

## RPM Support - Issue #1808

### exporting a sufficiently large repo with 'on\_demand' policy results in BSON error

04/04/2016 04:39 PM - cduryee

<b>Status:</b>	CLOSED - CURRENTRELEASE	<b>Start date:</b>	
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	mhrivnak	<b>Estimated time:</b>	0:00 hour
<b>Category:</b>			
<b>Sprint/Milestone:</b>			
<b>Severity:</b>	2. Medium	<b>Groomed:</b>	No
<b>Version:</b>		<b>Sprint Candidate:</b>	No
<b>Platform Release:</b>	2.8.3	<b>Tags:</b>	Pulp 2
<b>OS:</b>		<b>Sprint:</b>	Sprint 1
<b>Triaged:</b>	Yes	<b>Quarter:</b>	

**Description**

If you attempt to export a repo with many RPMs, you'll get a BSON document too large error.

To repro:

- create a large repo with "on\_demand" policy, such as Fedora 23 everything
- sync repo
- create repo group with said repo
- attempt to publish repo group

actual result:

BSON document too large error

expected result:

error that either a.) not all units are on-disk and repo/repo group cannot be exported, or b.) 'on\_demand' repos cannot be exported

#### Associated revisions

##### Revision dd5f8c7e - 04/08/2016 11:06 PM - mhrivnak

copytree no longer accumulates an unbounded number of errors

<https://pulp.plan.io/issues/1808>

fixes #1808

##### Revision dd5f8c7e - 04/08/2016 11:06 PM - mhrivnak

copytree no longer accumulates an unbounded number of errors

<https://pulp.plan.io/issues/1808>

fixes #1808

#### History

##### #1 - 04/08/2016 10:10 PM - mhrivnak

A patch such as this will prevent the copytree() function from accumulating an unbounded number of errors before giving up. That unbounded collection of error messages is what caused the mongodb document size limit to be exceeded.

```
diff --git a/server/pulp/server/util.py b/server/pulp/server/util.py
index 5411e87..98fede3 100644
--- a/server/pulp/server/util.py
+++ b/server/pulp/server/util.py
```

```
@@ -207,6 +207,8 @@ def copytree(src, dst, symlinks=False, ignore=None):
    os.makedirs(dst)
    errors = []
    for name in names:
+       if len(errors) >= 100:
+           raise Error(errors)
    if name in ignored_names:
        continue
    srcname = os.path.join(src, name)
```

**#2 - 04/08/2016 10:56 PM - mhrivnak**

- Status changed from NEW to ASSIGNED
- Assignee set to mhrivnak
- Sprint/Milestone set to 19

**#3 - 04/08/2016 11:09 PM - mhrivnak**

- Status changed from ASSIGNED to POST

<https://github.com/pulp/pulp/pull/2518>

**#4 - 04/11/2016 04:10 PM - mhrivnak**

- Platform Release set to 2.8.3

**#6 - 04/14/2016 03:18 PM - mhrivnak**

- Status changed from POST to MODIFIED
- % Done changed from 0 to 100

Applied in changeset [pulp:pulp|dd5f8c7ea40cbdc0f57c1aa1c88bbe3f4384dfff](https://github.com/pulp/pulp/pull/2518).

**#7 - 04/15/2016 04:40 PM - mhrivnak**

- Triaged changed from No to Yes

**#8 - 04/21/2016 09:04 PM - pthomas@redhat.com**

Verified

```
pre>
[root@mgmt5 ~]# pulp-admin rpm repo export run --repo-id rhel67+-----+
Publishing Repository [rhel67]
```

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

Task Failed

The repository cannot be exported because some units are not downloaded.

```
[root@mgmt5 ~]# pulp-admin rpm repo group export run --group-id grp1
Exporting Repository Group [grp1]
```

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

Task Failed

The repository group cannot be exported because these repos have units that are not downloaded: zoo, rhel67

```
[root@mgmt5 ~]#
```

**#9 - 04/27/2016 12:38 AM - semyers**

- Status changed from MODIFIED to 5

**#10 - 05/17/2016 09:31 PM - semyers**

- Status changed from 5 to CLOSED - CURRENTRELEASE

**#11 - 03/08/2018 07:21 PM - bmbouter**

- *Sprint set to Sprint 1*

**#12 - 03/08/2018 07:21 PM - bmbouter**

- *Sprint/Milestone deleted (19)*

**#13 - 04/15/2019 10:32 PM - bmbouter**

- *Tags Pulp 2 added*