Deniz Muratlı, 29480

I chose the tables “Labels” and “Albums” in my database. Since *Albums* had foreign keys referring to both *Labels* and *Artists*, I ended up having to populate 3 tables (*Labels, Artists, Albums*), each with a million dummy entries. I did this by modifying connect.py and populator.py from Recitation 8 to fit my database and tables. I had to maintain a separate populator script for each of the tables, as *Albums* had foreign key constraints that had to be satisfied, and changes to the tables could not be committed en masse.

I then moved on to benchmarking queries before and after indexing. My first query joined *Albums* and *Labels* and found lines where the label’s establishment year was before 2000 and the album’s release year was after 2000:

SELECT \*

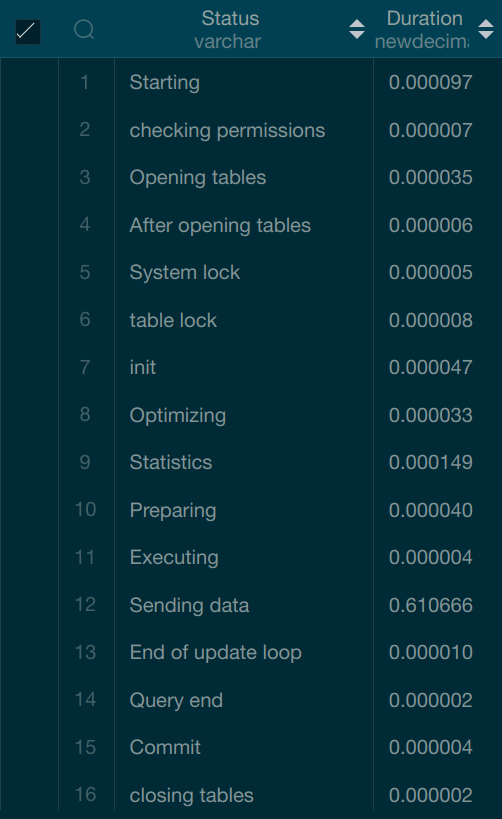
FROM Labels

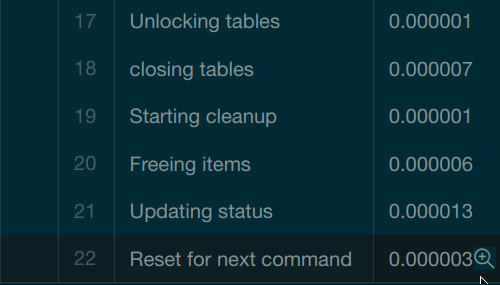
JOIN Albums ON Labels.lid = Albums.lid

WHERE Albums.rel\_year > 2000 and Labels.est\_year < 2000

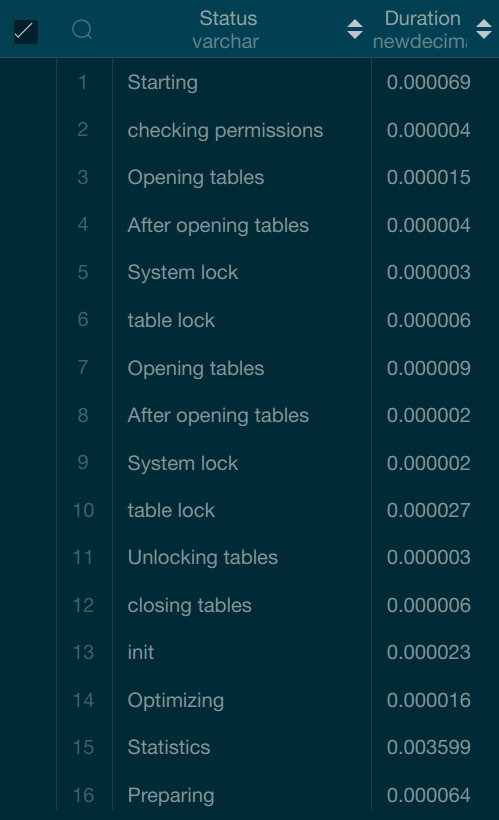
LIMIT 1000009;

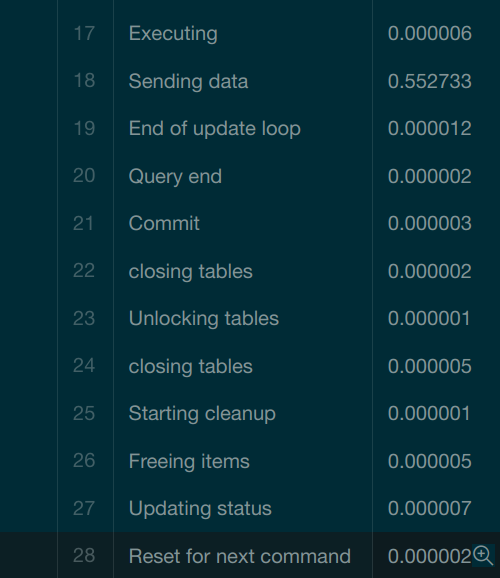
Its benchmark results **before** indexing are as follows:





Its benchmark results **after** indexing are as follows:



My second query joined *Albums* and *Labels* and found lines where the label was based in Sweden and the album name contained an “r” at any position:

SELECT \*

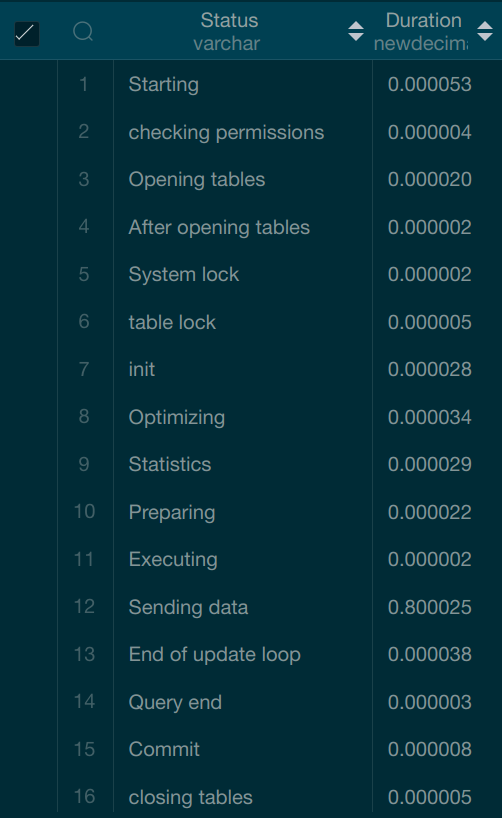
FROM Labels

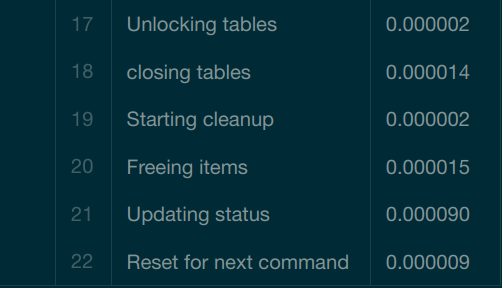
JOIN Albums ON Labels.lid = Albums.lid

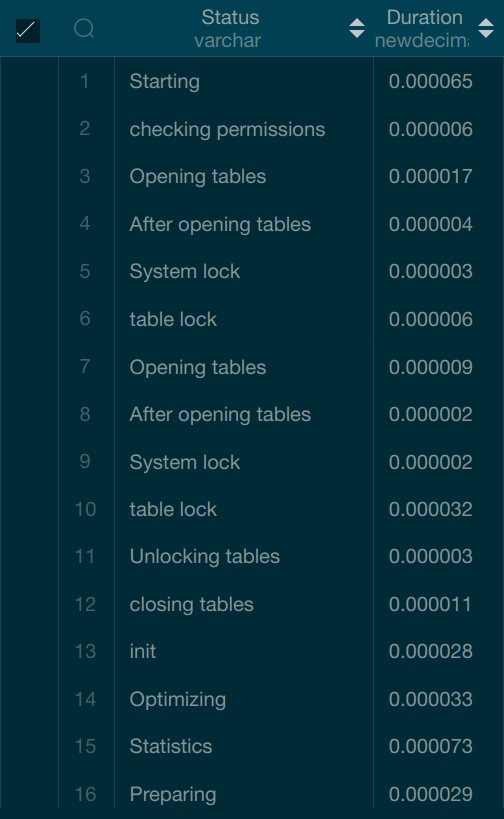
WHERE Labels.lcountry = "Sweden" AND Albums.atitle LIKE "%r%"

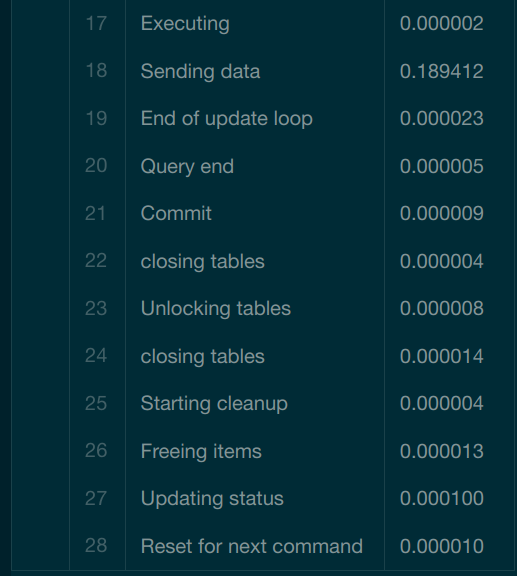
LIMIT 1000009;

Its benchmark results **before** indexing are as follows:



Its benchmark results **after** indexing are as follows:



  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
I used the code provided in Recitation 8 for indexing and conducting the benchmark tests. For indexing, I used the following commands for the two different cases:

1) CREATE INDEX rel\_year\_ind ON Albums(rel\_year);

CREATE INDEX est\_year\_ind ON Labels(est\_year);

2) CREATE INDEX atitle\_ind ON Albums(atitle);

CREATE INDEX lcountry\_ind ON Labels(lcountry);

In both cases, I benchmarked the query as follows:

SET profiling = 1;

[QUERY]

SHOW PROFILE;

All the code I used and the benchmark results can be found at <https://drive.google.com/drive/folders/1YGFVkIaMmsIkqLhvdsYigV07-m-Yg5FO?usp=sharing>.