Select count is slow

Last updated by | Lisa Liu | Nov 6, 2020 at 10:35 AM PST

Select count is slow

Monday, February 4, 2019
2:34 PM

As I have researched this issue and was looking into a solution I found plenty of documents discussing how slow select * from a PostgreSQL table is, The reason why this is slow is related to the MVCC implementation in PostgreSQL. The fact that multiple transactions can see different states of the data means that there can be no straightforward way for "COUNT(*)" to summarize data across the whole table; PostgreSQL must walk through all rows, in some sense. This normally results in a sequential scan reading information about every row in the table.

1- https://wiki.postgresql.org/wiki/Slow_Counting
2- https://www.postgresql.org/message-id/B9C9130B5D27D4119D5D00A0C9D3A987109583%40SERVER
3- https://stackoverflow.com/questions/7943233/fast-way-to-discover-the-row-count-of-a-table-in-postgresql
4- http://www.postgresqltutorial.com/postgresql-count-function/

Action Plan
========
1- I would highly suggest that you try our Memory Optimized servers instead of the general purpose as Select count(*) operation heavily utilizes the disk content.
2- try the recommendation in this webpage: https://stackoverflow.com/questions/7943233/fast-way-to-discover-the-row-count-of-a-table-in-postgresql
Instead of getting the *exact* count (**slow** with big tables):
SELECT count(*) AS exact_count FROM myschema.mytable;
You get a close estimate like this (**extremely fast**):
SELECT reltuples::bigint AS estimate FROM pg_class where relname='mytable';
How close the estimate is depends on whether you run ANALYZE enough. It is usually very close

Created with Microsoft OneNote 2016.

How good have you found this content?