Managing Metadata with Antelope –

Keeping up with the 2000+ station EarthScope USArray Transportable Array experiment

Jennifer Eakins

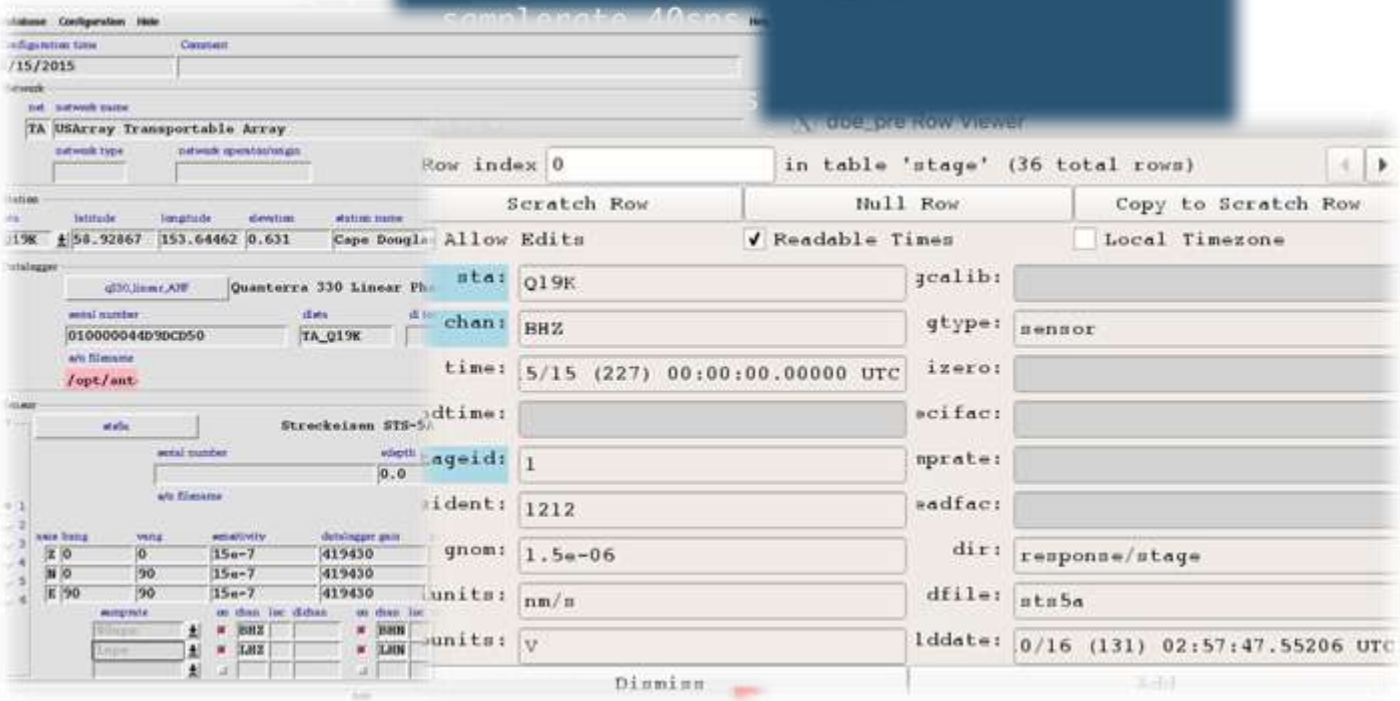
AUG – Rome, Italy

18 May, 2016

Available tools

- dbbuild
  - Batch mode
  - GUI
- seed2db
- dbe

```
time 08/15/2015   0:00:00.000

net TA USArray Transportable Array
sta Q19K 58.92867 -153.64462 0.631 Cape Dougl

# radio (UNAVCO/PBO)
datalogger q330_linear_ANF 010000044D9DCD50 TA

sensor sts5a 0.0 130724  # Streckeisen STS-5A

   axis2 Z 0   0 15e-7 419430 - - 1
   axis2 N 0  90 15e-7 419430 - - 2
   axis2 E 90 90 15e-7 419430 - - 3
```



Overview of metadata

- Rolling deployment since 2004
- Currently:
  - ***309*** stations (120 TA)
  - Field season April-Oct.
- At the peak:
  - ***5-20*** station services, installs, removals per week leading to interrupts 2x/wk
- Overall:
  - ***2004*** total stations

Why not use your dbmaster/ for updates?

Removing database rows for active programs == **problems**!

Could result in longer delays if you shutdown rtsystem:

- Possible errors in batch file
- Incoming dataless SEED may not be correct
- Wrong stations/channels chosen to add/update

Directory structure for dbmaster preparation

- Need working area that is not in-use rtsystem/dbmaster
- Create a pre-dbmaster area
- ta_dbuild area
  - active
  - closed
- Individual network area
- Staged merging zones:
  - CONTRIB…merge
  - all_merge

```
taops.ucsd.edu{rt}522% pwd
/anf/TA/dbs/pre-dbmaster
taops.ucsd.edu{rt}523% ls
CONTRIB_NETWORKS_merge/  iu_only/
ak_only/                 ld_only/
all_merge/               n4@
at_only/                 nn_only/
av_only/                 ny_only/
az_only/                 po_only/
bk_only/                 ta_dbbuild/
ci_only/                 usnsn_only/
cn_only/                 uu_only/
err.merge                wu_only/
ii_only/
taops.ucsd.edu{rt}524%
```

External stations – using seed2db on dataless SEED

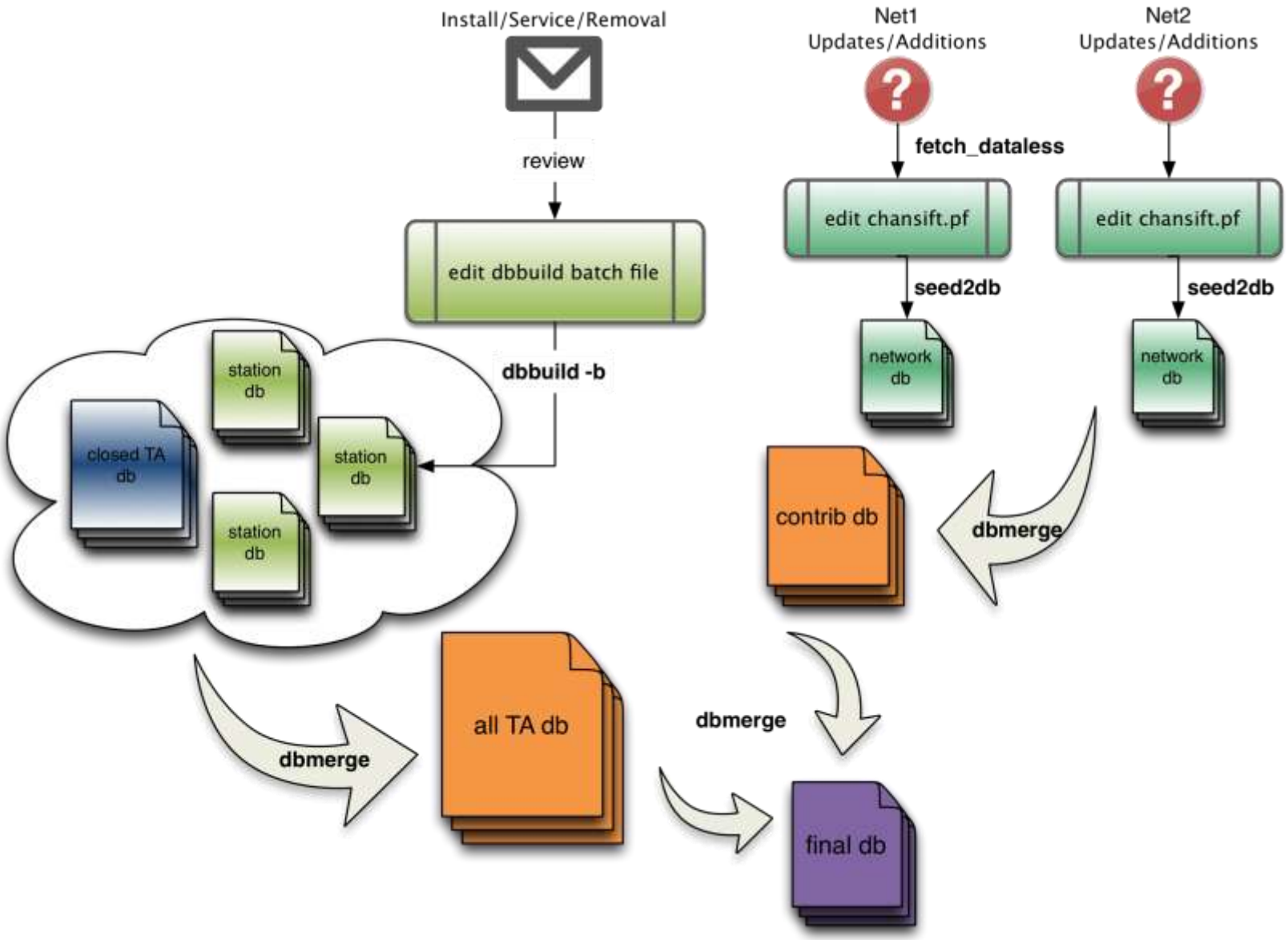
- Develop way to check for and download updated dataless files
- Use chansift.pf
  - Subset sta/chan

```
taops.ucsd.edu{rt}524% ls cn_only
2015-238/              cn_tmp.sensor
2016-015/              cn_tmp.site
CN.dataless*           cn_tmp.sitechan
chansift.pf            cn_tmp.snetsta
cn_tmp.calibration     cn_tmp.stage
cn_tmp.instrument      fetch_dataless*
cn_tmp.lastid          foo
cn_tmp.network         response/
cn_tmp.schanloc
```

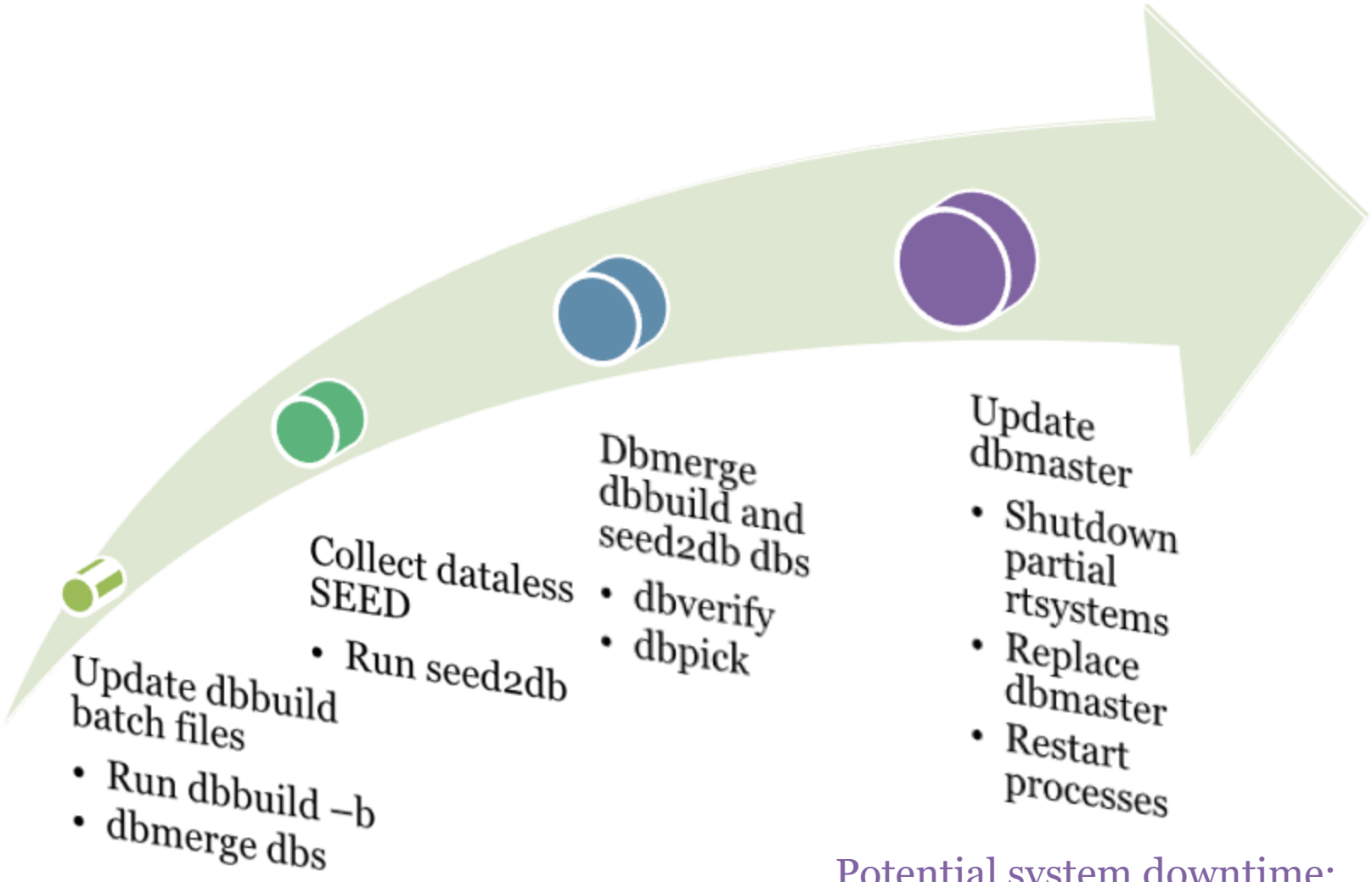
- Command line:

seed2db -respdir response -stagedir response/stage_CN -chansift chansift.pf CN.dataless cn_tmp

Flowchart of build process



Approximate times for prep vs. system downtime



Partial Shutdown – no rtexec -k

- acq host (q3302orb)

```
# turn these off during dbmaster updates
#
q3302orb_prelim yes
q3302orb_prelimBGANb    yes
#
q3302orb_AKCAN  yes
q3302orb_BGAN   yes
q3302orb_Low48  yes
q3302orb_Strays yes
q3302orb_EARN   yes
q3302orb_TPFO   yes
q3302orb_CASC   yes
q3302orb_DIP    yes
q3302orb_NoEP   yes

# collect datalogger logs
q330logs2db     yes
q330logs2db_CASC        yes
```

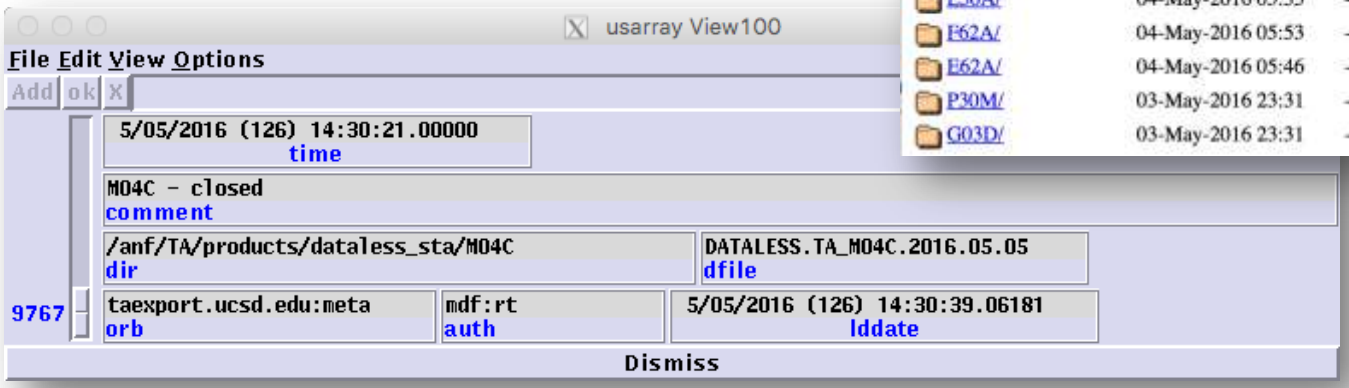
- ops host (db and wf writes)

```
 turn off for dbmaster updates
orbdetect       yes
orbassoc        yes
orbevproc       yes

orb2dbSeismic   yes
orb2dbInfraMet  yes
orb2db100       yes
orb2wf_reg      yes
orb2wf_AKinframet       yes
orb2wf_AKSOH    yes
orb2db_prelim   yes
orb2db_soh      yes

orb2dbt         yes
orb2dbt_orig    yes

orb2dbt_soh         yes
orb2dbt_prelim      yes

orb2logs        yes
orb2logs_prelim     yes

assoc_SCEC          yes
```

```
q3302orb_AKCAN  q3302orb  -v -noipresolv -S state/q3302orb_AKCAN  -pf q3302orb_AKCAN
 -calib_db $DBOPS/usarray tadataAKCAN wforb  $ORB  dataorb $ORB  cmdorb $CMDORB poco
rb $POCORB statusorb $STATUSORB
```

Exporting metadata

- mk_dataless_seed
  - Distribute via email, ftp, or orbxfer
  - Keep track of distribution with dmcfiles



- db2stationxml

Summary

- Antelope provides a top to bottom solution for metadata creation
- Antelope tools available for metadata import/build: *dbbuild* & *seed2db*
- Limit processing and possible acquisition downtime by creating a “pre-dbmaster” area

Extra – dbverify output, what can I ignore?

To be fixed:

- chanid
- calib/calper
- hang/vang
- arid/orid/evid/wfid/magid
- duplicate arrival/origins
- hang/vang

To ignore:

- hang/vang
- site lat/long matches

dbverify – things to ignore after dbmaster update

- Some hang/vang errors

```
dbopen sitechan
        dbsubset hang!=NULL
        dbsubset chan=~/.*Z/
        dbverify abs(hang)<5 sta ch
Record #857 : chanid =      6624
false E12A    BHZ         180.0
Record #858 : chanid =      6627
false E12A    LHZ         180.0
Record #859 : chanid =      6630
false E12A    VHZ         180.0
Record #860 : chanid =      6633
false E12A    UHZ         180.0
        4 failures of 'abs(hang)<5'.
```

```
false P25A    BDE         0.0
Record #8802 : chanid =     69340
false P25A    LDE         0.0
Record #8803 : chanid =     69341
false P25A    VDE         0.0
Record #8804 : chanid =     69342
false P25A    UDE         0.0
Record #9560 : chanid =     76208
false Y22D    BDE         0.0
Record #9561 : chanid =     76209
false Y22D    LDE         0.0
Record #9562 : chanid =     76210
false Y22D    VDE         0.0
Record #9563 : chanid =     76211
false Y22D    UDE         0.0
        408 failures of 'abs(hang-90)<5||abs(hang-270)<5'.
```

- site lat/long matches

```
dbopen site
        dbsort lat lon sta
        dbgroup lat lon sta
        dbfind_dups lat .0001 lon .0001

Records #391 and #392 match:
 lat            :       34.945500       34.945500       0.000000        0.000100
 lon            :     -106.460000     -106.460000       0.000000        0.000100
Records                         #  391                   #  392
 lat            :                  34.9455                    34.9455
 lon            :                 -106.4600                  -106.4600
 sta            :                    ASM                        TASN
 bundle         :
 bundletype     :                      1                          1
```

dbverify – beyond dbmaster: id issues, calib/calper

- Chanid

  dbfixchanids dbname

```
dbopen wfdisc
        dbjoin sensor
        comment check that chanids are set more or less correctly
        dbverify wfdisc.chanid==sensor.chanid wfdisc.chanid sensor.chanid
Record #129767 : sta = G03D
endtime =  4/26/2016 (117)
false    75906    75921
Record #129769 : sta = G03D
endtime =  4/26/2016 (117)
false    75907    75922
```

```
dbopen wfdisc
        dbjoin calibration
        comment check that wfdisc calib's are set (properly)
        dbverify wfdisc.calib==calibration.calib wfdisc.calib calibration.calib
Record #171686 : sta = P19K    chan = BHZ       time =  5/18/2016 (139) 10:24:04.525
endtime =  5/18/2016 (139) 10:42:29.500
false          1     1.5895
Record #171687 : sta = P19K    chan = BHN       time =  5/18/2016 (139) 10:24:04.525
endtime =  5/18/2016 (139) 10:42:29.500
false          1     1.5895
```

- Calib/calper

  dbjoin $db.wfdisc calibration | dbselect -s -

  "wfdisc.calib==calibration.calib" > /dev/null

dbverify – beyond dbmaster: del phases, arrival/origin matches

- del phases

  dbsubset –v dbname.arrival "iphase=='del'" |

  dbdelete –sv -

- arrival/assoc

- origin matches

```
dbopen assoc
        dbjoin arrival
        dbjoin origin
        dbjoin site
        comment verify that assoc station matches arrival station
        dbverify assoc.sta==arrival.sta arid orid assoc.sta arrival.sta
        comment check that assoc.delta corresponds to the computed distance
        dbverify abs(delta-distance(origin.lat,origin.lon,site.lat,site.lon))<.001 del
ta distance(origin.lat,origin.lon,site.lat,site.lon)
        comment check that the arrival follows the event
        dbverify arrival.time-origin.time>0  origin.time arrival.time
        comment check that computed arrivals are close to actual arrivals
        dbverify phase!~/P.*/||(abs(parrival()-arrival.time)<10) phase parrival()-arri
val.time
Record #3513 : arid =  2882922  orid =  3877500  sta = LRL     time =  4/04/2016 (095)
 11:48:58.656  lat =   30.2995  lon = -113.6842  depth =    0.0000  ndef =   19  nass
=   19  ondate =  1992211  offdate =       -1
false Pb          -16.576
Record #4912 : arid =  2883665  orid =  3960316  sta = AGMN    time =  4/03/2016 (094)
  8:37:56.954  lat =  -14.3519  lon =  166.8205  depth =   35.0000  ndef =    0  nass
=  245  ondate =  2006232  offdate =  2599365
false Pdiff        255.94
```