

Nectar - Issue #2696

Pulp streamer not 'streaming' content to squid

04/07/2017 06:26 PM - alexjfisher

Status:	CLOSED - CURRENTRELEASE	Start date:	
Priority:	Normal	Due date:	
Assignee:	alexjfisher	Estimated time:	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	2. Medium	Groomed:	No
Version - Nectar:	1.5.3	Sprint Candidate:	No
Platform Release:	2.13.1	Tags:	Pulp 2
Target Release - Nectar:	1.5.4	Sprint:	
OS:	CentOS 7	Quarter:	
Triaged:	No		
Description			
Hi			
I'm trying to get on_demand repos working.			
I've got a very slow internet connection (2Mbit shared), which isn't helping, but I think pulp streamer isn't actually streaming data to squid. Instead, it appears to send all the data to squid after it's fetched it all itself.			
When a client tries to download a large file (eg squashfs.img), the first problem it hits, is the 1 minute default proxy read timeout in apache. After increasing that (ProxySet connectiontimeout=60 timeout=3600), the next thing to timeout waiting for data is squid. It's default read_timeout is 10 minutes.			
My understanding is that pulp_streamer should be streaming data to squid and as long as bytes are being received, squid nor apache should abort with read timeouts.			
Related issues:			
Is duplicate of Nectar - Issue #2235: Download progress won't stream to clients		CLOSED - DUPLICATE	

Associated revisions

Revision f0f66d38 - 04/18/2017 09:45 AM - alexjfisher

Re-enable request streaming

This got accidentally turned off in <https://github.com/pulp/nectar/pull/54>

fixes #2696 <https://pulp.plan.io/issues/2696>

History

#1 - 04/07/2017 08:57 PM - alexjfisher

Sorry, was in a big rush earlier. A few more details...

Pulp version 2.12.1 on Centos 7. Installed using <https://github.com/Katello/puppet-pulp>. Puppet's configuration of apache and squid look reasonable. They match the documentation anyway!

Upstream feed URL is **https://www.mirrorservice.org/sites/mirror.centos.org/7.3.1611/os/x86_64/** and all internet access is via a corporate squid (non-caching, whitelist only) proxy.

Upstream proxy seems to be well behaved (manual downloads through it work fine, if slowly). I can't bypass it, so ruling it out 100% isn't going to be easy.

I've been using curl to test fetching on_demand (and not yet downloaded!) files. Using tcpdump, I can see data flowing from the corporate proxy into pulp_streamer, but I can't see anything between pulp_streamer and pulp's squid instance. Curl sits there receiving 0bps until it either suddenly starts to very quickly download the file, or a read timeout occurs (in apache or squid depending on how I've been fiddling with their read timeout settings.)

Meanwhile (and maybe not that important), every 10 minutes (default setting), the deferred download task thing triggers and tries to download the files

(expecting that squid will probably have cached them). It hasn't, so the task also waits for squid to (hopefully) eventually get given the content from pulp_streamer. Unlike curl, the worker times out after 27 seconds. It tries a few times before giving up. I'm not sure, but I'm hoping each of these tries hasn't actually triggered pulp_streamer into opening another connection to www.mirrorservice.org.

I also played with the download_concurrency setting.

Lowering the server.conf lazy download_concurrency from 5 to 2 seemed to help a bit. I guess I made a bit more out of my limited bandwidth.

Hopefully I've provided enough that someone might be able to reproduce some of these issues? Let me know if there's anything else I can provide or check. I'm afisher on freenode IRC.

Thanks,
Alex

#2 - 04/07/2017 09:07 PM - alexjfisher

- OS set to CentOS 7

#3 - 04/10/2017 11:58 AM - alexjfisher

I have also tried each of the following.

- Using an http feed URL.
- Using an internal feed URL and turning off use of the upstream corporate proxy.
- Using katello's version of python-liburl3 <https://github.com/Katello/katello-packaging/commit/b60f4e8362a66c7864d01c2c10c0c4ff77f30dbe>
- Upgrading python-requests and python-urllib3 packages to fedora 24 versions.
- Replacing squid with varnish.

No improvement. pulp-streamer still doesn't stream :(

#4 - 04/10/2017 01:54 PM - alexjfisher

Regression is in python-nectar I think.

yum downgrade https://repos.fedorapeople.org/repos/pulp/pulp/stable/2.8/7Server/x86_64/python-nectar-1.5.2-1.el7.noarch.rpm

and pulp_streamer streams again.

<https://github.com/pulp/nectar/pull/54/commits/4fe7327c4ad1bb2f8c040b75d80306ee047de615> I guess??

#5 - 04/10/2017 03:16 PM - alexjfisher

<https://github.com/pulp/nectar/pull/57>

#6 - 04/11/2017 04:00 PM - jortel@redhat.com

- Status changed from NEW to POST

- Assignee set to alexjfisher

#7 - 04/11/2017 04:05 PM - alexjfisher

- Version - Nectar set to 1.5.3

#8 - 04/11/2017 04:10 PM - alexjfisher

Other relevant versions are...

python-urllib3 1.10.2-2.el7_1

python-requests 2.6.0-1.el7_1

both from http://mirror.centos.org/centos/7/os/x86_64/Packages/

#9 - 04/11/2017 05:42 PM - alexjfisher

- Is duplicate of Issue #2235: Download progress won't stream to clients added

#11 - 04/18/2017 06:28 PM - alexjfisher

- Status changed from POST to MODIFIED

Applied in changeset [f0f66d38aaaba253b232eafbf5854813484101e1](https://github.com/pulp/pulp/commit/f0f66d38aaaba253b232eafbf5854813484101e1).

#12 - 05/16/2017 02:11 PM - bizhang

- Platform Release set to 2.13.1

- Target Release - Nectar set to 1.5.4

#13 - 05/16/2017 09:37 PM - bmbouter

- Status changed from MODIFIED to CLOSED - DUPLICATE

Since this does seem like a duplicate, I'm going to close it as such. Once the fix for [Issue 2696](#) is into a build, if the reporter (or anyone) can reproduce this issue, please post a comment and we can reopen.

#14 - 05/17/2017 03:14 PM - bmbouter

- Status changed from CLOSED - DUPLICATE to 5

This should not have been closed. It was accidentally closed because of the 'duplicates' relationship with Issue 2235. Hopefully reopening won't re-open both of them.

#15 - 05/17/2017 04:13 PM - alexjfisher

Tested fix by installing latest nectar rpm (on a pulp 2.12.2 system)

yum install https://repos.fedorapeople.org/repos/pulp/pulp/beta/2.13/7/x86_64/python-nectar-1.5.4-1.el7.noarch.rpm

New build contains fix and streaming works fine.

#16 - 05/24/2017 08:04 PM - bizhang

- Status changed from 5 to CLOSED - CURRENTRELEASE

#17 - 04/15/2019 10:19 PM - bmbouter

- Tags Pulp 2 added