Hardware Scaling Lessons from the TA

Frank Vernon

AUG - IRIS
7 June 2008
Skamania, Washington

ANF Operations Year 1 - 2004

- TA Field
  - 13 TA Stations
  - 66 Contributed stations (CI, BK, AZ)
- Server Hardware
  - Sun Blade 1000
- Software
  - Nagios - system monitoring

ANF Operations Year 2 - 2005

- TA Field
  - 66 TA Stations
  - 65 Contributed stations (CI, BK, AZ, NN)

- Server Hardware
  - Sun Blade 1000 removed in Fall - Load average too large
  - Sun V240 installed in Fall
  - 3.5 TByte SCSI storage array installed
  - Linux Webserver

- Software
  - Nagios - system monitoring
  - cfengine - system configuration

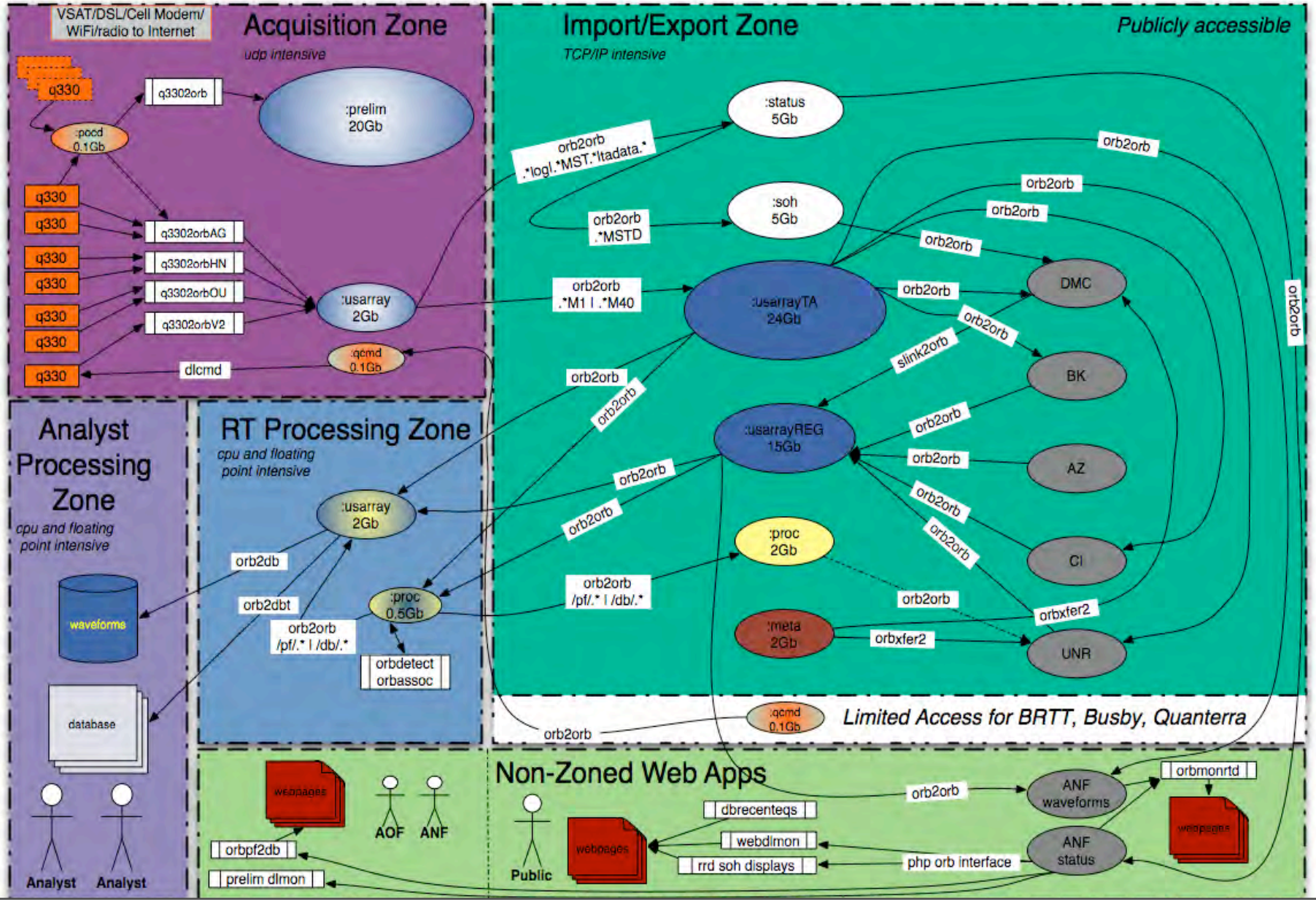
ANF Operations Year 3 - 2006

- TA Field
  - 232 TA Stations
  - 75 Contributed stations (CI, BK, AZ, NN, US)
- Server Hardware
  - Sun V240 - load average problems
  - Sun T2000 installed - moved acquisition
  - 3.5 TByte SCSI storage array
  - SRB Brick brought online
  - SDSC V240 backup brought online
  - DMC V240 backup brought online
  - ANF export V100 brought online
  - iSCSI Storage Area Network implemented
  - Linux Webserver
- Software
  - Nagios - system monitoring
  - cfengine - system configuration

ANF Operations Year 4 - 2007

- TA Field
  - 411 TA Stations
  - 77 Contributed stations (CI, BK, AZ, NN, US)
- Server Hardware
  - Sun Cluster Installed with zones - 2 V240s, 1 V245, and 3 T2000
  - SRB Brick
  - SDSC V240 backup replaced with T2000
  - DMC V240 backup replaced with T2000
  - iSCSI Storage Area Network - 3.5 TByte capacity expanded to 15 Tbytes
  - Veritas VxFS replaces ufs file system - inode problems
  - Linux Webserver
- Software
  - Nagios - system monitoring - Deprecated
  - intermapper - system monitoring installed
  - cfengine - system configuration

RT system at the Array Network Facility



Backup system at SDSC and DMC

ANF Operations Year 5 - 2008

- TA Field
  - 442 TA Stations
  - 57 Contributed stations (CI, AZ, NN, US)
- Server Hardware
  - Sun Cluster - 3 V240s and 3 T2000
  - SRB Brick
  - iSCSI Storage Area Network 15 TBytes
  - Decommissioned Sun Cluster
    - PxFS replaced with QFS
    - Discovered Sun Cluster does not support iSCSI
    - Kept zone functionality
  - Installed 3 T5220 for web support
- Software
  - intermapper - system monitoring installed
  - cfengine - system configuration
  - Confluence installed for ANF Wiki

ANF Operation Zones

- Real time
  - anfops
    - Q330 acquistion
  - anfexport
    - acquire Q330 field station data
    - acquire contributed regional/national network data
    - serve data to internal and external clients
  - anfproc
    - realtime event processing
  - anfwf
    - waveform writer
  - anfanalyst
    - analyst review

ANF Operations Zone

- Miscellaneous
  - anfdev
    - development and testing
  - anfpublic
    - limited public access
  - anfmon
    - intermapper
- Web
  - anfwebproc
    - backend web page production
  - anfwebproj
    - web page exports
  - anfwebtest
    - web content development and testing

Metrics and Applications from the TA

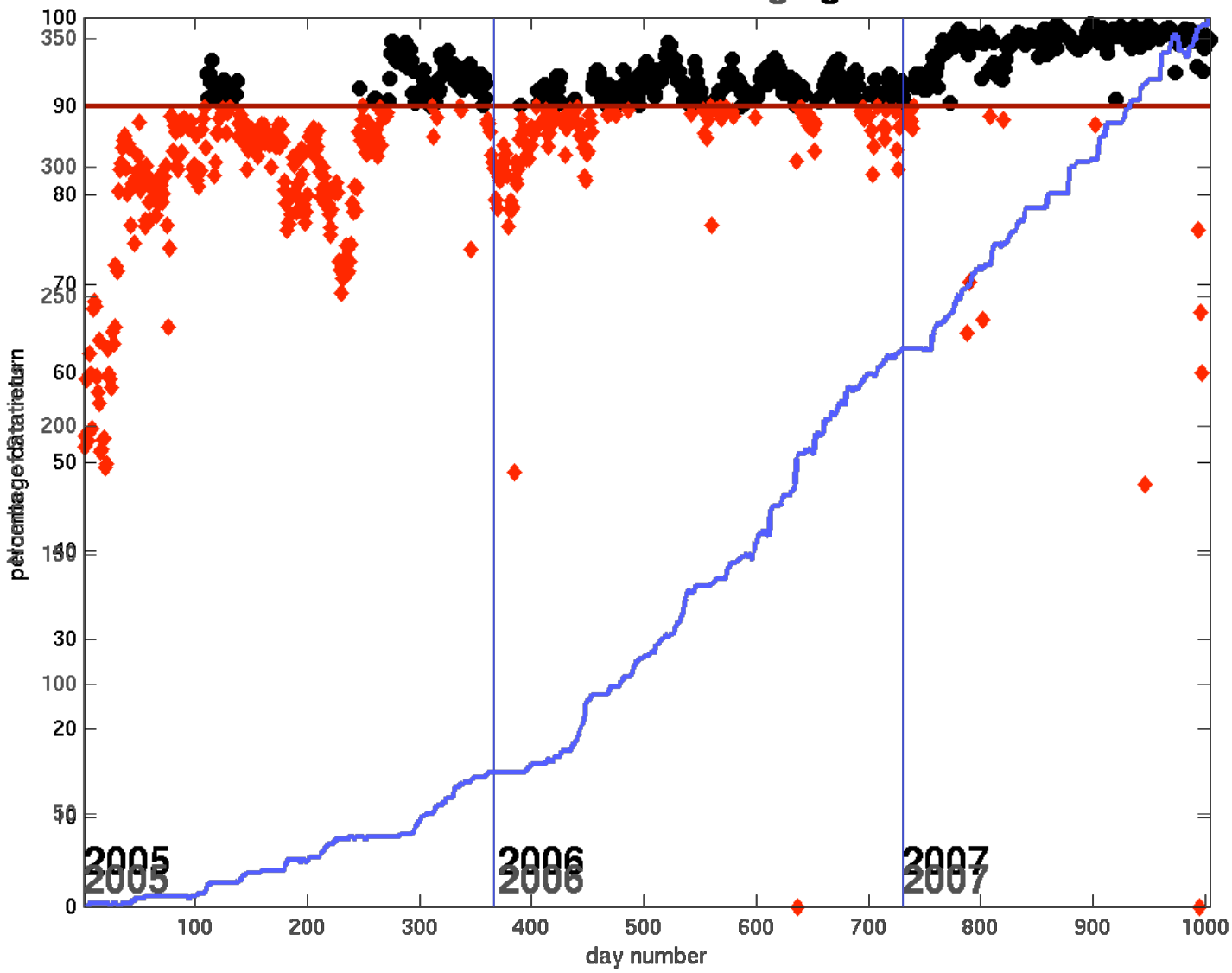
Frank Vernon

AUG - IRIS
7 June 2008
Skamania, Washington

USArray Data Flow at ANF

- 3 Tbytes of data - April 2004 - June 2008 (compressed)
- As of June 2008
  - 4 Gbytes/day compressed data
  - 2 Mbit/sec data export
  - 436 seismic stations
  - 2616 seismic channels
  - 13516 soh channels
  - 1.5M picks
  - 32K events

TA real-time data return starting 1 Jan 2005

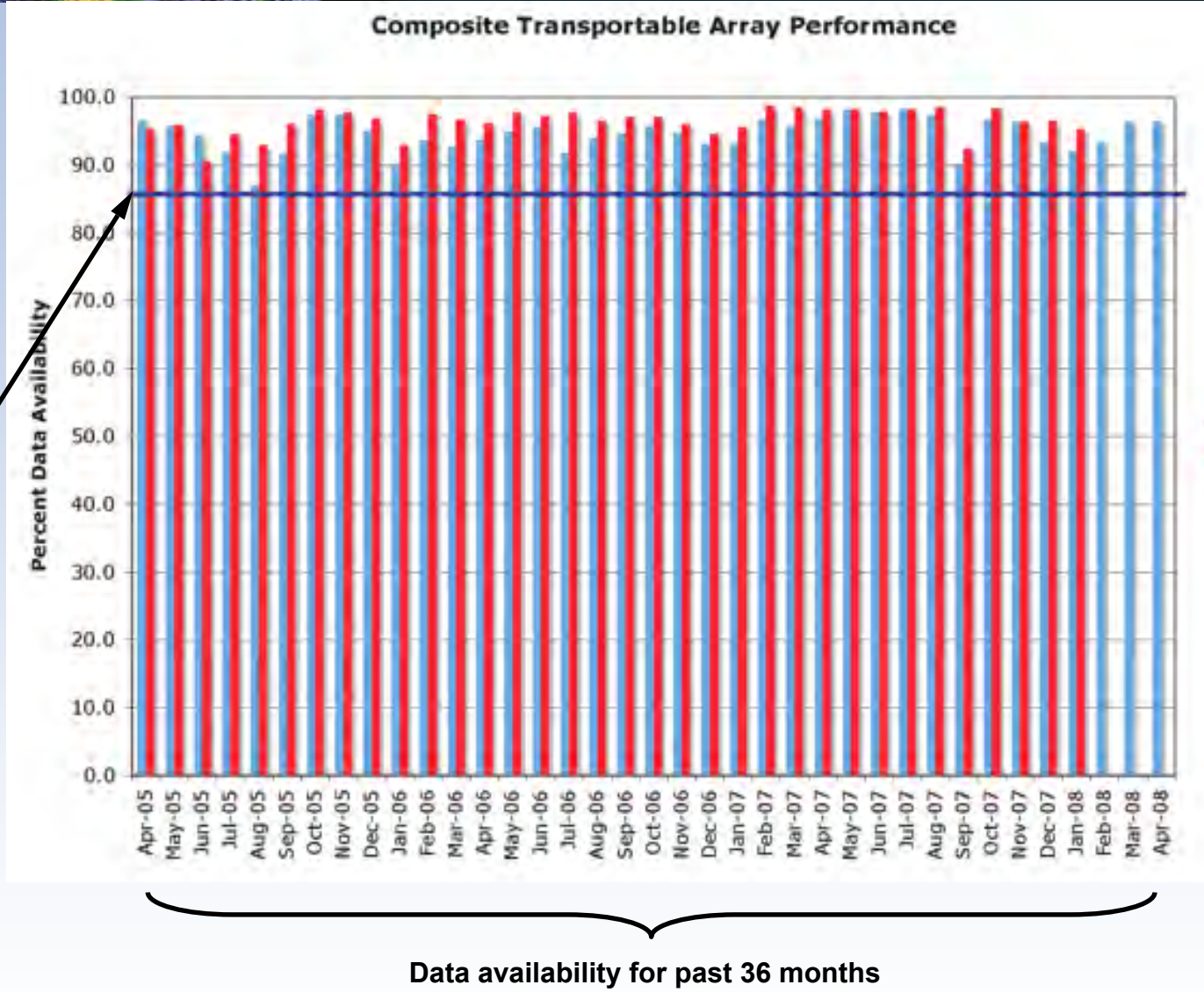


Operational Statistics

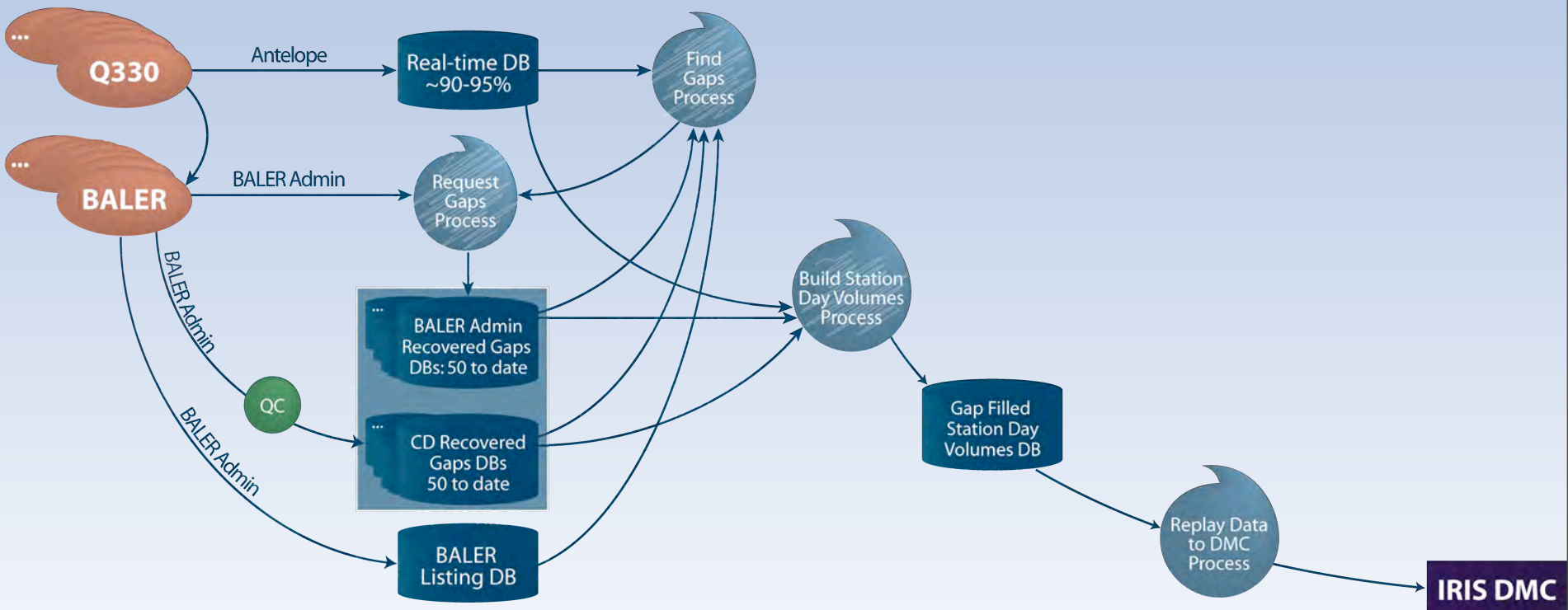
Availability (uptime)

- Consistently > 90%
- Redundant data recovery mechanisms
- Acquiring ~ 1.2 Tb/y (~5 Gb/day)

85% is official performance goal



Data gaps: recovery & integration to database



- Identification of data gaps;
- Recovery of data gaps using BalerAdmin software;
- Building databases from field cdroms;
- Rebuilding station-day volumes once all gaps are filled for each station-day;
- Replay updated gap filled station-day volumes to DMC for arc

Monthly gap processing - gathering data

- miniseed2db - build monthly rt wfdisc
- rt_daily_return - final gap identification
- dmcgap2db - convert DMC gap list into db
- baler loop
  - build_baler_data - build non-overlapped wfs
    - cdroms - multiple dbs
    - baler_admin - multiple dbs
  - gap_status - identified recovered gaps
  - baler_request - baler_admin input
  - baler_admin
- interate loop several times ~ 1-2 days

Monthly gap processing - sending data

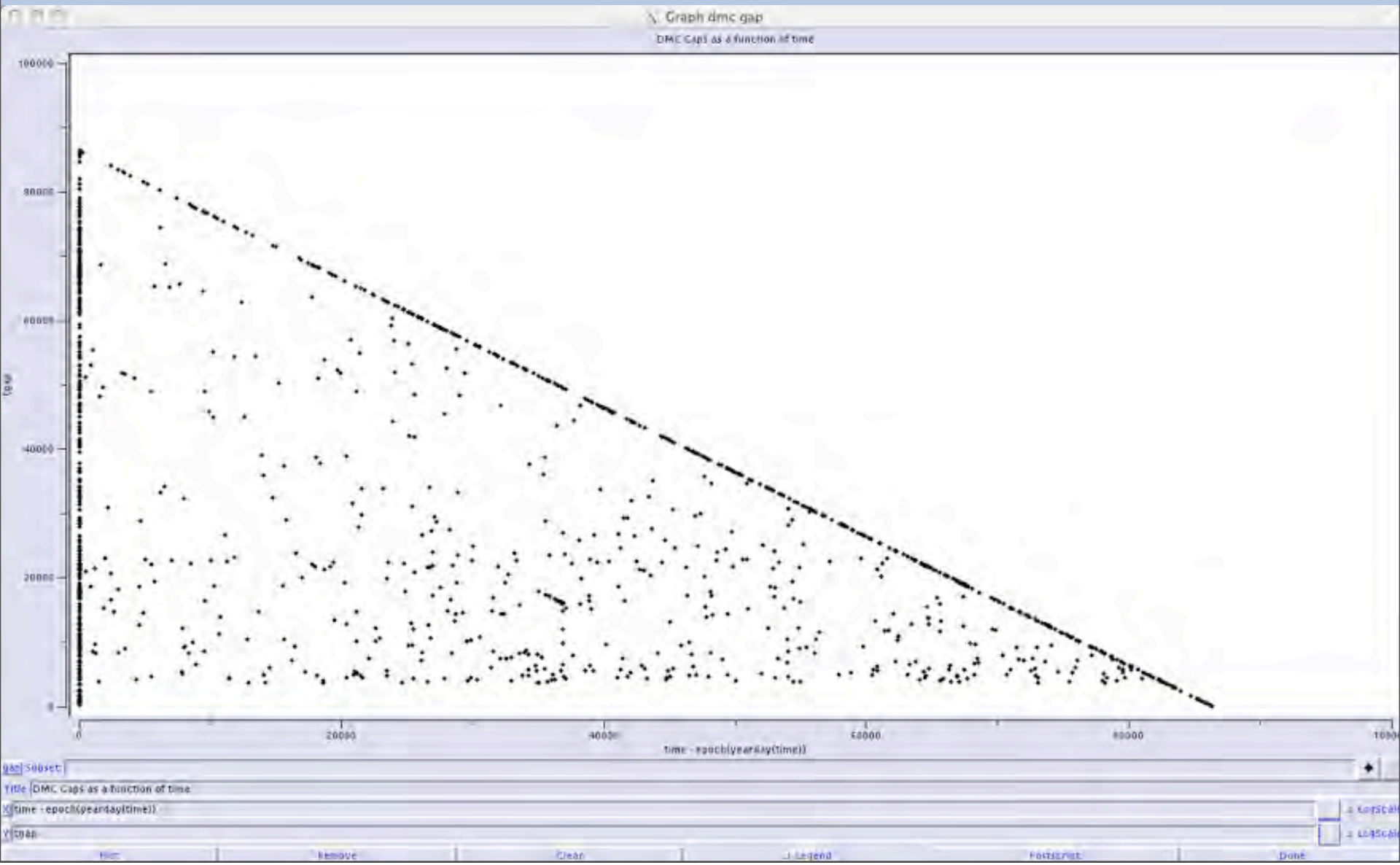
- build_baler_data - build non-overlapped wfs
  - cdroms - multiple dbs
  - baler_admin - multiple dbs
- fill_gaps
  - builds gap replaced station/day volumes
- dbreplay
  - sends repaired data to DMC

- START NEXT MONTHS PROCESSING!

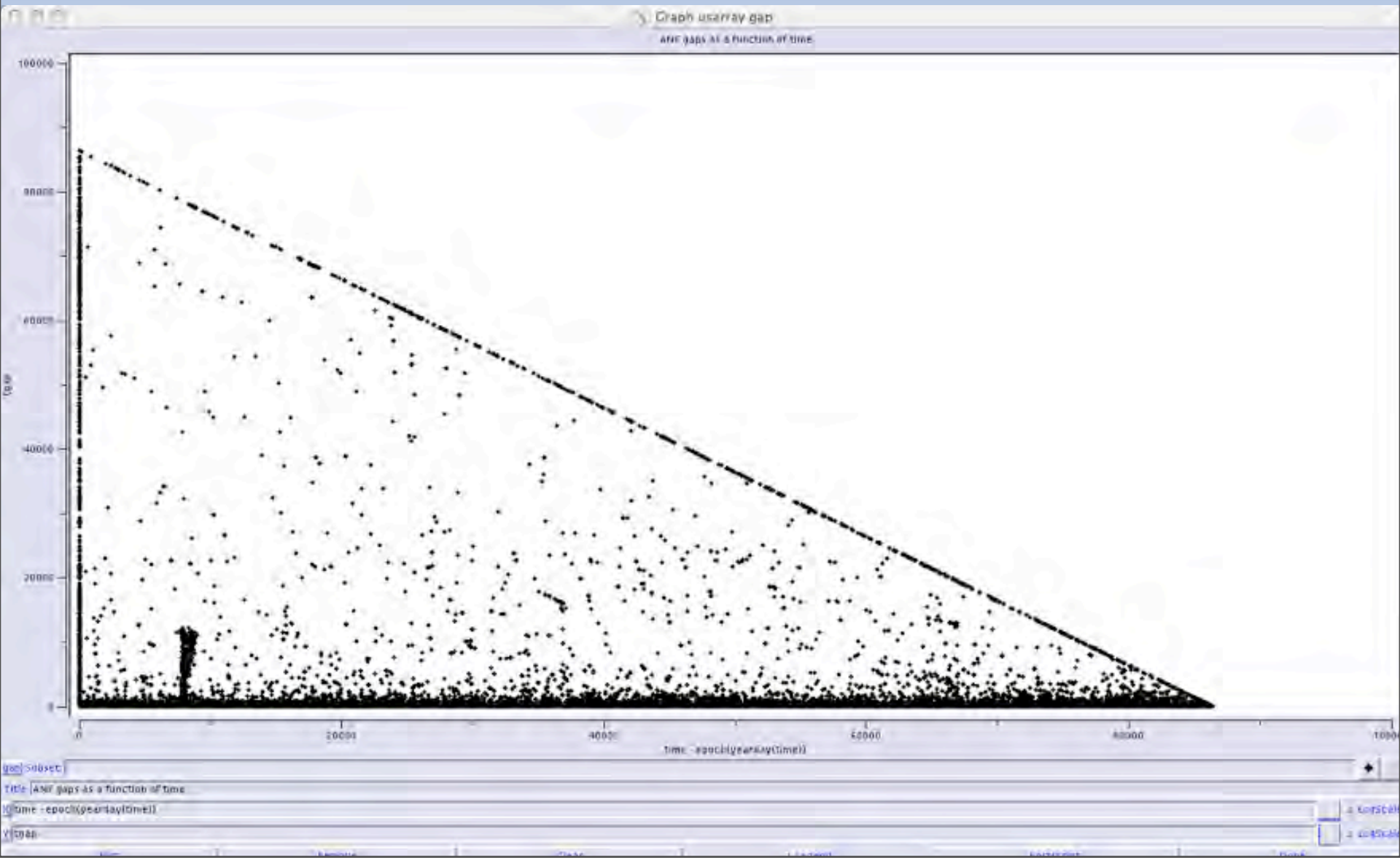
2008 Stats

- 1 Jan - 16 May
  - 50307 Station Days
  - 44585 Station Days 100% data return - 88%
  - 94.3% Total Data return
- 10 March - 16 May
  - 87% Station/Days 100% data return
  - 96.9% Total Data return
- 2398 DMC identified gaps
- 65431 ANF identified gaps

DMC


Graph dmc gap
DMC Gaps as a function of time
100000
80000
60000
40000
20000
0
gap
0
20000
40000
60000
80000
100000
time - epoch(yearday(time))
Subset
Title
DMC Gaps as a function of time
X time - epoch(yearday(time))
Y gap
Plot
Remove
Clear
Legend
Postscript
Done

ANF


Graph usarray gap
ANF gaps as a function of time
100000
80000
60000
40000
20000
0
0
20000
40000
60000
80000
100000
time - epoch(yearday(time))
Title ANF gaps as a function of time
time - epoch(yearday(time))
gap

Orientation

http://www.ldeo.columbia.edu/~ekstrom/Projects/USARRAY/POLARIZATION/

- Empirical orientation determination using surface and mantle wave polarization techniques of USArray and other networks

- TA made direct measurement of orientation of stations using fiber-optic gyroscope

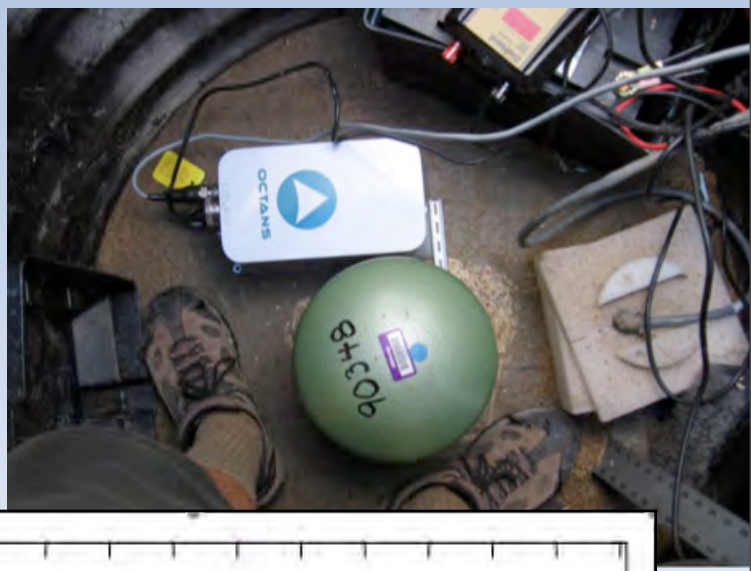
  IXSEA Octans IV

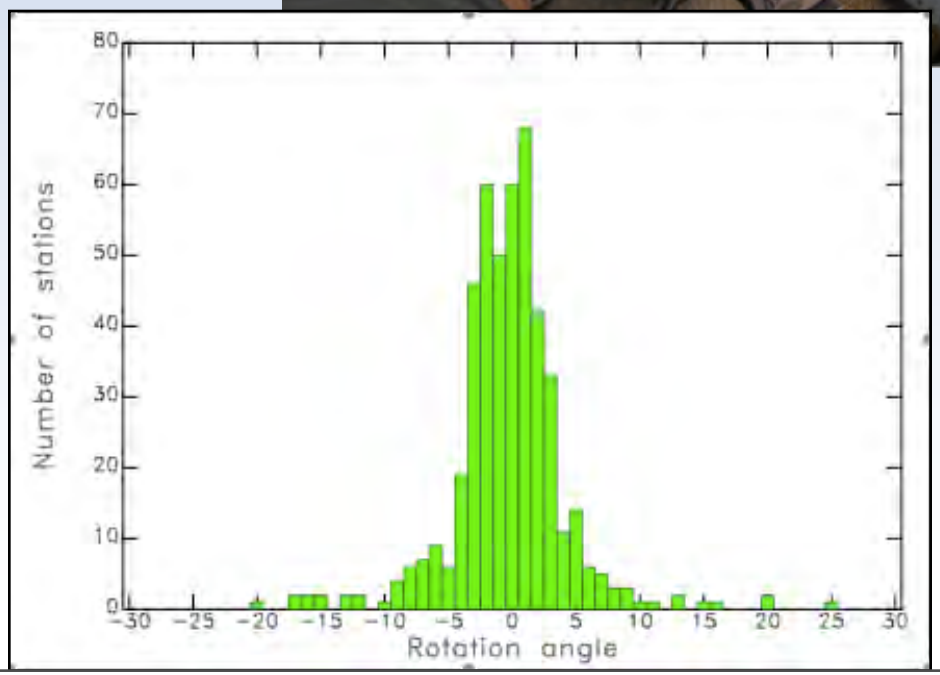
  (Nonmagnetic orientation accurate to < 0.2 degrees)

- Techniques agree to within 1.2 Degrees

- TA uses Octans at all new station installations and on station removal.

(Ekstrom, Busby submitted SRL 2008)

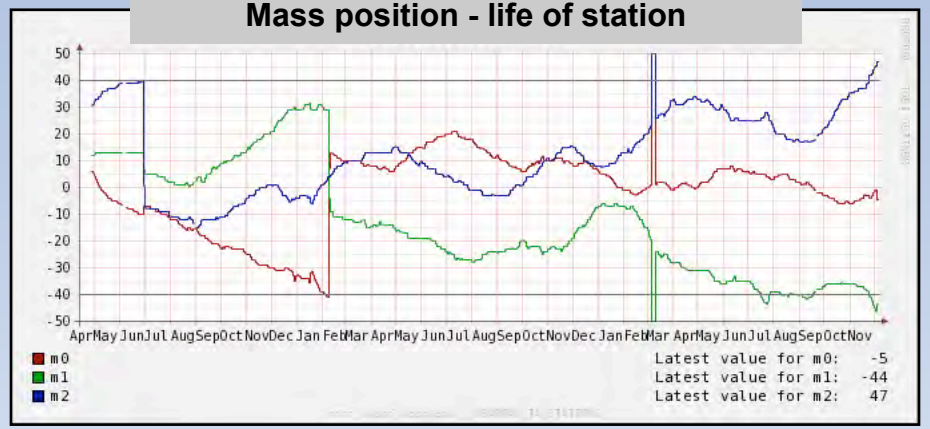




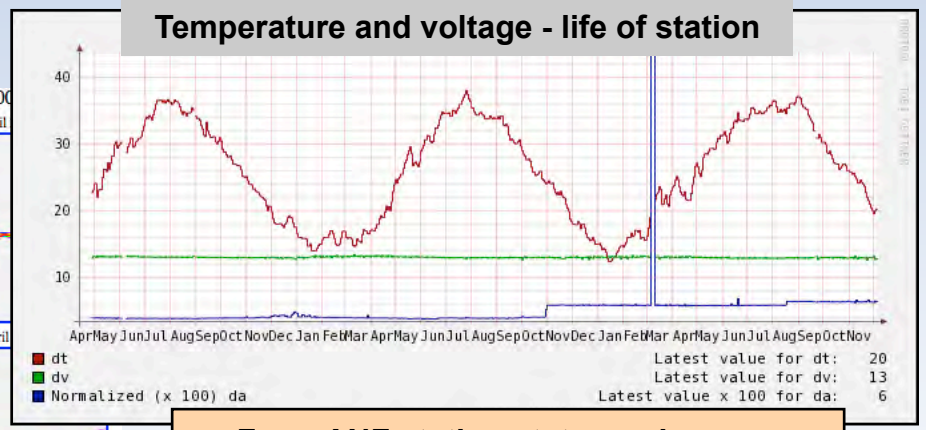
State of Health Review

- Real-time monitoring of SoH
  - Detect problems
  - Initiate corrective actions
- Station QC & SoH on the web
  - SoH channel displays for near-real-time and summary
  - Metrics for arbitrary time intervals

Mass position - life of station

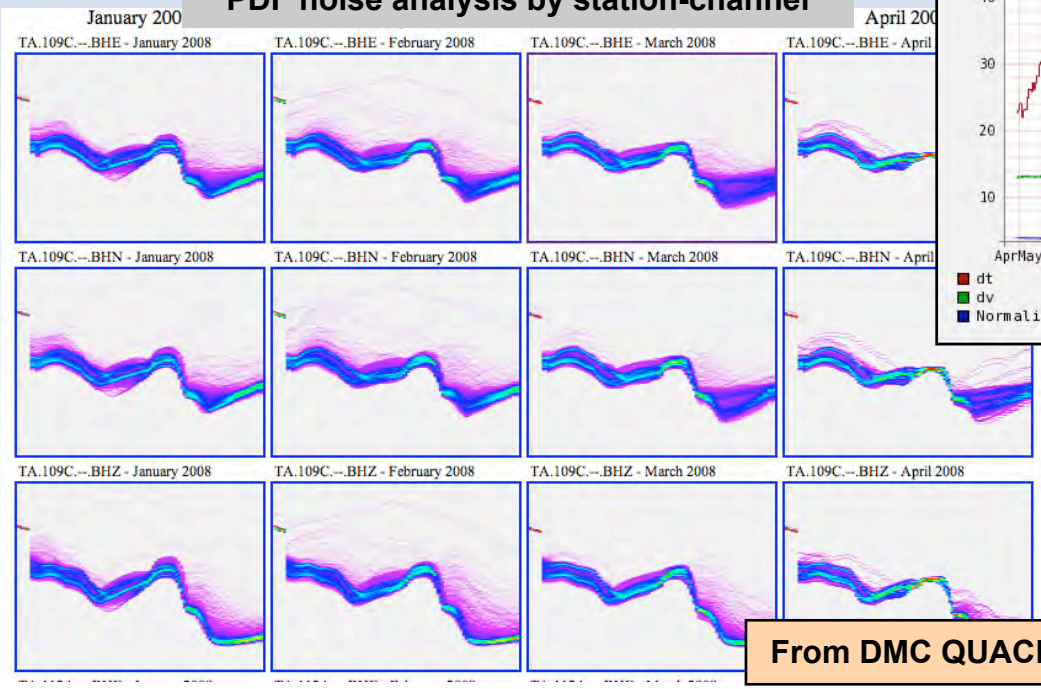


Temperature and voltage - life of station



From ANF station status web pages

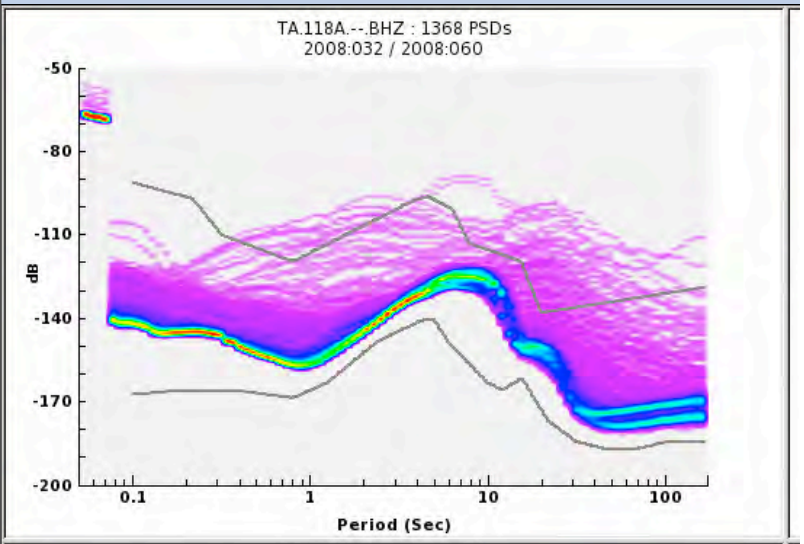
PDF noise analysis by station-channel

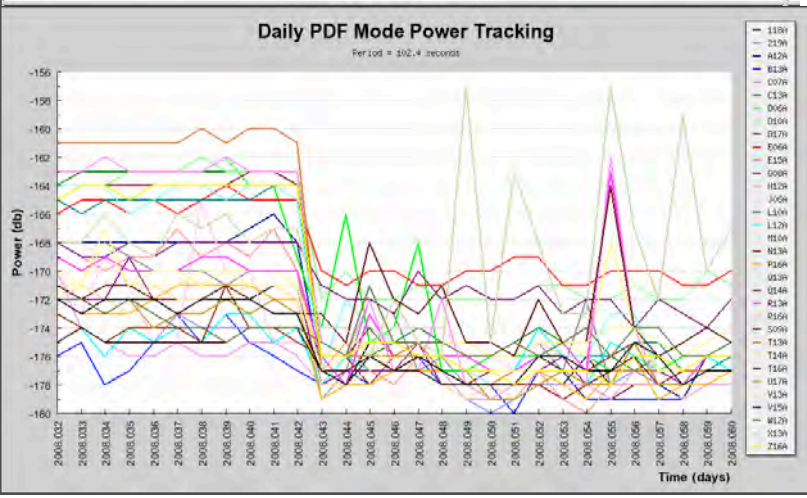


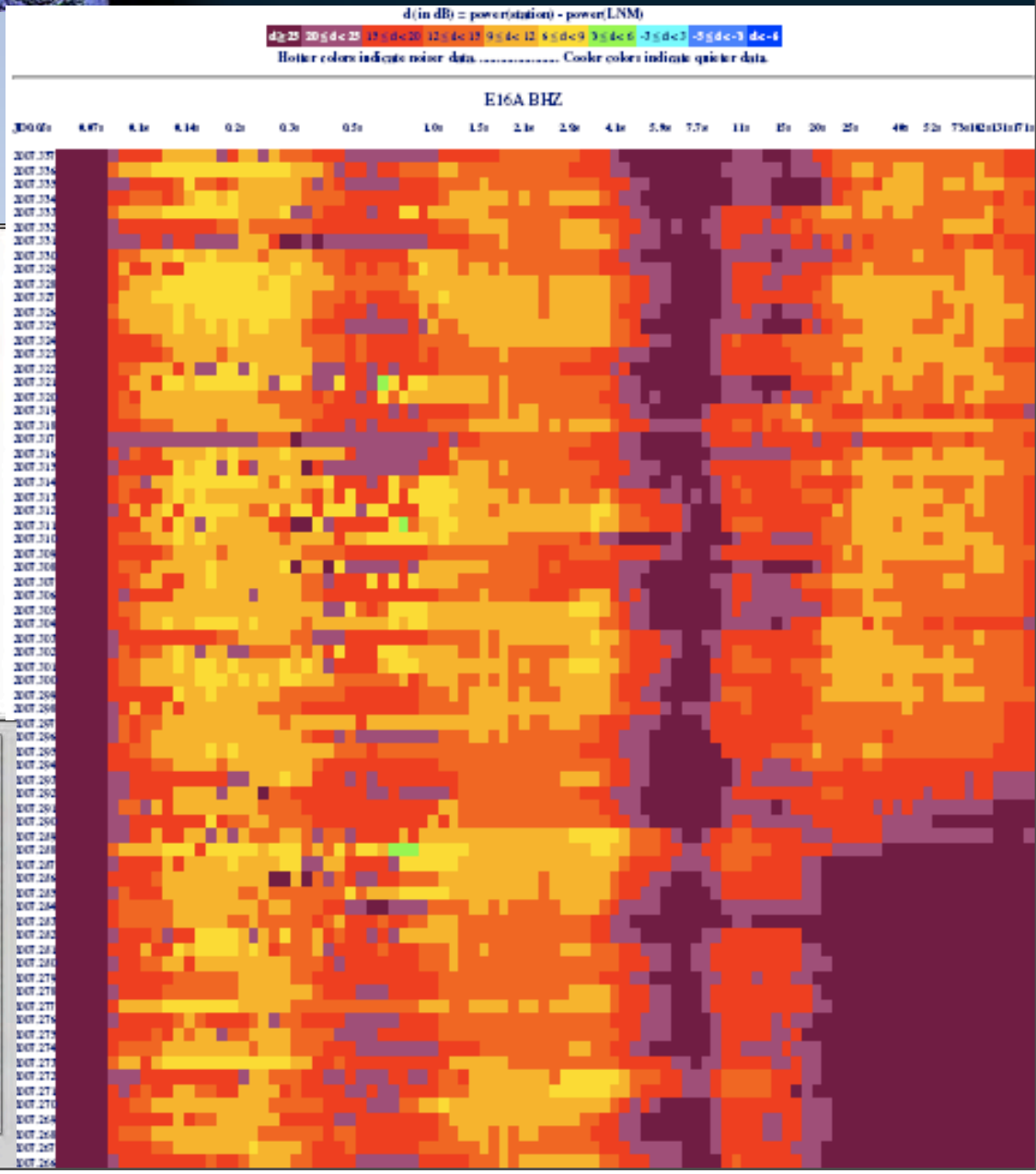
From DMC QUACK tool

Diagnostic view: discovery!

Guralp Vertical: Mass Pos vs noise

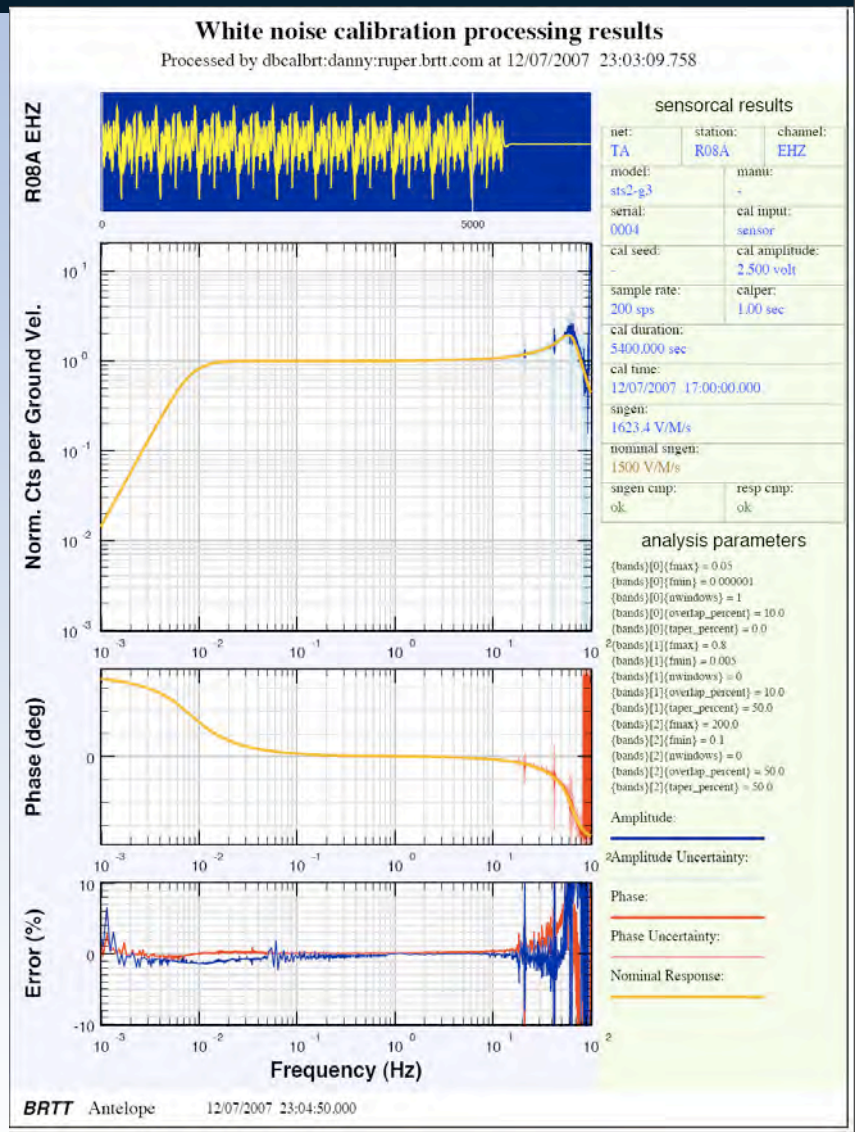






Calibration

- Automated process to command, capture and analyze cal signals applied in situ using Antelope.
- Interpret calibration analyses to verify amplitude and phase response, stationarity of sensor
- Will be applied to all stations at beginning and end of deployment.
- Will be archived as Data Product



User Commands DBCALIBRATE(1)

NAME

dbcalibrate – sensor and cross comparison calibration analysis program

SYNOPSIS

dbcalibrate [-out dbout] [-prm] [-p pfname] [-calper calper]
[-resp_dir resp_dir] [-resp_dfile resp_dfile]
[-resp_dfilee resp_dfilee] [-resp_dfilen resp_dfilen]
[-dlcalwf_sifter expr] [-sngen sngen] [-outreono] [-v]
[-error_at_calper] [-template name] [-dbcmp dbcmp]
[-noise tstart_noise] [-type {ratio|power|coherence}]
dbin [sequence_id [sequence_id_cmp[,chan_cmp]]]

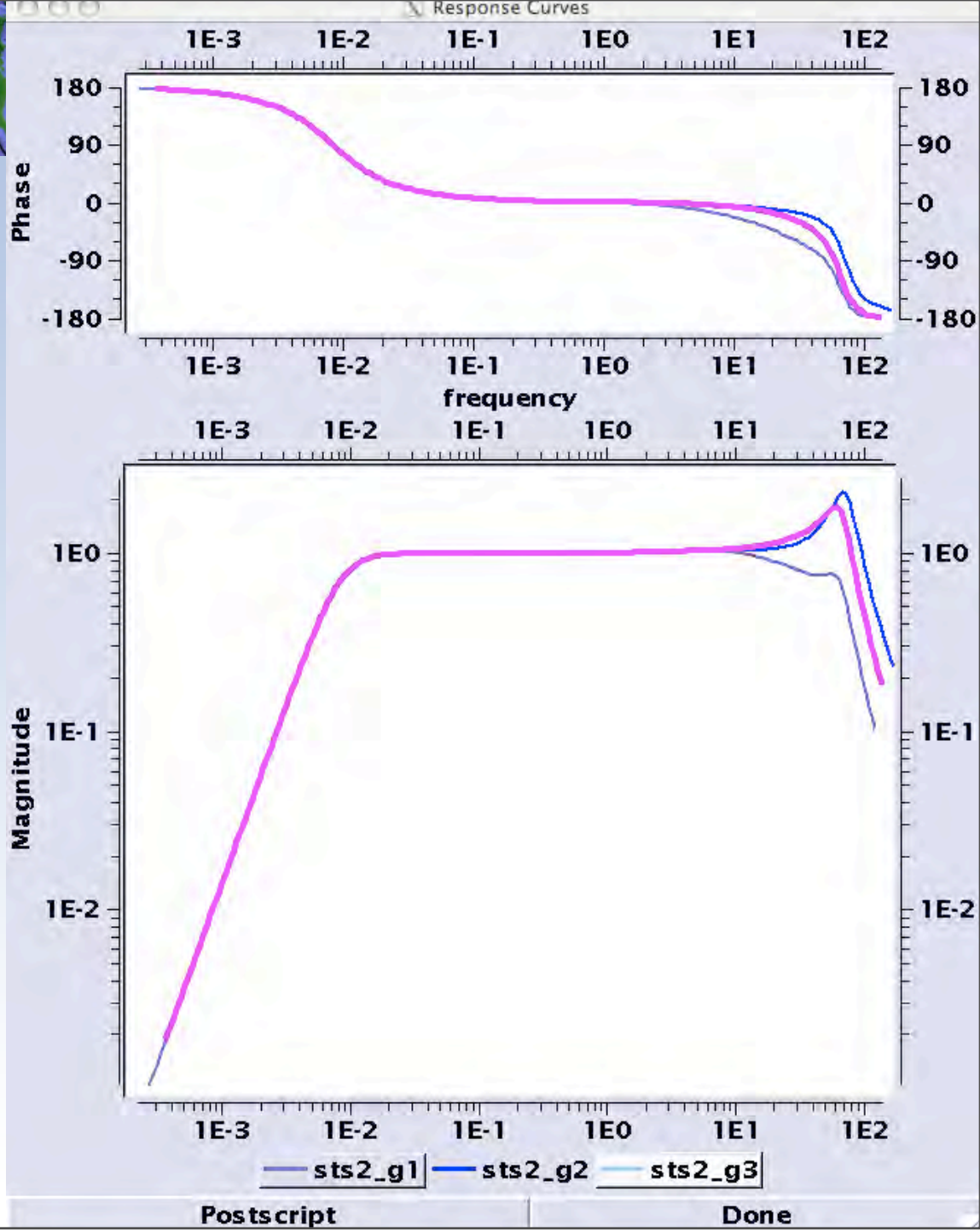
n BRTT Antelope software

Metadata:

Instrument responses

STS-2

3 Generations



Metadata:

Instrument responses

TA Sensors

3 Types