Web based portal for DESC simulation studies a.k.a. "ImSim on Demand" (a pre-implementation discussion)

Tony Johnson, SLAC tonyj@slac.stanford.edu 5th December 2013

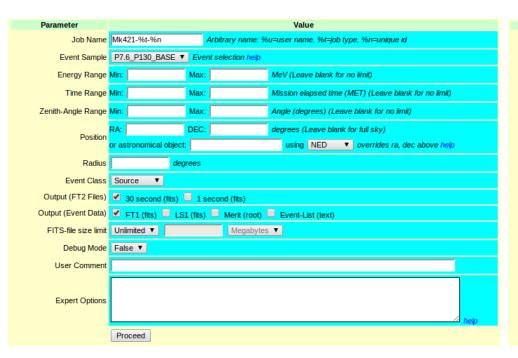
Contents

- Background
 - Fermi Gamma-Ray Space Telescope web tools
 - Skimmers
 - Pipeline
 - Data Catalog
 - Web based MC framework for EXO
- Potential simulation portal for DESC
 - Goals, potential implementation
 - Open questions

Fermi Gamma-Ray Space Telescope web tools

- Fermi developed an extensive web toolkit for data processing/quality monitoring/access
 - The toolkit is not Fermi specific and has been re-used by other experiments
- Skimmers allow web based data selection
 - Data selection is done in background by automated data processing system
 - Supports parallelization for efficient operation
 - User is notified by e-mail once data is ready
 - Data made available via web based data catalog

Fermi Data Skimmers



Parameter	Value		
	%u-%t-%n Arbitrary name: %u=user name, %t=job type, %n=unique id		
	/Data/Flight/Reprocess/P202/MERIT @ ▼		
Data Source			
TCut	CTBParticleType == 0 && CTBClassLevel > 1 && CTBBestEnergy > 20000 && abs(FT1B) < 2.5 && FT1ZenithTheta < 100		
Min Run Number			
Max Run Number			
Datacat Criteria	help		
	✓ Include all columns		
Tuple Columns	✓ Include Tkr* ✓ Include Vtx* ✓ Include Cal* ✓ Include Acd* □ Include Mc* ✓ Include CTB* ✓ Include FT1* ✓ Include Evt* ✓ Include Pt* □ Include Gt* □ Include Obf* □ Include Grb* □ Include FSw*		
	☐ Include Git* ☐ Include Obt* ☐ Include Grb* ☐ Include Fsw*		
Debug Mode	False ▼		
Max Root file size (MB)	200		
User Comment			
Expert Options	help		
	Proceed		

Web based MC framework for EXO

- Example of reuse of Fermi web toolkit
- Allows MC tasks to be configured via web
 - Supports hierarchy of tasks to make it easy to have many jobs which are similar but slightly different
 - Allows trial runs (limited number of events) for testing
 - Allows full task to be run for arbitrary number of events
 - Event count can be increased after initial run
 - Tasks are parallelized by automated processing pipeline
 - Generate summary plots for entire run

Web based MC framework for EXO

- Web interface shows collaborators what MC data is available, and exactly how it was created
- Provides access to data via data catalog
- For EXO creating new tasks is restricted to web czar
 - Prioritizes/evaluates requests
 - Combines similar requests where appropriate
 - Monitors and ensures runs are successful and that final plots look good

MC Web Interface screenshots

EXO Simulation Summary



MC Web Interface screenshots

Task P3_Source_S8_px257_py10

Properties

Edit

 Name
 P3_Source_S8_px257_py10

 Description
 Task cloned from P3_Source_S8_px257

Parameters

Edit

Name	Value	From task	Notation
DATACAT_DIR	/EXO/MC/Phase3c	Phase3_PARENT	
EXO_FILE	mc.exo	Phase3_PARENT	The .exo file used to run the mc job
EXOBASE	/nfs/slac/g/exo/software/hudson/builds-rhel5/svn-id/8274	Phase3_PARENT	The version of EXOAnalysis to use for simulation
MAC_FORBIDDEN	false	P3_Source_S8_px257_py10	
MAC_ISOTOPE	Th228	P3_Source_S8_px257_py10	
MAX_EVENTS_PER_JOB	40000	P3_Source_S8_px257_py10	The number of jobs in each parallel job
OUTPUT_DIR	/nfs/slac/g/exo_data2/exo_data/data/MC/Phase3a_svn7148/\${MC_TASK}	P3_Source_S8_px257_py10	
PIPELINE_TASK	EXOGenericMonteCarlo	Phase3_PARENT	The pipeline task to use
PrintModulo	500	P3_Source_S8_px257_py10	Diagnostic message rate
RECON_DRIFT_VELOCITY	.00171	Phase3_PARENT	/rec/drift_velocity_mm_per_ns
RUN_SIMULATION	BriansSimulation.py	Phase3_PARENT	The script to run on the main job
SIM_DRIFT_VELOCITY	0.171	Phase3_PARENT	/digitizer/driftVelocity
TOTAL_EVENTS_REQUESTED	4000000	P3_Source_S8_px257_py10	The total number of events requested

Templates

Add template

```
From task
mc.exo.template P3 Source S8 px257 pv10 Edit Rename Delete
                                                  load $EXOLIB/plugins/EXOGeant4Module.*
                                                  use exosim digitizer rec risetime uind wiregaindummy v-wiregaindummy cluster gridcorr toutput
                                                  /exosim/macro mc.mac
                                                  /exosim/SkipEmptyEvents true
                                                  /exosim/initial seed %(INITIALSEED)s
                                                  /exosim/run number %(PIPELINE STREAM)s
                                                  /exosim/SourcePosition 8 2.57 0.1 29.48
                                                  /digitizer/setDatabaseTime 1348000000
                                                  /digitizer/driftVelocity %(SIM DRIFT VELOCITY)s
                                                  /digitizer/collectionDriftVelocity .225
                                                  /digitizer/setDigitizationTime 2048 microsecond
                                                  /digitizer/setTriggerTime 1024 microsecond
                                                  /digitizer/LXeEnergyRes 0.0
                                                  /digitizer/wireNoise 800.
                                                  /digitizer/APDNoise 2000.
                                                  /digitizer/ElectronicsDBFlavor measured times
                                                  /cluster/drift velocity mm per ns %(RECON DRIFT VELOCITY)s
                                                  /rec/LowerFitBoundWire 40
                                                  /rec/UpperFitBoundWire 140
```

Potential simulation portal for DESC

- Evaluating the use of these tools for DESC simulation portal
 - Would allow users to request ImSim runs
 - Would use processing pipeline coupled with Panda to run at:
 - SLAC/Grid/NERSC/IN2P3/...
 - Would provide single place to find and access existing datasets
 - Perhaps provide skimmer for selecting interesting observations

Open Questions

- Would this be useful?
 - If so exactly what exact functionality would need to be accessible from the web interface
 - Are there simulation experts/physics group contacts who would be willing to work with us on this?
 - Would we need a MC czar to coordinate requests
 - Or could it be totally self-service