Setup Guide for

HDP Developer: Apache Pig and Hive

Overview

The Hortonworks Training Course that you are attending is taught using a virtual machine (VM) for the lab environment. Before attending the course, you must complete the steps in this document.

When Cutting and Pasting Commands and URLs...

This guide was originally generated using a tool that tries to be "helpful" by changing certain characters to improve formatting, including dashes, double quotation marks, and single quotation marks. Unfortunately, these auto-formatting tools can cause commands and URLs not to work when copied and pasted directly into a Linux command line or browsers. We have attempted to locate and correct each of these issues, but it is possible there are items we have missed. If you are copying and pasting a command or URL that results in an error, it is likely that one of the characters you entered has been interpreted as a special character rather than the intended dash or quote symbol, and is thus incorrect. To fix this problem, when copying and pasting, replace all dashes and quote symbols with manually-typed equivalents before executing the command.

We apologize for any inconvenience.

Steps

Here is an overview of the steps you will perform:

- Download the virtual machine file. The VM file is in a folder online that we will share with you. The file is very large (~ 5GB) and may take several hours to download, depending on your Internet connection.

- You must install **VMWare Player** to open the VM file on Windows. On Mac, you can download a free trial of **VMWare Fusion** at http://www.vmware.com/products/fusion/.

- You must start the virtual machine after it is imported into VMWare to verify that it can run on your machine's hardware and software.

- You must also run a MapReduce job to verify that your system has enough memory to perform the labs in the course.

- It is possible that you will need to configure the BIOS on your Windows machine to enable virtualization.

- **THE FOLLOWING SETUP INSTRUCTIONS MUST BE COMPLETED AT LEAST 24 HOURS PRIOR TO THE START OF THE CLASS!**

> **IMPORTANT**: It is critical for the success of the training that prior to the start of the class each student's computer has the VM file downloaded, imported into VMWare, and verified that the virtual machine starts on the student's computer and a MapReduce can execute successfully.

IMPORTANT: If any of the following steps fail, or if you have any issues or questions, please send an email to training-support@hortonworks.com.

System Requirements

The following is the recommended minimal system requirements:

- Reasonably powerful **x86/amd64 hardware.** Recent Intel (**3rd generation i5/i7 or above preferred**) or equivalent AMD processor should do.

- RAM: At least **8 GB**

- Available Hard-disk: at least **80 GB**

- **64-bit** OS (Windows 7, Windows 8 and Mac OSX)

- VMWare Player for Windows (or VMWare Fusion for Mac)

- **Virtualization must be enabled on BIOS** for Windows based machine
    - Procedures to turn on virtualization settings in your computer's BIOS vary depending on the BIOS manufacturer. Check the information that came with your computer or go to the computer manufacturer's website.
    - Alternatively, to find out whether your computer's CPU is capable of hardware-assisted virtualization, you can download and run the Hardware-Assisted Virtualization Detection Tool (http://www.microsoft.com/en-us/download/details.aspx?id=592) provided by Microsoft.

Install the Classroom Virtual Machine

Objective:	Import the VMWare image into VMWare Player and start the virtual machine successfully.
Before You Begin:	Make sure your hardware meets the necessary requirements.

Perform the following steps at least one day prior to the start of the course:

Step 1: Install VMWare Player

1.1. On Windows, download the latest 64-bit version of VMWare Player:

http://www.vmware.com/support/download-player.html

On a Mac, download a free trial of VMWare Fusion at http://www.vmware.com/products/fusion/.

1.2. Run the downloaded installation file.

Step 2: Download the VM

2.1. Using one of the following links, download the VM file. This file is over 5GB in size and may take a couple of hours to download, depending on your internet connection:

US East - http://tinyurl.com/us-east-da-vm

US West - http://tinyurl.com/us-west-da-vm

Europe - http://tinyurl.com/europe-da-vm

Asia - http://tinyurl.com/asia-da-vm

2.2. Unzip the file you just downloaded. Please make sure to use 7zip(free to download) utility for Windows platform. The default Windows Unzip utility does not work.

Step 3: Open the VM File

3.1. Start the VMWare Player application and select the option to **Open a Virtual Machine**:
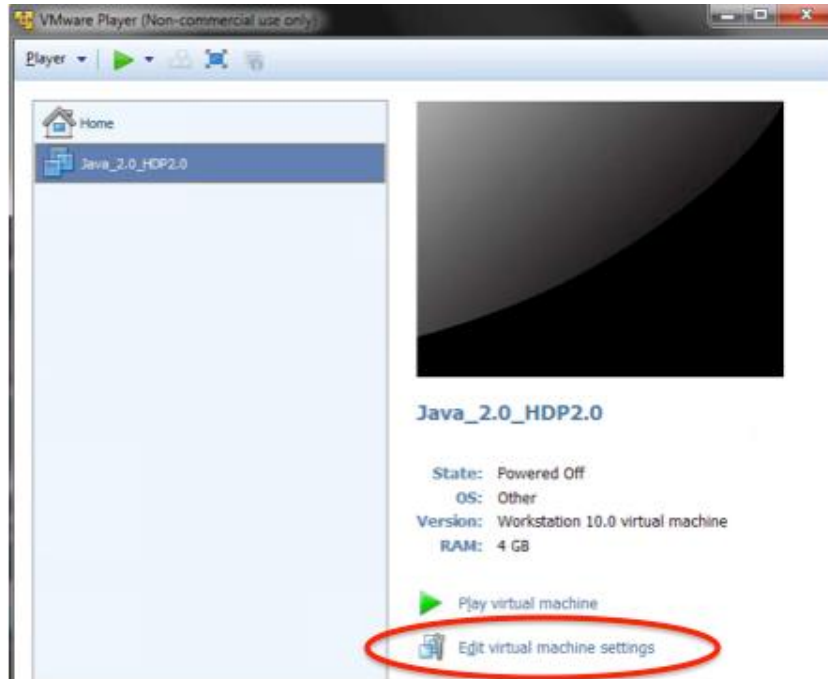

3.2. Browse to and open the unzipped VM file

3.3. It will take several minutes for the file to be imported into VMWare Player/Fusion:

3.4. When the import is complete, the VM will show up in the VMWare Player/Fusion library.

Step 4: Modify the VM's Memory

**4.1.** Click the **Edit virtual machine settings** link:

4.2. Set the **Memory** to a size reasonable for your hardware. For example, if your machine has 8GB of RAM, then set the base memory to 6144MB. If your machine has 16+GB of RAM, then you can increase the base memory to 8GB if you prefer.
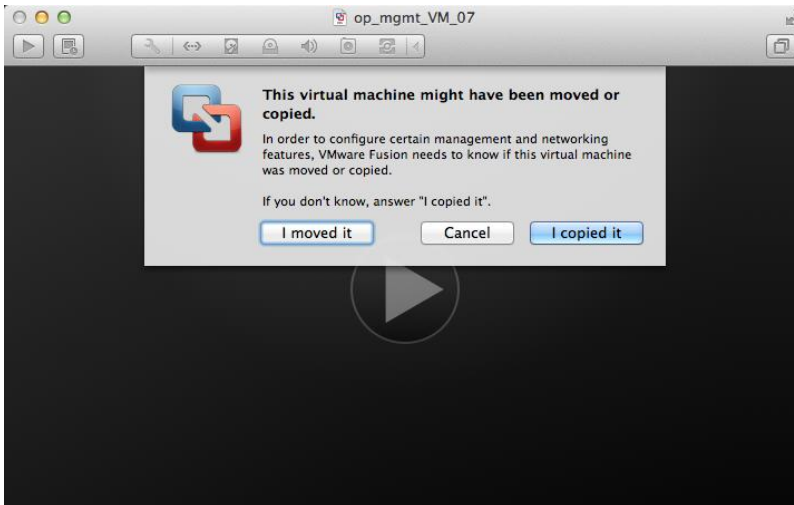
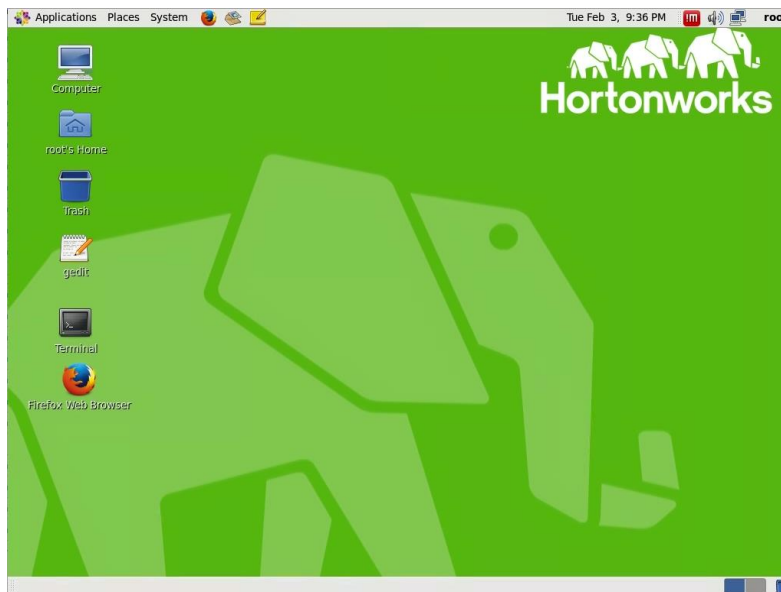**4.3.** Click **OK** to save any changes.

Step 5: Start the VM

5.1. Click the **Play virtual machine** link to start the VM.

5.2. If following windows pops up then click on "Don't Upgrade" & "I copied it"
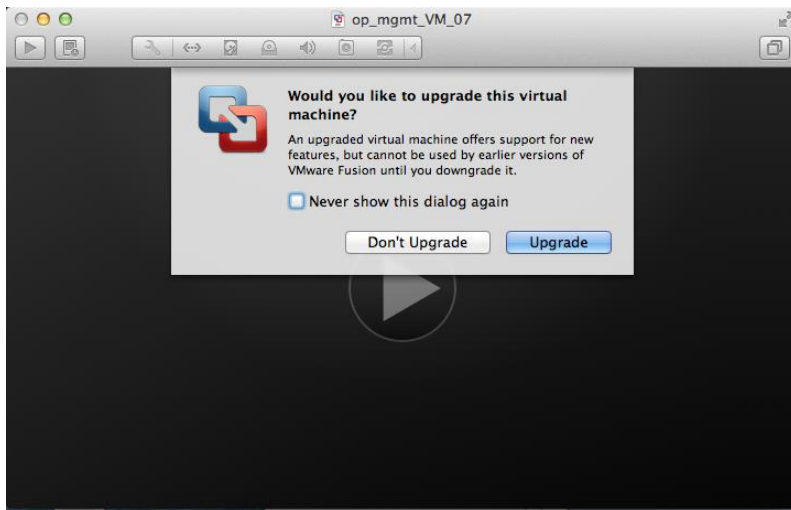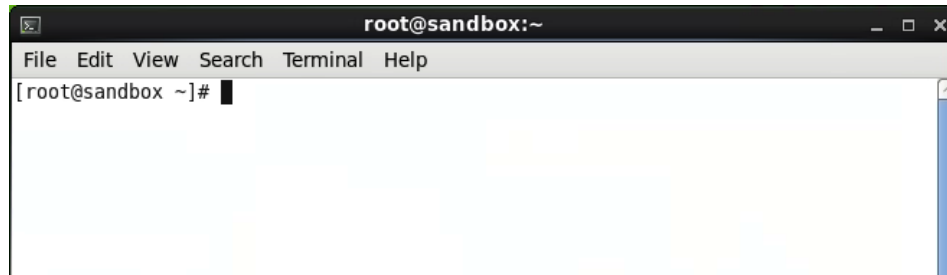
5.3. Wait for the VM to start - it may take several minutes. When the VM has started successfully, you should see the following screen:
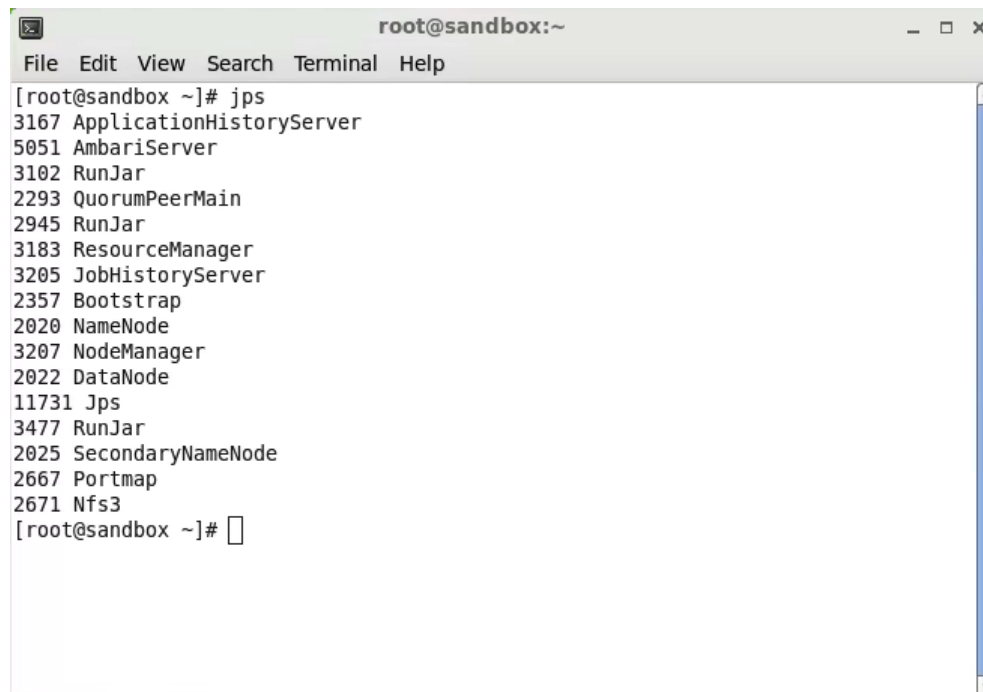

5.4. Double click on Terminal icon:

Copyright © 2015, Hortonworks, Inc.

5.5. Wait for few minutes and enter following commands to verify all the components of the cluster are running:

```
# jps
```

You should see a list of hadoop daemons running on this VM:

RESULT: You should be connected to the classroom VM running on AWS.

IMPORTANT: If any of these steps failed, or if you have any issues or questions, please contact to your instructor OR send an email to training-support@hortonworks.com.