E2 Data Integrity Investigation

John G. Zweizig
LIGO, Caltech

LSC collaboration meeting
Baton Rouge, March 14-17, 2001

LIGO-G010161-00-E

Overview

- Purpose of Investigation
  - Verify that all data are successfully collected and recorded
  - Exercise and verify data monitoring infrastructure.
  - Look for errors / anomalies in the data stream
  - Verify RDS and Archived data can be read back
- Integrity Checks
  - BitTest - Check channels for stuck words/bits
  - Slice2 - Look for repeated segments.
  - RDS tabulation
  - Archived Data tabulation
  - Data comparison
  - RDS re-run

Channel Monitor ($BitTest$)

Function

- Check ALL raw data (e.g. channels read from adcs)
- Look for stuck bits (always 0 / always 1)
- Look for repeated data words.
- Flag readout errors.

Results

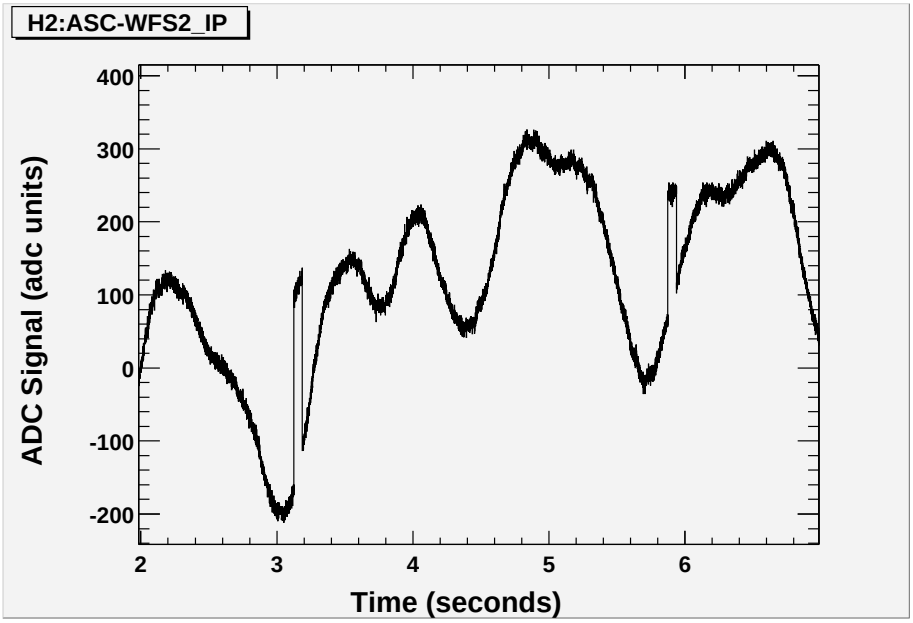
- 274 raw data channels checked continuously.
- Channel report produced every 20 minutes
- Operators check results 1-2 times per shift
- Some error messages - nothing significant.
- Majority of error repeat counts, often due to data overflows.

Typical BitTest Error Summary

```
Report produced for 1200 seconds to November 9, 2000 9:04:07

============================== Channels with errors

Error(s) in channel: H0:GDS-EX_TO1
Longest repeat group (173) excedes limit (100)

Error(s) in channel: H2:LSC-ETMX_CAL
Longest repeat group (108) excedes limit (100)

Error(s) in channel: H2:LSC-ETMY_CAL
Longest repeat group (574) excedes limit (200)

Error(s) in channel: H0:PEM-NBR_2K
Longest repeat group (157) excedes limit (100)

Error(s) in channel: H2:PSL-ISS_MC2_F
Longest repeat group (133) excedes limit (128)

============================== Channels with all one value

H2:PSL-PMC_RFPDDC_F H2:PSL-PMC_ERR_F

============================== Channels with error flags

============================== Channels not read out
```

Misplaced Slice Test

- Purpose: Look for repeated data segments.
- Why? Errors seen in E1 run:



- Settings: 62.5ms slices, <= 0.25s delay, > 1 channel.
- Results: accidentals only (dominated by ∼constant channels).

Archived Data Statistics

DAQ Run	First File Start Time	Last File Start Time	Number of Files	Missing Frames
1	657747696	657795496	240	923
2	657796619	657872019	378	7
3	657872226	657873826	9	50
4	657874076	657874076	1	34
5	657874310	657894110	100	1626
6	657895936	657932136	182	8
7	657932344	657933544	7	83
8	657933827	657934027	2	7
9	657934234	658038234	521	107
10	658038541	658046341	40	85
11	658046626	658165026	593	83
12	658165309	658308909	719	348
13	658309457	658323857	73	82
14	658324139	658353539	148	0
Total (%run)	657747696	658353739	3013 99.4%	3443 0.6%

Note: Data archived in HPSS as tar files named frame_archive/LHO/full/E2/H-< $start-time$ >.F.n200.tar

RDS Statistics

RDS Run[1]	Start GPS	Stop GPS	Frames Recorded	Frames Lost	Inter-Run Gap
1	657748786	657795671	46875	10	1007
2	657796678	658084277	287164	435	0
3	658084277	658196365	111841	247	343
4	658196708	658254979	58271	12	172
5[2]	658255151	658353665	98020	494	0
Total	657748786	658353665	602161	1198	1522
(%)			99.55%	0.20%	0.25%

Notes:
1) RDS Runs delimited by restart of RDS writer.
2) Run 5 consists of data from one-arm running.

Reduced Data Set Rerun

- Goals:
  - Read ALL RDS data frames
  - Regenerate triggers, trends, etc with consistent configurations.
  - Verify trigger logging mechanism
- Procedure:
  - Distribute all E2 RDS data via Shared memory mechanism.
  - Data replayed at $\sim$ 3.4 times faster than real-time.
  - Set shared memory manager to insure that all clients receive all data.
  - Run monitor processes:

BitTest	blrms_monitor	glitchMon
LockLoss	PSLmon	ServoMon
Slice2		

RDS Rerun Results

2-Arm running: 504,139 2-arm data frames were read, distributed to all monitors on 12/8/00 - 12/10/00.

Process	# Triggers[1]	Rate
BitTest	8177	16 mHz
EarthQuake	110	-
LockLoss	4217	8 mHz
PSLmon	193650	383 mHz
ServoMon	1559	3 mHz
Slice2	3764	7 mHz
Total	211477	418 mHz

1-Arm running: 98,020 1-arm data frames were read, distributed to all monitors on 12/19/00.

Process	# Triggers[1]	Rate
BitTest	160	1.632 mHz
EarthQuake	30	0.306 mHz
LockLoss	26	0.265 mHz
PSLmon	30715	313.4 mHz
ServoMon	2	0.020 mHz
Slice2	62	0.633 mHz
Total	31350	319.8 mHz

Notes:
1) (103 + 2693) triggers were not correctly entered into the meta-Database due to process table synchronization issues.

Comparison of RDS and Archive

- Purpose:
  - Verify that DMT is receiving the same data as the frame-builder/LDAS.
  - Make sure not data are being lost or modified in Frame generation, copying, archival, etc.
  - Verify that data can be retrieved from archive.
- Method:
  - Spot check data from different points during run.
  - Full archive and RDS frames read in with DMT Data accessor (Dacc class, based on FrameCPP).
  - Compare data from all common channels in RDS and Archive.
- Results:
  - Two sets of 10 seconds of data were selected, GPS 657799950-657799959 and GPS 657799950-657799959
  - Data compared in all 20 frames for 255 (261 for later data) channels.
  - No differences were found.