

Arvados - Story #14965

Port arv-mount to Python 3

03/13/2019 04:10 PM - Tom Morris

| | |
|--|---------------------------------------|
| Status: Resolved | Start date: 07/16/2019 |
| Priority: Normal | Due date: |
| Assigned To: Eric Biagiotti | % Done: 100% |
| Category: | Estimated time: 0.00 hour |
| Target version: 2019-07-31 Sprint | |
| Description | |
| Subtasks: | |
| Task # 15328: Review 14965-arv-mount-py-three | Resolved |
| Related issues: | |
| Related to Arvados Epics - Story #14532: [Epic] Port to Python 3 to for Pytho... | Resolved 01/01/2020 09/16/2020 |
| Has duplicate Arvados - Bug #15251: arv-mount supports python 3 | Duplicate |

Associated revisions

Revision caeb9d74 - 06/17/2019 02:50 PM - Eric Biagiotti

Merge remote-tracking branch 'origin/14965-arv-mount-py-three' into 14965-arv-mount-py-three

refs #14965
14965

Arvados-DCO-1.1-Signed-off-by: Eric Biagiotti <ebiagiotti@veritasgenetics.com>

Revision b629d9e8 - 06/26/2019 03:39 PM - Eric Biagiotti

Merge remote-tracking branch 'origin/master' into 14965-arv-mount-py-three

refs #14965

Arvados-DCO-1.1-Signed-off-by: Eric Biagiotti <ebiagiotti@veritasgenetics.com>

Revision 76530546 - 07/16/2019 04:29 PM - Eric Biagiotti

Merge branch 'master' into 14965-arv-mount-py-three

refs #14965

Arvados-DCO-1.1-Signed-off-by: Eric Biagiotti <ebiagiotti@veritasgenetics.com>

Revision 16d9c611 - 07/16/2019 06:44 PM - Eric Biagiotti

Merge branch '14965-arv-mount-py-three' of git.curoverse.com:arvados into 14965-arv-mount-py-three

refs #14965
14965

Arvados-DCO-1.1-Signed-off-by: Eric Biagiotti <ebiagiotti@veritasgenetics.com>

Revision 05c87aaf - 07/17/2019 02:38 PM - Eric Biagiotti

Merge branch 'master' into 14965-arv-mount-py-three

refs #14965

Arvados-DCO-1.1-Signed-off-by: Eric Biagiotti <ebiagiotti@veritasgenetics.com>

Revision c1cf71ba - 07/19/2019 02:56 PM - Eric Biagiotti

Merge branch '14965-arv-mount-py-three'

refs #14965

History

#1 - 05/21/2019 07:27 PM - Tom Morris

- Parent task deleted (#14532)

#2 - 05/21/2019 07:28 PM - Tom Morris

- Tracker changed from Task to Story

- Target version changed from Arvados Future Sprints to To Be Groomed

#3 - 05/21/2019 07:28 PM - Tom Morris

- Related to Story #14532: [Epic] Port to Python 3 to for Python 2 sunset in December 2019 added

#4 - 05/21/2019 07:28 PM - Tom Morris

- Has duplicate Bug #15251: arv-mount supports python 3 added

#5 - 05/29/2019 02:49 PM - Tom Morris

- Target version changed from To Be Groomed to Arvados Future Sprints

- Story points set to 2.0

#6 - 06/05/2019 03:37 PM - Eric Biagiotti

- Assigned To set to Eric Biagiotti

- Target version changed from Arvados Future Sprints to 2019-06-19 Sprint

#7 - 06/10/2019 02:16 PM - Ward Vandewege

- Release set to 22

#8 - 06/19/2019 01:13 PM - Eric Biagiotti

- Status changed from New to In Progress

#9 - 06/19/2019 01:14 PM - Eric Biagiotti

- Target version changed from 2019-06-19 Sprint to 2019-07-03 Sprint

#10 - 07/03/2019 01:15 PM - Eric Biagiotti

- Target version changed from 2019-07-03 Sprint to 2019-07-17 Sprint

#11 - 07/16/2019 07:13 PM - Eric Biagiotti

Latest at [16d9c611d5026e50692e1b3d0fefb951de5afa7b](https://github.com/Arvados/arv-mount/commit/16d9c611d5026e50692e1b3d0fefb951de5afa7b)

- arv-mount supports Python 3
- Besides the futurize changes, I mostly fixed test encoding issues. I tried to minimize changes to our actual fuse code.

Testing:

- <https://ci.curoverse.com/view/Developer/job/developer-run-tests/1400/>
- Also Manually tested arv-mount installation and some functionality in py3/py2 virtualenvs

#12 - 07/17/2019 01:46 PM - Eric Biagiotti

- Target version changed from 2019-07-17 Sprint to Arvados Future Sprints

#13 - 07/17/2019 01:48 PM - Eric Biagiotti

- Target version changed from Arvados Future Sprints to 2019-07-31 Sprint

#14 - 07/17/2019 07:00 PM - Lucas Di Pentima

Sorry for the delay, I was doing some manual testing on arvbox and it took me some time.

As we have had performance issues in the past with the fuse driver, some of my comments are headed towards that, following this Python 2/3 "cheat sheet": http://python-future.org/compatible_idioms.html

- When running tests, I'm getting lots of warning messages like this one: `/usr/lib/python3.5/json/decoder.py:355: ResourceWarning: unclosed <ssl.SSLSocket fd=21, family=AddressFamily.AF_INET, type=2049, proto=6, laddr=('127.0.0.1', 42742), raddr=('127.0.0.1', 47393)>` — as we talked on zoom, I'm worried about resources not being fully freed after a test run (may create issues on Jenkins slaves), but if that's not the case, I think it would still be good to avoid them for log readability purposes.
- Also got some deprecation warnings: `arvados/services/fuse/tests/test_exec.py:65: DeprecationWarning: Please use assertRegex instead.` (check <https://six.readthedocs.io/#unittest-assertions> for py2/py3 compatibility)
- `services/fuse/arvados_fuse/fusedir.py` - Lines 170 & 178: There're some `dict.itervalues()` that were translated to `dict.values()` calls, that would use extra memory on python2, as it's described on http://python-future.org/compatible_idioms.html#iterating-through-dict-keys-values-items, seems that the way to go is calling `itervalues(dict)`?
- `services/fuse/arvados_fuse/__init__.py:L276` — Using `dict.items()` is inefficient on Py2 as described on the previously mentioned guide list, maybe use `viewitems(dict)` or `iteritems(dict)`?
- `services/fuse/arvados_fuse/fusedir.py:L99` and `services/fuse/arvados_fuse/__init__.py:L727` — Using `list(dict.items())` is inefficient on Py2 as described on the previously mentioned guide list, maybe use `listitems(dict)`?

As for the manual tests, at first I initially got my arvbox shell instance to freeze when trying to write a file using py3 after trying it first with py2, but later on I did a proper test initially starting a shell and trying py2 and later closing the shell and re-starting it and doing the same with py3, and it worked without issues, so maybe I didn't unmounted the filesystem correctly at the first try.

Also, I think the documentation needs updating to state that it's compatible with Python3 <https://doc.arvados.org/sdk/python/arvados-fuse.html>

#15 - 07/19/2019 01:26 PM - Eric Biagiotti

Latest at <3ce47c5d012f9dbd6335ad0ae092eb9cc8a3022e>

Unit tests: <https://ci.curoverse.com/job/developer-run-tests/1404/> - Some workbench integration failures, but I ran them locally and they passed.
Conformance tests: <https://ci.curoverse.com/view/CWL/job/arvados-cwl-conformance-tests/193/>

Also retested manually in virtualenvs.

Lucas Di Pentima wrote:

- When running tests, I'm getting lots of warning messages like this one: `/usr/lib/python3.5/json/decoder.py:355: ResourceWarning: unclosed <ssl.SSLSocket fd=21, family=AddressFamily.AF_INET, type=2049, proto=6, laddr=('127.0.0.1', 42742), raddr=('127.0.0.1', 47393)>` — as we talked on zoom, I'm worried about resources not being fully freed after a test run (may create issues on Jenkins slaves), but if that's not the case, I think it would still be good to avoid them for log readability purposes.

I was able to reduce some of these with <a1a1a8756149d5d05526503fd1e85115ea083fac>, but the json and yaml ones have been difficult to pin down. I'm fairly certain these pertain to http2libs keep-alive feature <https://github.com/httplib2/httplib2#keep-alive>. Similar issues have been reported here <https://github.com/kennethreitz/requests/issues/3912>. The reason these only show up in testing is because they are normally suppressed by default in python, but unittest resets warning suppression.

They don't seem to be causing problems, but their impact on jenkins slaves is worth keeping in mind. It's also worth noting that these warnings show up in majority of our py3 test suites (`sdk/python`, `sdk/cwl`). For these reasons, I would like to leave the output as is and make a separate issue to investigate further. Let me know if you think this is satisfactory and I will make the issue.

- Also got some deprecation warnings: `arvados/services/fuse/tests/test_exec.py:65: DeprecationWarning: Please use assertRegex instead.` (check <https://six.readthedocs.io/#unittest-assertions> for py2/py3 compatibility)

Thanks for finding that! Fixed in latest.

- `services/fuse/arvados_fuse/fusedir.py` - Lines 170 & 178: There're some `dict.itervalues()` that were translated to `dict.values()` calls, that would use extra memory on python2, as it's described on http://python-future.org/compatible_idioms.html#iterating-through-dict-keys-values-items, seems that the way to go is calling `itervalues(dict)`?
- `services/fuse/arvados_fuse/__init__.py:L276` — Using `dict.items()` is inefficient on Py2 as described on the previously mentioned guide list, maybe use `viewitems(dict)` or `iteritems(dict)`?
- `services/fuse/arvados_fuse/fusedir.py:L99` and `services/fuse/arvados_fuse/__init__.py:L727` — Using `list(dict.items())` is inefficient on Py2 as described on the previously mentioned guide list, maybe use `listitems(dict)`?

Fixed all three in latest.

Also, I think the documentation needs updating to state that it's compatible with Python3 <https://doc.arvados.org/sdk/python/arvados-fuse.html>

That page will be updated once we create the py3 arv-mount package. I'm tracking it here: <https://dev.arvados.org/issues/15473>

#16 - 07/19/2019 02:02 PM - Lucas Di Pentima

Eric Biagiotti wrote:

Latest at <3ce47c5d012f9dbd6335ad0ae092eb9cc8a3022e>
[...]

They don't seem to be causing problems, but their impact on jenkins slaves is worth keeping in mind. It's also worth noting that these warnings show up in majority of our py3 test suites (sdk/python, sdk/cwl). For these reasons, I would like to leave the output as is and make a separate issue to investigate further. Let me know if you think this is satisfactory and I will make the issue.

Yes, thanks for doing the research! I wasn't aware of the warnings happening on other py3 test runs.

[...]

Also, I think the documentation needs updating to state that it's compatible with Python3
<https://doc.arvados.org/sdk/python/arvados-fuse.html>

That page will be updated once we create the py3 arv-mount package. I'm tracking it here: <https://dev.arvados.org/issues/15473>

Great, so this LGTM, thanks!

#18 - 07/19/2019 03:34 PM - Eric Biagiotti

- *Status changed from In Progress to Resolved*