

## art - Support #13174

### Some issues have been found with the RootOutput automatic file closing changes

07/08/2016 01:57 PM - Kurt Biery

<b>Status:</b>	Closed	<b>Start date:</b>	07/18/2016
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Kyle Knoepfel	<b>% Done:</b>	100%
<b>Category:</b>	I/O	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	2.02.02	<b>Spent time:</b>	49.50 hours
<b>Scope:</b>	Internal	<b>SSI Package:</b>	art
<b>Experiment:</b>	-		

#### Description

I have run some artdaq tests of the new RootOutput automatic-file-closing features that are available in art v2.x.y, and I would like to share some observations and ideas.

But first, I should apologize for not being more explicit in some of our earlier discussions. In the online, it is very important to always close all open files at EndRun time, independent of whatever other file-closing criteria are in effect for a particular run. I've noticed some situations in which this doesn't happen with the new functionality, and I'll describe those below. (For reference, I should also point out that multiple online data-taking runs can take place using single instantiations of art.)

As an example of the behavior that is desired, if the file-closing condition is set to 100 events, and we have two runs of 325 events, we would expect 8 files to be created by artdaq. For each run: three files with 100 events and one file with 25 events.

My first observation is that the only fileSwitch.boundary that seems useful online is "Event". In all of the cases that I can think of at the moment, we want the requested condition (file size <= X kB, number of events <= 123) to be acted upon immediately, and it seems like "fileSwitch.boundary: Event" is the correct way to configure that. If I've misunderstood that, it would be great to have that clarified.

Additional observations are based on specific test cases:

### File size and event count conditions can store events from multiple runs in a single file.

#### Example 1:

Here is the RootOutput configuration:

```
normalOutput: {
  module_type: RootOutput
  fileName: "/tmp/artdaqdemo_r%06r_sr%02s_%#_%to.root"
  maxEventsPerFile : 200
  #maxSize : 500
  fileSwitch : {
    boundary : Event
    #force : true
  }
  compressionLevel: 0
  #tmpDir : "/home/biery"
}
```

Run 708 had 523 events, run 709 had 473 events.

- One issue is that the first events from run 709 went into one of the same files as events from run 708.
- Another issue is that the last file did not get renamed until the system was shutdown.

Here is the file list after run 709 was ended but before the system was shutdown.

```
[biery@mu2edeq01 tmp]$ ls -altF | head
total 49040
drwxrwxrwt. 151 root   root   159744 Jul  8 13:11 ./
```

```

drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:08 masterControl/
-rw-r--r--   1 biery   mu2e   127427 Jul  8 13:08 RootOutput-5416-0713-11a5-3290.root
-rw-r--r--   1 biery   mu2e   573430 Jul  8 13:08 artdaqdemo_r000709_sr01_4_20160708T180749.root
-rw-r--r--   1 biery   mu2e   574505 Jul  8 13:07 artdaqdemo_r000708_sr01_3_20160708T180701.root
-rw-r--r--   1 biery   mu2e   573430 Jul  8 13:07 artdaqdemo_r000708_sr01_2_20160708T180641.root
-rw-r--r--   1 biery   mu2e   573430 Jul  8 13:06 artdaqdemo_r000708_sr01_1_20160708T180622.root
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:06 aggregator/
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:06 eventbuilder/

```

## Example 2:

FHiCL configuration:

```

normalOutput: {
  module_type: RootOutput
  fileName: "/tmp/artdaqdemo_r%06r_sr%02s_%#_%to.root"
  #maxEventsPerFile : 200
  maxSize : 500
  fileSwitch : {
    boundary : Event
    #force : true
  }
  compressionLevel: 0
  #tmpDir : "/home/biery"
}

```

Similar behavior:

```

[biery@mu2edaq01 tmp]$ ls -altF | head
total 54036
-rw-r--r--   1 biery   mu2e   190415 Jul  8 13:46 RootOutput-ce0b-3fef-1779-ae2b.root
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:46 masterControl/
drwxrwxrwt. 152 root     root   159744 Jul  8 13:46 ./
-rw-r--r--   1 biery   mu2e   1112307 Jul  8 13:45 artdaqdemo_r000711_sr01_4_20160708T184443.root
-rw-r--r--   1 biery   mu2e   1113170 Jul  8 13:44 artdaqdemo_r000710_sr01_3_20160708T184330.root
-rw-r--r--   1 biery   mu2e   1112307 Jul  8 13:43 artdaqdemo_r000710_sr01_2_20160708T184231.root
-rw-r--r--   1 biery   mu2e   1112307 Jul  8 13:42 artdaqdemo_r000710_sr01_1_20160708T184133.root

```

## A Run-based boundary produces extra small empty files.

Here is the FHiCL:

```

normalOutput: {
  module_type: RootOutput
  fileName: "/tmp/artdaqdemo_r%06r_sr%02s_%#_%to.root"
  #maxEventsPerFile : 200
  #maxSize : 500
  fileSwitch : {
    boundary : Run
    force : true
  }
  compressionLevel: 0
  #tmpDir : "/home/biery"
}

```

I took two runs, numbers 712 and 713. Here are the files on disk:

```

[biery@mu2edaq01 tmp]$ ls -altF | head
total 56068
drwxrwxrwt. 153 root     root   159744 Jul  8 13:50 ./
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:50 masterControl/
-rw-r--r--   1 biery   mu2e   186172 Jul  8 13:49 artdaqdemo_r-_sr-_4_20160708T184953.root
-rw-r--r--   1 biery   mu2e   1006811 Jul  8 13:49 artdaqdemo_r000713_sr01_3_20160708T184901.root
-rw-r--r--   1 biery   mu2e   186172 Jul  8 13:48 artdaqdemo_r-_sr-_2_20160708T184846.root
-rw-r--r--   1 biery   mu2e   699186 Jul  8 13:48 artdaqdemo_r000712_sr01_1_20160708T184817.root
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:48 aggregator/
drwxrwxrwx   2 biery   mu2e   12288 Jul  8 13:48 eventbuilder/

```

**Subtasks:**

Bug # 13272: Spurious output files saved in artdaq context

**Closed**

Feature # 13273: Allow maxSubRuns and maxRuns for output-file handling

**Closed****History****#1 - 07/08/2016 02:30 PM - Kurt Biery**

I've also tested a SubRun-based closing model, and we get lots of spurious files there, also:

```
[biery@mu2edeq01 tmp]$ ls -altF | head -16
total 61028
drwxrwxrwt. 154 root root 159744 Jul 8 14:27 ./
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:22 artdaqdemo_r_sr_12_20160708T192253.root
-rw-r--r-- 1 biery mu2e 388344 Jul 8 14:22 artdaqdemo_r000715_sr06_11_20160708T192250.root
drwxrwxrwx 2 biery mu2e 12288 Jul 8 14:22 masterControl/
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:22 artdaqdemo_r_sr_10_20160708T192247.root
-rw-r--r-- 1 biery mu2e 499339 Jul 8 14:22 artdaqdemo_r000715_sr05_9_20160708T192235.root
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:22 artdaqdemo_r_sr_8_20160708T192233.root
-rw-r--r-- 1 biery mu2e 487974 Jul 8 14:22 artdaqdemo_r000715_sr04_7_20160708T192221.root
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:22 artdaqdemo_r_sr_6_20160708T192219.root
-rw-r--r-- 1 biery mu2e 492512 Jul 8 14:22 artdaqdemo_r000715_sr03_5_20160708T192208.root
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:22 artdaqdemo_r_sr_4_20160708T192207.root
-rw-r--r-- 1 biery mu2e 465688 Jul 8 14:22 artdaqdemo_r000715_sr02_3_20160708T192155.root
-rw-r--r-- 1 biery mu2e 186172 Jul 8 14:21 artdaqdemo_r_sr_2_20160708T192155.root
-rw-r--r-- 1 biery mu2e 438875 Jul 8 14:21 artdaqdemo_r000715_sr01_1_20160708T192143.root
drwxrwxrwx 2 biery mu2e 12288 Jul 8 14:21 eventbuilder/
```

**#2 - 07/08/2016 02:43 PM - Kyle Knoepfel**

- Description updated

**#3 - 07/11/2016 11:33 AM - Kyle Knoepfel**

- Status changed from New to Feedback

We would like to meet with you to discuss more specifically your needs.

**#4 - 07/11/2016 11:44 AM - Kurt Biery**

Sure! If something informal is OK, I can stop by WH9 sometime this afternoon or tomorrow afternoon. If something more formal is needed, please suggest a venue.

**#5 - 07/13/2016 02:05 PM - Kurt Biery**

I've copied four samples of the extra small "empty" data files to cluck:/scratch/biery/data.

**#6 - 07/13/2016 02:44 PM - Marc Paterno**

I've taken a quick look at one of the "empty" files. It seems to be a properly structured file (all the expected TTrees are there, as is an embedded SQLite database). However, the file contains 0 events, 0 subruns, and 0 runs. The branches for a variety of different data products are all there (most of the branches are set to carry Fragments). However, there are no entries on the branches.

**#7 - 07/14/2016 09:21 AM - Kyle Knoepfel**

Thank you for the files, Kurt. Indeed, the files look well-formed, and I concur that they are entirely empty of runs, subruns and events. I have attempted to create a spurious file within the art test suite using a configuration similar to yours -- I have been unable to replicate the error.

Can you supply us with instructions as to how to recreate the spurious output files from within artdaq? Thank you.

**#8 - 07/18/2016 06:29 AM - Kurt Biery**

Hi Kyle,

I tried installing the artdaq-demo (<https://cdcvns.fnal.gov/redmine/projects/artdaq-demo/wiki>) on woof and cluck, but ran into problems there. We'll look into those; in the meantime, I installed it on the mu2e DAQ Pilot cluster and added you to the k5login for the mu2edeq account.

To see the problem, open two shell windows and log into mu2edeq01.fnal.gov in both of them ("ssh mu2edeq01.fnal.gov -l mu2edeq"). In the first window:

- 'cd Issue13174'
- 'source ./setupARTDAQDEMO'
- 'start2x2x2System.sh'

In the second window:

- 'cd Issue13174'
- 'source ./setupARTDAQDEMO'
- 'manage2x2x2System.sh init'
  - this initializes (configures) the artdaq processes
  - when this command is run, copies of the FHiCL files that were sent to the artdaq processes are created in the current directory
- 'manage2x2x2System.sh -N <run number> start'
- [wait a bit]
- 'manage2x2x2System.sh stop'
- at this point, you can look in /tmp to see the data files that were created

To kill the "start2x2x2System.sh" script, either <ctrl-c> in the first shell, or send it "manage2x2x2System.sh exit" from the second shell.

**#9 - 07/18/2016 08:39 AM - Kyle Knoepfel**

Thank you, Kurt, for the instructions. For management reasons, I will break this issue into subcomponents.

**#10 - 07/22/2016 12:17 PM - Kyle Knoepfel**

Are the requested configuration enhancements necessary for output modules other than RootOutput? Currently, the file-switching capabilities are tied to RootOutput. It is not difficult *per se* to incorporate the enhancements for other output modules, but whether or not this is necessary could influence the design.

**#11 - 08/04/2016 08:47 AM - Kyle Knoepfel**

- *Category set to I/O*
- *Status changed from Feedback to Resolved*
- *Assignee set to Kyle Knoepfel*
- *SSI Package art added*
- *SSI Package deleted ()*

**#12 - 08/09/2016 08:23 AM - Kyle Knoepfel**

- *Status changed from Resolved to Closed*
- *Target version set to 2.02.02*