

Pulp - Story #4209

Story # 3693 (CLOSED - CURRENTRELEASE): Lazy for Pulp3

As a plugin writer, I can implement lazy sync for plugins with multiple metadata files

11/28/2018 08:04 PM - amacdona@redhat.com

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:	mdellweg	% Done:	0%
Category:		Estimated time:	0:00 hour
Sprint/Milestone:	3.0.0	Tags:	
Platform Release:		Sprint:	
Groomed:	No	Quarter:	
Sprint Candidate:	No		

Description

Problem

For plugins such as Python and Docker that have many metadata files, we may need a different way to implement lazy sync than the current plan, which is to omit the ArtifactDownloader stage and the ArtifactSaver stage.

Because there are many metadata files, (and with Docker, these metadata files are also content units), we are using the ArtifactDownloader stage to download metadata. In this case, the metadata files need to be downloaded, and related content saved, but the non-metadata files should be skipped for downloading and saving.

Solution

The DeclarativeArtifact gets a new flag called 'deferred_download' defaulting to False to specify, that the artifact stages stop processing it (making it lazy).

At the same time the 'download_artifacts' flag of DeclarativeVersion is removed.

So the plugin must either use the deferred_download flag, or tailor its own Pipeline.

Related issues:

Related to Python Support - Story #1884: As a user, I can lazily sync python ...	MODIFIED
Related to Container Support - Story #4174: As a user, I can lazy sync a dock...	CLOSED - CURRENTRELEASE
Related to RPM Support - Issue #4446: multiple test failures on master branch	CLOSED - CURRENTRELEASE

Associated revisions

Revision 0cd2645f - 02/21/2019 11:57 AM - mdellweg

Use deferred_download flag for lazy feature

Required PR: <https://github.com/pulp/pulpcore-plugin/pull/50>

<https://pulp.plan.io/issues/4209> re #4209

Revision 354aa520 - 02/21/2019 10:27 PM - bmbouter

Updating lazy for pulp_rpm

The changes to the plugin API have changed the DeclarativeVersion interface. We now set deferred_download on the DeclarativeArtifact directly.

<https://pulp.plan.io/issues/4209> re #4209

History

#1 - 12/04/2018 11:01 PM - dalley

- Sprint set to Sprint 46

- Tags Pulp 3 RC Blocker added

I'm pretty sure we meant to make this an RC blocker but forgot to do so. Adding to the sprint because that's what we said we would do with all RC blockers IIRC.

#2 - 12/04/2018 11:05 PM - dalley

- Related to Story #1884: As a user, I can lazily sync python packages added

#3 - 12/05/2018 04:04 PM - amacdona@redhat.com

- Related to Story #4174: As a user, I can lazy sync a docker repository added

#4 - 12/12/2018 05:17 PM - bmbouter

- Sprint deleted (Sprint 46)

I have to comment on this as part of the lazy epic, but since there are two options and it's not clear which we are doing I'm going to remove from the sprint for right now. At the RC blocker meeting we wanted work on the sprint to be ready-to-work. Once some more comments happen we will put it back on the sprint (very soon).

#5 - 01/03/2019 05:38 PM - bmbouter

I agree with the problem this story identifies. We need a way for plugin to add lazy support while still using the ArtifactDownloader and ArtifactSaver stages in the pipeline.

There is a related problem that we need to consider at the same times as this one because I believe one solution will resolve all of these issues. Plugins that discover content to download from content downloaded by the pipeline would need additional ArtifactDownloader and ArtifactSaver stages later in the pipeline. Adding more stages later fixes that issue, but since the pipeline size has to be fixed and the number of "discoveries" does not this solution can't address all plugin needs. For example the Maven sync will have this issue.

I'm writing ^ issue up as a separate ticket now and after that I can show how that solution will also resolve this problem as well.

#6 - 01/10/2019 05:57 PM - jortel@redhat.com

Option 1 seems simplest only let's name it something like *DeclarativeArtifact.always_download* instead. This way the attribute describes the desired behavior instead of describing the artifact which has the side effect of the desired behavior.

#7 - 01/11/2019 02:08 PM - mdellweg

Just to be explicit here: Please ensure, that this is handled per Artifact. In gem, a content unit consists of two artifacts: One containing the actual gem (subject to lazy) and one containing Metadata (which in a future version might be desirable to be available immediately for dependency resolution).

BTW: Option 1 +1

#8 - 01/11/2019 08:34 PM - mdellweg

If you decide for Option 1, i have a solution ready:

https://github.com/ATIX-AG/pulpcore-plugin/tree/non_lazy_artifacts

#9 - 02/13/2019 03:49 PM - bmbouter

- Status changed from NEW to ASSIGNED

- Assignee set to mdellweg

Even though the full design isn't written out here, it is being discussed here: <https://github.com/pulp/pulpcore-plugin/pull/50>

So others don't start on the work I'm marking it as assigned now.

#10 - 02/21/2019 03:03 PM - mdellweg

The solution agreed upon now is:

The DeclarativeArtifact gets a new flag called 'deferred_download' defaulting to False to specify, that the artifact stages stop processing it (making it lazy).

At the same time the 'download_artifacts' flag of DeclarativeVersion is removed.

So the plugin must either use the deferred_download flag, or tailor its own Pipeline.

#11 - 02/21/2019 05:04 PM - bmbouter

- Description updated

#12 - 02/22/2019 04:41 PM - bmbouter

- Status changed from ASSIGNED to POST

PRs available: https://github.com/pulp/pulp_file/pull/177 and <https://github.com/pulp/pulpcore-plugin/pull/50>

#13 - 02/22/2019 04:46 PM - bmbouter

- Status changed from POST to MODIFIED

[mdellweg](#) can you closes #4209 with the docs PR you are making please?

#14 - 02/26/2019 02:03 PM - ttereshc

- Related to Issue #4446: multiple test failures on master branch added

#15 - 03/04/2019 05:26 PM - mdellweg

[bmbouter](#) Adding the closes tag brings travis to fail with:

```
Error: issue #4209 has invalid status of MODIFIED. Status must be one of NEW, ASSIGNED, POST.
```

Removed it again.

#16 - 04/25/2019 06:44 PM - daviddavis

- Sprint/Milestone set to 3.0.0

#17 - 04/26/2019 10:33 PM - bmbouter

- Tags deleted (Pulp 3, Pulp 3 RC Blocker)

#18 - 12/13/2019 06:27 PM - bmbouter

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE