

art - Bug #13663

art 2.1 is dangerously creative with output file names

08/24/2016 04:06 PM - Andrei Gaponenko

Status:	Closed	Start date:	08/24/2016
Priority:	Normal	Due date:	
Assignee:	Kyle Knoepfel	% Done:	100%
Category:	I/O	Estimated time:	4.00 hours
Target version:	2.04.00	Spent time:	6.00 hours
Occurs In:		Experiment:	Mu2e
Scope:	Internal	SSI Package:	art

Description

Hello,

From a discussion on the art-users list I learned that in art v2.1 output files may or may not be written with the names specified by the user. One of the comments was

This is another manifestation of the frequent conflict between what's the right thing when you're dealing with valuable data, or in a batch-processing environment, and what's convenient in a rapid development/debugging cycle.

I think the current behavior is worse than that. In a "development cycle" the old behavior of overwriting existing files was convenient. If you don't want to overwrite "valuable data" on the other hand, the correct behavior is to stop the program with an error message. If my fcl configuration directs art to write out a file named "a.root", my script will be reading "a.root" in the next step. If art ignores my configuration and "creatively" comes with a different output file name, I (or my script) will likely be looking at a stale file and perhaps getting wrong physics conclusion.

I am filing this as a bug, because the newly implemented behavior breaks the basic agreement that art software does what it is told to do. It should either write user-specified files, or give an error if it can't. It would be fine to have an option to switch between "do-not-overwrite" and "ok-to-overwrite" modes, but making up new output file names is just wrong.

Andrei

Associated revisions

Revision f87df13f - 08/31/2016 08:19 AM - Kyle Knoepfel

Resolve issue #13663: restore original file-renaming behavior.

- Revert PostCloseFileRenamer behavior to enable file overwriting.
- Enforce '%#' file-name placeholder for output modules that specify fileProperties.
- Update fileMode deprecation behavior/messages.

History

#1 - 08/24/2016 04:23 PM - Thomas Junk

There is a noclobber option to mv, and the noclobber environment variable can be queried (mv clobbers by default, even with set noclobber). The unmoved new file would have an uninformative name, RootOutputNNNNNN.root, instead of one based on what the user wanted. But even then, there is a hazard that a user may ignore the error message and read a stale file.

#2 - 08/24/2016 04:26 PM - Kyle Knoepfel

Note that this issue is slated for discussion at tomorrow's art-stakeholders meeting.

#3 - 08/24/2016 04:32 PM - Kyle Knoepfel

Note also that this behavior change was documented in the [art 2.1 release notes](#) under "Behavior changes".

#4 - 08/29/2016 11:31 AM - Kyle Knoepfel

- Status changed from New to Assigned
- Assignee set to Kyle Knoepfel
- Estimated time set to 4.00 h
- SSI Package art added
- SSI Package deleted ()

This issue was discussed at last week's stakeholders meeting. The resolution proposed at that meeting was:

Restore the old behavior (i.e. previously-written files are allowed to be overwritten); however, for configurations that invoke the new output file-handling facilities, require the presence of the sequence-number placeholder in the filename. If the placeholder is not present, throw an exception notifying the user of the expected filename requirements.

This proposal will be sent to the stakeholders' list.

#5 - 08/29/2016 03:50 PM - Andrei Gaponenko

The proposed resolution looks fine. Thank you!

Andrei

#6 - 08/31/2016 08:29 AM - Kyle Knoepfel

- Category set to I/O
- Status changed from Assigned to Resolved
- % Done changed from 0 to 100

Resolved with commit [art:f87df13](#) according to proposal.

#7 - 09/14/2016 02:10 PM - Kyle Knoepfel

- Status changed from Resolved to Closed
- Target version set to 2.04.00