

## Migration Plugin - Issue #7625

### RPM Migration results in incomplete package metadata

09/30/2020 07:57 AM - adam.winberg@smhi.se

<b>Status:</b> CLOSED - CURRENTRELEASE	<b>Start date:</b>
<b>Priority:</b> High	<b>Due date:</b>
<b>Assignee:</b> dalley	<b>Estimated time:</b> 0:00 hour
<b>Category:</b>	
<b>Sprint/Milestone:</b> 0.5.1	
<b>Severity:</b> 3. High	<b>Sprint Candidate:</b> No
<b>Platform Release:</b>	<b>Tags:</b>
<b>OS:</b>	<b>Sprint:</b> Sprint 84
<b>Triaged:</b> Yes	<b>Quarter:</b>
<b>Groomed:</b> No	

**Description**

edit:

Migration does not properly populate the filelist metadata on the resulting RPMs, which impacts subsequent usage of the RPM copy APIs, which read this metadata as part of the process to determine dependency relationships.

original:

I am running pulp on RHEL 8 with a rpm-based installation based on packages from the Katello nightly repo.

I have succeeded in migrating my pulp2 content to pulp3 and all repos look complete (i've reran the migration and no new repoversions are generated).

In an attempt to test the multirepo copy functionality described at <https://pulp-rpm.readthedocs.io/en/latest/workflows/copy.html#recipes>

, I created a couple of new, empty repos and tried to copy errata from my migrated RHEL8-appstream repo to them:

```
POST /pulp/api/v3/rpm/copy/
config:=[
  {"source_repo_version": "<rhel8-appstream-version-href>", "dest_repo": "my-new-repo1", "content": [{"ADVISORY_HREF1"}]},
  {"source_repo_version": "<rhel8-baseos-version-href>", "dest_repo": "my-new-repo2", "content": []},
]
dependency_solving=False
```

This works (after applying the patch from [https://github.com/pulp/pulp\\_rpm/commit/712abdf1abb95c969b54fd2968a573189b77bcba](https://github.com/pulp/pulp_rpm/commit/712abdf1abb95c969b54fd2968a573189b77bcba)), or at least it executes without any errors.

I'm a bit confused though by the result. I copied 16 advisories from the appstreams repo to my new empty repo and the copy ended up with copying all modulemds and almost all packages (11466 of 12053). That doesnt seem right to me. And my other new repo ('my-new-repo2') is still empty, meaning that of those 11000 packages there were none that had any dependencies from the baseos repo which also strikes me as odd. If i set 'dependency\_solving=True' (contrary to the documentation) I end up with the same amount of packages in my 'my-new-repo1' and 2 packages in 'my-new-repo2'.

Am I misunderstanding something about this functionality?

**Related issues:**

Related to RPM Support - Issue #7705: Advanced-copy overview can be misleading **CLOSED - CURRENTRELEASE**

#### Associated revisions

Revision d88dadd0 - 10/26/2020 08:28 PM - dalley

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625>

**Revision d88dadd0 - 10/26/2020 08:28 PM - dalley**

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625>

**Revision d88dadd0 - 10/26/2020 08:28 PM - dalley**

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625>

**Revision 7084fe30 - 10/27/2020 01:54 PM - dalley**

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625> (cherry picked from commit d88dadd0b4ad40a112c250b0a78cc79593b3c8e9)

**Revision 7084fe30 - 10/27/2020 01:54 PM - dalley**

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625> (cherry picked from commit d88dadd0b4ad40a112c250b0a78cc79593b3c8e9)

**Revision 7084fe30 - 10/27/2020 01:54 PM - dalley**

Incomplete migration of filelist and 'other' metadata

closes: #7625 <https://pulp.plan.io/issues/7625> (cherry picked from commit d88dadd0b4ad40a112c250b0a78cc79593b3c8e9)

## History

---

**#1 - 09/30/2020 12:10 PM - adam.winberg@smhi.se**

I have also tried this with using new source repos instead of the migrated repos and the copy between the first source/dest repo pairs then looks correct, but there are no dependencies copied from the baseos repo. Considering that the copy from the appstream repo contained in this test contained 610 rpm packages there really should be some dependencies from the baseos repo.

**#2 - 10/01/2020 05:17 PM - ttereshc**

- *Triaged changed from No to Yes*

- *Sprint set to Sprint 82*

**#3 - 10/02/2020 06:03 PM - rchan**

- *Sprint changed from Sprint 82 to Sprint 83*

**#4 - 10/07/2020 10:26 PM - ggaaney**

@adam.winberg@smhl.se - I'm looking into this., Could you tell me the specific advisories you're copying? I want to make sure I'm reproducing your scenario correctly. Thanks!

**#5 - 10/08/2020 07:06 AM - adam.winberg@smhi.se**

I've been using 'RHS-2020:4059' in the red hat appstream repo in my recent testing. Just a single advisory that is, but it should contain a bunch of dependencies to packages from the baseos repo.

**#6 - 10/08/2020 05:06 PM - ggaaney**

- *Assignee set to ggaaney*

**#7 - 10/08/2020 05:07 PM - ggaaney**

- *Status changed from NEW to ASSIGNED*

**#8 - 10/12/2020 11:08 PM - ggaaney**

OK, a quick update before this all falls out of my head.

I'm running on the pulp2-nightly-pulp3-source-centos7 dev-box, so it has both pulp2 and pulp3 installed. To start with, tho, I'm just working on the Pulp3 'side'.

I'm syncing RHEL8-BASEOS and RHEL-APPSTREAM from these URLs:

```
RHEL8_BASE_URL="https://cdn.redhat.com/content/dist/rhel8/8/x86_64/baseos/os/"
RHEL8_APPS_URL="https://cdn.redhat.com/content/dist/rhel8/8/x86_64/appstream/os/"
```

I am using the HREF for the advisory RHSA-2020:4059

After initial syncing, looking at that advisory shows 94 packages using this command:

```
$ export ADVISORY="RHSA-2020:4059"
$ http :/pulp/api/v3/content/rpm/advisories/?id=${ADVISORY} | jq -r ".results[0] | .pkglist[0].packages | length"
```

I created 2 repos, a new-base and a new-apps repo, and then copied '4059 into new-apps using the following command and config:

```
http POST :/pulp/api/v3/rpm/copy/ @config.json
{
  "config": [
    {
      "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/d09a1ae0-6cd6-4810-b861-86dac40bfe48/versions/1/",
      "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/",
      "content": ["/pulp/api/v3/content/rpm/advisories/9ce5b0da-897c-4f86-b09e-c431dee197a1/"],
      {
        "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/a0e73e8e-3876-4a06-a315-7bf276d28944/versions/1/",
        "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/",
        "content": []
      }
    ],
    "dependency_solving": false
  ]
}
```

This left me with 94 packages in new-appstream and nothing in new-base, as expected - since I had set dep-solv to false, and had limited baseos-to-new-baseos to be {} (ie nothing)

Repeating the experiment with dependency\_solving=true, I ended up with 9 modules and 194 packages added to new-appstream, and 288 packages added to new-baseos.

I have not (and am **not** planning to! :) ) go through a complete dependency-analysis of '4059, but this is pretty much what I'd expect. If you do **not** require depsolving, you get only those entities directly referenced for the advisories referenced in the content: list. If you **do** ask for depsolving, the process starts with those, and then depsolves against the contents of the provided source-repos. As noted in the last example in the [copy-workflow-recipes](#), DEST2 will only get the entities from SRC2 if that's the only place it can fulfill the dependencies from.

Reading the docs, I do see your confusion around needing dependency\_solve=true to make this happen - the introduction to copy suggests it happens all the time. Please open a docs-issue, and we'll fix it next docs-day (which, as it turns out, is tomorrow!)

Anyway - just working form Pulp3, the behavior I'm seeing is what I'd expect.

[dalley](#) - can you look this over and give us/me any deeper insight?

#### #9 - 10/13/2020 12:10 AM - dalley

I agree that the first result, where zero packages were copied into new-baseos repository if dependency\_solving=False, is entirely expected and that if our documentation suggests otherwise then our documentation needs some improvement. +1 to that.

@Grant, but I think we should probably look a little bit closer at the dependency\_solving result, it feels at least very questionable to me that 16 advisories would copy nearly every package in appstream but also would only result in 2 dependent packages from baseos being copied. Especially given your experiment where 288 were copied into new-baseos from a the dependencies of a single advisory.

I guess it's not impossible (maybe this one advisory wasn't one of the 16 Adam copied, and all of the others are 'shaped' significantly differently), but it feels kind of implausible.

@Adam, I think it would really help if you could try two things:

- Try copying RHSA-2020:4059 as Grant did, and see if you get a similar result to Grant, or if you still have a mostly-empty new-baseos repo
- Find 1-4 advisories from your original group of 16 which still reproduce a sizable cascade of packages copied between appstream -> new-appstream, but doesn't copy many into baseos -- and then let us know the advisory IDs so that we can do an exact reproduction of what you see more easily.

And @Grant, did you wipe the repositories between your two experiments? If so, perhaps you could try the two copies (without depsolving, then with depsolving) back to back and make sure if it behaves the way it's expected to.

#### #10 - 10/13/2020 01:04 AM - ggaaney

dalley wrote:

I agree that the first result, where zero packages were copied into new-baseos repository if dependency\_solving=False, is entirely expected and that if our documentation suggests otherwise then our documentation needs some improvement. +1 to that.

The docs issue is this one line from the [Advanced Copy workflow](#) doc:

When copying a RPM packages from one repository to another, all of the RPM packages that those packages depend on will also be copied.

This is before the in-depth explanation points out that dep-solve is **not** the default.

@Grant, but I think we should probably look a little bit closer at the dependency\_solving result, it feels at least very questionable to me that 16 advisories would copy nearly every package in appstream but also would only result in 2 dependent packages from baseos being copied. Especially given your experiment where 288 were copied into new-baseos from a the dependencies of a single advisory.

Concur, it def feels like there is something odd going on here.

I guess it's not impossible (maybe this one advisory wasn't one of the 16 Adam copied, and all of the others are 'shaped' significantly differently), but it feels kind of implausible.

@Adam, I think it would really help if you could try two things:

- Try copying RHSA-2020:4059 as Grant did, and see if you get a similar result to Grant, or if you still have a mostly-empty new-baseos repo
- Find 1-4 advisories from your original group of 16 which still reproduce a sizable cascade of packages copied between appstream -> new-appstream, but doesn't copy many into baseos.

And @Grant, did you wipe the repositories between your two experiments? If so, perhaps you could try the two copies (without depsolving, then with depsolving) back to back and make sure if it behaves the way it's expected to.

The experiment I ran was "copy no-deps, then repeat with deps=true", without deleting/recreating repos in the middle. Planning to try a second deps=true run into clean repos as my next step, which will be tomorrow.

Right now, I am working at this from the pure-Pulp3 viewpoint. If Adam can recreate the problem with a small set of advisories, and we **cannot**, then I'll move to "sync rhel8 in pulp2, migrate, and start again," to see if something in the migration path left a landmine somehow (alho off the top of my head I am having trouble figuring out how that might be. Stranger things have happened, tho :))

#### #11 - 10/13/2020 01:09 AM - dalley

Ah, right, I forgot momentarily that there was an extra level here w/ the migration plugin.

#### #12 - 10/13/2020 12:17 PM - adam.winberg@smhi.se

I have now copied RHSA-2020:4059 with "dependency\_solving=True", this results in 94 rpms in the appstreams dest-repo but still zero packages in my baseos dest repo.

I also picked 3 advisories from my original copy and copied those with "dependency\_solving=True". This added 199 modulemd's and 958 packages to my appstream dest repo and nothing at all in my baseos dest repo. The advisories I copied was RHSA-2020:3623 RHSA-2020:3634 RHSA-2020:3662

This is using pulp2-migrated repos as source repos.

#### #13 - 10/13/2020 02:31 PM - ggainey

[adam.winberg@smhi.se](mailto:adam.winberg@smhi.se) wrote:

I did a pure pulp3 test as well with RHSA-2020:4059, i.e. created new repos and synced them against my appstream/baseos remotes, and then used these as sources in my copy. Still the same result, 94 rpms in the appstreams dest-repo and nothing in baseos. This sounds like the results [ggainey](#) got when using "dependency\_solving=false", but i am very sure I am setting that parameter to 'true'.

This is my copy command:

OK, this is...exciting. Right now I'm concentrating on only Pulp3 being involved here, not triggered by anything in the migration - thanks Adam for the experiment in #c13 !

Let me get all my details on the table here. If anybody spots something obviously different from Adam's setup, sing out please!

- vagrant box pulp2-nightly-pulp3-source-centos7
- pulpcore version: 4ce583
- pulp\_rpm version: 6b78e3

These are a couple commits behind behind current-master; I will update to current-master this morning and retry (alho nothing in the commits I'm missing appear to impact rpm-copy)

Sync'd content from:

```
RHEL8_BASE_URL="https://cdn.redhat.com/content/dist/rhel8/8/x86_64/baseos/os/"
RHEL8_APPS_URL="https://cdn.redhat.com/content/dist/rhel8/8/x86_64/appstream/os/"
```

Remotes are set up as 'on\_demand':

```
BASE_REMOTE_HREF=$(http POST :/pulp/api/v3/remotes/rpm/rpm/ name=$RHEL8_BASE_NAME url=$RHEL8_BASE_URL policy='on_demand' client_cert=@/home/vagrant/devel/pulp_startup/CDN_cert/cdn.crt client_key=@/home/vagrant/devel/pulp_startup/CDN_cert/cdn.key ca_cert=@/home/vagrant/devel/pulp_startup/CDN_cert/redhat-uep.pem | jq -r '.pulp_href')
```

```
APPS_REMOTE_HREF=$(http POST :/pulp/api/v3/remotes/rpm/rpm/ name=$RHEL8_APPS_NAME url=$RHEL8_APPS_URL policy='on_demand' client_cert=@/home/vagrant/devel/pulp_startup/CDN_cert/cdn.crt client_key=@/home/vagrant/devel/pulp_startup/CDN_cert/cdn.key ca_cert=@/home/vagrant/devel/pulp_startup/CDN_cert/redhat-uep.pem | jq -r '.pulp_href')
```

### Config and command

```
http POST :/pulp/api/v3/rpm/copy/ @config.json
```

```
{
  "config": [
    {
      "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/d09a1ae0-6cd6-4810-b861-86dac40bfe48/versions/1/",
      "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/",
      "content": ["/pulp/api/v3/content/rpm/advisories/9ce5b0da-897c-4f86-b09e-c431dee197a1/"],
      {
        "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/a0e73e8e-3876-4a06-a315-7bf276d28944/versions/1/",
        "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/",
        "content": []
      }
    ],
    "dependency_solving": true
  ]
}
```

Results after second-execution with `dependency_solving: true` - first run had `dependency_solving: false` and put 94 RPMs into `appstream-new3` and nothing in `base-new3`.

### Results for APPSTREAM:

```
(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 7625_multicopy]$ http :/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/ HTTP/1.1 200 OK
...
```

```
{
  "description": null,
  "latest_version_href": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/",
  "metadata_signing_service": null,
  "name": "rhel8-apps-new3",
  "pulp_created": "2020-10-12T20:11:54.519172Z",
  "pulp_href": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/",
  "remote": null,
  "retain_package_versions": 0,
  "versions_href": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/"
}
```

```
(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 7625_multicopy]$ http :/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/ HTTP/1.1 200 OK
...
```

```
{
  "base_version": null,
  "content_summary": {
    "added": {
      "rpm.modulemd": {
        "count": 9,
        "href": "/pulp/api/v3/content/rpm/modulemds/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
      },
      "rpm.package": {
        "count": 194,
        "href": "/pulp/api/v3/content/rpm/packages/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
      }
    },
    "present": {
      "rpm.advisory": {
        "count": 1,
        "href": "/pulp/api/v3/content/rpm/advisories/?repository_version=/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
      }
    }
  }
}
```

```

    },
    "rpm.modulemd": {
      "count": 10,
      "href": "/pulp/api/v3/content/rpm/modulemds/?repository_version=/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
    },
    "rpm.package": {
      "count": 288,
      "href": "/pulp/api/v3/content/rpm/packages/?repository_version=/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
    }
  },
  "removed": {}
},
"number": 2,
"pulp_created": "2020-10-12T20:39:00.224864Z",
"pulp_href": "/pulp/api/v3/repositories/rpm/rpm/c58d5724-20ea-481a-b227-13a35b1a3ca5/versions/2/"
}

```

#### Results for **BASE**:

```

(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 7625_multicopy]$ http :/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/
HTTP/1.1 200 OK
...

```

```

{
  "description": null,
  "latest_version_href": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/1/"
},
  "metadata_signing_service": null,
  "name": "rhel8-base-new3",
  "pulp_created": "2020-10-12T20:11:54.015856Z",
  "pulp_href": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/",
  "remote": null,
  "retain_package_versions": 0,
  "versions_href": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/"
}

```

```

(pulp) [vagrant@pulp2-nightly-pulp3-source-centos7 7625_multicopy]$ http :/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/1/
HTTP/1.1 200 OK
...

```

```

{
  "base_version": null,
  "content_summary": {
    "added": {
      "rpm.package": {
        "count": 370,
        "href": "/pulp/api/v3/content/rpm/packages/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/1/"
      }
    },
    "present": {
      "rpm.package": {
        "count": 370,
        "href": "/pulp/api/v3/content/rpm/packages/?repository_version=/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/1/"
      }
    },
    "removed": {}
  },
  "number": 1,
  "pulp_created": "2020-10-12T20:39:01.874543Z",
  "pulp_href": "/pulp/api/v3/repositories/rpm/rpm/46fca451-bf12-4e98-ae12-880c3d789f81/versions/1/"
}

```

#### #14 - 10/13/2020 02:57 PM - adam.winberg@smhi.se

I am on an rpm-based installation, with

- python3-pulpcore-3.4.1-1.el8.noarch
- python3-pulp-2to3-migration-0.2.0-0.1.b7.el8.noarch
- python3-pulp-rpm-3.5.1-1.el8.noarch

**#15 - 10/13/2020 04:40 PM - ggaaney**

[adam.winberg@smhi.se](mailto:adam.winberg@smhi.se) wrote:

I am on an rpm-based installation, with

- python3-pulpcore-3.4.1-1.el8.noarch
- python3-pulp-2to3-migration-0.2.0-0.1.b7.el8.noarch
- python3-pulp-rpm-3.5.1-1.el8.noarch

Goodness - in pulp3 terms, those are Olde. Current from [http://yum.theforeman.org/pulpcore/3.7/el8/x86\\_64/](http://yum.theforeman.org/pulpcore/3.7/el8/x86_64/) are

- python3-pulpcore-3.7.1-3.el8.noarch.rpm
- python3-pulp-2to3-migration-0.4.0-1.el8.noarch.rpm
- python3-pulp-rpm-3.7.0-1.el8.noarch.rpm

There have been A LOT of fixes for pulpcore/pulp\_rpm in that time, including at least 6 in pulp\_rpm thta involve copy and/or depsolving. And I know that a ton of work has gone on in the migration area as well recently.

Adam, can you upgrade your pulp3 pieces to at least 3.7? We may be fighting a problem that's already been fixed sometime since June.

**#16 - 10/13/2020 05:02 PM - adam.winberg@smhi.se**

Goodness - in pulp3 terms, those are Olde. Current from [http://yum.theforeman.org/pulpcore/3.7/el8/x86\\_64/](http://yum.theforeman.org/pulpcore/3.7/el8/x86_64/) are

Whaaat - I've been hunting around for pulp3 rpm builds and finally was directed to the Katello repo on the pulp mailing list. So that's what I'm using and I've been waiting for a new Katello release just to get new pulp3 versions. I had no knowledge of the foreman repos you reference at all - are they new?

I will absolutely upgrade, and rerun my migration and tests. Will get back to you.

**#17 - 10/13/2020 07:53 PM - ggaaney**

- Related to Issue #7705: *Advanced-copy overview can be misleading added*

**#18 - 10/13/2020 09:24 PM - dalley**

I think the latest currently-released version of Katello only supports 3.4.1, but since they're working with newer versions of Pulp in development, they have those repositories built as well.

I'm guessing that you're not actually using Katello, but installing from RPM is more convenient?

**#19 - 10/14/2020 09:53 PM - ggaaney**

@Adam I made some updates to the advanced-copy docs to try and add some clarity - see [https://github.com/pulp/pulp\\_rpm/pull/1871](https://github.com/pulp/pulp_rpm/pull/1871) for the changes

**#20 - 10/15/2020 07:20 AM - adam.winberg@smhi.se**

I'm guessing that you're not actually using Katello, but installing from RPM is more convenient?

Exactly!

I made some updates to the advanced-copy docs to try and add some clarity - see [https://github.com/pulp/pulp\\_rpm/pull/1871](https://github.com/pulp/pulp_rpm/pull/1871) for the changes

That looks really good, a big improvement.

After some db problems after updating to 3.7 I've now got it working and I'm running a migration. The initial run took about 3 times as long to finish as my previous migrations, maybe a good sign? Unfortunately it failed in the end, but I'm rerunning it right now, it might have been a temporary problem with the tmp storage.

**#21 - 10/15/2020 03:59 PM - ggaaney**

[adam.winberg@smhi.se](mailto:adam.winberg@smhi.se) wrote:

Results in 94 copied packages to the appstream-new repo and zero packages in the baseos-new repo.

That is...SO odd. OK, give me the pulp-versions you now have, I will reset my env to use those releases and see if I can recreate.

Also - is this with migrated repos, or sync-into-pulp3-from-scratch repos? Want to see if we can isolate to just Pulp3, or if migration is somehow involved.

**#22 - 10/15/2020 04:04 PM - adam.winberg@smhi.se**

Now I have the following versions:

```
python3-pulpcore-3.7.1-3.el8.noarch  
python3-pulp-rpm-3.7.0-1.el8.noarch  
python3-pulp-2to3-migration-0.4.0-1.el8.noarch
```

This is with migrated repos. I will do a test with pulp3-from-scratch repos as well.

**#23 - 10/15/2020 08:31 PM - adam.winberg@smhi.se**

Exactly the same result when using newly created and synced pulp3 repos as source repos - 94 packages in appstream-new repo, 0 packages in baseos-new.

**#24 - 10/15/2020 08:48 PM - ggainey**

[adam.winberg@smhi.se](mailto:adam.winberg@smhi.se) wrote:

Exactly the same result when using newly created and synced pulp3 repos as source repos - 94 packages in appstream-new repo, 0 packages in baseos-new.

Well, that's mildly heartening - it means whatever is going on, isn't triggered by migration leaving data in an odd state.

With scratch-repos, running against the 3.7 branch for both pulpcore and pulp\_rpm, I still can't reproduce here. There's something Going On in your environment, I think; between us, we'll figure it out (he said hopefully)

I may not have more results right away, I'm on PTO tomorrow and Monday. Will monitor this issue in case anyone has a brilliant idea.

**#25 - 10/15/2020 11:48 PM - rchan**

- Sprint changed from Sprint 83 to Sprint 84

**#26 - 10/22/2020 03:52 PM - adam.winberg@smhi.se**

I've tried again but same result. Since my result is exactly the same as [ggainey](#) ends up with when using 'dependency\_solving=false' there is obviously something about the dependency solving that doesn't work in my environment.

The setting does do something though, when I run with 'dependency\_solving=true' the copy takes ~50s to complete. Running the copy with 'dependency\_solving=false' takes ~8s. And running with 'true' copies 7 modulemds to my new-appstreams repo whereas running with 'false' results in no modulemds at all (but 94 rpms and 1 advisory).

How can I further debug this?

**#27 - 10/22/2020 04:18 PM - dalley**

Exactly the same result when using newly created and synced pulp3 repos as source repos - 94 packages in appstream-new repo, 0 packages in baseos-new.

Well, that's mildly heartening - it means whatever is going on, isn't triggered by migration leaving data in an odd state.

Well, not necessarily. It could be that the content has been migrated improperly (the dependency information being wrong or something), and Pulp 3 isn't noticing this and is re-using the existing, incorrect content, with incorrect dependency information.

I'm not sure if this is actually the case, it's just an avenue of investigation.

@Adam Have you tried a complete wipe of the database yet, or have you just been deleting the repositories? (If you've already mentioned this I apologize),

**#28 - 10/22/2020 04:30 PM - adam.winberg@smhi.se**

It could be that the content is being migrated improperly (the dependency information being wrong or something), and Pulp 3 isn't noticing this



and is re-using the existing, incorrect content

Yes, I've actually thought about this as well.

Have you tried a complete wipe of the database yet, or have you just been deleting the repositories?

I've flushed my db between every sync attempt, but I have not tried to create a pulp3 repo and synced it against red hat *before* running the migration. I will try this (drop and recreate db, create pulp3 RHEL repos and remotes and sync, then try copy. Without running any 2to3-migration).

#### #29 - 10/22/2020 08:33 PM - ggaaney

I am getting back to this issue, sorry for the delays. I am working on the exact reverse of Adam's latest, I am sync'ing RHEL8 in the "Pulp2 Side" of my box, and will then migrate it to Pulp3 and retry my test-scenario, to see if I can recreate Adam's system's (mis)behavior.

#### #30 - 10/23/2020 09:02 AM - adam.winberg@smhi.se

My test is done:

- drop pulp db and create it + run migrations
- create repo and remotes for RHEL baseos and appstream and sync
- create baseos-new and appstream-new repos as copy targets
- no 2to3 migrations involved at all

It works! Or at least it produces content in the baseos-new repo, 381 packages. It will be interesting to see if [ggaaney](#) can reproduce my issue with his test.

#### #31 - 10/25/2020 04:09 PM - adam.winberg@smhi.se

Reran the migration with 2to3-migration 0.5.0 (was at 0.4.0) before and it resulted in a number of errors showing inconsistencies in my pulp2 db. Packages which was marked as downloaded in pulp2 was missing on my filesystem. These errors did not show up with 0.4.0. After fixing my pulp2 db the migration completed without errors but I still get the same behaviour on copy.

BUT, this time I get errors in the log while doing the copy:

```
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by polkit-0:0.115-x86_64.x8
6_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libiscsi-0:1.18.0-x86_64
.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libstdc++-0:8.2.1-x86_64
.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libguestfs-1:1.38.4-x86_
64.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libguestfs-1:1.38.4-x86_
64.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by hivex-0:1.3.15-x86_64.x8
6_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /usr/bin/bash needed by glibc-common-0:2.28-x86_6
4.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /usr/bin/pkg-config needed by netcf-devel-0:0.2.8
-x86_64.x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libtirpc-0:1.1.4-x86_64.
x86_64
Oct 25 14:58:07 lxserv2285.smhi.se rq[3097790]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems so
lving dependencies, copy may be incomplete: nothing provides /usr/bin/perl needed by perl-hivex-0:1.3.15-x86_6
4.x86_64
```

#### #32 - 10/25/2020 06:14 PM - ggaaney

@adam.winberg@smhi.se - congratulations, you and I are now getting the same results (he says, sadly)

Detailed results are at the end of this, but the net is a) reproduced Adam's results, but only after migrating RHEL8 from pulp2, b) there are a BUNCH of dependency failures, and c) they are all related to file-dependencies. I am probably not going to get all the way through that today, as I have other commitments - but at least we seem to be on the right trail now!

[dalley](#) - could this be related to <https://pulp.plan.io/issues/7202> ? That bug is fixed in 3.7, though.

Next step is to start clean against pulpcore/pulp\_rpm master, and see if it continues to happen.

Full testing details:

After restarting a clean environment I can reproduce the problem, including the errors from [#31](#). I did the following:

- cd pulpcore; git checkout 3.7
- cd ../pulp\_rpm; git checkout 3.7
- Use pulplift to create a new pulp2-nightly-pulp3-source-centos7
- Yum update to get pulp2.21.4 (this was a mistake, I forgot 2.21.4 was in staging repos (which is weird, since I'm the fin release-nanny for 2.21.4...)) - anyway, harmless
- Create and sync rhel8-base and rhel8-appstream into pulp2:

```
export BASE="rhel8-base"
BASE_REMOTE=https://cdn.redhat.com/content/dist/rhel8/8/x86_64/baseos/os/
export APPS="rhel8-appstream"
export APPS_REMOTE=https://cdn.redhat.com/content/dist/rhel8/8/x86_64/appstream/os/
pulp-admin rpm repo create --serve-http=true --repo-id=$BASE --relative-url=$BASE --feed=$BASE_REMOTE --download-policy on_demand --feed-key /home/vagrant/devel/pulp_startup/CDN_cert/cdn-key-and.pem --feed-cert /home/vagrant/devel/pulp_startup/CDN_cert/cdn-key-and.pem --feed-ca-cert /home/vagrant/devel/pulp_startup/CDN_cert/redhat-uep.pem
pulp-admin rpm repo create --serve-http=true --repo-id=$APPS --relative-url=$APPS --feed=$APPS_REMOTE --download-policy on_demand --feed-key /home/vagrant/devel/pulp_startup/CDN_cert/cdn-key-and.pem --feed-cert /home/vagrant/devel/pulp_startup/CDN_cert/cdn-key-and.pem --feed-ca-cert /home/vagrant/devel/pulp_startup/CDN_cert/redhat-uep.pem
pulp-admin rpm repo sync run --repo-id=$BASE
pulp-admin rpm repo sync run --repo-id=$APPS
```

- Migrate all RPM content to Pulp3:

```
http POST :/pulp/api/v3/migration-plans/ plan='{ "plugins": [ { "type": "rpm" } ] }'
```

- Create rhel8-base-new and rhel8-appstream-new

```
RHEL8_NAME="rhel8"
RHEL8_BASE_NAME="rhel8-base"
RHEL8_APPS_NAME="rhel8-apps"
ADVISORY="RHSA-2020:4059"
## create new repos
NEW_BASE_HREF=$(http POST :/pulp/api/v3/repositories/rpm/rpm/ name=${RHEL8_BASE_NAME}-new3 | jq -r '.pulp_href')
echo "${RHEL8_BASE_NAME}-new : " $NEW_BASE_HREF
if [ -z "$NEW_BASE_HREF" ]; then exit; fi
NEW_APPS_HREF=$(http POST :/pulp/api/v3/repositories/rpm/rpm/ name=${RHEL8_APPS_NAME}-new3 | jq -r '.pulp_href')
echo "${RHEL8_APPS_NAME}-new : " $NEW_APPS_HREF
if [ -z "$NEW_APPS_HREF" ]; then exit; fi
## find advisory-href
ADVISORY_HREF=$(http :/pulp/api/v3/content/rpm/advisories/?id=${ADVISORY} | jq -r '.results[] | .pulp_href')
echo "ADVISORY ${ADVISORY_HREF}"
```

- Copy from rhel8-appstream to rhel8-appstream-new, just advisory , with rhel8-base as a secondary repo, using this config:

```
cat config4.json
{
  "config": [
    {
      "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/871fea49-f1bb-4b32-871b-f2d0b739b89e/versions/1/",
      "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/",
      "content": ["/pulp/api/v3/content/rpm/advisories/2e3556bf-14ff-4690-9efc-096c97858cfe/"],
      {
        "source_repo_version": "/pulp/api/v3/repositories/rpm/rpm/a7c3836c-65cc-4d4a-b1b4-3d2cd1766c1f/versions/1/",
        "dest_repo": "/pulp/api/v3/repositories/rpm/rpm/e93b9628-0bf0-4dc7-bee7-b5151dae9a8e/",
        "content": []
      }
    ],
    "dependency_solving": true
  ]
}
http POST :/pulp/api/v3/rpm/copy/ @config4.json
```

- Ask what we got into rhel8-appstream-new:

```
http GET :/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/
{
  "base_version": null,
  "content_summary": {
    "added": {
```

```

    "rpm.advisory": {
      "count": 1,
      "href": "/pulp/api/v3/content/rpm/advisories/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    },
    "rpm.modulemd": {
      "count": 7,
      "href": "/pulp/api/v3/content/rpm/modulemds/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    },
    "rpm.package": {
      "count": 94,
      "href": "/pulp/api/v3/content/rpm/packages/?repository_version_added=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    }
  },
  "present": {
    "rpm.advisory": {
      "count": 1,
      "href": "/pulp/api/v3/content/rpm/advisories/?repository_version=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    },
    "rpm.modulemd": {
      "count": 7,
      "href": "/pulp/api/v3/content/rpm/modulemds/?repository_version=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    },
    "rpm.package": {
      "count": 94,
      "href": "/pulp/api/v3/content/rpm/packages/?repository_version=/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
    }
  },
  "removed": {}
},
"number": 1,
"pulp_created": "2020-10-25T16:17:45.762510Z",
"pulp_href": "/pulp/api/v3/repositories/rpm/rpm/1e2c5ffb-1eb6-4fc9-a2c2-830a1989da4b/versions/1/"
}

```

- The following errors in the logs:

```

Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/bin/bash needed by glibc-common-0:2.28-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by cyrus-sasl-lib-0:2.1.27-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libguestfs-1:1.38.4-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by cyrus-sasl-lib-0:2.1.27-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libguestfs-1:1.38.4-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by libgcrypt-0:1.8.3-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /sbin/ldconfig needed by cyrus-sasl-lib-0:2.1.27-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/sbin/gluster needed by libvirt-daemon-driver-storage-gluster-0:4.5.0-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/bin/pkg-config needed by bash-completion-1:2.7-noarch.noarch

```







```
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/sbin/mke2fs needed by supermin-0:5.1.19-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/bin/perl needed by libguestfs-benchmarking-1:1.38.4-x86_64.x86_64
Oct 25 16:17:45 pulp2-nightly-pulp3-source-centos7.padre-fedora.example.com rq[26935]: pulp: pulp_rpm.app.depsolving:WARNING: Encountered problems solving dependencies, copy may be incomplete: nothing provides /usr/bin/pkg-config needed by libiscsi-devel-0:1.18.0-x86_64.x86_64
```

### #33 - 10/26/2020 04:08 AM - dalley

- Project changed from RPM Support to Migration Plugin

[ggainey](#) It's not the same issue but it's similar(ish). The effects are the same. The migration plugin isn't migrating the filelist metadata correctly, confirmed by comparing the content metadata created for the "bash" package when migrated against the metadata created when saved as part of a pure Pulp 3 sync.

migrated from pulp2:

```
'files': [],
```

fresh w/ pulp 3:

```
'files': [
  [None, '/etc/skel/', '.bash_logout'],
  [None, '/etc/skel/', '.bash_profile'],
  [None, '/etc/skel/', '.bashrc'],
  [None, '/usr/bin/', 'alias'],
  [None, '/usr/bin/', 'bash'],
  [None, '/usr/bin/', 'bashbug'],
  [None, '/usr/bin/', 'bashbug-64'],
  [None, '/usr/bin/', 'bg'],
  [None, '/usr/bin/', 'cd'],
  [None, '/usr/bin/', 'command'],
  [None, '/usr/bin/', 'fc'],
  [None, '/usr/bin/', 'fg'],
  [None, '/usr/bin/', 'getopts'],
  [None, '/usr/bin/', 'hash'],
  [None, '/usr/bin/', 'jobs'],
  [None, '/usr/bin/', 'read'],
  [None, '/usr/bin/', 'sh'],
  [None, '/usr/bin/', 'type'],
  [None, '/usr/bin/', 'ulimit'],
  [None, '/usr/bin/', 'umask'],
  [None, '/usr/bin/', 'unalias'],
  [None, '/usr/bin/', 'wait'],
  ['dir', '/usr/lib/', '.build-id'],
  ['dir', '/usr/lib/.build-id/', 'c6'],
  [None, '/usr/lib/.build-id/c6/', '651902765f4cb2f5f0b2dd9c136b4486083087'],
  ['dir', '/usr/share/doc/', 'bash'],
  ...
  ...
  (there's a lot of them)
```

### #34 - 10/26/2020 04:19 AM - dalley

- Priority changed from Normal to High

- Severity changed from 2. Medium to 3. High

### #35 - 10/26/2020 05:30 AM - dalley

I narrowed it down to a misbehavior here [0]. The XML going in is fine and matches the standard XML (except for the order of the params), the `createrepo_c` object coming out is incomplete.

This is mostly identical to the code in the RPM sync task, with the exception of using a slightly different function provided by `createrepo_c` [1]

The standard form `cr.xml_parse_filelists()` works fine, the form used by the migration plugin `cr.xml_parse_filelists_snippet()` appears not to. I'm not sure if this is due to improper usage or bugginess. We are almost certainly the only ones who've ever used this function (it was added for our use case) so it could easily be either. Looks like there are a couple of upstream tests for it though [2]

[0] [https://github.com/pulp/pulp-2to3-migration/blob/master/pulp\\_2to3\\_migration/app/plugin/rpm/xml\\_utils.py#L233](https://github.com/pulp/pulp-2to3-migration/blob/master/pulp_2to3_migration/app/plugin/rpm/xml_utils.py#L233)

[1] [https://github.com/pulp/pulp\\_rpm/blob/master/pulp\\_rpm/app/tasks/synchronizing.py#L410](https://github.com/pulp/pulp_rpm/blob/master/pulp_rpm/app/tasks/synchronizing.py#L410)

[2] [https://github.com/rpm-software-management/createrepo\\_c/blob/master/tests/python/tests/test\\_xml\\_parser.py#L364-L512](https://github.com/rpm-software-management/createrepo_c/blob/master/tests/python/tests/test_xml_parser.py#L364-L512)

**#36 - 10/26/2020 04:39 PM - dalley**

- Assignee changed from ggainey to dalley

**#37 - 10/26/2020 04:48 PM - dalley**

- Subject changed from Unexpected results from multi-repository copy to RPM Migration results in incomplete package metadata

- Description updated

**#38 - 10/26/2020 06:53 PM - dalley**

- Status changed from ASSIGNED to POST

<https://github.com/pulp/pulp-2to3-migration/pull/257>

**#39 - 10/26/2020 08:29 PM - dalley**

- Status changed from POST to MODIFIED

Applied in changeset [pulp:pulp-2to3-migration|d88dadd0b4ad40a112c250b0a78cc79593b3c8e9](https://github.com/pulp/pulp-2to3-migration/commit/d88dadd0b4ad40a112c250b0a78cc79593b3c8e9).

**#40 - 10/26/2020 08:33 PM - dalley**

@Adam We'll push out a 0.5.1 with the fix, it may take a few days to get pulled into the Foreman repos. Once the update hits, you need to wipe your Pulp 3 database again, and then after one more migration from Pulp 2 everything should work *crosses fingers*.

**#41 - 10/27/2020 01:26 PM - adam.winberg@smhi.se**

Thats great, but right now i am reluctant to run the 2to3 migration due to issue [#7744](#) (when I run the migration several thousand rpm files are deleted from my pulp2 content dir).

**#42 - 10/27/2020 02:40 PM - ttereshc**

- Sprint/Milestone set to 0.5.1

**#43 - 10/27/2020 08:14 PM - pulpbot**

- Status changed from MODIFIED to CLOSED - CURRENTRELEASE

**#44 - 11/07/2020 05:22 PM - adam.winberg@smhi.se**

I have now run migration with 0.5.1 and pulpcore 3.7.3 and the depsolving does indeed seem to work now.