

QGIS Application - Bug report #17043

Browser keeps scanning directory with gpkg file

2017-08-21 12:50 PM - Richard Duivenvoorde

Status:	Closed	
Priority:	High	
Assignee:	Alessandro Pasotti	
Category:	Browser	
Affected QGIS version:	master	Regression?: No
Operating System:	Debian Testing	Easy fix?: No
Pull Request or Patch applied:	Yes	Resolution:
Crashes QGIS or corrupts data:	No	Copied to github as #: 24942
Description		
<p>Using 2.99 / current master (20 aug 2017) running self compiled QGIS on Debian Testing, when I run QGIS from cli the following messages keep flying over my (debug) terminal:</p> <pre>src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [29ms] [thread:0x7f4fb4002620] GDALOpen error # 4 : `/home/richard/geodata/prov.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [15ms] [thread:0x7f4fb4002620] GDALOpen error # 4 : `/home/richard/geodata/world.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [141ms] [thread:0x7f4fb4001fc0] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [10ms] [thread:0x7f4fb4001fc0] GDALOpen error # 4 : `/home/richard/geodata/prov.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [17ms] [thread:0x7f4fb4001fc0] GDALOpen error # 4 : `/home/richard/geodata/world.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [148ms] [thread:0x561ed4366150] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [24ms] [thread:0x561ed4366150] GDALOpen error # 4 : `/home/richard/geodata/prov.gpkg' not recognized as a supported file format. src/providers/gdal/qgsgdaldataitems.cpp: 282: (dataItem) [18ms] [thread:0x561ed4366150] GDALOpen error # 4 : `/home/richard/geodata/world.gpkg' not recognized as a supported file format.</pre> <p>It also takes a lot of resources at that moment.</p> <p>I see this when the Browser Panel is open.</p> <p>As you see I have a directory with a couple of gpkg files in it (besides other files), and (to me) it looks like that after concluding that GDAL cannot open one or more of the (sub) layers, instead of stopping it starts all over again to see if it can open it...</p> <p>At that moment qgis eats 40% of cpu over a couple of cores on my laptop. Note that this never ends.</p>		

Associated revisions

Revision 1d4c6061 - 2017-09-12 03:00 PM - Alessandro Pasotti

[bugfix] Browser keeps scanning directory with gpkg file

Fixes #17043 by introducing a check for last directory scan timestamp and skipping the directoryChanged signal is the last scan was less than 10 seconds in the past.

Tested with > 300 gpkg files in a single directory.

History

#1 - 2017-08-21 04:40 PM - Giovanni Manghi

- *Priority changed from Normal to High*
- *Regression? changed from No to Yes*

#2 - 2017-08-21 06:48 PM - Alessandro Pasotti

- *Assignee set to Alessandro Pasotti*

I saw this in Nodebo and I thought it was fixed, would you please give me a test scenario?

The gpkg files are valid?

What are your options for the browser?

#3 - 2017-08-21 06:49 PM - Alessandro Pasotti

... also, do you have them in multiple favorites?

#4 - 2017-08-22 11:25 AM - Richard Duivenvoorde

- *File settings.png added*

Hi Alessandro,

I will sent you the two gpkg packages I have by email.

They are only available in /home/richard/geodata

And I have 1 favourite: /home/richard/geodata

It is not clearly when it happens, sometimes it does NOT and other times, it is when I open the gpkg node of one of them. I also run QGIS-master and QGIS-2.18.11 concurrently, and thought that that would be the problem.

I observe it with both versions.

I add a screenshot of my settings, right is 2.18.11 and left is master

#5 - 2017-08-22 11:27 AM - Alessandro Pasotti

with "I observe it with both versions." do you mean that this bug is not a regression that was introduced recently on master?

#6 - 2017-08-22 11:30 AM - Giovanni Manghi

- *Status changed from Open to Feedback*

#7 - 2017-08-22 12:08 PM - Richard Duivenvoorde

I mean that I can see it in both windows, either the master terminal or the 2.18 terminal.

I just compiled 2.18.12 and it still happens, as said not ALWAYS, but with some fiddling and trying it happens at a certain point...

Slightly other output?

```
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [7ms] [thread:0x55963379e9f0] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg'
not recognized as a supported file format.
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [12ms] item favourites:/home/richard/geodata state changed 1 -> 2
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [2ms] item favourites:/home/richard/geodata state changed 2 -> 1
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [9ms] [thread:0x55963379e9f0] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg'
not recognized as a supported file format.
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [12ms] [thread:0x55963379e9f0] GDALOpen error # 4 :
`/home/richard/geodata/prov.gpkg' not recognized as a supported file format.
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [6ms] item favourites:/home/richard/geodata state changed 1 -> 2
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [1ms] item favourites:/home/richard/geodata state changed 2 -> 1
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [5ms] [thread:0x55963379e9f0] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg'
not recognized as a supported file format.
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [11ms] [thread:0x55963379e9f0] GDALOpen error # 4 :
`/home/richard/geodata/prov.gpkg' not recognized as a supported file format.
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [11ms] [thread:0x55963379e9f0] GDALOpen error # 4 :
`/home/richard/geodata/world.gpkg' not recognized as a supported file format.
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [9ms] item favourites:/home/richard/geodata state changed 1 -> 2
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [67ms] item favourites:/home/richard/geodata state changed 2 -> 1
src/providers/gdal/qgsgdaldataitems.cpp: 288: (dataItem) [6ms] [thread:0x55963379e9f0] GDALOpen error # 4 : `/home/richard/geodata/nl.gpkg'
not recognized as a supported file format.
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [8ms] item favourites:/home/richard/geodata state changed 1 -> 2
src/core/qgsbrowsermodel.cpp: 418: (itemStateChanged) [2ms] item favourites:/home/richard/geodata state changed 2 -> 1
```

#8 - 2017-08-22 12:13 PM - Giovanni Manghi

Richard Duivenvoorde wrote:

I mean that I can see it in both windows, either the master terminal or the 2.18 terminal.

Hi Richard, I think that Alessandro's question is because he needs to understand of this behavior was introduced by the work he is doing on the Browser and GPKG support in QGIS 3, but it seems that after all this is also happening in QGIS 2.18.* so after all is not a regression, correct?

cheers!

#9 - 2017-08-22 12:22 PM - Richard Duivenvoorde

- File *qgis218terminal.png* added

Ok, I found a more reproducible case now:

Given you have the prov.gpkg (with prov layer in it):

- open 2.18.12 in terminal and make sure you have browser panel open
- open master in terminal and make sure you have browser panel open

No need to do anything in QGIS 2.18, just observe the terminal

Now open the layer prov in prov.gpkg, immediately in my 2.18.12 terminal you get the messages above.

AND when you unload the layer in QGIS-master, the running messages stop!

So I think(!) that it has something to do with maybe concurrently opening/scanning a gpkg file with two QGIS processes?

#10 - 2017-08-22 12:33 PM - Alessandro Pasotti

As I said, I've seen this before and I think I know why it happens:

1. when opening a GPKG sqlite creates two temporary files (wal and shm), this triggers a new re-scan of the folder, re-opening the files and re-creating the shm and wal, which triggers a new re-scan of the folder infinite loop! (an attempt to fix this was done in commit:bcb48f2bdc8f274cbc68f376489430f45e0b17d1)
2. worse: each file is opened twice with GDAL and OGR providers, that could be completely avoided in GDAL 2.x (but this is somewhere down in my TODO list)

There is no perfect solution to this issue but I can try to mitigate the effect.

BTW, I'd be surprised if this was a regression, I believe that the bug was just sitting there and wait for someone to test the gpkg.

#11 - 2017-08-22 01:19 PM - Richard Duivenvoorde

- Regression? changed from Yes to No

Hi Ale,

I agree, it NOT a regression: removing the Regression flag. I'm pretty sure I've seen this already for a long time. But always thought that it was something in my setup or so.

Is it maybe possible to add a 'filter' to the 'scan'-function? Some candidates for this filter:

- *.shm
- *.wal
- /initrd.img <- always on linux machines?
- all kind of temporary files (like the .qgs~, vim also has something like that etc)

I would also be ok, with some kind of 'empty' browser, so that you really have to add directories to be able to browse them.

OR only scan the favourites, and not the full tree?

#12 - 2017-08-22 03:54 PM - Alessandro Pasotti

No way to add filters, that was my first idea too..

The signal is fired any time something changes in the directory, my fix was to disconnect the signal while the children of the tree are populated (that's when the providers try to open the files - and where GDAL fails) and reconnecting them at the end.

And as an additional issue, if you start creating a file (for example a new shape or gpkg or whatever) the signal is fired and the provider fails because the file is still being written (but this is another issue).

BTW, let me try to reproduce it, because I thought it was fixed and I couldn't reproduce it anymore, maybe having many gpkg in the same folder triggers the error.

#13 - 2017-08-22 05:45 PM - Alessandro Pasotti

I cannot reproduce in master. Can you please confirm that on master it is fixed?

If you run simultaneously 2.8, then you might see the error in master because of the bug in 2.18 that keeps opening the files and creating the temporary shm and wal and trigger the re-scan.

#14 - 2017-08-25 02:58 PM - Richard Duivenvoorde

Tested with two master QGIS versions concurrently running, and then it seems fine!

As soon as I start a 2.18, master still starts freaking out.

So: close this one?

#15 - 2017-09-08 11:00 AM - Alessandro Pasotti

No, please leave this open, I need to further investigate if are there other conditions in which it can happen.

#16 - 2017-09-08 11:01 AM - Giovanni Manghi

- Status changed from Feedback to Open

#17 - 2017-09-08 10:17 PM - Alessandro Pasotti

- Status changed from Open to In Progress

Now I'm able to reproduce it consistently in master. I'm looking for a more complex yet robust solution.

#18 - 2017-09-11 08:40 AM - Alessandro Pasotti

- Status changed from In Progress to Feedback

- Pull Request or Patch supplied changed from No to Yes

This PR <https://github.com/qgis/QGIS/pull/5164> should definitely fix the issue (and bring better performances too)

#19 - 2017-09-12 05:20 PM - Anonymous

- Status changed from Feedback to Closed

- % Done changed from 0 to 100

Applied in changeset commit:qgis|1d4c6061a8a5512092041aafea14a8a32247d046.

Files

settings.png	264 KB	2017-08-22	Richard Duivenvoorde
qgis218terminal.png	715 KB	2017-08-22	Richard Duivenvoorde