Sparkst Demo Instructions

1. Set up ssh tunneling to your sandbox
2. Use scp to copy the sparkst.zip file to your maria\_dev account (/home/maria\_dev)
3. Log on to the maria\_dev account
4. Execute: unzip sparkst.zip
5. Enter: su root
6. Enter the password you set a while ago for your root account
7. Enter: vi /usr/hdp/current/spark2-client/conf/log4j.properties
8. Use vi to change the line

log4j.rootCategory=INFO, console

to

log4j.rootCategory=WARN, console

9) Save the change and exit vi

1. To leave root enter: exit
2. Enter: cd sparkst
3. Enter the following line to enable use of python 2.7 which you should have installed for assignment #3

scl enable python27 bash

1. Open up a second terminal window to the maria\_dev account
2. Enter the following line to enable use of python 2.7 which you should have installed for assignment #3

scl enable python27 bash

1. Enter: cd sparkst
2. In one maria\_dev terminal window enter ./produce.sh
3. In a second terminal window enter ./consume.sh
4. Go back to the first terminal window and enter some sequence of words and press the ‘return/enter’ key
5. Observe the spark streaming consumer executing on the other terminal window execute a word count on the sequence of words it has received in real time
6. Repeat steps 18 and 19 as you wish