Bako Hama Earvin Big Data Final project In this project, we will be using SparkStreaming to read a series of line on a local port (9999) of our computer. I opened one putty with an Amazon EC2 instances to launch spark. I used the same putty to launch another terminal from the same EC2 instances. I then connected it to the same port (9999) by typing nc lk -9999 in our second terminal. Finally, you can start the stream and start getting counting words. Once you are done typing your lines, you can interrupt kernel and check your results. Be careful to have that port free before running the program.

In [10]:
```
import findspark
```

In [11]:
```
findspark.init('/home/ubuntu/spark-2.4.4-bin-hadoop2.7')
```

In [12]:
```
import pyspark
```

In [13]:
```
from pyspark import SparkContext
```

In [14]:
```
from pyspark.streaming import StreamingContext
```

In [ ]:
```
sc = SparkContext('local[2]', 'NetworkWordCount')
#Our application will be working with two local threads
```

In [16]:
```
ssc = StreamingContext(sc,1)
#We will be taking the content in batch of 1 second.
```

In [17]:
```
lines = ssc.socketTextStream("localhost", 9999)
#This the connection to the local port 9999
```

In [18]:
```
words = lines.flatMap(lambda line: line.split(' '))
#This functions to break each line into individual words
```

In [19]:
```
pairs = words.map(lambda word:(word,1))
#We use pairs to count the words
```

In [20]:
```
word_counts = pairs.reduceByKey(lambda num1,num2: num1+num2)
#This reduce by key will help us get the number of time the same word is repeated in one lin
e
```

In [21]:
```
word_counts.pprint()
```

In [22]:
```
ssc.start()
#We can see the reading gowaring here once we start the stream,
#And if we want to stop it, we can interupt kernel and see the result
#We will be using another console to type in our lines.
```

```
-------------------------------------------
Time: 2020-05-04 19:06:26
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:27
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:28
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:29
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:30
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:31
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:32
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:33
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:34
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:35
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:36
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:37
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:38
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:39
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:40
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:41
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:42
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:43
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:44
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:45
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:46
-------------------------------------------
('name', 1)
('is', 1)
('My', 1)
('Earvin', 1)

-------------------------------------------
Time: 2020-05-04 19:06:47
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:48
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:49
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:50
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:51
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:52
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:53
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:54
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:55
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:56
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:57
-------------------------------------------
('love', 1)
('basketball', 1)
('anime', 1)
('I', 1)
('and', 1)

-------------------------------------------
Time: 2020-05-04 19:06:58
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:06:59
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:00
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:01
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:02
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:03
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:04
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:05
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:06
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:07
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:08
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:09
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:10
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:11
-------------------------------------------
('live', 1)
('brother', 1)
('', 1)
('I', 1)
('with', 1)
('my', 1)

-------------------------------------------
Time: 2020-05-04 19:07:12
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:13
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:14
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:15
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:16
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:17
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:18
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:19
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:20
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:21
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:22
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:23
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:24
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:25
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:26
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:27
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:28
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:29
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:30
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:31
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:32
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:33
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:34
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:35
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:36
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:37
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:38
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:39
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:40
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:41
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:42
-------------------------------------------
('Hunter', 2)
('is', 1)
('favorite', 1)
('anime', 1)
('A', 1)
('one', 1)
('my', 1)

-------------------------------------------
Time: 2020-05-04 19:07:43
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:44
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:45
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:46
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:47
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:48
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:49
-------------------------------------------

-------------------------------------------
Time: 2020-05-04 19:07:50
-------------------------------------------
```