

RPM Support - Issue #6476

Repomd compression behaves differently for S3

04/09/2020 09:15 PM - fao89

Status:	NEW	Start date:	
Priority:	Normal	Due date:	
Assignee:		Estimated time:	0:00 hour
Category:			
Sprint/Milestone:	Priority items (outside of planned milestones/releases)		
Severity:	2. Medium	Groomed:	No
Version:		Sprint Candidate:	No
Platform Release:		Tags:	
OS:		Sprint:	
Triaged:	Yes	Quarter:	

Description

on repomd we have: primary.xml.gz update_info.xml.gz When testing xml.gz behaves like it is not compressed On the other hand, it behaves as compressed on S3 testing

https://github.com/pulp/pulp_rpm/blob/master/pulp_rpm/tests/functional/api/test_publish.py#L51-L67

```
def read_xml_gz(content):
    """
    Read xml and xml.gz.
    Tests work normally but fails for S3 due '.gz'
    Why is it only compressed for S3?
    """
    with NamedTemporaryFile() as temp_file:
        temp_file.write(content)
        temp_file.seek(0)

    try:
        content_xml = gzip.open(temp_file.name).read()
    except OSError:
        # FIXME: fix this as in CI primary/update_info.xml has '.gz' but it is not gzipped
        content_xml = temp_file.read()
    return content_xml
```

History

#1 - 04/16/2020 05:39 PM - fao89

- Description updated

#2 - 04/16/2020 05:40 PM - fao89

- Description updated

#3 - 04/23/2020 08:05 PM - fao89

I tested with real S3 bucket and got the same behavior, with:

```
DEFAULT_FILE_STORAGE = 'storages.backends.s3boto3.S3Boto3Storage'
MEDIA_ROOT = ''
```

primary.xml.gz is compressed, as we expect but for:

```
DEFAULT_FILE_STORAGE = 'pulpcore.app.models.storage.FileSystem'
MEDIA_ROOT = '/var/lib/pulp/'
```

primary.xml.gz does have the gz suffix, but it is not compressed

#4 - 04/23/2020 08:31 PM - ttereshc

- *Sprint/Milestone set to Priority items (outside of planned milestones/releases)*
- *Triaged changed from No to Yes*