Analysis of Failed Runs in IU Mini Data Challenge

Kei Moriya
Indiana University
GlueX Offline Meeting
November 13, 2013

Brief Summary

- Processed 50M bggen events - 10 hrs worth of data at $10^7\gamma$/s
- 10k events x 5000 files
- All events had primary vertex smeared (smear_thrown_vertex), E_γ range 8.4 - 9.0 GeV, no EM background added
- Using 192 nodes on IU cluster, can process 2000 files/day = 20M events/day (hdgeant, mcsmear, REST)
- 4739 files had usable REST output (94.8%)
- 161 files got stuck at REST, remaining 100 finished REST but had unusable output (more detailed logs available)

WHY?

Re-submit Failed Jobs

- First time processed 5000 runs and had the 192-cluster node full
- Resubmitted only the 261 failed jobs (at once)
- Did not update repository in between (same code)
- Several possible issues:
  - persistent problem of GlueX code stopping (have seen for years)
  - track reconstruction enters inifinite loop (introduced recently?)
  - multiple jobs simultaneously accessing databases
  - memory boundary issues

Results of Re-submission

today's topics

Re-run result	REST stopped	REST unusable
success	89	84
stopped	72	1
small file size	0	15
TOTAL	**161**	**100**

- Out of 161 runs that stopped before, 72 stopped, none had small file size
- Out of 100 unusable runs, 1 stopped, and 15 had small file size
- The two problems seem rather isolated from each other
- For runs that stopped both times, final event being processed is always same
- For runs that finished REST but had small file sizes both times, final event being processed is always same
- Out of 261 failed initial runs, able to reproduce 72 + 15 = 87 runs (33.3%)

Runs where REST stopped both times

- 161 runs in initial try where REST stopped
- Jobs will usually end within 2 hours, but for these runs REST gets stuck at same event and job is killed by hand
- 72 runs where this was reproduced in re-run
- Look at runs where the failure was fairly early:

run number	final event
0626	34
1303	198
2447	194
3182	70
4316	158
4361	13

Run 0626, event 34

- REST stopped at event 34
- Tried running REST again, does stop at this event
- Used hddm_cull_events to pull out event 34 only from original smeared file
- Running hd_dump on this single event hangs when using option -A (if no factory to dump out is specified, will exit) able to isolate single bad event!!
- For 98 possible factories, look at hd_dump output using hd_dump -f -p -D[factory name]
- hd_dump stops at DTrackWireBased, DTrackTimeBased, DChargedTrack, DNeutralParticle, DNeutralParticleHypothesis, DNeutralShower, DVertexIndependentResults, DVertex, DParticleSet, DPhysicsEvent, DPiPlus, DPiMinus, DParticleCombo, DMCThrownMatching

Bad Factories in run 0626, event 34

- Event is $\pi^0 \pi^0 \pi^+ \omega$ n, $\omega \rightarrow \pi^+ \pi^- \pi^0$

charged tracks

Event: 34
DMCThrown:

PID:	Name:	q:	x(cm):	y(cm):	z(cm):	E(GeV):	t(ns):	p(GeV/c):	theta(deg):	phi(deg):	pdgtype:	myid:	parentid:	mech:	geantId:
0	Unknown	+0	-0.6	0.2	63.0	6.1988	0.000	6.190	4.033	27.223	2	0	0	12	1543941296
0	Unknown	+0	-0.6	0.2	63.0	0.6385	0.000	0.269	66.335	92.306	2101	0	0	11	2
0	Unknown	+0	-0.6	0.2	63.0	2.3933	0.000	2.370	10.584	-152.777	-2	0	0	12	1543986432
0	Unknown	+0	-0.6	0.2	63.0	0.4246	0.000	0.267	66.999	-87.694	2	0	0	11	0
0	Unknown	+0	-0.6	0.2	63.0	2.8180	0.000	2.503	13.475	-130.295	91	0	0	11	2
7	Pi0	+0	-0.6	0.2	63.0	0.4870	0.000	0.468	49.868	31.818	111	1	0	1	1543941696
7	Pi0	+0	-0.6	0.2	63.0	2.3309	0.000	2.327	23.566	-137.079	111	2	0	1	0
0	Unknown	+0	-0.6	0.2	63.0	6.8373	0.000	6.309	5.305	49.705	92	0	0	11	2
8	Pi+	+1	-0.6	0.2	63.0	0.1789	0.000	0.112	122.015	-103.316	211	3	0	1	1543917952
33	omega	+0	-0.6	0.2	63.0	1.7376	0.000	1.552	2.287	26.611	223	4	0	1	0
13	Neutron	+0	-0.6	0.2	63.0	4.9207	0.000	4.830	7.311	55.996	2112	5	0	1	2
1	Photon	+0	6.0	10.1	156.0	0.1438	0.000	0.144	42.818	7.699	22	28	12	1497449284	0
1	Photon	+0	6.0	10.1	156.0	0.1422	0.000	0.142	20.051	132.962	22	29	12	1497449284	2
1	Photon	+0	6.0	10.1	156.0	0.4305	0.000	0.430	63.299	67.094	22	81	10	1497449284	1544001040
1	Photon	+0	6.0	10.1	156.0	0.0375	0.000	0.037	123.299	91.130	22	82	10	1497449284	0
1	Photon	+0	6.0	10.1	156.0	0.0808	0.000	0.081	55.096	-52.660	22	84	9	1497449284	2
1	Photon	+0	6.0	10.1	156.0	0.1709	0.000	0.171	105.084	-104.527	22	85	9	1497449284	1544004368
1	Photon	+0	6.0	10.2	156.3	0.5701	0.000	0.570	42.825	-102.357	22	198	122	1497449284	2
1	Photon	+0	6.0	10.2	156.3	0.0096	0.000	0.010	144.349	-3.471	22	199	122	1497449284	1544007312
9	Pi-	-1	-0.6	0.2	63.0	0.2487	-0.000	0.206	56.529	164.371	-211	206	4	1497449284	0
8	Pi+	+1	-0.6	0.2	63.0	0.5983	-0.000	0.582	16.000	-11.031	211	207	4	1497449284	2
7	Pi0	+0	-0.6	0.2	63.0	0.8908	-0.000	0.881	4.200	10.875	111	208	4	1497449284	1544010640
1	Photon	+0	-0.6	0.2	63.0	0.3301	-0.000	0.330	7.708	155.012	22	209	208	1497449284	0
1	Photon	+0	-0.6	0.2	63.0	0.5607	-0.000	0.561	10.655	-3.615	22	210	208	1497449284	2
5	Muon+	+1	5.0	-16.3	50.7	0.1303	0.000	0.076	128.924	-63.047	-13	232	3	1497449284	1544013968
4	Neutrino	+0	5.0	-16.3	50.7	0.0432	0.000	0.043	102.978	3.284	121416	231	3	1497449284	0
2	Positron	+1	23.7	-8.4	17.5	0.0256	0.000	0.026	75.429	-63.688	-11	233	232	1497449284	2
4	Neutrino	+0	23.7	-8.4	17.5	0.0456	0.000	0.046	136.067	160.316	121416	231	232	1497449284	1543917312
4	Neutrino	+0	23.7	-8.4	17.5	0.0344	0.000	0.034	39.912	31.641	121416	228	232	1497449284	0
1	Photon	+0	-0.6	0.2	63.0	1.1895	-0.000	1.190	26.114	-132.263	22	240	2	1497449284	2
1	Photon	+0	-0.6	0.2	63.0	1.1414	-0.000	1.141	21.104	-143.218	22	241	2	1497449284	1543942336
1	Photon	+0	-0.6	0.2	63.0	0.3231	-0.000	0.323	38.543	32.129	22	254	1	1497449284	0
1	Photon	+0	-0.6	0.2	63.0	0.1639	-0.000	0.164	72.643	31.416	22	255	1	1497449284	0

- First factory to fail: DTrackWireBased
- Factory DFDCHit has 2169 entries, contrast to 393 entries in event 32

Run on isolated single event: run0626, event 34

- Isolate single event 34 from bggen file
- Run hdgeant, mcsmear, REST processing on this
- REST finishes in this case - due to random numbers within hdgeant not being offset in the same way as the crashing case (all decay momenta are different)

need "buildup" by other events to reproduce crash

- I can run over the single-event REST file with analysis processors, but when I open up the analysis tree ROOT crashes

some other issue?

- Running hd_dump -A on the REST file does NOT stop

Run 4361, event 13

- REST stopped at event 13
- Tried running REST again, does stop at this event
- Used hddm_cull_events to pull out event 13 only from original smeared file
- REST does NOT stop for this event! can run $K^+\Lambda$ processor and can create ROOT trees, no crash
- Used hddm_cull_events to pull out first 15 events from original smeared file
- REST DOES stop now! unable to isolate single bad event!!
- Running hd_dump on these 15 events will hang at event 13 when using option -A

Isolated Event Files

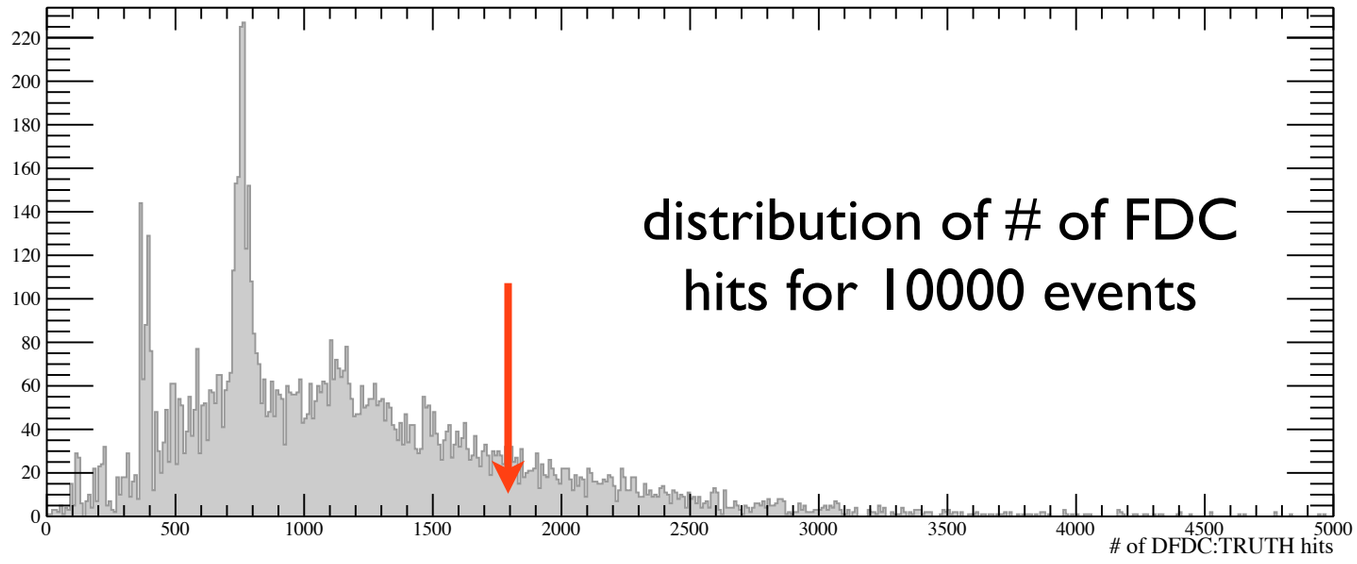
- For 21 runs, I have files that isolate the final event or all events up to that event. Files for 12 smaller runs at http://dustbunny.physics.indiana.edu/~kmoriya/2013-11-12-failed_hddm_files/
- If I knew how to change how the GlueX ticker shows, I could isolate more files, but 21 runs should be good for now
- For 14 runs, running hd_dump -A stops
- For 4 runs, hd_dump finishes
- For 2 runs, the single event file finishes, but the consecutive events file stops
- For 1 run, the single file stops, but the consecutive events file finishes

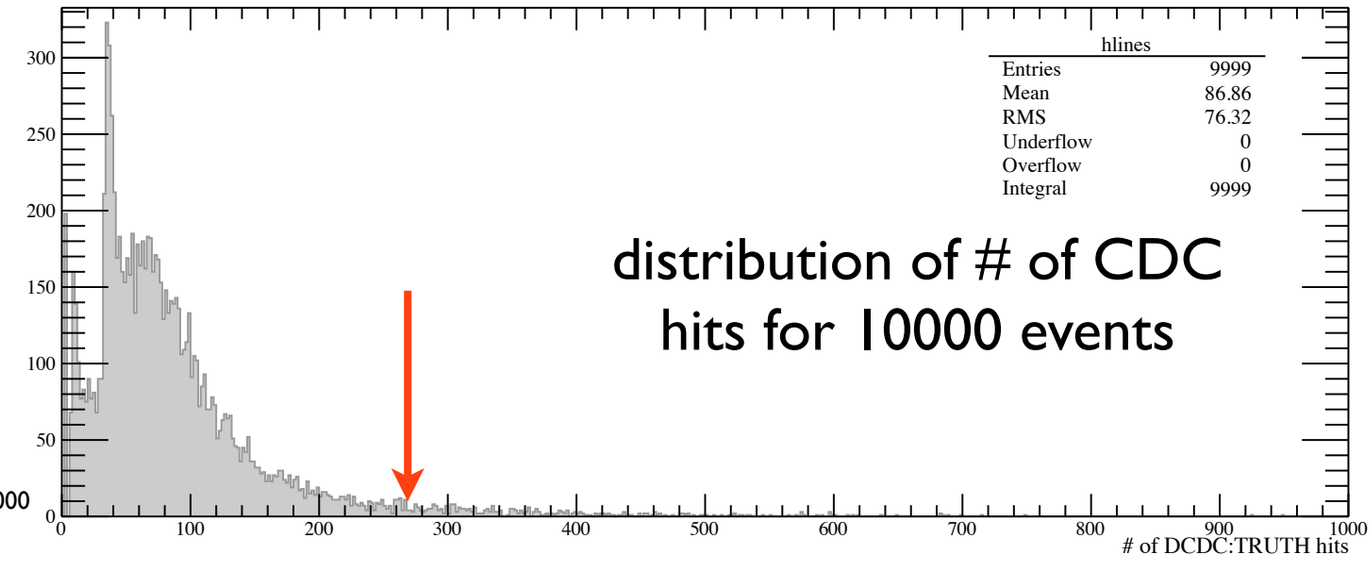
DMCThrown for 21 bad events

- 11 events have ω in them (leads to low momentum track pairs?)
- 1 event with Λ, 1 event with Σ^0 - the probability of finding either within 21 events is extremely low
- Are low-momentum opposite-charge pairs from ω and Λ events causing these crashes?

DCHit:TRUTH for each event

- Truth information is always available
- For 21 bad events, average of ~1700 FDC hits
- If I run over 10,000 events from a good run, we get1137 hits / event.
- For CDC, average of 270 hits
- Average for 10,000 good events is 87 hits/ event





used 2013-11-08 run 0000 for this study

Second re-run

	1st re-run		2nd re-run	
Re-run result	REST stopped	REST unusable	REST stopped	REST unusable
success	89	84	92	82
stopped	72	1	69	3
small file size	0	15	0	15
TOTAL	**161**	**100**	**161**	**100**

161 REST Failed

100 small file size

consistently bad

bad only the first time

- 60 files of stopx3
- 78 files of stop→good→good
- 9 files of stop→good→stop
- 12 files stop→stop→good
- 2 files of stop→20MB→19MB

- 15 files of small→small→small
- 80 files of stop→good→good
- 2 files of small→20MB→20MB
- 1 file of small→stop→stop
- 2 files of small→good→stop

majority of failures due to database access?

What happened to small size files?

- For 100 small filesize runs, 84 show that mcsmear processed 10k events
- For these 84 files, REST processed 10k events
- For 16 runs where REST did not process 10k events, the number of events processed is the same as where mcsmear ended
- For 16 runs where mcsmear did not process 10k events, 5 runs (1971,2863,3754,4574,4897) had following message:
- 15 runs where small file size was reproducible are from these 16 runs

```
JANA ERROR>> Thread 0 hasn't responded in 30 seconds. (run:event=2:5505) Cancelling ...
JANA ERROR>> Caught HUP signal for thread 0x2ad90c534700 thread exiting...
JANA ERROR>>
JANA ERROR>> Last thread to lock output file mutex: 0x2ad90c534700
JANA ERROR>> Attempting to unlock mutex to avoid deadlock.
JANA ERROR>> However, the output file is likely corrupt at
JANA ERROR>> this point and the process should be restarted ...
JANA ERROR>>
JANA ERROR>>
JANA ERROR>> Automatic relaunching of threads is disabled. If you wish to
JANA ERROR>> have the program relaunch a replacement thread when a stalled
JANA ERROR>> one is killed, set the JANA:MAX_RELAUNCH_THREADS configuration
JANA ERROR>> parameter to a value greater than zero. E.g.:
JANA ERROR>>
JANA ERROR>>     jana -PJANA:MAX_RELAUNCH_THREADS=10
JANA ERROR>>
JANA ERROR>> The program will quit now.
JANA >>
JANA >>Telling all threads to quit ...
JANA >>Merging thread 0 ...

Closed HDDM file
 5504 event written to bggen_8.4_9.0-2013-10-22_1971_hdgeant_smeared.hddm
JANA >>Merging event reader thread ...
JANA >> 5504 events processed  (5515 events read) Average rate: 5.6Hz
```

Why did hdgeant/mcsmear fail?

- For the 15 files where mcsmear processed less than 10k events, 10 hdgeant logs had the following message
- These runs are the ones that did not have the HUP signal from mcsmear

```
      6000  events simulated

!!!!! ZFATAL called from MZPUSH

!!!!! ZFATAL reached from MZPUSH    for Case=  3

       IQUEST(11) =    1502105          16EB99
       IQUEST(12) =  ********          454E494B   KINE
       IQUEST(13) =      64000           FA00
       IQUEST(14) =      64000           FA00
       IQUEST(15) =          1              1
       IQUEST(16) =          0              0
       IQUEST(17) =         10              A
       IQUEST(18) =          0              0

       Current Store number =  0  (JQDIVI= 2)
1ZEBRA SYSTEM Post-Mortem from ZPOSTM.

 /QUEST/
```

MZPUSH - function to alter the size of a bank →memory issues?

no other runs (even ones that stopped) had ZFATAL error

Reproducible runs with less than 10k events either come from ZFATAL called in hdgeant, or HUP called in mcsmear

Could the unreproducible bad runs be from database access?

$b_1\pi$ events

- Check if $b_1\pi$ events, which have $b_1 \rightarrow \omega\pi$ in them, has a high failure rate
- Generate 200 files of 10k events each
- genr8 input file is one in repository
- No events with hdgeant ZFATAL, 2 files with mcsmear HUP

channel	total	stop	small
bggen	5000	161(3.22%)	100 (2%)
$b_1\pi$	200	6 (3%)	11 (5.5%)

Summary

- Starting to delve into why some jobs fail
- Overall failure reproducibility is 29% ((60 + 10+5) / 261)
- For 60 runs where REST stops, many ω, Λ events - oppositely charged track pairs causing crash? Large number of CDC, FDC hits.
- Consistent hdgeant failures for 10 runs is due to call of ZFATAL from MZPUSH, mcsmear failures for 5 runs are call of HUP
- Are majority of unreproducible failures due to database access?
- It would be nice if I could confirm these failed runs on an independent machine (do these problems depend on architecture?). Provided scripts and auxiliary files (see generation.README.txt)
  http://dustbunny.physics.indiana.edu/~kmoriya/2013-11-12-generation/
- Have provided hddm files with these events (see failed.README.txt)
  http://dustbunny.physics.indiana.edu/~kmoriya/2013-11-12-failed_hddm_files/
- I can provide scripts that specify configuration at each stage if people are interested - you only need the bggen configuration file and the hdgeant control.in file

Backup

Run 4316, event 158

- REST stopped at event 158
- Same as run 4361, event 13: running on this isolated single event will NOT cause programs to not stop
- Running hd_dump on the first 158 events will hang