QGIS Application - Bug report #19938
GDAL/OGR vector geoprocessing algorithms not working with GPKG, SQLite, FileGDB, etc inputs

2018-09-25 01:27 PM - Andrea Giudiceandrea

Status:	Closed		
Priority:	High		
Assignee:	Nyall Dawson		
Category:	Processing/OGR		
Affected QGIS version:	3.3(master)	**Regression?:**	Yes
Operating System:		**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	27760

Description

See also #19938-4 for a better understanding of the issue.

Original description:

The GDAL/OGR vector geoprocessing algorithms do not work with GPKG, SQLite, FileGDB, (probably other formats), as input vector layer.

See the following error for GDAL "Buffer vector" alg. as example, with a GPKG input vector layer:

```
Processing algorithm…
Algorithm 'Buffer vectors' starting…
Input parameters:
{ 'DISSOLVE' : False, 'DISTANCE' : 10, 'EXPLODE_COLLECTIONS' : False, 'FIELD' : None, 'GEOMETRY' : 'geometry',
*'INPUT' : 'C:\\test\\demo\\test.gpkg|layername=points_layer'*, 'OPTIONS' : '', 'OUTPUT' :
'C:/Users/Andrea/AppData/Local/Temp/processing_a2e42c246f5c42e7b34ff0883d38d602/d685014df90f4511a2d43058c7d3cc2e/OUTPUT
e/OUTPUT.shp' }

GDAL command:
ogr2ogr
C:/Users/Andrea/AppData/Local/Temp/processing_a2e42c246f5c42e7b34ff0883d38d602/d685014df90f4511a2d43058c7d3cc2e/OUTPUT.
e/OUTPUT.shp
*C:/Users/Andrea/AppData/Local/Temp/processing_a2e42c246f5c42e7b34ff0883d38d602/2e08c2dab752494ebc42763c1f331c53/INPUT.s
53/INPUT.shp* -dialect sqlite -sql "SELECT ST_Buffer(geometry, 10.0) AS geometry, fid,field1 FROM *'points_layer'*" -f "ESRI
Shapefile"
GDAL command output:
ERROR 1: In ExecuteSQL(): sqlite3_prepare_v2(SELECT ST_Buffer(geometry, 10.0) AS geometry, fid,field1 FROM
'points_layer'):

*no such table: points_layer*
```

It seems that before the GDAL/OGR vector geoprocessing algs run their tasks, **every input layer contained in GPKG, SQLite, FileGDB, is converted in a temporary INPUT.SHP** shapefile which is used as new input layer for the ogr2ogr command.

This happens because in QgsProcessingUtils::convertToCompatibleFormat the layer suffix (QFileInfo.suffix()) for those layer types is something like:

```
'gpkg|layername=points_layer'
```

and this string does not match with any of the QgsVectorFileWriter.supportedFormatExtensions(), due to the presence of "|layername=points_layer", leading to always convert (translate) those types of layers in the temporary shapefile INPUT.SHP.
I don't know if this is a normal behaviour or a bug, but I think they can be normally used by ogr2ogr as input layer without conversion, provided we specify the right geometry column name.

If this is a bug, I think it could be fixed in QgsProcessingUtils::convertToCompatibleFormat just changing

```
QFileInfo fi( vl->source() );
requiresTranslation = !compatibleFormats.contains( fi.suffix(), Qt::CaseInsensitive );
```

with

```
QFileInfo fi( vl->source() );
requiresTranslation = !compatibleFormats.contains( fi.suffix().split('|')[0], Qt::CaseInsensitive );
```

Anyway, algorithms fail to properly build the sql statement: in fact if INPUT.SHP must be used as new input layer, then the table name after FROM should be 'INPUT' and not the original layer name 'points_layer'.

This problem could be fixed in
https://github.com/qgis/QGIS/blob/master/python/plugins/processing/algs/gdal/GdalAlgorithm.py#L104-L109

```
# parameter is a vector layer, with OGR data provider
# so extract selection if required
ogr_data_path = self.parameterAsCompatibleSourceLayerPath(parameters, parameter_name, context,
                                                          QgsVectorFileWriter.supportedFormatExtensions(),
                                                          feedback=feedback)
ogr_layer_name = GdalUtils.ogrLayerName(input_layer.dataProvider().dataSourceUri())
```

where

```
ogr_layer_name = GdalUtils.ogrLayerName(input_layer.dataProvider().dataSourceUri())
```

should be

```
ogr_layer_name = GdalUtils.ogrLayerName(ogr_data_path)
```

This works for me, changing my local copy of GdalAlgorithm.py, but I cannot say if this is the right way to solve the problem or if it's possible to completely avoid the needing of convert GPKG, SQLite, FileGDB layers, and probably other formats, to a temporary shapefile.

Moreover, it is not possible to add the output layer to an exiting GPKG, SQLite... container without completely overwrite it.

Furthermore, it is not possible to properly select a layer ("Select file" in the algorithm window) from GPKG, SQLite, FileGDB containers in the case it is not already present in the map.

Related issues:

Related to QGIS Application - Bug report # 19945: ogr based tools cannot outp...	**Closed**	**2018-09-25**
Related to QGIS Application - Bug report # 19946: ogr based tools do not work...	**Closed**	**2018-09-25**

Associated revisions

Revision 737ab306 - 2018-09-28 06:25 AM - Nyall Dawson

[processing][gdal] Correctly handle geopackage paths with layername argument

Fixes #19938

Revision 308d7136 - 2018-09-28 06:25 AM - Nyall Dawson

[processing] More fixes to convertToCompatibleFormat, lots of unit
tests

Fixes #19938

History

#1 - 2018-09-25 06:44 PM - Giovanni Manghi

- Status changed from Open to Feedback

- Priority changed from Normal to High

- Regression? changed from No to Yes

I have nticed this too today, while trying to import data in PostGIS using the ogr based tools. Your analysis is correct, anyway there are cases (not yet sure how/why) where the datasources like gpkg, sqlite, shapes are NOT translated to a temporary INPUT.shp, and so the tools works.

On the other hand it seems that PostGIS inputs are completely broken for such tools.

More notes to come.

#2 - 2018-09-25 08:02 PM - Giovanni Manghi

Possibly related issue #1 #19945

#3 - 2018-09-25 08:36 PM - Giovanni Manghi

Possibly related issue #2 #19946

#4 - 2018-09-25 09:15 PM - Giovanni Manghi

Processing seems to work in different ways, depending on how a datasource like a gpkg, SL or shp was loaded in the project:

1) If the layer is loaded by drag and drop (from an operating system file navigator window), or using the browser or by using the datasource manager, then ogr based tools are OK (minus https://issues.qgis.org/issues/19945) as the input of the ogr2ogr calls is a path that points to the datasource, example:

```
GDAL command:
ogr2ogr /home/giovanni/Desktop/diss2.gpkg /home/giovanni/Desktop/teste.shp -dialect sqlite -sql "SELECT ST_Union(geometry) AS geometry, region FROM teste GROUP BY region" -f "GPKG"
GDAL command output:
```

Execution completed in 47.39 seconds
Results:{'OUTPUT': <QgsProcessingOutputLayerDefinition {'sink':/home/giovanni/Desktop/diss2.gpkg, 'createOptions': {'fileEncoding': 'System'}}>}@

2) If the input layer was generated and added from within QGIS itself (i.e. a "save as" operation) the the input for ogr based tools seems to be a temporary shapefile as described in this ticket, with all the consequences here described.

#5 - 2018-09-25 09:16 PM - Giovanni Manghi

- Description updated

- Status changed from Feedback to Open

- Subject changed from GDAL/OGR vector geoprocessing algorithms not working with GPKG, SQLite, FileGDB, ... to GDAL/OGR vector geoprocessing algorithms not working with GPKG, SQLite, FileGDB, etc inputs

#6 - 2018-09-25 10:56 PM - Andrea Giudiceandrea

Giovanni Manghi wrote:

> *there are cases (not yet sure how/why) where the datasources like gpkg, sqlite, shapes are NOT translated to a temporary INPUT.shp, and so the tools works.*

During my tests, this only happens when there is only 1 layer in a gpkg/sqlite container and it is added to the map through the Data Source Manager / Vector or dragging and dropping it in the map from outside.

In fact, in this case (e.g. test_polygons.gpkg with only 1 layer named 'polygons') the layer is added with a source like 'test_polygons.gpkg' (instead of a more correct 'test_polygons.gpkg|layername=polygons') and so the suffix string will be 'gpkg' which is in QgsVectorFileWriter.supportedFormatExtensions() thus QgsProcessingUtils::convertToCompatibleFormat will not covert it in a temporary shapefile.

When the layer in a single layer gpkg/sqlite container is added to the map through the Browser panel instead, its source will be 'test_polygons.gpkg|layername=polygons'.

#7 - 2018-09-25 10:58 PM - Nyall Dawson

- Status changed from Open to In Progress

- Assignee set to Nyall Dawson

#8 - 2018-09-26 02:11 AM - Nyall Dawson

See https://github.com/qgis/QGIS/pull/8028 -- testing would be appreciated!

#9 - 2018-09-26 11:28 PM - Jürgen Fischer

- Description updated

#10 - 2018-09-27 02:39 AM - Andrea Giudiceandrea

Nyall Dawson wrote:

> *See https://github.com/qgis/QGIS/pull/8028 -- testing would be appreciated!*

Unfortunately, I cannot test it before it is merged.

#11 - 2018-09-27 08:44 AM - Jürgen Fischer

- Related to Bug report #19945: ogr based tools cannot output anymore to temp files added

#12 - 2018-09-27 08:44 AM - Jürgen Fischer

- Related to Bug report #19946: ogr based tools do not work anymore with PostGIS inputs (possibly also other rdbms datasources) added

#13 - 2018-09-27 09:33 AM - Giovanni Manghi

Andrea Giudiceandrea wrote:

> *Nyall Dawson wrote:*
>
> > *See https://github.com/qgis/QGIS/pull/8028 -- testing would be appreciated!*
>
> *Unfortunately, I cannot test it before it is merged.*

just extract the processing folder from here https://github.com/nyalldawson/QGIS/archive/gdal_layers.zip and drop it in the QGIS plugins folder. Then restart QGIS and test it, after it remove the processing folder inside the plugins one, otherwise this will mask processing as it is shipped by qgis installer.

#14 - 2018-09-27 10:23 AM - Nyall Dawson

That process won't work here - there's core changes too

#15 - 2018-09-28 02:53 AM - Andrea Giudiceandrea

Andrea Giudiceandrea wrote:

> *Moreover, it is not possible to add the output layer to an exiting GPKG, SQLite... container without completely overwrite it.*
>
> *Furthermore, it is not possible to properly select a layer ("Select file" in the algorithm window) from GPKG, SQLite, FileGDB containers in the case it is not already present in the map.*

Nyall, will your PR also address those two additional problems or should I issue as many new separate bug reports?

#16 - 2018-09-28 06:25 AM - Nyall Dawson

- % Done changed from 0 to 100

- Status changed from In Progress to Closed

Applied in changeset commit:qgis|737ab3069a6ce279a3ddb7f9df38baed481c10f5.

#17 - 2018-10-05 02:50 PM - Andrea Giudiceandrea

Andrea Giudiceandrea wrote:

> *Andrea Giudiceandrea wrote:*

> > *Moreover, it is not possible to add the output layer to an exiting GPKG, SQLite... container without completely overwrite it.*
> >
> > *Furthermore, it is not possible to properly select a layer ("Select file" in the algorithm window) from GPKG, SQLite, FileGDB containers in the case it is not already present in the map.*
>
> *Nyall, will your PR also address those two additional problems or should I issue as many new separate bug reports?*

See
#20026 "Processing algorithms are not able to add the output layer to an exiting GPKG or SQLite container without completely overwrite it"
#20027 "In processing algorithms is not possible to properly select an input layer not in the map from GPKG, SQLite, FileGDB containers"

Andrea Giudiceandrea wrote:

> *in this case (e.g. test_polygons.gpkg with only 1 layer named 'polygons') the layer is added with a source like 'test_polygons.gpkg' (instead of a more correct 'test_polygons.gpkg|layername=polygons')*

See #20031 "The only layer contained in a singlelayer GPKG / SQLite / FileGDB is incorrectly added to the map without "|layername=" in some circumstances"