DS 256 - SSDS
Assignment 1

Diksha Chaudhary
PhD Student(15934)
Department of Computational and Data Sciences
Indian Institute of Science
Bangalore, Karnataka 560012, India

March 14, 2019

1 Spark

Implemented Pagerank in Spark. But not able to access YarnLogs for analysis. The application ids are as follows:

- application_1552131679281_0026

- application_1552131679281_0027

- application_1552214754151_0035

- application_1552214754151_0035

2 Giraph

Implemented Pagerank in Giraph. I think code is fine but job was failing for the input. Also implemeted WCC in giraph but facing the following error:

```
2019-03-14 16:27:59,826 INFO yarn.GiraphYarnClient: Obtained new Application ID: application_1552214754151_0263
2019-03-14 16:27:59,826 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
2019-03-14 16:27:59,832 INFO yarn.GiraphYarnClient: Set the environment for the application master
2019-03-14 16:27:59,835 INFO yarn.GiraphYarnClient: Environment for AM :{CLASSPATH=${CLASSPATH}:./*:$HADOOP_CONF_DIR:${HADOOP_COMMON_HOME}/sha
e/hadoop/common/*:${HADOOP_COMMON_HOME}/share/hadoop/common/lib/*:${HADOOP_HDFS_HOME}/share/hadoop/hdfs/*:${HADOOP_HDFS_HOME}/share/hadoop/hdf
/lib/*:${HADOOP_MAPRED_HOME}/share/hadoop/mapreduce/*:${HADOOP_MAPRED_HOME}/share/hadoop/mapreduce/lib/*:${HADOOP_YARN_HOME}/share/hadoop/yarn
*:${HADOOP_YARN_HOME}/share/hadoop/yarn/lib/*:$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*:$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*}
2019-03-14 16:27:59,835 INFO yarn.GiraphYarnClient: buildLocalResourceMap ....
2019-03-14 16:27:59,835 WARN yarn.YarnUtils: Unable to delete file /tmp/giraph-conf.xml
Exception in thread "main" java.lang.IllegalStateException: Failed to build LocalResouce map.
        at org.apache.giraph.yarn.GiraphYarnClient.buildLocalResourceMap(GiraphYarnClient.java:519)
        at org.apache.giraph.yarn.GiraphYarnClient.buildContainerLaunchContext(GiraphYarnClient.java:351)
        at org.apache.giraph.yarn.GiraphYarnClient.run(GiraphYarnClient.java:136)
        at org.apache.giraph.GiraphRunner.run(GiraphRunner.java:96)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:90)
        at org.apache.giraph.GiraphRunner.main(GiraphRunner.java:126)
        at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
        at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
        at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
        at java.lang.reflect.Method.invoke(Method.java:498)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:318)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:232)
Caused by: java.io.FileNotFoundException: /tmp/giraph-conf.xml (Permission denied)
        at java.io.FileOutputStream.open0(Native Method)
        at java.io.FileOutputStream.open(FileOutputStream.java:270)
        at java.io.FileOutputStream.<init>(FileOutputStream.java:213)
        at java.io.FileOutputStream.<init>(FileOutputStream.java:101)
        at org.apache.giraph.yarn.YarnUtils.exportGiraphConfiguration(YarnUtils.java:235)
        at org.apache.giraph.yarn.GiraphYarnClient.buildLocalResourceMap(GiraphYarnClient.java:511)
```