

CprE419 Lab 4
Neha Maddali

Experiment 1:

Run exp1.pig with: `pig -x local exp1.pig`

Top 10 states according to land area:

part-r-00000	
~/Downloads/lab4_exp1/output	
AK	1477953211577
TX	676586997978
CA	403466310059
MT	376961878670
NM	314160748240
AZ	294207314414
NV	284331937541
CO	268431246426
WY	251470069067
OR	248607802255

Experiment 2:

Create lab 3 folder: `hadoop fs -mkdir /lab4`

Move network_trace to lab4 folder in HDFS: `hadoop fs -put ~/Downloads/network_trace /lab4`

Run mapreduce on exp2.pig: `pig -x mapreduce exp2.pig`

```
cpre419@cpre419-VirtualBox:~/Downloads$ hadoop fs -cat /lab4/exp2/output/*
2023-03-11 21:39:49,150 INFO sasl.SaslDataTransferClient: SASL encryption trust
check: localhostTrusted = false, remoteHostTrusted = false
10.31.24.241      273
192.168.1.4       114
10.31.4.211       70
10.31.4.188       52
107.152.26.197    4
17.134.127.250    4
162.125.34.137    4
162.125.33.7      4
162.125.18.133    4
107.152.25.197    4
```

Experiment 3:

Move ip_trace to lab4 folder in HDFS: `hadoop fs -put ~/Downloads/ip_trace /lab4`

Move raw_block to lab4 folder in HDFS: `hadoop fs -put ~/Downloads/raw_block /lab4`

Run mapreduce on exp3.pig: `pig -x mapreduce exp3.pig`

Firewall:

```
cpre419@cpre419-VirtualBox:~/Downloads$ hadoop fs -cat /lab4/exp3/firewall/part-r-00000 | head -n 10
2023-03-11 22:07:00,262 INFO sasl.SaslDataTransferClient: SASL encryption trust
check: localhostTrusted = false, remoteHostTrusted = false
15:28:46.539147 0 10.31.27.252.1985 224.0.0.102.1985: Blocked
15:28:47.479137 1 10.31.26.186.5353 224.0.0.251.5353: Blocked
15:28:47.479217 2 10.31.26.216.5353 224.0.0.251.5353: Blocked
15:28:48.105882 3 10.31.27.253.1985 224.0.0.102.1985: Blocked
15:28:48.628287 5 10.10.251.232 224.0.0.1: Blocked
15:28:49.672578 8 10.31.26.216.5353 224.0.0.251.5353: Blocked
15:28:50.194885 10 10.31.24.175.5353 224.0.0.251.5353: Blocked
15:28:50.299328 11 10.33.14.68.5353 224.0.0.251.5353: Blocked
15:28:50.403757 12 10.33.14.68.5353 224.0.0.251.5353: Blocked
15:28:50.508240 13 10.31.26.216.5353 224.0.0.251.5353: Blocked
cat: Unable to write to output stream.
```

Output:

```
cpre419@cpre419-VirtualBox:~/Downloads$ hadoop fs -cat /lab4/exp3/output/part-r-00000 | head -n 10
2023-03-11 22:10:58,306 INFO sasl.SaslDataTransferClient: SASL encryption trust
check: localhostTrusted = false, remoteHostTrusted = false
13.249.140.62.443 73223
23.35.205.78.443 46495
10.31.24.241.64121 18179
151.101.148.133.443 11473
130.211.46.84.443 10590
10.31.24.241.63030 10361
100.24.180.47.443 3864
10.31.27.253.1985 3809
10.31.27.252.1985 3764
10.31.24.241.64037 3357
cat: Unable to write to output stream.
```