

RPM Support - Issue #3585

Pulp sync/publish relies on unique filenames per repo, corrupting repository

04/20/2018 04:24 PM - dekimsey

Status:	CLOSED - WONTFIX	Start date:	
Priority:	Normal	Due date:	
Assignee:		Estimated time:	0:00 hour
Category:			
Sprint/Milestone:			
Severity:	3. High	Groomed:	No
Version:	2.12.2	Sprint Candidate:	No
Platform Release:		Tags:	Pulp 2
OS:	RHEL 7	Sprint:	
Triaged:	Yes	Quarter:	
Description			
<p>We have an upstream vendor, Centrify, that releases unique (nevra) packages in their published yum repo but reuses the filename in their yum repo. Effectively unpublishing the old copy. While this is poor practice on their part, it's breaking Pulp's rpm content management.</p> <p>On publish, Pulp doesn't maintain these packages as separate entities, symlinks and metadata do not match. They are stored correctly on sync in pulp's content/units/rpm directory tree. <i>However, the published symlinks use the upstream filename and not a computed nevra. This results in "content does not match metadata" errors from yum and a general corruption of the repository.</i> Not only are multiple versions not only accessible but the sync process doesn't seem to handle this gracefully either. New metadata, old package on disk.</p> <pre># Sync'd content organized by digest (yay). # sync'd filenames aren't nevra (whatevs, we got us some digests) [root@katello ~]\$ find /var/lib/pulp/content/units/rpm -name 'CentrifyDC-5.4.3-*x86_64.rpm' /var/lib/pulp/content/units/rpm/56/0e778319ec52a6aa75e7b72345bfd1fa11a91f4b112b36d30ed2c256ce15e5/ CentrifyDC-5.4.3-rhel5.x86_64.rpm /var/lib/pulp/content/units/rpm/a4/39ee653ddd3a4203ec949de3c8ff69494b5bf5d00798536bbbbb5eaeb46863/ CentrifyDC-5.4.3-rhel5.x86_64.rpm [root@katello ~]\$ find /var/lib/pulp/content/units/rpm -name 'CentrifyDC-5.4.3-*x86_64.rpm' -execd ir rpm -qp {} + CentrifyDC-5.4.3-905.x86_64 CentrifyDC-5.4.3-887.x86_64 # Published content symlinks are named by filename (boo) [root@katello ~]# sudo find -L /var/lib/pulp/published/yum/https/repos/Trustwave/Library -name 'Ce ntrifyDC-5.4.3*.x86_64.rpm' /var/lib/pulp/published/yum/https/repos/Trustwave/Library/custom/centrify/centrify-centrifydc-rpms /Packages/c/CentrifyDC-5.4.3-rhel5.x86_64.rpm [root@katello ~]# readlink /var/lib/pulp/published/yum/https/repos/Trustwave/Library/custom/centri fy/centrify-centrifydc-rpms/Packages/c/CentrifyDC-5.4.3-rhel5.x86_64.rpm /var/lib/pulp/content/units/rpm/56/0e778319ec52a6aa75e7b72345bfd1fa11a91f4b112b36d30ed2c256ce15e5/ CentrifyDC-5.4.3-rhel5.x86_64.rpm [root@katello ~]# rpm -qp /var/lib/pulp/published/yum/https/repos/Trustwave/Library/custom/centrif y/centrify-centrifydc-rpms/Packages/c/CentrifyDC-5.4.3-rhel5.x86_64.rpm CentrifyDC-5.4.3-905.x86_64 # Example yum repoquery attempt (though this only shows two versions, sorry) [root@sandbox ~]# repoquery --location -a CentrifyDC https://smartproxy.com/pulp/repos/Trustwave/development/ccv-biz-portal-el7/custom/centrify/centrif y-centrifydc-rpms/Packages/c/CentrifyDC-5.4.1-rhel4.i386.rpm https://smartprox.com/pulp/repos/Trustwave/development/ccv-biz-portal-el7/custom/centrify/centrify -centrifydc-rpms/Packages/c/CentrifyDC-5.4.3-rhel5.x86_64.rpm [root@sandbox ~]# repoquery -a CentrifyDC CentrifyDC-0:5.4.1-455.i386 CentrifyDC-0:5.4.3-905.x86_64</pre>			

While I've reported the issue to Centrifly, Pulp should not be depending on correct package naming standards from third-parties. I think published content should have symlinks renamed to "\$nevra.rpm" to ensure correctness.

In our case, this has caused our content to become inaccessible. I don't know when content from the units directory will be removed. But we've basically lost access to the older packages and yum cannot handle the upgrading since the published files are incorrect packages.

In case anyone else runs into this with Centrifly, feel free to reference the case we filed, "180419-159679: rpm repo doesn't use unique rpm names for packages".

History

#1 - 04/20/2018 04:30 PM - dekimsey

Just some version information for posterity.

```
$ cat /etc/redhat-release
CentOS Linux release 7.3.1611 (Core)
$ rpm -q katello
katello-3.4.2-1.el7.noarch
$ rpm -qa | grep pulp | sort
pulp-admin-client-2.12.2-1.el7.noarch
pulp-client-1.0-3.noarch
pulp-docker-plugins-2.3.0-1.el7.noarch
pulp-katello-1.0.2-1.el7.noarch
pulp-puppet-plugins-2.12.2-1.el7.noarch
pulp-puppet-tools-2.12.2-1.el7.noarch
pulp-rpm-admin-extensions-2.12.2-1.el7.noarch
pulp-rpm-plugins-2.12.2-1.el7.noarch
pulp-selinux-2.12.2-1.el7.noarch
pulp-server-2.12.2-1.el7.noarch
python-isodate-0.5.0-4.pulp.el7.noarch
python-kombu-3.0.33-6.pulp.el7.noarch
python-pulp-bindings-2.12.2-1.el7.noarch
python-pulp-client-lib-2.12.2-1.el7.noarch
python-pulp-common-2.12.2-1.el7.noarch
python-pulp-docker-common-2.3.0-1.el7.noarch
python-pulp-oid_validation-2.12.2-1.el7.noarch
python-pulp-puppet-common-2.12.2-1.el7.noarch
python-pulp-repoauth-2.12.2-1.el7.noarch
python-pulp-rpm-common-2.12.2-1.el7.noarch
python-pulp-streamer-2.12.2-1.el7.noarch
rubygem-smart_proxy_pulp-1.3.0-1.el7.noarch
```

#2 - 04/20/2018 04:35 PM - dekimsey

Looking up the packages for CentriflyDC-5.4.3-887

```
$ curl https://katello.trustwave.com//katello/api/packages/40d9a24c-b9c4-48e4-a9b2-6a56a31f0e94
{
  id: 96963,
  uuid: "40d9a24c-b9c4-48e4-a9b2-6a56a31f0e94",
  name: "CentriflyDC",
  version: "5.4.3",
  release: "887",
  arch: "x86_64",
  epoch: "0",
  summary: "Centrifly DirectControl Agent",
  filename: "CentriflyDC-5.4.3-rhel5.x86_64.rpm",
  sourcerpm: "CentriflyDC-5.4.3-887.src.rpm",
  checksum: "cec7c8d4f67fdf40e76cc42e6c59452d46f59d2d",
  nvrea: "CentriflyDC-5.4.3-887.x86_64",
  nvra: "CentriflyDC-5.4.3-887.x86_64",
  hosts_available_count: 0,
  hosts_applicable_count: 0,
  description: "RPM to install Centrifly DirectControl on Linux x86 platforms.",
  license: "BSD with portions copyright (c) Centrifly Corporation 2006-2017 and licensed under Centrifly End User License Agreement",
  buildhost: "build75-rhel5-x86-64.centrifly.com",
  vendor: null,
  relativepath: "CentriflyDC-5.4.3-rhel5.x86_64.rpm",
  children: { },
  checksumtype: "sha1",
  size: 22285042,
```

```
url: "http://www.centrifly.com/",
build_time: 1512566040,
group: "Applications/System",
requires: [],
provides: [
"CentriflyDC-5.4.3-887"
],
files: [],
human_readable_size: "21.3 MB"
}
```

Looking up the packages for CentriflyDC-5.4.3-905

```
$ curl https://katello.trustwave.com/katello/api/packages/05ee84d8-8328-4646-83d3-d4869f4f9efe
{
id: 101950,
uuid: "05ee84d8-8328-4646-83d3-d4869f4f9efe",
name: "CentriflyDC",
version: "5.4.3",
release: "905",
arch: "x86_64",
epoch: "0",
summary: "Centrifly DirectControl Agent",
filename: "CentriflyDC-5.4.3-rhel5.x86_64.rpm",
sourcerpm: "CentriflyDC-5.4.3-905.src.rpm",
checksum: "5a7806d50b6847d4b6b713ca25dde8d71aa32012",
nvrea: "CentriflyDC-5.4.3-905.x86_64",
nvra: "CentriflyDC-5.4.3-905.x86_64",
hosts_available_count: 111,
hosts_applicable_count: 813,
description: "RPM to install Centrifly DirectControl on Linux x86 platforms.",
license: "BSD with portions copyright (c) Centrifly Corporation 2006-2017 and licensed under Centrifly End User License Agreement",
buildhost: "build75-rhel5-x86-64.centrifly.com",
vendor: null,
relativepath: "CentriflyDC-5.4.3-rhel5.x86_64.rpm",
children: { },
checksumtype: "sha1",
size: 22267710,
url: "http://www.centrifly.com/",
build_time: 1522287156,
group: "Applications/System",
requires: [],
provides: [],
files: [],
human_readable_size: "21.2 MB"
}
```

#3 - 04/20/2018 04:51 PM - dalley

- Triaged changed from No to Yes

#4 - 04/12/2019 10:20 PM - bmbouter

- Status changed from NEW to CLOSED - WONTFIX

Pulp 2 is approaching maintenance mode, and this Pulp 2 ticket is not being actively worked on. As such, it is being closed as WONTFIX. Pulp 2 is still accepting contributions though, so if you want to contribute a fix for this ticket, please reopen or comment on it. If you don't have permissions to reopen this ticket, or you want to discuss an issue, please reach out via the [developer mailing list](#).

#5 - 04/15/2019 10:12 PM - bmbouter

- Tags Pulp 2 added