

File Support - Task #4028

Prevent duplicate files in repositories

09/19/2018 09:24 PM - daviddavis

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:	bmbouter	% Done:	100%
Category:		Estimated time:	0:00 hour
Sprint/Milestone:	0.1.0	Tags:	
Platform Release:		Sprint:	Sprint 60
Groomed:	No	Quarter:	
Sprint Candidate:	No		

Description

Prevent the possibility of duplicate files (ie files with the same relative_path) in a single repo version. I think the places we need to check for dupes are after sync and associate.

Steps to reproduce:

```
http POST http://localhost:24817/pulp/api/v3/repositories/ name=foo
export REPO_HREF=$(http :24817/pulp/api/v3/repositories/ | jq -r '.results[] | select(.name == "foo") | ._href')
echo "blah1" > 1.iso
echo "blah2" > 2.iso
export ART1=$(http --form POST http://localhost:24817/pulp/api/v3/artifacts/ file@./1.iso | jq -r '_href')
export ART2=$(http --form POST http://localhost:24817/pulp/api/v3/artifacts/ file@./2.iso | jq -r '_href')
export CU1=$(http POST http://localhost:24817/pulp/api/v3/content/file/files/ relative_path=1.iso _artifact="$ART1" | jq -r '_href')
export CU2=$(http POST http://localhost:24817/pulp/api/v3/content/file/files/ relative_path=1.iso _artifact="$ART2" | jq -r '_href')
http POST ':24817'$REPO_HREF'versions/' add_content_units="[\ "$CU1\", \ "$CU2\"]"
sleep 5
http ":24817${REPO_HREF}versions/1/"
```

You'll see the count is 2 which indicates that there are two 1.iso files in the repository.

Related issues:

Related to File Support - Test #4738: File repositories shouldn't have duplic...

NEW

Related to Pulp - Story #5008: As a user, I can enforce content specific uniq...

CLOSED - CURRENTRELEASE

Blocked by Pulp - Story #3934: As a plugin writer, I can have a stage that re...

CLOSED - CURRENTRELEASE

Associated revisions

Revision 2551ede9 - 10/08/2019 08:54 PM - bmbouter

Prevent duplicate relative_path files

When a RepositoryVersion has another file.file added with the same relative_path the original one should be removed. This is shown correctly on the RepositoryVersion added, removed fields.

This adds a test which uploads two distinct Artifacts, makes them into File content, and then associates one, and then another with a Repo. The second RepositoryVersion should have only 1 file.file content unit in it.

Required PR: <https://github.com/pulp/pulpcore/pull/331>

<https://pulp.plan.io/issues/4028> closes #4028

History

#1 - 09/19/2018 09:24 PM - daviddavis

- Blocked by Story #3934: As a plugin writer, I can have a stage that removes duplicates added

#2 - 09/20/2018 02:19 PM - daviddavis

- Tracker changed from Issue to Task
- % Done set to 0

#3 - 09/20/2018 09:53 PM - daviddavis

- Subject changed from Remove duplicate files from repositories after sync to Prevent duplicate files in repositories
- Description updated

#4 - 04/25/2019 02:12 PM - daviddavis

- Description updated

#5 - 04/25/2019 02:16 PM - daviddavis

When publishing we sort by created when building the manifest file[0] as [#4731](#) points out. We should look at removing this code since it won't be needed anymore.

[0] https://github.com/pulp/pulp_file/blob/dd366601de3ae8741a7f0c2ee8f288f90f74d142/pulp_file/app/tasks/publishing.py#L73-L75

#6 - 04/25/2019 02:23 PM - daviddavis

- Related to Test #4738: File repositories shouldn't have duplicate files added

#7 - 04/25/2019 10:33 PM - daviddavis

- Blocked by Issue #4740: Pulpcore doesn't provide a way to guarantee uniqueness in repo versions added

#8 - 04/26/2019 12:42 AM - daviddavis

- Blocked by deleted (Issue #4740: Pulpcore doesn't provide a way to guarantee uniqueness in repo versions)

#9 - 04/26/2019 12:42 AM - daviddavis

- Blocked by Issue #3541: Core should not add/remove content to a repository or create a repository_version without plugin input added

#10 - 04/26/2019 10:33 PM - bmbouter

- Tags deleted (Pulp 3)

#11 - 06/21/2019 04:34 PM - ttereshc

- Related to Story #5008: As a user, I can enforce content specific uniqueness constraints in a repo version added

#12 - 06/21/2019 04:48 PM - rchan

Did not move from Sprint 54 to 55 because this issue needs some planning/grooming and not high enough priority to be done in next 3 week sprint.

#13 - 10/01/2019 06:48 PM - bmbouter

- Status changed from NEW to ASSIGNED
- Assignee set to bmbouter
- Sprint set to Sprint 60

Taking as assigned so I can prove pulpcore's implementation as part of fixing 5008. Adding to sprint so everyone is aware.

#14 - 10/01/2019 07:23 PM - daviddavis

- Sprint/Milestone set to 0.1.0

#15 - 10/03/2019 06:08 PM - daviddavis

- Blocked by deleted (Issue #3541: Core should not add/remove content to a repository or create a repository_version without plugin input)

#16 - 10/08/2019 12:05 AM - bmbouter

- Status changed from ASSIGNED to POST

PR available at: https://github.com/pulp/pulp_file/pull/290

#17 - 10/08/2019 09:52 PM - bmbouter

- Status changed from POST to MODIFIED

- % Done changed from 0 to 100

Applied in changeset [2551ede93a5a21df4a1ff7beb114e24cef377a79](#).

#18 - 12/13/2019 06:38 PM - bmbouter

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE