Avoiding the Tower of Babel Syndrome:

An integrated issue-based quality assurance system

Gene Van Buren

Brookhaven National Laboratory
STAR Experiment

Levente Hajdu, Jerome Lauret, and Michael DePhillips (BNL)
Lanny Ray (University of Texas - Austin)

Outline

- **Primary Goal: ensure effective and efficient progress towards achieving physics-quality data**

- **How do we achieve this:**

  - **Identification of issues in data through Quality Assurance**

    - Identification occurs at first opportunity possible

      - Staged QA:
        Online, initial reconstruction ("Fast Offline"), Production

  } **Browsers**

  - **Reporting and archiving of issues**

  - **One-stop-shopping for access to issues / QA information**

  } **Electronic Logs**

QA Stages: Goals

- **Online**
  - Immediate assessment of hardware functionality and performance for experiment operators (shift crew)
  - Trouble-shooting aid for other run-time issues

- **Fast Offline**
  - Short time-scale assessment on general validity of the data (can it be reconstructed?)
  - Feedback to experiment shift crew on effectiveness of operations / conditions

- **Production**
  - Ensure use of proper reconstruction chain
  - Validation of calibrations

Online

- **Browsing:**
  - **Histograms of raw (unreconstructed) information from detectors in data stream**
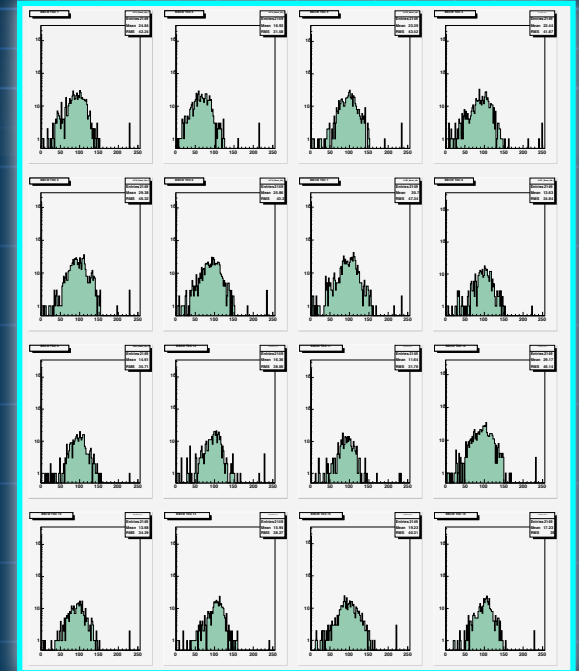 - **Use full events, but only a small fraction of all events**
    - Use online Event Pool
    - Continual updating (per run basis)
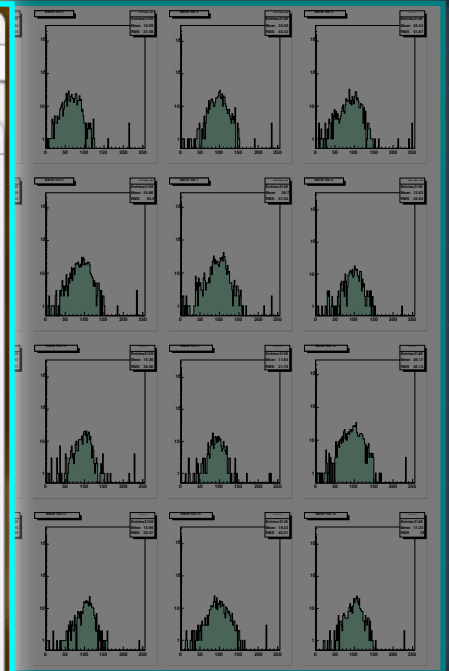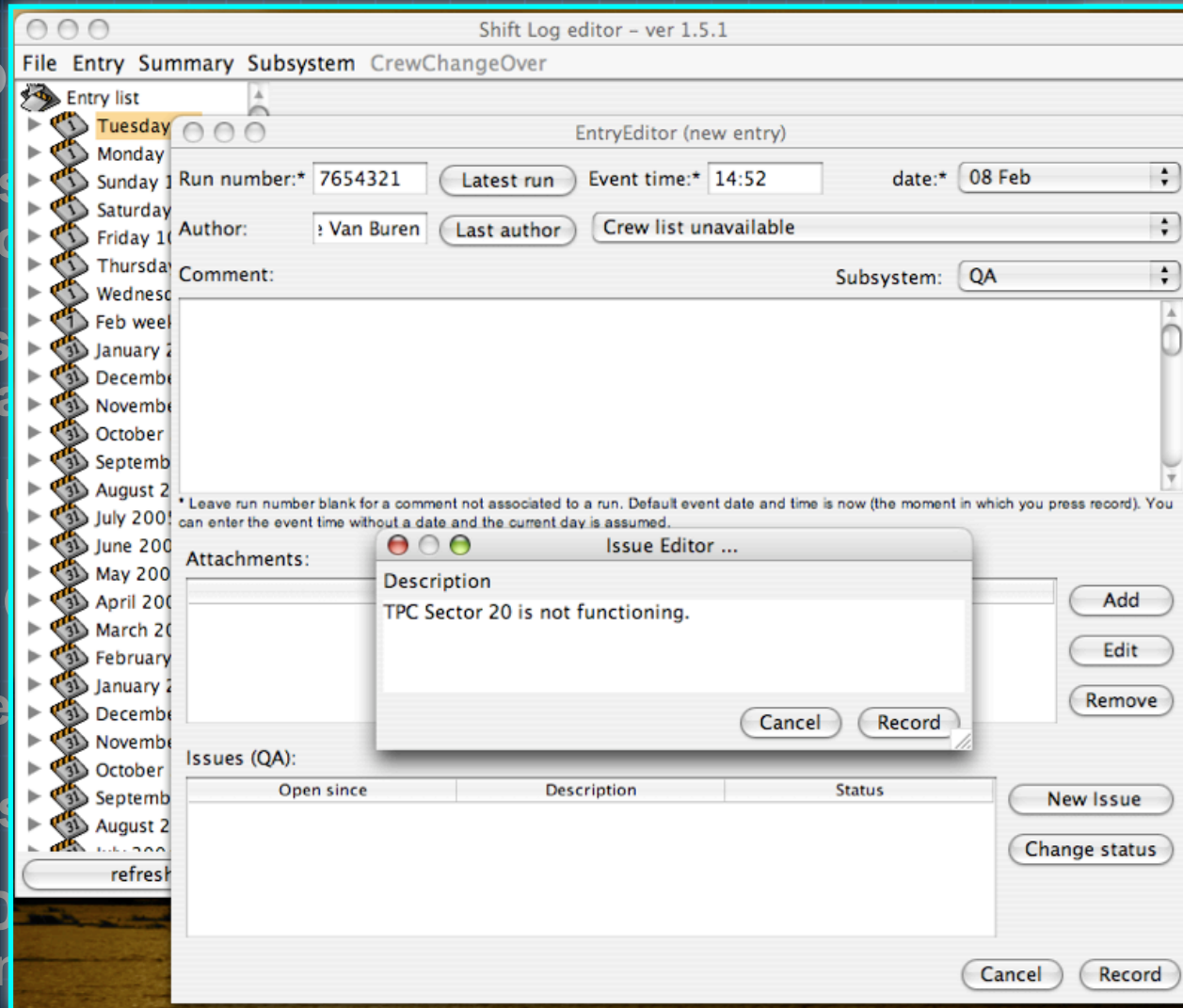 - **Reference histograms are available for comparison**
  - **Histograms from each DAQ run are kept on disk**
  - **PDFs of histograms are saved into database if the run's data is considered "good" (operator's discretion)**
- **Reporting: issues are entered directly into Electronic ShiftLog**

Online

Offline:
Fast Offline & Production

- **Browsing:**
  - **Histograms of reconstructed information** (hits, tracks, physics...)
    - Fast Offline: not necessarily well-calibrated
  - **Use full events, but only a small fraction of all files analyzed**
    - Fast Offline: only a fraction of all files get processed
    - Histograms generated (and kept) for each DAQ file processed (per file basis)
  - **QA done *on-demand***
    - PostScript/PDF of histograms generated then and there
- **Reporting: dedicated mechanism...**

Offline Reporting

- **Distinction between data types:**
  Fast Offline, Production, Simulations, Nightly Tests

- **Reporting mechanism is web-based**
  - Offline QA shifts can be performed from anywhere over internet

- **Users establish sessions for working on reports**
  - Flexibility to leave and come back, or even switch machines
  - Multiple simultaneous users

- **Reports contain multiple "data entries" for each file examined**

- **Report content is focused on issues...**

- **Report summary goes into ESL!!!**

QA Shift Report Form: Real Data Production

Add one entry for each QA job examined!

Run ID: [____] (usually a 7 digit number) (Lookup) other QA Reports for this Run
File Sequence number: [____] (leave as NA if none available)
Production Job ID: [NA]
Production job status: ○OK ○Crashed
Number of events in this file: [____]
Number of events with reconstructed primary vertex: [____]
QA job status: ○OK ○Crashed

If you would like to enter some additional comments beyond what is described by the **Active Issues** below, please do so here.

[_____]

(Save/Continue) (Reset)

Active Issues for this Data Entry:

(Remove) StE soft...   (Edit/Examine)
(Remove) FTPCW...   (Edit/Examine)
(Remove) StEMB...   (Edit/Examine)
(Remove) primary... , 19A) (Edit/Examine)

Menu
Session   Gene1
View Current Contents
Re-Enter Shift Info
Add A Data Entry For...
Fast Offline Data
Real Data Production
Real Data Nightly Test
Monte Carlo Data Production
Monte Carlo Data Nightly Test
Finish & Submit
Open Issue Browser/Editor
Open Report Archive
Erase Session & Start Over

Further down, options to add new issues or currently inactive issues
||
V

Offline Issues Management

- **May be a continuation from an online issue**

- **Tagged by an ID# and a brief description**

- **Each file examined is tied to its issues**
  - Issues may persist for several runs, or appear sporadically
  - Existing issues are listed when filing a QA Report
  - By default, active issues stay active, inactive stay inactive

- **Each issue can be amended with comments, edited for correctness, and eventually closed/resolved using an Issue Browser/Editor**

- **Changes to status of issues** (i.e. newly active or newly inactive) **are monitored for Report Summaries**

Offline Issues Management

QA Issue Browser/Editor

Issue ID: 1102 [Lookup]

Opened/Created: **12:44 04/19/05 EDT**
Last used/modified: **16:30 04/26/05 EDT**
Currently allowed types:
Last used in a Real Data Production entry: **16:30 04/26/05 EDT**

Name (short description): **No V0 and Xi vertices**
Full Description And Notes:

No vertices found for V0 and Xi (only for Central and High Tower

12:19 04/20/05 EDT :
run 5080031-4040041 : No vertices found for V0 and Xi in Central

06:07 04/21/05 EDT :
run 5079014-2010001 : No vertices found for V0 and Xi in Central

06:51 04/25/05 EDT :
run 5065074-1080001 : No vertices found for V0 and Xi in Central

02:37 04/26/05 EDT :
run 5073001-2040030 : No vertices found for V0 and Xi in Central

[Edit Descriptions and Notes] (Please use only for correcting errors!)

Note to add (or **resolution** if closing):

[   ]

[Add Note] [Close/Resolve]
Allow issue to be used for data entries of: [----------] [Allow Type]

[Open/Create New Issue]

- **May be a** ... **sue**
- **Tagged** ...
- **Each file** ...
  - Issues ... **poradically**
  - Existing ...
  - By defau... **inactive**
- **Each iss** ... **ents, edited for corre** ... **esolved using an** ...
- **Changes** ... **tive or newly inactive) a** ... **aries**

Offline Report Archives

- At the conclusion of a shift, the QA crew submits their Report

- Report summary is generated which details changed issues (new/gone)

  - Limiting the information in the summary to what is most significant is important in communicating effectively!!!

  - Emailed to operations shift crew for immediate perusal

  - Submitted into Electronic ShiftLog

- Full Report is archived, where it is indexed by filing time, and by the runs examined

Report Archives

○ **Indexed...**

   ○ **...by runs examined**

   ○ **...by when reported (report numbers)**

QA Reports for Run 6161047

Lookup report by any of:

- By Run:
  - Run Year: 6   Day: 161   Number: 6161047
  - Run Number: [   ] Lookup

- By Report: OFFLINE
  - Report Year: Select
  - Report Number: [   ] Lookup

Reports available for file sequences:

(Numbers in parentheses refer to multiple reports for one file sequence.)

- **2090001** (1)   Report_2005_06_0667.html   [FAST OFFLINE]
- **2090001** (2)   Report_2005_06_0667.html   [FAST OFFLINE]

Report Archives

○ **Indexed…**

○ **…by runs examined**

○ **…by when reported (report numbers)**

Data Report for Fast Offline Data:

```
Run ID:                                      6161047
File Sequence number:                        2090001
Production Job ID:                           NA
Production job status (OK or crashed?):      ok
Number of events in this file:               197
Number of events with reconstructed
    primary vertex:                          192
QA job status (OK or crashed?):              ok
```

Issues:

- [ID:1126] StEHT point: z
- [ID:1082] EmcCat4 delta no data, 6A to 6C
- [ID:1085] TPC: padrow dist. of hits has anamolous spikes
- [ID:1075] SVT phi distribution of hits very jagged 2F
- [ID:1070] FTPC: total charge in E << W
- [ID:1031] FTPC charge step flat to 200
- [ID:1096] FTPC cluster radial position: anomalous spikes
- [ID:1135] # hits in FTPC, East vs. West (1C)
- [ID:1018] FTPC hits saturate some anodes
- [ID:1161] SVT z distribution of hits very jagged 2E
- [ID:1127] TPC hit phi distribution
- [ID:1143] no SVT laser spot (1E)
- [ID:1016] Hot TPC Anodes
- [ID:1162] no FTPC West track information!

6161047

Number: 6161047

Reports available for file sequences:

(Numbers in parentheses refer to multiple reports for one file sequence.)

- **2090001** (1) Report_2005_06_0667.html   [FAST OFFLINE]
- **2090001** (2) Report_2005_06_0667.html   [FAST OFFLINE]

Information Hub: ESL

The Electronic ShiftLog (ESL) serves as a repository of...

- **Summaries of each operation shift**

- **Brief descriptions of purpose and Online QA for each run**

- **Fast Offline QA Report summaries**

- **Links to RunLog for every run**
  - Detailed information about the run (time, conditions, active detectors, event counts by file and trigger)
  - Subsequent links to Online QA plots and Offline QA reports

- **All other operations comments**
  - Tagged by subdetector if appropriate

Information Hub: ESL

STAR Shift log

Sat - Apr 09, 2005
Go to bottom

Sun - Apr 10, 2005

Mon - Apr 11, 2005

Sunday - April 10, 2005

00:29 (00:29) DetOps	**Summary Report** - Night Shift No useful beam until midnight. At midnight MCR called to say that they would deliver a useful beam but it was aborted to due Quench Link Interlock. Alexei tried to fix the laser. See his message. Multiple issues with canbus when he was doing that. The anode appeared to have turn in auto-ramp after a reboot. They had to be brought down from full voltage. FTPC crate 71 shows bad voltage reading. Red alarm. Ignore. Still in full control of the FTPC. Interlock alarm around midnight. DAQ connection. Cleared itself (called B. Christie). - Fabrice Retiere
03:38 DetOps	Run 6100005 - running DAQ/TRG while bringing up other detectors running up trip on TPC anode BBC and: 35000 Yellow back: 8500 Blue back:9100 BBC and/(Yellow back + Blue back) = 2.0 - Howard Wieman
03:41 DetOps	Run 6100007 - Finally a ppProductionMinBias SVT and BSMD not included a short run for Akio - Howard Wieman
04:00	When powering on the SVT HV the system after ramping up to 791V automatically rumped down to 154V and wents into an infinite loop unfinnishing thus the HV ramp up. I have stopped the sequence from experts pannel and ther I have found that HV maximum was

Gene Van Buren - BNL - STAR   **CHEP '06 - Mumbai**   15

RunLog

ESL Browses/Searches

- **Browse by time period, online QA entries, subsystem entries, or some combinations**

- **Full text searches**

Full text search ID: 1162

Submit   Reset

Previous Month	**June 2005**		Next Month			
Sun	**Mon**	**Tues**	**Wed**	**Thurs**	**Fri**	**Sat**
			1 (152)	2 (153)	3 (154)	4 (155)
5 (156)	6 (157)	7 (158)	8 (159)	9 (160)	10 (161)	11 (162)
12 (163)	13 (164)	14 (165)	15 (166)	16 (167)	17 (168)	18 (169)
19 (170)	20 (171)	21 (172)	22 (173)	23 (174)	24 (175)	25 (176)
26 (177)	27 (178)	28 (179)	29 (180)	30 (181)		

period:

(mm/dd/yyyy)

Start date : _____

(mm/dd/yyyy)

End date: _____

Submit   Reset

Subsystem:

Include all entries tagged with:

- ☑ General
- ☐ BBC
- ☐ BEMC
- ☐ DAQ
- ☐ DetOps
- ☐ EEMC
- ☐ FPD
- ☐ FTPC
- ☐ L3
- ☐ PMD
- ☐ SVT
- ☐ SSD
- ☐ TPC
- ☐ TOC
- ☐ TRG
- ☐ QA

STAR Shift log

Searching "ID:1162". Entries sorted by relevance.

Go to bottom

Saturday - June 11, 2005

| 21:16 (00:00) QA | STAR QA Shift Report Summary (FAST OFFLINE)

SUMMARY OF CHANGED ISSUES FOR Fast Offline Data (+ new / - gone):

+ [ID:1162] no FTPC West track information!
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1162

+ [ID:1158] Difference in x-y position of primary vertex from FTPC and TPC
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1158

SUMMARY OF RUNS / FILE SEQUENCES EXAMINED FOR Fast Offline Data
6161046 / 2090001
6161047 / 2090001 |

...

Sunday - June 12, 2005

| 19:03 (00:00) QA | STAR QA Shift Report Summary (FAST OFFLINE)

SUMMARY OF CHANGED ISSUES FOR Fast Offline Data (+ new / - gone):

+ [ID:1157] TPC: total charge in E < W
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1157

+ [ID:1163] Difference between z position of primary vertices by SVT and TPC
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1163

+ [ID:1084] No SVT data
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1084

- [ID:1162] no FTPC West track information!
http://www.star.bnl.gov/STAR/comp/qa/QAShiftReport/issueEditor.php?iid=1162 |

...

Issue came...
...and went

Summary

- ⦿ **QA issues are identified and reported shortly after the data is available**
  - ◯ QA is staged to provide increased degree of inspection with increased extent of data reconstruction
- ⦿ **Tracking of issue evolution is achieved through management of issues as entities at all stages**
  - ◯ Also provide ability to associate issues with specific datasets
- ⦿ **Electronic ShiftLog collects QA information along with other operation details - "hub of QA"**
  - ◯ Online reports and offline summaries are stored directly in ESL
  - ◯ Links to RunLog provide access to further online/run details
  - ◯ Links to full offline reports and issue histories provide more indepth details of issues noted offline
  - ◯ Good browse/search tools make navigation reasonable (feasible!)