QGIS Application - Feature request #19895
Garbage-collection is not performed after deletion of vector layer from geopackage

2018-09-19 07:39 PM - Andrea Giudiceandrea

Status:	Closed		
Priority:	Normal		
Assignee:	Alessandro Pasotti		
Category:	Browser		
Pull Request or Patch supplied:	Yes	**Resolution:**	
Easy fix?:	No	**Copied to github as #:**	27719

Description

With QGIS 3.3.0-master (d99d506e75), deleting a vector layer from a geopackage, through the "Browser" panel or the "DB Manager", is not followed by a proper garbage-collection (shrink/compact) of the geopackage.

I think a VACUUM sql command should be automatically executed after deleting a layer from a geopackage or a contextual menu option should be added in the "Browser" panel and in the "DB Manager" to let the user perform such operation.

Associated revisions

Revision b51cb21d - 2018-09-21 10:38 AM - Alessandro Pasotti

GPKG Browser VACUUM menu item

Fixes #19895 - Garbage-collection is not performed after deletion of vector layer from geopackage

By design, VACUUM (being a potentially time-consuming operation)
was automatically performed only after deleting a raster
while when deleting a vector layer it was not executed.

This commit adds a menu item in the browser that allows
the user to perform this operation on the DB.

History

#1 - 2018-09-20 01:01 PM - Alessandro Pasotti

- Status changed from Open to Feedback

- Assignee set to Alessandro Pasotti

I might be wrong but IIRC this was intentional, because a VACUUM can take a very long time on large layers, I think that a possible solution would be to add a "Run vacuum" option to the browser data item of a geopackage layer.

What do you think?

#2 - 2018-09-20 03:13 PM - Andrea Giudiceandrea

It could be OK for me to add "Run vacuum" or "Compact geopackage" or something similar in the Browser panel for geopackages (along with "Add Connection"/"Remove connection" and "Create a New Layer or Table...") and eventually also in DB Manager (along with "Re-connect" and "Remove").

However, a VACUUM sql command is still executed automatically after a raster layer has been deleted from a geopackage, as you can see in
https://github.com/qgis/QGIS/blob/master/src/providers/ogr/qgsgeopackagedataitems.cpp#L441-L450

#3 - 2018-09-20 03:45 PM - Alessandro Pasotti

mmm looks like I wrote that code :) So, if the VACUUM is already executed I don't really know what to do here.

#4 - 2018-09-20 05:05 PM - Andrea Giudiceandrea

It seems that the VACUUM is only executed (in deleteGeoPackageRasterLayer) after a **RASTER layers** is deleted. I will confirm this when a new build of qgis-dev with your commit commit:e62c4eb941c5b2e290675f8cab75758857c2a502 will be available on osgeo4w

For **VECTOR layer** I cannot trace down in the source code if and where VACUUM is executed. The fact is that the geopackage is not "vacuumed" after a vector layer is deleted.

#5 - 2018-09-21 08:48 AM - Alessandro Pasotti

- Tracker changed from Bug report to Feature request

Yes, you are right of course, as I mentioned in my first reply IIRC it was not implemented on purpose and the idea was to add a menu item to do that.

Btw, that's technically a new feature, but it could also be considered a bug, so let's see if I can do it!

#6 - 2018-09-21 09:46 AM - Andrea Giudiceandrea

Thank you Alessandro! I'm not so skilled to add such an options by myself (without the risk of breaking something).

I was writing this before reading your last comment, so now it may not be so useful anymore:

I confirm that VACUUM is executed after a RASTER layer is deleted, and that is not executed after a VECTOR layer is deleted.

As I can see in the source code (qgsgeopackagedataitems.ccp):

- for **QgsGeoPackageRasterLayerItem**, executeDeleteLayer() calls deleteGeoPackageRasterLayer() which executes all the sql commands in order to delete a RASTER layer and its references and **executes the VACUUM sql command at the end**;

- for **QgsGeoPackageVectorLayerItem**, executeDeleteLayer() calls deleteLayer() so relying on GDAL (GDALGeoPackageDataset::DeleteLayer) to execute all the sql commands required to delete a VECTOR layer and its references: DeleteLayer() **does not executes a VACUUM sql command at the end**.

So I think we could:

- **add a "Run vacuum"** or "Compact geopackage" options as previously said (which is anyway needed to allow users to easily shrink geopackages)

and one of the following

1) add the VACUUM sql command execution after ::deleteLayer() in QgsGeoPackageVectorLayerItem::executeDeleteLayer() or in a little deleteGeoPackageVectorLayer() (for symmetry)

or

2) remove the VACUUM sql command from deleteGeoPackageRasterLayer (provided we add the "Run vacuum" option).

By the way - you probably know it already - I realised now that since GDAL 2.3.0 there are two new functions in GDALGeoPackageDataset: DeleteRasterLayer() and DeleteVectorOrRasterLayer() along with DeleteLayer().

#7 - 2018-09-21 10:46 AM - Alessandro Pasotti

- Status changed from Feedback to In Progress

- Pull Request or Patch supplied changed from No to Yes

PR https://github.com/qgis/QGIS/pull/7976

#8 - 2018-09-21 12:16 PM - Anonymous

- % Done changed from 0 to 100

- Status changed from In Progress to Closed

Applied in changeset commit:qgis|b51cb21d1aeecb2cc89bc8057ccc8d649306fe8d.