Computing Practices and Challenges for COHERENT

Jacob Daughhetee

On behalf of the COHERENT Collaboration

11/15/2021

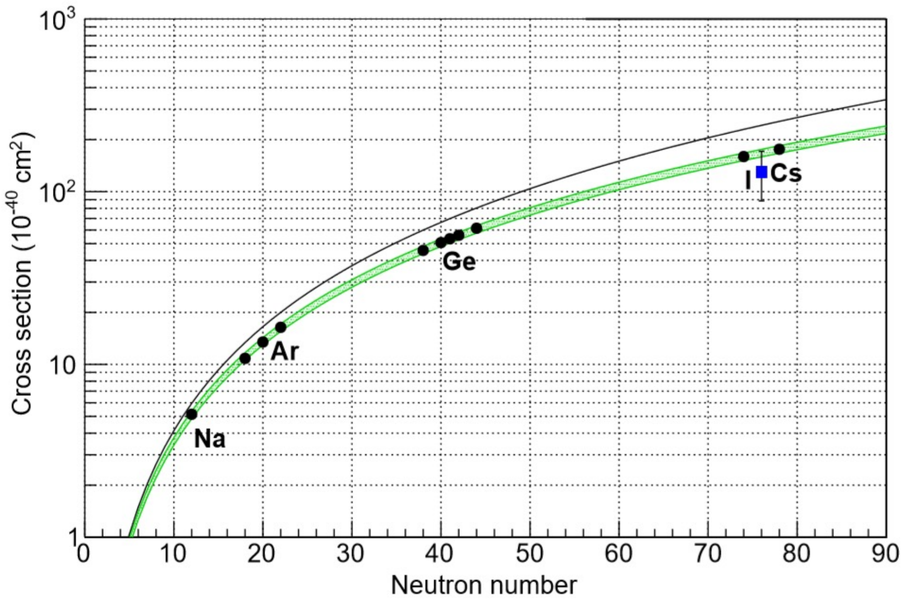
The COHERENT collaboration

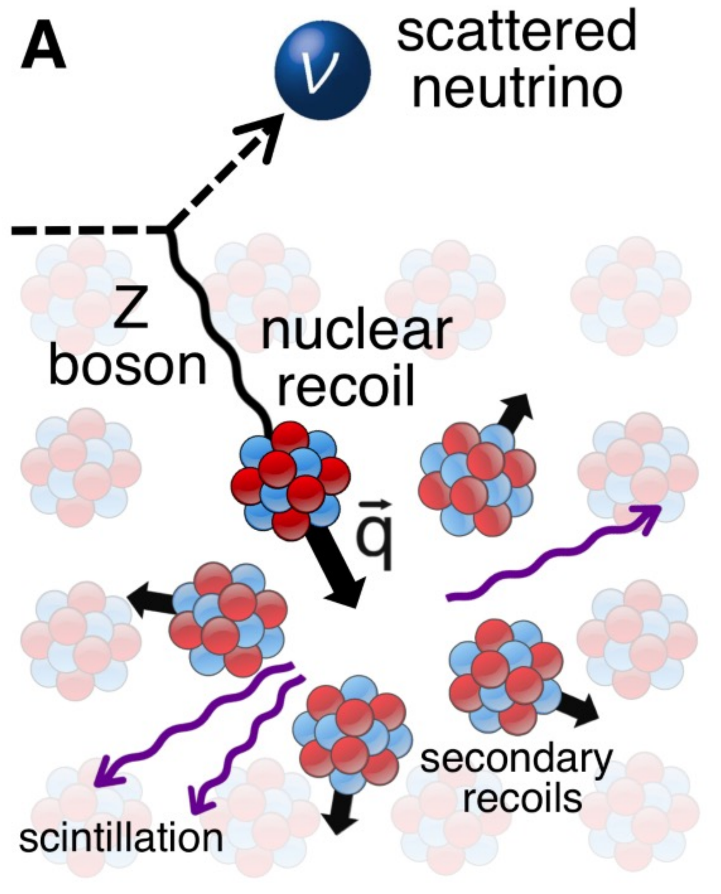
http://sites.duke.edu/coherent

~90 members,
19 institutions
4 countries

COHERENT

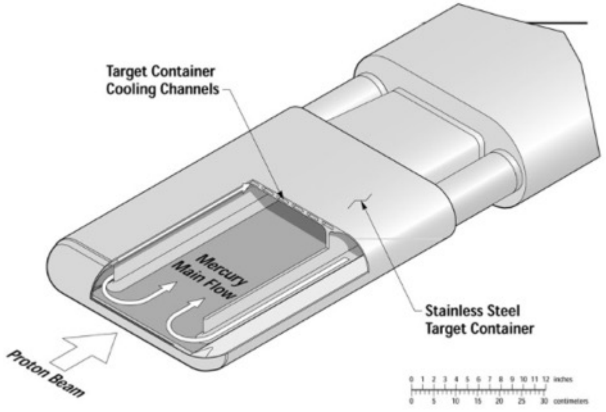
$$\sigma_{tot} = \frac{G_F^2 E_\nu^2}{4\pi}\left[Z\left(1-4\sin^2\theta_W\right)-\boxed{N}\right]^2 F^2(Q^2)$$

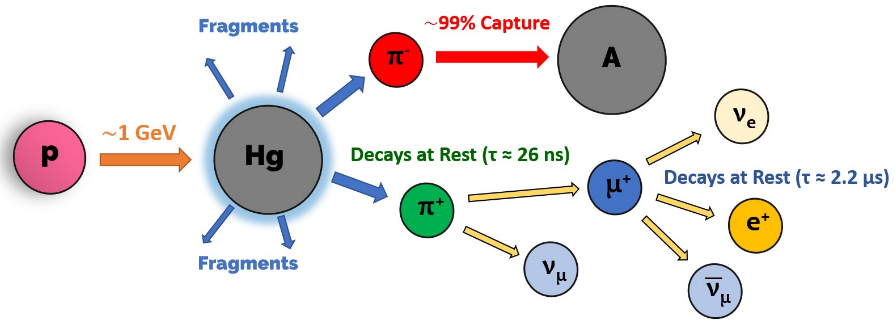


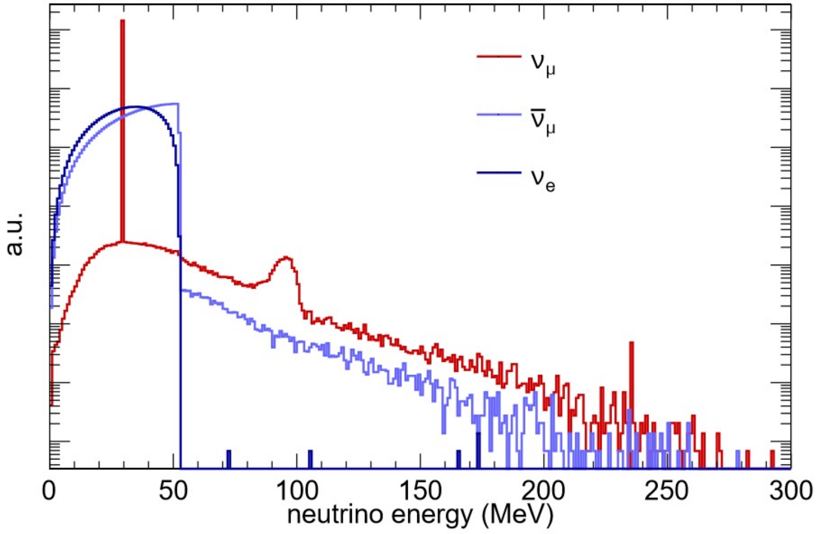


- **Coherent Elastic Neutrino-Nucleus Scattering**
- Clean prediction from the Standard Model – D. Freedman 1974
- Cross-section increases with energy as long as coherence condition is satisfied ($q \leq \sim R^{-1}$)
- Largest of all SM neutrino cross-sections at 1-100 MeV scale
- NC mediated: all flavors of neutrino can scatter via CEνNS

COHERENT



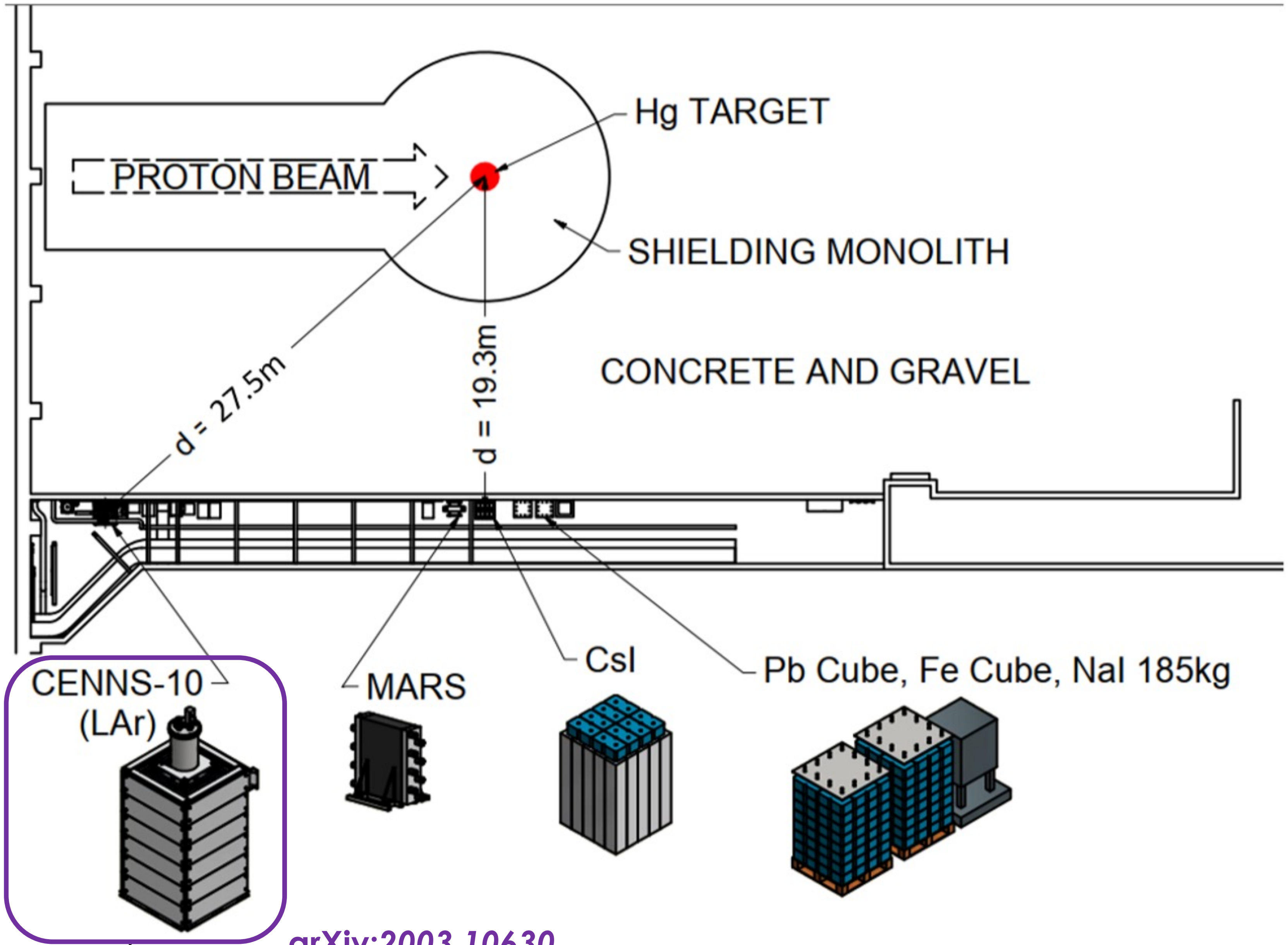


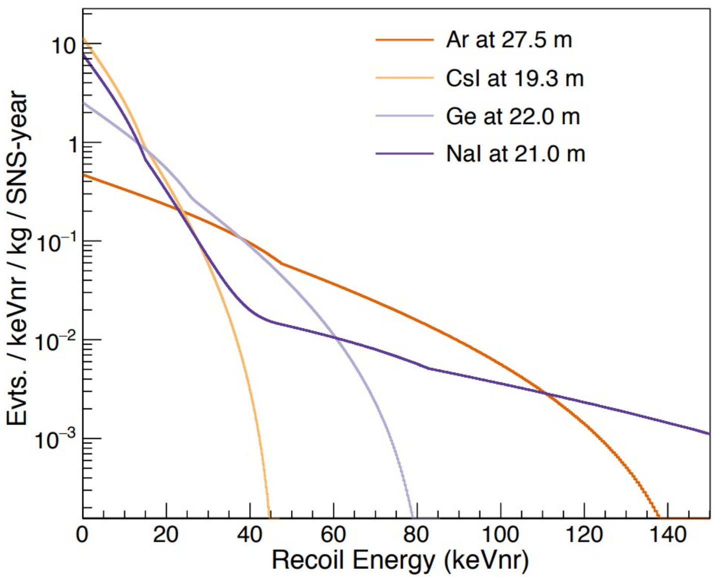


~1 GeV protons on Hg target at 60 Hz in 400 ns FWHM bunches

Oak Ridge National Laboratory

Neutrino Alley

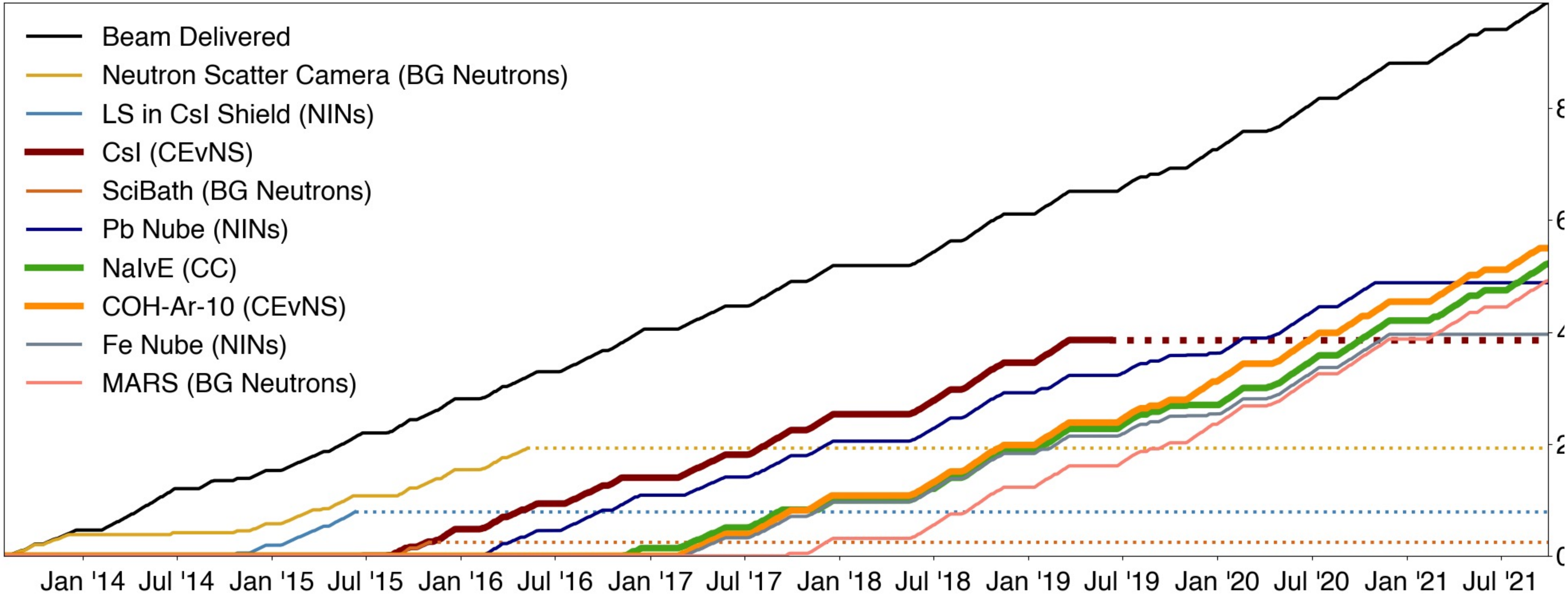




arXiv:2003.10630

Collecting Data

Beam delivered to COHERENT detectors



Data Handling

- Several continuously operating detectors produces a large amount of data!

- Fortunately, we've been able to make use of ORNL Physics computing resources.

- HCData cluster available for collaborators – multiple data drives attached and computing nodes available for sim/processing.

- However, space fills up quickly!

  - ➢ Raw data
  - ➢ Various stages of processing
  - ➢ Simulations

```
Size  Used Avail Use% Mounted on
139T  136T  2.8T  99% /data3

Size  Used Avail Use% Mounted on
124T  113T   12T  91% /data2

Size  Used Avail Use% Mounted on
124T  117T  7.0T  95% /data1

Size  Used Avail Use% Mounted on
139T  126T   13T  91% /data4

Size  Used Avail Use% Mounted on
139T  136T  3.0T  98% /data5

Size  Used Avail Use% Mounted on
110T   94T   16T  86% /data6
```

Data Handling

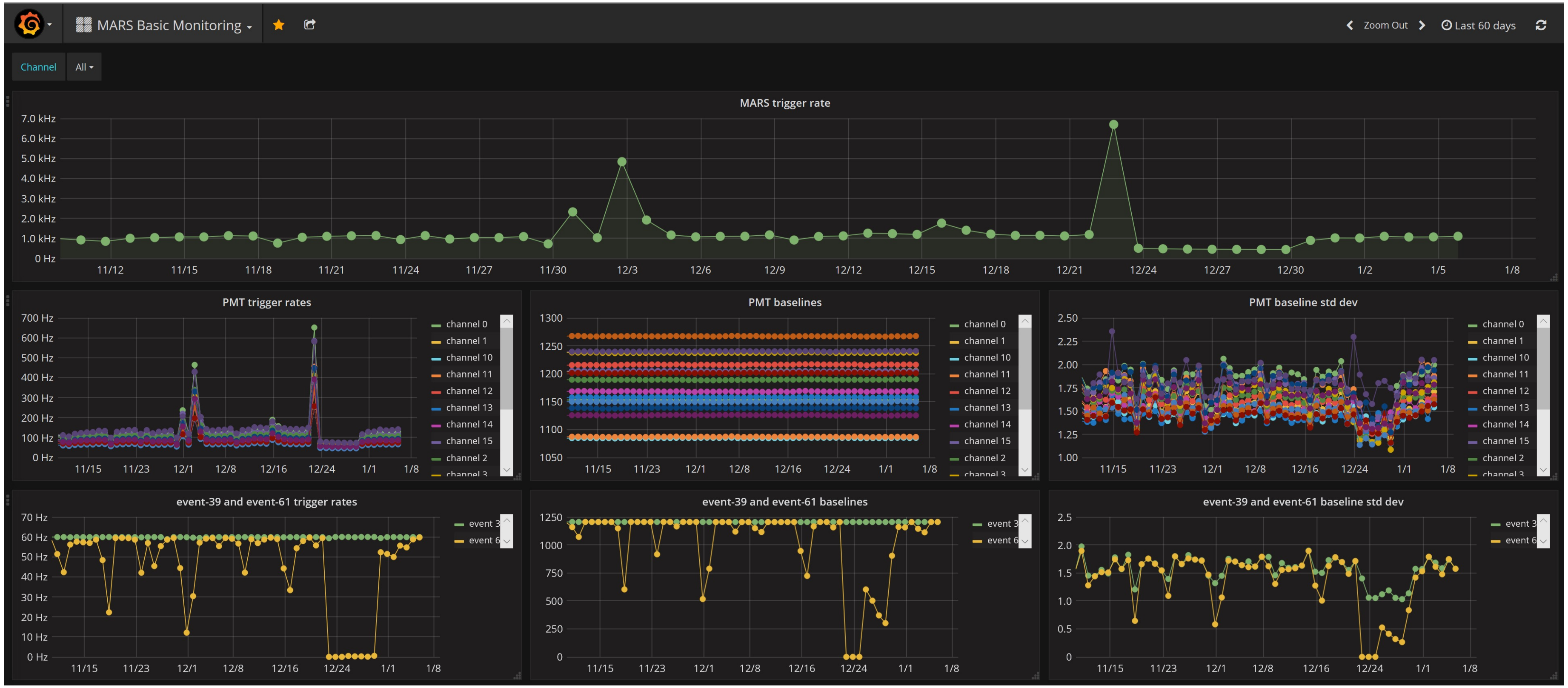
- Monitoring data streams from detectors often falls onto select individuals working on the specific experiments.

- Attempt to keep a handle on things by designating a data coordinator (position currently held by Rebecca Rapp).

- Tough to stay ahead of problems without personnel designated to keep track of drive fill rate, data transfers, etc.!

- Strategies for reducing data usage (and confusion)*:
    - Data processing acceptable for multiple analyzers
    - Unified simulations for each detector
    - Keeping run lists up to date in standardized format/location

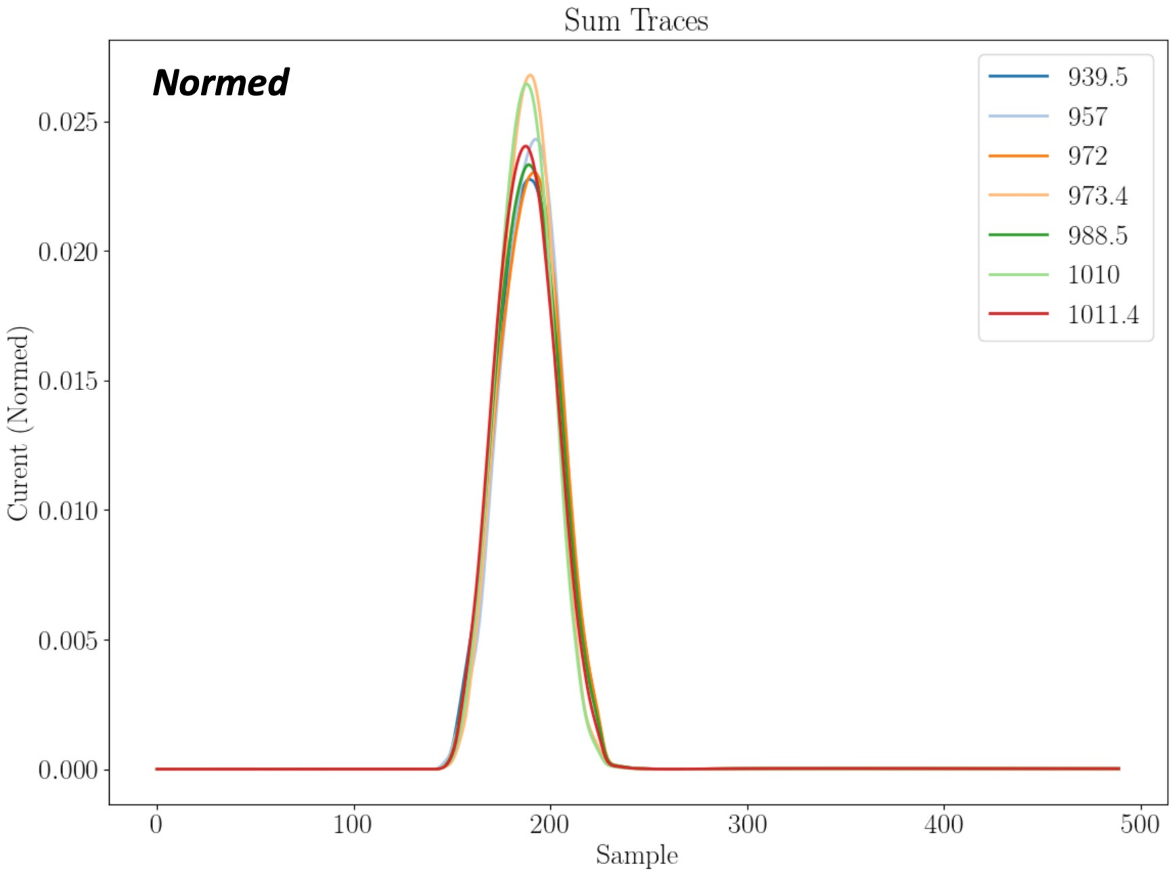
Online Monitoring



- Periodic checks on detector health and function necessary for to ensure high uptime.
- Access to information for several systems through Grafana
- 24/7 coverage for cryogenic LAr
- ThinLinc client allows external users to access LAr DAQ machines; status panel website displays all relevant operations info (LAr levels, temps, pressures, DAQ status, etc.)

Archiver

- Precise knowledge of beam performance reduces systematic uncertainties in time profiles and protons on target.
- SNS provides this information and broadcasts several variables related to accelerator functions.
- EPICS database can be queried for specific process variables (beam power, energy, radiation level in Nu Alley, etc.)
- COHERENT makes use of its own archiver to store relevant variable information for posterity.



Average POT trace waveforms generated from archiver

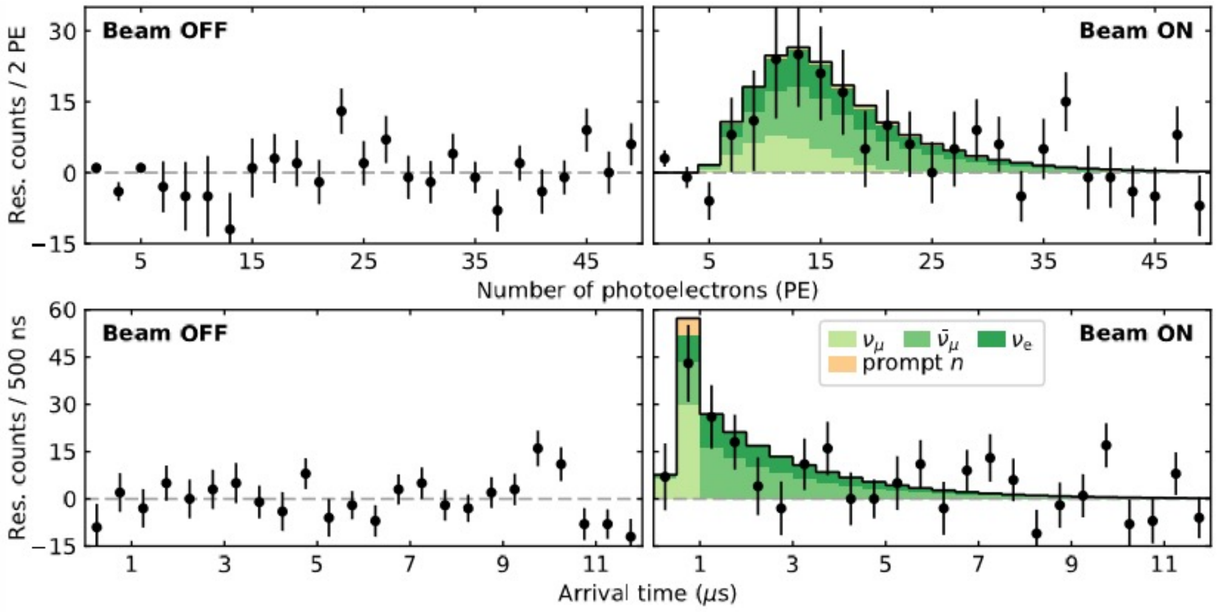
Data Releases

- Theory and phenomenological community are often *very* interested in obtaining experiment data at a rawer level
- Having the actual dataset is far preferable to data grabbing from plots
- There are many benefits to providing this with publication of results:
  - Builds confidence within the community
  - Prevents misunderstanding of analysis if relevant scripts / components provided

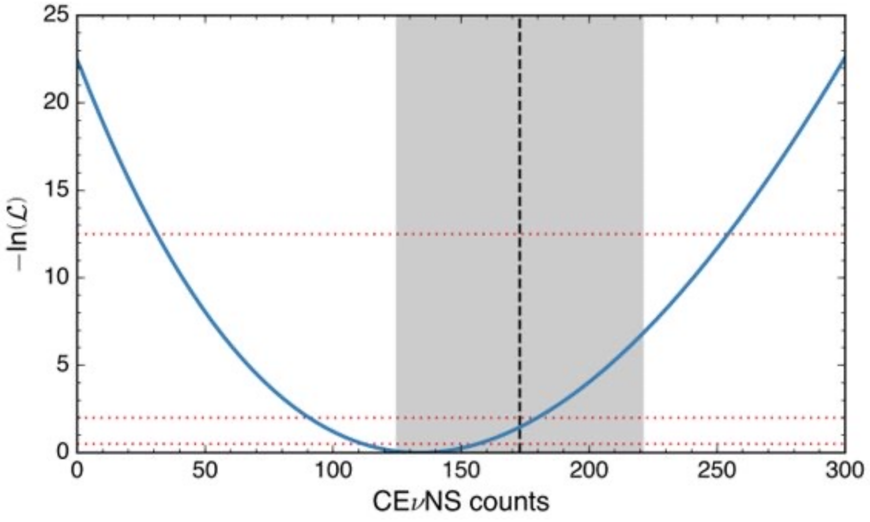
Data releases provided for CEvNS search results:

CsI 2018 -- https://zenodo.org/record/1228631

LAr 2020 -- https://zenodo.org/record/3903810



https://arxiv.org/abs/1708.01294



Data Releases

- Reproducibility of results is essential!
- Planning for data release forces use of best practices with version control.
- Provides a ‘frozen’ version of the analysis that can always be returned to.
- Release of containerized environment (e.g., Singularity) would be ideal. Hope to include in future releases.

zenodo Search Upload Communities

June 22, 2020 Dataset Open Access

COHERENT Collaboration data release from the first detection of coherent elastic neutrino-nucleus scattering on argon

Akimov, D.; Albert, J.B.; An, P.; Awe, C.; Barbeau, P.S.; Becker, B.; Belov, V.; Blackston, M.A.; Blokland, L.; Bolozdynya, A.; Cabrera-Palmer, B.; Chen, N.; Chernyak, D.; Conley, E.; Cooper, R.L; Daughhetee, J.; del Valle Coello, M.; Detwiler, J.A.; Durand, M.R.; Efremenko, Y.; Elliott, S.R.; Fabris, L.; Febbraro, M.; Fox, W.; Galindo-Uribarri, A.; Green, M.P.; Hansen, K.S.; Heath, M.R.; Hedges, S.; Hughes, M.; Johnson, T.; Kaemingk, M.; Kaufman, L.J.; Khromov, A.; Konovalov, A.; Kozlova, E.; Kumpan, A.; Li, L.; Librande, J.T.; Link, J.M.; Liu, J.; Mann, K.; Markoff, D.M.; McGoldrick, O.; Moreno, H.; Mueller, P.E.; Newby, J.; Parno, D.S.; Penttila, S.; Pershey, D.; Radford, D.; Rapp, R.; Ray, H.; Raybern, J.; Razuvaeva, O.; Reyna, D.; Rich, G.C.; Rudik, D.; Runge, J.; Salvat, D.J.; Scholberg, K.; Shakirov, A.; Simakov, G.; Sinev, G.; Snow, W.M.; Sosnovtsev, V.; Suh, B.; Tayloe, R.; Tellez-Giron-Flores, K.; Thornton, R.T.; Tolstukhin, I.; Vanderwerp, J.; Varner, R.L.; Virtue, C.J.; Visser, G.; Wiseman, C.; Wongjirad, T.; Yang, J.; Yen, Y.-R.; Yoo, J.; Yu. C.-H.; Zettlemoyer, J.

Release of COHERENT collaboration data from the first detection of coherent elastic neutrino-nucleus scattering (CEvNS) on argon. This data release corresponds with the results of "Analysis A" published in arXiv:2003.10630[nucl-ex]. The data release enables further studies of CEvNS.

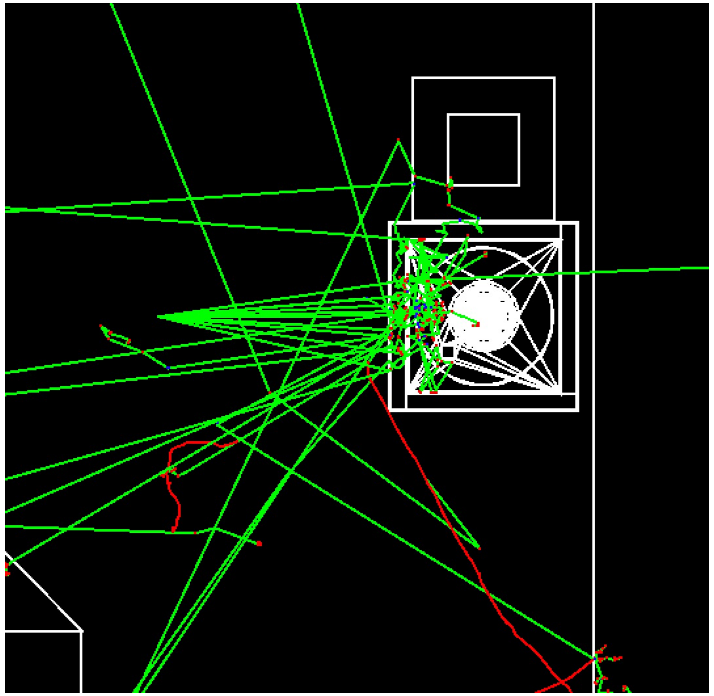
Use of the data release is presented in the accompanying pdf document within this submission. Example code is included within the release as part of this submission. The materials here are also available at http://coherent.ornl.gov/data/, which preserves the directory structure used within the accompanying document. Note the use of the example code in this release expects the directory structure written within the accompanying pdf document.

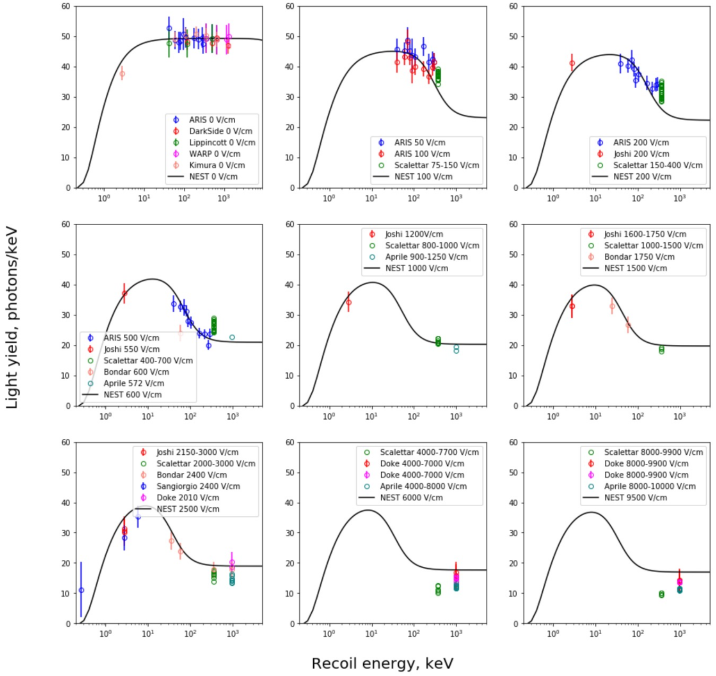
Preparing data releases requires proper accounting of data and software used for analysis!

Simulations

- Geant4 predominantly used for detector simulations.
- Cross-checks on neutron simulations using MCNP.
- Interest in using NEST as we push down to lower energies in LAr.
- Scintillation behavior parameterized at higher energies begins to deviate at energies approaching current detection threshold.

DT Neutron simulation for COH-LAr-10

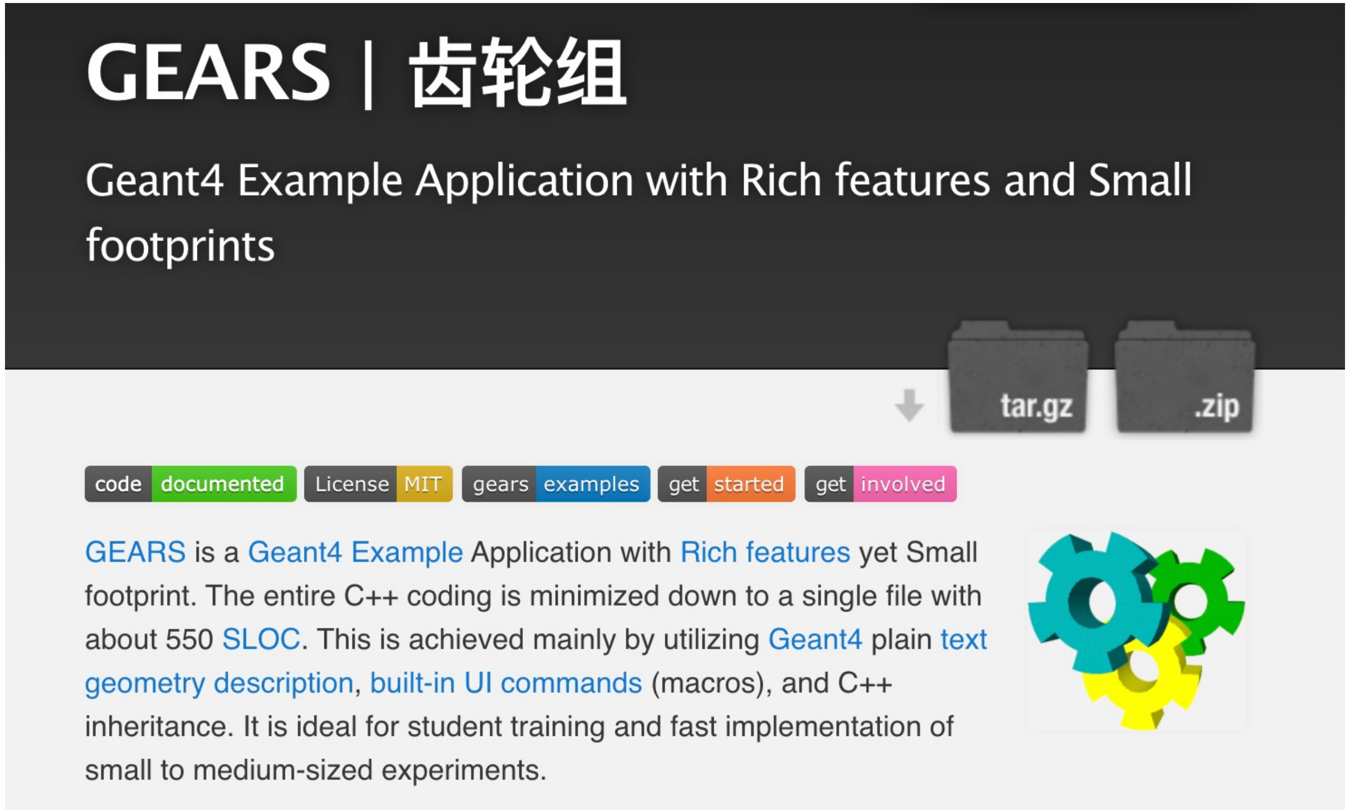


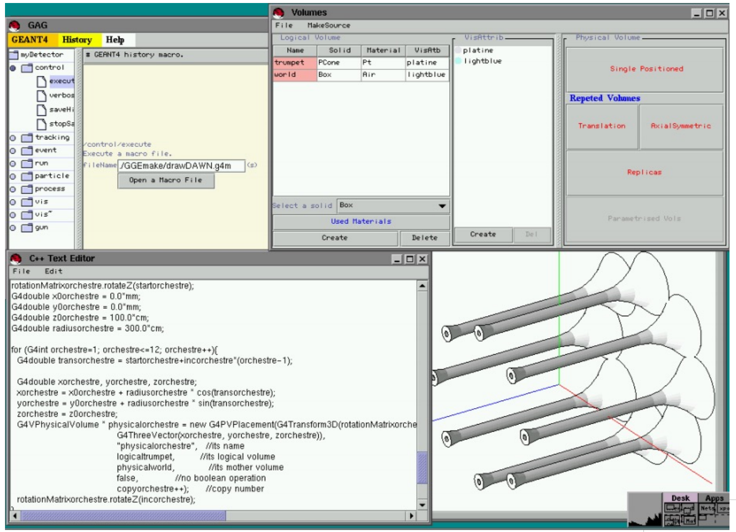


- Argon NEST is under-development version of NEST for argon (currently in alpha version)
- Can simulate Ly/Qy and also comparable with G4 models (as library/separate functions)

GEARS

- Simplified implementation of G4 simulation ideal for new initiates.
- Graphical geometry editor
- Application has large degree of versatility for experienced users
- Not widely adopted by collaboration at this time but hope to support more in near future for new students.





OAK RIDGE National Laboratory

Software Challenges

Managing software can present many difficulties in small to mid-sized experiments:

- Consensus on specific software tools to use (ROOT, Python, etc) – often have groups splinter along preferred languages

- For COHERENT, different detector systems use different DAQ and analysis software.

- Adherence or even agreement on best use practices

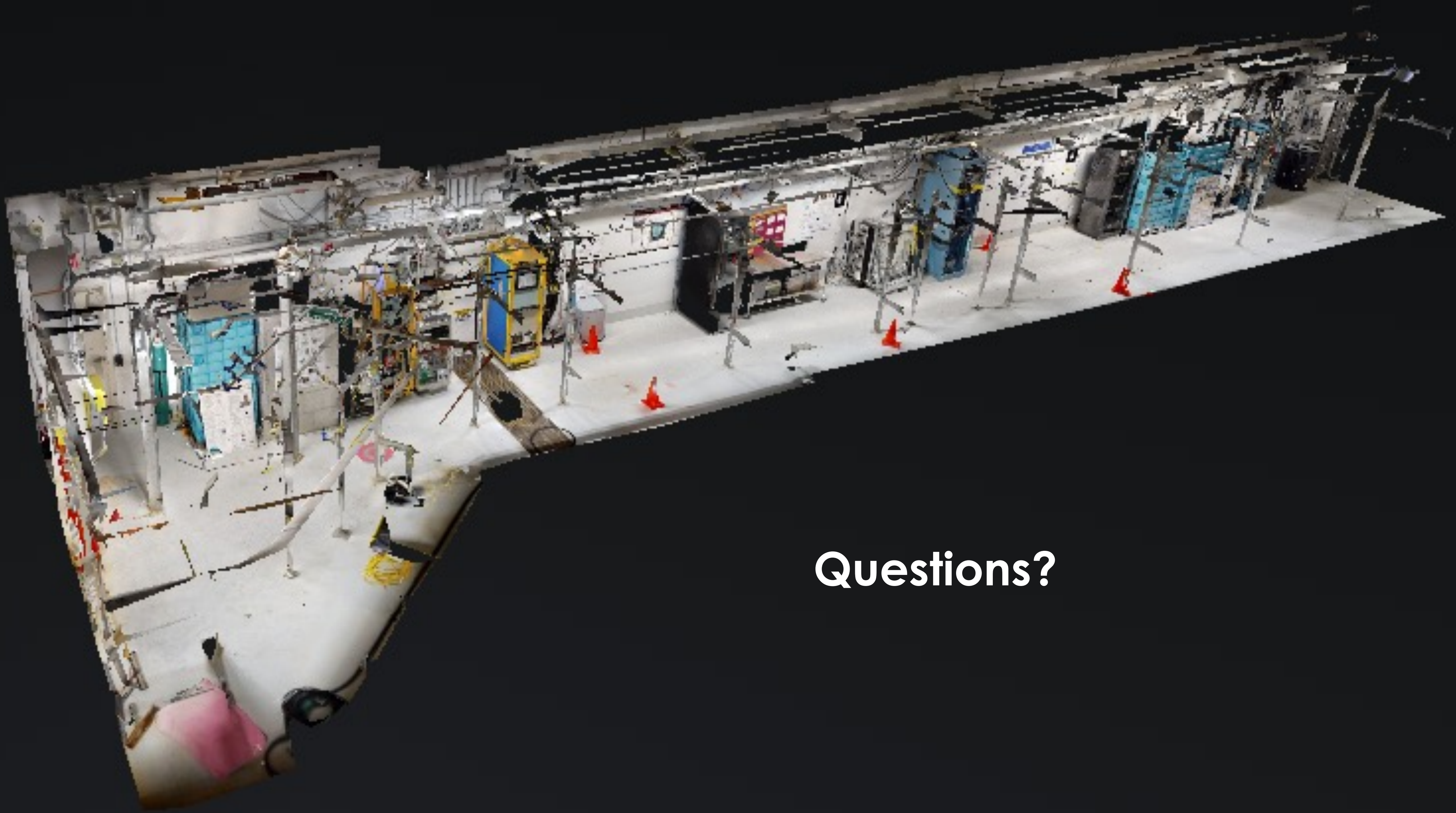
- Turnover rate of graduate students and postdocs can disrupt proposed unified software tools

- Many benefits to unifying software tools – but it is hard to overcome inertia!

    - ➢ Prevents reproduction of previous efforts
    - ➢ Wider base of users for resolving problems
    - ➢ Facilitates moving between detector data without being totally lost

Summary

- COHERENT faces several challenges in software usage due to data from multiple detectors analyzed by many different users.
- The collaboration has implemented several checks to ensure reproducibility of results and accounting of data.
- Starting these internal controls can be difficult, but they are very useful.
- Bringing everyone under one umbrella is difficult. Especially true with lack of continuity in organizational positions held by younger members.
- The collaboration has established Data and Software Coordinator positions to try to prevent problems arising and to unify efforts.
- Suggestions on tools or strategies welcome!



Questions?