

**WA2914 Introduction PySpark**

**Lab Server VM Setup Guide**

**Web Age Solutions**

## Table of Contents

|   |   |
|---|---|
| Part 1 - Minimum Hardware Requirements to host the VMs.....             | 3 |
| Part 2 - Software Provided.....   | 3 |
| Part 3 - Installing the VMWare image.....                               | 3 |
| Part 4 - Running the VM.....  | 4 |
| Part 5 - Verifying Python3 and Checking Gedit.....                      | 4 |
| Part 6 - Setting Up the IP Address of the Lab Server VM (optional)..... | 5 |
| Part 7 - Summary.....   | 9 |

## Part 1 - Minimum Hardware Requirements to host the VMs

- 64-bit x86 CPU that supports hardware virtualization (Intel-VT or AMD-V).
- Hardware virtualization enabled in the BIOS.
- 8 GB RAM, 16 GB recommended
- 80 GB Hard Disk
- VMWare player or workstation 14

## Part 2 - Software Provided

You will receive the following file (further referred to as the **VM ZIP** file) containing the VMware player-compatible virtual machine:

- **VM\_WA2914\_REL\_1\_1.zip**

## Part 3 - Installing the VMWare image

If have VMware player 15 already installed then skip this part.

Download the VMware player **15** or higher from the VMware website using the following link:

<http://www.vmware.com>

- \_\_\_ 1. Install the VMware player accepting all the defaults during the installation.
- \_\_\_ 2. Restart the computer.
- \_\_\_ 3. From the Start menu, select **All Programs > VMware > VMware Player**.
- \_\_\_ 4. If prompted to download a new version of VMware player decline the update.

## Part 4 - Running the VM

\_\_1. Extract the **VM\_WA2914\_REL\_1\_1.zip** file into your computer.

**Note:** Every student in the class will need a dedicated Lab Server. So the class setup will require as many Lab Servers as there are students in the class. In other words, you will need to perform this setup as many times.

It is recommended to have each Lab Server VM installed on a separate physical machine, although they can be collocated as long as their Network Connectivity is setup with the Bridged option (see details further in the document).

\_\_2. Open the folder where you extracted the VM and a short cut in the desktop for **cloudera-quickstart-vm-5.13.0-0-vmware.vmx** and rename it as **VM\_WA2914\_REL\_1\_1**

\_\_3. Double click on the shortcut to open it.

\_\_4. Click "**I moved it**", if prompted.

\_\_5. If you are prompted for credentials, enter cloudera for user/password.

\_\_6. If you are promoted to download and install the VMware Tools for Linux, reject the option.

VM bootstrapping may take some time, and when it completes, you should be automatically logged in the Lab Server VM as the cloudera user and presented with the Cloudera Desktop.

## Part 5 - Verifying Python3 and Checking Gedit

The version of python installed here will be used for labs 1-8. Gedit should be installed by default. Follow instructions below to verify that it is working.

\_\_1. In the VM, open a terminal window. You should find an icon in the desktop.

\_\_2. Test that python3 have been installed by executing the command below:

```
python3 --version
```

You should see that the version is 3.4.10

\_\_3. Test that pip3 have been installed by executing the command below:

```
pip3 --version
```

You should see that the version is 19.2.3

\_\_4. To confirm that gedit editor is available, from the terminal window execute the following:

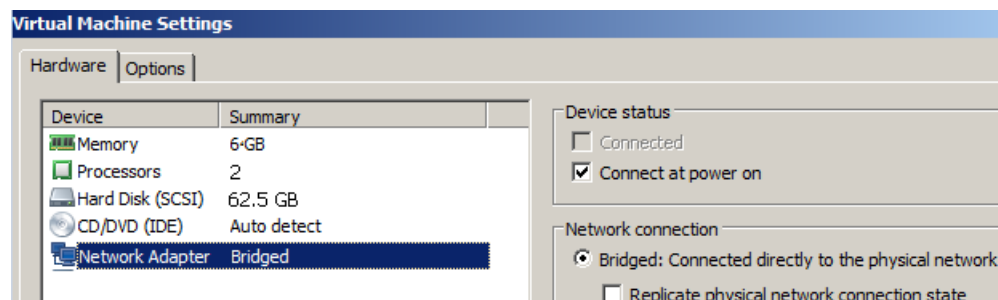
```
gedit
```

The gedit editor window should open.

\_\_5. Close gedit and the Terminal.

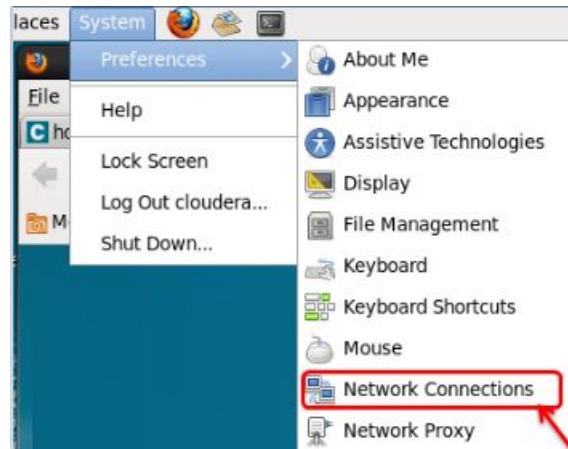
## Part 6 - Setting Up the IP Address of the Lab Server VM (optional)

This step applies if your organization requires to setup individual IPs. If you setup your VM Network Adapter with the *Bridged* option as shown in the screen-shoot below, by default, you will have a DHCP leased IP address assigned to the Lab Server. It may be a convenient feature from the administration perspective, but will affect student SSH connections during the class as they will always be required to change the Lab Server IP address whenever the IP address of the Lab Server changes (IP lease may be configured to expire every day and the class runs for four days).



Considering the inconvenience to the students, it may be worthwhile to assign each Lab Server a unique IP address.

\_\_1. From the Lab Server toolbar, select *System > Preferences > Network Connections*.



The *Network Connections* Dialog opens.

\_\_2. Select **Wired / Auto eth1** and click **Edit ...**



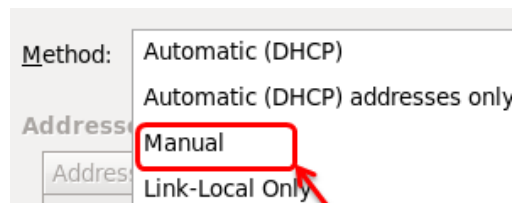
The **Editing Auto eth1** Dialog opens.

\_\_3. Select the **IPv4 Settings** Tab.



In the screen-shoot above the network adapter is configured to receive IP address from the DHCP server.

\_\_4. For setting up the static IP address, select **Manual** from the *Method*: drop-down.



\_\_5. Click **Add** to add an *IP Address*, *Netmask*, *Gateway* and *DNS* as per your network settings.

Sample input screen is shown below.

Editing Auto eth1

Connection name: Auto eth1

☒ Connect automatically

☒ Available to all users

Wired 802.1x Security IPv4 Settings IPv6 Settings

Method: Manual

**Addresses**

| Address       | Netmask | Gateway     |
|---------------|---------|-------------|
| 192.168.0.101 | 24      | 192.168.0.1 |

Add Delete

DNS servers: 192.168.0.1

Search domains:

DHCP client ID:

☒ Require IPv4 addressing for this connection to complete

Routes...

Cancel Apply...

\_\_6. Click **Apply ...**

The *Authenticate* Dialog opens up.

\_\_7. In the *Password for root:* text window, enter **cloudera** and click **Authenticate**.

You should be returned to the *Network Connections* Dialog.

\_\_8. Click **Close**.



## **Part 7 - Summary**

**You have successfully installed the software for this course!**

If you have any question please contact us by email at [support@webagesolutions.com](mailto:support@webagesolutions.com)

From US and Canada call: 1-877-812-8887 ext. 26

International call: 416-406-3994 ext. 26