Data Processing Workflows at Scattering User Facilities

Joe Osborn

Computer Science and Mathematics Division, ORNL

8/14/20

Outline

- Two sections:
- ORNL
  - ORNL neutron scattering facilities
  - Workflow needs
  - Current implementation
- BNL
  - BNL Relativistic Heavy Ion Collider
  - Workflow needs
  - Current implementation
- Future plans

General Questions To Address

- Scattering facilities have thousands of users with many different scientific needs
  - How to design software to address these needs?
- Scattering facilities have rapidly growing data sets
  - How to process them in a timely fashion, given the various needs by users?
- Users desire ability to monitor data processing and analysis
  - How to keep software transparent to users while also hiding primary functionality?

ORNL Neutron Facilities

- ORNL is one of the premier neutron scattering research facilities in the world
- Two major user facilities

Spallation Neutron Source (SNS)

High Flux Isotope Reactor (HFIR)

Spallation Neutron Source

- SNS produces neutrons via pulsed proton bunches on a mercury target
- Spallation neutrons are guided down beam lines to ~20 instruments utilized by ~3000 users in physics, chemistry, biology, and materials science
- While resulting data is not enormous (O(10) GB), the broad variety of needs from users presents a challenge for workflows!

High Flux Isotope Reactor

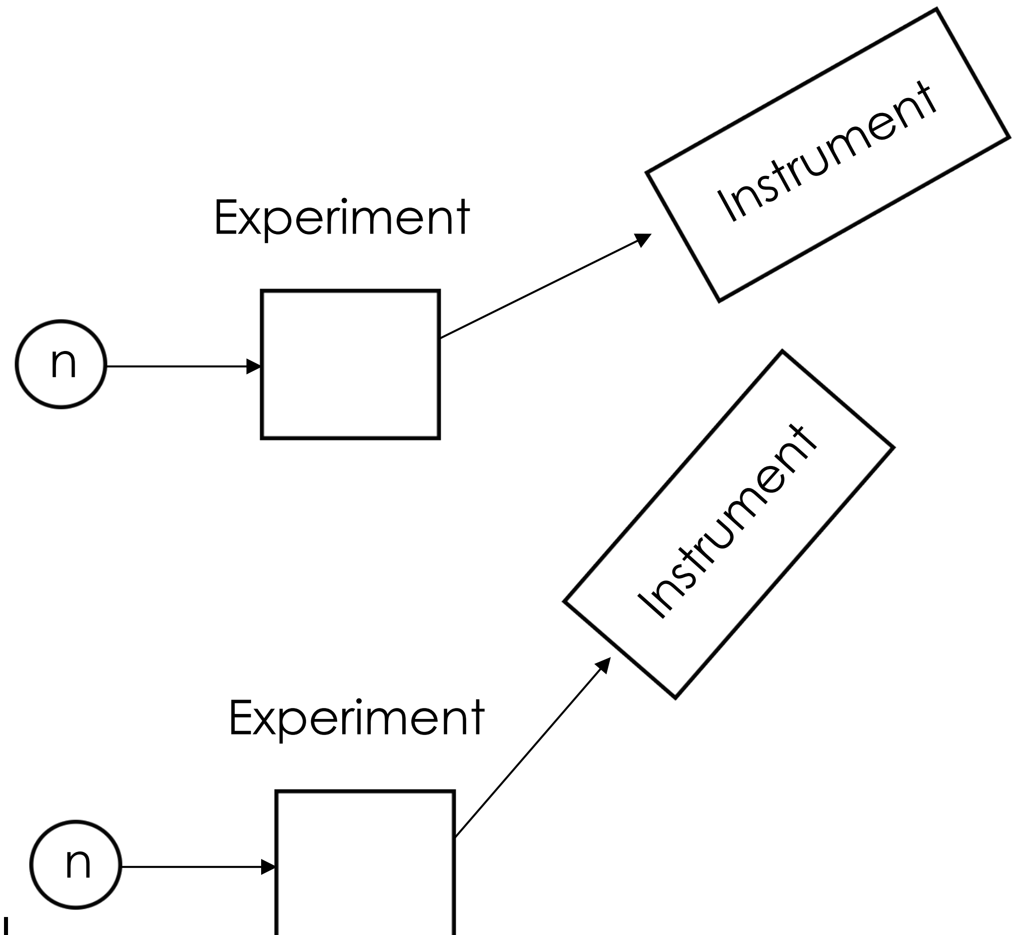
- HFIR produces neutrons via enriched U-235, where thermalized neutrons are ”beamed” out of the reactor to experiments
- HFIR has 12 instruments and is used by ~500 people annually to conduct physics, chemistry, biology, materials science and engineering
- While resulting data is not enormous (O(10) GB), the broad variety of needs from users presents a challenge for workflows!

Overview of Workflow

- Users bring their experiment to a particular instrument to collect data
- May collect a variety of data sets
  - Calibration
  - Background
  - Real data
  - ...
- Collect data in runs (intervals of time) given certain beam conditions, experiment conditions, etc.
- "Reduce" data to get instrument independent results to take home for further analysis

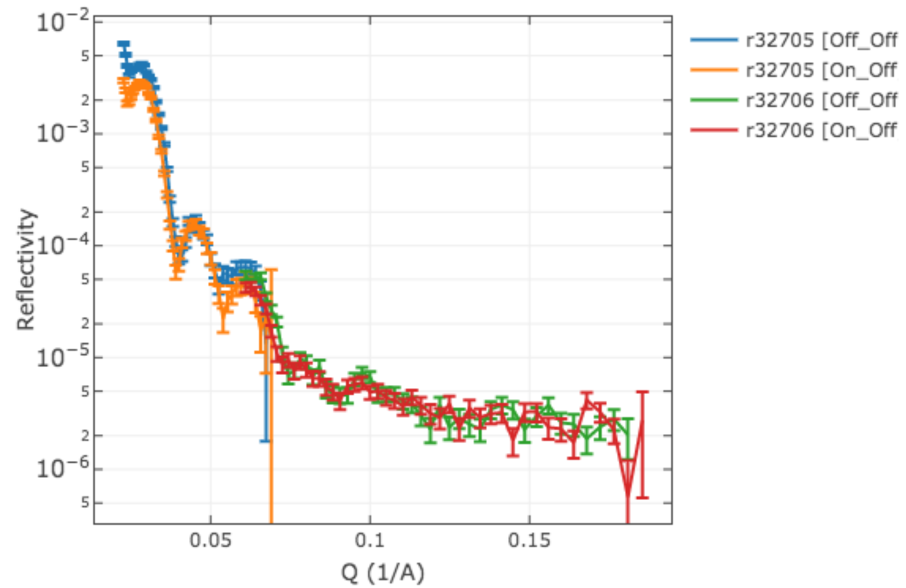
Overview of Workflow - Data Reduction

- Runs are collected at, for example, different scattering angles
- Collect the intensity of neutrons scattered at a given angle
- Correct for effects due to limited instrument capabilities
  - Reconstruction efficiencies
  - Geometry
  - …
- Combine all runs to get a final intensity as a function of scattering angle



Overview of Workflow - Data Reduction

- Runs are collected at, for example, different scattering angles
- Collect the intensity of neutrons scattered at a given angle
- Correct for effects due to limited instrument capabilities
  - Reconstruction efficiencies
  - Geometry
  - …
- Combine all runs to get a final intensity as a function of scattering angle



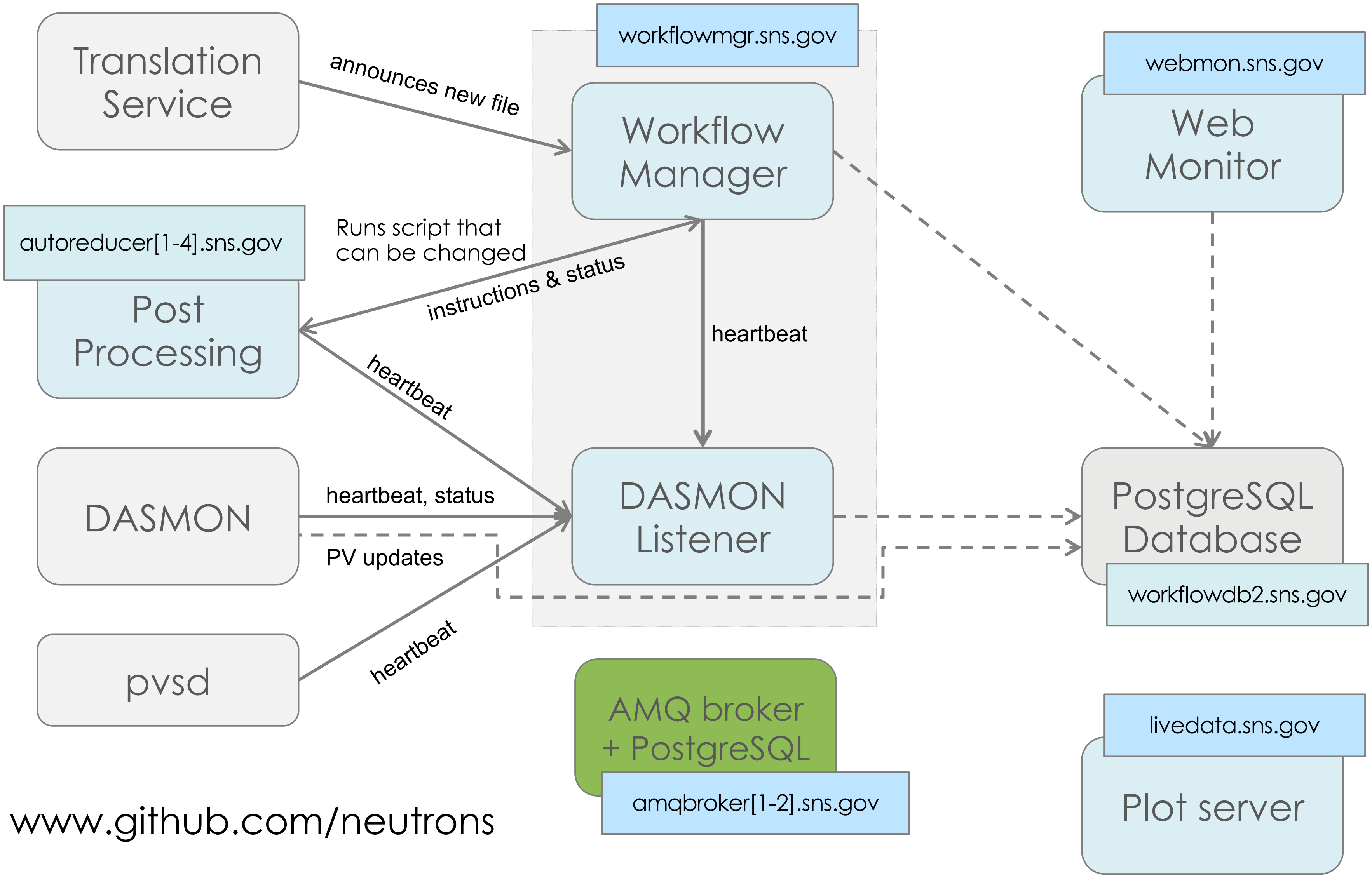
Workflow Needs

- Highly flexible (!)
  - Simple ways to tweak and/or change a workflow given users variable needs
- A way to batch set of runs together for collective processing
- Automated data reduction
  - However, would like ability to also manually control when needed
- Error reporting

OAK RIDGE National Laboratory

Current Workflow Implementation

Processing Architecture



www.github.com/neutrons

Workflow Manager

- State machine
- States and tasks are defined in a database
- All transactions that occur in the manager are logged
- Only message received is a file path!
- File path contains necessary (meta) data to process file

Web Monitor

- A live web monitor displays instrument status and data processing status
- Users utilize this heavily!

OAK RIDGE National Laboratory | HIGH FLUX ISOTOPE REACTOR | SPALLATION NEUTRON SOURCE

m2d | admin | logout

Instrument Status

home › dashboard — dashboard | extended dashboard | latest runs

Central systems: Workflow

List of instruments:

Instrument	Status	Instrument	Status
ARCS	Stopped	HYS	Stopped
BL0	-	MANDI	Recording
BSS	Recording	NOM	Recording
CG1D	-	NOW4	-
CG2	Stopped	PG3	Recording
CG3	Stopped	REF_L	Stopped
CNCS	Recording	REF_M	Recording
CORELLI	Recording	SEQ	Recording
EQSANS	Stopped	SNAP	Recording
HB2A	-	TOPAZ	Recording
HB2B	Recording	USANS	Recording
HB2C	Stopped	VIS	Stopped
HB3A	-	VULCAN	Recording

Web Monitor

- A live web monitor displays instrument status and data processing status
- Users utilize this heavily!
- Can search for run specific information, see status of reduction jobs, etc.

Align:P3_Emptybeam_02128020

Proposal: IPTS-23459 Run: 0

Status: Stopped Count rate: 0

Systems: Workflow

Last run: 32773 from IPTS-23459 created on Feb. 19, 2020, 12:20 p.m.

List of latest runs:

Show: 25 Search:

Run	Created on ▼	Status
32773	Feb. 19, 2020, 12:20 p.m.	complete
32772	Feb. 19, 2020, 11:49 a.m.	complete
32771	Feb. 19, 2020, 11:33 a.m.	complete
32770	Feb. 19, 2020, 10:35 a.m.	complete
32769	Feb. 19, 2020, 10:19 a.m.	complete
32768	Feb. 19, 2020, 9:42 a.m.	complete
32767	Feb. 19, 2020, 3:26 a.m.	complete
32766	Feb. 19, 2020, 2:54 a.m.	complete
32765	Feb. 19, 2020, 2:39 a.m.	complete
32764	Feb. 19, 2020, 1:41 a.m.	complete
32763	Feb. 19, 2020, 1:24 a.m.	complete
32762	Feb. 19, 2020, 12:48 a.m.	complete
32761	Feb. 18, 2020, 6:21 p.m.	complete
32760	Feb. 18, 2020, 5:49 p.m.	complete
32759	Feb. 18, 2020, 5:33 p.m.	complete
32758	Feb. 18, 2020, 4:35 p.m.	complete
32757	Feb. 18, 2020, 4:19 p.m.	complete
32756	Feb. 18, 2020, 3:46 p.m.	complete
32755	Feb. 18, 2020, 3:36 p.m.	complete
32754	Feb. 18, 2020, 9:18 a.m.	complete
32753	Feb. 18, 2020, 8:45 a.m.	complete
32752	Feb. 18, 2020, 8:29 a.m.	complete
32751	Feb. 18, 2020, 7:31 a.m.	complete
32750	Feb. 18, 2020, 7:14 a.m.	complete
32749	Feb. 18, 2020, 6:38 a.m.	complete

Run Status

- Status of a run of can be viewed by users for that experiment
- Data is pulled live from the database
- Allows users to see in real time the progress of their experiment and make possible changes while taking data

Meta data from data catalog

Reduced data

AR workflow AMQ log

Some tasks can be requested

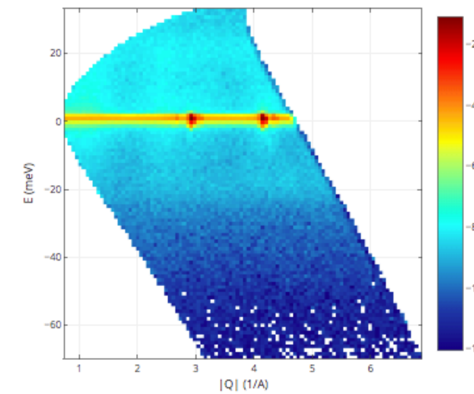
OAK RIDGE National Laboratory | HIGH FLUX ISOTOPE REACTOR | SPALLATION NEUTRON SOURCE — m2d | admin | logout

HYS Run 245701

home › hys › ipts-22351 › run 245701 — live monitoring: status | runs | PVs

previous | next

Run title	Hi Entropy, UnPol, 35 meV Ei 360 Hz S2 +40, 300 K
Run start	Jan. 10, 2020, 7:09 a.m.
Run end	Jan. 10, 2020, 8:09 a.m.
Duration	3606.06591797
Total counts	1735036
Proton charge	5.00115147116e+12



Data files:

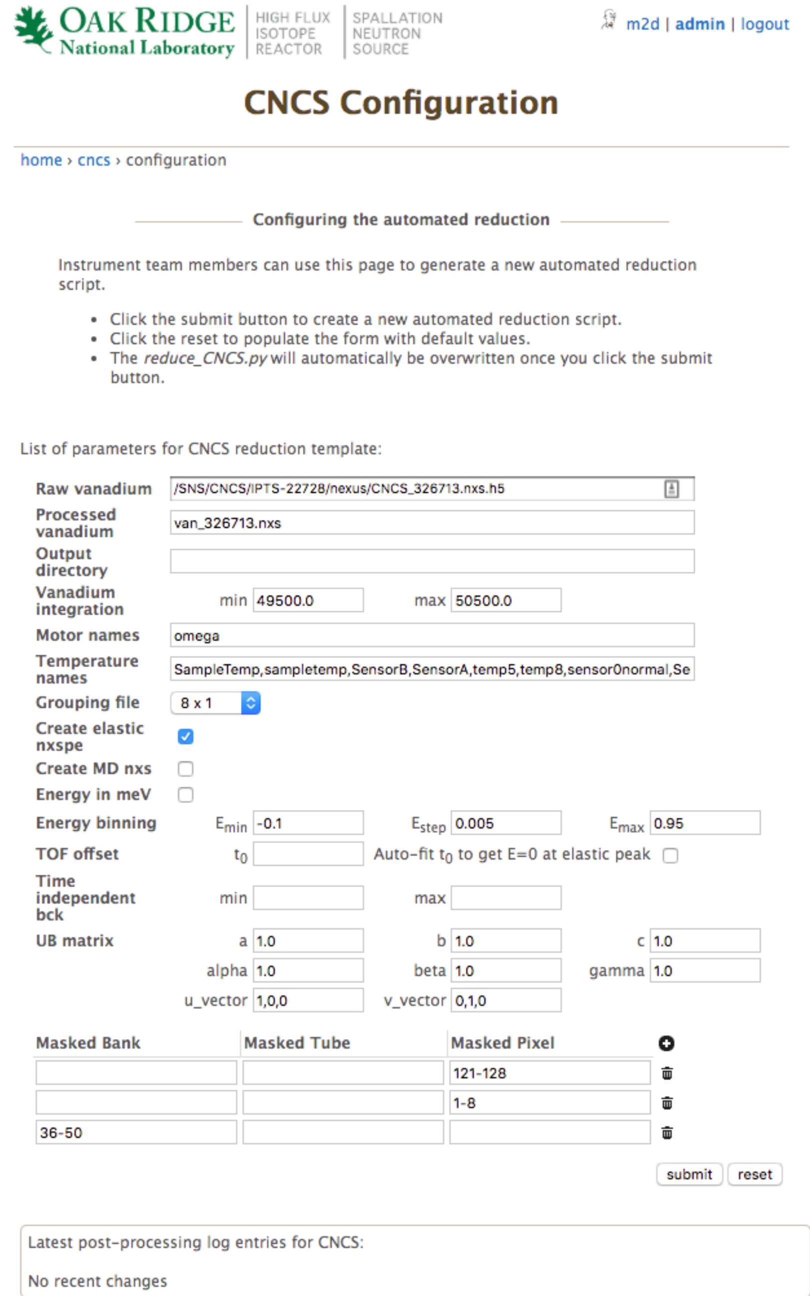
/SNS/HYS/IPTS-22351/nexus/HYS_245701.nxs.h5

Message	Information	Time
reduction_catalog.comple	autoreducer1.sns.gov	Jan. 10, 2020, 8:10 a.m.
reduction_catalog.started	autoreducer1.sns.gov	Jan. 10, 2020, 8:10 a.m.
reduction_catalog.data_re		Jan. 10, 2020, 8:10 a.m.
reduction.complete	Unverified HTTPS request is being made. Adding certificate verification is strongly advised. See: https://urllib3.readthedocs.org/en/latest/secu	Jan. 10, 2020, 8:10 a.m.
catalog.oncat.complete	autoreducer2.sns.gov	Jan. 10, 2020, 8:10 a.m.
catalog.oncat.started	autoreducer2.sns.gov	Jan. 10, 2020, 8:10 a.m.
reduction.started	autoreducer1.sns.gov	Jan. 10, 2020, 8:10 a.m.
catalog.complete		Jan. 10, 2020, 8:10 a.m.
reduction_catalog.comple		Jan. 10, 2020, 8:10 a.m.
catalog.oncat.data_ready		Jan. 10, 2020, 8:10 a.m.
catalog.complete		Jan. 10, 2020, 8:10 a.m.
reduction.data_ready		Jan. 10, 2020, 8:10 a.m.
reduction_catalog.comple		Jan. 10, 2020, 8:10 a.m.
postprocess.data_ready		Jan. 10, 2020, 8:10 a.m.
sms	Translation Succeeded - Run 245701 successfully translated	Jan. 10, 2020, 8:10 a.m.
sms	SMS run stopped	Jan. 10, 2020, 8:09 a.m.
sms	SMS Start Run Sent to STC	Jan. 10, 2020, 7:09 a.m.
sms	SMS run started	Jan. 10, 2020, 7:09 a.m.

Submit for post-processing: catalog | reduction | all post-processing

Reduction Flexibility

- Requirement that workflow be flexible
- Scientists for a given instrument can modify their reduction script
- There is also a custom form on the web monitor for more well defined workflows



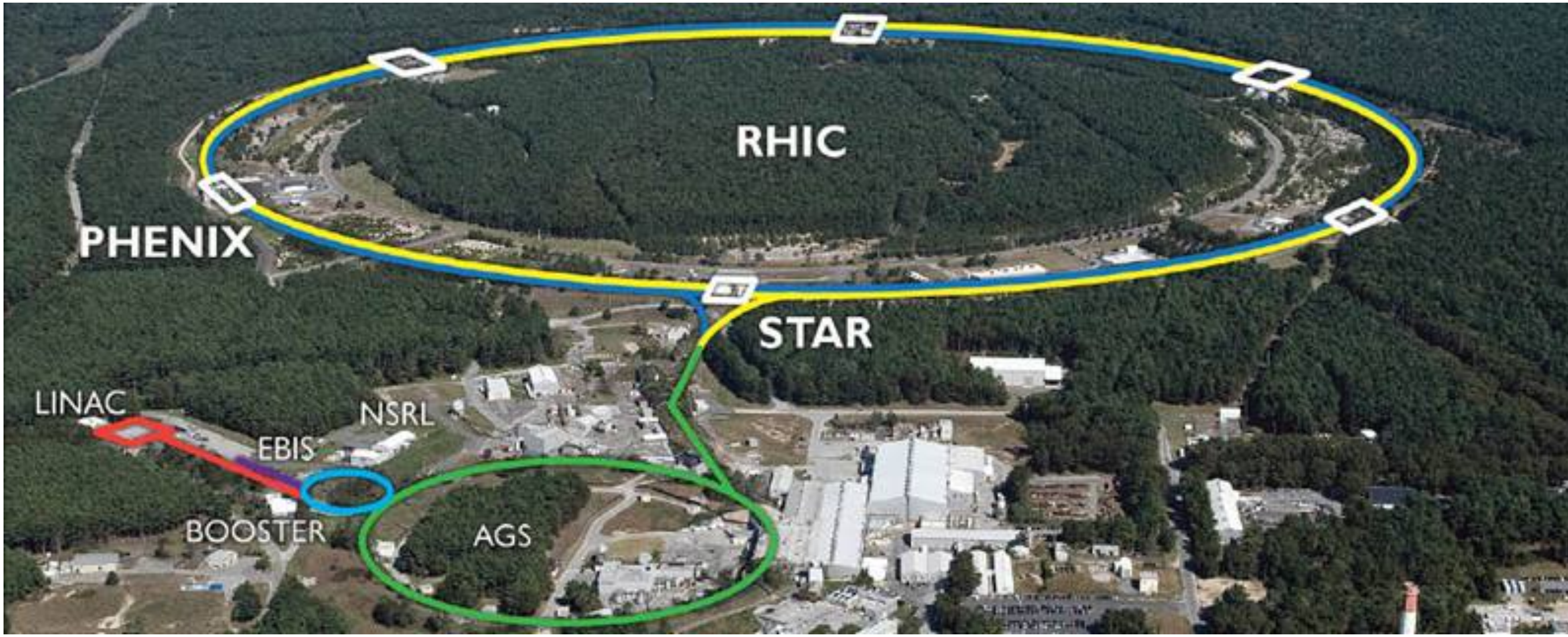
OAK RIDGE National Laboratory

Joe Osborn (ORNL)

Oak Ridge National Laboratory

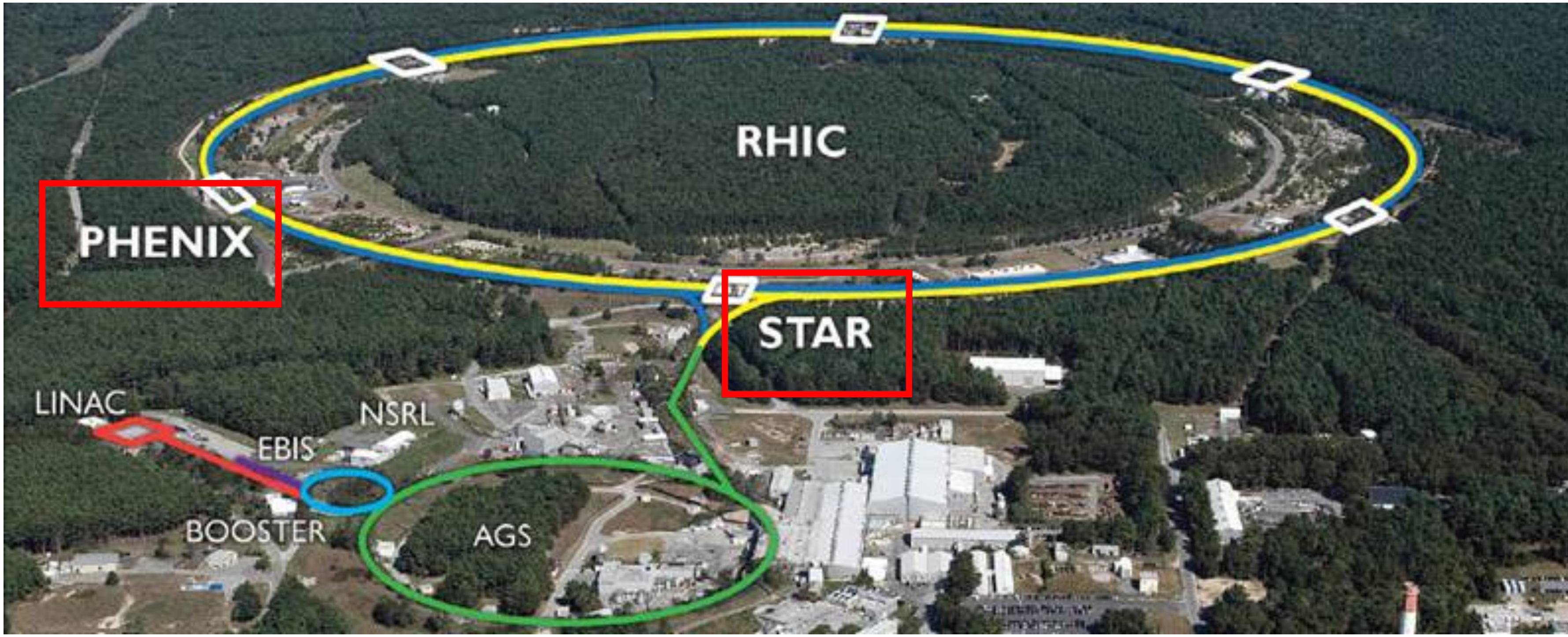
BNL Relativistic Heavy Ion Collider

Relativistic Heavy Ion Collider (RHIC)



- RHIC accelerates and collides protons and nuclei from deuteron to uranium
- Protons/nuclei are scraped off targets and accelerated up to ~99.999% the speed of light
- Collisions occur at two interaction points – STAR and (s)PHENIX

Relativistic Heavy Ion Collider (RHIC)



- RHIC has a different experimental model than SNS or HFIR
  - Two experiments with thousands of users
  - Accelerator division provides different types of collisions (e.g. proton-proton or nucleus-nucleus), determined by the two experiments
  - Users are almost exclusively accelerator and nuclear/particle physicists

Overview of Workflow

- Experiment collects data from a certain species of collisions (e.g. proton-proton)
- Measure particles in a given solid angle area
- Curate data with calibrations
- Users responsible for writing software to apply their own methods to curated data

Workflow Needs

- Highly flexible (!)
  - Simple ways to tweak and/or change a workflow given users variable needs
- A way to batch set of runs together for collective processing
- Process many PBs of data
- Be usable by a wide variety of people with various levels of software expertise

RHIC vs. SNS/HFIR Needs

RHIC

- Only two experimental facilities
- Collision species is fixed for a data taking period
  - Obtain one massive data set to be shared amongst users
- Analysis workflow depends on the user
  - E.g. corrections/efficiencies may be different analysis-to-analysis

Neutrons

- Many different instruments
- Users bring target material to study
  - Data depends on users target needs
- Analysis workflow depends on instrument/target

Data reduction workflows must be flexible to accommodate different user needs!

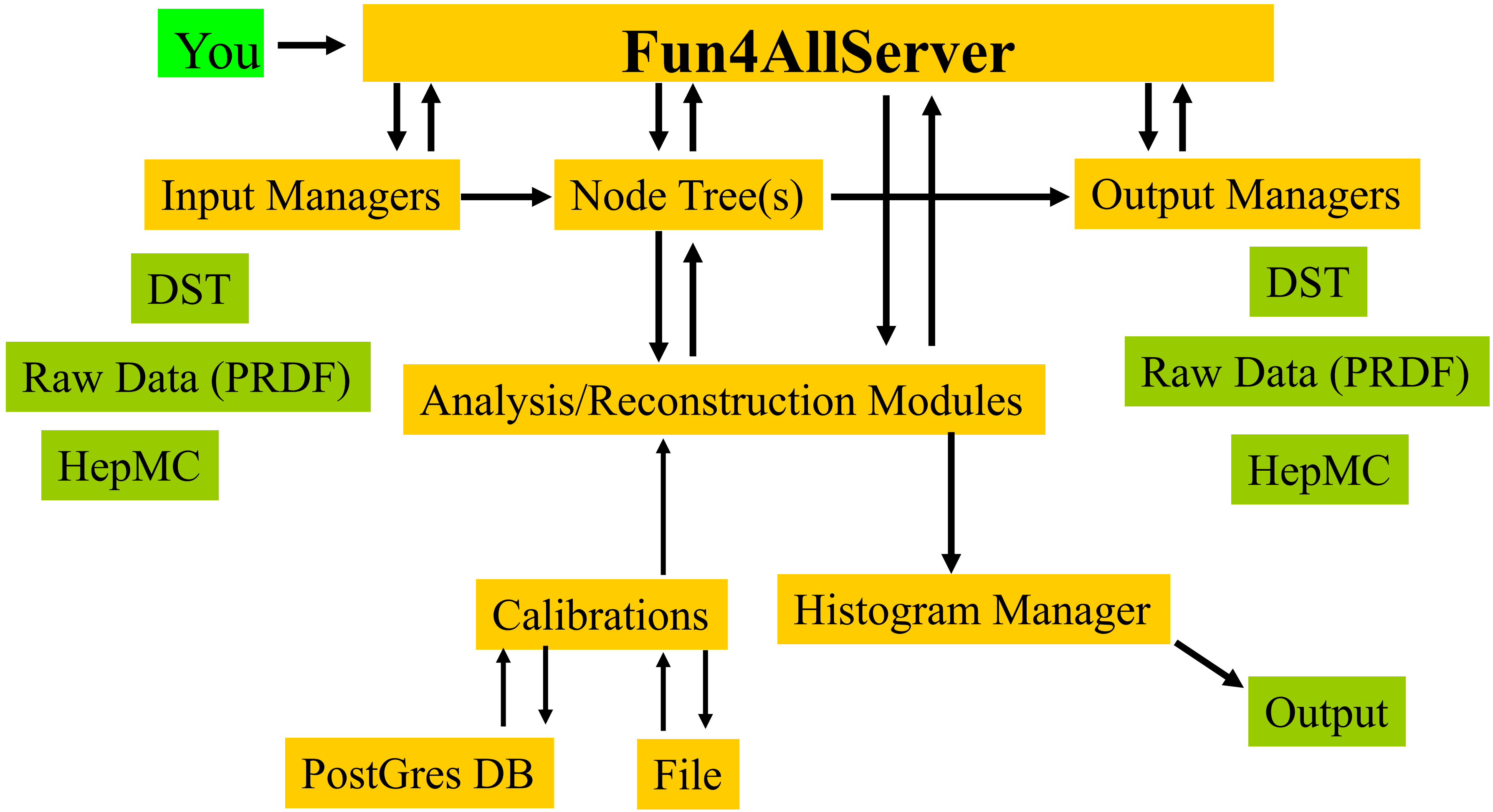
Oak Ridge National Laboratory

Current Workflow Implementation

Processing Architecture: Fun4All

- Fun4All is an open source workflow tool being used by several nuclear physics experiments worldwide
  - www.github.com/sPHENIX-Collaboration/coresoftware/tree/master/offline/framework
- Allows maximal flexibility while maintaining as little interface with the user as possible
- Development started in 2002, but has undergone many changes since then
  - Example – recent containerization with Singularity and ability to run jobs on the Open Science Grid
- Many thanks to Chris Pinkenburg (BNL) for the material for some of these slides

Fun4All


You
Fun4AllServer
Input Managers
Node Tree(s)
Output Managers
DST
Raw Data (PRDF)
HepMC
Analysis/Reconstruction Modules
DST
Raw Data (PRDF)
HepMC
Calibrations
Histogram Manager
PostGres DB
File
Output

Fun4All Node Tree

- Node tree contains data nodes from which modules access information
- Any kind of data can be stored on the node tree – as long as it is a defined class!

```
Node Tree under TopNode TOP
TOP (PHCompositeNode)/
   DST (PHCompositeNode)/
      PHG4INEVENT (PHDataNode)
      PIPE (PHCompositeNode)/
         G4HIT_PIPE (IO,PHG4HitContainer)
      MVTX (PHCompositeNode)/
         G4HIT_MVTX (IO,PHG4HitContainer)
      INTT (PHCompositeNode)/
         G4HIT_INTT (IO,PHG4HitContainer)
      TPC (PHCompositeNode)/
         G4HIT_ABSORBER_TPC (IO,PHG4HitContainer)
         G4HIT_TPC (IO,PHG4HitContainer)
      CEMC_ELECTRONICS (PHCompositeNode)/
         G4HIT_CEMC_ELECTRONICS
(IO,PHG4HitContainer)
      CEMC_SPT (PHCompositeNode)/
         G4HIT_CEMC_SPT (IO,PHG4HitContainer)
      G4HIT_CEMC (IO,PHG4HitContainer)
      G4HIT_ABSORBER_CEMC (IO,PHG4HitContainer)
      HCALIN (PHCompositeNode)/
         G4HIT_ABSORBER_HCALIN
(IO,PHG4HitContainer)

        …
```

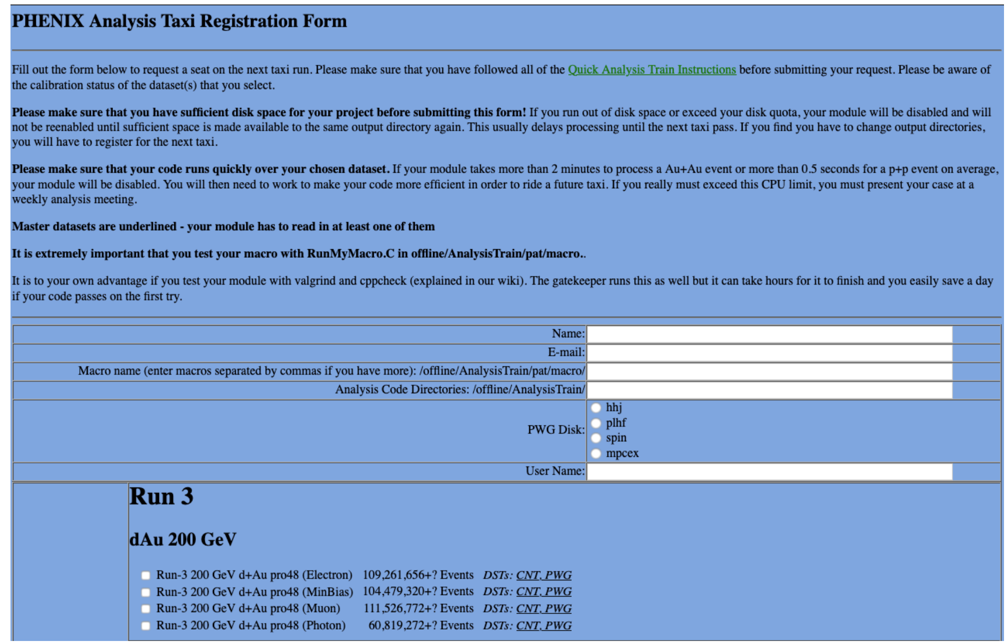
User Defined Analysis Modules

- Init(PHCompositeNode *topNode)
- InitRun(PHCompositeNode *topNode)
- Process_event (PHCompositeNode *topNode)
- ResetEvent(PHCompositeNode *topNode)
- End(PHCompositeNode *topNode)

- Users write various analysis modules to serve their specific needs
- Modules must inherit from SubsysReco which defines the workflow of the Fun4All system
- Otherwise, the rest is left to the user to ensure maximum flexibility

Batch Processing System

- Primarily use Condor as a batch processing system
- Users can submit an "analysis train" form
- Select their data set/calibrations/etc
- Submits jobs with a user defined script/module
- Job status can be monitored



PHENIX Analysis Taxi Registration Form

Fill out the form below to request a seat on the next taxi run. Please make sure that you have followed all of the Quick Analysis Train Instructions before submitting your request. Please be aware of the calibration status of the dataset(s) that you select.

Please make sure that you have sufficient disk space for your project before submitting this form! If you run out of disk space or exceed your disk quota, your module will be disabled and will not be reenabled until sufficient space is made available to the same output directory again. This usually delays processing until the next taxi pass. If you find you have to change output directories, you will have to register for the next taxi.

Please make sure that your code runs quickly over your chosen dataset. If your module takes more than 2 minutes to process a Au+Au event or more than 0.5 seconds for a p+p event on average, your module will be disabled. You will then need to work to make your code more efficient in order to ride a future taxi. If you really must exceed this CPU limit, you must present your case at a weekly analysis meeting.

Master datasets are underlined - your module has to read in at least one of them

It is extremely important that you test your macro with RunMyMacro.C in offline/AnalysisTrain/pat/macro..

It is to your own advantage if you test your module with valgrind and cppcheck (explained in our wiki). The gatekeeper runs this as well but it can take hours for it to finish and you easily save a day if your code passes on the first try.

Name:
E-mail:
Macro name (enter macros separated by commas if you have more): /offline/AnalysisTrain/pat/macro/
Analysis Code Directories: /offline/AnalysisTrain/
PWG Disk: hhj / plhf / spin / mpcex
User Name:

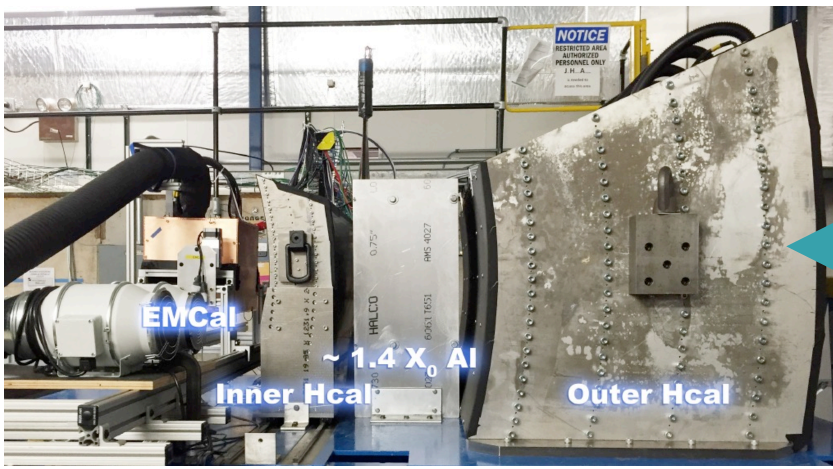
Run 3

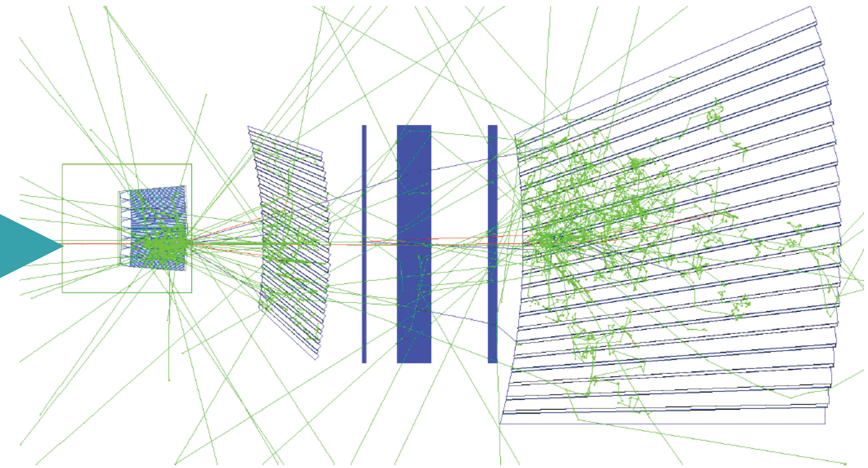
dAu 200 GeV

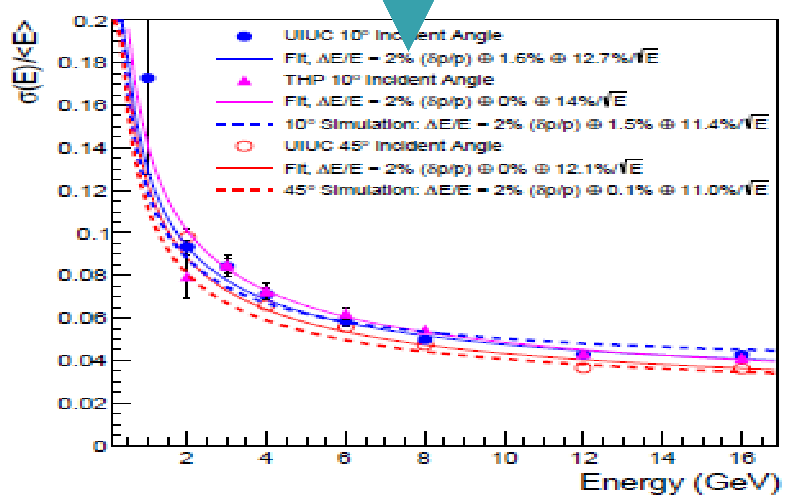
- Run-3 200 GeV d+Au pro48 (Electron) 109,261,656+? Events *DSTs: CNT, PWG*
- Run-3 200 GeV d+Au pro48 (MinBias) 104,479,320+? Events *DSTs: CNT, PWG*
- Run-3 200 GeV d+Au pro48 (Muon) 111,526,772+? Events *DSTs: CNT, PWG*
- Run-3 200 GeV d+Au pro48 (Photon) 60,819,272+? Events *DSTs: CNT, PWG*

Workflow Validation

- Allows for full validation of simulations and data, treated the same by Fun4All







Future Directions

Future Auto-Reduction Directions

- Planning a complete rewrite of the auto-reduction software
- Why?
  - Software is somewhat old – first developed in 2012
    - Remove some pieces that are now outdated
  - Scope of how we use auto-reduction has changed
  - What we want to do has changed
- Desires
  - Many of the same desires for original software, but with higher degree of flexibility and capability for modification as needs change

Serverless in Auto Reduction

- Recent push in industry towards ”serverless” architectures
  - Serverless definition – cloud computing model where a provider runs the server and dynamically manages the allocation of machine resources based on the needs of an application
- Greater level of flexibility – only use resources necessary
- Containerized execution – fully stateless implementation

Future Fun4All Directions

- New major nuclear physics initiative, the Electron Ion Collider, CD0 just approved by DOE
- Will have similar experimental situation to RHIC – one or two major experiments with thousands of users
- Currently there is a lot of discussion about the best approach to software frameworks
  - Use existing tools as base? Start as a Greenfield project?
- Other labs have explored workflows in particle physics as well
  - See for example the ART framework from Fermilab, art.fnal.gov

Conclusions

- Scattering facilities at ORNL and BNL have a huge depth of scientific needs
- Workflow systems for data reduction and analysis reflect the flexibility and variety of its users
- Workflow systems like auto reduction (ORNL) and Fun4All (BNL) allow for data curation for users to perform world leading science
- The needs from users are constantly changing
  – Software rewrites offer opportunities to re-evaluate effectiveness and assess the possibility for new features