Running Spark Streaming example (Socket Streaming)

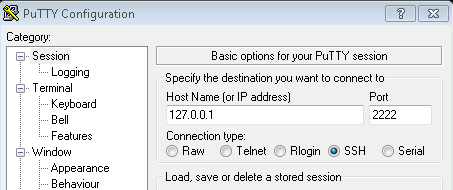
First Session:

* Connect to the sandbox using the document:
* Sandbox\_Opening\_Tutorial
* Once connected to the sandbox host, navigate to the spark directory.

cd /root/spark-1.1.0.2.1.5.0-695-bin-2.4.0.2.1.5.0-695

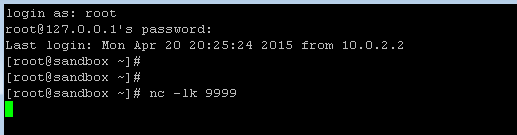
Second session:

* Open another putty session for the sandbox
* Host id: 127.0.0.1
* Port: 2222
* User: root
* Password: hadoop



* In the new terminal start a server using the Netcat

[root@sandbox ~]# nc -lk 9999

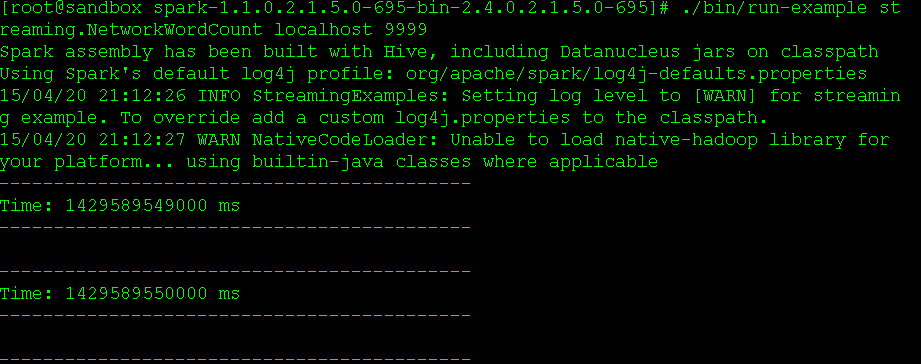


* It acts as a data service.

Running example code:

* Go to first terminal execute the below command to start the spark streaming example.

./bin/run-example streaming.NetworkWordCount localhost 9999



* The streaming example will start running for every second to receive the streaming data on port 9999

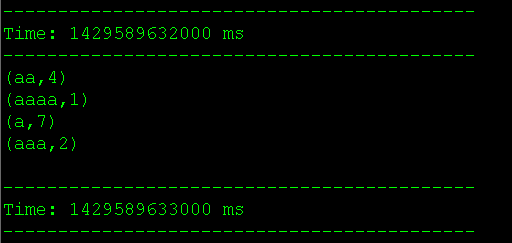
Input Data

* From the second terminal start entering data.



Program Execution

* The data will be received by the streaming program and will start counting the number of same words entered in a line.



Overview:

