CS 6381 Final Project Cassandra Setup:

General:

1. Spin up a Ubuntu VM (e. g. 18.04).
2. Install Cassandra.
   1. sudo apt install openjdk-8-jre-headless
   2. cd /usr/lib/jvm/java-8-openjdk-amd64
   3. sudo nano ~/.bashrc
   4. Scroll to bottom of page and add in “JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64”
   5. Save the file using Ctrl + S on Windows or Command + S on Mac and then exit out using Ctrl + X on Windows or Command + X on Mac.
   6. Reopen a new terminal.
   7. echo $JAVA\_HOME = See if above setup works.
   8. sudo apt install curl
   9. echo "deb http://downloads.apache.org/cassandra/debian 39x main" | sudo tee -a /etc/apt/sources.list.d/cassandra.sources.list
   10. curl https://downloads.apache.org/cassandra/KEYS | sudo apt-key add -
   11. sudo apt-get update
   12. sudo apt-get install cassandra
3. Run Cassandra.
   1. Start: sudo service cassandra start
   2. Check status: sudo service cassandra status (should see “active (running)”)
   3. Stop: sudo service cassandra stop
   4. If Cassandra stops after some time (e. g. Python application stops connecting to Cassandra), do the stop command then start command.
4. Create keyspace and table.
   1. Open Cassandra terminal: cqlsh
   2. Create keyspace (kind of like a schema): Create KeySpace {{keyspace name (e. g. cs6381finalprojectks1)}} with replication = {‘class’: ‘SimpleStrategy’, ‘replication\_factor’: 3};
   3. Show all keyspaces: SELECT \* FROM system\_schema.keyspaces;
   4. Select keyspace: use {{your keyspace name}}
   5. Create table: Create Table table1(id {{id or uuid depending on use}} Primary Key, message text);
   6. Drop table: drop table {{table name (e. g. table1}}
5. Run code.
   1. Manually
      1. Navigate to Code folder and run python3 cassandraRW.py accordingly to run writers/readers.
      2. Run python3 calculateStatistics.py to calculate statistics of time differences.
      3. Run python3 capAnalysis.py to get CAP-related times.
   2. Tests
      1. Navigate to Code/tests folder and run a test (e. g. to run all.sh, which runs all the tests in that folder, run “./all.sh”).

Procedure:

1. Run hardwareMeasure.py, which will continuously run so the user can constantly see the latest max and average statistics. After running all conditions as follows, add latest statistics to base file table.
2. cassandraRW.py
   1. 1 VM x 1writerTest
   2. 1 VM x 1readerTest
   3. 10 VM x 10writer1readerTest
   4. 10 VM x 1writer10readerTest
   5. 5 VM x 25writer25readerTest
3. capAnalysis.py - Run test in 1 of the following ways.
   1. Way 1
      1. Create table, run server, insert, and continuously update.
      2. On each of 10 VM’s: Run readIteration.py with a parameter of 10 for 10 iterations.
   2. Way 2
      1. 10 VM x 1insert1write10readTest