

QGIS Application - Bug report #4788

Add PostGIS layer - connect: Slow performance (>2 minutes) when geometry_columns or geography_columns table has many records

2012-01-11 03:48 PM - Kirill Mueller

| | | |
|--|------------|-------------------------------------|
| Status: | Closed | |
| Priority: | Normal | |
| Assignee: | Tim Sutton | |
| Category: | | |
| Affected QGIS version: | 1.7.3 | Regression?: No |
| Operating System: | | Easy fix?: No |
| Pull Request or Patch supplied: | | Resolution: fixed |
| Crashes QGIS or corrupts data: | | Copied to github as #: 14643 |

Description

Connecting to a Postgres database that has many entries in the geometry_columns or geography_columns table is horribly slow. In our case, with more than 10000 tables with a geometry column, it takes slightly less than three minutes.

There are no performance problems when choosing "Only look in the geometry_columns".

The reason is the following query, issued by QgsPostgresConn::getTableInfo(), line 392 ff.
(89d22a65b630f80047f96d1c36b5885f26cc88ee):

```
SELECT pg_class.relname,pg_namespace.nspname,pg_attribute.attname,pg_class.relkind
FROM pg_attribute,pg_class,pg_namespace
WHERE pg_namespace.oid=pg_class.relnamespace AND pg_attribute.attrelid = pg_class.oid
AND ( EXISTS (SELECT * FROM pg_type WHERE pg_type.oid=pg_attribute.atttypid AND pg_type.typname IN
('geometry','geography','topogeometry')) OR pg_attribute.atttypid IN (SELECT oid FROM pg_type a WHERE EXISTS (SELECT * FROM
pg_type b WHERE a.typbasetype=b.oid AND b.typname IN ('geometry','geography','topogeometry'))))
AND has_schema_privilege( pg_namespace.nspname, 'usage' )
AND has_table_privilege( "" || pg_namespace.nspname || "."" || pg_class.relname || "", 'select' )
AND NOT EXISTS (SELECT * FROM geometry_columns WHERE pg_namespace.nspname=f_table_schema AND
pg_class.relname=f_table_name)
AND pg_class.relkind IN ('v','r');
```

When leaving out the last "AND NOT EXISTS..." part, the query finishes in no time. I have attached the output of EXPLAIN ANALYZE -- if I understand the execution tree correctly, the time is burnt in repeated sequential scans of the geometry_columns table (line 38).?

Rewriting the "AND NOT EXISTS" part using WITH solves the performance issues here, but works only from Postgres 8.4. I have mailed the pqsql-general list for advice on how to speed up this query for older versions, see the discussion at

<http://archives.postgresql.org/pqsql-general/2012-01/msg00265.php> .

Here's the improved version of the problematic query:

```
WITH forbidden_geometry_columns AS (
SELECT
DISTINCT
pg_class.oid AS id
FROM pg_class,pg_namespace,geometry_columns
WHERE pg_namespace.oid=pg_class.relnamespace
AND pg_namespace.nspname=f_table_schema AND pg_class.relname=f_table_name
)
SELECT pg_class.relname,pg_namespace.nspname,pg_attribute.attname,pg_class.relkind
FROM pg_attribute,pg_class,pg_namespace
```

```
WHERE pg_namespace.oid=pg_class.relnamespace AND pg_attribute.attrelid = pg_class.oid
AND ( EXISTS (SELECT * FROM pg_type WHERE pg_type.oid=pg_attribute.atttypid AND pg_type.typname IN
('geometry','geography','topogeometry')) OR pg_attribute.atttypid IN (SELECT oid FROM pg_type a WHERE EXISTS (SELECT * FROM
pg_type b WHERE a.typbasetype=b.oid AND b.typname IN ('geometry','geography','topogeometry'))))
AND has_schema_privilege( pg_namespace.nspname, 'usage' )
AND has_table_privilege( "" || pg_namespace.nspname || "." || pg_class.relname || "", 'select' )
AND NOT EXISTS (SELECT * FROM forbidden_geometry_columns WHERE forbidden_geometry_columns.id=pg_class.oid)
AND pg_class.relkind IN ('v','r');
```

History

#1 - 2012-01-11 04:25 PM - Kirill Mueller

- *File 0001-improve-query-that-lists-all-PostGIS-columns-fix-478.patch added*

See <http://archives.postgresql.org/pgsql-general/2012-01/msg00306.php> . Patch attached.

#2 - 2012-01-11 05:33 PM - Jürgen Fischer

- *Status changed from Open to Closed*
- *Resolution set to fixed*

Kirill Mueller wrote:

See <http://archives.postgresql.org/pgsql-general/2012-01/msg00306.php> . Patch attached.

Nice. Thanks. Applied in commit:a646ff6f.

#3 - 2012-01-11 09:13 PM - Sandro Santilli

looks like a good candidate for backporting to 1.7 and 1.8

#4 - 2012-01-12 12:26 AM - Kirill Mueller

Actually, a backport would be nice :-)

#5 - 2012-01-12 01:12 AM - Sandro Santilli

- *Status changed from Closed to Reopened*
- *Affected QGIS version changed from master to 1.7.3*
- *Target version set to Version 1.7.4*

Backport in 1.7 is commit:c6eb2364019f0ee2af1949249dcd79c3cae3b9e (not referencing this ticket but only the commit in master and with a missing letter, ouch).

Still needs to be backported to 1.8

#6 - 2012-01-12 01:13 AM - Sandro Santilli

- Assignee set to Tim Sutton

Tim: I've been told you're the one taking care of backports to 1.8 :)

#7 - 2012-04-16 06:32 AM - Paolo Cavallini

- Target version changed from Version 1.7.4 to Version 1.8.0

#8 - 2012-09-04 11:54 AM - Paolo Cavallini

- Target version changed from Version 1.8.0 to Version 2.0.0

#9 - 2012-09-11 02:22 PM - Jürgen Fischer

- Status changed from Reopened to Closed

no backports to 1.8 - closing.

Files

| | | | |
|---|---------|------------|----------------|
| qgis_list_geometry_columns.txt | 5.96 KB | 2012-01-11 | Kirill Mueller |
| 0001-improve-query-that-lists-all-PostGIS-columns-fix-478.patch | 1.52 KB | 2012-01-11 | Kirill Mueller |