

Arvados - Bug #3825

[Crunch] crunch-job should save (very) large log files without filling up tmp space

09/05/2014 02:39 PM - Ward Vandewege

Status:	Resolved	Start date:	10/02/2014
Priority:	Normal	Due date:	
Assigned To:	Tim Pierce	% Done:	100%
Category:	Crunch	Estimated time:	0.00 hour
Target version:	2014-10-08 sprint		
Description			
Instead of writing to a temporary file and then calling arv-put, pipe the log directly to an arv-put process.			
Subtasks:			
Task # 3913: Test crunch-job log output			Resolved
Task # 3912: Rewrite crunch_job::Log to pipe output to arv-put			Resolved
Task # 4078: Review 3825-crunch-pipe-to-arv-put-final			Resolved

Associated revisions

Revision d6596397 - 10/06/2014 06:22 PM - Tim Pierce

Merge branch '3825-crunch-pipe-to-arv-put-final'

Closes #3825.

History

#1 - 09/12/2014 02:56 PM - Ward Vandewege

- Description updated

#2 - 09/15/2014 11:03 AM - Tom Clegg

- Subject changed from [Crunch] fix crunch-dispatch so it does not fill up tmp space when there a job has a (very) large log file to [Crunch] crunch-job should save (very) large log files without filling up tmp space

- Description updated

- Category set to Crunch

#3 - 09/17/2014 03:13 PM - Tom Clegg

- Target version changed from Arvados Future Sprints to 2014-10-08 sprint

#4 - 09/17/2014 03:46 PM - Tim Pierce

- Assigned To set to Tim Pierce

#5 - 10/02/2014 05:36 PM - Tim Pierce

- Status changed from New to In Progress

#6 - 10/03/2014 08:47 PM - Tom Clegg

Reviewing 3825-crunch-pipe-to-arv-put at [c6a6231](#)

- I think it would read better if `output_log_*` were renamed to `log_writer_*`
 - The word "output" usually means "output of a task or job" around here, and that's already a bit overloaded and confusing.
 - Perhaps `write_output_log` should rename to something slightly different than `write_log_writer`, though. Maybe `write_to_log_writer` or `log_writer_send`?
 - I do like the idea of using the whole substring `log_writer` (née `output_log`) in all of the related functions. It makes a really clear distinction between "this log pipe" and "some mysterious perl feature" and "some other logging-related thing we do here" etc.

I think a semicolon, although not strictly necessary, would be better form here:

- ```
+ return $log_pipe_pid
+}
```

Other than that, LGTM!

**#7 - 10/06/2014 06:25 PM - Tim Pierce**

- *Status changed from In Progress to Resolved*

- *% Done changed from 67 to 100*

Applied in changeset arvados|commit:d65963976174bd9c94fdb0b91eeb8a281f01e7b3.

**#8 - 10/08/2014 07:12 PM - Ward Vandewege**

- *Target version changed from 2014-10-08 sprint to 2014-10-29 sprint*

**#9 - 10/08/2014 07:19 PM - Ward Vandewege**

- *Target version changed from 2014-10-29 sprint to 2014-10-08 sprint*