

Pulp - Issue #8670

URL escaping in yarl alters escaped characters

04/30/2021 12:36 AM - dannysauer

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:		Estimated time:	0:00 hour
Category:			
Sprint/Milestone:	3.14.0		
Severity:	2. Medium	Groomed:	No
Version:	Master	Sprint Candidate:	No
Platform Release:		Tags:	
OS:		Sprint:	Sprint 98
Triaged:	Yes	Quarter:	

Description

After signing the URL when using the S3 backend (or presumably Azure), the call to generate the HTTP redirect re-parses the URL returned, sometimes resulting in characters being encoded as HTML entities or previously-encoded entities being decoded. While S3 is ok with ignoring an altered query string, CloudFront expects the entire URL to match what was signed exactly. This wasn't an issue *before* I patched django-storages to actually include the parameters passed in to the `artifact_file.storage.url()` call, but once that patch was applied, query strings were sent back with the URL for content-disposition.

It appears that providing a Yarl URL object to the `raise HTTPFound()` prevents re-parsing, resulting in the query string being left alone. If the object is created using `URL()` with `encoded=True`, that bypasses the reparsing in Yarl, and everyone's much happier. It's also verified to not break regular S3 operation.

Associated revisions

Revision 6066e5b1 - 06/15/2021 03:12 PM - dannysauer

Stabilize URLs before generating redirects

The aio URL object mangles query strings when they're passed in. Specifically, it unescapes http entities which are escaped but which don't need escaped. This is normally ok, but in the case of signed URLs, changing anything about the URL invalidates the signature. By pre-creating a yarl.URL object before raising the redirect exception, the parsing doesn't happen in aiohttp. Passing in the "encoding=True" stops the reparsing when the object is created.

fixes #8670 <https://pulp.plan.io/issues/8670>

History

#1 - 04/30/2021 12:52 AM - pulpbot

- Status changed from NEW to POST

PR: <https://github.com/pulp/pulpcore/pull/1295>

#2 - 04/30/2021 04:38 PM - fao89

- Triaged changed from No to Yes

- Sprint set to Sprint 95

#3 - 04/30/2021 06:14 PM - rchan

- Sprint changed from Sprint 95 to Sprint 96

#4 - 05/14/2021 05:34 PM - rchan

- Sprint changed from Sprint 96 to Sprint 97

#5 - 06/02/2021 06:03 PM - rchan

- Sprint changed from Sprint 97 to Sprint 98

#6 - 06/15/2021 03:12 PM - dannysauer

- Status changed from POST to MODIFIED

Applied in changeset [pulpcore|6066e5b11d8864cddb91164bf56f561230d33aed](#).

#7 - 07/01/2021 04:00 PM - pulpbot

- Sprint/Milestone set to 3.14.0

#8 - 07/01/2021 04:43 PM - pulpbot

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE