QGIS Application - Bug report #18402
QGIS 3 crashed - Trying to add data from a large (81Gb) GeoPackage on USGS NHD flattened data via the 'catalog'...

2018-03-09 11:15 PM - Greg Cocks

Status:	Closed		
Priority:	Normal		
Assignee:	Even Rouault		
Category:	Data Provider/OGR		
Affected QGIS version:	3.0.0	**Regression?:**	No
Operating System:	Windows 7	**Easy fix?:**	No
Pull Request or Patch supplied:	No	**Resolution:**	
Crashes QGIS or corrupts data:	Yes	**Copied to github as #:**	26292

Description

User Feedback

Trying to add data from a large (81Gb) GeoPackage on USGS NHD flattened data via the 'catalog'... can't 'see into' the geopackage

Note that I AM able to see into the corresponding FGDB

Thanks...

---
h2. Report Details

Crash ID: 042ffec7f8c64f2dd8646773af79b1430ec6537e

Stack Trace

```
GDALDatasetGetLayerCount :
std::basic_string<char,std::char_traits<char>,std::allocator<char> >::insert :
CPLStringList::size :
wildcards :
handlesDirectoryPath :
QgsDataItem::runCreateChildren :
QgsDataItem::runCreateChildren :
QgsDataItem::rowCount :
QThreadPoolPrivate::reset :
QThread::start :
BaseThreadInitThunk :
RtlUserThreadStart :
```

QGIS Info
QGIS Version: 3.0.0-Girona
QGIS code revision: commit:001c80b0c3
Compiled against Qt: 5.9.2
Running against Qt: 5.9.2
Compiled against GDAL: 2.2.3
Running against GDAL: 2.2.3

System Info
CPU Type: x86_64
Kernel Type: winnt
Kernel Version: 6.1.7601

Associated revisions

Revision 4f531359 - 2018-06-04 02:23 PM - Even Rouault

[OGR provider] Performance improvements on huge GeoPackage database (fixes #18402)

- Introduce an approximate feature count for GeoPackage in case there are
  at least 100,000 rows in a table.
- Add a super fast implementation of GetExtent() for GeoPackage when there
  is an RTree
- Do not require feature count when enumerating layers from the browser

Revision 6160e066 - 2018-06-04 09:34 PM - Even Rouault

Merge pull request #7167 from rouault/gpkg_perf_improvements

[OGR provider] Performance improvements on huge GeoPackage database (fixes #18402)

History

#1 - 2018-03-09 11:31 PM - Greg Cocks

- File small.jpg added

UPDATE: baulking on smaller GeoPackages as well, this one is ½Gb as an example (see screen shot), 'the wheel just spins', finally opened after about 4 minutes and could see contents

Various examples at ftp://rockyftp.cr.usgs.gov/ngtoc/hydro/outgoing/GeoPackage_hydrography_examples/ if you wanted to try

#2 - 2018-03-09 11:45 PM - Greg Cocks

And then crashed again with smaller GeoPackage layers added, just FYI

User Feedback

Report Details

Crash ID: 042ffec7f8c64f2dd8646773af79b1430ec6537e

Stack Trace

```
GDALDatasetGetLayerCount :
std::basic_string<char,std::char_traits<char>,std::allocator<char> >::insert :
CPLStringList::size :
wildcards :
handlesDirectoryPath :
QgsDataItem::runCreateChildren :
QgsDataItem::runCreateChildren :
QgsDataItem::rowCount :
QThreadPoolPrivate::reset :
QThread::start :
BaseThreadInitThunk :
RtlUserThreadStart :
```

QGIS Info
QGIS Version: 3.0.0-Girona
QGIS code revision: commit:001c80b0c3
Compiled against Qt: 5.9.2
Running against Qt: 5.9.2
Compiled against GDAL: 2.2.3

Running against GDAL: 2.2.3

System Info
CPU Type: x86_64
Kernel Type: winnt
Kernel Version: 6.1.7601

#3 - 2018-03-11 04:45 PM - Alessandro Pasotti

- Status changed from Open to Feedback

Can you please describe the exact step-by-step procedure to reproduce this crash?
There are multiple ways to add a gpkg layer in QGIS: are you using the QGIS file browser?
Also, can you paste the URL of a file that can be used to reproduce the issue? (there are hundreds under the two folders in the URL you provided)

#4 - 2018-03-11 04:51 PM - Alessandro Pasotti

Tested on Linux and it works fine, I picked one at random: HRNHDPlus0101.gpkg

Are you by chance loading the file from a network drive?

#5 - 2018-03-11 04:57 PM - Alessandro Pasotti

Also tried on win 10 64 bits without issues, loading from a network drive also worked.

Can you reproduce it on win 10 too?

#6 - 2018-03-12 09:15 PM - Greg Cocks

- File ss_number_2.jpg added
- File ss_number_1.jpg added

Alessandro Pasotti wrote:

> *Can you please describe the exact step-by-step procedure to reproduce this crash?*
> *There are multiple ways to add a gpkg layer in QGIS: are you using the QGIS file browser?*
> *Also, can you paste the URL of a file that can be used to reproduce the issue? (there are hundreds under the two folders in the URL you provided)*

Alessandro,

Of course, sorry for any delay in replying:

- Open QGIS3
- do NOT open a new or existing project
- Using the Browser window (see attached screen shot #1), browse to the flattened geopackage (81Gb or so, here: ftp://rockyftp.cr.usgs.gov/ngtoc/hydro/outgoing/GeoPackage_hydrography_examples/whole_flattened_NHD_current/NHDPlusHR_flattened_2011103_gpkg_20171106.7z)
- click the arrow to open in the browser window for contents - the wheel spins for 20+ minutes (see attached screen shot #2) - and then QGIS crashes out with the error provided originally.

QGIS 3.x (g7.4.0) is only using a consistent 167,008lk of memory during this whole process (I have 16Gb) and << 5% of the the 'CPU'

The geoPackage is sitting locally on an external 1Tb hard drive, USB 2 connection, that has solid IO performance with other tasks.

Please let me know if you need any other information.

Regards,
GREG...

#7 - 2018-03-12 09:16 PM - Greg Cocks

Alessandro Pasotti wrote:

> *Also tried on win 10 64 bits without issues, loading from a network drive also worked.*
>
> *Can you reproduce it on win 10 too?*

I will download QGIS 3 and the large dataset at ftp://rockyftp.cr.usgs.gov/ngtoc/hydro/outgoing/GeoPackage_hydrography_examples/whole_flattened_NHD_current/NHDPlusHR_flattened_20171103_pkg_20171106.7z and try at home tonight on Windows 10

#8 - 2018-03-12 09:26 PM - Greg Cocks

Alessandro Pasotti wrote:

> *Tested on Linux and it works fine, I picked one at random: HRNHDPlus0101.gpkg*
>
> *Are you by chance loading the file from a network drive?*

Hello,

Windows 7 Enterprise and a file on a USB 2 attached 1Tb drive.

Just to confirm, there are **two** (2) errors I am seeing here:

1) as per more detailed reply above - with the large dataset ftp://rockyftp.cr.usgs.gov/ngtoc/hydro/outgoing/GeoPackage_hydrography_examples/whole_flattened_NHD_current/NHDPlusHR_flattened_20171103_gpkg_20171106.7z, the QGIS Browser never shows a preview of the GPKG contents and eventually crashes out after 15+ minutes (see items #1 and #7 in this thread.)

2) using a small example of the Geopackage, say the HRNHDPlus0101.gpkg you used, you can get a preview in the QGIS Browser, add to a project, save the project, start to navigate around - and the project / QGIS **then** crashes (see item #2 in this thread)

Sorry for any confusion.

Regards,
GREG...

#9 - 2018-03-12 10:41 PM - Greg Cocks

- File fred.qgs added
- File DxDiag.txt added
- File ss_number_3.jpg added

Hello,

Finally got it to load, FYI, after a couple of reboots, the large 80Gb GPKG (i.e., #1 and #7 in this thread)

- took 48 minutes on my system to give a preview of GPKG contents
- took 35 minutes on my system after right clicking one of the vector datasets and add to canvas (screen shot #3 attached) for it to start drawing (8,475,5572 vectors in this dataset in the gpkg)
- took about ½ Gb more of RAM once the vector dataset has loaded

I attached some specs on 'my' system in a text file

I also attached the screen shot and the simple *.qgs file

Let me know if you need anything else...

Regards,
GREG...

Regards,
GREG...

#10 - 2018-03-14 06:24 PM - Greg Cocks

- File DxDiag_xps13_windows10.txt added

Alessandro Pasotti wrote:

> *Also tried on win 10 64 bits without issues, loading from a network drive also worked.*
>
> *Can you reproduce it on win 10 too?*

Tried the 80Gb GPKG on Windows 10

About 6 minutes to open gpkg for preview, twice as fast to draw and open attribute table for 8½ million vectors

PC not as capable (see attached) - BUT data coming off internal SSD rather than an external old school HD with USB 2

#11 - 2018-06-04 01:38 PM - Even Rouault

- Category changed from Data Provider to Data Provider/OGR

#12 - 2018-06-04 02:17 PM - Even Rouault

- Assignee set to Even Rouault

#13 - 2018-06-04 09:34 PM - Even Rouault

- Status changed from Feedback to Closed

- % Done changed from 0 to 100

Applied in changeset commit:qgis|4f531359aee9e861af92394072917add8cb80ef5.

Files

small.jpg	36.2 KB	2018-03-09	Greg Cocks
ss_number_1.jpg	130 KB	2018-03-12	Greg Cocks
ss_number_2.jpg	78 KB	2018-03-12	Greg Cocks
fred.qgs	57.9 KB	2018-03-12	Greg Cocks
DxDiag.txt	5.42 KB	2018-03-12	Greg Cocks
ss_number_3.jpg	83.3 KB	2018-03-12	Greg Cocks
DxDiag_xps13_windows10.txt	83 KB	2018-03-14	Greg Cocks