

QGIS Application - Bug report #15698

[Processing] Import Vector into PostGIS database doesn't work with file geodatabase

2016-10-13 03:17 AM - Jérôme Guélat

Status:	Closed	
Priority:	Low	
Assignee:	Sandro Santilli	
Category:	Data Provider/OGR	
Affected QGIS version:	2.16.3	Regression?: No
Operating System:		Easy fix?: No
Pull Request or Patch supplied:		Resolution: fixed/implemented
Crashes QGIS or corrupts data:		Copied to github as #: 23621
Description		
<p>The Processing tool "Import Vector into PostGIS database" (in GDAL/OGR) doesn't work if the vector layer comes from a file geodatabase. This is a regression since the tool is working with QGIS 2.14.</p> <p>Steps to reproduce:</p> <ol style="list-style-type: none">1. Open a vector layer stored in a file geodatabase (I'm using the ESRI FileGDB driver)2. Use this layer in the "Import Vector into PostGIS database" tool3. You will see that Processing generates a wrong ogr2ogr console call, the geodatabase name is used instead of the dataset name		
Related issues:		
Related to QGIS Application - Bug report # 16135: Selection by layerid instead of layername		Closed 2017-01-30

Associated revisions

Revision 1b1b2388 - 2016-10-15 09:11 AM - Sandro Santilli

Fix extraction of ogr LayerName from multi-layer dataset URIs

Adds supports for "layerid" when present.

Drop special handling for "table=" portions found in URI, making the code more generic.

Includes testcase.

Fixes #15698 - import geodatabase to postgis via processing

Revision 979d6197 - 2016-10-15 09:12 AM - Sandro Santilli

Fix extraction of ogr LayerName from multi-layer dataset URIs

Adds supports for "layerid" when present.

Drop special handling for "table=" portions found in URI, making the code more generic.

Includes testcase.

Fixes #15698 - import geodatabase to postgis via processing

Revision 6c536418 - 2016-10-15 09:13 AM - Sandro Santilli

Fix extraction of ogr LayerName from multi-layer dataset URIs

Adds supports for "layerid" when present.

Drop special handling for "table=" portions found in URI,
making the code more generic.

Includes testcase.

Fixes #15698 - import geodatabase to postgis via processing

Revision c4222687 - 2016-10-15 09:31 AM - Sandro Santilli

Fix extraction of ogr LayerName from multi-layer dataset URIs

Adds supports for "layerid" when present.

Drop special handling for "table=" portions found in URI,
making the code more generic.

Includes testcase.

Fixes #15698 - import geodatabase to postgis via processing

Revision 52a0082a - 2016-10-17 11:55 AM - Sandro Santilli

Fix extraction of ogr LayerName from database dataset URIs

See <https://github.com/qgis/QGIS/commit/6c5364186dd8d45ac51e5bd1a72c6a542f032cb1#commitcomment-19439676>

Includes testcase.

REF #15698

Revision 8843de88 - 2016-10-18 11:05 AM - Sandro Santilli

Fix ogrLayerName handling of PostgreSQL dataset URIs

Also document "uri" parameter semantic, and add more tests.

See for background

<https://lists.osgeo.org/pipermail/qgis-developer/2016-October/045311.html>

REF #15698

History

#1 - 2016-10-13 08:25 AM - Sandro Santilli

- Assignee set to Sandro Santilli

- Target version set to Version 2.18

Could you please attach a minimal geodatabase to reproduce this ?

#2 - 2016-10-13 11:29 PM - Jérôme Guélat

- File *vector.gdb.zip* added

Of course! Here's one...

#3 - 2016-10-14 12:34 AM - Jérôme Guélat

- File *vector2.gdb.zip* added

More details: apparently QGIS 2.14 is doing it correctly only if at least 2 datasets are stored in the file geodatabase. If there's a single dataset (like in the geodatabase I posted earlier), then we also get a wrong ogr2ogr console call.

I've attached another minimal geodatabase with 2 datasets...

Thanks for your help!

#4 - 2016-10-14 01:28 AM - Sandro Santilli

Sorry for being so lame, but could you please explain step "1. Open a vector layer stored in a file geodatabase (I'm using the ESRI FileGDB driver)" ? How do I do that ?

#5 - 2016-10-14 01:33 AM - Sandro Santilli

Figured, I could open it (<http://gis.stackexchange.com/questions/26285/file-geodatabase-gdb-support-in-qgis>) -- but I only have the OpenFileGDB driver.

#6 - 2016-10-14 01:38 AM - Sandro Santilli

So even with OpenFileGDB driver I confirm the issue, but the issue for me happens with both input datasets. This is with current master_2 branch (2.17)

#7 - 2016-10-14 01:41 AM - Sandro Santilli

- Status changed from *Open* to *In Progress*

Also confirmed with 2.16.3, while 2.14.7 is fine

#8 - 2016-10-14 02:45 AM - Sandro Santilli

Actually, I cannot confirm 2.14 is fine.

Following the code it seems that the "layername" is extracted from the input layer uri, which for a geodatabase can look like this:

```
/home/strk/lab/qgis/bugs/b15698/vector2.gdb|layerid=1 (I get this for vector2 testpoint)
/home/strk/lab/qgis/bugs/b15698/vector2.gdb (I get this for vector2 testpolygon)
/home/strk/lab/qgis/bugs/b15698/vector.gdb (I get this for vector testpolygon MultiPolygon)
```

Either the URI is wrong or just not enough to be used to reference the content.

In none of the above cases the ogr2ogr call gets either "vector2" or "vector" for the dataset name (also with 2.14.7 here)

#9 - 2016-10-14 02:51 AM - Sandro Santilli

- Category set to Data Provider/OGR

This boils down to the "Layer source" being incomplete for the loaded layers

(double click on the loaded layers to verify, check General.Layer_source, it's missing the dataset name)

#10 - 2016-10-14 03:13 AM - Jérôme Guélat

Thanks for having a look at this problem... Here's what I get (using Windows, btw.):

2.14.7: layer source complete if I use the data in the "vector2" geodatabase

2.14.7: layer source incomplete if I use the data in the "vector" geodatabase

2.16.3: layer source incomplete for both geodatabases

Apparently the OGR data provider in 2.14.7 behaves differently if there's more than one dataset in a file geodatabase.

#11 - 2016-10-14 03:14 AM - Jürgen Fischer

Sandro Santilli wrote:

*This boils down to the "Layer source" being incomplete for the loaded layers
(double click on the loaded layers to verify, check General.Layer_source, it's missing the dataset name)*

does the datasource only have one layer? then the layer name/id is optional.

#12 - 2016-10-14 04:37 AM - Sandro Santilli

vector2 has 2 layers, loading both results in one having 'layerid=1' and the other not having any 'layerid' in the DatasourceUri (I guess it would default to layerid=0).

#13 - 2016-10-14 04:44 AM - Sandro Santilli

Jérôme Guélat wrote:

*Thanks for having a look at this problem... Here's what I get (using Windows, btw.):
2.14.7: layer source complete if I use the data in the "vector2" geodatabase*

For **both** layers in that file ?

Which GDAL/OGR version ?

Also, can you show the complete layer source ?

And, is the ogr2ogr commandline correct when you import the layer with the complete "layer source" ?

| Apparently the OGR data provider in 2.14.7 behaves differently if there's more than one dataset in a file geodatabase.

Do you know the exact commit you're using of 2.14.7 ?

#14 - 2016-10-14 05:13 AM - Sandro Santilli

Ok now I see the 2.14 thing.

I was confused by the fact that I was opening a project saved with 2.16.

When you open a geodatabase with 2.14, the "Layer source" (DatasourceURI) does indeed contain a "|layername=xxx" suffix for the multi-layer input, and no suffix for the single-layer input.

In 2.16 and later, instead, you get "|layerid=1" for datasets **other_than_the_first_one** in any geodatabase.

Having "layername" in the URI is great as you can get to the dataset name from the URI itself. I'm not sure why the Ogr provider changed that as the new convention seems less useful.

In any case, fixing the provider isn't enough as it would still leave us with datasets that may still have the URI encoded in the bogus way. Unless we make the provider re-write the URI right after reading (recommended).

#15 - 2016-10-14 05:17 AM - Jérôme Guélat

*Results for 2.14.7 with GDAL/OGR 2.1.1 (installed with osgeo4w) using the following commit:ed8807c

Layer source for the *vector* geodatabase (source is incomplete and ogr2ogr command line is wrong):

polygon layer: C:\\Users\\jgu\\Desktop\\vector.gdb

Layer sources for the *vector2* geodatabase (both layers have complete layer sources, ogr2ogr command lines are both correct):

point layer: C:\\Users\\jgu\\Desktop\\vector2.gdb|layername=testpoint

polygon layer: C:\\Users\\jgu\\Desktop\\vector2.gdb|layername=testpolygon

Results for 2.16.3 with GDAL/OGR 2.1.1 (installed with osgeo4w)

Layer source for the *vector* geodatabase (source is incomplete and ogr2ogr command line is wrong):

polygon layer: C:\\Users\\jgu\\Desktop\\vector.gdb

Layer sources for the *vector2* geodatabase (sources are incomplete, ogr2ogr command lines are both wrong):

point layer: C:\\Users\\jgu\\Desktop\\vector2.gdb|layerid=1

polygon layer: C:\\Users\\jgu\\Desktop\\vector2.gdb

#16 - 2016-10-14 05:27 AM - Sandro Santilli

NOTE: the "layername=" to "layerid=" change was introduced in 2016-06-30, according to ChangeLog, to fix #15168

#17 - 2016-10-14 05:33 AM - Sandro Santilli

Taking it back, it doesn't look like #15168 has to do with this. Must have been a change in the geodatabase loader, to use layerid instead of layername.

OGR provider supports both and never rewrites the uri. Even allows having both in the uri, giving precedence to layername if present.

#18 - 2016-10-14 06:12 AM - Sandro Santilli

Pull request ready for tests: <https://github.com/qgis/QGIS/pull/3605>

#19 - 2016-10-14 06:13 AM - Jérôme Guélat

Great! Thanks a lot!

#20 - 2016-10-14 06:34 AM - Sandro Santilli

- % Done changed from 0 to 50

Jérôme can we include the geodatabase files you provided in our testsuite ?

#21 - 2016-10-14 06:36 AM - Jérôme Guélat

Of course you can.

#22 - 2016-10-14 06:44 AM - Sandro Santilli

btw, it looks like "convert format" tool is still suffering the same bug - are you aware of a separate ticket for that ?

#23 - 2016-10-14 06:52 AM - Jérôme Guélat

Apparently most of the OGR tools in Processing are also having the same problem (I've just tried with buffer, clip, convert, etc.)

I'm not aware of a separate ticket...

#24 - 2016-10-14 01:41 PM - Sandro Santilli

- % Done changed from 50 to 80

I've just tried and with the patch "Buffer vectors" works fine.

"Clip vectors by extent" and "Convert format" seem to have a different issue, so I suggest a new ticket is filed for them (after checking each with the updated code)

#25 - 2016-10-15 12:15 AM - Sandro Santilli

- Status changed from In Progress to Closed

Fixed in changeset commit:"1b1b238836162e5088fb58670da169a6f352d9a8".

#26 - 2016-10-15 12:24 AM - Sandro Santilli

- Resolution set to fixed/implemented
- Target version changed from Version 2.18 to Version 2.16
- % Done changed from 80 to 100

Fixed for 2.16.4, 2.18.0, master

#27 - 2016-10-15 12:33 AM - Sandro Santilli

Also fixed for 2.14.8

#28 - 2016-10-28 06:53 AM - Alexander Bruy

- Status changed from Closed to Reopened
- Resolution deleted (fixed/implemented)

This fix breaks some Processing OGR algorithms (tested with buffer, but other may be affected too).

To reproduce:

1. open polys.gml from Processing test dataset
2. try to execute "Buffer vectors" from GDAL/OGR -> [OGR] Geoprocessing -> Buffer vectors
3. algorithm execution will fail with error

Warning 1: layer names ignored in combination with -sql.

ERROR 1: In ExecuteSQL(): sqlite3_prepare(SELECT ST_Buffer(geometry, 1000), * FROM 'polys2'):
no such table: polys2

Also related tests in ToolsTest.py fail if /tmp located on another partition.

#29 - 2017-01-05 01:10 AM - Giovanni Manghi

- Status changed from Reopened to Closed
- Resolution set to fixed/implemented

AlexB reverted and applied a different fix.

#30 - 2017-06-06 03:47 PM - Sandro Santilli

- Description updated
- Priority changed from Severe/Regression to Low

Were also tests reverted ? Which revision had the revert ? (Associated Revisions tab doesn't say).

A new ticket was filed which might be related to this: #16135

#31 - 2017-06-06 03:47 PM - Sandro Santilli

- Related to Bug report #16135: Selection by layerid instead layername from .gdb added

test introduced to fix this issue started to fail due different factors:

1) it fail in case data and tmp are in different filesystem (due to hard link nature).

A proposed patch here: <https://github.com/qgis/QGIS/pull/5151>

2) in <https://github.com/qgis/QGIS/pull/5144> you can find reports that fail in Travis and in different test configurations

Files			
vector.gdb.zip	18.9 KB	2016-10-13	Jérôme Guélat
vector2.gdb.zip	20.9 KB	2016-10-13	Jérôme Guélat