QGIS Application - Bug report #1718
Incorrect unique key interpretation on PostGIS view - view doesn't load as expected

2009-05-29 02:23 PM - Mike Taves

Status:	Closed		
Priority:	Low		
Assignee:			
Category:	Data Provider		
Affected QGIS version:	master	**Regression?:**	No
Operating System:	All	**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	fixed
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	11778

Description

I have three tables, each with zero-to-many relations to connect from spot (providing location) -> sample -> analysis. I would like to show the analyses at each point using a view. But when I try to add the view to QGIS, I get a message with a error message lecture about unique keys, and I cannot add the view.

This bug can be reproduced using this example schema and data:

Schema

```
CREATE TABLE spot
(
  gid serial PRIMARY KEY NOT NULL,
  geometry geometry,
  id character varying(50) UNIQUE NOT NULL
);

CREATE TABLE sample
(
  sid serial PRIMARY KEY NOT NULL,
  identifier character varying(50),
  spot_id character varying(50),
  CONSTRAINT sample_spot_id_fkey FOREIGN KEY (spot_id)
      REFERENCES spot (id) MATCH SIMPLE
      ON UPDATE CASCADE ON DELETE NO ACTION
);

CREATE TABLE analysis
(
  sid serial PRIMARY KEY NOT NULL,
  sample_sid integer,
  parameter character varying,
  result real,
  CONSTRAINT analysis_sample_sid FOREIGN KEY (sample_sid)
      REFERENCES sample (sid) MATCH SIMPLE
      ON UPDATE NO ACTION ON DELETE CASCADE
);

CREATE OR REPLACE VIEW spot_analysis AS
 SELECT ana.sid AS analysis_sid, spt.geometry, spt.id AS spot_id, smp.sid AS sample_sid, ana.parameter, ana.result
   FROM spot spt
   JOIN sample smp ON spt.id::text = smp.spot_id::text
   JOIN analysis ana ON smp.sid = ana.sample_sid;
```

Data

```
INSERT INTO geometry_columns(f_table_catalog, f_table_schema, f_table_name, f_geometry_column, coord_dimension, srid,
"type")
```

```
    VALUES (_,_,'spot','geometry',2,-1,'POINT');

INSERT INTO geometry_columns(f_table_catalog, f_table_schema, f_table_name, f_geometry_column, coord_dimension, srid,
"type")
    VALUES (_,_,'spot_analysis','geometry',2,-1,'POINT');

INSERT INTO spot (geometry, id) VALUES ('0101000000000000000000F03F000000000000F03F', 'spot a');
INSERT INTO spot (geometry, id) VALUES ('0101000000000000000000F03F0000000000000040', 'spot b');

INSERT INTO sample (identifier, spot_id) VALUES ('samp 1', 'spot a');
INSERT INTO sample (identifier, spot_id) VALUES ('samp 2', 'spot b');
INSERT INTO sample (identifier, spot_id) VALUES ('samp 3', 'spot a');
INSERT INTO sample (identifier, spot_id) VALUES ('samp 4', 'spot b');

INSERT INTO analysis (sample_sid, parameter, result) VALUES (1, 'foo', 3.4);
INSERT INTO analysis (sample_sid, parameter, result) VALUES (1, 'bla', 4.1);
INSERT INTO analysis (sample_sid, parameter, result) VALUES (2, 'foo', 3.0);
INSERT INTO analysis (sample_sid, parameter, result) VALUES (2, 'lol', 54.2);
INSERT INTO analysis (sample_sid, parameter, result) VALUES (3, 'lol', 65.2);
INSERT INTO analysis (sample_sid, parameter, result) VALUES (3, 'foo', 2.0);
```

Description of behaviour

Here is the full text of the error message shown after trying to add the [[PostGIS]] vector "spot_analysis":

No suitable key column in view

- The view 'public.spot_analysis' has no column suitable for use as a unique key.
- Qgis requires that the view has a column that can be used as a unique key. Such a column should be derived from a table column of type int4 and be a primary key, have a unique constraint on it, or be a [[PostgreSQL]] oid column. To improve performance the column should also be indexed.
- The view you selected has the following columns, none of which satisfy the above conditions:
- 'geometry' derives from 'public.spot.geometry' and is not suitable (type is geometry) and does not have a suitable constraint)
- 'id' derives from 'public.spot.id' and is not suitable (type is varchar) and has a suitable constraint)
- 'parameter' derives from 'public.analysis.parameter' and is not suitable (type is varchar) and does not have a suitable constraint)
- 'result' derives from 'public.analysis.result' and is not suitable (type is float4) and does not have a suitable constraint)
- 'sample_sid' derives from 'public.analysis.sample_sid' and is not suitable (type is int4) and does not have a suitable constraint)
- 'sid' derives from 'public.analysis.sid' and is suitable.
- 'spot_id' derives from 'public.sample.spot_id' and is not suitable (type is varchar) and does not have a suitable constraint)
- Note: 'sid' initially appeared suitable but does not contain unique data, so is not suitable.

This analysis is a bit off. It references column names used from the original tables used to construct the view. In this case, I have column names "sid" in sample and analysis. I routinely use similar views that stop at the sample join, a case where there is only one column named "sid" in the mix. This problem appears when several "sid" columns exist in the relations.
The analysis_sid key in the view is unique, and will always be unique given the constraints and use of joins. Here is what the data look like (all columns except geometry):

	analysis_sid		spot_id		sample_sid		parameter		result	
	1		spot a		1		foo		3.4	
	2		spot a		1		bla		4.1	
	3		spot b		2		foo		3	
	4		spot b		2		lol		54.2	
	5		spot a		3		lol		65.2	
	6		spot a		3		foo		2	

I can reproduce the behaviour with QGIS 1.0.2 and 1.2.0 via OSGeo4W. I've tried this on different [[PostGIS]] servers (versions/platforms), so I'm pretty confident [[PostGIS]] has nothing to do with it.

This may be related to #1535

History

#1 - 2009-05-29 02:51 PM - Mike Taves

I didn't catch this in the error message above until now (emphasis mine):

'sid' derives from 'public.analysis.sid' **and is suitable**.

This is a correct interpretation, but somehow this logic is not put to use and the error message is shown and the GIS view is not.

#2 - 2009-05-30 04:27 AM - Jürgen Fischer

Replying to [comment:1 mwtoews]:

> *I didn't catch this in the error message above until now (emphasis mine):*
> *'sid' derives from 'public.analysis.sid'* ***and is suitable****.*
>
> *This is a correct interpretation, but somehow this logic is not put to use and the error message is shown and the GIS view is not.*

#3 - 2009-05-30 11:00 AM - Mike Taves

Replying to [comment:2 jef]:

Thanks jef, that's useful info. This means a simple workaround is to avoid table aliases in the PG view, e.g.:
<pre>
CREATE OR REPLACE VIEW spot_analysis AS
 SELECT analysis.sid AS analysis_sid, spot.geometry, spot.id AS spot_id, sample.sid AS sample_sid, analysis.parameter, analysis.result
   FROM spot
   JOIN sample ON spot.id::text = sample.spot_id::text
   JOIN analysis ON sample.sid = analysis.sample_sid;
</pre>
works fine in both 1.0.2 and 1.2.0

#4 - 2009-07-18 09:31 AM - Giovanni Manghi

Can this be considered a solution to the problem? Should the ticket be left open, or can be closed?

#5 - 2009-07-18 10:37 AM - Paolo Cavallini

It seems that the problem lies in using both table aliases AND column aliases; if you use only one of the two, the problem disappears.

```
CREATE OR REPLACE VIEW spot_analysis AS
 SELECT ana.sid, spt.geometry, spt.id AS spot_id, smp.sid AS sample_sid, ana.parameter, ana.result
   FROM spot spt
   JOIN sample smp ON spt.id::text = smp.spot_id::text
   JOIN analysis ana ON smp.sid = ana.sample_sid;
```

How do other clients behave in such a case?

Aha, more info: the problem seems to depend on aliases in primary keys. With OIDs as primary key the problem seems to disappear.
Probably #1417 is a duplicate of this.

Thanks Emilia Venturato for tracking it down.

#6 - 2009-07-27 03:36 PM - Giovanni Manghi

Replying to [comment:5 pcav]:

> *How do other clients behave in such a case?*

Well... I made a couple of tests with uDIG and gvSIG which are both programs I'm not really used to.

uDIG shows correctly the table of attributes of the view (we are speaking about the one in the description of this ticket) and the points, and so does gvSIG (in this case not before having added the view in the geometry_columns table).

#7 - 2011-12-16 01:58 PM - Giovanni Manghi

- Target version changed from Version 1.7.0 to Version 1.7.4

#8 - 2012-04-16 06:22 AM - Paolo Cavallini

- Crashes QGIS or corrupts data set to No
- Target version changed from Version 1.7.4 to Version 1.8.0
- Affected QGIS version set to master

#9 - 2012-04-18 01:13 AM - Jürgen Fischer

- Pull Request or Patch supplied set to No
- Status changed from Open to Closed
- Assignee deleted (nobody -)
- Resolution set to fixed

QGIS now requires the user to select the unique column.