

Pulp - Task #1803

Plan replacement of mongodb with postgres

03/31/2016 04:53 PM - Anonymous

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:	semyers	% Done:	0%
Category:	API	Estimated time:	0:00 hour
Sprint/Milestone:		QA Contact:	
Platform Release:		Complexity:	
Blocks Release:		Smash Test:	
Backwards Incompatible:	Yes	Verified:	No
Groomed:	No	Verification Required:	No
Sprint Candidate:	No	Sprint:	Sprint 5
Tags:	Documentation, Pulp 2		
Description			
Related issues:			
Related to Pulp - Task #1872: Profile Django ORM instantiation cost		CLOSED - CURRENTRELEASE	
Related to Pulp - Task #1873: Plan REST API for 3.0		CLOSED - CURRENTRELEASE	
Related to Pulp - Task #1874: Plan User/Auth system for 3.0		CLOSED - DUPLICATE	

History

#1 - 03/31/2016 05:11 PM - ipanova@redhat.com

- Status changed from NEW to ASSIGNED

#2 - 03/31/2016 06:21 PM - semyers

I'm in the process of writing all this in more detail in an etherpad, but for now I'll outline the work so far.

pcreech and I started by basically just writing out all of models we had, and trying to organize them in a diagram.

This went poorly, because pulp has four distinct categories of models, and the tool we were using didn't make it very easy to diagram them all together. With all the pulp releasing I've been doing, and all the RHUI pcreech has been doing, all of this finally came together toward the end of last week while jortel was on PTO. Before unleashing the confusing and questionably useful diagrams on the team, I wanted to get jortel's feedback. I was not disappointed.

With his help, we've identified the aforementioned data categories: Repositories, Content Units, RBAC, Tasks. I'm writing up a doc right now that explains what collections we currently have in mongo, to which category they belong, how they currently relate to each other, and some speculation about how we can migrate the data in that category to a more relational design. Since I'll be speculating, this doc will go up in an etherpad for review and improvement by the small team focused on this at the moment (jortel, pcreech, me), before being submitted to the team at large.

Depending on how quickly I get this done and reviewed, the link to that etherpad should most likely be appearing here tomorrow.

#3 - 04/05/2016 07:02 PM - semyers

- Private changed from No to Yes

The "Relational Pulp" etherpad has been drafted and edited. It is available for internal comment at <http://pulp.etherpad.corp.redhat.com/244> by the entirety of the pulp team.

#4 - 04/05/2016 07:02 PM - semyers

- Private changed from Yes to No

#5 - 04/27/2016 07:14 PM - semyers

- Related to Task #1872: Profile Django ORM instantiation cost added

#6 - 04/28/2016 03:37 PM - mhrivnak

- Related to Task #1873: Plan REST API for 3.0 added

#7 - 04/28/2016 03:48 PM - mhrivnak

- Related to Task #1874: Plan User/Auth system for 3.0 added

#8 - 05/02/2016 05:30 PM - mhrivnak

- Sprint/Milestone changed from 19 to 20

#9 - 05/10/2016 04:24 PM - semyers

The migration plan is largely "ratified", in that objections to it have been (or are currently being) addressed. I feel like we've reached a point where what we don't know outweighs what we do know, and the best way I can think of to bridge the knowledge gap is to finishing modeling pulp out on postgres and start trying to use it. I'll be modifying my relational pulp project to this end over the coming days.

<https://github.com/seandst/relational-pulp/>

#10 - 05/23/2016 03:33 PM - mhrivnak

- Sprint/Milestone changed from 20 to 21

#11 - 06/06/2016 06:44 PM - semyers

Quick update, progress is still being made. pcrech has converted the project from docker to Vagrant, which is awesome. The pulp platform models related to repos and units have been written down, and I'm currently porting RPM's units and repo-related models now so we can start looking at migrating data from nonrel-pulp to rel-pulp to see what explodes and needs to be revisited. :)

#12 - 06/13/2016 03:56 PM - mhrivnak

- Sprint/Milestone changed from 21 to 22

#13 - 07/05/2016 02:31 PM - mhrivnak

- Sprint/Milestone changed from 22 to 23

#14 - 07/07/2016 11:55 PM - semyers

- Status changed from ASSIGNED to CLOSED - CURRENTRELEASE

This is largely done (and by done, I mean now we can start to do it? :D). The rel-pulp doc and related repo are generally accepted by the team, and are now available in the pulp namespace on github:

<https://github.com/pulp/relational-pulp> - check out the db-translation-guide.md doc.

Most of the work was done by a cabal made up of jortel, pcreech, and myself with lots of great ideas and insights coming from all over the place, including stakeholders not on the pulp team. Thanks to all involved. More meetings will ensue, followed by more redmine tasks.

#15 - 03/08/2018 08:03 PM - bmbouter

- *Sprint set to Sprint 5*

#16 - 03/08/2018 08:05 PM - bmbouter

- *Sprint/Milestone deleted (23)*

#17 - 04/15/2019 10:32 PM - bmbouter

- *Tags Pulp 2 added*