BATCH - Support #6086

Minos candidate file removal

04/30/2014 04:28 PM - Arthur Kreymer

Status:	Assigned	**Start date:**	05/13/2014
Priority:	Normal	**Due date:**	05/20/2014
Assignee:	Arthur Kreymer	**% Done:**	90%
Category:		**Estimated time:**	8.00 hours
Target version:		**Spent time:**	15.40 hours

Description

This tracks removal of Minos candidate files from tape and disk.

History

#1 - 04/30/2014 04:34 PM - Arthur Kreymer

Date: Wed, 30 Apr 2014 16:29:33 -0500
From: Arthur Kreymer <kreymer@fnal.gov>
To: minos-users@fnal.gov
Cc: minos_authors@fnal.gov
Subject: Minos candidate file removals May 13 2014

As announced and discussed since 2012,
Minos will remove from disk and tape all the Candidate files
for all reconstruction releases, both data and simulation.

These files are no longer used, and fill about half of our tapes,
roughly 430 TB in 1.6 million files.

A migration to new media is about to start,
and we do not want to move files which we will never use.

This does not affect Minos+, as we no longer write cand files.

The file removals will start May 13, 2014.
This work is tracked in https://cdcvs.fnal.gov/redmine/issues/6086

#2 - 04/30/2014 04:36 PM - Arthur Kreymer

From http://nusoft.fnal.gov/minos/dcache/CFL/CFLSUM

```
      KF   TB
rn   410  142
rf   740   66
mc   480  226
NET 1630  430
```

#3 - 05/15/2014 03:13 PM - Arthur Kreymer

```
Assessed cand files in DCache read pools

    mindata@minos27

mkdir /minos/app/mindata/log/candx

MDC=/minos/app/mindata/log/candx
MDCP=/minos/app/mindata/log/candx/mdcp

    Files in DCache pools

curl -o ${MDCP} http://fndca.fnal.gov/dcache/files/minos

wc -l ${MDCP}
343040 /minos/app/mindata/log/candx/mdcp

grep cand ${MDCP} | wc -l
472
```

```
grep cand ${MDCP} > ${MDC}/mdcp-cand

grep dogwood0 ${MDC}/mdcp-cand | wc -l
119

cat ${MDC}/mdcp-cand | cut -f 6,7 -d / | sort -u

mcout_data/cedar_phy_bhcurv
mcout_data/cedar_phy_bhhi
mcout_data/cedar_phy_linfix
mcout_data/dogwood5
mcout_data/elm3
reco_far/cedar
reco_far/dogwood0
reco_far/R1_18_2
reco_near/cedar
reco_near/dogwood0

for DIR in ${DIRS} ; do
printf "%30s " ${DIR}  ; grep ${DIR} ${MDC}/mdcp-cand | wc -l ; done

   mcout_data/cedar_phy_bhcurv 32
     mcout_data/cedar_phy_bhhi 8
   mcout_data/cedar_phy_linfix 16
           mcout_data/dogwood5 144
               mcout_data/elm3 107
                reco_far/cedar 25
             reco_far/dogwood0 116
              reco_far/R1_18_2 1
               reco_near/cedar 20
            reco_near/dogwood0 3
```

#4 - 05/15/2014 03:15 PM - Arthur Kreymer

```
Surveyed cand files in the CFL complete file listing

    Surveyed cand files in the CFL complete file listing

CFL=/nusoft/app/web/htdoc/minos/dcache/CFL/CFL
CFLCAND=${MDC}/cflcand
CFLCANDS=${MDC}/cflcands

grep cand ${CFL} | tr -s / > ${CFLCAND}

wc -l  ${CFLCAND}
1646801 /minos/app/mindata/log/candx/cflcand

   List with just paths

cat ${CFLCAND} | cut -f 9 -d ' ' > ${CFLCANDS}

   Files not in /cand_data/

grep -v /cand_data/ ${CFLCANDS} | grep -v BAD | wc -l
425

FPS=`grep -v /cand_data/ ${CFLCANDS} | grep -v BAD`
for FP in ${FPS} ; do
dirname ${FP} ; done | sort -u

/pnfs/minos/alignment/neardet/d59/NewAlignment
/pnfs/minos/mcout_data/cedar/far/daikon_00/L010185N
/pnfs/minos/mcout_data/cedar_phy_bhhi/near/daikon_04/L010185N
/pnfs/minos/mcout_data/cedar_phy_bhlo/near/daikon_04/L010185N
/pnfs/minos/mcout_data/R1.14/far
/pnfs/minos/mcout_data/R1.7
/pnfs/minos/reco_far/R1_17/bcnd_data/2005-05
/pnfs/minos/reco_far/R1_17/bntp_data/2005-05
/pnfs/minos/reco_far/R1_17/bnts_data/2005-05
/pnfs/minos/reco_far/R1_17/sntp_data/2005-05
/pnfs/minos/reco_far/R1_17/snts_data/2005-05
/pnfs/minos/sim_reco/far/atmos/cand
/pnfs/minos/sim_reco/far/cosmic/cand

 grep -v '\.cand\.' ${CFLCANDS}  | wc -l
```

```
177

 grep '\.bcnd\.' ${CFLCANDS}  | wc -l
24

    Proceeding with the mainline reco_near/far mcout near/far files in cand_data and name .cand.
```

#5 - 05/15/2014 03:17 PM - Arthur Kreymer

Opened service desk ticket to advise Enstore admins.

RITM0100001 05/15 FYI - removal of 430 TB, 1.6 Million files from /pnfs/minos

We have started the long promised removal of Minos 'candidate' files from /pnfs/minos.
The old files are no longer in use,
and new files of this sort are no longer being produced.

I am removing files at the rate of under 5 files per second,
to avoid overloading the Enstore servers.

Please contact minos-admin if there are any operational issues.

There will be four major removals:

reco_near
reco_far
mcout_data/*/near
mcout_data/*/far

#6 - 05/15/2014 03:22 PM - Arthur Kreymer

```
NEARCAND=${MDC}/nearcand
NEARCANDL=${MDC}/nearcand-list

   reco_near cand_data .cand.

grep /reco_near/ ${CFLCANDS} | grep /cand_data/ | grep '\.cand\.' > ${NEARCAND}

wc -l ${NEARCAND}
408710 /minos/app/mindata/log/candx/nearcand

   Directory summary

DIRS=`cat ${NEARCAND} | cut -f 5 -d / | sort -u`
for DIR in ${DIRS} ; do
printf "%-25s " ${DIR} ; grep /${DIR} ${NEARCAND} | wc -l
done > ${MDC}/nearcand-count

    Get listing of NEARCAND files

cat ${NEARCAND} | while read FP ; do
ls -l ${FP} ; usleep 200000
done 2>&1 | tee ${NEARCANDL}
date

Thu May 15 11:09:08 CDT 2014

    Remove the files

    root@if-admin-minos

MDC=/minos/app/mindata/log/candx
NEARCAND=${MDC}/nearcand
{
date
cat ${NEARCAND} | while read FP ; do
echo /bin/rm -f ${FP} ; /bin/rm -f ${FP} ;usleep 200000
done
date
}  2>&1 | tee      ${MDC}/nearcand-rm
chown mindata.e875 ${MDC}/nearcand-rm

Thu May 15 15:00:41 CDT 2014
/bin/rm -f /pnfs/minos/reco_near/S06-06-22-R1-22/cand_data/2005-11/N00009107_0002.cosmic.cand.S06-06-22-R1-22.
0.root
```

```
/bin/rm -f /pnfs/minos/reco_near/S06-06-22-R1-22/cand_data/2005-11/N00009104_0019.cosmic.cand.S06-06-22-R1-22.
0.root
...

/bin/rm -f /pnfs/minos/reco_near/dogwood7/cand_data/2011-11/N00021294_0009.spill.cand.dogwood7.0.root
Fri May 16 20:43:09 CDT 2014
```

#7 - 05/16/2014 09:34 PM - Arthur Kreymer

```
    Moving on to reco_far cand

    mindata@minos27

MDC=/minos/app/mindata/log/candx
CFLCANDS=${MDC}/cflcands

FARCAND=${MDC}/farcand
FARCANDL=${MDC}/farcand-list

grep /reco_far/ ${CFLCANDS} | grep /cand_data/ | grep '\.cand\.' > ${FARCAND}

wc -l ${FARCAND}
728429 /minos/app/mindata/log/candx/farcand

    Directory summary

DIRS=`cat ${FARCAND} | cut -f 5 -d / | sort -u`
for DIR in ${DIRS} ; do
printf "%-25s " ${DIR} ; grep /${DIR} ${FARCAND} | wc -l
done > ${MDC}/farcand-count

less ${MDC}/farcand-count

    Get listing of FARCAND files

date
cat ${FARCAND} | while read FP ; do
ls -l ${FP} ; usleep 200000
done 2>&1 | tee ${FARCANDL}
date

Fri May 16 21:30:03 CDT 2014
-rw-r--r-- 1 rubin e875 9135413 Jul  3  2006 /pnfs/minos/reco_far/S06-06-22-R1-22/cand_data/2005-11/F00033083_
0003.spill.cand.S06-06-22-R1-22.0.root
-rw-r--r-- 1 rubin e875 87455920 Jul  3  2006 /pnfs/minos/reco_far/S06-06-22-R1-22/cand_data/2005-11/F00033083
_0015.all.cand.S06-06-22-R1-22.0.root
...
-rw-r--r-- 1 minospro e875 11347783153 Jan 30  2012 /pnfs/minos/reco_far/dogwood7/cand_data/2011-03/F00047421_
0000.cosmic.cand.dogwood7.0.root
Sun May 18 21:49:32 CDT 2014
```

#8 - 05/19/2014 08:28 AM - Arthur Kreymer

```
    root@if-admin-minos

MDC=/minos/app/mindata/log/candx
FARCAND=${MDC}/farcand

{
date
cat ${FARCAND} | while read FP ; do
echo /bin/rm -f ${FP} ; /bin/rm -f ${FP} ;usleep 200000
done
date
}  2>&1 | tee      ${MDC}/farcand-rm
chown mindata.e875 ${MDC}/farcand-rm

Mon May 19 08:26:45 CDT 2014
/bin/rm -f /pnfs/minos/reco_far/S06-06-22-R1-22/cand_data/2005-11/F00033083_0003.spill.cand.S06-06-22-R1-22.0.
root
/bin/rm -f /pnfs/minos/reco_far/S06-06-22-R1-22/cand_data/2005-11/F00033083_0015.all.cand.S06-06-22-R1-22.0.ro
ot
...
/bin/rm -f /pnfs/minos/reco_far/dogwood7/cand_data/2011-03/F00047409_0009.cosmic.cand.dogwood7.0.root
/bin/rm -f /pnfs/minos/reco_far/dogwood7/cand_data/2011-03/F00047421_0000.cosmic.cand.dogwood7.0.root
```

```
Wed May 21 14:22:23 CDT 2014

   Oops, ran this as mindata, not root,
   ran second pass as root to pick up the 10K files not yet removed

grep 'cannot remove' ${MDC}/farcand-rm | cut -f 2 -d \` | cut -f 1 -d \' \
  > ${MDC}/farcand-rmfail

wc -l ${MDC}/farcand-rmfail
10039 /minos/app/mindata/log/candx/farcand-rmfail

Wed May 21 14:37:51 CDT 2014
/bin/rm -f /pnfs/minos/reco_far/cedar_phy_bhcurv/cand_data/2004-10/F00027712_0007.all.cand.cedar_phy_bhcurv.0.
root
/bin/rm -f /pnfs/minos/reco_far/cedar_phy_bhcurv/cand_data/2003-07/F00018033_0000.all.cand.cedar_phy_bhcurv.0.
root
...

/bin/rm -f /pnfs/minos/BAD/timedog3/reco_far/dogwood3/cand_data/2008-05/F00040933_0008.cosmic.cand.dogwood3.0.
root
/bin/rm -f /pnfs/minos/BAD/timedog3/reco_far/dogwood3/cand_data/2008-05/F00040933_0008.spill.cand.dogwood3.0.r
oot
Wed May 21 15:27:32 CDT 2014
```

#9 - 05/21/2014 04:32 PM - Arthur Kreymer

```
    MC NEAR CAND

    mindata@minos27

MDC=/minos/app/mindata/log/candx
CFLCANDS=${MDC}/cflcands

MCNEARCAND=${MDC}/mcnearcand
MCNEARCANDL=${MDC}/mcnearcand-list

grep /mcout_data/ ${CFLCANDS} | grep /near/ | grep /cand_data/ | grep '\.cand\.' > ${MCNEARCAND}

wc -l ${MCNEARCAND}
390085 /minos/app/mindata/log/candx/mcnearcand

    Directory summary

DIRS=`cat ${MCNEARCAND} | cut -f 5 -d / | sort -u`
for DIR in ${DIRS} ; do
printf "%-25s " ${DIR} ; grep /${DIR} ${MCNEARCAND} | wc -l
done > ${MDC}/mcnearcand-count

less ${MDC}/mcnearcand-count

    Get listing of MCNEARCAND files

date
cat ${MCNEARCAND} | while read FP ; do
ls -l ${FP} ; usleep 200000
done 2>&1 | tee ${MCNEARCANDL}
date

Mon May 19 17:11:25 CDT 2014
-rw-r--r-- 1 rubin e875 312645083 Jul  4  2006 /pnfs/minos/mcout_data/S06-06-22-R1-22/near/cand_data/n13011003
_0000_L010185.cand.S06-06-22-R1-22.root
-rw-r--r-- 1 rubin e875 319099234 Jul  4  2006 /pnfs/minos/mcout_data/S06-06-22-R1-22/near/cand_data/n13011009
_0000_L010185.cand.S06-06-22-R1-22.root
...
rw-r--r-- 1 rubin e875 314638488 Sep 23  2006 /pnfs/minos/mcout_data/cedar/near/carrot_06/L010185/cand_data/32
9/n13013293_0000_L010185.cand.cedar.root
-rw-r--r-- 1 rubin e875 614650647 Mar  1  2007 /pnfs/minos/mcout_data/cedar/near/daikon_00/L010185N/cand_data/
112/n13011120_0009_L010185N_D00.cand.cedar.root
Tue May 20 22:37:32 CDT 2014

    Removed the files

    root@if-admin-minos

MDC=/minos/app/mindata/log/candx
```

```
MCNEARCAND=${MDC}/mcnearcand

{
date
cat ${MCNEARCAND} | while read FP ; do
echo /bin/rm -f ${FP} ; /bin/rm -f ${FP} ;usleep 200000
done
date
}  2>&1 | tee      ${MDC}/mcnearcand-rm
chown mindata.e875 ${MDC}/mcnearcand-rm

Wed May 21 16:31:26 CDT 2014
/bin/rm -f /pnfs/minos/mcout_data/S06-06-22-R1-22/near/cand_data/n13011003_0000_L010185.cand.S06-06-22-R1-22.r
oot
/bin/rm -f /pnfs/minos/mcout_data/S06-06-22-R1-22/near/cand_data/n13011009_0000_L010185.cand.S06-06-22-R1-22.r
oot
...
/bin/rm -f /pnfs/minos/mcout_data/cedar/near/carrot_06/L010185/cand_data/329/n13013293_0000_L010185.cand.cedar
.root
/bin/rm -f /pnfs/minos/mcout_data/cedar/near/daikon_00/L010185N/cand_data/112/n13011120_0009_L010185N_D00.cand
.cedar.root
Fri May 23 00:28:21 CDT 2014

wc -l ${MDC}/mcnearcand-rm
390087 /minos/app/mindata/log/candx/mcnearcand-rm

grep -v /bin/rm ${MDC}/mcnearcand-rm
Wed May 21 16:31:26 CDT 2014
Fri May 23 00:28:21 CDT 2014

    Count are correct, this purge is done
```

#10 - 05/23/2014 07:16 AM - Arthur Kreymer

```
    MC FAR CAND

    mindata@minos27

MDC=/minos/app/mindata/log/candx
CFLCANDS=${MDC}/cflcands

MCFARCAND=${MDC}/mcfarcand
MCFARCANDL=${MDC}/mcfarcand-list

grep /mcout_data/ ${CFLCANDS} | grep /far/ | grep /cand_data/ | grep '\.cand\.' > ${MCFARCAND}

wc -l ${MCFARCAND}
79806 /minos/app/mindata/log/candx/mcfarcand

    Directory summary

DIRS=`cat ${MCFARCAND} | cut -f 5 -d / | sort -u`
for DIR in ${DIRS} ; do
printf "%-25s " ${DIR} ; grep /${DIR} ${MCFARCAND} | wc -l
done > ${MDC}/mcfarcand-count

less ${MDC}/mcfarcand-count

    Get listing of MCFARCAND files

date
cat ${MCFARCAND} | while read FP ; do
ls -l ${FP} ; usleep 200000
done 2>&1 | tee ${MCFARCANDL}
date

Wed May 21 07:30:53 CDT 2014
-rw-r--r-- 1 rubin e875 342776869 Jul 12  2006 /pnfs/minos/mcout_data/S06-05-25-R1-22/far/cand_data/f21001008_
0000_L010185.cand.S06-05-25-R1-22.root
-rw-r--r-- 1 rubin e875 339281368 Jul 12  2006 /pnfs/minos/mcout_data/S06-05-25-R1-22/far/cand_data/f21001002_
0000_L010185.cand.S06-05-25-R1-22.root
...

rw-r--r-- 1 rubin e875 802536135 Jul 27  2007 /pnfs/minos/mcout_data/cedar_phy_srsafitter/far/daikon_02/Cosmic
Mu/cand_data/112/f20011124_0001_CosmicMu_D02.cand.cedar_phy_srsafitter.root
```

```
-rw-r--r-- 1 rubin e875 810706447 Jul 27  2007 /pnfs/minos/mcout_data/cedar_phy_srsafitter/far/daikon_02/Cosmi
cMu/cand_data/112/f20011125_0008_CosmicMu_D02.cand.cedar_phy_srsafitter.root
Wed May 21 13:45:15 CDT 2014

    Remove the files

    root@if-admin-minos

MDC=/minos/app/mindata/log/candx
MCFARCAND=${MDC}/mcfarcand

{
date
cat ${MCFARCAND} | while read FP ; do
echo /bin/rm -f ${FP} ; /bin/rm -f ${FP} ;usleep 200000
done
date
}  2>&1 | tee      ${MDC}/mcfarcand-rm
chown mindata.e875 ${MDC}/mcfarcand-rm

Fri May 23 07:10:00 CDT 2014
/bin/rm -f /pnfs/minos/mcout_data/S06-05-25-R1-22/far/cand_data/f21001008_0000_L010185.cand.S06-05-25-R1-22.ro
ot
/bin/rm -f /pnfs/minos/mcout_data/S06-05-25-R1-22/far/cand_data/f21001002_0000_L010185.cand.S06-05-25-R1-22.ro
ot
...
/bin/rm -f /pnfs/minos/BAD/HOTDOG5/cand_data/N00019245_0003.cosmic.cand.dogwood5.0.root
/bin/rm -f /pnfs/minos/BAD/HOTDOG5/cand_data/N00019245_0003.spill.cand.dogwood5.0.root
Fri May 23 17:44:44 CDT 2014

# grep -v /bin/rm  ${MDC}/badcand-rm
Fri May 23 14:54:20 CDT 2014
Fri May 23 17:44:44 CDT 2014

# grep /bin/rm  ${MDC}/badcand-rm | wc -l
31703

  The file count is correct.
```

#11 - 05/23/2014 06:25 PM - Arthur Kreymer

```
    SMALLCAND FILES

There are 1604 files left in small directories.
Most of these are hidden behind symlinks,
and/or have been mv'd without enmv,
making the Complete File Listing out of skew.

DATA > grep -v /pnfs/minos/reco_near/ cflcands | grep -v /pnfs/minos/reco_far/ | grep -v /pnfs/minos/mcout_dat
a/ | grep -v /BAD/ | wc -l
1604

grep -v /pnfs/minos/reco_near/ cflcands | \
grep -v /pnfs/minos/reco_far/ | \
grep -v /pnfs/minos/mcout_data/ | \
grep -v /BAD/ > smallcand

DIRS=`cat ${SMALLCAND} | cut -f 4 -d / | sort -u`

for DIR in ${DIRS} ; do
printf "%-25s " ${DIR} ; grep /${DIR} ${SMALLCAND} | wc -l
done > ${MDC}/smallcand-count

alignment                 1
reco_data                 1582
sim_reco                  21

    Get listing of SMALLCAND files

date
cat ${SMALLCAND} | while read FP ; do
ls -l ${FP} ; usleep 200000
done 2>&1 | tee ${SMALLCANDL}
date
```

```
Fri May 23 18:06:00 CDT 2014
ls: /pnfs/minos/sim_reco/far/cosmic/cand/cosmic_mu_r230.cand.R0.8.0.root: No such file or directory
ls: /pnfs/minos/sim_reco/far/atmos/cand/atmos_nu_r110.cand.R0.8.0.root: No such file or directory
ls: /pnfs/minos/sim_reco/far/cosmic/cand/up_mu_r300.cand.R0.8.0.root: No such file or directory
ls: /pnfs/minos/sim_reco/far/cosmic/cand/cosmic_mu_r213.cand.R0.8.0.root: No such file or directory
...
-rw-r--r-- 1 minospro e875 113714953 Dec 14  2008 /pnfs/minos/BAD/CDATA/reco_near/cedar/c_data/2008-09/N000148
71_0006.cosmic.cand.cedar.0.root
-rw-r--r-- 1 minospro e875 111966429 Dec 14  2008 /pnfs/minos/BAD/CDATA/reco_near/cedar/c_data/2008-10/N000151
22_0021.cosmic.cand.cedar.0.root
-rw-r--r-- 1 minospro e875 449981679 Dec 14  2008 /pnfs/minos/BAD/CDATA/reco_near/cedar/c_data/2008-09/N000148
71_0006.spill.cand.cedar.0.root
-rw-r--r-- 1 minospro e875 111887950 Dec 14  2008 /pnfs/minos/BAD/CDATA/reco_near/cedar/c_data/2008-10/N000151
22_0022.cosmic.cand.cedar.0.root
...
ls: /pnfs/minos/reco_data/R1.11/cand_data/2004-09/F00027036_0003.cand.R1.11.root: No such file or directory
ls: /pnfs/minos/reco_data/R1.11/cand_data/2004-09/F00027048_0003.cand.R1.11.root: No such file or directory
Fri May 23 18:11:38 CDT 2014

pnfs/minos/mcout_data/R0.8.0 -> /pnfs/minos/sim_reco/far
Paths here changed from cand to cand_data without enmv

Will clean this up next week.
```

#12 - 04/11/2015 07:51 AM - Arthur Kreymer

- Status changed from New to Assigned

- % Done changed from 0 to 90

Set status to Assigned.