Data Access and Data Merging Studies for the E3 and E4 Engineering Runs

Peter Shawhan
LIGO / Caltech

Patrick Sutton, Gabriela Gonzalez, Vijay Chickarmane
Penn State University

LIGO Scientific Collaboration Meeting
August 15, 2001

Overview

Nominal description: "Verify that an analysis using corresponding channels from the two sites sees sensible results (e.g. correct channel names, sampling rates, calibration scale factors)."

Goals:

- Provide a guide to data access
- Provide instructions for merging frame files
- Put information about channel lists and time intervals on the web
- Compare channel naming conventions and sampling rates
- Compare means and variances for selected channels

Final report has DCC number: LIGO-T010092-00-D

Data Access

The HPSS archive at Caltech has all "full" frame data from E3 and E4, plus second- and minute-trends

E4 RDS data is still on disk at LLO, but E4 RDS data at LHO and E3 RDS data at both sites has been deleted

List of available datasets (with list of channels and valid time intervals) is reachable from the LIGOtools home page (http://www.ldas-sw.ligo.caltech.edu/ligotools) by clicking on "Universal Dataset Names (UDNs), time ranges, channel lists, and links to servers"

The "getFrames" utility may be used to retrieve data subsets, e.g:

```
getFrames -d //ligo/raw/llo/e4 -t 673800000-673800100 -c "L1:LSC-AS*" -o
```

This is described in a LIGOtools FAQ
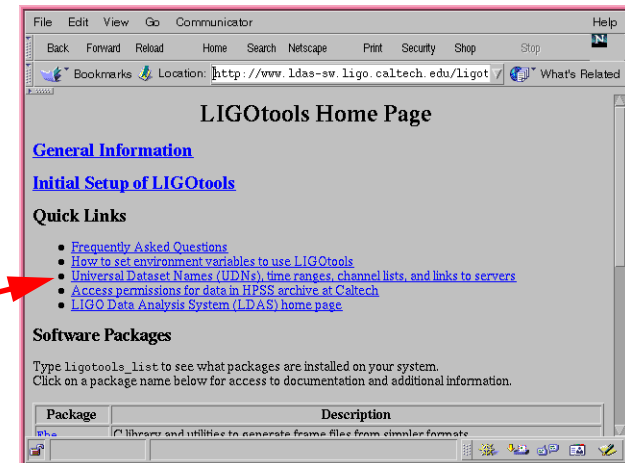
How to Merge Frame Files

Basic technique:

```
FrCopy -i H-123456789.F -i L-123456789.F -o myout.F
```

Note that each of the input files is preceded by `-i`

You can use a list of files as input, *e.g.*:

```
FrCopy -i H-1234567*.F -i L-1234567*.F -o myout.F
```

If you want to merge just a subset of channels, use the `-a` flag:

```
FrCopy -i H-1234567*.F -a H0:PEM-LVEA_SEISX \
     -i L-1234567*.F -a "L0:PEM-*_SEIS*" -o myout.F
```

To look at data with DTT or Data Viewer, you will need to split the data into individual one-second-long frame files, *e.g.*:

```
FrSplit -i myout.F -o datadir -n 1000
```

(There are LIGOtools FAQs with this information)

Channel List Comparisons

Only one channel was in RDS but not "full" data: H0:PEM-NBR_2K

Found some duplicated channels in RDS datasets:
H1:ASC-ETMX_OptLevYaw, L0:PEM-EX_V1

Noted some differences in channel naming conventions, *e.g.*:

- LHO magnetometer channel names are inconsistent:
  BSC9_MAG*, BSC10_MAG*, BSC1_MAG1*, MY_MAG*

- Voltage monitoring: L0:PEM-LVEA_V1 vs. H0:PEM-LVEA2_V1

- Several weather channels named differently at LHO vs. LLO

Sampling rates were consistent, with only a few exceptions

- L1:LSC-ETMX_CAL sampled at 16384 Hz; all other *_CAL channels
  sampled at 2048 Hz

- As of E4, certain H1 channels had sampling rates different from H2
  and L1 (H1:IOO-MC_L; OSEM sensors)

Comparisons of Channel Means and Variances

Picked four 10-second times during E4 (one per day) while locked
Calculated mean and variance in 37 overlapping 1-second sections
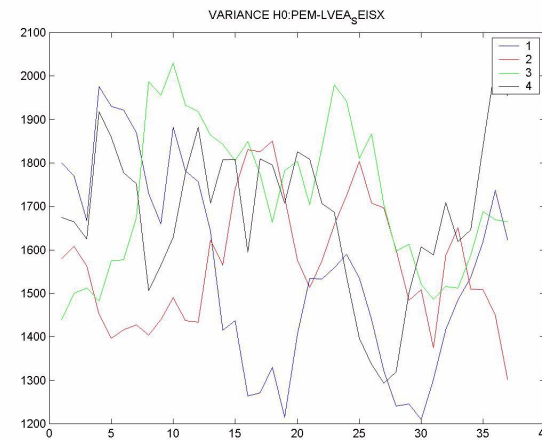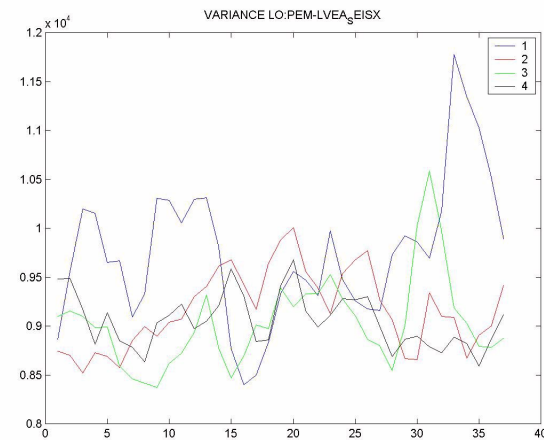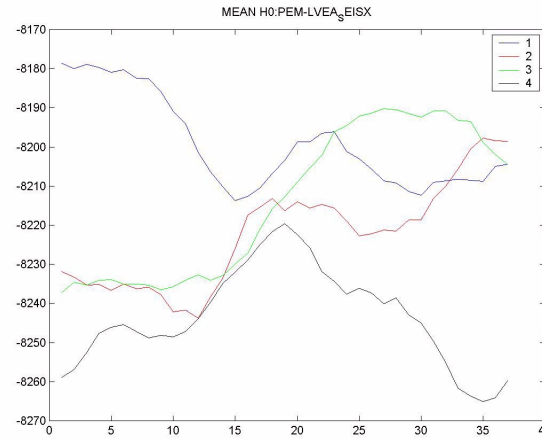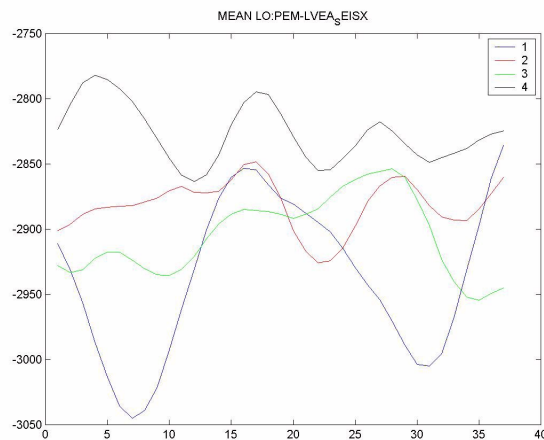
(Done for a few selected channels)

Notes

It's hard to conclude much from comparisons of mean and variance

It would be very nice to have a description of each channel, including the physical location of PEM sensors