

Pulp - Issue #9215

Cache should be invalidated for reclaimed repos

08/06/2021 03:24 PM - ipanova@redhat.com

Status: MODIFIED	Start date:
Priority: Normal	Due date:
Assignee: gerrod	Estimated time: 0:00 hour
Category:	
Sprint/Milestone:	
Severity: 2. Medium	Groomed: No
Version:	Sprint Candidate: No
Platform Release:	Tags:
OS:	Sprint: Sprint 107
Triaged: Yes	Quarter:

Description

1. create, sync, distribute repo
2. fetch some content from it
3. run reclaim space on that repo
4. fetch same content from it
5. observe 502

```
$ pwd
/home/vagrant/devel/pulp_rpm/docs/_scripts
(pulp) [vagrant@pulp3-source-fedora34 _scripts]$ source docs_check_sync.sh

pulp rpm repository list
[
  {
    "pulp_href": "/pulp/api/v3/repositories/rpm/rpm/ff61861d-fc67-4741-8086-7ccbea2632a6/",
    "pulp_created": "2021-08-06T13:01:12.513113Z",
    "versions_href": "/pulp/api/v3/repositories/rpm/rpm/ff61861d-fc67-4741-8086-7ccbea2632a6/versions/",
    "pulp_labels": {},
    "latest_version_href": "/pulp/api/v3/repositories/rpm/rpm/ff61861d-fc67-4741-8086-7ccbea2632a6/versions/1/",
    "name": "sync-repo",
    "description": null,
    "retain_repo_versions": null,
    "remote": null,
    "autopublish": false,
    "metadata_signing_service": null,
    "retain_package_versions": 0,
    "metadata_checksum_type": null,
    "package_checksum_type": null,
    "gpgcheck": 0,
    "repo_gpgcheck": 0,
    "sqlite_metadata": false
  }
]
```

```
(pulp) [vagrant@pulp3-source-fedora34 _scripts]$ http POST :24817/pulp/api/v3/repositories/reclaim_space/ repo_hrefs=["/pulp/api/v3/repositories/rpm/rpm/ff61861d-fc67-4741-8086-7ccbea2632a6/\"]

$ wget https://pulp3-source-fedora34.fluffy.example.com/pulp/content/sync-repo/Packages/f/fox-1.1-2.noarch.rpm --no-check-certificate
--2021-08-06 13:07:54-- https://pulp3-source-fedora34.fluffy.example.com/pulp/content/sync-repo/Packages/f/fox-1.1-2.noarch.rpm
Resolving pulp3-source-fedora34.fluffy.example.com (pulp3-source-fedora34.fluffy.example.com)... 1
```

```
27.0.0.1, 192.168.122.186
Connecting to pulp3-source-fedora34.fluffy.example.com (pulp3-source-fedora34.fluffy.example.com) |
127.0.0.1|:443... connected.
WARNING: The certificate of 'pulp3-source-fedora34.fluffy.example.com' is not trusted.
WARNING: The certificate of 'pulp3-source-fedora34.fluffy.example.com' doesn't have a known issuer
.
HTTP request sent, awaiting response... 502 Bad Gateway
2021-08-06 13:07:54 ERROR 502: Bad Gateway.
```

502 will be received for previously created distributions until the cache will expire. After 10 mins I can normally retrieve content. 502 is received regardless of the download policy. For newly created distributions after reclaim content is normally served, only old distributions are affected.

```
wget https://pulp3-source-fedora34.fluffy.example.com/pulp/content/sync-repo/Packages/f/fox-1.1-2
.noarch.rpm --no-check-certificate
--2021-08-06 13:08:21-- https://pulp3-source-fedora34.fluffy.example.com/pulp/content/sync-repo/P
ackages/f/fox-1.1-2.noarch.rpm
Resolving pulp3-source-fedora34.fluffy.example.com (pulp3-source-fedora34.fluffy.example.com)... 1
27.0.0.1, 192.168.122.186
Connecting to pulp3-source-fedora34.fluffy.example.com (pulp3-source-fedora34.fluffy.example.com) |
127.0.0.1|:443... connected.
WARNING: The certificate of 'pulp3-source-fedora34.fluffy.example.com' is not trusted.
WARNING: The certificate of 'pulp3-source-fedora34.fluffy.example.com' doesn't have a known issuer
.
HTTP request sent, awaiting response... 200 OK
```

From IRC conversations:

```
ipanova
do i understand correctly that for repos which are not auto-distributed I need to find all the rep
o_versions, find all it's publications and distributions that are serving the version, retrieve th
e base_path and only then i can invalidate the cache?
repo.distributions.all() for such repos will be empty qs
and so that's why all these dancing around ^^
basically this for each such repo that is being touched by reclaim task:
```

```
versions = repo.versions.all()
pubs = Publication.objects.filter(repository_version__in=versions, complete=True)
paths = Distribution.objects.filter(publication__in=pubs).values_list("base_path", flat=True)
Cache().delete(base_key=paths)
```

```
I mean, this does work
iI get 404 right away
it's just i was hoping repo.invalidate_cache() would do it
but nope
```

```
dralley
right :/ yeah, it's irritating
```

Associated revisions

Revision 7ea58ed1 - 10/18/2021 11:17 PM - gerrod

Invalidate cache after reclaim disk

fixes: #9215

History

#1 - 08/10/2021 04:47 PM - dkliban@redhat.com

- Triaged changed from No to Yes

- Sprint set to Sprint 102

#2 - 08/11/2021 01:25 PM - ipanova@redhat.com

- Status changed from NEW to ASSIGNED

- Assignee set to ipanova@redhat.com

#3 - 08/12/2021 05:23 PM - rchan

- Sprint changed from Sprint 102 to Sprint 103

#4 - 08/17/2021 02:22 PM - ipanova@redhat.com

- Status changed from ASSIGNED to NEW

- Assignee deleted (ipanova@redhat.com)

#5 - 08/27/2021 05:08 PM - rchan

- Sprint changed from Sprint 103 to Sprint 104

#6 - 09/10/2021 12:28 AM - rchan

- Sprint changed from Sprint 104 to Sprint 105

#7 - 09/23/2021 11:54 PM - rchan

- Sprint changed from Sprint 105 to Sprint 106

#8 - 10/08/2021 03:16 PM - rchan

- Sprint changed from Sprint 106 to Sprint 107

#9 - 10/12/2021 04:40 PM - gerrod

- Assignee set to gerrod

#10 - 10/12/2021 05:27 PM - pulpbot

- Status changed from NEW to POST

PR: <https://github.com/pulp/pulpcore/pull/1682>

#11 - 10/18/2021 11:17 PM - gerrod

- Status changed from POST to MODIFIED

Applied in changeset [pulpcore|7ea58ed1cfccc57a54d78f39a657865c0d7bb202](#).