Pulp - Issue #7445

Pulp2 overrides permissions set by pulp3

09/02/2020 12:31 PM - ipanova@redhat.com

Status:	CLOSED - CURRENTRELEASE	**Start date:**	
Priority:	High	**Due date:**	
Assignee:	ggainey	**Estimated time:**	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	2. Medium	**Groomed:**	No
Version:		**Sprint Candidate:**	No
Platform Release:	2.21.4	**Tags:**	Pulp 2
OS:		**Sprint:**	Sprint 83
Triaged:	Yes	**Quarter:**	

Description

Imagine you have pulp2 and some data(in this case i had only ISO content), then pulp3 gets installed. Write permission for the pulp group is granted and setgid is set [0] Issue comes when new content gets synced into pulp2(in this case I synced an rpm repo) .Pulp2 overrides permissions, this makes is impossible to create a hardlink

```
$  ll /var/lib/pulp/content/units/
total 12
drwxrwsr-x. 173 apache pulp 4096 Sep  2 08:34 iso    <-----------------   existing content by th
e time pulp3 installed
drwxr-sr-x.  12 apache pulp  106 Sep  2 08:35 modulemd  <--------------------   new content afte
r pulp3 installed
drwxr-sr-x.   5 apache pulp   36 Sep  2 08:35 modulemd_defaults
drwxr-sr-x.  35 apache pulp 4096 Sep  2 08:35 rpm
(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 ~

$  ll /var/lib/pulp/content/units/modulemd
total 0
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 00
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 04
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 1b
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 41
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 66
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 78
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 8a
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 90
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 a8
drwxr-sr-x. 2 apache pulp 76 Sep  2 08:35 ea


Then i synced a new ISO repo.


$ ll /var//lib//pulp/content/units/iso/12
total 0
drwxr-sr-x. 2 apache pulp 19 Sep  2 08:33 3f7c65dc3598a59bbb867425c4e52cc54ecb66ff0f6db4656d349799
b96594
(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 ~]$ ll /var//lib//pulp/content/units/iso/12/3f7
c65dc3598a59bbb867425c4e52cc54ecb66ff0f6db4656d349799b96594/
```

All the content that appears after pulp3 is installed, does not have write permission for the pulp group. This makes it impossible to create hard link during the migration https://pulp.plan.io/issues/7244

[0] https://github.com/pulp/pulp_installer/blob/master/roles/pulp_common/tasks/install.yml#L107-L133

Related issues:

Is duplicate of Migration Plugin - Issue #7244: Hard links not being used dur...	**CLOSED - DUPLICATE**

Copied to RPM Support - Issue #7653: Pulp2 overrides permissions set by pulp3	**CLOSED - CURRENTRELEASE**
Copied to Docker Support - Issue #7654: Pulp2 overrides permissions set by pulp3	**CLOSED - CURRENTRELEASE**

Associated revisions

Revision e067d28a - 10/08/2020 02:55 PM - ggainey

Taught util.misc.mkdir() to set umask to allow group-w for makedirs calls.

Hunted down a LOT of makedirs calls and replaced with call to misc.mkdir()

fixes #7445

Revision 034c6baf - 10/13/2020 10:00 PM - ggainey

Taught util.misc.mkdir() to set umask to allow group-w for makedirs calls.

Hunted down a LOT of makedirs calls and replaced with call to misc.mkdir()

fixes #7445

(cherry picked from commit e067d28a00bdec989dd03170d599e58cde4fba13)

History

#1 - 09/02/2020 12:36 PM - ipanova@redhat.com

- Description updated

#2 - 09/02/2020 12:36 PM - ipanova@redhat.com

- Description updated

#3 - 09/02/2020 12:56 PM - ipanova@redhat.com

- Description updated

#4 - 09/02/2020 01:11 PM - ipanova@redhat.com

- Description updated

#5 - 09/02/2020 05:30 PM - ipanova@redhat.com

- Subject changed from Pulp2 ovverides permissions set by pulp3 to Pulp2 overrides permissions set by pulp3

#6 - 09/02/2020 08:59 PM - fao89

- Triaged changed from No to Yes

- Sprint set to Sprint 80

#7 - 09/03/2020 04:53 PM - ttereshc

- Priority changed from Normal to High

- Tags Katello added

#8 - 09/03/2020 04:54 PM - ipanova@redhat.com

- Is duplicate of Issue #7244: Hard links not being used during RPM content type migration added

#9 - 09/04/2020 05:24 PM - rchan

- Sprint changed from Sprint 80 to Sprint 81

#10 - 09/18/2020 09:29 PM - rchan

- Sprint changed from Sprint 81 to Sprint 82

#11 - 09/22/2020 08:44 PM - fao89

- Category set to Installer - Moved to GitHub issues

#12 - 09/24/2020 02:27 PM - ggainey

- Status changed from NEW to ASSIGNED

- Assignee set to ggainey

#13 - 09/28/2020 05:44 PM - ggainey

- Category deleted (Installer - Moved to GitHub issues)

- Tags Pulp 2 added

- Tags deleted (Katello)

This isn't an installer or migration issue. What's happening here is that Pulp2 uses os.makedirs() to create new paths when syncing. makedirs(), even when mode= is specified, uses current-umask to set permissions (see the [os.makedirs](#) documentation). The umask for pulp is 022. As a result, directories created post-pulp3-install end up 755, and Things Break.

os.makedir() is used in a number of places in pulp2, any of which will result in directories not ending up with the permissions the code is expecting (especially the places where pulp2 calls makedirs() with mode=770, which clearly expect the result to be what was specified at least for the leaf-node).

Pulp2 has a utility method in base-pulp, [pulp.plugins.util.misc.mkdir()](#). We probably should a) fix this method to Do The Right Thing in terms of insuring the leaf-node ends up with permissions specified by mode=, if any, and b) replace existing explicit os.makedirs() calls with calls to this utility routine.

#14 - 09/28/2020 08:06 PM - ggainey

Is the problem here *just* the leaf-node? If it's just the leaf-node, then we can 'fix' this just by insuring we directly chmod() the end-result to the perms we want/need/expect. If any/all internediate dirs need the permissions as well, then there is some more-interesting work that has to happen.

[ipanova@redhat.com](#) , [ttereshc](#) - thoughts?

#15 - 09/28/2020 08:43 PM - mdepaulo@redhat.com

How about we use the [systemd unit files to set the umask](#) to 002?

Note about setgid:

Those perms drwxr-sr-x are 2755, where the leading 2 is setgid.

The setgid is getting inherited from the parent dir, so pulp_installer did its job of setting it. [setgid cannot be controlled by umask](#), so it will not complicate this approach.

#16 - 09/29/2020 02:42 PM - ggainey

[mdepaulo@redhat.com](#) wrote:

> How about we use the [systemd unit files to set the umask](#) to 002?

Interesting idea. We're explicitly setting umask for pulp-resource-manager in Pulp2 now:

system/pulp_resource_manager.service:
ExecStart=/usr/bin/celery worker -A pulp.server.async.app -n resource_manager@%%h
-Q resource_manager -c 1 --events --umask 18 --pidfile=/var/run/pulp/resource_manager.pid

> Note about setgid:
>
> Those perms drwxr-sr-x are 2755, where the leading 2 is setgid.
>
> The setgid is getting inherited from the parent dir, so pulp_installer did its job of setting it. [setgid cannot be controlled by umask](#), so it will not complicate this approach.

Yeah, concur, ownership should be fine, it's just the default-directory-perms that need to be addressed.

#17 - 09/29/2020 06:35 PM - ggainey

[mdepaulo@redhat.com](#) wrote:

> How about we use the [systemd unit files to set the umask](#) to 002?

The problem with setting it outside code is, that umask can be/is set:

- at the service level in /etc/systemd/system/.service, which can be overridden by
- at the user-level, in /etc/systemd/user.conf and users/, which can be overridden by
- the system-level shell/env files /etc/bashrc /etc/profile and /etc/cshrc

Setting it explicitly in code, of course, trumps all of the above.

I'm experimenting on a pulp2 box. So far, I have not figured out which umask-default controls this particular scenario, and it makes me sad. I do have a fix in code that works. Experimentation continues.

#18 - 09/30/2020 03:40 PM - ggainey

Commentary/discussion with jsherrill:

On Wed, Sep 30, 2020 at 9:04 AM Justin Sherrill jsherril@redhat.com wrote:

> Wouldn't this first approach also require a script to fix existing files/directories?

The pulp3 installer does exactly tyhat at install-time :

https://github.com/pulp/pulp_installer/blob/master/roles/pulp_common/tasks/install.yml#L107-L133

> The difference I think is that this first approach would only need the script run once prior to migration. While the 2nd approach would require it to be run before every migration. I could see this taking a long time on a filesystem with 100s of thousands of files, or millions even. More so on a high latency file system. Since subsequent migrations are aimed to be as 'quick' as possible, i think we have to go with the first option?

The 'what if this is a HYOOGE filesystem" argument is persuasive :)

#19 - 09/30/2020 03:42 PM - ggainey

(Addendum to #c18 - "first approach" is "explicitly set/reset umask prior to makedirs() calls", "second approach" is "run script after any pulp2-activity post-pulp3-install")

#20 - 09/30/2020 04:29 PM - ggainey

One final(?) complicating factor to be aware of: in python2, os.makedirs(dir, mode=) will try to set the permissions on **all** directories it creates on the way to the leaf. In python3, that is no longer true - intermediate directories'permissions are driven just by umask.

This means that a sequence like

```
import os
mask = os.umask(0)
os.makedirs("/tmp/foo/bar/blech", mode=0770)
```

results in each of foo, bar, and blech ending up with permissions 770.

The same code in python3 results in foo and bar being 777, and only blech being 770.

Especially in testing, it's important to remember the environment the code will be running under. Pulp2 runs under Python2 and can therefore 'assume' the "mode= is for perms for any/all created dirs" behavior, where Pulp3 cannot.

Hopefully this note will save someone else from the 10 minutes of confusion I just suffered....

#21 - 10/01/2020 06:40 PM - mdepaulo@redhat.com

ggainey wrote:

> mdepaulo@redhat.com wrote:
>
>> How about we use the systemd unit files to set the umask to 002?
>
> The problem with setting it outside code is, that umask can be/is set:
>
> - at the service level in /etc/systemd/system/.service, which can be overridden by
> - at the user-level, in /etc/systemd/user.conf and users/, which can be overridden by

No, this affects systemd user mode, not system mode, which Pulp runs in.

> - the system-level shell/env files /etc/bashrc /etc/profile and /etc/cshrc

No, one of the designs behind systemd is to prevent these from bleeding in. And for umask, that's actually a different issue entirely than environment variables.

> Setting it explicitly in code, of course, trumps all of the above.

You're right, I just thought I'd propose the systemd unit files because:

1. it would be quicker/easier to implement.
2. If there was ever a need to configure it at runtime/install-time, it would be easier.

   I'm experimenting on a pulp2 box. So far, I have not figured out which umask-default controls this particular scenario, and it makes me sad. I do have a fix in code that works. Experimentation continues.

To clarify, this is a variant of the 1st approach.

#22 - 10/02/2020 06:03 PM - rchan

- Sprint changed from Sprint 82 to Sprint 83

#23 - 10/05/2020 07:17 PM - ggainey

- Copied to Issue #7653: Pulp2 overrides permissions set by pulp3 added

#24 - 10/05/2020 07:24 PM - ggainey

- Copied to Issue #7654: Pulp2 overrides permissions set by pulp3 added

#25 - 10/05/2020 07:46 PM - ggainey

- Status changed from ASSIGNED to POST

PR: https://github.com/pulp/pulp/pull/4000

#26 - 10/08/2020 02:56 PM - ggainey

- Status changed from POST to MODIFIED

Applied in changeset pulp|e067d28a00bdec989dd03170d599e58cde4fba13.

#27 - 10/08/2020 04:48 PM - ggainey

- Platform Release set to 2.21.4

#28 - 11/02/2020 04:06 PM - ggainey

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE