Lab 2: HDFS namenode images

(DSCI 551, Fall 2022)

Due: September 23, 2022

Points: 10

1. [4 points] Create at least two levels of hierarchy (e.g., /user/john) and at least one file under a directory (e.g., /user/john/hello.txt) in your HDFS server. For example,

```
Starting secondary namenodes [ip-172-31-22-2.us-west-1.compute.internal]
[ec2-user@ip-172-31-22-2 ~]$ hdfs dfs -ls /user/john
Found 3 items
-rw-r--r--   1 ec2-user supergroup       3269 2022-09-06 20:05 /user/john/WordCount.java
drwxr-xr-x   - ec2-user supergroup          0 2022-09-06 20:07 /user/john/a
-rw-r--r--   1 ec2-user supergroup         37 2022-09-13 23:04 /user/john/dept.csv
[ec2-user@ip-172-31-22-2 ~]$
```

2. [6 points] Use hdfs oiv command to export your HDFS file system image in both XML and TSV (tab-separated value) format. Recall that the images are stored in the /tmp/hadoop-ec2-user/dfs/name/current directory.

   Examples:

   hdfs oiv -i fsimage_0000000000000000160 -o fsimage160.xml -p XML

   hdfs oiv -i fsimage_0000000000000000160 -o fsimage160.tsv -p Delimited

Submission:

1. screenshot for task 1 (similar to the example above)
2. XML file of your HDFS image for task2
3. TSV file of your HDFS image for task2