

Pulp - Story #9574

The validate_duplicate_content function should provide more information in the failure case.

11/17/2021 11:28 AM - quba42

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

Description

Ticket moved to GitHub: "pulp/pulpcore/2062":<https://github.com/pulp/pulpcore/issues/2062>

When the validate_duplicate_content function finds illegal duplicate content in a repo version being created, the output is (pulp_deb example):

```
"Cannot create repository version. More than one deb.package content with the duplicate values for package, version, architecture."
```

For users to have any chance of debugging this situation, it would be vital for the error to provide them with a list of the offending duplicate units, preferably the pulp_href, so they can go and look at them in detail.

Without this information I just know "I have duplicate units somewhere in the potentially tens of thousands of units in the repo version being created". (Since the repo version is then not created, I can't even go hunting for the duplicate units myself...) Right now, I can't even distinguish a situation where two packages are clashing, from one where all my packages are double (for example).

User reported backtrace for the error they encountered:

```
File \"/usr/lib/python3.6/site-packages/pulpcore/tasking/pulpcore_worker.py\", line 317, in _perform_task
    result = func(*args, **kwargs)
File \"/usr/lib/python3.6/site-packages/pulpcore/app/tasks/repository.py\", line 219, in add_and_remove
    new_version.add_content(models.Content.objects.filter(pk__in=add_content_units))
File \"/usr/lib/python3.6/site-packages/pulpcore/app/models/repository.py\", line 963, in __exit__
    repository.finalize_new_version(self)
File \"/usr/lib/python3.6/site-packages/pulp_deb/app/models/repository.py\", line 57, in finalize_new_version
    validate_repo_version(new_version)
File \"/usr/lib/python3.6/site-packages/pulpcore/plugin/repo_version_utils.py\", line 137, in validate_repo_version
    validate_duplicate_content(version)
File \"/usr/lib/python3.6/site-packages/pulpcore/plugin/repo_version_utils.py\", line 108, in validate_duplicate_content
    _(\"Cannot create repository version. {msg}\").format(msg=\"\", \".join(error_messages))
```

History

#1 - 11/17/2021 11:35 AM - quba42

Directly printing the "package, version, architecture" (in the above example) along with their pulp_href for the duplicate units would be even nicer. Otherwise users will report a bunch of pulp_href to us, and we still won't be any wiser unless they also go and retrieve those units for us.

#2 - 11/23/2021 05:00 PM - dkliban@redhat.com

- Tracker changed from Issue to Story
- % Done set to 0
- Severity deleted (2. Medium)
- Triaged deleted (No)

#3 - 01/17/2022 05:45 PM - fao89

- *Description updated*

- *Status changed from NEW to CLOSED - DUPLICATE*