Data and the software for Neutron Scattering at the SNS and HFIR

Garrett E. Granroth
Neutron Data Sciences Group Leader
Neutron Scattering Division
How 2019 March 18, 2019
Outline

• Our Customers (Users)
• Overview of our Instruments
• Scope of Data
• Software Ecosystem
• Reduction Software
• Comparison of Models
Users

• ~1000 unique users every year

• Usually groups of 2 – 5

• Broad scientific interest
  – Quantum materials, thermoelectrics, catalysts, battery materials, superconductors, polymers, macro molecules, bio fuels, engineering, materials, and many more.
Spallation Neutron Source instruments

- 18 Instruments
  - In User Program
  - For studying materials
- Pulsed; 60 Hz; 1.4MW
- Operates ~ 200 days per year
- experiment changes ~ 1 - 5 days per instrument
High Flux Isotope Reactor Instruments

- **12 Instruments**
  - In User Program
  - Materials studies
- **Continuous 85 MW**
- **Operates 7, 24 day cycles**
- **Experiment changes every ~1 - 5 days per instrument**
Neutron data has expanded in size and complexity
Data

• < 5M events /s /instrument ~ 60 MB/s/instrument of raw data on the stream

• Saved to HDF5 files using NeXus schema (https://www.nexusformat.org/)

• Capturing as much metadata as we can

• Cataloging using MongoDB based solution Oncat. (Oncat.ornl.gov)

• Users usually
  – work with large data on our machines downloading results via sftp client
  – Small data they download via sftp client to their own machine
  – Can use globus to transfer large data (Leveraging ORNL computing)
Software Ecosystem

• Modular codes easily glued together
  – Some critical codes not easily glueable

• Workflow engines
  – can leverage HPC, and Cloud
  – User provided compute is harder
  – Easily configure novel workflows
Mantid – Community Code for Reducing neutron data

- Convert from events to histograms (reduce data)
  - In some cases data size grows.
  - Format for that users can work with

- Event filtering for time varying stimuli (J. Appl. Cryst. 51, 616 (2018))

- Facility for live reduction

- Stores a history in metadata

- Python API allows,
  - Autoreduction, custom guis, expanded workflows

- C++/Python framework

- Collaboration of STFC, ORNL, ESS, ILL

- PSI, ANSTO, Juelich contribute

- Mantidproject.org
Materials Models need More compute

• Density Functional Theory (DFT)
• Molecular Dynamics (MD)
• Big Box Structure calculations

Molecules of the heavy hydrogen isotopes deuterium and tritium preferentially bind to copper atoms in a metal-organic framework compound.

Nature Comm. 8, 14496 (2017)


• Javelin
  - https://github.com/rosswhitfield/javelin
Instrument Models need more compute

- Complex multiple scattering understood by Monte Carlo ray tracing

- Super resolution techniques are being investigated to remove instrumental effects

Workflow engines can help

• Pegasus (Panorama project) https://panorama360.github.io/
  – MD and quasi elastic scattering
  – Big box models and diffuse scattering
  – Monte Carlo ray tracing of direct deometry spectrometers

• ICE (Iceman project) https://wiki.eclipse.org/ICE
  – DFT and molecular spectroscopy

• Beam https://doi.org/10.1016/j.procs.2016.05.410
  – MD and quasi elastic scattering
  – Polymer modeling for reflectivity data
Conclusions

• A modular software infrastructure is critical
• We use Python to glue codes together.
• Mantid is a key framework for our community
• Workflow engines are helpful, but the ideal is not there yet.