Pulp - Story #3907

As a plugin writer, the HttpDownloader provides auto_decompress=True like aiohttp

08/08/2018 12:04 AM - bmbouter

Status:	CLOSED - WONTFIX		**Start date:**	
Priority:	Normal		**Due date:**	
Assignee:	bmbouter		**% Done:**	100%
Category:			**Estimated time:**	0:00 hour
Sprint/Milestone:	3.0.0			
Platform Release:			**Tags:**	
Groomed:	No		**Sprint:**	Sprint 41
Sprint Candidate:	No		**Quarter:**	

Description

Problem Statement

aiohttp by default decompresses everything it downloads, e.g. gzip, or other compression types. This is great for typical clients which would otherwise have to decompress downloaded data, e.g. downloading a foo.tar.gz file, when reading the data from aiohttp you'regetting the actual contents of foo, not the foo.tar.gz compressed data.

Pulp needs to store data exactly as it was stored remotely. To uncompress the data and save it in Pulp's backend is incorrect because Pulp things of that file as foo.tar.gz, not foo. When it presents it to client's it's presented as foo.tar.gz. If we uncompress it, but then save it as foo.tar.gz (incorrect) then clients will in turn be receiving foo.tar.gz and go to uncompress it (a second time) and that will fail.

There are other situations though where we want the auto_decompression. When downloading temporary metadata for example, often times remote metadata is compressed, and we need to download, uncompress it, and then have our code parse it. So auto-decompression is really useful in some cases.

Solution

Enable Pulp's downloaders to have auto decompression by default, but have the option to turn it off and download binary data.

Associated revisions

Revision b1fed0c4 - 08/08/2018 01:03 AM - bmbouter

Add auto_decompress option everywhere

https://pulp.plan.io/issues/3907 closes #3907

Revision b1fed0c4 - 08/08/2018 01:03 AM - bmbouter

Add auto_decompress option everywhere

https://pulp.plan.io/issues/3907 closes #3907

Revision 655d8e1f - 08/13/2018 09:40 PM - dalley

Revert "Add auto_decompress option everywhere"

History

#1 - 08/08/2018 01:04 AM - dalley

- Status changed from NEW to POST

- Assignee set to bmbouter

https://github.com/pulp/pull/3582

#2 - 08/08/2018 02:56 PM - bmbouter

- Sprint set to Sprint 41

I'm adding to the sprint because this is required to have plugins use DeclarativeVersion.

#3 - 08/08/2018 03:21 PM - bmbouter

- Description updated

#4 - 08/08/2018 04:52 PM - bmbouter

- Status changed from POST to MODIFIED

- % Done changed from 0 to 100

2/2

Applied in changeset pulp|b1fed0c47c2a1407eca90cf4ab41572bc1eb28c2.

#5 - 08/13/2018 09:34 PM - bmbouter

- Status changed from MODIFIED to ASSIGNED

Actually after investigating the original symptoms that motivated this issue I believe we do **not** need to offer this feature. The Content-Encoding response header indicates the encoding of the response by the webserver just before sending it. In all cases we want to have aiohttp automatically decompress this data.

I'm going to negative commit this feature.

#6 - 08/13/2018 09:44 PM - bmbouter

- Status changed from ASSIGNED to POST

negative commit PR: https://github.com/pulp/pulp/pull/3597

#7 - 08/13/2018 09:44 PM - bmbouter

- Status changed from POST to CLOSED - WONTFIX

closing as wontfix since we never implemented it.

#8 - 04/25/2019 06:45 PM - daviddavis

- Sprint/Milestone set to 3.0.0

#9 - 04/26/2019 10:34 PM - bmbouter

- Tags deleted (Pulp 3)