~]# ssh root@127.0.0.1 -p 2222
root@127.0.0.1's password:
You are required to change your password immediately (root enforced)
Last login: Thu Mar 23 10:48:19 2017 from 172.17.0.1
Changing password for root.
(current) UNIX password:
New password:
Retype new password:
```

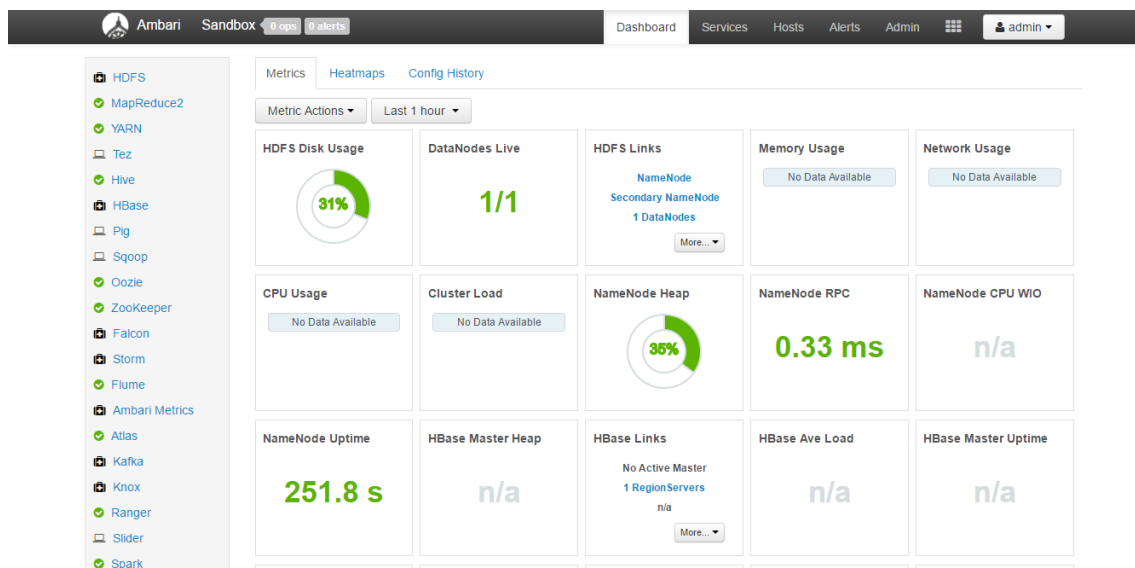
Reset your Ambari admin password by doing the following.

```
[root@sandbox ~]# 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
Using python /usr/bin/python
Stopping ambari-server
Ambari Server stopped
Using python /usr/bin/python
Starting ambari-server
Ambari Server running with administrator privileges.
Organizing resource files at /var/lib/ambari-server/resources...
Ambari database consistency check started...
No errors were found.
Ambari database consistency check finished
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.....
Ambari Server 'start' completed successfully.
```

Visit the Ambari home page at <ip_address>:8080 and give the username: admin and the password you just set up.



Checking Hadoop Service Accounts

1. In your putty window type the following to see the accounts listed in the `/etc/passwd` file. You should see Hadoop service account names such as `hive`, `zookeeper`, `ambari`, `hdfs`, etc, among many others.

```
# cat /etc/passwd
```

2. Now in the Ambari Web UI click on the 'Admin' button (not the admin account button), then select 'Service Accounts'. You should see the same accounts here listed from the command above.

Adding Ambari User Accounts

Use the Ambari Web UI to create some new user accounts for each of the permission levels.

1. Click 'Manage Ambari' from the 'admin' user account menu
2. Click Users to check out the user pane. Right now it should list a number of accounts that have been set up as part of the sandbox, including the admin that you are logged in as.
3. Click '+Create Local User'
4. Create a local user called `npuser`, with the password `npuser`, then click Save. We will not assign this account any permissions. Ensure the user is 'Active' and is not an admin.
5. Create two more accounts with the password as their name called `rouser` and `opuser` the same as in the last step. We will not assign permissions yet.
6. Finally, create an account called `aduser`, with the password `aduser`, and configure the Ambari Admin option to Yes.

Assigning User Permissions

1. Click roles
2. In the Cluster Operator section beneath 'Assign roles to these users' click Add User, then New, then type in `opuser`. Click the tick to save.
3. In the Cluster User section add the `rouser`.

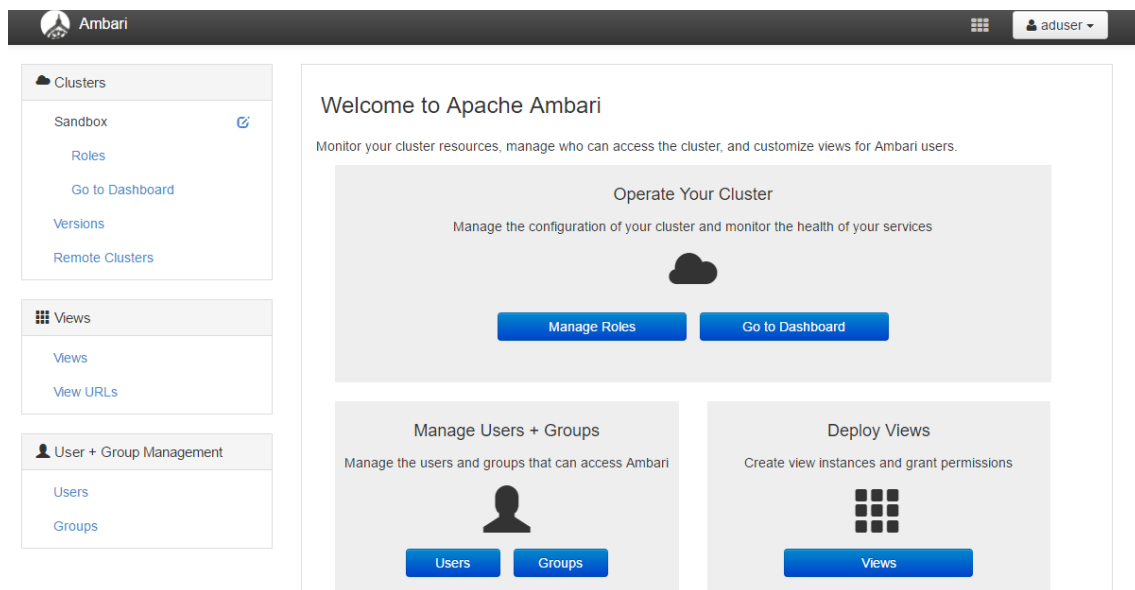
Testing the Users

You will now test the permissions of the users you have created.

1. Log out of the Ambari Web UI by selecting Sign out from the admin menu.
2. Sign in to the Ambari Web UI using the user name npuser and the password npuser. The npuser is known to Ambari but has no permissions in Ambari.
 - a. Were you able to log in?
 - b. What do you see in the browser window?
 - c. Do you see the options to navigate to the Dashboard, Services, Hosts, or Alerts pages?
 - d. Do you see the Ambari Views option?
 - e. Are there any Views assigned to the npuser?
3. Log out of the Ambari Web UI by selecting Sign out from the npuser menu.
4. Sign in to the Ambari Web UI using the user name rouser and the password rouser. The rouser account has been assigned read-only permission in Ambari.
 - a. Were you able to log in?
 - b. What do you see in the browser window?
 - c. Do you see the options to navigate to the Dashboard, Services, Hosts, or Alerts pages?
 - d. Do you see the Ambari Views option?
 - e. Are there any Views assigned to the rouser?
5. While logged in as rouser, click Dashboard
 - a. Are you able to view Dashboard information?
6. While logged in as rouser, select HDFS from the Services menu.
 - a. Are you able to view HDFS service information on the Summary page?
 - b. Do you see the Actions and Service Actions buttons on the Summary page? (Why not?)
7. While logged in as rouser, click Hosts to open the Hosts page.
 - a. Are you able to view host information on the Hosts page?
 - b. Do you see the Actions button on the Hosts page? (Why not?)
8. While logged in as rouser, click Alerts to open the Alerts page.
 - a. Are you able to view host information on the Alerts page?
 - b. Do you see the Actions button on the Alerts page? (Why not?)
9. Log out of the Ambari Web UI by selecting Sign out from the rouser menu.
10. Sign in to the Ambari Web UI using the user name opuser and the password opuser. The opuser has been assigned operator permission in Ambari.

- a. Were you able to log in?
 - b. What do you see in the browser window?
 - c. Do you see the options to navigate to the Dashboard, Services, Hosts, or Alerts pages?
 - d. Do you see the Ambari Views option?
 - e. Are there any Views assigned to the opuser?
 - f. Are you able to see the Admin page listed at the top of the browser window? (This should be the first time that this page is available as an option so far)
11. While logged in as opuser, click Dashboard
 - a. Are you able to view Dashboard information?
12. Select HDFS from the Services menu
 - a. Are you able to view HDFS service information on the Summary page?
 - b. Do you see the Actions and Service Actions buttons? (Why?)
13. While logged in as opuser, click Hosts to open the Hosts page.
 - a. Are you able to view host information on the Hosts page?
 - b. On the Hosts page, do you see the Actions button? (Why?)
14. While logged in as opuser, click Alerts to open the Alerts page.
 - a. Are you able to view alert information on the Alerts page?
 - b. On the Alerts page, do you see the Actions button? (Why?)
15. Open the opuser menu.
 - a. What are the available choices?
 - b. Specifically, notice that there is no Manage Ambari menu option.
16. Log out of the Ambari Web UI by selecting Sign out from the opuser menu.
17. Sign in to the Ambari Web UI using the user name aduser and the password aduser. The aduser has been assigned admin permission in Ambari.
 - a. Were you able to log in?
 - b. What do you see in the browser window?
 - c. Do you see the options to navigate to the Dashboard, Services, Hosts, or Alerts pages?
 - d. Do you see the Ambari Views option?
 - e. Are there any Views assigned to the aduser?
 - f. Are you able to see the Admin page listed at the top of the browser window?
18. While logged in as aduser, click Dashboard.
 - a. Are you able to view Dashboard information?
19. While logged in as aduser, select HDFS from the Services menu.

- a. Are you able to view HDFS service information on the Summary page?
 - b. On the Summary page, do you see the Actions and Service Actions buttons? (Why?)
20. While logged in as aduser, click Hosts to open the Hosts page.
- a. Are you able to view host information on the Hosts page?
 - b. On the Hosts page do you see the Actions button? (Why?)
21. While logged in as aduser, click Alerts to open the Alerts page.
- a. Are you able to view alert information on the Alerts page?
 - b. On the Alerts page, do you see the Actions button? (Why?)
22. Open the aduser menu.
- a. What are the available choices?
 - b. Notice the Manage Ambari menu option, which is only available to user accounts with admin permission.
23. Select Manage Ambari from the aduser menu. You should see the following



24. Remain logged in as the aduser.

Creating Groups

1. Click 'Groups' then '+Create Local Group'
2. Type 'opgroup' as a name then click Save
3. Click the opgroup to add group members
4. Click 'New' and type 'npuser'. Click the blue check box to confirm.

5. Repeat the previous step to add rouser, opuser and aduser to the opgroup. All four user accounts you created earlier should now be part of the opuser group.
6. Click Roles to add permissions to the opgroup
7. In the Cluster Operator section add the opgroup beneath 'Assign roles to these groups'. Remember to click the blue check box.

Testing Group and User Interaction

1. Remain logged in as aduser.
2. Click 'Go to Dashboard' to return to the main Ambari Web UI window
3. Open the aduser menu.
 - a. Has the aduser lost their admin permissions because of being added to the opgroup? Can you still see the Manage Ambari option?
4. Sign out of aduser and sign in to npuser. The npuser is known to Ambari but has no permissions. However, the npuser account was added as a member of the opgroup group, which has operator permissions.
 - a. What do you see in the browser window?
 - b. Do you see the option to navigate to the Dashboard, Services, Hosts, or Alerts pages?
 - c. Do you see the Ambari Views option?
 - d. Does the npuser account now appear to have some permissions in Ambari?
5. While logged in as npuser, select HDFS from the Services menu.
 - a. Are you able to view HDFS service information on the Summary page?
 - b. On the Summary page, do you see the Actions and Service Actions buttons? (Why?)
 - c. Would this indicate that the npuser account now has operator permission?
6. Sign out of the npuser account.

Task 4 – Updating Configuration

1. Open your putty terminal
2. Type in 'ls /etc/hadoop/conf' to see the Hadoop configuration files. Some of the more common are core-site.xml, hadoop-site.xml, and mapred-site.xml.
3. There will be a number of configuration files for other services not found here. Run the commands below to see them.
 - a. ls /etc/ambari-server/conf
 - b. ls /etc/ambari-agent/conf
 - c. ls /etc/hive/conf
 - d. ls /etc/pig/conf
 - e. ls /etc/zookeeper/conf
4. Now connect to your Ambari Web UI at <ip_address>:8080 and log in as an administrator user
5. Click 'Services'
6. Click 'Actions' under the list of installed services on the left. We will not stop and start all of the services right now as this takes a long time, but be aware this is where you can do so.
7. Select the HDFS service and check out the options on the Service Actions menu as we will be performing some as we go on
8. In the HDFS service click the Quick Links and select the NameNode UI. This shows the NameNode UI. This should open a new link to your NameNode UI. Close this and return to the Ambari HDFS service page.
9. Click on the Configs tab of the HDFS service
10. On the Settings page find and change the NameNode Java heap size setting. Drag the slider bar to 2GB but DO NOT save the change.
11. Now click the grey circular arrow to change this value to the one recommended by the Ambari guided configuration.
12. Now click in to the Advanced tab and scroll down to find the General section.
13. Find the parameter called Block replication and change it to 2 then click Save. You can optionally add notes to this configuration change on the next step.

14. If Ambari makes some recommendations, accept them and then click 'Proceed Anyway'.
15. Return to the HDFS service page
16. You should see a yellow-orange bar at the top of the page alerting you of the need to restart services. Click the drop down Restart menu and select Restart All Affected. A Background Operation Running window will then open where you can observe progress.
17. Ambari lets you know of any other services that also need restarting by having a yellow symbol next to them, so, if need be, select each one that has the restart icon and follow the same procedure.
18. Reselect the HDFS service and click Configs. You should now have a configuration history! You can hover over the configurations and click Compare to see the changes.
19. You can change back to the previous configuration by hovering over the configuration and selecting 'Make Current'
20. Change back to the previous configuration and restart all services as necessary

Task 5 – Interacting with HDFS

1. Open your putty instance
2. Run the following commands to see information on your instance
3. Display the help information on the HDFS shell commands

```
# hdfs dfs -help
```
4. Use the HDFS shell to list local files and directories

```
# hdfs dfs -ls file:///usr
```
5. Use the HDFS shell to list files and directories in HDFS

```
# hdfs dfs -ls hdfs:///user
# hdfs dfs -ls /user
```
6. List the files and directories in the root HDFS directory

```
# hdfs dfs -ls /
```
7. List your user HDFS home directory by not supplying a directory path argument

```
# hdfs dfs -ls
```

8. You'll notice you can't currently do this as your home directory doesn't exist in HDFS yet. Let's create one.

```
# hdfs dfs -mkdir /user/root
```

9. If you rerun the argument from step 7 the command will now print nothing, this is not an error, it is because there is nothing in your directory yet.

10. Before we do anything with our new directory, let's check the ownership of it.

```
# hdfs dfs -ls /user
```

11. You should see a new directory called /user/root that wasn't there before. All of these directories you see are owned by the 'hdfs' group, as you should see this listed to the left of the directory name and date. However the owner is a unique user for each directory.

12. Let's change it so both the owner and group are from your account.

```
# hdfs dfs -chown root:root /user/root
```

13. Check the ownerships again and notice it has changed from 'hdfs' to 'root'.

14. Create a subdirectory hierarchy within your HDFS home directory then list the contents to verify it was created.

```
# hdfs dfs -mkdir -p dir1/dir2/dir3
# hdfs dfs -ls -R
```

15. Create a .txt file in your local directory on the virtual machine and add some text to it. When using vi, first type 'i' for insert, that will allow you to start typing. When you are finished, type escape to stop typing. Then press the colon ':' which will give you a new prompt. Enter 'wq' to save and quit.

```
# vi test.txt
```

16. Copy this file to your user directory in HDFS then list the directory to ensure it worked.

```
# hdfs dfs -put test.txt test.txt
# hdfs dfs -ls
```

17. Display the contents of the test.txt file in the HDFS directory.

```
# hdfs dfs -cat test.txt
```

18. Move the test.txt file into the directory structure you created earlier then check the move was successful.

```
# hdfs dfs -mv test.txt dir1/dir2/dir3
# hdfs dfs -ls - should show just dir1
# hdfs dfs -ls dir1/dir2/dir3 - should show test.txt
```

19. Now let's view HDFS in the NameNode web UI. Go to the web UI page at <your_url>:50070 (or alternatively, use the quick links we saw earlier in Ambari)

20. On the main page, click Utilities then Browse the file system
21. Click on the user directory name following by root, you should then be able to click through the file structure you made earlier and see information on the test.txt file you added. The only way to view file contents is to download it.
22. View the current HDFS block size. It is defined by the dfs.blocksize property in the /etc/hadoop/conf/hdfs-site.xml file.

```
# cat /etc/hadoop/conf/hdfs-site.xml
```


It should be 134217728 bytes, 128 MB, the same as what you should see on the file you just viewing in the NameNode web UI.

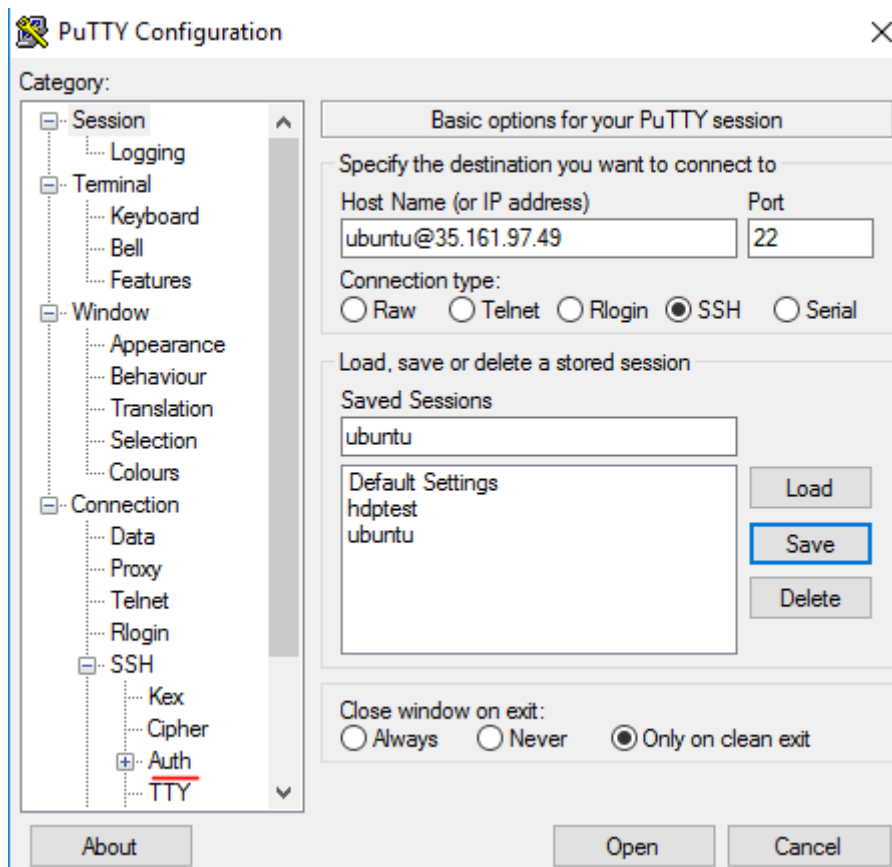
You can now destroy this Hortonworks Sandbox instance. We have explored some initial concepts within HDP and will now move on to creating our own instance.

Go and view your instance on AWS in the list of running instances. Click on it so it is highlighted (and only your instance is highlighted). Click on Actions -> Instance State -> Terminate.

Task 6 – Creating Your Own

We will now install our own setup using Apache Ambari and Hortonworks. First you will need to create a new virtual machine in AWS. We will only be creating a single node cluster for this demonstration.

1. Go to create a new EC2 instance in AWS, start in the 'Quick Start' menu and find the Ubuntu 14.04 server instance.
2. Go through similar steps to create a virtual machine as you did earlier.
3. In order to access this virtual machine through putty you will need to perform an additional step. Before signing in, on the putty interface navigate to SSH -> Auth. In the section for 'Private key for authorisation', browse to the .ppk file you have been provided.
4. Use putty to access your virtual machine once you have the IP address and log in with the user name 'ubuntu'.



5. Change the root password to something you will remember

```
# sudo passwd root
```
6. Get the repository file for the Ambari install (no newline in the web address)

```
# cd /etc/apt/sources.list.d  
# sudo wget http://public-repo-1.hortonworks.com/ambari/ubuntu14/2.x/updates/2.2.0.0/ambari.list
```

7. Update apt-get

```
# sudo apt-get update
```

8. Use the apt-get command to install Ambari Server. This will also download and install Java. This may take a few minutes.

```
# sudo apt-get install ambari-server
```

9. Now you must initialise the Ambari Server, the -s option is the silent option which will install Ambari with default settings.

```
# sudo ambari-server setup -s
```

10. Once it has been initialised you should start the Ambari Server. You should see a 'successful' message.

```
# sudo ambari-server start
```

11. Now you should be able to connect to the Ambari Web UI at <host_ip>:8080

12. Log in using the username admin and password admin

13. Because the Ambari Server cannot find an installed cluster, it will present the option to install a cluster. Click [Launch Install Wizard](#)

14. Pick a cluster name that is easy to remember and click Next

15. Ensure that only the HDP 2.3 radio button is selected and click Advanced Repository Options

16. Deselect all options aside from ubuntu14

17. Type 'hostname' in your putty connection to get your machine's name.

18. Type your machine's name as the target host. You are currently only installing a single cluster node, but you can add multiple target hosts, one per line, or add more hosts later.

19. These next steps will create an ssh file for your to provide the private SSH key.

20. On your virtual machine, create a new key. Do not give it a name, simply press enter to each question.

```
# cd  
# ssh-keygen
```

21. This will create a hidden folder called .ssh. cd into this folder.

```
# cd .ssh
```

22. Copy and paste the contents from id_rsa into the Ambari console. Delete the newline at the bottom if it adds one.

23. Perform the following command in your machine

```
# cat id_rsa.pub > authorized_keys
```

24. Ensure the user is filled in as ubuntu.

25. Click 'Register and Confirm'

Install Options

Enter the list of hosts to be included in the cluster and provide your SSH key.

Target Hosts

Enter a list of hosts using the Fully Qualified Domain Name (FQDN), one per line. Or use [Pattern Expressions](#)

```
ip-172-31-6-246.us-west-2.compute.internal
```

Host Registration Information

☒ Provide your [SSH Private Key](#) to automatically register hosts

Choose file

No file chosen

```
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAAKCAQEAu3P1Ue311f30v+XCW9LHTQExndvSJiSfcTNovV7nFLK
N7n0e
```

SSH User Account

ubuntu

☐ Perform [manual registration](#) on hosts and do not use SSH

[← Back](#)

[Register and Confirm →](#)

26. Watch the installation and click through to confirm. You should see something saying 'Success' if it worked. It may take some time to check the hosts.

27. For now, we will install just a few of the available services, so select HDFS, YARN + MapReduce2, Zookeeper, and Ambari Metrics (and untick anything else), then click Next.

28. In Assign Masters you would choose which cluster nodes would run which master service components. Normally you would balance out the workloads so no single machine is overworked, but since there is only a single node they will all run from there for now. Similarly, the Assign Slaves and Clients window will assign which nodes will store and process data.

29. When you move on to Customize Services, you're able to make configuration changes before you deploy, or fix any alerts if necessary. For now, you should be able to leave this as it is.

30. Click Next then review and deploy (if you get some configuration warnings at this point, you should be fine to skip them).

31. Monitor the installation and read the summary window once complete. The Ambari Web UI will be displayed once complete. If you run into any warnings, check these, but you may be fine to continue on.
32. Explore the Ambari Web UI for your installation. How does it differ to the sandbox you worked with earlier? You'll notice that you have a couple of alerts – that's because when we set up this instance we gave the machine just a small amount of storage, and we've basically used all that up on the installation! Don't worry about it for now, we'll increase that in the next exercise book...
33. Run the following command in your virtual machine console to see the top level structure of HDFS that has been created by Ambari.

```
# hadoop fs -ls /
```
34. Destroy the instance of this set up.