QGIS Application - Bug report #14965
wfs provider copies features when zoomin in and out

2016-06-06 08:21 AM - Richard Duivenvoorde

Status:	Closed			
Priority:	Normal			
Assignee:	Even Rouault			
Category:	Unknown			
Affected QGIS version:	master	**Regression?:**	No	
Operating System:		**Easy fix?:**	No	
Pull Request or Patch supplied:	No	**Resolution:**	wontfix	
Crashes QGIS or corrupts data:	No	**Copied to github as #:**	22914	

Description

WFS provider multiplies features when you zoom in and out.

0

Create a project in epsg:28992 (dutch grid)

1

Add our national border ('landsgrens') from this wfs:

https://geodata.nationaalgeoregister.nl/bestuurlijkegrenzen/wfs?request=GetCapabilities

2

Tick 'Show feature count' in legend.

It will show you 1 feature (multipolygon)

3

Now: zoom in and out a few times covering and not covering the country.. I use the mouse-scroll for this, but zoomin tool also works :-)

4

The number of features will add up after a while (not EVERY zoom...)

You can also see it in the feature table.

First I had 3 features at a certain moment and thought it had something to do with GML parsing from wfs 2.0 request... but it is not...

History

#1 - 2016-06-12 06:32 AM - Even Rouault

This is an issue with the server. The gml:id (or fid in WFS 1.0) that it returns are not constant over multiple GetFeature calls.

For example, if you paste the following URL in your browser, and refresh it multiple times, you will see the gml:id changing

https://geodata.nationaalgeoregister.nl/bestuurlijkegrenzen/wfs?SERVICE=WFS&REQUEST=GetFeature&VERSION=2.0.0&TYPENAMES
bestuurlijkegrenzen:landsgrens

True in WFS 1.0 & 1.1 too.

2 workarounds:

- uncheck "Only requests features overlapping the view extent"

- or set a custom SQL using the DISTINCT keyword, such as "SELECT DISTINCT * FROM landsgrens"

I cannot really think to a way of auto-detecting this behaviour on client side.

#2 - 2016-06-12 11:17 AM - Richard Duivenvoorde

Just brainstorming: what about a md5sum of the geom as wkt? This would be a lot of overhead maybe? But maybe can be done in a separate thread (after the loading of the features?)?
Would this maybe be a way to make such clients less dependent on the different server implementations.

Another point: SHOULD this be always the same between requests? Because I know it is Geoserver they are using, and they are also very picky on standards, so IF standards declare that the (f)id's should stay the same, I'm pretty sure they will try to make it that way :-)

#3 - 2016-06-12 11:35 AM - Even Rouault

Computing a MD5 is completely doable and actually currently implemented in the case where you do a SELECT DISTINCT. It is done on all selected fields (attributes+geometry). But I'm not completely sure it is a good idea to do it in the general case as you could potentially have different features with same values (for non spatial tables).

The WFS standards require a FID to remain constant over time, otherwise requests that work on feature ids will not work, WFS-T either, etc...

WFS 1.1 spec :
"7.1 Feature and Element Identifiers
This document assumes that every feature instance that a particular WFS implementation can operate upon is uniquely identifiable. That is to say, when a WFS implementation reports a feature identifier for a feature instance, that feature identifier is unique to the server and can be used to repeatedly reference the same feature instance (assuming it has not been deleted)."

WFS 2.0 spec:
"7.2.1 Assigning resource identifiers
Each feature instance in a WFS shall be assigned a persistent unique resource identifier which is assigned by the server when the feature is created.
This identifier shall be invariant under all WFS operations including delete which means that a resource identifier cannot be reused once it has been assigned.
Resource identifiers are not intended to associate WFS resources with real world objects and the values do not have to be meaningful outside the scope of a web feature service instance.

7.2.2 Encoding resource identifiers
For features encoded using GML, the resource identifier shall be encoded in the XML attribute gml:id."

#4 - 2016-06-12 12:47 PM - Jukka Rahkonen

If you see ids like gml:id="landsgrens.fid--6b30a5d1_155461cb2a6_5502" coming from Geoserver it means that the source datastore is configured without pointing to a primary key which could be used for generating good, unique, and constant fids. The ending part "--6b30a5d1_155461cb2a6_5502" is at least unique within each GetFeature.

Contact nationaalgeoregister.nl. This issue is not at all hard to fix for them. I think that trying to handle the case in QGIS side is waste of time.

#5 - 2016-06-24 07:34 AM - Even Rouault
- Resolution set to wontfix

Closing. This can/should be normally easily fixed on server side.

#6 - 2016-06-24 01:29 PM - Nyall Dawson

- Status changed from Open to Closed

#7 - 2017-09-22 10:05 AM - Jürgen Fischer

- Category set to Unknown

Files

wsfmultiplier.png	187 KB	2016-06-06	Richard Duivenvoorde