

Pulp - Issue #4604

Remove duplicates does not work properly if there are multiple repeated content types in the repo

03/29/2019 03:08 PM - ipanova@redhat.com

Status:	MODIFIED	Start date:	
Priority:	Normal	Due date:	
Assignee:	ipanova@redhat.com	Estimated time:	0:00 hour
Category:		Sprint Candidate:	No
Sprint/Milestone:	3.0	Tags:	
Severity:	2. Medium	QA Contact:	
Version:		Complexity:	
Platform Release:		Smash Test:	
Blocks Release:		Verified:	No
OS:		Verification Required:	No
Backwards Incompatible:	No	Sprint:	Sprint 51
Triaged:	Yes		
Groomed:	No		

Description

It got manifested in docker plugin.

Snippet code:

```
async def run(self):
    """
    The coroutine for this stage.
    Returns:
        The coroutine for this stage.
    """
    async for batch in self.batches():
        rm_q = Q()
        for d_content in batch:
            if isinstance(d_content.content, self.model):
                unit_q_dict = {
                    field: getattr(d_content.content, field) for field in self.field_names
                }
                # Don't remove *this* object if it is already in the repository version.
                not_this = ~Q(pk=d_content.content.pk)
                dupe = Q(**unit_q_dict)
                rm_q |= Q(dupe & not_this)
        queryset_to_unassociate = self.model.objects.filter(rm_q)
        self.new_version.remove_content(queryset_to_unassociate)

        for d_content in batch:
            await self.put(d_content)
```

When we process `declarative_content.content` in the batch that is different from the model we are supposed to remove the dupes, we end up to have `rm_q` empty which leads to removal of whole content of the model.

Associated revisions

Revision 01b5dd2a - 04/01/2019 06:48 PM - ipanova@redhat.com

Remove duplicates does not work properly if there is multiple content types in the repo.

closes #4604

<https://pulp.plan.io/issues/4604>

History

#1 - 03/29/2019 03:22 PM - ipanova@redhat.com

- File Screenshot from 2019-03-29 15-22-09.png added

#2 - 03/29/2019 03:40 PM - CodeHeeler

- Triaged changed from No to Yes

- Sprint set to Sprint 51

#3 - 03/29/2019 03:48 PM - ipanova@redhat.com

- Status changed from NEW to POST

- Assignee set to ipanova@redhat.com

<https://github.com/pulp/pulpcore-plugin/pull/81>

#4 - 03/29/2019 03:55 PM - ipanova@redhat.com

- Subject changed from Remove duplicates does not work properly if there is multiple content types in the repo to Remove duplicates does not work properly if there are multiple repeated content types in the repo

#5 - 04/01/2019 08:41 PM - bmbouter

@ipanova I love the screenshot! It shows the problem exactly. I'm looking at this PR now.

#6 - 04/01/2019 08:52 PM - ipanova@redhat.com

- Status changed from POST to MODIFIED

Applied in changeset [pulpcore-plugin|01b5dd2a38c4f55fc710fb86b8fde8c92d4175bf](#).

#7 - 04/04/2019 12:29 PM - ipanova@redhat.com

- Tags Pulp 3 added

#8 - 04/25/2019 06:44 PM - daviddavis

- Sprint/Milestone set to 3.0

#9 - 04/26/2019 10:31 PM - bmbouter

- Tags deleted (Pulp 3)

Files

Screenshot from 2019-03-29 15-22-09.png

259 KB

03/29/2019

ipanova@redhat.com