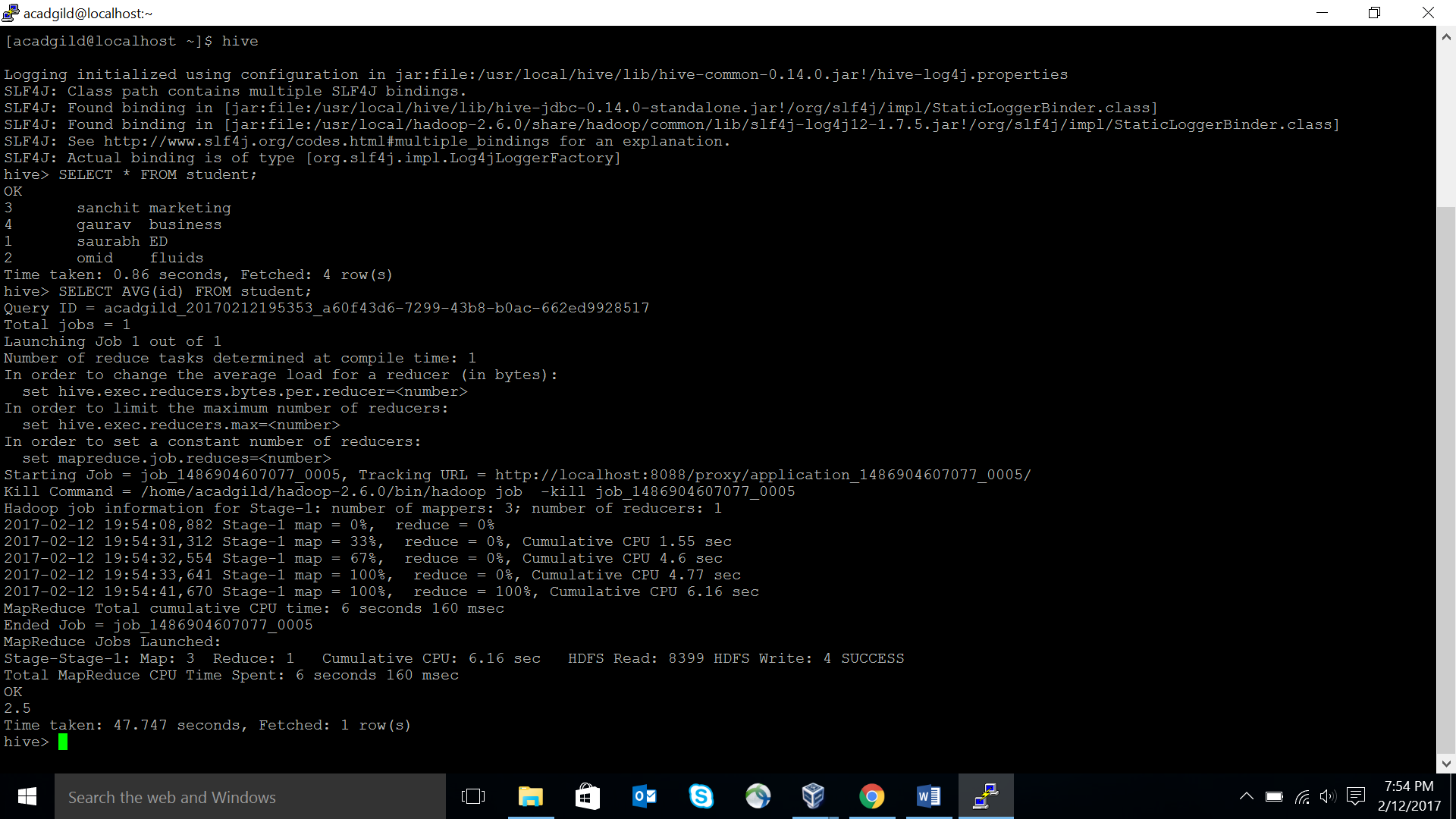
**SESSION 16: UDF DEMO, THRIFT SERVER**

**ASSIGNMENT 2**

**TABLE DATA:**

Hive>SELECT \* FROM student;

Hive>SELECT AVG(id) FROM student;



**--Time taken: 47.747 seconds**

**CREATING COMPACT INDEX IN HIVE:**

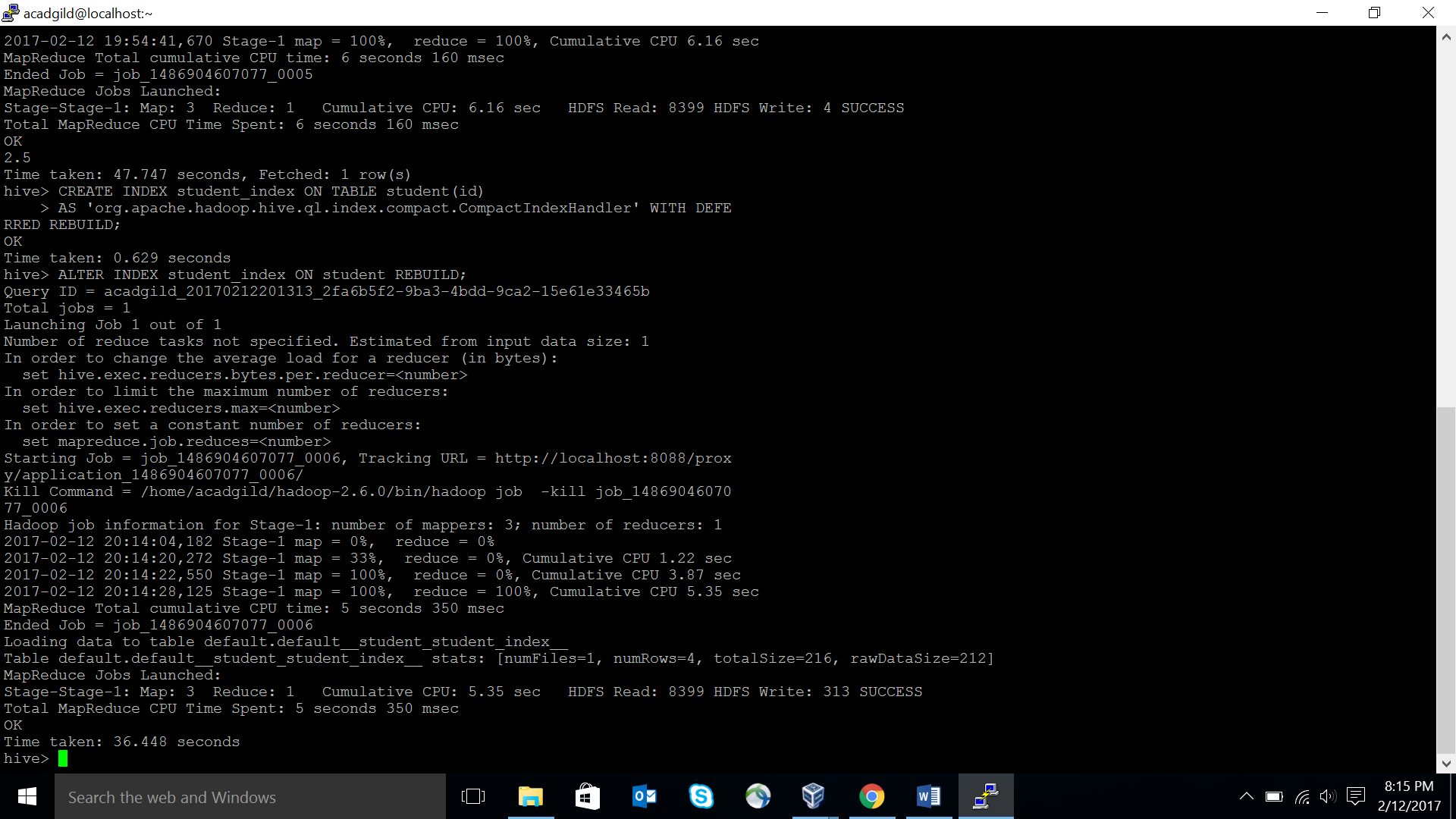
Hive>CREATE INDEX student\_index ON TABLE student(id)

>AS ‘org.apacha.hadoop.hive.ql.index.compact.CompactIndexHandler’ WITH DEFERRED

REBUILD;

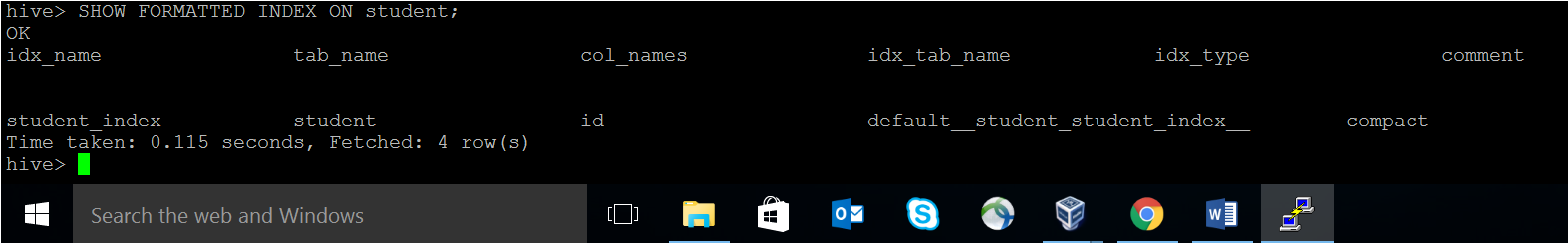
--ALTER INDEX

Hive>ALTER INDEX student\_index ON student REBUILD;



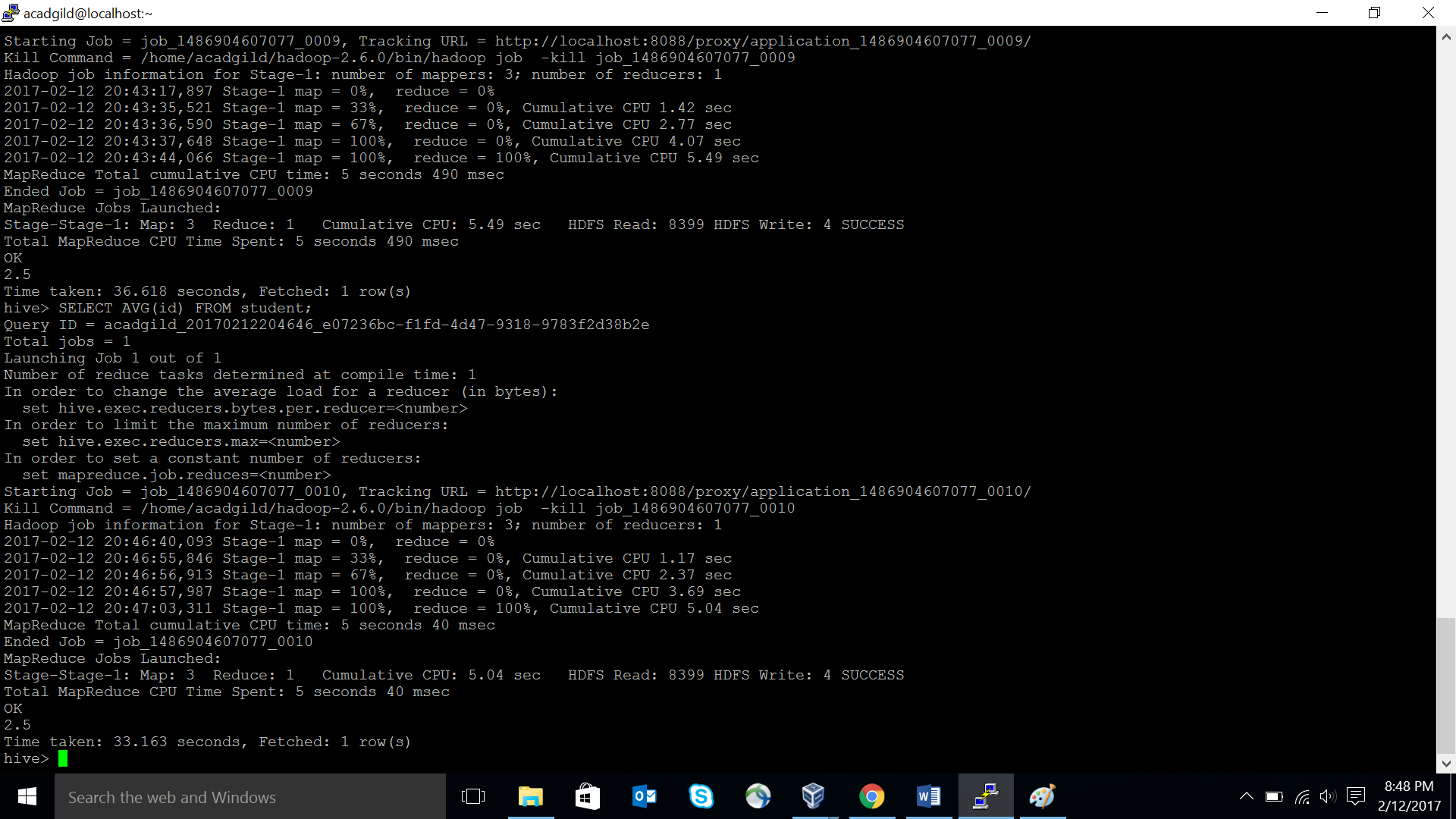
--SHOW INDEX

Hive> SHOW FORMATTED INDEX ON student;



--COMPARE PERFORMANCE BY PERFORMING SAME AVG FUNCTION

Hive>SELECT AVG(id) FROM student;



**--Time taken: 33.163 seconds, as compared to 47.747 seconds when there was no index.**

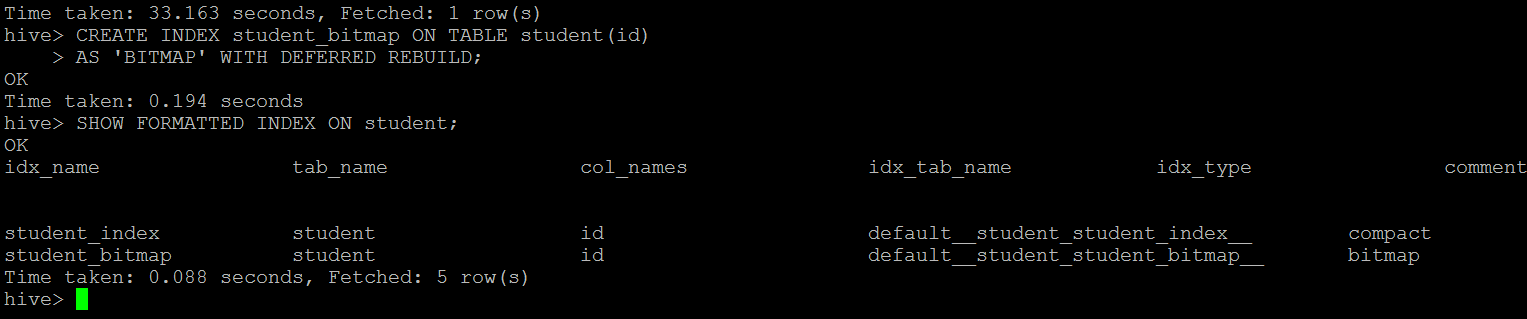
**CREATING BITMAP INDEX IN HIVE:**

Hive>CREATE INDEX student\_bitmap ON TABLE student(id)

>AS ‘BITMAP’ WITH DEFERRED REBUILD;

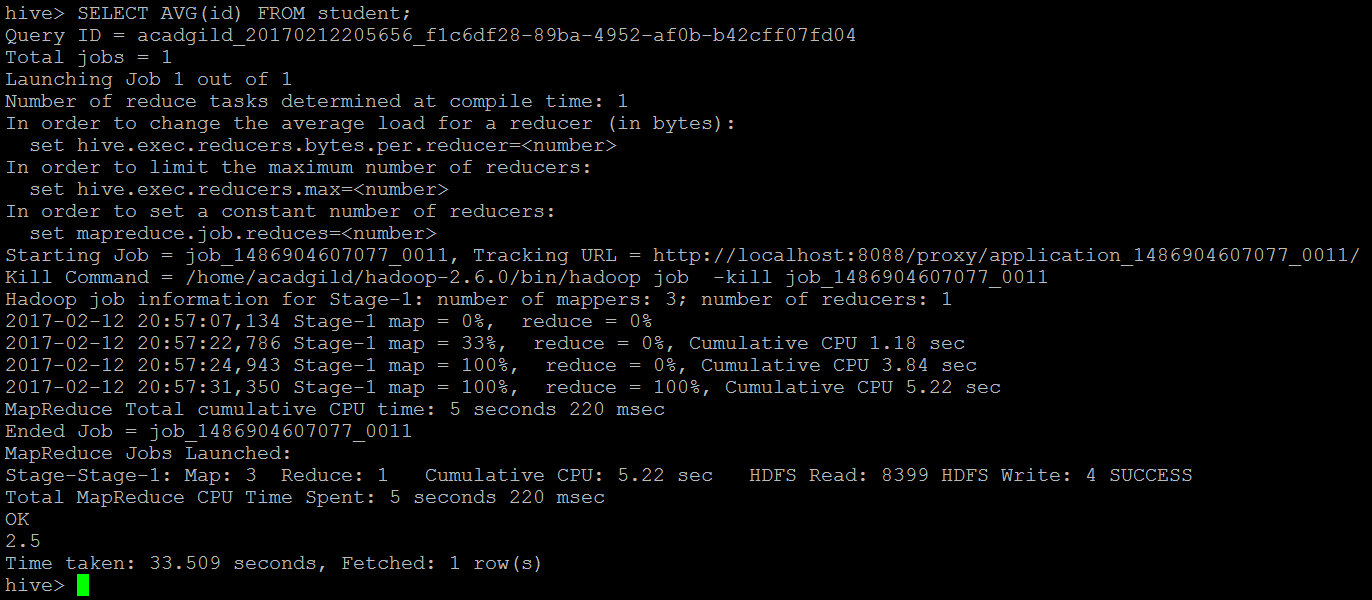
--SHOW INDEX

Hive> SHOW FORMATTED INDEX ON student;

****

**AVERAGE OPERATION WITH TWO INDEXES:**

Hive>SELECT AVG(id) FROM student;

****

**--Time taken: 33.509 seconds, which is almost same as in case of compact index.**

**AVERAGE OPERATION WITH ONLY BITMAP INDEX:**

-- DROP COMPACT INDEX

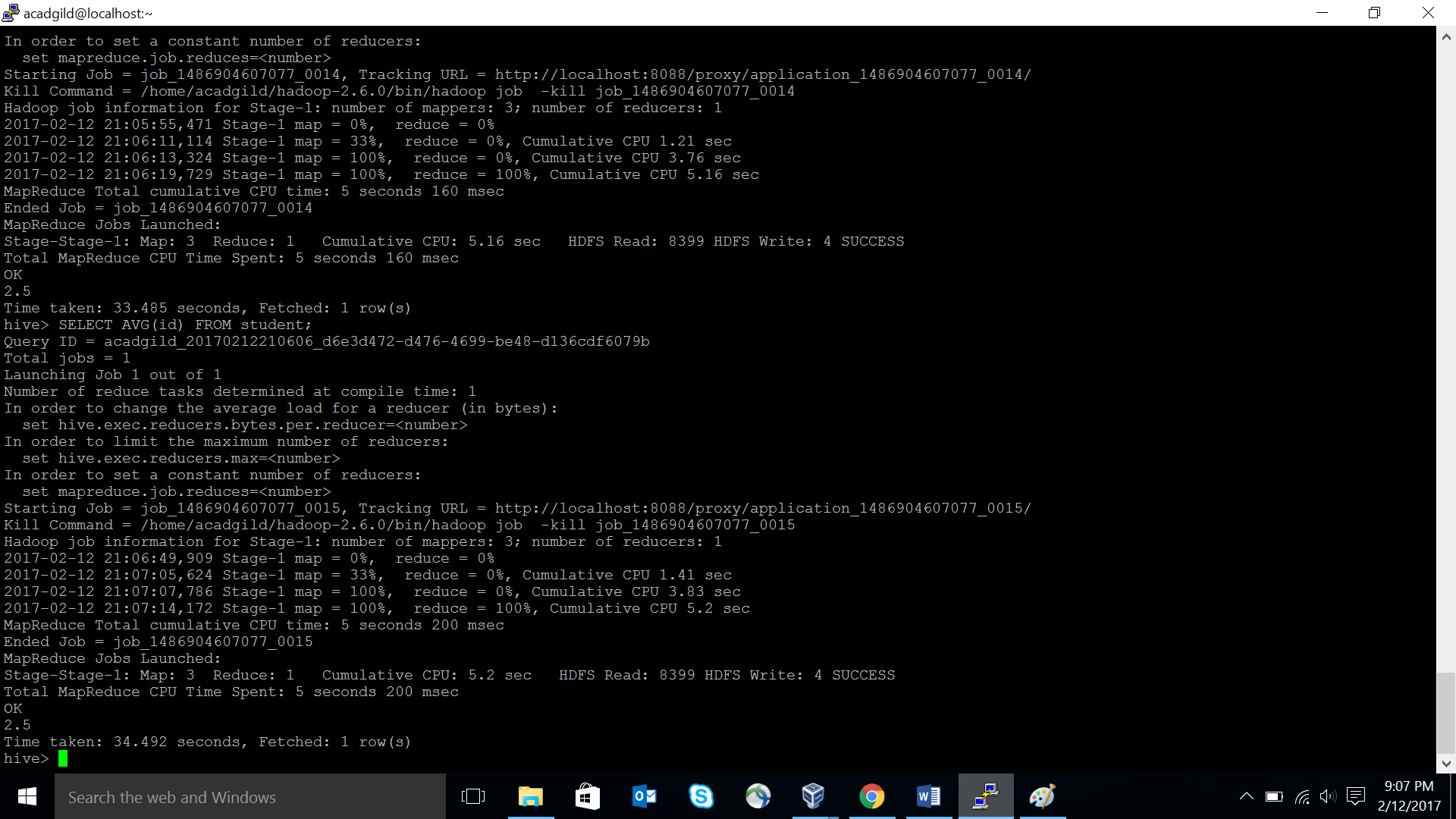
Hive> DROP INDEX student\_index ON student;

--SHOW INDEX

Hive> SHOW FORMATTED INDEX ON student;

--PERFORM AVG FUNCTION

Hive>SELECT AVG(id) FROM student;



**--Time taken: 34.492 seconds, which is marginally more than the above two cases.**