QGIS Application - Bug report #17031
getFeatures (by Fids) extremely slow on large PostgreSQL table WITH PK (OK if NO PK)

2017-08-17 05:20 PM - Vincent Dionne

Status:	Closed		
Priority:	High		
Assignee:			
Category:	Data Provider/PostGIS		
Affected QGIS version:	2.18.3	**Regression?:**	Yes
Operating System:		**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	not reproducable
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	24930

Description

Hello, here is the case:

I have a PostGis Table containing 5 millions GPS point. I want to selected some of them and then call getFeatures in Python.

The table has a standard serial PK.

The call to loLayer.getFeatures(loRequest) is extremely slow. First I did a loRequest.setFilterFids(loLayer.selectedFeaturesIds()) to get the selection.

Removing the PK on the table solves the problem => super fast.

After some research, I found that the Postgre Data Provider use CTID in its WHERE clause (FAST) when no PK is found (disabling editing features, wich ,in my case is okay). And in all othe cases, the fids are cast into strings in the where clause... (SLOW - the index is not use i guess)

I understand it is related with the fix of #14262.

At last there is a huge performance regression when working with large tables. At some point, when no editing is needed (ie select statement), the where clause should be working with a integer PK.

I hope I was clear enough describing the problem... thanks.

History

#1 - 2017-08-17 05:24 PM - Giovanni Manghi

- Status changed from Open to Feedback

- Priority changed from Normal to High

Was ok on a previous LTR like 2.8 or 2.14?

#2 - 2017-08-17 05:59 PM - Vincent Dionne

Tested in 2.14 and it is OK

Here is the postgreSQL log showing the WHERE clause in 2.14 : (exec time : ~ 2 seconds)

DECLARE qgis_4 BINARY CURSOR FOR SELECT

...
FROM "scad"."20170322_ibus" WHERE "id" IN
(3453850,3453851,5800408,3453852,5800409,5800410,5800411,1121748,1121749,1121750,1121751,1121 ...

in 2.18: (exec time : i don't know , I stopped it after 15 minutes)

DECLARE qgis_5 BINARY CURSOR FOR SELECT
...
FROM "scad"."20170322_ibus" WHERE ("id"::text='5174976' OR "id"::text='5064624' OR "id"::text='5174977' OR "id"::text='5064625' OR "id"::text='5174978' OR "id"::text='5174979' OR "id"::text='5174980' OR "id"::text='5174981' OR "id"::text='5174982' OR "id"::text='5174983' OR "id"::text='5174984' OR "id"::text='5174985' ...

#3 - 2017-08-17 10:14 PM - Giovanni Manghi

- Status changed from Feedback to Open

#4 - 2017-08-24 09:20 AM - Jürgen Fischer

- Description updated

#5 - 2017-08-24 09:44 AM - Jürgen Fischer

What type does your primary key have?

#6 - 2017-08-24 11:31 AM - Giovanni Manghi

- Status changed from Open to Feedback

#7 - 2017-09-08 01:35 PM - Jürgen Fischer

- Assignee deleted (Jürgen Fischer)

#8 - 2017-09-25 04:19 PM - Giovanni Manghi

Please leave further feedback.

#9 - 2018-11-11 12:57 AM - Jürgen Fischer

- Resolution set to not reproducable

- Status changed from Feedback to Closed

closed for the lack of feedback.