

## Migration Plugin - Issue #7994

### Serialization of pulp3\_repository\_version is wrong or failing at times

12/11/2020 08:28 PM - ttereshc

<b>Status:</b>	CLOSED - CURRENTRELEASE	<b>Start date:</b>	
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	ipanova@redhat.com	<b>Estimated time:</b>	0:00 hour
<b>Category:</b>		<b>Sprint Candidate:</b>	No
<b>Sprint/Milestone:</b>	0.7.0	<b>Tags:</b>	
<b>Severity:</b>	2. Medium	<b>Sprint:</b>	Sprint 87
<b>Platform Release:</b>		<b>Quarter:</b>	
<b>OS:</b>			
<b>Triaged:</b>	Yes		
<b>Groomed:</b>	No		

#### Description

There is [an attempt to exclude pulp3\\_repository\\_version if it's not set](#).

This causes a failure if I try to limit the output to certain fields. Even if I ask to include the field itself.

```
[vagrant@pulp2-nightly-pulp3-source-centos7 ~]$ http ':/pulp/api/v3/pulp2content/?limit=1&fields=["pulp3_repository_version"]'
HTTP/1.1 500 Internal Server Error
Connection: close
Content-Length: 27
Content-Type: text/html; charset=UTF-8
Date: Fri, 11 Dec 2020 19:19:36 GMT
Server: gunicorn/20.0.4
Vary: Cookie
X-Frame-Options: SAMEORIGIN

<h1>Server Error (500)</h1>
```

#### Logs

```
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: pulp: djan
go.request:ERROR: Internal Server Error: /pulp/api/v3/pulp2content/
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: Traceback
(most recent call last):
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr
/local/lib/pulp/lib64/python3.6/site-packages/django/core/handlers/exception.py", line 34, in inne
r
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: response =
_get_response(request)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr
/local/lib/pulp/lib64/python3.6/site-packages/django/core/handlers/base.py", line 115, in _get_res
ponse
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: response =
self.process_exception_by_middleware(e, request)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr
/local/lib/pulp/lib64/python3.6/site-packages/django/core/handlers/base.py", line 113, in _get_res
ponse
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: response =
wrapped_callback(request, *callback_args, **callback_kwargs)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr
/local/lib/pulp/lib64/python3.6/site-packages/django/views/decorators/csrf.py", line 54, in wrappe
d_view
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: return vie
w_func(*args, **kwargs)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr
```

```
/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/viewsets.py", line 114, in view
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: return self.dispatch(request, *args, **kwargs)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/views.py", line 505, in dispatch
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: response = self.handle_exception(exc)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/views.py", line 465, in handle_exception
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: self.raise_uncaught_exception(exc)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/views.py", line 476, in raise_uncaught_exception
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: raise exc
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/views.py", line 502, in dispatch
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: response = handler(request, *args, **kwargs)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/mixins.py", line 43, in list
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: return self.get_paginated_response(serializer.data)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/serializers.py", line 760, in data
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: ret = super().data
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/serializers.py", line 260, in data
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: self._data = self.to_representation(self.instance)
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/serializers.py", line 678, in to_representation
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: self.child.to_representation(item) for item in iterable
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/usr/local/lib/pulp/lib64/python3.6/site-packages/rest_framework/serializers.py", line 678, in <listcomp>
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: self.child.to_representation(item) for item in iterable
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: File "/home/vagrant/devel/pulp-2to3-migration/pulp_2to3_migration/app/serializers.py", line 172, in to_representation
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: del result['pulp3_repository_version']
Dec 11 19:08:34 pulp2-nightly-pulp3-source-centos7.rhgoose.example.com gunicorn[32087]: KeyError: 'pulp3_repository_version'
```

It's easy to fix by checking if the key is there. However I wonder if there are more issues here. If I make a request using httpie without any fields specified, I get the expected format, no `pulp3_repository_version`, because it's `None`.

```
$ http ':/pulp/api/v3/pulp2content/?limit=1'
HTTP/1.1 200 OK
Allow: GET, HEAD, OPTIONS
Connection: Keep-Alive
Content-Length: 586
Content-Type: application/json
Date: Fri, 11 Dec 2020 19:26:08 GMT
Keep-Alive: timeout=5, max=10000
Server: gunicorn/20.0.4
Vary: Accept, Cookie
X-Frame-Options: SAMEORIGIN

{
  "count": 266,
```

```

"next": "http://127.0.0.1:24817/pulp/api/v3/pulp2content/?limit=1&offset=1",
"previous": null,
"results": [
  {
    "downloaded": true,
    "pulp2_content_type_id": "iso",
    "pulp2_id": "eedf4b-d1d9-49b6-be5d-5d8c002ac3ab",
    "pulp2_last_updated": 1572546758,
    "pulp2_storage_path": "/var/lib/pulp/content/units/iso/6c/855d3e0d05e36b074280d7bac3ae14c5f3fcd5a0b50117332370bd614d72bd/5.iso",
    "pulp3_content": "/pulp/api/v3/content/file/files/6b9174ef-72ed-44d4-a61f-7458c7e2f1fd/",
    "pulp_created": "2020-12-11T18:29:43.890958Z",
    "pulp_href": "/pulp/api/v3/pulp2content/757d4bcf-ef54-4519-9a20-8b1d369f6798/"
  }
]
}

```

When I use bindings, I see it there with value None. Is it intentional? I thought that bindings are expected to produce the same result as API.

```

In [22]: pulp2content_api.list(ordering='pulp2_id', limit=1)
Out[22]:
{'count': 266,
 'next': 'http://127.0.0.1:24817/pulp/api/v3/pulp2content/?limit=1&offset=1&ordering=pulp2_id',
 'previous': None,
 'results': [{'downloaded': False,
              'pulp2_content_type_id': 'iso',
              'pulp2_id': '006d116f-aaa0-498b-82d4-f97aeb9f340c',
              'pulp2_last_updated': 1572546754,
              'pulp2_storage_path': '/var/lib/pulp/content/units/iso/2d/ffc2fa86aff8ebe792baee3fa7df0daee50888db60d9bba32c05ed68901288',
              'pulp3_content': '/pulp/api/v3/content/file/files/f816fb17-320d-496e-a16d-a8b5e17686cd/',
              'pulp3_repository_version': None,
              'pulp_created': datetime.datetime(2020, 12, 11, 18, 29, 43, 886139, tzinfo=tzlocal())
            },
            {'pulp_href': '/pulp/api/v3/pulp2content/8b94653f-9e74-4911-b258-178906a77cc5/'}]}

```

And a separate question, do we need to exclude it from the output at all? It seems that it would be better if API is consistent and always has certain fields present. Any reasons to exclude it?

## Associated revisions

### Revision 406f2315 - 12/14/2020 08:50 PM - ipanova@redhat.com

Fixed Pulp2Content serialization when filters are applied.

closes #7994 <https://pulp.plan.io/issues/7994>

### Revision 406f2315 - 12/14/2020 08:50 PM - ipanova@redhat.com

Fixed Pulp2Content serialization when filters are applied.

closes #7994 <https://pulp.plan.io/issues/7994>

### Revision 406f2315 - 12/14/2020 08:50 PM - ipanova@redhat.com

Fixed Pulp2Content serialization when filters are applied.

closes #7994 <https://pulp.plan.io/issues/7994>

## History

### #1 - 12/11/2020 09:07 PM - ipanova@redhat.com

I managed to reproduce this in case of migrated orphaned content which in this case does not belong to any of the repo versions We need to check if repo\_version is present among the keys as you have suggested. Also i think this would be a proper way on how to display fields

```
$ http GET ':24817/pulp/api/v3/pulp2content/?limit=1&fields=pulp3_content,pulp2_id'
```

```
HTTP/1.1 200 OK
Access-Control-Expose-Headers: Correlation-ID
Allow: GET, HEAD, OPTIONS
Connection: close
Content-Length: 288
Content-Type: application/json
Correlation-ID: 790c603695d64e339b10cf50790cc810
Date: Fri, 11 Dec 2020 20:15:18 GMT
Server: gunicorn/20.0.4
Vary: Accept, Cookie
X-Frame-Options: SAMEORIGIN
```

```
{
  "count": 532,
  "next": "http://localhost:24817/pulp/api/v3/pulp2content/?fields=pulp3_content%2Cpulp2_id&limit=1&offset=1",
  "previous": null,
  "results": [
    {
      "pulp2_id": "7f9cf757-ed35-4f0b-85d7-af4cf4f50103",
      "pulp3_content": "/pulp/api/v3/content/file/files/3fbc169d-9a88-4dec-884f-5a36a2c51b87/"
    }
  ]
}
```

versus

```
$ http GET ':24817/pulp/api/v3/pulp2content/?limit=1&fields=["pulp3_content"]'
HTTP/1.1 200 OK
Access-Control-Expose-Headers: Correlation-ID
Allow: GET, HEAD, OPTIONS
Connection: close
Content-Length: 152
Content-Type: application/json
Correlation-ID: 94d4e64a516d47bbb12756f25dd12785
Date: Fri, 11 Dec 2020 20:17:09 GMT
Server: gunicorn/20.0.4
Vary: Accept, Cookie
X-Frame-Options: SAMEORIGIN
```

```
{
  "count": 532,
  "next": "http://localhost:24817/pulp/api/v3/pulp2content/?fields=%5B%22pulp3_content%22%5D&limit=1&offset=1",
  "previous": null,
  "results": [
    {}
  ]
}
```

Honestly saying I don't remember what were the reasons to not serialize `repo_version` if not available. It probably makes sense to serialize that field in any case, this way end user will know to say which content is orphaned. I do remember why we have decided to serialize only when `repo_version` is available. It is available only if an erratum has been migrated, otherwise `repo_version` will always be none.

## #2 - 12/14/2020 06:35 PM - ipanova@redhat.com

I have looked some into this. Forget what i have said in the previous comment. The 500 was a result of misuse of filters.

This is the correct way

```
$ http GET ':24817/pulp/api/v3/pulp2content/?limit=5&fields=pulp3_repository_version'
HTTP/1.1 200 OK
Access-Control-Expose-Headers: Correlation-ID
Allow: GET, HEAD, OPTIONS
Connection: close
Content-Length: 606
Content-Type: application/json
Correlation-ID: 9a4a1cf8f92940f0bcdd7cb6c9dfe3b3
Date: Mon, 14 Dec 2020 17:41:52 GMT
Server: gunicorn/20.0.4
Vary: Accept, Cookie
X-Frame-Options: SAMEORIGIN
```

```
{
  "count": 43,
```

```

    "next": "http://localhost:24817/pulp/api/v3/pulp2content/?fields=pulp3_repository_version&limit=5&offset=5",
    "previous": null,
    "results": [
      {
        "pulp3_repository_version": "/pulp/api/v3/repositories/rpm/rpm/bbc95d78-6401-4ff7-bf7b-7ffb060f22b1/versions/1/"
      },
      {
        "pulp3_repository_version": "/pulp/api/v3/repositories/rpm/rpm/bbc95d78-6401-4ff7-bf7b-7ffb060f22b1/versions/1/"
      },
      {
        "pulp3_repository_version": "/pulp/api/v3/repositories/rpm/rpm/bbc95d78-6401-4ff7-bf7b-7ffb060f22b1/versions/1/"
      },
      {
        "pulp3_repository_version": "/pulp/api/v3/repositories/rpm/rpm/bbc95d78-6401-4ff7-bf7b-7ffb060f22b1/versions/1/"
      },
      {}
    ]
  }
}

```

The dict will always contain 'pulp3\_repository\_version' key and it will be none in all of the cases, except in case of migrated errata. Only errata has the relationship set to the pulp2\_repo from which we are extracting the repo\_version.

### #3 - 12/14/2020 06:45 PM - ttereshc

Thanks for sharing the correct syntax for requesting fields. I'm still trying to find the right way to do it with bindings.

http GET ':24817/pulp/api/v3/pulp2content/?limit=1&fields=pulp2\_id,pulp\_created' This will give HTTP 500, because pulp3\_repository\_version is not in the list. At least it makes sense that it's not failing when it's specified.

### #4 - 12/14/2020 06:48 PM - ipanova@redhat.com

ttereshc wrote:

Thanks for sharing the correct syntax for requesting fields. I'm still trying to find the right way to do it with bindings.

http GET ':24817/pulp/api/v3/pulp2content/?limit=1&fields=pulp2\_id,pulp\_created' This will give HTTP 500, because pulp3\_repository\_version is not in the list. At least it makes sense that it's not failing when it's specified.

you are right, in that case the dict will look like OrderedDict([('pulp\_created', '2020-12-14T17:41:03.874603Z'), ('pulp2\_id', '93c4a2ff-1ff8-435b-a57d-41ae4f05e4a2')])

### #5 - 12/14/2020 07:11 PM - ipanova@redhat.com

- Status changed from *NEW* to *ASSIGNED*
- Assignee set to *ipanova@redhat.com*
- Triaged changed from *No* to *Yes*
- Sprint set to *Sprint 87*

### #6 - 12/14/2020 08:50 PM - ipanova@redhat.com

- Status changed from *ASSIGNED* to *MODIFIED*

Applied in changeset [pulp:pulp-2to3-migration|406f23156e7c340f7fe5b94d8e06806daafdc953](https://pulp.plan.io/revisions/406f23156e7c340f7fe5b94d8e06806daafdc953).

### #7 - 01/22/2021 09:20 PM - ttereshc

- Sprint/Milestone set to *0.7.0*

### #8 - 02/04/2021 12:39 PM - pulpbot

- Status changed from *MODIFIED* to *CLOSED - CURRENTRELEASE*