Software Sustainability in Experimental Facilities

Anshu Dubey

Software Engineering and Reuse in Modeling, Simulation, and Data Analytics for Science and Engineering (ISC 2022 BOF)

License, Citation and Acknowledgements

License and Citation

Acknowledgements

- This work was supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research (ASCR), and by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.
- This work was performed in part at the Argonne National Laboratory, which is managed by UChicago Argonne, LLC for the U.S. Department of Energy under Contract No. DE-AC02-06CH11357.
- This work was performed in part at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

U.S. DEPARTMENT OF ENERGY | Office of Science

Stewardship and Sustainability of Office of Science Software

A working group drawn from research divisions, computing facilities and experimental/observational facilities

Experimental and observational facilities have some similar and some unique challenges

Sustainability is a multi-pronged issue for them

Argonne NATIONAL LABORATORY

BERKELEY LAB
Lawrence Berkeley National Laboratory

OAK RIDGE National Laboratory

The Team

- **Anshu Dubey**, Mathematics and Computer Science, Argonne National Laboratory
- **Katherine Riley**, Argonne Leadership Computing Facility, Argonne National Laboratory
- **Nicholas Schwarz**, Advanced Photon Source, Argonne National Laboratory
- **David E. Bernholdt**, Computer Science and Mathematics and Oak Ridge Leadership Computing Facility, Oak Ridge National Laboratory
- **Bronson Messer**, Oak Ridge Leadership Computing Facility, Oak Ridge National Laboratory
- **Mathieu Doucet,** Neutron Scattering Division, Oak Ridge National Laboratory
- **Rama K. Vasudevan**, Center for Nanophase Materials Sciences, Oak Ridge National Laboratory
- **Deborah Agrawal**, Computing Research Division, Lawrence Berkeley National Laboratory
- **Katerina Antypas**, National Energy Research Scientific Computing**,** Lawrence Berkeley National Laboratory
- **Harinarayan Krishnan**, Advanced Light Source/Computing Research Division, Lawrence Berkeley National Laboratory
- **Edward Balas**, Energy Sciences Network, Lawrence Berkeley National Laboratory

IDEAS productivity

ECP EXASCALE COMPUTING PROJECT

Experimental Facilities Software

Operational software
– runs the equipment
/ experiment
Manage allocations

Experimental Facilities Software

Operational software
– runs the equipment / experiment
Manage allocations

Data acquisition software
from instruments
from sensors and diagnostic equipment

Data management software
curation
archival
compression

Experimental Facilities Software

Operational software
– runs the equipment / experiment
Manage allocations

Data acquisition software
from instruments
from sensors and diagnostic equipment

Data management software
curation
archival
compression

Data analysis software

Experimental Facilities Software

Operational software
– runs the equipment / experiment
Manage allocations

Data acquisition software
from instruments
from sensors and diagnostic equipment

Data management software
curation
archival
compression

Data analysis software

Simulations for experiment design or validation

Experimental Facilities Software

Operational software – runs the equipment / experiment
Manage allocations

Could be vendor provided

Variable complexity

→ training to use

Critical to be up all the time -- fault tolerance

Usually resources for sustaining built into the operational cost of the facility

Experimental Facilities Software

Data acquisition software
from instruments
from sensors and diagnostic equipment

Could be vendor provided – maybe platform specific
Could be custom written -- many pain points
Custom solutions don't always interoperate, limited testing

Data management software
curation
archival
compression

General lack of information about good solutions
Could be vendor provided, platform specific
Pipelines for moving data around i.e. from experimental facility to HPC center

Experimental Facilities Software

Data analysis software

Simulations for experiment design or validation

These challenges tend to be similar to other fields.

Typically developed less sustainably than other computational software
Many facilities still in the earliest stages