SciSoft - Support #22991

Rogue events in dual phase simulation

07/25/2019 09:03 AM - Heidi Schellman

Status:	Feedback	Start date:	07/25/2019
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour

Description

Ken Herner has identified the following event as running for a very long time (and then terminating correctly). This is dual-phase protoDUNE simulation. It's an example of a reasonably common problem.

You can set this up by doing

ssh to dunegpvm1X.fnal.gov # (may not need this)

source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh
setup -B dunetpc v08_03_00 -q debug:e17

cd /dune/data/users/${USER}

lar -c rawhitfinding_dune10ktdphase_workspace4x2.fcl -n1 --nskip=302
root://fndca1.fnal.gov:1094/pnfs/fnal.gov/usr/dune/tape_backed/dunepro/protodune-dp/raw/2019/mc/cosmics/DC3/00/00/11/84/1184_1_a_dc.cosmics

and then wait a few hrs.

If you can point us to how you approach this issue we may be able to get people who can follow up in future.

Thanks

Heidi

History

#1 - 07/25/2019 09:12 AM - Kyle Knoepfel

- Tracker changed from Bug to Support

Thanks, Heidi. We'll take a look.

#2 - 07/25/2019 09:29 AM - Kyle Knoepfel

- Status changed from New to Feedback

Heidi, I'm unable to use XRootD to access the file. I get the following error:

```
secgsi_InitProxy: cannot access private key file: /nashome/k/knoepfel/.globus/userkey.pem
```

which indicates that I'm not part of the DUNE VO. Is it possible to copy the file to a locally accessible directory from dunegpvm0X?

#3 - 07/25/2019 11:02 AM - Thomas Junk

If you have setup dunetpc, type

```
setup_fnal_security
```

to get your x509 certificate and grid proxy for interactive use. Then try running lar.

#4 - 07/25/2019 11:16 AM - Thomas Junk

Oh, that voms_proxy_init in the setup_fnal_security script (in duneutil) will fail if you are not a member of the DUNE VO. We make DUNE VO membership part of the onboarding script for collaborators, but you may need special dispensation as a computing professional who's not otherwise on DUNE? If you'd like to join DUNE, you can talk to the IB reps -- I belive Gina Rameik and Alberto Marchionni are the reps. But it involves a

contribution to the common fund, and if you're just providing services like this, maybe that's overkill. You could submit a Service Desk ticket to be added to the DUNE VO.

#5 - 07/25/2019 11:32 AM - Thomas Junk

Just running that job in the debugger on slf7, I notice that it spends a lot of time in MINUIT.

/cvmfs/larsoft.opensciencegrid.org/products/larreco/v08_02_01/source/larreco/HitFinder/DPRawHitFinder_module.cc

has a method hit::DPRawHitFinder::CreateFitFunction that makes a funtion as a string that is made into a TF1. ROOT will interpret string-functions more slowly than a compiled function. That won't make execution time infinite, just slower than it has to be.

#6 - 07/25/2019 11:36 AM - Kyle Knoepfel

Tom, you've identified the correct module. The ARM forge map profiler shows that almost all of the time is spent in calling the FitExponentials function, which, in turn, invokes ROOT's fitting routines. The practical consequence is that in order to avoid processing these types of events, either LArSoft will need to be changed, or DUNE will need to place a filter before the DPRawHitFinder module that can query certain features of the event and fail the event based on the relevant features, before hit-finding commences.

#7 - 07/25/2019 11:38 AM - Kyle Knoepfel

- File Screen Shot 2019-07-25 at 11.37.30 AM.png added

See attached image.

#8 - 07/25/2019 12:36 PM - Thomas Junk

Yup, I've been running it for the last 50 minutes using the ARM forge debugger, and I poked the value fLogLevel to 6 from 0, and it now prints out progress as it goes along. It seems to frequently fit over 20 peaks at a time, and these fits take a few seconds to complete, only for the module to discover that it has to add another peak to the fitting function. There does seem to be a cap on this so it doesn't just end up fitting infinite numbers of peaks, but it sure is slow.

This event may have an EM shower on it, just a hunch, if this module works better on other events. If we skip around events that might trigger this problem, we'd probably bias some physics. Might be better just to limit the number of peaks in a fit, and divide the waveform into shorter ROI's with fewer peaks per ROI if it starts getting too populated. I'll keep it running for a while and see if it gets truly stuck.

#9 - 07/25/2019 01:21 PM - Thomas Junk

Still making "progress". I looked at the event in the event display:

```
lar -c evd_protodunedp.fcl root://fndca1.fnal.gov:1094/pnfs/fnal.gov/usr/dune/tape_backed/dunepro/protodune-dp
/raw/2019/mc/cosmics/DC3/00/00/11/84/1184_1_a_dc.cosmics
```

and selected the "MC Truth" checkbox after navigating to event 10293. Seems unremarkable. Some channels have >500 sim::IDE's on them, one per tick per particle, and some waveforms do look filled up for 500 to 1000 ticks at a time. Looks like the hit finder is not handling these gracefully.

Files

Screen Shot 2019-07-25 at 11.37.30 AM.png	153 KB	07/25/2019	Kyle Knoepfel