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)
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

- Reporting and archiving of issues
- One-stop-shopping for access to issues / QA information

Browsers

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**
Reporting: issues are entered directly into **Electronic ShiftLog**
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*

- **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!!!
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

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

No V0 and Xi vertices

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

.run 5080031-4040041 : No vertices found for V0 and Xi in Central

.run 5079014-2010001 : No vertices found for V0 and Xi in Central

.run 5065074-1080001 : No vertices found for V0 and Xi in Central

.run 5073001-2040030 : No vertices found for V0 and Xi in Central

Note to add (or resolution if closing):

Add Note Close/Resolve

Allow issue to be used for data entries of: 

Open/Create New Issue

Gene Van Buren - BNL - STAR

CHEP '06 - Mumbai
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:  

- By Report: ONLINE  
  - Report Year:  
  - Report Number:  

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

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] SteEHT point: z
- [ID:1082] FmCc4 delta no data_6A to 6C
- [ID:1085] TPC: padrow dist. of hits has anomolous 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!

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**  
**Sun - Apr 10, 2005**  
**Mon - Apr 11, 2005**

**Sunday - April 10, 2005**

<table>
<thead>
<tr>
<th>Time</th>
<th>Description</th>
</tr>
</thead>
</table>
| **00:29 (00:29)** | **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 due to 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). |
| **03:38** | **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 |
| **03:41** | **Run 6100007** - Finally a ppProductionMinBias  
SVT and BSMD not included  
a short run for Akio |
| **04:00** | When powering on the SVT HV the system after ramping up to 791V automatically ramped down to 154V and went into an infinite loop unfinishing thus the HV ramp up. I have stopped the sequence from experts carrel and then I have found that HV maximum was |
RunLog

Run Period: CuCu @ 62 GeV

**Setup:**
- tpc, svt, tof, emc, ftd, ftdc, ssc, pmd, ssd, eemc, fdp2, bsm, emc

**MagF:**
- all

**DAQType:**
- phys, ped, laser
- pulser
- RTS
- ShiftLeader

Filter Bad: RTS

**Files (Daq):**
- tpc, sv, tof, emc, ftdc, pmd, ssd, eemc, bsm

**Files (Scaler):**
- emc, bsm, esmc

**Detectors:**
- tpc, sv, tof, emc, ftdc, pmd, ssd, eemc, bsm, emc

**Star Magnet:**
- 4,506.4 Amps, Polarity A Full Field

**QA Info:**
- 6080050.pdf

**DAQ Info:**
- DAQ Rates

**RHIC Summary:**

<table>
<thead>
<tr>
<th>RHIC Ring</th>
<th>Species</th>
<th>Energy</th>
<th>Ions E09</th>
<th>Fill-Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>Cu</td>
<td>31.064</td>
<td>120.59</td>
<td>6,531</td>
</tr>
<tr>
<td>Yellow</td>
<td>Cu</td>
<td>31.064</td>
<td>117.52</td>
<td>6,531</td>
</tr>
</tbody>
</table>

**Trigger Summary:**

| ClockFrequency [Start / End] | 9.382 / 9.382 MHz |
| TIER1 File Name             | trg_050111.bin   |

<table>
<thead>
<tr>
<th>Name</th>
<th>bit (daqTrgld)</th>
<th>offlineTrgld</th>
<th>prescal</th>
<th>NumberOfEvents</th>
</tr>
</thead>
<tbody>
<tr>
<td>cu62-zdc</td>
<td>1</td>
<td>76001</td>
<td>400.0</td>
<td>314</td>
</tr>
<tr>
<td>cu62-zdc-tacs</td>
<td>10</td>
<td>75002</td>
<td>100.0</td>
<td>1123</td>
</tr>
<tr>
<td>cu62-zdc-narrow</td>
<td>100</td>
<td>75007</td>
<td>1.0</td>
<td>72105</td>
</tr>
<tr>
<td>cu62-bbc-by</td>
<td>1000</td>
<td>75012</td>
<td>500.0</td>
<td>244</td>
</tr>
<tr>
<td>cu62-bbc-narrow</td>
<td>10000</td>
<td>75011</td>
<td>1.0</td>
<td>72233</td>
</tr>
<tr>
<td>cu62-zerobias</td>
<td>100000</td>
<td>76300</td>
<td>1.0</td>
<td>1317</td>
</tr>
</tbody>
</table>

See poster by W. Betts
ESL Browses/Searches

- Browse by time period, online QA entries, subsystem entries, or some combinations
- Full text searches

Tick QA option.
### STAR Shift log

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

Go to bottom

Saturday - June 11, 2005

<table>
<thead>
<tr>
<th>Time</th>
<th>Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>21:16</td>
<td>STAR QA Shift Report Summary (FAST OFFLINE)</td>
</tr>
<tr>
<td></td>
<td>SUMMARY OF CHANGED ISSUES FOR Fast Offline Data (+ new / - gone):</td>
</tr>
<tr>
<td></td>
<td>+ [ID:1162] no FTPC West track information!</td>
</tr>
<tr>
<td></td>
<td>+ [ID:1158] Difference in x-y position of primary vertex from FTPC and TPC</td>
</tr>
<tr>
<td></td>
<td>SUMMARY OF RUNS / FILE SEQUENCES EXAMINED FOR Fast Offline Data</td>
</tr>
<tr>
<td></td>
<td>6161046 / 2090001</td>
</tr>
<tr>
<td></td>
<td>6161047 / 2090001</td>
</tr>
</tbody>
</table>

Sunday - June 12, 2005

<table>
<thead>
<tr>
<th>Time</th>
<th>Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>19:03</td>
<td>STAR QA Shift Report Summary (FAST OFFLINE)</td>
</tr>
<tr>
<td></td>
<td>SUMMARY OF CHANGED ISSUES FOR Fast Offline Data (+ new / - gone):</td>
</tr>
<tr>
<td></td>
<td>+ [ID:1157] TPC: total charge in E &lt; W</td>
</tr>
<tr>
<td></td>
<td>+ [ID:1163] Difference between z position of primary vertices by SVT and TPC</td>
</tr>
<tr>
<td></td>
<td>+ [ID:1084] No SVT data</td>
</tr>
<tr>
<td></td>
<td>- [ID:1162] no FTPC West track information!</td>
</tr>
</tbody>
</table>

---

Issue came... 

...and went
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 in-depth details of issues noted offline. Good browse/search tools make navigation reasonable (feasible!).