

Assignment 1

2014CS50280 - Ashley Jain
2013ME10654 - Sri Ram Bandi

Bulk load :-

Command :- \copy registers(student_id,course_id) from '<global_path>/tuples.csv'
DELIMITER ',' CSV HEADER;

Time To Load : 19.78 sec

It is fastest loader among three approaches, due to its single insert, execute and commit nature. It loads the huge chunk of data in a single insert and commit step.

Batch Insert :-

SQL command :- INSERT INTO registers(student_id,course_id) VALUES (
student_id1, course_id1) , (student_id2, course_id2) , 500000 tuples ;

Time To Load : 27.51 sec

It takes longer than first approach since this approach inserts, logs and commits after every insertion, having more overhead than Bulk insertion.

JDBC :-

Program :- Insert Batch statement into database using jdbc driver with java code
(given in "code" folder)

Time To Load : 39.84 sec

It takes longer time than other approaches due to overhead time in connection to database and load csv(500000 tuples) files. We implemented Batch insert model in this JDBC loading. So, it adds server connectivity and programmatic(for loops) overheads to the already poor performance of batch insert.

*For reproducing the similar results, use this machine :
Mac OSX, High Sierra, 4GB RAM, i5 Core