Carsten Hood
UIN: 922009787

Project 3 – Playing With Hadoop

TURNING IN 1 DAY LATE AND USING 1 OF MY LATE DAY PASSES

Turning in one day late and using one of my late date passes.

Step 1: Prepare Linux Environment

(a) I spent 4 minutes on Step 1.

(b) No problems or questions. (I already had an Ubuntu distribution set up for CSCE 465 Computer & Network Security so I simply opened VirtualBox, checked my Ubuntu version, and took a screenshot.)

(c) Screenshot of Ubuntu running on my Macbook Pro using Oracle VirtualBox:

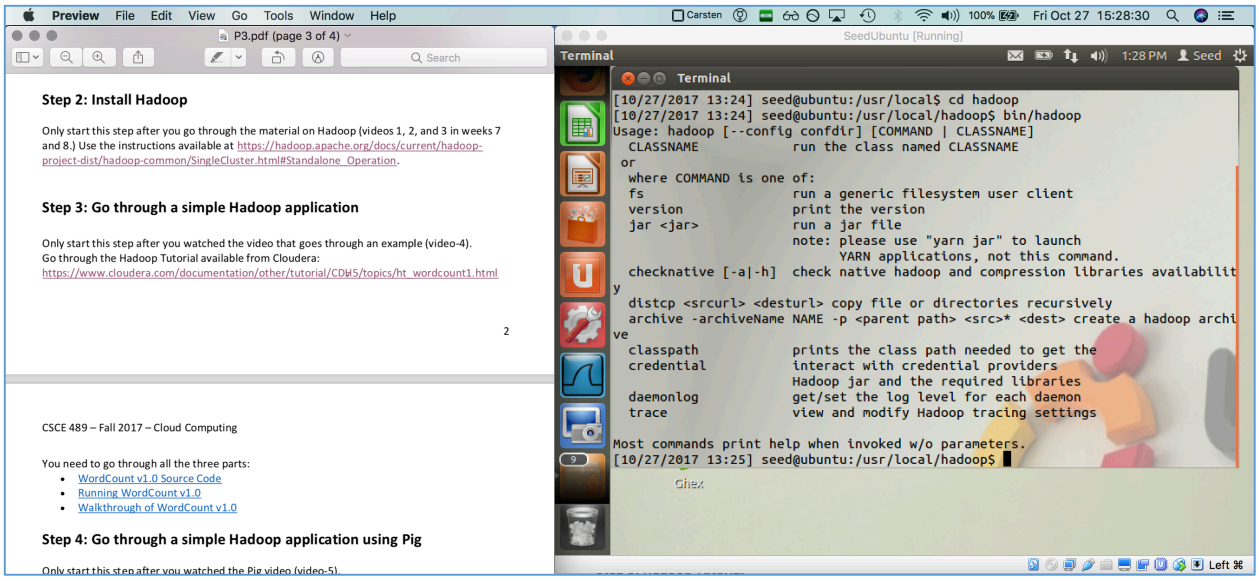
Step 2: Install Hadoop

(a) I spent 55 minutes on Step 2.

(b) Comments:

- My Ubuntu distribution (via VirtualBox) failed to develop an Internet connection at first, which precluded downloading *pdsh* and the Hadoop distribution. I spent a while struggling with this problem before it was resolved (apparently on its own).
- It was unclear which Hadoop version to install through the download link. I used several additional online tutorials to figure this out and install Hadoop:
    - https://www.digitalocean.com/community/tutorials/how-to-install-hadoop-in-stand-alone-mode-on-ubuntu-16-04
    - http://www.bogotobogo.com/Hadoop/BigData_hadoop_Install_on_ubuntu_single_node_cluster.php)

(c) Screenshot of successfully running `$ bin/hadoop`:

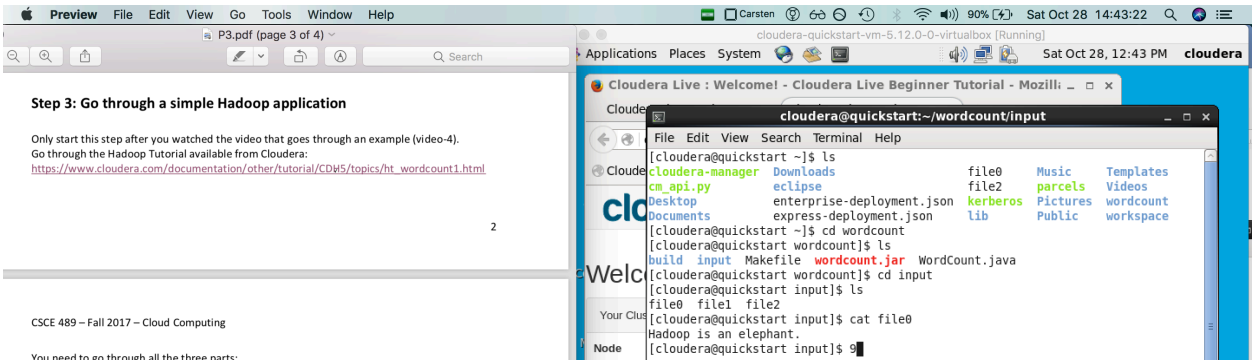


Step 3: Hadoop Tutorial

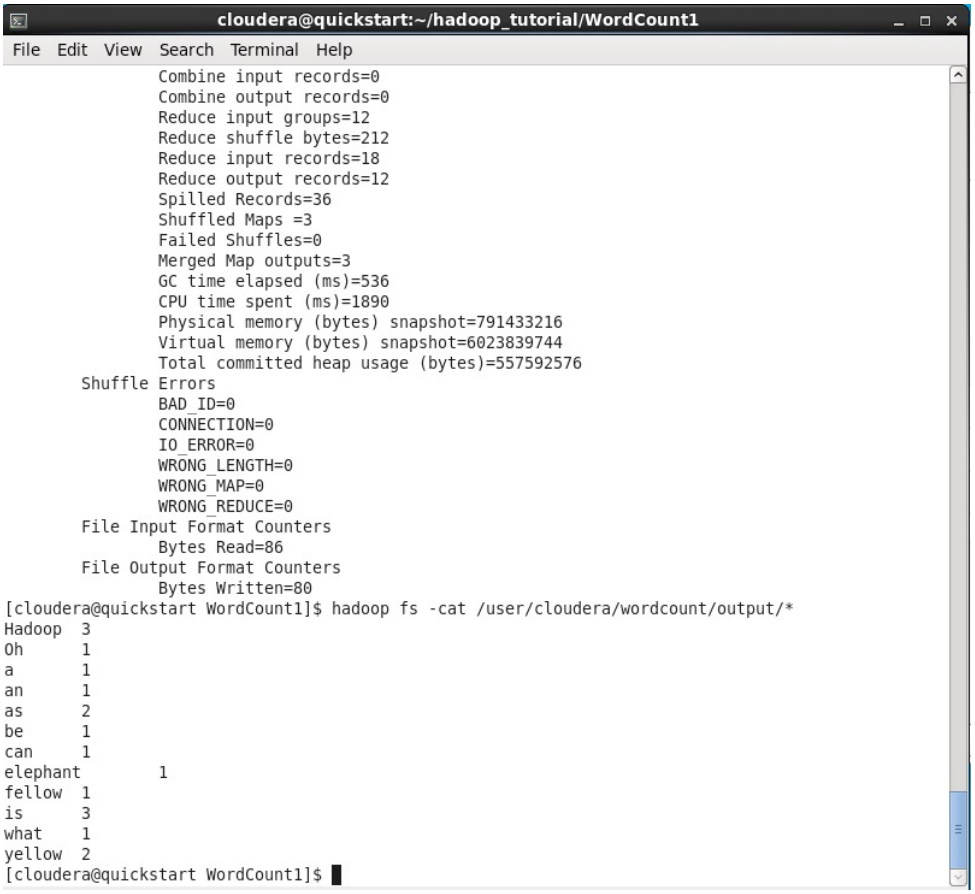
(a) I spent 110 minutes on Step 3.

(b) I started by attempting to run the word-count program on the Ubuntu machine set up in the first step. Since the tutorial is written for the Cloudera QuickStart VM, this led to a lot of confusion and wasted time (as I tried to figure out how to adapt the given commands to similar commands on the Ubuntu system. Eventually it was suggested on Piazza to use the Cloudera QuickStart VM and I was able to get the program running.

(c) Screenshot of Cloudera QuickStart VM setup and prep for step 3:



Screenshot of running word-count program and resulting output file:



Step 4: Pig Tutorial

(a) I spent 40 minutes on Step 4.

(b) I kept getting errors early on following the Pig setup tutorial; eventually I was advised to install Java 8 (instead of using Java 6) but I may not have don't this correctly since the errors remained.

(c) Screenshots of updating java installation to do Pig tutorial.