Pulp - Issue #1018

nux-dextop-el7-x86_64 repo sync gets stuck

06/01/2015 10:58 AM - Ben.Stanley

Status:	CLOSED - WORKSFORME	**Start date:**	
Priority:	High	**Due date:**	
Assignee:	ipanova@redhat.com	**Estimated time:**	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	3. High	**Groomed:**	No
Version:	2.6.1	**Sprint Candidate:**	No
Platform Release:		**Tags:**	Pulp 2
OS:	RHEL 7	**Sprint:**	
Triaged:	Yes	**Quarter:**	

Description

I have observed that attempts to sync my repo nux-dextop-el7-x86_64 do not complete.
The feed comes from
http://mirror.li.nux.ro/li.nux.ro/nux/dextop/el7/x86_64/

[root@bumblebee]# pulp-admin rpm repo sync run --repo-id=nux-dextop-el7-x86_64
───
Synchronizing Repository [nux-dextop-el7-x86_64]
───

This command may be exited via ctrl+c without affecting the request.

[-]
Waiting to begin...

Nothing appears in the logs.

I know that no sync is progressing because nload tells me there is no network traffic.

The tasks list says

[root@bumblebee ~]# pulp-admin tasks list
Operations: sync
Resources: nux-dextop-el7-x86_64 (repository)
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: b621280e-75b5-42eb-932c-c2e3f8a23cfa

Other tasks referencing this repository are either Cancelled or Successful.

Listing the details of the repo reveals the following:

[root@bumblebee ~]# pulp-admin rpm repo list --details
Id: nux-dextop-el7-x86_64
Display Name: nux-dextop-el7-x86_64
Description: None
Content Unit Counts:
Rpm: 1834
Notes:
Importers:
Config:
Feed: http://mirror.li.nux.ro/li.nux.ro/nux//dextop/el7/x86_64

Validate: True
Id: yum_importer
Importer Type Id: yum_importer
Last Sync: None
Repo Id: nux-dextop-el7-x86_64
Scheduled Syncs:
Distributors:
Auto Publish: True
Config:
Http: True
Https: False
Relative URL: nux/dextop/el7/x86_64
Distributor Type Id: yum_distributor
Id: yum_distributor
Last Publish: None
Repo Id: nux-dextop-el7-x86_64
Scheduled Publishes:
Auto Publish: False
Config:
Http: True
Https: False
Distributor Type Id: export_distributor
Id: export_distributor
Last Publish: None
Repo Id: nux-dextop-el7-x86_64
Scheduled Publishes:

How can I find out what is preventing this sync from running?

I have found that I can fix errors of this type by deleting the repo and re-creating it.

I am preserving the state of this problem so that we can investigate it.

History

#1 - 06/03/2015 01:47 AM - Ben.Stanley

Additional observations:

[root@bumblebee ~]# pulp-admin tasks list | grep Waiting | wc -l
5
[root@bumblebee ~]# pulp-admin tasks list | grep Running | wc -l
0

[root@bumblebee ~]# pulp-admin tasks list | less

Operations: reaper
Resources:
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: 19832c86-bff5-4bba-a2c8-900678c612f6

Operations: sync
Resources: nux-dextop-el7-x86_64 (repository)
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: b621280e-75b5-42eb-9e2c-c2e3f8a23cfa

Operations: reaper
Resources:
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: e3cdff0d-0cf8-4924-92ae-20a6b5bdf5d0

Operations: reaper
Resources:
State: Waiting

Start Time: Unstarted
Finish Time: Incomplete
Task Id: fb368d3f-4de1-4dc8-ac57-074947ee6af6

Operations: reaper
Resources:
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: c3bdcb56-01b7-45f4-97f1-da7849a118b6

What are all those reapers waiting for?

What happens if I cancel them?

[root@bumblebee ~]# pulp-admin tasks cancel --task-id=19832c86-bff5-4bba-a2c8-900678c612f6
Task cancel is successfully initiated.

[root@bumblebee ~]# pulp-admin tasks cancel --task-id=e3cdff0d-0cf8-4924-92ae-20a6b5bdf5d0
Task cancel is successfully initiated.

[root@bumblebee ~]# pulp-admin tasks cancel --task-id=fb368d3f-4de1-4dc8-ac57-074947ee6af6
Task cancel is successfully initiated.

[root@bumblebee ~]# pulp-admin tasks cancel --task-id=c3bdcb56-01b7-45f4-97f1-7849a118b6
Task cancel is successfully initiated.

[root@bumblebee ~]# pulp-admin tasks list | grep Waiting | wc -l
1
[root@bumblebee ~]# pulp-admin tasks list | grep Running | wc -l
0
[root@bumblebee ~]# pulp-admin tasks list

Operations: sync
Resources: nux-dextop-el7-x86_64 (repository)
State: Waiting
Start Time: Unstarted
Finish Time: Incomplete
Task Id: b621280e-75b5-42eb-9e2c-c2e3f8a23cfa

So now I know that there is only the sync tasks waiting, and that nothing else is running.

Does re-starting pulp have any effect?

pulp-stop <- this is taking a very long time to complete.

pulp-start

The pulp-stop and pulp-start commands are scripts as follows:
--------------------------- pulp-stop ---------------------------------
[root@bumblebee ~]# cat /usr/local/bin/pulp-stop
#!/bin/bash

   1. pulp-stop
   2. stops the pulp daemons

   1. Complete stop/start also includes httpd (between pulp_worksrs and mongod)
   2. This is not included here as it interrupts customer service.

SERVICES=(pulp_resource_manager pulp_celerybeat pulp_workers mongod)

for ((i=0; i<${#SERVICES[@]}; ++i))
do
systemctl stop ${SERVICES[i]}
done

--------------------------- pulp-start ---------------------------------
[root@bumblebee ~]# cat /usr/local/bin/pulp-start
#!/bin/bash

   1. pulp-start
   2. starts up the pulp daemons

1. Complete stop/start also includes httpd (between pulp_worksrs and mongod)
2. This is not included here as it interrupts customer service.

SERVICES=(pulp_resource_manager pulp_celerybeat pulp_workers mongod)

```
for ((i=${#SERVICES[@]}-1; i>=0; --i ))
do
systemctl start ${SERVICES[i]} || exit 1
done
```

Hmmm... Now things are busy
[root@bumblebee ~]# pulp-admin tasks list | grep Running | wc -l
3
[root@bumblebee ~]# pulp-admin tasks list | grep Waiting | wc -l
525

It looks like it has decided to sync the whole world after I re-started pulp. That was unexpected.

Lets stop that so that we can test the repo under investigation.
[root@bumblebee ~]# while true; do sleep 5; date; CancelAllRunningTasks; done
This calls a bash function

```
function CancelAllRunningTasks() {
local AllTaskIds=( $(pulp-admin tasks list | perl 0777 -pe 's/Operations:[^\n]*\nResources:[^\n]*\nState: \nStart Time: +[^ \n]\nFinish Time: [^ \n]\nTask Id: ([^ \n])/\2/igs' | grep -e "^[a-f0-9\]\+\$") )
local i
for ((i=0; i<${#AllTaskIds[@]}; ++i ))
do
pulp-admin tasks cancel --task-id ${AllTaskIds[i]}
done
}
```

When this finally finishes killing tasks, we are ready to proceed.

Now attempt to sync the problem repo.

[root@bumblebee ~]# pulp-admin rpm repo sync run --repo-id=nux-dextop-el7-x86_64
━━

Synchronizing Repository [nux-dextop-el7-x86_64]
━━

This command may be exited via ctrl+c without affecting the request.

Downloading metadata...
[-]
... completed

Downloading repository content...
[==] 100%
RPMs: 117/117 items
Delta RPMs: 0/0 items

... completed

Individual package errors encountered during sync:

Package:
http://mirror.li.nux.ro/li.nux.ro/nux//dextop/el7/x86_64/mythtv-debuginfo-0.27.1
-3.el7.nux.x86_64.rpm
Error:

An unexpected error has occurred. More information can be found in the client
log file ~/.pulp/admin.log.

[root@bumblebee ~]# less ~/.pulp/admin.log
2015-06-02 19:38:49,172 - ERROR - Client-side exception occurred
Traceback (most recent call last):
File "/usr/lib/python2.7/site-packages/pulp/client/extensions/core.py", line 478, in run
exit_code = Cli.run(self, args)
File "/usr/lib/python2.7/site-packages/okaara/cli.py", line 974, in run
exit_code = command_or_section.execute(self.prompt, remaining_args)
File "/usr/lib/python2.7/site-packages/pulp/client/extensions/extensions.py", line 224, in execute
return self.method(*arg_list, **clean_kwargs)
File "/usr/lib/python2.7/site-packages/pulp/client/commands/repo/sync_publish.py", line 126, in run

self.poll([sync_task], kwargs)
File "/usr/lib/python2.7/site-packages/pulp/client/commands/polling.py", line 120, in poll
task = self._poll_task(task)
File "/usr/lib/python2.7/site-packages/pulp/client/commands/polling.py", line 211, in _poll_task
self.progress(task, running_spinner)
File "/usr/lib/python2.7/site-packages/pulp/client/commands/repo/sync_publish.py", line 71, in progress
self.renderer.display_report(task.progress_report)
File "/usr/lib/python2.7/site-packages/pulp_rpm/extensions/admin/status.py", line 72, in display_report
self.render_download_step(progress_report)
File "/usr/lib/python2.7/site-packages/pulp_rpm/extensions/admin/status.py", line 249, in render_download_step
'name': error[constants.NAME],
KeyError: 'name'

The sync has run without getting stuck, but something else has gone wrong as well

Observations:
1) Something is going wrong with pulp such that synchs can get stuck. Re-starting pulp seems to fix it.
2) Observed a KeyError .

So it seems that something in pulp's internal state is getting upset, and re-setting pulp fixes it.

#2 - 06/05/2015 04:36 PM - jortel@redhat.com

- Priority changed from Normal to High

- Severity changed from 2. Medium to 3. High

- Triaged changed from No to Yes

Please look through the logs for and report any stack traces or errors.

#3 - 07/06/2015 04:33 AM - mhrivnak

- Tags deleted (Easy Fix)

I assume it was a mistake for the "Easy Fix" tag to be put on this.

#4 - 07/06/2015 07:46 AM - Ben.Stanley

Yes, the easy fix tag was a mistake. I could not figure out how to remove it.

On 6 July 2015 12:33:18 pm Pulp <dropbox+pulp+c71e@plan.io> wrote:

#5 - 10/14/2015 04:55 PM - ipanova@redhat.com

- Status changed from NEW to ASSIGNED

- Assignee set to ipanova@redhat.com

#6 - 10/14/2015 05:38 PM - ipanova@redhat.com

- Status changed from ASSIGNED to CLOSED - WORKSFORME

I've been trying to reproduce this issue, nothing that happened to you happened during steps of reproduce.
One really strange thing that i noticed in your repo details output is that you did not synced repo so the Last sync is None but nevertheless there are
Content unit Counts that usually appear after sync.
I am closing this as Worksforme, in case it will happen again please reopen.

#7 - 10/14/2015 10:04 PM - Ben.Stanley

It also works for me now.

I recently attempted to find the root of this kind of problem, and I found
that I could make problems happen by synchronizing multiple repos
simultaneously. However, if I serialize the syncs, then all 580 repos sync
successfully. It is the first time I have been able to sync all the repos.

I used to have all 580 repos sync on a schedule at 10pm on Friday night,
assuming that pulp could handle some concurrency and scheduling. It seems
that it can't.

Ben Stanley

On 15 October 2015 2:39:13 am Pulp <dropbox+pulp+c71e+pulp@plan.io> wrote:

#8 - 04/15/2019 10:48 PM - bmbouter

- Tags Pulp 2 added