

RhodeCode CE/EE - Bug #5633

Moderately large pull requests fail because inefficient use of reviewer_data_json column in pull_requests table

22.08.2020 10:09 - Yechen Qiao

Status:	Resolved	Start date:	22.08.2020
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Sorting:		Affected Version:	
Commit Number:			

Description

When we are trying to merge sufficiently large number of changes, internal service error occurs while attempting to make the pull requests.

On further investigation, it is because reviewer_data_json is too large.

But most of the space taken is due to user's base64 encoded images. The text in there easily fits if I remove the image data.

Is there a purpose why those user images are stored there? If not, would it be better to simply not store user images at all and parse them in real-time? :)

History

#1 - 22.08.2020 10:10 - Yechen Qiao

"But most of the space taken is due to user's base64 encoded images." to clarify this seems to be user photo retrieved from rhodecode, which seems to not need to be stored.

#2 - 22.08.2020 11:38 - Marcin Kuzminski [CTO]

- Status changed from New to In Progress

Hi,

The reason why those images are there is that there are two backends for the user images: gravatar, which is a link to the user gravatar, or the built-in initials gravatar. In the case of reviewer data we store only the LINK to the image, but because Previously RhodeCode didn't have a way to store artifacts, the links in case of the second backend were actually built base64 images.

Since now we have an artifact store, i guess it would be a good time to replace the on-the-fly generated base64 svg images to an actually stored link in the artifacts.

But it's odd that the DB cannot fit the data in that column, maybe there's some problem with the DB and the column limit ? I'm guessing you're using MySQL as a backend which has those odd limitations for some columns.

#3 - 22.08.2020 17:35 - Yechen Qiao

Yes exactly! we are using MariaDB. The column has TEXT type. If we increase to mediumtext manually store is fine but when retrieve it causes extreme slowdown to load the closed tab of the pull request list in project. Until eventually this simply became an Ajax error when loading it. We had to revert back by altering the column to have no reviewer info for the offending PRs. Interestingly after doing so reviewer pane still seems to have all the info it had before?

#4 - 22.08.2020 22:17 - Marcin Kuzminski [CTO]

I Think you'd need to stop all traffic from RhodeCode (stop it completely) and then do the column alter

#5 - 12.10.2020 23:13 - Redmine Integration

- Status changed from In Progress to Resolved

Commit b5299f6ded6a by Marcin Kuzminski marcin@rhodecode.com on stable branch changed this issue.

<https://code.rhodecode.com/rhodecode-enterprise-ce/changeset/b5299f6ded6a6201fce67f03369db52de0016fb1>