

QGIS Application - Feature request #4330

Spatial Indexing of Shapefiles on first load

2011-09-28 03:17 AM - Jonathan Moules

Status: Open	
Priority: Normal	
Assignee:	
Category: Vectors	
Pull Request or Patch Supplied:	Resolution:
Easy fix?: No	Copied to github as #: 14267
Description	
Hi, Earlier today I loaded three shapefiles into my 1.7.0 QGIS with a combined 5million+ features. QGIS defaults to "visible" for new layers, so these hung the program for several minutes as they don't have spatial indexes. My suggestion therefore is that whenever a user adds a shapefile with more than x features (x = 10,000? I dunno) that doesn't have a QGIS spatial index, QGIS pops up a dialog asking the user if they want to create an index. For bonus points it could have checkboxes for "remember my answer for this session" and "remember my answer always". ArcGIS does this with Rasters and pyramids - it's very useful.	

History

#1 - 2011-10-02 11:12 AM - Giovanni Manghi

I'm not sure this issue (of slow vector loading) goes away by creating spatial indexes, from my tests it doesn't.

The real issue is this: #1978

Have you made a test and noticed a noticable difference (in a big shapefile) with indexes and without?

#2 - 2011-10-02 11:12 AM - Giovanni Manghi

- Category set to Vectors

#3 - 2011-10-03 02:05 AM - Jonathan Moules

Ah, it depends what you mean. I actually encountered this by loading a shapefile with ~3million line features so I can answer your question.

A spatial index will MASSIVELY increase loading speed if you're only looking at part of the entire dataset; that's what they're designed for. I.e., my 3million features cover the entirety of the English county of Warwickshire (about 2 thousand sq km), if I'm zoomed to a small area, say a town or village, it loads very fast when I have a spatial index (a few seconds). If I don't have the index, looking at that same small area takes forever (several minutes).

On the other hand, I doubt (but haven't tested it) that a spatial index will speed viewing of the entire dataset at once, which is what you may be testing. It shouldn't, that's not how they work. But you don't typically use a dataset with that many features by viewing it all at once (you won't be able to see the features after all).

#4 - 2011-10-03 02:23 AM - Giovanni Manghi

Jonathan Moules wrote:

Ah, it depends what you mean. I actually encountered this by loading a shapefile with ~3million line features so I can answer your question. A spatial index will MASSIVELY increase loading speed if you're only looking at part of the entire dataset; that's what they're designed for. I.e., my 3million features cover the entirety of the English county of Warwickshire (about 2 thousand sq km), if I'm zoomed to a small area, say a town or village, it loads very fast when I have a spatial index (a few seconds). If I don't have the index, looking at that same small area takes forever

(several minutes).

On the other hand, I doubt (but haven't tested it) that a spatial index will speed viewing of the entire dataset at once, which is what you may be testing. It shouldn't, that's not how they work. But you don't typically use a dataset with that many features by viewing it all at once (you won't be able to see the features after all).

Of course, you are right. The feature request (to add spatial indexes on load) make perfectly sense.

#5 - 2011-12-16 02:10 PM - Giovanni Manghi

- Target version set to Version 1.7.4

#6 - 2011-12-25 08:28 AM - Alexander Bruy

- Subject changed from *Spatial Indexing of Shapefiles* to *Spatial Indexing of Shapefiles on first load*

#7 - 2012-04-15 10:09 AM - Giovanni Manghi

- Target version changed from Version 1.7.4 to Version 2.0.0

#8 - 2012-10-06 02:15 AM - Pirmin Kalberer

- Target version changed from Version 2.0.0 to Future Release - Nice to have

#9 - 2017-05-01 12:50 AM - Giovanni Manghi

- Easy fix? set to No