

## Arvados - Story #13429

### [API] [arvados-cwl-runner] Save workflow outputs to desired storage classes

05/02/2018 02:21 PM - Tom Morris

<b>Status:</b> Closed	<b>Start date:</b> 06/04/2018
<b>Priority:</b> Normal	<b>Due date:</b>
<b>Assigned To:</b> Fuad Muhic	<b>% Done:</b> 100%
<b>Category:</b>	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b> 2018-06-20 Sprint	
<b>Description</b> arvados-cwl-runner should accept a command line option that determines the storage classes used to save the workflow output.  The minimal implementation is to update the storage_classes_desired field of the output collection. It's OK that this wastes storage resources by first writing the output data to the default class, then copying it to the desired class, and then (after blobSignatureTTL) deleting it from default.	
<b>Subtasks:</b> Task # 13524: Review 13429-cwl-runner-storage-classes-support <b>Resolved</b>	
<b>Related issues:</b> Blocks Arvados - Feature #11184: [Keep] Support multiple storage classes <b>In Progress</b>	

#### Associated revisions

##### Revision 3fbabc1b - 06/18/2018 02:47 PM - Fuad Muhic

Merge branch '13429-cwl-runner-storage-classes-support'

closes #13429

Arvados-DCO-1.1-Signed-off-by: Fuad Muhic <[fmuhic@capeannenterprises.com](mailto:fmuhic@capeannenterprises.com)>

#### History

##### #1 - 05/02/2018 02:24 PM - Tom Morris

- Related to Feature #11184: [Keep] Support multiple storage classes added

##### #2 - 05/02/2018 02:26 PM - Tom Morris

- Target version set to To Be Groomed

##### #3 - 05/02/2018 04:18 PM - Tom Clegg

- Subject changed from arvados-cwl-runner support storage tiers to [API] [arvados-cwl-runner] Save workflow outputs to desired storage classes  
- Description updated

##### #4 - 05/02/2018 04:19 PM - Tom Clegg

- Related to deleted (Feature #11184: [Keep] Support multiple storage classes)

##### #5 - 05/02/2018 04:19 PM - Tom Clegg

- Blocks Feature #11184: [Keep] Support multiple storage classes added

##### #6 - 05/02/2018 05:08 PM - Tom Clegg

- Description updated

##### #7 - 05/02/2018 05:10 PM - Tom Morris

- Target version changed from To Be Groomed to Arvados Future Sprints  
- Story points set to 1.0

##### #8 - 05/23/2018 02:57 PM - Tom Morris

- Target version changed from Arvados Future Sprints to 2018-06-06 Sprint

**#9 - 05/23/2018 03:08 PM - Peter Amstutz**

- Assigned To set to Fuad Muhic

**#10 - 06/04/2018 05:54 PM - Lucas Di Pentima**

Reviewing [95c01d22571043d7b6405f928538204cf930d453](https://github.com/curoverse/arvados/pull/95c01d22571043d7b6405f928538204cf930d453)

- File sdk/cwl/arvados\_cwl/\_\_init\_\_.py
  - Line 604: I think the default storage classes value could be set on the parser.add\_argument() call
  - Line 606: On 13430 one of the requirements was to accept just one storage class for the moment, it that also desirable on this story?
  - Line 727: Typo on help message: 'workflow'
- I think some tests are needed for cases when the required storage classes is not the default value, and to check that it errors out when passed multiple storage classes.

**#11 - 06/06/2018 12:27 PM - Fuad Muhic**

- Status changed from New to In Progress

**#12 - 06/06/2018 01:10 PM - Fuad Muhic**

I added check and tests for multiple storage classes. I think it would be nice to have a test that checks if output collection contains storage\_classes\_desired field set to proper value defined by --storage-classes argument. However I'm really not sure how to write that test, I tried but I guess I'm not experienced enough with cwl-runner.

**#13 - 06/06/2018 02:48 PM - Fuad Muhic**

- Target version changed from 2018-06-06 Sprint to 2018-06-20 Sprint

**#14 - 06/06/2018 05:36 PM - Lucas Di Pentima**

Reviewing [b01b480adb45c3bbcfab13e343e08c16854dedc](https://github.com/curoverse/arvados/pull/b01b480adb45c3bbcfab13e343e08c16854dedc)

- File sdk/cwl/tests/test\_submit.py \* Line 338: capture\_stdout is not being used after the main() call so I suppose it can avoided or read from it to assure that the correct error message is being printed?
- File sdk/cwl/arvados\_cwl/\_\_init\_\_.py \* Line 789: On error I think the way to go is to log it and return 1 (as in lines 804 & 824)
- I think a test that assures an output collection has correct correct storage\_classes\_desired field is not necessary as the Python SDK Collection's test already proves that. OTOH I think having a test that confirms that default & non-default --storage-classes values are passed correctly to make\_output\_collection() will be very useful. I think this can be done using a similar approach as your test\_error\_when\_multiple\_storage\_classes\_specified test and mocking make\_output\_collection to check how it was called.
- I realized that the storage\_classes argument is being used only when --submit is not added. When it is used, a-c-r submits a container that runs another a-c-r with all necessary params (see test at test\_submit.py line 591, for example), so we'll need a similar test that assures that the correct@ --storage-classes@ param is being passed along. I'm not sure where this needs to be added, please talk to Peter if you don't find it.
- Also, remember to merge master into this branch so Jenkins test don't fail because of ciso8601

**#15 - 06/18/2018 10:02 AM - Fuad Muhic**

tests: <https://ci.curoverse.com/job/developer-run-tests/759/> (with merged master)

- Added test which check that storage\_classes parameters is propagated correctly to make\_output\_collection
- In --submit mode, storage\_classes are passes to submitted container. (I opted to pass --storage-classes parameter only if user specified if in the first place. If user don't specify --storage-classes, submitted container will not have --storage-classes=default. Everything still works correctly and we don't have to add 100+ lines to test\_submit.py). I added tests for this.

**#16 - 06/18/2018 01:41 PM - Lucas Di Pentima**

This LGTM. Thanks!

**#17 - 06/18/2018 02:50 PM - Fuad Muhic**

- Status changed from In Progress to Resolved

- % Done changed from 0 to 100

Applied in changeset [arvados%3fbabc1b5236b6667f453a2949849d96f6e683df](https://github.com/curoverse/arvados/pull/arvados%3fbabc1b5236b6667f453a2949849d96f6e683df).

**#18 - 06/18/2018 02:52 PM - Fuad Muhic**

- Status changed from Resolved to Closed

**#19 - 07/23/2018 06:41 PM - Tom Morris**

- Release set to 13