

Arvados - Story #11419

[SDKs] support text-mode open() in Python 3

04/04/2017 03:41 PM - Tom Clegg

| | |
|----------------------------------------------------------------------------------|---------------------------------------|
| Status: Resolved | Start date: 03/21/2017 |
| Priority: Normal | Due date: |
| Assigned To: Tom Clegg | % Done: 100% |
| Category: | Estimated time: 0.00 hour |
| Target version: 2018-12-21 Sprint | |
| Description | |
| Subtasks: | |
| Task # 14547: Review 11419-text-mode | Resolved |
| Related issues: | |
| Related to Arvados - Story #11770: [Python SDK] Implement support for univers... | New 05/26/2017 |
| Related to Arvados - Story #11308: Support Python3 for arvados-python-client ... | Resolved 03/21/2017 |
| Blocks Arvados Epics - Story #14532: [Epic] Port to Python 3 to for Python 2 ... | Resolved 01/01/2020 09/16/2020 |

Associated revisions

Revision 7317196f - 12/10/2018 09:18 PM - Tom Clegg

Merge branch '11419-text-mode'

closes #11419

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

Revision d7d074d7 - 12/17/2018 02:58 PM - Tom Clegg

11419: Read JSON files in binary mode.

Files opened in text mode do not have a size() method.

refs #11419

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

History

#1 - 04/06/2017 02:03 PM - Tom Clegg

- Subject changed from open files in text or binary mode to support text-mode open() in Python 3

#2 - 05/10/2017 07:19 PM - Tom Clegg

- Subject changed from support text-mode open() in Python 3 to [SDKs] support text-mode open() in Python 3

- Assigned To deleted (Tom Clegg)

- Parent task deleted (#11308)

#3 - 05/10/2017 07:20 PM - Tom Clegg

- Tracker changed from Task to Feature

- Target version deleted (2017-05-10 sprint)

#4 - 08/28/2017 09:48 PM - Tom Morris

- Target version set to Arvados Future Sprints

#5 - 07/31/2018 03:11 PM - Tom Morris

- Related to Story #11308: Support Python3 for arvados-python-client & command line utilities added

#6 - 07/31/2018 03:18 PM - Tom Morris

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

#7 - 07/31/2018 03:41 PM - Tom Morris

- Release set to 14

#8 - 08/01/2018 03:56 PM - Tom Morris

- Tracker changed from Feature to Story

#9 - 11/13/2018 09:48 PM - Tom Morris

- Release deleted (14)

#10 - 11/27/2018 12:58 AM - Tom Morris

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

#11 - 11/28/2018 04:40 PM - Tom Morris

- Target version changed from To Be Groomed to 2018-12-12 Sprint

#12 - 11/28/2018 05:05 PM - Tom Clegg

- Assigned To set to Tom Clegg

#13 - 11/29/2018 10:00 PM - Tom Clegg

11419-text-mode @ [0565661b2c4fb220428ea37c2e0728e4b07d2465https://ci.curoverse.com/view/Developer/job/developer-run-tests/989/](https://ci.curoverse.com/view/Developer/job/developer-run-tests/989/)

This change mostly comes down to wrapping the ArvadosFileReader/Writer in an io.BufferedReader/Random and then in an io.TextIOWrapper, if text mode is requested.

With the previous code, the "block boundary isn't on a character boundary" test crashed in Python 2 when using text mode, and the same wrapper fixes it, so I figured it would make sense to fix both at once. However, this does cause a slight API change that could conceivably break existing code. In text mode, the file-like object returned by Collection.open()...

1. ...is not an ArvadosFileReader/Writer. Code that tests isinstance(f, ...) will behave differently. I found no examples of this in our own repo.
2. ...cannot seek to arbitrary positions (see [TextIOBase seek restrictions](#)). The only occurrence of this I found in our repo was in the "open in text mode in Python 2" test.
3. ...does not accept close(flush=False). This revealed (via sdk/cwl test failures) that [source:sdk/python/arvados/commands/run.py](#) was accidentally using text mode to copy arbitrary files.

#14 - 11/30/2018 05:30 PM - Tom Clegg

This broke some things: <https://ci.curoverse.com/job/developer-run-tests-sdk-python-ruby/275/console>

#15 - 11/30/2018 05:30 PM - Tom Clegg

- Status changed from New to In Progress

#16 - 11/30/2018 08:35 PM - Tom Clegg

11419-text-mode @ [127e57b8ea812b9d03b67c0118230a475b727158https://ci.curoverse.com/view/Developer/job/developer-run-tests/992/](https://ci.curoverse.com/view/Developer/job/developer-run-tests/992/)

#17 - 12/03/2018 03:13 PM - Lucas Di Pentima

This LGTM, please merge.
Thanks!

#18 - 12/10/2018 10:16 PM - Tom Clegg

- Status changed from In Progress to Resolved
- % Done changed from 0 to 100

Applied in changeset [arvados|7317196f547cf8c5e0ec87d1526322c305faba30](#).

#19 - 12/14/2018 08:38 PM - Peter Amstutz

This has a regression:

<https://ci.curoverse.com/view/CWL/job/run-cwl-test-4xphq/799/console>

```
2018-12-14 20:28:39 arvados.cwl-runner ERROR: [step-valuefrom5-wf.cwl] While getting final output object: 'TextIOWrapper' object has no attribute 'size'
Traceback (most recent call last):
  File "/usr/lib/python2.7/dist-packages/arvados_cwl/runner.py", line 465, in done
    if f.size() > 0:
AttributeError: 'TextIOWrapper' object has no attribute 'size'
```

#20 - 12/14/2018 08:38 PM - Peter Amstutz

- Status changed from Resolved to Feedback
- Target version changed from 2018-12-12 Sprint to 2018-12-21 Sprint

#21 - 12/14/2018 09:00 PM - Tom Clegg

11419-cwl-read-json-binary-mode @ [fef817dc5d6d1e0160c6c14ba4f41581e241f33a](https://doi.org/10.1093/bioinformatics/bty433)

#22 - 12/18/2018 02:55 PM - Tom Clegg

- Status changed from Feedback to Resolved

#23 - 03/01/2019 06:30 PM - Tom Morris

- Release set to 15