

## RPM Support - Issue #4617

### Migration of modular content can fail (upgrade from <=2.16 to 2.17+)

04/01/2019 05:55 PM - kfiresmith

<b>Status:</b>	CLOSED - CURRENTRELEASE	<b>Start date:</b>	
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	dkliban@redhat.com	<b>Estimated time:</b>	0:00 hour
<b>Category:</b>			
<b>Sprint/Milestone:</b>	2.19.1		
<b>Severity:</b>	3. High	<b>Groomed:</b>	No
<b>Version:</b>		<b>Sprint Candidate:</b>	No
<b>Platform Release:</b>	2.19.1	<b>Tags:</b>	Pulp 2
<b>OS:</b>	RHEL 7	<b>Sprint:</b>	Sprint 51
<b>Triaged:</b>	No	<b>Quarter:</b>	
<b>Description</b>			
<p>Hello folks, Our upgrade of a (albeit smaller but otherwise identical) beta Pulp server went just fine, so this morning I upgraded production which went fine until the database migration.</p> <p>The migration failed very early at 'Applying migration pulp_rpm.plugins.migrations.0043_add_modulemd_modulemd-defaults failed.' The underlying traceback shows 'IOError: Not a gzipped file'.</p> <p>Full paste lives here: <a href="https://paste.fedoraproject.org/paste/wElu5a3Tf8OejEzMNxXN9w">https://paste.fedoraproject.org/paste/wElu5a3Tf8OejEzMNxXN9w</a></p> <p>Packages installed are:</p> <pre>\$rpm -qa   egrep 'pulp qpid gofer celery mongo'   sort mongodb-2.6.12-6.el7.x86_64 mongodb-server-2.6.12-6.el7.x86_64 pulp-admin-client-2.18.1-2.el7.noarch pulp-docker-admin-extensions-3.2.2-1.el7.noarch pulp-docker-plugins-3.2.2-1.el7.noarch pulp-puppet-admin-extensions-2.18.1-1.el7.noarch pulp-puppet-plugins-2.18.1-1.el7.noarch pulp-rpm-admin-extensions-2.18.1-1.el7.noarch pulp-rpm-plugins-2.18.1-1.el7.noarch pulp-selinux-2.18.1-2.el7.noarch pulp-server-2.18.1-2.el7.noarch python2-celery-4.0.2-7.el7.noarch python2-qpid-1.37.0-4.el7.noarch python2-qpid-qmf-1.39.0-1.el7.x86_64 python-gofer-2.12.5-1.el7.noarch python-gofer-qpid-2.12.5-1.el7.noarch python-isodate-0.5.0-4.pulp.el7.noarch python-mongoengine-0.10.5-1.el7.noarch python-pulp-bindings-2.18.1-2.el7.noarch python-pulp-client-lib-2.18.1-2.el7.noarch python-pulp-common-2.18.1-2.el7.noarch python-pulp-docker-common-3.2.2-1.el7.noarch python-pulp-oid_validation-2.18.1-2.el7.noarch python-pulp-puppet-common-2.18.1-1.el7.noarch python-pulp-repoauth-2.18.1-2.el7.noarch python-pulp-rpm-common-2.18.1-1.el7.noarch python-pymongo-3.2-2.el7.x86_64 python-pymongo-gridfs-3.2-2.el7.x86_64 qpid-cpp-client-1.39.0-1.el7.x86_64 qpid-cpp-server-1.39.0-1.el7.x86_64 qpid-cpp-server-linearstore-1.39.0-1.el7.x86_64 qpid-proton-c-0.26.0-1.el7.x86_64 qpid-qmf-1.39.0-1.el7.x86_64</pre>			

I'm a bit stumped because I'm not a developer, or a database wiz, so any help in getting our prod instance back up would be very much appreciated! Thanks!

## Summary

Some repos unexpectedly contain uncompressed modules.yaml metadata.  
Pulp should not fail but try to work with uncompressed data instead.

```
Applying pulp_rpm.plugins.migrations version 43
Loading unit model: puppet_module = pulp_puppet.plugins.db.models:Module
Loading unit model: docker_manifest = pulp_docker.plugins.models:Manifest
Loading unit model: docker_blob = pulp_docker.plugins.models:Blob
Loading unit model: docker_manifest_list = pulp_docker.plugins.models:ManifestList
Loading unit model: docker_image = pulp_docker.plugins.models:Image
Loading unit model: docker_tag = pulp_docker.plugins.models:Tag
Loading unit model: modulemd_defaults = pulp_rpm.plugins.db.models:ModulemdDefaults
Loading unit model: erratum = pulp_rpm.plugins.db.models:Errata
Loading unit model: distribution = pulp_rpm.plugins.db.models:Distribution
Loading unit model: srpm = pulp_rpm.plugins.db.models:SRPM
Loading unit model: package_group = pulp_rpm.plugins.db.models:PackageGroup
Loading unit model: package_category = pulp_rpm.plugins.db.models:PackageCategory
Loading unit model: iso = pulp_rpm.plugins.db.models:ISO
Loading unit model: package_environment = pulp_rpm.plugins.db.models:PackageEnvironment
Loading unit model: drpm = pulp_rpm.plugins.db.models:DRPM
Loading unit model: package_langpacks = pulp_rpm.plugins.db.models:PackageLangpacks
Loading unit model: modulemd = pulp_rpm.plugins.db.models:Modulemd
Loading unit model: rpm = pulp_rpm.plugins.db.models:RPM
Loading unit model: yum_repo_metadata_file = pulp_rpm.plugins.db.models:YumMetadataFile
Loading auxiliary model: erratum_pkglists = pulp_rpm.plugins.db.models:ErratumPkglis
Loaded plugin puppet_distributor for types: puppet_module
Loaded plugin puppet_file_distributor for types: puppet_module
Loaded plugin puppet_install_distributor for types: puppet_module
Loaded plugin docker_distributor_web for types: docker_image,docker_blob,docker_manifest,docker_tag
Loaded plugin docker_rsync_distributor for types: docker_image,docker_blob,docker_manifest,docker_tag
Loaded plugin docker_distributor_export for types: docker_image
Loaded plugin yum_distributor for types: rpm,srpm,drpm,erratum,package_group,package_category,distribution,yum_repo_metadata_file
Loaded plugin export_distributor for types: rpm,srpm,drpm,erratum,distribution,package_category,package_group
Loaded plugin rpm_rsync_distributor for types: rpm,srpm,drpm,erratum,package_group,package_category,distribution,yum_repo_metadata_file
Loaded plugin iso_distributor for types: iso
Loaded plugin iso_rsync_distributor for types: iso
Loaded plugin group_export_distributor for types: rpm,srpm,drpm,erratum,distribution,package_category,package_group
Loaded plugin puppet_importer for types: puppet_module
Loaded plugin docker_importer for types: docker_blob,docker_image,docker_manifest,docker_manifest_list,docker_tag
Loaded plugin yum_importer for types: distribution,drpm,erratum,package_group,package_category,rpm,srpm,yum_repo_metadata_file,package_environment,package_langpacks,modulemd,modulemd_defaults
Loaded plugin iso_importer for types: iso
Loaded plugin puppet_whole_repo_profiler for types: puppet_module
Loaded plugin yum_profiler for types: rpm,erratum,modulemd
Loaded plugin yum for types: rpm
Loaded plugin rhui for types: rpm
Applying migration pulp_rpm.plugins.migrations.0043_add_modulemd_modulemd-defaults failed.

Halting migrations due to a migration failure.
Not a gzipped file
Traceback (most recent call last):
  File "/usr/lib/python2.7/site-packages/pulp/server/db/manage.py", line 239, in main
    return _auto_manage_db(options)
  File "/usr/lib/python2.7/site-packages/pulp/server/db/manage.py", line 306, in _auto_manage_db
```

```
migrate_database(options)
File "/usr/lib/python2.7/site-packages/pulp/server/db/manage.py", line 135, in migrate_database
    update_current_version=not options.test)
File "/usr/lib/python2.7/site-packages/pulp/server/db/migrate/models.py", line 189, in apply_migration
    migration.migrate()
File "/usr/lib/python2.7/site-packages/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py", line 341, in migrate
    modulemds, defaults = load(fp, working_dir)
File "/usr/lib/python2.7/site-packages/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py", line 317, in load
    bfr = metadata.read(1024000)
File "/usr/lib64/python2.7/gzip.py", line 261, in read
    self._read(readsize)
File "/usr/lib64/python2.7/gzip.py", line 296, in _read
    self._read_gzip_header()
File "/usr/lib64/python2.7/gzip.py", line 190, in _read_gzip_header
    raise IOError, 'Not a gzipped file'
IOError: Not a gzipped file
```

#### Related issues:

Has duplicate RPM Support - Issue #4619: Migration of modular content can fail...

**CLOSED - DUPLICATE**

Copied to RPM Support - Test #4729: Migration of modular content can fail (up...

**CLOSED - WONTFIX**

#### Associated revisions

##### Revision 4e4a6b4a - 04/02/2019 03:14 PM - dkliban@redhat.com

Problem: migration fails for uncompressed modules files

Solution: handle IOError and read the uncompressed file

fixes: #4617 <https://pulp.plan.io/issues/4617>

##### Revision dc9cb3be - 05/16/2019 07:56 PM - dkliban@redhat.com

Problem: migration fails for uncompressed modules files

Solution: handle IOError and read the uncompressed file

fixes: #4617 <https://pulp.plan.io/issues/4617> (cherry picked from commit 4e4a6b4ad4514df16f07ce43fd2e197695f4cbf5)

#### History

##### #1 - 04/01/2019 09:18 PM - dkliban@redhat.com

Looks like the problem was with some of the content on the CDN. Here is a patch that fixes the problem. We'll include this fix in 2.19.1.

```
diff --git a/plugins/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py b/plugins/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py
index 7736a5a8..bb2bcb55 100644
--- a/plugins/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py
+++ b/plugins/pulp_rpm/plugins/migrations/0043_add_modulemd_modulemd-defaults.py
@@ -336,14 +336,26 @@ def migrate(*args, **kwargs):
     repos_to_republish = set()
     for file in metadatafiles:
         repository = Repository.objects.get(repo_id=file.repo_id)
-        with gzip.open(file._storage_path, 'r') as fp:
-            working_dir = "/var/cache/pulp"
-            modulemds, defaults = load(fp, working_dir)
-            add_modulemds(repository, modulemds, working_dir)
-            add_defaults(repository, defaults, working_dir)
-            repository_controller.disassociate_units(repository, [file])
-            repository_controller.rebuild_content_unit_counts(repository)
-            repos_to_republish.add(file.repo_id)
+        try:
+            with gzip.open(file._storage_path, 'r') as fp:
+                working_dir = "/var/cache/pulp"
+                modulemds, defaults = load(fp, working_dir)
+                add_modulemds(repository, modulemds, working_dir)
+                add_defaults(repository, defaults, working_dir)
+                repository_controller.disassociate_units(repository, [file])
+                repository_controller.rebuild_content_unit_counts(repository)
+            repos_to_republish.add(file.repo_id)
+        except IOError:
+            # Not a gzipped file, read uncompressed
+            with open(file._storage_path, 'r') as fp:
+                working_dir = "/var/cache/pulp"
+                modulemds, defaults = load(fp, working_dir)
+                add_modulemds(repository, modulemds, working_dir)
+                add_defaults(repository, defaults, working_dir)
+                repository_controller.disassociate_units(repository, [file])
+                repository_controller.rebuild_content_unit_counts(repository)
```

```

+         repos_to_republish.add(file.repo_id)
+     except IOError:
+         _logger.info(file.unit_key)
+         with open(file._storage_path, 'r') as fp:
+             working_dir = "/var/cache/pulp"
+             modulemds, defaults = load(fp, working_dir)
+             add_modulemds(repository, modulemds, working_dir)
+             add_defaults(repository, defaults, working_dir)
+             repository_controller.disassociate_units(repository, [file])
+             repository_controller.rebuild_content_unit_counts(repository)
+             repos_to_republish.add(file.repo_id)
+
+     if repos_to_republish:
+         with open('/var/lib/pulp/0043_add_modulemd-modulemd-defaults.txt', 'w') as f:
+             f.write(str(list(repos_to_republish)))

```

## #2 - 04/01/2019 09:28 PM - ttereshc

- Project changed from 31 to RPM Support

- Subject changed from Pulp failed upgrade on production 2.16.3 --> 2.18.1 (IO Error: not a gzipped file) to Migration of modular content can fail (upgrade from <=2.16 to 2.17+)

- Description updated

## #3 - 04/01/2019 09:28 PM - ttereshc

- Has duplicate Issue #4619: Migration of modular content can fail (upgrade from <=2.16 to 2.17+) added

## #4 - 04/01/2019 10:24 PM - kfiresmith

Thanks very much DKliban. I'm back in business. For what it's worth for the rest of Pulp 2.x's lifespan, trapping exceptions as done in the patch with some logging goodness to stdout at bail-time would be very much welcome to all those who may encounter such a traceback.

## #6 - 04/02/2019 03:13 PM - dkliban@redhat.com

- Status changed from NEW to ASSIGNED

- Assignee set to dkliban@redhat.com

## #7 - 04/02/2019 03:58 PM - dkliban@redhat.com

- Status changed from ASSIGNED to MODIFIED

Applied in changeset [4e4a6b4ad4514df16f07ce43fd2e197695f4cbf5](#).

## #8 - 04/12/2019 11:56 AM - ttereshc

- Sprint set to Sprint 51

## #9 - 04/15/2019 09:57 PM - bmbouter

- Tags Pulp 2 added

## #10 - 04/23/2019 09:44 PM - bherring

- Copied to Test #4729: Migration of modular content can fail (upgrade from <=2.16 to 2.17+) added

## #11 - 05/02/2019 02:54 PM - dkliban@redhat.com

- Platform Release set to 2.19.1

## #12 - 05/03/2019 09:33 PM - dkliban@redhat.com

- Sprint/Milestone set to 2.19.1

## #13 - 05/21/2019 02:15 PM - dkliban@redhat.com

Applied in changeset [dc9cb3bee2f183b80ea35bff5084957bc1be93fa](#).

## #14 - 05/22/2019 09:26 PM - dkliban@redhat.com

- Status changed from MODIFIED to 5

## #15 - 05/30/2019 06:42 PM - dkliban@redhat.com

- Status changed from 5 to CLOSED - CURRENTRELEASE