

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: QA

- **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

- Communication: reporting and archiving of issues

- Integrated access: one-stop-shopping for retrieval of 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

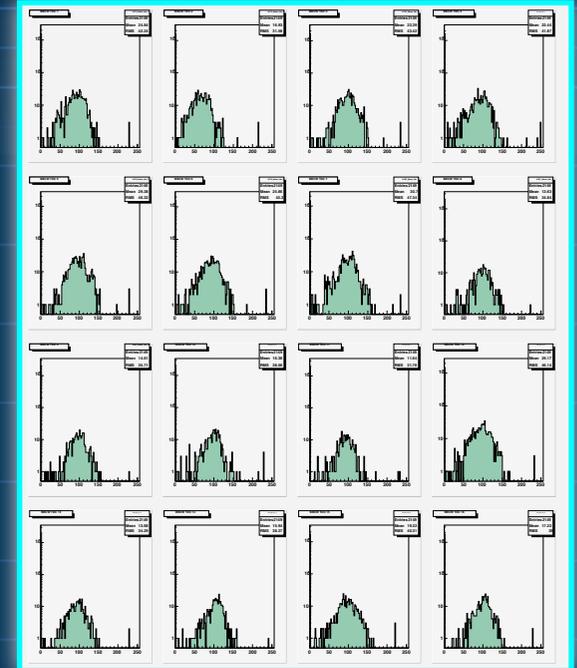
- Continual updating (per run basis)

- 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**



Online

Shift Log editor - ver 1.5.1

File Entry Summary Subsystem CrewChangeOver

Entry list

Tuesday

Monday

Sunday

Saturday

Friday

Thursday

Wednesday

Feb week

January

December

November

October

September

August

July

June

May

April

March

February

January

December

November

October

September

August

July

refresh

EntryEditor (new entry)

Run number:* 7654321 Latest run Event time:* 14:52 date:* 08 Feb

Author: Van Buren Last author Crew list unavailable

Comment:

Subsystem: QA

* Leave run number blank for a comment not associated to a run. Default event date and time is now (the moment in which you press record). You can enter the event time without a date and the current time will be used.

Attachments:

Description

TPC Sector 20 is not functioning.

Add

Edit

Remove

Cancel Record

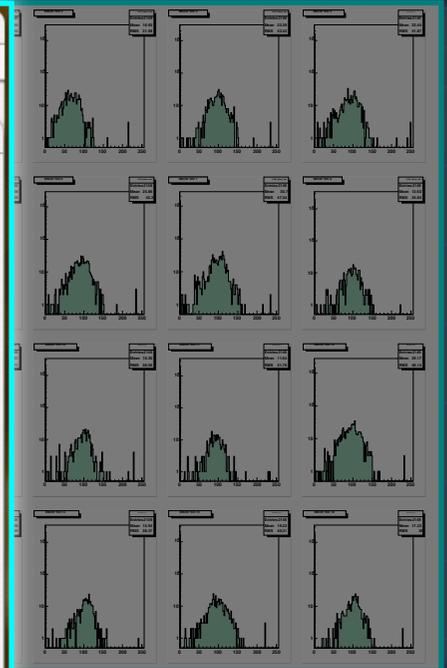
Issues (QA):

Open since	Description	Status
------------	-------------	--------

New Issue

Change status

Cancel Record



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

Electronic ShiftLog (ESL)

STAR Shift log

[Sun - Jan 16, 2005](#)

Mon - Jan 17, 2005

[Tue - Jan 18, 2005](#)

[Go to bottom](#)

Monday - January 17, 2005

00:15 (00:15)	43 Bunch Store. injected almost 3/4 hour ago, backgrounds still bad. (BBC bkg Y+B) /ZDC ~ 2. ZDC ~ 27K. not yet marked for physics.
General	- Dave Underwood
00:28 (00:29)	Summary Report - Night Shift
General	The only store started before this shift. It was 37 bunches instead of 28, and the BBC bkg looked terrible.(BBC B+Y bkg)/ZDC around 2.5 to 3. I don't know if this was completely real. I called MCR and found that the unbunched beam was 10×10^9 , a
	- David Underwood
	<pre>run number status nEvents glbSetupName ----- ----- ----- ----- 6016047 good 99744 cuProductionMinBias 6016048 bad 0 pedestal_rhicclock 6017002 1 pedestal_rhicclock</pre>
00:48	Run 6017004 - took over shift with physics run (6017004) just started. As Dave noted, backgrounds are much higher than we've had so far (B+Y bg)~57, (ZDC and)~22. MCR is aware. Also, running is about 20% slower (~70Hz)-is this the addition of SSD? In data are tp
General	- Evan Finch
00:49	Run 6017004 - (last comment got cut off)-In data are tpc,ftpc,ssd,btow,etow,esmd,pmd
General	- Evan Finch
00:54	Run 6017004 - FTPC had an anode trip-didn't end run for this.
FTPC	- Evan Finch
01:41	Run 6017006 - Run stopped after 83K events for TPC anode trip
General	- Evan Finch

Offline:

Fast Offline & Production

○ Browsing:

- Histograms of reconstructed information (hits, tracks, vertices...)

- 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 by the offline browser

○ 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!!!**

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

QA Shift Report Form: Real Data Production

Add one entry for each QA job examined!

Run ID: (usually a 7 digit number) other QA Reports for this Run

File Sequence number: (leave as NA if none available)

Production Job ID:

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.

Active Issues for this Data Entry:

 StE soft FTPCW StEMB p primary

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



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
 - List of existing issues and their status provided to users 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:

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

(Please use only for correcting errors!)

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

Allow issue to be used for data entries of:

- May be a
- Tagged
- Each file
- Issues m
- List of e
- filing a Q
- By defau
- Each iss
- for corre
- using an
- Changes
- (inactive) a

sue
n
poradically
to users when
y inactive
ents, edited
esolved
tive or newly
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: Day: Number:
- Run Number:

- By Report:

- 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:

- [\[ID:1126\] StEHT point: z](#)
- [\[ID:1082\] EmcCat4 delta no data, 6A to 6C](#)
- [\[ID:1085\] TPC: padrow dist. of hits has anomalous 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

○ Indexed...

○ ...by runs examined

○ ...by when reported (report numbers)

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

RunLog

Run Period: **CuCu @ 62GeV** Select Reset

Setup: **all** MagF: **all** DAQType: phys ped laser Filter Bad: RTS ShiftLeader

tpc svt tof emc fpd ftpc l3 pmd ssd eemc fpd2 bsmc esmd

STAR Run5

- Mar.20-Mar.26
 - Sunday,20 79
 - Monday,21 80
 - 6080006, Physics
 - 6080009, Physics
 - 6080010, Physics
 - 6080011, Physics
 - 6080013, Physics
 - 6080014, Physics
 - 6080015, Physics
 - 6080016, Physics
 - 6080017, Physics
 - 6080018, Physics
 - 6080020, Physics
 - 6080021, Physics
 - 6080028, Physics
 - 6080041, Physics
 - 6080043, Physics
 - 6080046, Physics
 - 6080047, Physics
 - 6080049, Physics
 - 6080050, Physics
 - 6080052, Physics
 - Tuesday,22 81
- Mar.13-Mar.19
- Mar.6-Mar.12

Run Number 6080050
Start Time Mon Mar 21 23:01:20 EST [2005-03-22 04:01:20 GMT]
Stop Time Mon Mar 21 23:24:24 EST [2005-03-22 04:24:24 GMT]
Completion Status Successful
GlbSetup/DaqRunType cu62productionMinBias / Physics
Events 99026
Files (Daq): 161
Files (scaler):
Detectors tpc svt tof emc ftpc pmd ssd eemc bsmc esmd
Star Magnet 4,506.4 Amps Polarity A Full Field
QA Info : [6080050.pdf](#) [Check Offline QA](#)
DAQ Info : [DAQ Rates](#) [Trigger details](#)

RHIC Summary

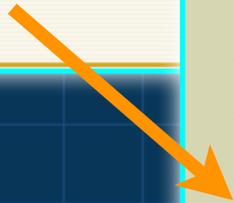
RHIC Ring	Species	Energy	Ions E09	Fill-Number
Blue	Cu	31.064	120.59	6,531
Yellow	Cu	31.064	117.52	6,531

Trigger Summary

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

name	bit(daqTrgId)	onlineT	rescale	numberOfEvents
cu62-zdc	1	76001	400.0	314
cu62-zdc-tacs	10	76002	100.0	1123
cu62-zdc-narrow	100	76007	1.0	72105
cu62-bbc-by	1000	76012	500.0	244
cu62-bbc-narrow	10000	76011	1.0	72233
cu62-zerobias	100000	76300	1.0	1317

22:39	Run 6080048
General	
23:01	Run 6080049
General	
23:23	Run 6080050
General	



See poster by W. Betts

ESL Browsers/Searches

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

Full text search ID:

period:

(mm/dd/yyyy)

Start date :

(mm/dd/yyyy)

End date:

Subsystem:

Include all entries tagged with:

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

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)			

ESL Browses/Searches

STAR Shift log

Entries from Jan 16, 2005 to Jan 17, 2005

[Go to bottom](#)

Sunday - January 16, 2005

08:06 SSD	One of the channels in SSD is not holding voltage Module 8, at the top (neither N nor P, sort of in the middle), we brought the voltages down as instructed.	- Manuel Calderon
--------------	--	-------------------

Monday - January 17, 2005

12:12 SSD	Run 6017053 - SSD is dominating the deadtime (~80%), and we're running with 80Hz. The Panitkin plots for SSD seem to show a high occupancy, which could be a problem. Jerome B. is looking into it.	- Manuel Calderon
12:19 SSD	Run 6017054 - Pedestal run for SSD, see if it helps with occupancy.	- Manuel Calderon
12:37 SSD	Run 6017057 - Running with tpc emc ftpc pmd ssd eemc esmd, rate is 85 Hz. SSD looks better now after pedestal run, the occupancy was reduced. Jerome will try to see if he can figure out why.	- Manuel Calderon
13:13 SSD	Ladder 8, which has no data on it, might have been responsible for the occupancy problem in ssd. The bias on this ladder was removed after the pedestal run (6017040), which might have affected the pedestals, even though there's no data from this ladder.	- Manuel Calderon

[Go to top](#)

Issue topic
started...
...and
followed up



STAR Shift log

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

[Go to bottom](#)

Saturday - June 11, 2005

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