

art - Feature #4357

Feature # 4031 (Closed): For DS-50 DAQ/artdaq, support customizable output files

Implement on-close file configurable file renaming [ds50daq-related]

07/17/2013 10:13 AM - Kurt Biery

Status: Closed	Start date:
Priority: Normal	Due date:
Assignee: Christopher Green	% Done: 100%
Category: I/O	Estimated time: 0.00 hour
Target version:	Spent time: 21.25 hours
Scope: External	SSI Package:
Experiment: DarkSide	
Description	
At the moment, we have a cron job at LNGS to rename the data files that are written on disk from a somewhat generic file name to a name that has the run number, subrun number, and a timestamp in it. The request here is to have the file have a configurable name as soon as it is closed.	
We have confirmed with Alessandro that the initial filename is irrelevant - it is the name of the complete, closed file that matters.	
In some of our discussions, Marc and Chris suggested that the filename could be arbitrary initially, but once some characteristics of the data have been determine, it might be possible to change the name to something like <pattern-with-run-and-subrun-number>_incomplete.root. Then when it is close, it could be changed to <pattern>_complete.root (or something like that). [Apologies if I got this wrong.]	
Subtasks:	
Feature # 4625: Implement configurable temporary directory for in-progress output file.	Closed
Related issues:	
Related to art - Feature #3774: Improved ability to generate output file names.	Closed 04/26/2013

History

#1 - 07/17/2013 10:15 AM - Kurt Biery

- Parent task set to #4031

#2 - 07/18/2013 11:33 AM - Christopher Green

- Category set to I/O

- Status changed from New to Feedback

For proposed specifiers, please see [Output file renaming](#). Comments, additions, etc. appreciated.

#3 - 07/19/2013 07:04 AM - Christopher Green

- Status changed from Feedback to Assigned

On 7/18/13 3:53 PM, Kurt Biery wrote:

Hi Chris,

Hi Kurt,

This looks great to me. A couple of minor questions:

- For the "timestamp of lowest/highest run/subrun", would that be the beginning or end of the run/subrun?

The timestamp is currently (according to `artdaq/ArtModules/detail/RawEventQueueReader.cc`) the beginning, as set by the time the art process in the event builder sees the first `RawEvent`, /not/ anything to do with the DAQ. If you wish the event

timestamp to follow the DAQ, then this code (and possibly the RawEvent itself) would need to be enhanced.

- Would the Process name be the one that is specified in the FHiCL configuration? An alternative that I might imagine would be something from the Linux process table. If it is the former, then it might be helpful to make that more obvious to readers who aren't as familiar with the parts of the art configuration.

It would be the process_name as specified in the FHiCL configuration.

A possibly unrelated question would be "how would the multi-step renaming that Marc has mentioned a couple of times be handled?" From what I recall, this is something like <totally random filename> to <something_{format_specifiers}_inprogress.root> to <something_{format_specifiers}_complete.root>. Are the "inprogress" and "complete" substrings configurable? Or, has this whole idea gone by the wayside? I don't have a strong preference in any case, just curious.

I thought that since Alessandro claimed that the in-progress name of the file did not matter, then we would have a unique in-progress name (processing timestamp and unique salt, for example) since continually updating the format specifier values served no requested purpose.

Alessandro,
Do the format specifiers that Chris has listed on this page [Output file renaming](#) look good to you?

Comments to Alessandro's email under separate cover.

Thanks,
Chris.

#4 - 07/19/2013 07:05 AM - Christopher Green

Hi Alessandro,
On 7/18/13 4:08 PM, Alessandro Razeto wrote:

that is great!
does the name can include even a path?
like /data/complete/phase2-double-phase/Run-XXX-YYY.root?

Yup.

while unfinished file would be in /data/incoming/xyz.root?

Yup.

Thank you very much

You're welcome,
Chris.

Alessandro

#5 - 07/19/2013 07:08 AM - Christopher Green

- Scope set to External
- Experiment DarkSide added

[Output file renaming](#) updated per Kurt's request to clarify process name.

#6 - 07/22/2013 09:12 AM - Kurt Biery

On Friday, Chris explained that providing parameters that use the time of events is not as straightforward as he originally thought. Given that, and the fact that the times of the file open and file close are still available, I recommended that we defer the implementation of the event-based time parameters to some later time.

I just updated the Wiki page that Chris created with the list of parameters to show these parameters as "deferred".

If any of this was a wrong thing to do, please let me know, and I will switch things back.
Kurt

#7 - 07/23/2013 02:51 PM - Christopher Green

- *Status changed from Assigned to Resolved*
- *% Done changed from 0 to 100*
- *Estimated time set to 20.00 h*

Implemented with commits [6c1e2bb](#) and [d0a5478](#).

#8 - 03/03/2014 11:44 AM - Lynn Garren

- *Status changed from Resolved to Closed*