

Arvados - Story #14939

Update crunchstat-summary to Python 3

03/08/2019 10:08 PM - Tom Morris

Status: Resolved	Start date: 03/08/2019
Priority: Normal	Due date:
Assigned To: Tom Morris	% Done: 100%
Category:	Estimated time: 0.00 hour
Target version: 2019-04-24 Sprint	
Description	
Subtasks:	
Task # 14940: Review 14939-crunchstat-summary-python3 -> 14939-crunchstat-summary-python3	Resolved
Task # 15127: Review 14939-crunchstat-summary-python3-fixup	Closed
Related issues:	
Blocks Arvados Epics - Story #14532: [Epic] Port to Python 3 to for Python 2 ...	Resolved 01/01/2020 09/16/2020

Associated revisions

Revision 5d77a55a - 03/08/2019 11:40 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision cd33e7da - 03/13/2019 11:00 PM - Tom Morris

Make iterator Python 2 compatible. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 22ed89be - 03/13/2019 11:02 PM - Tom Morris

Make log file reading Python 3 compatible. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 031a3005 - 03/13/2019 11:03 PM - Tom Morris

Fix test file handling to support Python 3. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 03fd14e6 - 03/13/2019 11:33 PM - Tom Morris

Add crunchstat-summary to Python 3 tests. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 6bac55d2 - 03/14/2019 10:53 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision c1589536 - 03/14/2019 10:53 PM - Tom Morris

Add crunchstat-summary to Python 3 tests. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 8a0bc9d2 - 03/14/2019 11:08 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision fdaf820a - 03/14/2019 11:08 PM - Tom Morris

Add crunchstat-summary to Python 3 tests. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 7c3413ae - 03/18/2019 09:01 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 77c137e8 - 03/18/2019 09:01 PM - Tom Morris

Add crunchstat-summary to Python 3 tests. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 44aa8df7 - 03/18/2019 09:06 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 0b2a6fed - 03/21/2019 07:35 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision d0dc4cc1 - 03/21/2019 07:42 PM - Tom Morris

Add crunchstat-summary to Python 3 tests. refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 00592ebe - 03/21/2019 07:47 PM - Tom Morris

Merge branch '14939-crunchstat-summary-python3'

Fixes #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 5950f909 - 03/27/2019 03:15 PM - Tom Clegg

Accept .py3 modifier for any Python suite in interactive mode.

(even ones that aren't normally run with python3 by jenkins)

refs #14939

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

Revision dbf8c571 - 03/27/2019 05:19 PM - Tom Morris

Add crunchstat-summary Python 3 to all tests.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision ae232084 - 04/19/2019 09:13 PM - Tom Morris

Add crunchstat-summary Python 3 to all tests.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 6f3c2bed - 04/19/2019 09:15 PM - Tom Morris

Update crunchstat-summary to Python 3. Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 8c345b8d - 04/19/2019 09:16 PM - Tom Morris

Add crunchstat-summary Python 3 to all tests.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision cf8a4757 - 04/22/2019 07:35 PM - Tom Morris

Add crunchstat-summary Python 3 to all tests.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 5254ba76 - 04/22/2019 07:35 PM - Tom Morris

Fix iteration order dependency.

Causes tests to fail in Python 3.5, but not Python 3.7 due to dependency on iteration order of dicts.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 2266cd0b - 04/24/2019 04:17 PM - Tom Morris

Add crunchstat-summary Python 3 to all tests.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision 6c1b0fb3 - 04/24/2019 04:17 PM - Tom Morris

Fix iteration order dependency.

Causes tests to fail in Python 3.5, but not Python 3.7 due to dependency on iteration order of dicts.

Refs #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

Revision ed35acd2 - 04/24/2019 04:19 PM - Tom Morris

Merge branch '14939-crunchstat-summary-python3-fixup'

Fixes #14939

Arvados-DCO-1.1-Signed-off-by: Tom Morris <tfmorris@veritasgenetics.com>

History

#1 - 03/08/2019 11:35 PM - Tom Morris

- Status changed from New to In Progress

- Start date set to 03/08/2019

#2 - 03/08/2019 11:35 PM - Tom Morris

- Parent task deleted (#14532)

#3 - 03/08/2019 11:35 PM - Tom Morris

- Tracker changed from Task to Story

#4 - 03/08/2019 11:36 PM - Tom Morris

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

#5 - 03/11/2019 04:17 PM - Tom Clegg

LiveLogReader threads are crashing in the test suite, which seems to indicate two broken things:

- Crashing threads don't cause the tests to fail, because it's Python and we don't explicitly catch exceptions in the thread and check/propagate

- them after calling join()... I guess we need to do that.
- LiveLogReader.next() is still needed by Python 2

Should update run-tests.sh to test crunchstat-summary with Python 3 as well (add "tools/crunchstat-summary:py3" to pythonstuff like "sdk/cwl:py3"). It looks like there are some bytes-vs.-str things to sort out.

#6 - 03/12/2019 08:17 PM - Tom Morris

- Target version changed from Arvados Future Sprints to 2019-03-13 Sprint

#7 - 03/13/2019 03:18 PM - Tom Morris

- Target version changed from 2019-03-13 Sprint to 2019-03-27 Sprint

#8 - 03/13/2019 11:48 PM - Tom Morris

Tom Clegg wrote:

LiveLogReader threads are crashing in the test suite, which seems to indicate two broken things:

- Crashing threads don't cause the tests to fail, because it's Python and we don't explicitly catch exceptions in the thread and check/propagate them after calling join()... I guess we need to do that.
- LiveLogReader.next() is still needed by Python 2

Fixed the iterator and other thread crashers, but not the general thread exception catching.

Should update run-tests.sh to test crunchstat-summary with Python 3 as well (add "tools/crunchstat-summary:py3" to pythonstuff like "sdk/cwl:py3"). It looks like there are some bytes-vs.-str things to sort out.

Done.

Also fixed Python 3 log file reading, test failure caused by integer division diffs, and a couple of other things.

All tests pass locally in both Python 2 & 3 and CI job running here: <https://ci.curoverse.com/view/Developer/job/developer-run-tests/1122/>

#9 - 03/14/2019 08:38 PM - Tom Clegg

Could tests/test_examples.py use crunchstat_summary.command.GzipText instead of having its own copy of it?

Nit: usually kwargs are foo=bar but there are a couple of foo = bar here. meh

The rest LGTM, thanks.

#10 - 03/15/2019 02:06 AM - Tom Morris

I force pushed a new branch 14939-crunchstat-summary-python3 at 6cf111041f450c5d33665615ac915869c4279f9f which addresses your two comments as well as fixes:

- resource leak warned by Python 3 (only)
- deprecation warnings in Python 3

I also eliminated some unnecessary u" additions and squashed some of the minor commits together and gave the UTF8 decoding wrapper class a better name.

If you could take a quick look to make sure it looks better, not worse, that'd be great.

#11 - 03/15/2019 05:23 AM - Tom Morris

I've put up another branch based on this one with a number of small fixes for [#12026](#), [#10570](#), & [#14451](#) which all have Eric as the reviewer. Please coordinate with him to see if it makes more sense for one person to review all the changes as a batch.

#12 - 03/18/2019 06:55 PM - Eric Biagiotti

Calling .items() to iterate on a dict is inefficient in Python 2 according to the following: http://python-future.org/compatible_idioms.html#iterating-through-dict-keys-values-items. It recommends using viewitems(), though this will require a future.utils import. When working on the ACR, I ran into packaging problems after adding the future import. If you add it, its probably worth running the package building and tests locally.

Also a nit about an extra line of whitespace here:

https://dev.arvados.org/projects/arvados/repository/revisions/14939-crunchstat-summary-python3/entry/tools/crunchstat-summary/tests/test_example_s.py#L63

Looks like this change from cpu to cp was unintentional:

https://dev.arvados.org/projects/arvados/repository/revisions/14939-crunchstat-summary-python3/entry/tools/crunchstat-summary/crunchstat_summary/webchart.py#L48

Also, do we use crunchstat-summary output to run cwl integration/conformance tests at all? If so, its worth running those as well.

#13 - 03/18/2019 08:56 PM - Tom Morris

Eric Biagiotti wrote:

Calling `.items()` to iterate on a dict is inefficient in Python 2 according to the following: http://python-future.org/compatible_idioms.html#iterating-through-dict-keys-values-items. It recommends using `viewitems()`, though this will require a `future.utils` import. When working on the ACR, I ran into packaging problems after adding the `future` import. If you add it, its probably worth running the package building and tests locally.

I think I'm willing to live with a few months of Python 2 only (unquantified) inefficiency in a non-performance critical section of code.

Also a nit about an extra line of whitespace here: https://dev.arvados.org/projects/arvados/repository/revisions/14939-crunchstat-summary-python3/entry/tools/crunchstat-summary/tests/test_examples.py#L63

Fixed.

Looks like this change from `cpu` to `cp` was unintentional: https://dev.arvados.org/projects/arvados/repository/revisions/14939-crunchstat-summary-python3/entry/tools/crunchstat-summary/crunchstat_summary/webchart.py#L48

Oops! Good catch. Fixed.

Also, do we use crunchstat-summary output to run cwl integration/conformance tests at all? If so, its worth running those as well.

No. I'm pretty sure there are no interactions with other components.

#14 - 03/21/2019 07:52 PM - Tom Morris

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

Applied in changeset [arvados|00592e8e08406f1b7649f758d4ed25a2882672ef](https://dev.arvados.org/projects/arvados/repository/revisions/14939-crunchstat-summary-python3/entry/tools/crunchstat-summary/crunchstat_summary/webchart.py#L48).

#15 - 03/27/2019 05:15 PM - Tom Morris

- Status changed from *Resolved* to *In Progress*
- Target version changed from *2019-03-27 Sprint* to *2019-04-10 Sprint*

#16 - 04/10/2019 03:02 PM - Tom Morris

- Target version changed from *2019-04-10 Sprint* to *2019-04-24 Sprint*

#17 - 04/22/2019 08:18 PM - Tom Morris

- Story points set to 1.0

Tom noticed that I failed to include this in the Python 3 tests (each component needs to be added individually). Adding it caused it to fail in CI because there was a latent bug in the code that was dependent on dictionary iteration order. It worked fine on Python 2.7 & 3.7, but failed on Python 3.5.

CI build at <https://ci.curoverse.com/view/Developer/job/developer-run-tests/1201/> (successful except for failing fuse tests)

#18 - 04/23/2019 01:14 PM - Eric Biagiotti

I'm trying to test locally (run `crunchstat-summary` with the test data) in a 3.5.7 virtualenv, and the `arvados-python-client` fails installation because of the `ciso8601` dependency. Is there a specific way I should be testing this?

#19 - 04/23/2019 04:05 PM - Tom Morris

I don't think I did anything special to set up my conda environment, but it doesn't have any specific `arvados-api-client` dependency, so you can use whatever relatively recent version that works for you.

Here's what I'm testing with:

python	3.5.5	h5001a0f_2	conda-forge
arvados-python-client	1.3.1.20190308210512		<pip>
ciso8601	2.1.1		<pip>

I can't tell what versions are in use on Jenkins, but it runs fine there too, so perhaps replicating that environment is the best way to go.

#20 - 04/23/2019 06:15 PM - Eric Biagiotti

If our code is relying on a dict being ordered, we should probably just use an OrderedDict.

From [cf8a475732f7a62108704b6d7803400681a69eba](#), changing line 187 in summarizer.py to the following got tests passing in 2.7, 3.5, and 3.7

```
stats = collections.OrderedDict()
```

Since python 3.7, dicts are officially ordered, but it is still recommended to use OrderedDict if we are supporting multiple python versions.

#21 - 04/23/2019 07:45 PM - Eric Biagiotti

As discussed in chat, there is no real need to require the dict to be ordered in this case, and the changes in [5254ba765ef759bb59584d7304ec22b77238772c](#) remove the order requirement. Subsequently, this LGTM. Thanks!

#22 - 04/24/2019 04:21 PM - Tom Morris

- Status changed from *In Progress* to *Resolved*

- % Done changed from 50 to 100

Applied in changeset [arvadosled35acd2e173b4fae4499d6ec21211cd2076e91c](#).

#23 - 05/21/2019 10:27 PM - Tom Morris

- Release set to 15