

RPM Support - Issue #4798

Rpm file uploaded/published successfully but not accessible

05/09/2019 01:10 AM - kravir

Status:	NEW	Start date:	
Priority:	Normal	Due date:	
Assignee:		Estimated time:	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	2. Medium	Groomed:	No
Version:		Sprint Candidate:	No
Platform Release:		Tags:	Pulp 2
OS:		Sprint:	
Triaged:	Yes	Quarter:	

Description

Description of problem: I am new to Pulp, uploaded and published an rpm file but the rpm is not accessible. Why upload and publish will be successful if file is not accessible.

Steps to Reproduce:

1. pulp-admin rpm repo uploads rpm --repo-id dev-epel-7-x86_64 -f python2-markdown-2.4.1-4.el7.noarch.rpm -d .
2. pulp-admin rpm repo publish run --repo-id dev-epel-7-x86_64
3. pulp-admin rpm repo content rpm --repo-id dev-epel-7-x86_64 --match 'filename=python2-markdown-2.4.1-4.el7.noarch.rpm'

Actual results:

```
lrwxrwxrwx. 1 apache apache 137 May  8 17:45 /var/www/pub/yum/https/repos/dev-epel-7-x86_64/Packages/p/python2-markdown-2.4.1-4.el7.noarch.rpm -> /var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4a89d5e5/python2-markdown-2.4.1-4.el7.noarch.rpm
```

```
ls -l /var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4a89d5e5/python2-markdown-2.4.1-4.el7.noarch.rpm
```

```
ls: cannot access /var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4a89d5e5/python2-markdown-2.4.1-4.el7.noarch.rpm: No such file or directory
```

```
10.222.253.xx -- [08/May/2019:18:30:59 +0000] "GET /pulp/repos/general-prod-epel-7-x86_64/Packages/p/python2-markdown-2.4.1-4.el7.noarch.rpm HTTP/1.1" 302 -
```

```
10.222.253.xx -- [08/May/2019:18:31:00 +0000] "GET /streamer/var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4a89d5e5/python2-markdown-2.4.1-4.el7.noarch.rpm?policy=eyJleHRlbnNpb25zIjogeyJyZWlvdGVfaXAiOiAiMTAuMjIyY2MDEwN2JiMjBmMmJmNmNlMGNhMzRkZWZmZjEwY2ZnN2M0YTg5ZDVlNS9weXR0b24yLW1hcmtkb3duLTUuNC4xLTQuZWw3Lm5vYXJjaC5ycG0iLCAiZXhwYXJhdGlvbiI6IDE1NTczNDZlNTB9;signature=A3sIQRBfSxuDFmQunLrQkqXNpgV1PxxWhYFP44daSsPV1Gd2zvEsYYy_ur5FFYt-5NWDAlDfVtjynI8u5vdCygm26y8xJ6fqoJw70vGjOa6Zp-_9chzdyrTemNG8LGx2r1L_798HCT9k5F8g8FkctlYxdrmttrO7IHJpiZAwlhR8qHqTwmBldx3LBVcRS1MvJVdN5EkPQBttAQpd9OnYKTeLEDHpe3Tjk7M2GLjshc1bTpfH0y4W-2NYGoX7_aDYJg5d6YcmogVQpri-IsfjLV0EwalGo_dpOm4_dMfX1Q54ckUx40Zj5ne2bfftWbh26HiizdlCC0fptPuAs3--lQ%3D%3D HTTP/1.1" 404
```

Expected results:

HTTP status code 200

Additional info:

in /var/log/messages

```
May 8 22:42:17 ip-10-222-253-xx pulp_streamer: pulp.streamer.server:INFO: Download failed [404]:
http://epel.mirrors.ovh.net/epel/7/x86_64/Packages/p/python2-markdown-2.4.1-4.el7.noarch.rpm
May 8 22:42:17 ip-10-222-253-xx pulp_streamer: pulp.streamer.server:ERROR: All download attempts
failed: /var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4
a89d5e5/python2-markdown-2.4.1-4.el7.noarch.rpm
```

```
May 8 22:42:17 ip-10-222-253-xx pulp_streamer: [-] 127.0.0.1 - - [08/May/2019:22:42:16 +0000] "GE
T /var/lib/pulp/content/units/rpm/54/a8b444eca08f471ef257f60107bb20f1bc2ce0ca34defff10cfa7c4a89d5e
5/python2-markdown-2.4.1-4.el7.noarch.rpm HTTP/1.1" 404 - "-" "urlgrabber/3.10 yum/3.4.3"
```

Related issues:

Related to RPM Support - Issue #4059: During rpm upload filename should be p...

CLOSED - CURRENTRELEASE

History

#1 - 05/09/2019 05:52 PM - mdellweg

- Project changed from Debian Support to RPM Support
- Category changed from 16 to 22

#2 - 05/10/2019 04:43 PM - amacdona@redhat.com

- Tags Pulp 2 added

#3 - 05/10/2019 05:09 PM - ttereshc

- Description updated

#4 - 05/10/2019 05:14 PM - ttereshc

Please, provide details about your repo.

```
pulp-admin rpm repo list --repo-id dev-epel-7-x86_64 --details
```

And what did you do before trying to upload? E.g. created repo <this way>, ran sync, then removed repo and recreated it.

#5 - 05/13/2019 02:30 PM - ipanova@redhat.com

Steps to reproduce:

- 1 create repo with on_demand policy
- 2 sync repo
- 3 upload rpm that synced repo contains

When rpms are uploaded the calculated storage path contains the upload_id instead of filename (Example

```
:/var/lib/pulp/content/units/rpm/26/fb397d0d1335dc6f72e55e05a625eab848aeadbef24289e763a0c26a4ca49d5/271dd2f4-80a5-4200-ac74-9f9996ef20
0a)
```

Storage path would still point to the old storage path and the flag that indicates whether the content is downloaded will not be flipped to True.

```
> db.units_rpm.find()[0]
{
  "_id" : "90e6899e-01d8-4988-bad1-6ddfd0ae3387",
  "pulp_user_metadata" : {
  },
  "_last_updated" : 1557742882,
  "_storage_path" : "/var/lib/pulp/content/units/rpm/26/fb397d0d1335dc6f72e55e05a625eab848aeadbef24289e763a0c
26a4ca49d5/wolf-9.4-2.noarch.rpm",
```

```

"downloaded" : false,
"checksum" : "d61925ae8f51feccc2f1bcc2ecd6a83dcc95c3e8c5392f43aa4775c246a5edf2",
"checksumtype" : "sha256",
"checksums" : {
  "sha256" : "d61925ae8f51feccc2f1bcc2ecd6a83dcc95c3e8c5392f43aa4775c246a5edf2"
},
"version_sort_index" : "01-9.01-4",
"release_sort_index" : "01-2",
"name" : "wolf",
"epoch" : "0",
"version" : "9.4",
"release" : "2",
"arch" : "noarch",
"build_time" : 1331831362,
"buildhost" : "smqe-ws15",
"size" : 2439,
"filename" : "wolf-9.4-2.noarch.rpm",
"relativepath" : "wolf-9.4-2.noarch.rpm",
"group" : "Internet/Applications",
"provides" : [
  {
    "release" : "2",
    "epoch" : "0",
    "version" : "9.4",
    "flags" : "EQ",
    "name" : "wolf"
  }
],
"files" : {
  "file" : [
    "/tmp/wolf.txt"
  ],
  "dir" : [ ]
},
"repodata" : {
  "filelists" : BinData(0,"eJwtTlsKwyAQ/PcUyxwglTkgfudtJN2kEqNipQ2E3L3adh9mmBdrshsXNwu5Mj4ZMXUERbcK45PCB
MrL7B+MfT8ZHQesonbmLeXlUyTJqZcvoCKBcQUlh3Efbt1RmefBCr65plnx3qVo3+acro/xMt+gX/mixB"),
  "other" : BinData(0,"eJyzKUHmzk5MT1VLErOsFXKywFRSgpp5ibmptkrl+TlpSgoF2emZKbZK1dUQlkJtrZIdlwIQ2JSlFhVn5
ucppBbkgzQbKCKUpebYKhkpKQBlbJUs9UyUFPTtuLhs9KG2ANKAnrUiXw=="),
  "primary" : BinData(0,"eJx1U8Fu2zAMvfcrBN1jLU7WJYUtoIdiGLBDh5121GQ6FiJbGiUnyIr++yLhoMBPZ16fKQeH+XKK31
WJ2Dx5qHm6HsuHxirBtWDvDrbVmIKE6ZQd3Jw6VOJ6ZDQC2AwbmDgne5q/okzBFvzkjPK1PxY7DkTE1N3om9h7Jk/n0xT818vP/1879sbW7IjY
O/vXN5hdK7Ecph60bdXeJPPrKHoxpYxXMuy6oWQyA0EjczHkvlxwT0pFc0EnMWPagUXo38SIOaISncuqqKFxqHy4LyFwuGpEomX+NH0wH6PxtK
g29lue9htd49kSmssLEi5T0juH8xfYm1Uc6F8eXzKzAwHkmuBOuyJNwui5H53nMoq67RKg1mH0Nb8NTOCuIo004a835RF3liRVpukah32KqaQD
gQ/WaNHCCC/vn6/lJW4hlbSBQaaNaudoRO60ctvQwQcIIpn76lsEhRy10xYCyY/OheIDP0f2FzD9nMmromVHNyIGiisD8ME1GmSXLUy1qoOVAO
4QTXQckky2VUeyC6yEmPND18Wx2e+R3cxtPsMzSAMEW93L7q16hRq/vKDs/Q31Dwp+vChUxPxf+tKLK5XY16kfPgHDOQdSQ==")
},
"description" : "A dummy package of wolf",
"header_range" : {
  "start" : 872,
  "end" : 2289
},
"sourcerpm" : "wolf-9.4-2.src.rpm",
"license" : "GPLv2",
"changelog" : [ ],
"url" : "http://tstrachota.fedorapeople.org",
"summary" : "A dummy package of wolf",
"time" : 1331832462,
"requires" : [ ],
"recommends" : [ ],
"_ns" : "units_rpm",
"_content_type_id" : "rpm",
"is_modular" : false
}

```

This issue might affect manifested behaviour <https://pulp.plan.io/issues/4059>

#6 - 05/14/2019 04:44 PM - kravir

The repo dev-epel-7-x86_64 is an existing repo, I did not create this repo

1. pulp-admin rpm repo list --repo-id dev-epel-7-x86_64 --details

RPM Repositories

Id: dev-epel-7-x86_64
Display Name: None
Description: None
Content Unit Counts:
Drpm: 5732
Erratum: 6391
Package Category: 5
Package Environment: 2
Package Group: 210
Rpm: 22362
Notes:
Scratchpad:
Importers:
Config:
Remove Missing: True
Id: yum_importer
Importer Type Id: yum_importer
Last Override Config:
Last Sync: 2019-05-08T18:05:38Z
Last Updated: 2017-12-29T03:35:04Z
Repo Id: dev-epel-7-x86_64
Scratchpad: None
Distributors:
Auto Publish: True
Config:
Http: False
Https: True
Relative URL: dev-epel-7-x86_64
Distributor Type Id: yum_distributor
Id: yum_distributor
Last Override Config:
Last Publish: 2019-05-08T19:18:13Z
Last Updated: 2017-12-29T03:35:04Z
Repo Id: dev-epel-7-x86_64
Scratchpad:
Auto Publish: False
Config:
Http: False
Https: True
Relative URL: dev-epel-7-x86_64
Distributor Type Id: export_distributor
Id: export_distributor
Last Override Config:
Last Publish: None
Last Updated: 2017-12-29T03:35:04Z
Repo Id: dev-epel-7-x86_64
Scratchpad:

#7 - 05/15/2019 04:47 PM - ttereshc

- Triaged changed from No to Yes

#8 - 05/15/2019 04:51 PM - ttereshc

- Category deleted (22)

#9 - 05/15/2019 04:57 PM - ttereshc

- Sprint set to Sprint 53

#10 - 05/15/2019 10:21 PM - ggainey

Been working on trying to reproduce this, and have had no luck to date. ipanova, here's the set of steps I'm using to attempt to reproduce - am I missing something critical?

```
# pstatus
# pulp-admin -u admin -p admin rpm repo create --repo-id 4798 --relative-url 4798 --feed \
  https://repos.fedorapeople.org/repos/pulp/pulp/fixtures/rpm-with-modules/ --download-policy on_demand
# pulp-admin rpm repo sync run --repo-id 4798
# curl -L -O -k -v https://repos.fedorapeople.org/repos/pulp/pulp/fixtures/rpm-with-modules/wolf-9.4-2.noarch.
rpm
# pulp-admin -u admin -p admin rpm repo uploads rpm --file ./wolf-9.4-2.noarch.rpm --repo-id 4798
# pulp-admin rpm repo content rpm --repo-id 4798 --match 'filename=wolf-9.4-2.noarch.rpm'
Arch:          noarch
Buildhost:     smqe-ws15
Checksum:      a42b42020d3f3eefc4421d89ce341a7b9f9293c1b4bdea33bb855a6fd7cce6f2
Checksumtype: sha256
Description:   A dummy package of wolf
Epoch:        0
Filename:      wolf-9.4-2.noarch.rpm
License:       GPLv2
Name:          wolf
Provides:      wolf = 9.4-2-0
Recommends:
Release:       2
Requires:
Version:       9.4
```

kravir - you said you hadn't created the dev-epel repo. Is this a new instance of pulp, or has it been up for some time? Specifically, has it been upgraded over time?

#11 - 05/16/2019 04:08 PM - kravir

ggainey - Our Pulp server was setup by an employee who recently left the company, the server has been up for a long time. The server was patched/updated recently (a month ago) and I think that has caused this issue. What we can do to fix it?

#12 - 05/16/2019 07:53 PM - ggainey

- Related to Issue #4059: During rpm upload filename should be preserved in the storage path added

#13 - 05/16/2019 09:22 PM - ggainey

kravir - investigation continues, but here's the current theory:

The first problem is that the feed-url being used for this repo is no longer serving content (<http://epel.mirrors.ovh.net> is alive but apparently empty) This is then exacerbated by issue#4059 - when you upload, the upload-process fails at the convert-repo-id-to-rpm-name step. These two together cause the retrieve to fail. I've linked these two issues, and will be working on 4059.

One can use `pulp-admin rpm repo update --repo-id dev-epel-7-x86_64 --feed <new-working-epel-url>` to change the feed-url - HOWEVER, that won't quite do for an `on_demand` repo due to issue#4265. Fun!

You can fix things up, after you set a valid feed-url on that repo, using the steps described in 4265 to cause streamer to forget everything it thinks it knows about that repo and rebuild:

1. Change the feed-url: `pulp-admin rpm repo update --repo-id dev-epel-7-x86_64 --feed <new-working-epel-url>`
2. Enter mongo-db shell: `mongo pulp_database`
3. Find the importer-id of the repo: `db.repo_importers.find({'repo_id': 'dev-epel-7-x86_64'})` - look for the ObjectId
4. Delete the lazy_content-catalog entries for that importer: `db.lazy_content_catalog.deleteMany({'importer_id': '<object-id>'})` On my system, for example, the cmd and results look like this:

```
> db.lazy_content_catalog.deleteMany({'importer_id': '5cdc701c30f25233c9d60c5d'})
{ "acknowledged" : true, "deletedCount" : 35 }
```

5. Force a full resync of of the repo: `pulp-admin -u admin -p admin rpm repo sync run --force-full --repo-id dev-epel-7-x86_64`

This **should** leave you with a working, `on_demand` repo. Let us know here how it goes!

ipanova - more eyes please, on the instructions above

#14 - 05/17/2019 02:13 PM - ipanova@redhat.com

[ggainey](#) that sounds correct to me.

#15 - 05/17/2019 07:28 PM - kravir

[ggainey](#) thanks for your reply and steps to fix the issue. I remember running sync for my repo but I had got this

```
# pulp-admin rpm repo sync run --repo-id dev-epel-7-x86_64
+-----+
|           Synchronizing Repository [dev-epel-7-x86_64]           |
+-----+
```

This command may be exited via ctrl+c without affecting the request.

Task Failed

Unable to sync a repository that has no feed.

My understanding is if repo is a non-feed repo then it should not download from an upstream repo but why is that happening here in first place. I thought if I am getting an rpm file from any xyz source and uploading/publishing that file to a non-feed repo then that rpm should just become available for download by my machines, but that is not happening. Maybe I'm missing some Pulp concepts here.

#16 - 05/30/2019 10:40 PM - ggainey

kravir wrote:

[ggainey](#) thanks for your reply and steps to fix the issue. I remember running sync for my repo but I had got this

[...]

My understanding is if repo is a non-feed repo then it should not download from an upstream repo but why is that happening here in first place. I thought if I am getting an rpm file from any xyz source and uploading/publishing that file to a non-feed repo then that rpm should just become available for download by my machines, but that is not happening. Maybe I'm missing some Pulp concepts here.

I think your understanding of expected-behavior is correct, but this particular repo is in a seriously not-typical state, so all bets are off.

Your --details report from c6 shows no Importer stanza - i.e., no feed. However, the errors you reported in the initial report, from /var/log/messages, includes

```
May  8 22:42:17 ip-10-222-253-xx *pulp_streamer*: pulp.streamer.server:INFO: Download failed [404]:
*http://epel.mirrors.ovh.net/epel/7/x86_64/Packages/p/python2-markdown-2.4.1-4.el7.noarch.rpm*
```

With pulp_streamer is involved, it means that the repo was initially set up as on-demand and pointing to epel.mirrors.ovh.net

My **new** current theory (which includes guesswork on what people not-you may have done in your shop two years ago, so take it with a certain number of grains of salt), is that:

- initial setup (sometime in 2017) was as an on-demand feed pointing at epel.mirrors.ovh.net
- epel.mirrors.ovh.net stopped serving content (which is the state it is in now)
- feed was set to null - but, as noted in c13, **that won't change pulp_streamer's idea of where to get content** because of [#4265](#)
- You uploaded the RPM 'directly', and here [#4059](#) fails us.
- attempting to access the uploaded-RPM-NEVRA passed to pulp_streamer, who ignores the upload and tries to access from epel.mirrors.ovh.net, and 404s happen.

Now, this is a fine theory, and would explain what I'm seeing from the logs and details and cmd-output you've shared with us. **However**, it doesn't explain why the current --details from this repo shows no Importer stanza at all - I have not been able to recreate that output, starting from a repo that has ever had a feed. So I'm still puzzled, and not certain we've really got our hands around all the issues we're having.

My current thinking basically boils down to "this repo is borked, and I'm not sure we're going to know exactly how it all happened". In addition, the 4059/4265 pair are going to make it hard to recover it 'correctly'. Assuming all of the/our/my assumptions are correct, **and** assuming that you really do just want a local mirror of epel-7-x86_64, you are probably best served by recreating the repo from scratch, using commands something like the following, which worked in my test-env:

```
# pulp-admin -u admin -p admin rpm repo create \
  --repo-id epel-7-x86_64 \
  --relative-url epel-7-x86_64 \
  --feed https://dl.fedoraproject.org/pub/epel/7/x86_64/ \
  --download-policy on_demand
# pulp-admin rpm repo sync run --repo-id epel-7-x86_64
# pulp-admin rpm repo content rpm --repo-id epel-7-x86_64 \
  --match 'filename=python2-markdown-2.4.1-4.el7.noarch.rpm'
Arch:          noarch
Buildhost:     buildvm-ppc64-09.ppc.fedoraproject.org
Checksum:      797ab5e852d39cfbf130d28f8cdd0dd52045c696dce883662abda7b8475ae9bd
Checksumtype:  sha256
Description:   This is a Python implementation of John Gruber's Markdown. It is
               almost completely compliant with the reference implementation,
               though there are a few known issues.
Epoch:        0
Filename:      python2-markdown-2.4.1-4.el7.noarch.rpm
License:       BSD
Name:          python2-markdown
Provides:      python-markdown = 2.4.1-4.el7-0, python2-markdown = 2.4.1-4.el7-0
Recommends:
Release:       4.el7
Requires:      /usr/bin/python2, python(abi) = 2.7-0, python2
Vendor:        Fedora Project
Version:       2.4.1
# curl -L -k -v -o markdown.rpm https://localhost/pulp/repos/epel-7-x86_64/Packages/p/python2-markdown-2.4.1-4
.el7.noarch.rpm
[...lots of curl-debug here...]
# file markdown.rpm
markdown.rpm: RPM v3.0 bin noarch python2-markdown-2.4.1-4.el7
#
```

Meanwhile, I'll start working on fixes for the two bugs that are actually biting us here.

@dkliban, @ipanova, more eyes would be welcome please.

[note: apologies for not responding sooner. I failed to watch this issue when responding, so I didn't get a notification about your last comment - and then went on PTO for a week]

#17 - 06/03/2019 04:29 PM - ipanova@redhat.com

Your understanding is correct.

From what i see, I could tell that: the repo was created with on demand policy and a feed. It got synced, lazy catalog got created. Catalog is bound to the importer.

Importer got updated, as Grant said, feed got removed, probably because it stopped working? Note that the repo was created in 2017! Lazy catalog is still bound to the importer.

Your clients get 404 because lazy catalog points to invalid URLs, as a bonus, because of the filename issue you hit this

<https://pulp.plan.io/issues/4798#note-5>, so upload does not help.

As suggested previously , if you would like to sync in the content - remove the catalog and update the feed option with a working URL. Trigger sync with a force-full option.

Meanwhile we will fix the filename and lazy catalog issues.

#18 - 06/04/2019 04:48 PM - amacdona@redhat.com

- Sprint changed from Sprint 53 to Sprint 54

#19 - 06/21/2019 03:26 PM - ttereshc

- Sprint changed from Sprint 54 to Sprint 55

#20 - 07/12/2019 03:27 PM - dkliban@redhat.com

- Sprint changed from Sprint 55 to Sprint 56

#21 - 08/02/2019 03:14 PM - rchan

- Sprint changed from Sprint 56 to Sprint 57

#22 - 08/26/2019 04:30 PM - rchan

- Sprint changed from Sprint 57 to Sprint 58

#23 - 09/17/2019 09:02 PM - rchan

- Sprint deleted (Sprint 58)