

Pulp - Issue #7876

NoneType' object has no attribute 'pk'

11/19/2020 08:18 PM - adam.winberg@smhi.se

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	High	Due date:	
Assignee:	bmbouter	Estimated time:	0:00 hour
Category:		Groomed:	No
Sprint/Milestone:	3.11.0	Sprint Candidate:	No
Severity:	2. Medium	Tags:	Katello
Version:		Sprint:	Sprint 90
Platform Release:		Quarter:	Q1-2021
OS:			
Triaged:	Yes		
Description			
On a fresh pulp3 installation, running consecutive 2to3 migrations results in following error (first migration ran without error):			
<pre>Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: Traceback (most recent call last): Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/rq/worker .py", line 936, in perform_job Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: rv = job.perform() Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/rq/job.py ", line 684, in perform Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: self._result = self._execute() Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/rq/job.py ", line 690, in _execute Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: return self.func(*self.args, **self.kwargs) Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulp_2to3 _migration/app/tasks/migrate.py", line 139, in migrate_from_pulp2 Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: migrate_content(plan) Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulp_2to3 _migration/app/migration.py", line 36, in migrate_content Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: plugin.migrator.migrate_content_to_pulp3() Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulp_2to3 _migration/app/plugin/rpm/migrator.py", line 140, in migrate_content_to_pulp3 Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: loop.run_until_complete(dm.create()) Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib64/python3.6/asyncio/base_events.p y", line 484, in run_until_complete Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: return future.result() Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulp_2to3 _migration/app/plugin/content.py", line 88, in create Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: await pipeline Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulpcore/ plugin/stages/api.py", line 225, in create_pipeline Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: await asyncio.gather(*futures) Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulpcore/ plugin/stages/api.py", line 43, in __call__ Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: await self.run() Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulpcore/ plugin/stages/artifact_stages.py", line 244, in run Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: RemoteArtifact.objects.bulk_get_or_create(self ._needed_remote_artifacts(batch)) Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: File "/usr/lib/python3.6/site-packages/pulpcore/ plugin/stages/artifact_stages.py", line 296, in _needed_remote_artifacts Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: if remote_artifact.remote_id == d_artifact.rem ote.pk: Nov 19 19:04:08 lxserv2285.smhi.se rq[1208174]: AttributeError: 'NoneType' object has no attribute 'pk'</pre>			

The first couple of migrations resulted in [#7842](#). Next day, new migration attempt and now I get this, which is produced earlier in the migration process than the previous error.

Running in an rpm-based installation on RHEL8:

```
python3-pulp-rpm-3.7.0-1.e18.noarch
python3-pulpcore-3.7.3-1.e18.noarch
python3-pulp-2to3-migration-0.5.1-1.e18.noarch
```

Associated revisions

Revision 25745275 - 02/16/2021 03:23 PM - bmbouter

Adds additional d_artifact.remote check

Not all d_artifact objects have remotes, e.g. the migration plugin does not. Other places in the code account for this, but there is one area that does not. This adds an additional check to one codepath that checks for equality of d_artifact.remote.pk only if d_artifact.remote exists.

closes #7876

History

#1 - 11/23/2020 06:07 PM - ttereshc

Hi Adam,

Do you have a large pulp2 setup? Many repositories? I think this issue depends on pulp2 data and it would be helpful to know some details, so we can reproduce the problem.

How do you get content in pulp2? Is it usually sync? On_demand? Do you upload your own content into pulp2? Do you copy content between repositories? Is it possible that there are repositories without any feed in the importer?

#2 - 11/23/2020 06:09 PM - ttereshc

FWIW, I haven't tracked it down yet. I think it might be an issue in pulpcore for the case when the same content is migrated with the remote and without.

#3 - 11/23/2020 06:35 PM - adam.winberg@smhi.se

We have about 80 repos, where 50 of them have a feed-url, and 38 of those are synced on a regular basis. The other 12 are synced manually when needed. And the repos without a feed-url are used as frozen repo images, which we copy rpm/erratas to from the synced repos in a more controlled manner. For example we have the RHEL8 baseos repo which we sync every day, and once a week we copy important erratas from that repo to our 'frozen-baseos' repo which is then used to serve our clients.

We very seldom upload our own content, it is almost all synced from external repos.

Some outputs from migration:

```
{
  "code": "migrating.rpm.content",
  "done": 153807,
  "message": "Migrating rpm content to Pulp 3 rpm",
  "state": "completed",
  "suffix": null,
  "total": 153807
},
{
  "code": "migrating.rpm.content",
  "done": 90653,
  "message": "Migrating rpm content to Pulp 3 erratum",
  "state": "completed",
  "suffix": null,
  "total": 90653
},
```

#4 - 12/01/2020 12:35 PM - ttereshc

- Project changed from Migration Plugin to Pulp

Agreed on the pulpcore meeting last week that this should be fixed in pulpcore by adding a check if a remote for an artifact exists https://github.com/pulp/pulpcore/blob/50c6705941c5f249fdacda6a52f0bd0c0efa453c/pulpcore/plugin/stages/artifact_stages.py#L297.

Likely will need a backport for katello.

Moving to the pulpcore tracker.

#5 - 12/01/2020 01:32 PM - adam.winberg@smhi.se

So I tested this a bit more. Redid my whole migration, worked without errors. Reran migration, worked. Reran again the next day, when pulp2 repos had been updated and published, and content had been copied to our frozen repos (repos without feed), but the frozen repos had NOT been published - worked. Reran again the next day after that, when the frozen repos had been published - did NOT work, got the "object has no attribute 'pk'" error.

#6 - 12/01/2020 05:01 PM - fao89

- Triaged changed from No to Yes
- Sprint set to Sprint 86

#7 - 12/04/2020 06:40 PM - rchan

- Sprint changed from Sprint 86 to Sprint 87

#8 - 12/18/2020 07:35 PM - rchan

- Sprint changed from Sprint 87 to Sprint 88

#9 - 01/14/2021 04:43 PM - dkliban@redhat.com

- Project changed from Pulp to Migration Plugin

#10 - 01/19/2021 03:42 PM - dalley

- Priority changed from Normal to High

#11 - 01/19/2021 08:16 PM - ttereshc

- Sprint/Milestone set to 0.9.0

#12 - 01/22/2021 09:32 PM - rchan

- Sprint changed from Sprint 88 to Sprint 89

#13 - 02/02/2021 05:53 PM - jsherril@redhat.com

- Tags Katello added

#14 - 02/04/2021 11:09 PM - rchan

- Sprint changed from Sprint 89 to Sprint 90

#15 - 02/12/2021 01:33 PM - bmbouter

ttereshc wrote:

Agreed on the pulpcore meeting last week that this should be fixed in pulpcore by adding a check if a remote for an artifact exists https://github.com/pulp/pulpcore/blob/50c6705941c5f249fdacda6a52f0bd0c0efa453c/pulpcore/plugin/stages/artifact_stages.py#L297.

+1 to moving this back to pulpcore.

Can you remind me of the situation where d_artifact.remote is None. Does the migration plugin just not have a Remote for some d_artifact types? Which types are those?

#16 - 02/12/2021 01:46 PM - ttereshc

Currently it's purely the migration case:

It can be an uploaded content unit (RPM or File), or a content for which pulp2 importer has been removed, or user is not interested in migrating some importer/remote, so it's not available on pulp3 side.

Another reason was inconsistency in pulpcore code.

There are parts which already check if remote is available or not but this one place just works under assumption that a remote is always there.

I don't recall any other use cases. I think someone came up with potential future use cases when a remote will be null there but I don't remember them.

#17 - 02/12/2021 06:10 PM - ttereshc

- Project changed from Migration Plugin to Pulp

- *Sprint/Milestone deleted (0.9.0)*

- *Quarter set to Q1-2021*

#18 - 02/12/2021 08:55 PM - bmbouter

- *Status changed from NEW to ASSIGNED*

- *Assignee set to bmbouter*

#19 - 02/12/2021 11:23 PM - pulpbot

- *Status changed from ASSIGNED to POST*

PR: <https://github.com/pulp/pulpcore/pull/1125>

#20 - 02/16/2021 03:23 PM - bmbouter

- *Status changed from POST to MODIFIED*

Applied in changeset [pulpcore|25745275c0f3a7bd832946749ed27381694d612f](#).

#21 - 03/15/2021 01:39 PM - ipanova@redhat.com

- *Sprint/Milestone set to 3.11.0*

#22 - 03/16/2021 07:55 PM - ipanova@redhat.com

- *Status changed from MODIFIED to CLOSED - CURRENTRELEASE*