ANNIE - Task #14890

remove older files from /pnfs/annie/persistent/raw

12/26/2016 07:57 AM - Arthur Kreymer

Status:	Resolved	**Start date:**	12/26/2016
Priority:	High	**Due date:**	12/27/2016
Assignee:	Arthur Kreymer	**% Done:**	100%
Category:		**Estimated time:**	10.00 hours
Target version:		**Spent time:**	7.10 hours

Description

RITM0500227 12/26 reduce usage of /pnfs/annie/persistent

As reported by arossi Dec 18,
space is running low in /pnfs/annie/persistent, under 10% free.

Free space is down to 4.7% on Dec 26.
See current usage at http://fndca.fnal.gov/cgi-bin/du_cgi.py

For performance, we need to reduce usage to under 90%
or ask for addititonal space.

History

#1 - 12/26/2016 08:00 AM - Arthur Kreymer

- % Done changed from 10 to 30

- Estimated time changed from 2.00 h to 4.00 h

Date: Thu, 22 Dec 2016 20:28:24 +0000
From: Benjamin Richards <b.richards@qmul.ac.uk>
To: Mayly Sanchez <mayly.sanchez@gmail.com>
Cc: Arthur Kreymer <kreymer@fnal.gov>, annie_ops <annie_ops@fnal.gov>

Hi Art, et al.
I wasn't at the ops meeting so dont know the full details. but from what i gleamed from the slack general channel this is not the tap storage but the strage used locally for analysis.

If that is the case i would think files labelled V2 and below and the corresponding raw files are pretty safe to delete. Matt should be able to indicate if he still needs these?

the V3 and V4 files need reprocessing into V5 files. So its useful to keep those at the moment.

#2 - 12/26/2016 08:01 AM - Arthur Kreymer

- % Done changed from 30 to 40

Date: Fri, 23 Dec 2016 15:40:20 +0000
From: Arthur Kreymer <kreymer@fnal.gov>
To: Benjamin Richards <b.richards@qmul.ac.uk>
Cc: Mayly Sanchez <mayly.sanchez@gmail.com>, annie_ops <annie_ops@fnal.gov>
Subject: Re: Fw: Annie persistent space more than 90% full (fwd)

```
Thanks Ben !

My reading of run numbers in /pnfs/annie/persistent raw
  indicates that we can remove raw Runs up to 247 for persistent.

The raw files remain available in /pnfs/annie/raw/raw

I will set up a script early next week to verify checksums
  and do the removals.

I will start with the older files.
  Runs through 247 occupy 21 TBytes of space,
  and we only need to remove 5 or so to buy time.
```

for RUN in `ls /pnfs/annie/persistent/raw/raw` ; do
[${RUN} -le 247] && du -sm /pnfs/annie/persistent/raw/raw/${RUN}
done | cut -f 1 | /nashome/k/kreymer/minos/scripts/count

```
Enter numbers to be added :
 Got 135 /tmp/FOO numbers
21284647
```

#3 - 12/26/2016 08:03 AM - Arthur Kreymer

- % Done changed from 40 to 60

SLACK daq channel

Arthur Kreymer [8:26 AM] 12/25
Thanks Ben ! Almost all of run 384 is in v5data, lacking just S10p19.

[8:29]
I have created script ~annieraw/archiver/transient to remove selected runs from the persistent area Will run it on v2data Monday (Boxing Day)

#4 - 12/26/2016 08:15 AM - Arthur Kreymer

- % Done changed from 60 to 80

- Estimated time changed from 4.00 h to 10.00 h

Checked access times for files in data and v2data 12/25

```
PR=/pnfs/annie/persistent/raw

find $PR/data -type f -atime -205 | wc -l
7991

   data
days files
205   7991
140    198
135      5
100      3

   V2data
days files
158  11580
120   7188
 80     31
 25      4
 24      0
```

#5 - 12/26/2016 08:50 AM - Arthur Kreymer

- % Done changed from 80 to 90

```
annieraw@anniegpvm02.fnal.gov

PD=/pnfs/annie/persistent/raw/v2data

Removed run 58 and all but the remaining first 5 and last 5 runs,
Removed runs 101 through 238

cd archiver

RUNS=`ls $PD | head -n -5 | tail -n +5`

date
for RUN in ${RUNS} ; do ./transient v2data ${RUN} do ; done
date

Mon Dec 26 08:37:03 CST 2016
Mon Dec 26 08:42:02 CST 2016

Removed additional large runs

1.4T    /pnfs/annie/persistent/raw/v2data/0243
351G    /pnfs/annie/persistent/raw/v2data/0244
```

```
764G    /pnfs/annie/persistent/raw/v2data/0245

for RUN in 0243 0244 0245 ; do ./transient v2data ${RUN} do ; done

du -sh ${PD}
266G    /pnfs/annie/persistent/raw/v2data
```

#6 - 01/04/2017 08:01 PM - Arthur Kreymer

- % Done changed from 90 to 100

```
annieraw@anniegpvm02

PD=/pnfs/annie/persistent/raw/data

du -sm $PD/*
27276    /pnfs/annie/persistent/raw/data/0001
18307    /pnfs/annie/persistent/raw/data/0002
2856    /pnfs/annie/persistent/raw/data/0003
1039    /pnfs/annie/persistent/raw/data/0004
531    /pnfs/annie/persistent/raw/data/0005
855    /pnfs/annie/persistent/raw/data/0006
60651    /pnfs/annie/persistent/raw/data/0007
1    /pnfs/annie/persistent/raw/data/0008
...
24494    /pnfs/annie/persistent/raw/data/0157
14284    /pnfs/annie/persistent/raw/data/0158
13417    /pnfs/annie/persistent/raw/data/0159
38628    /pnfs/annie/persistent/raw/data/0160
38963    /pnfs/annie/persistent/raw/data/0161
35895    /pnfs/annie/persistent/raw/data/0162
36667    /pnfs/annie/persistent/raw/data/0163

Few files used recently

find $PD -type f -atime -150 -exec ls -lu {} \;
-rw-r--r-- 1 annieraw annie 479003019 Sep  5 08:09 /pnfs/annie/persistent/raw/data/0162/DataR162S0p20T3Jul_17_
0541.root
-rw-r--r-- 1 annieraw annie 783032760 Aug 16 11:15 /pnfs/annie/persistent/raw/data/0093/DataR93S0p0T1Jun_11_11
39.root
-rw-r--r-- 1 annieraw annie 782626723 Nov  1 10:37 /pnfs/annie/persistent/raw/data/0023/DataR23S0p0T0May_10_14
19.root
-rw-r--r-- 1 annieraw annie 761470824 Dec 24 08:31 /pnfs/annie/persistent/raw/data/0001/DataR1S0p0T0Apr_14_155
3.root
-rw-r--r-- 1 annieraw annie 785918201 Nov 18 16:41 /pnfs/annie/persistent/raw/data/0100/DataR100S0p0T1Jun_14_1
418.root

Removed all but runs 1-5 and 159-163

cd archiver

RUNS=`ls $PD | head -n -5 | tail -n +6`

printf "${RUNS}"
0006
0007
0008
...
0156
0157
0158

printf "${RUNS}\n" | wc -l
138

for RUN in ${RUNS} ; do ./transient data ${RUN} ; done | less
for RUN in ${RUNS} ; do ./transient data ${RUN} ; done | grep -v ^rm

http://fndca3a.fnal.gov/cgi-bin/du_cgi.py
Data generated on  2017-01-04 11:53:27
Used 44,291 GiB, Usage 89.74 %
/pnfs/fs/usr/annie/persistent/raw/data     4,822

for RUN in ${RUNS} ; do ./transient data ${RUN} do ; done | less
```

```
less transient.log

RUN transient data 0006
Version 20161225 Wed Jan  4 19:51:59 CST 2017
855     /pnfs/annie/persistent/raw/data/0006
removed `/pnfs/annie/persistent/raw/data/0006/DataR6S0p0T0Apr_16_1002.root'
removed `/pnfs/annie/persistent/raw/data/0006/DataR6S0p1T0Apr_16_1004.root'
DONE Wed Jan  4 19:51:59 CST 2017
...
removed `/pnfs/annie/persistent/raw/data/0158/DataR158S1p9T3Jul_14_1410.root'
removed `/pnfs/annie/persistent/raw/data/0158/DataR158S2p0T3Jul_14_1554.root'
DONE Wed Jan  4 19:56:47 CST 2017

Will verify improved persistent space tomorrow 01/05

01/05 09:38 AnnieAnalysisPools Used 39,781 Usage 80.60 %
```

#7 - 01/05/2017 09:41 AM - Arthur Kreymer

- Status changed from New to Resolved

Will also close RITM0500227