Toward “Real-Time” Analysis in Neutron Science

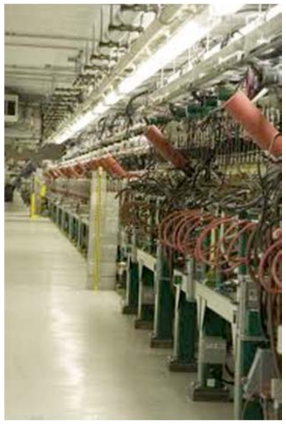
Ryan Adamson, Jose Borreguero, Blake Caldwell, Stuart Campbell, Rich Crompton, Olivier Delaire, David Dillow, Mathieu Doucet, Dave Giles, Monojoy Goswami, Melissa Graves-Brook, Gayle Greene, Matthew Gyurgyik, Mark Hagen, Steven Hartman, Matthew Hermanson, Steve Hicks, Kay Kasemir, Jason Kincl, Scott Koch, James Kohl, Chris Layton, Vickie Lynch, Ross Miller, Dude Neergaard, Andre Parizzi, Daniel Pelfrey, Pete Peterson, Thomas Proffen, Shelly Ren, Michael Reuter, Charles Roberts, Andrei Savici, **Galen Shipman**, Sergey Shpanskiy, Dale Stansberry, Bobby Sumpter, Madhan Sundaram, John Quigley, Carol Tang, Kevin Thach, Tara Thompson, Bogdan Vacaliuc, Sudharshan Vazhkudai, Pedro Vicente, Karen White, Barry Winn, Russell Taylor, Marie Yao

U.S. DEPARTMENT OF ENERGY

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

From Ion Source to Neutron Detectors

ION Source + Linac → H- → Proton Accumulator → H+ → Mercury Target → Neutrons! → Instrument Detectors

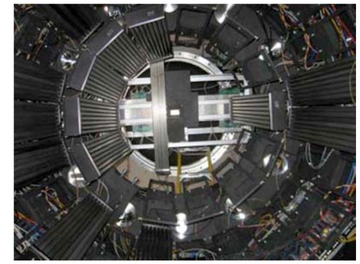


Ion source produces negatively charged hydrogen ions formed into a pulsed beam at 2.5 MeV

Linear accelerator accelerates the H- beam from 2.5 MeV to 1 GeV

H- pulse from linac is wrapped into a ring through a stripper foil, removing the electrons to produce protons H+. Protons are accumulated in the ring and then kicked out at 60 Hz

These pulsed, high-energy protons are directed to a mercury target, when a proton hits the target, 20 to 30 neutrons are released, moderated and directed to individual beam lines

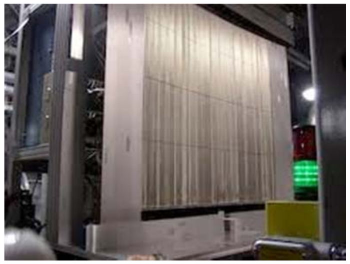


Neutron detectors surrounding a sample environment detect collisions of neutrons with the detector. The TOF and angle of incidence of the neutron is calculated.

OAK RIDGE National Laboratory

Many different types of instruments

- Inelastic Spectroscopy – ARCS, CNCS, SEQUOIA, HYSPEC
  - Resolves the change in kinetic energy when neutron/sample collisions are inelastic
  - Commonly used in condensed matter research to study atomic motion, magnetic and crystal field excitations
- Diffraction (Elastic) – POWGEN, NOMAD, VULCAN
  - Resolves the structure of a material
  - Use to study nanoscale structure and magnetic structure
- Small-angle neutron scattering (Elastic) – EQ-SANS, USANS
  - Uses small scattering angles to investigate the structure of materials at the mesoscopic scale
  - Use to study proteins, DNA, and other biological molecules; polymers, crystalline structures, etc.



SNS Data Life Cycle

Acquisition

- Neutron events
- Events from sample environment
- Other triggers

Reduction

- Corrected reduced data (histograms, S(Q,E), ..)
- Merging, reconstruction of data
- Instrument/technique dependent
- Need for ’real’ time reduction

Analysis

- Multi dimensional fitting
- Advanced visualization
- Comparison to simulation / feedback
- Field dependent, large variety of approaches

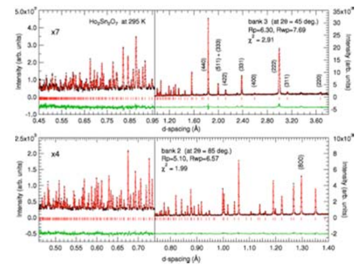
Simulation Modeling

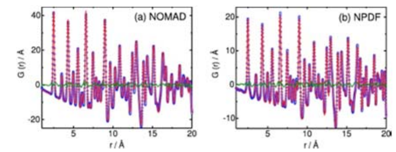
- Multitude of techniques (DFT, MD, ..)
- Advanced simulation of experiments
- ‘Refinement’ using experimental data
- Multiple experiments / probes

User Facility

- Variety of experiments, topics, methods and ‘computer literacy’ of users are significant challenge.

Example: NOMAD Diffractometer





Raw Data: up to 10^12 events per second

Acquisition →

Translated Data: Gigabytes to Terabytes

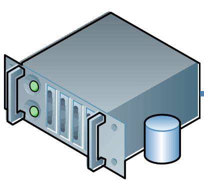
Reduction →

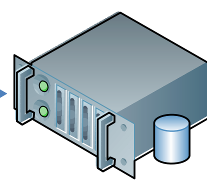
Reduced Data: e.g. Powder Diffraction Pattern

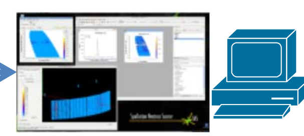
Analysis & Simulation →

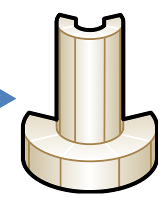
Analysis: PDF, MD simulation, etc.

Feedback guiding changes to the experiment setup









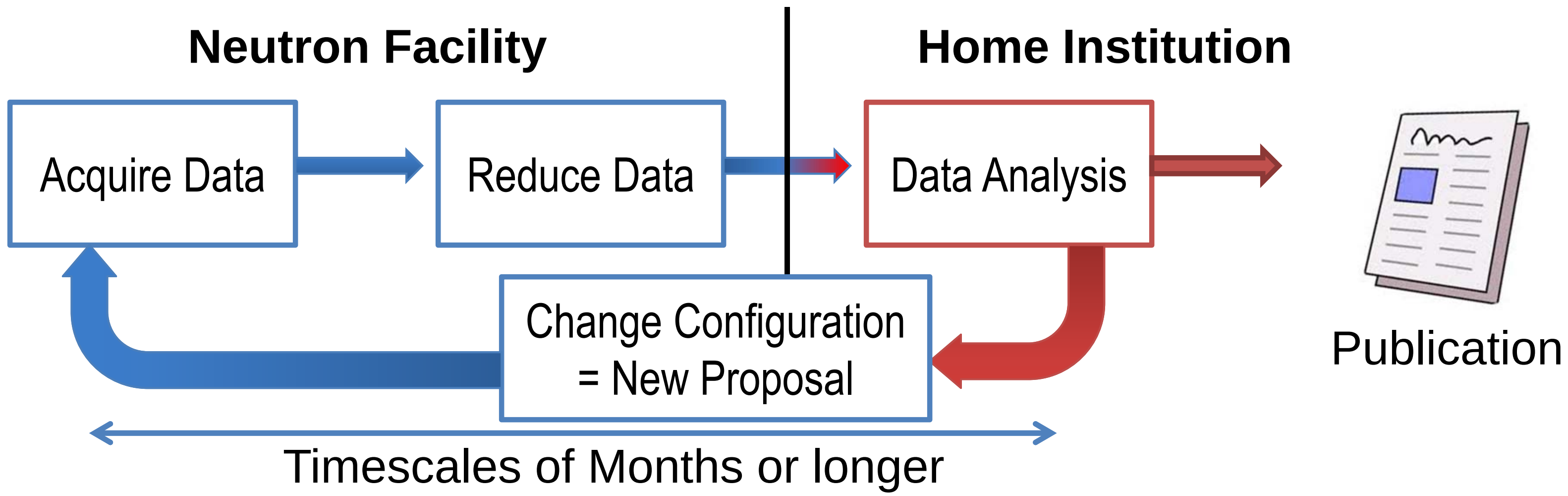
Data captured and stored on multiple systems at the beamline

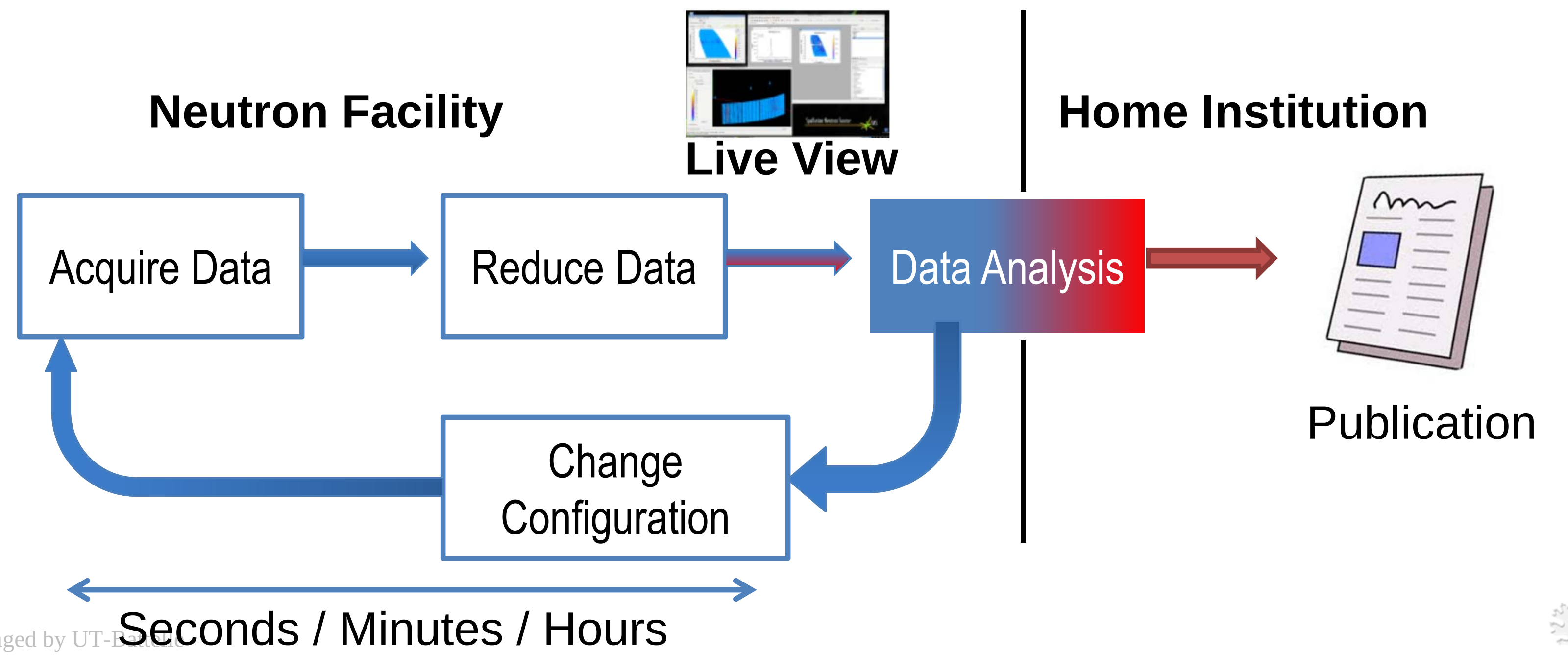
After completion of a "run" data is aggregated on a single system, translation begins

Once data is aggregated reduction begins using a workstation

Analysis and Simulation using mid-scale compute

Improving Productivity = Changing the Workflow

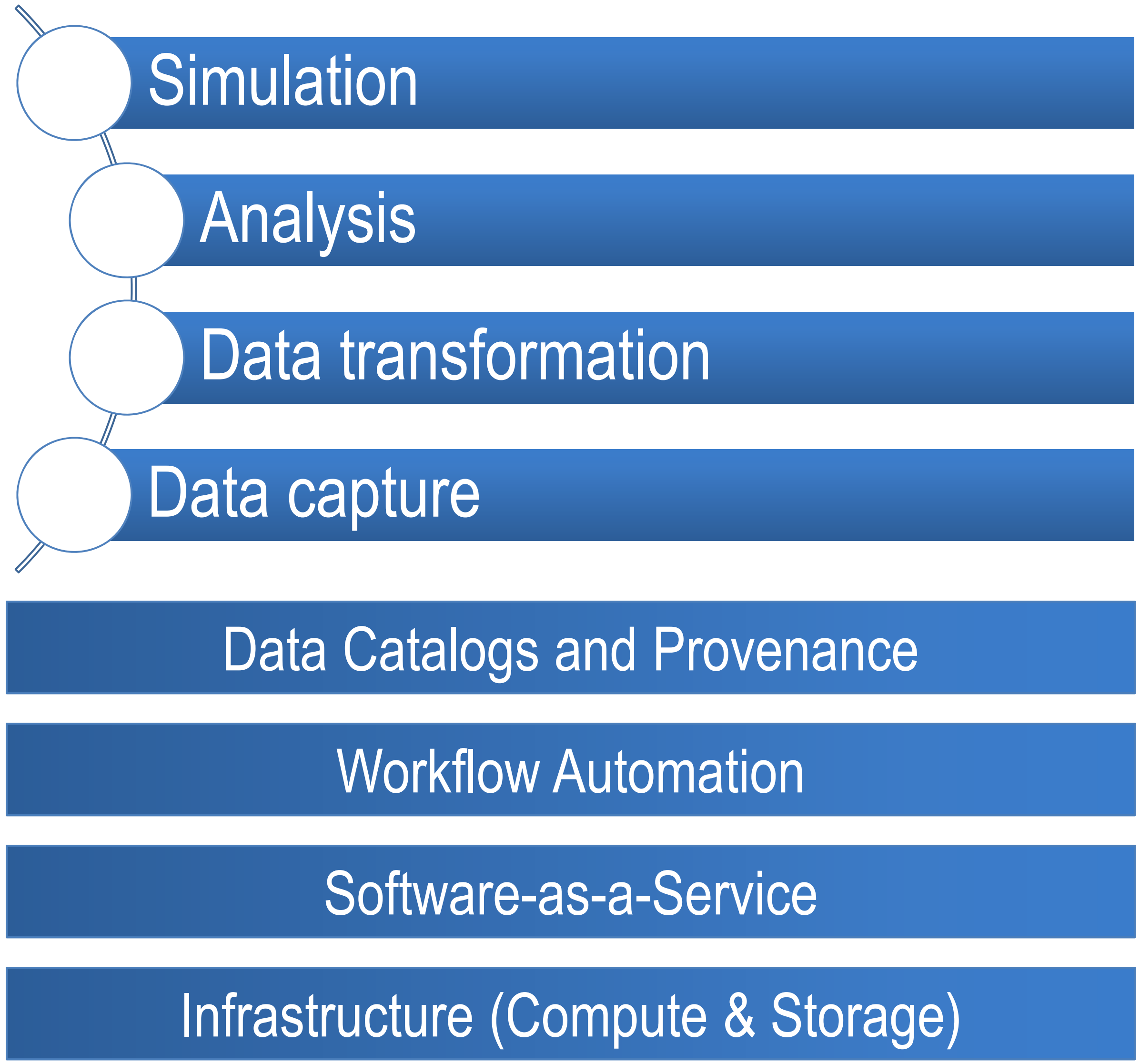




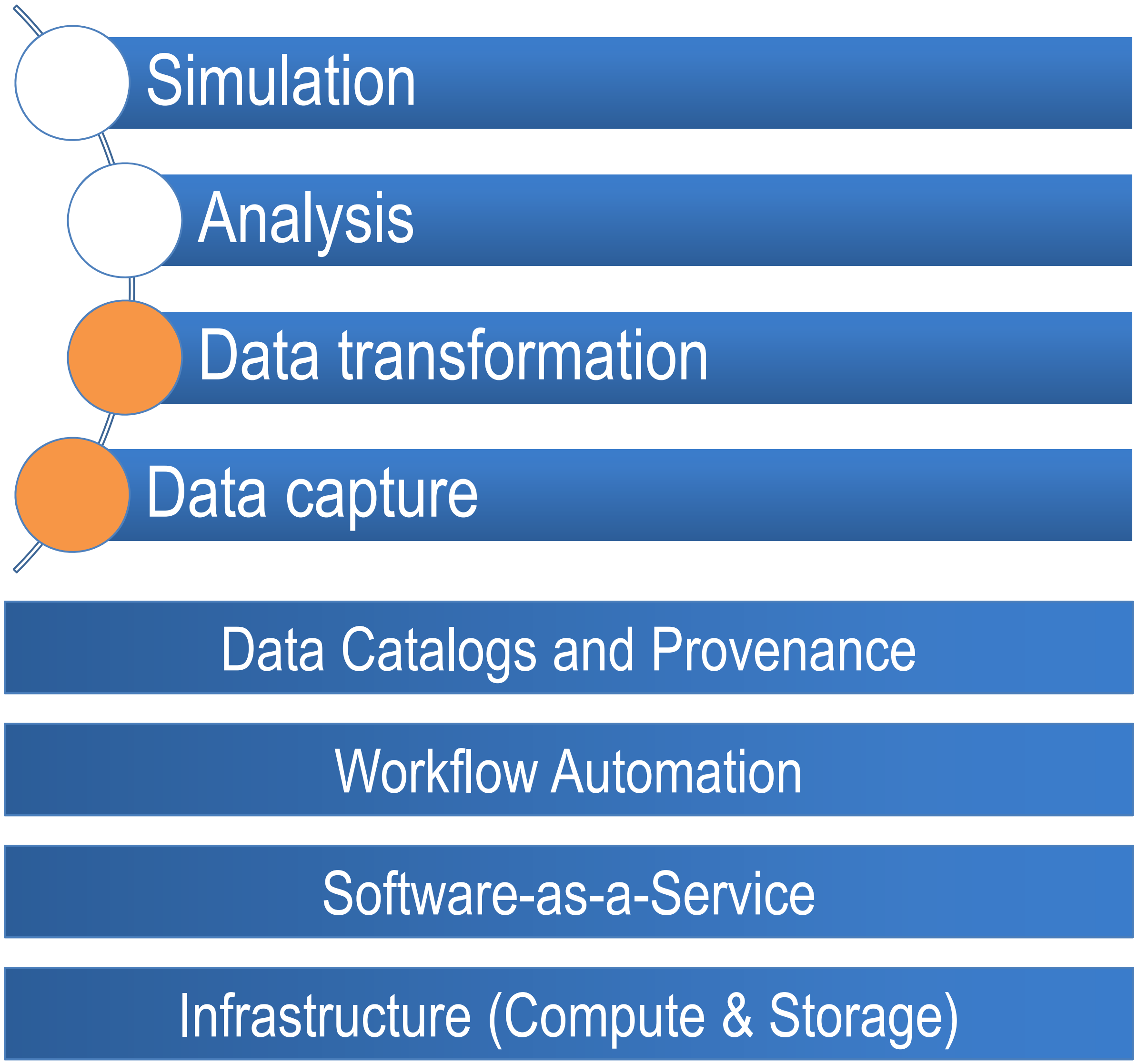
Challenges

- Advanced data infrastructure and the staff for operations
- Data capture and aggregation from a diverse sensors
- Data transformation/reduction in near real-time for live view
- Coupling of advanced analysis/simulation
- Cataloging of experiment results to include provenance
- Remote access to data and compute resources
- Long-term stewardship and public access

A Knowledge Discovery Architecture for Neutron Science



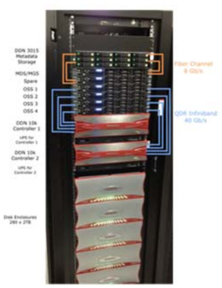
A Knowledge Discovery Architecture for Neutron Science

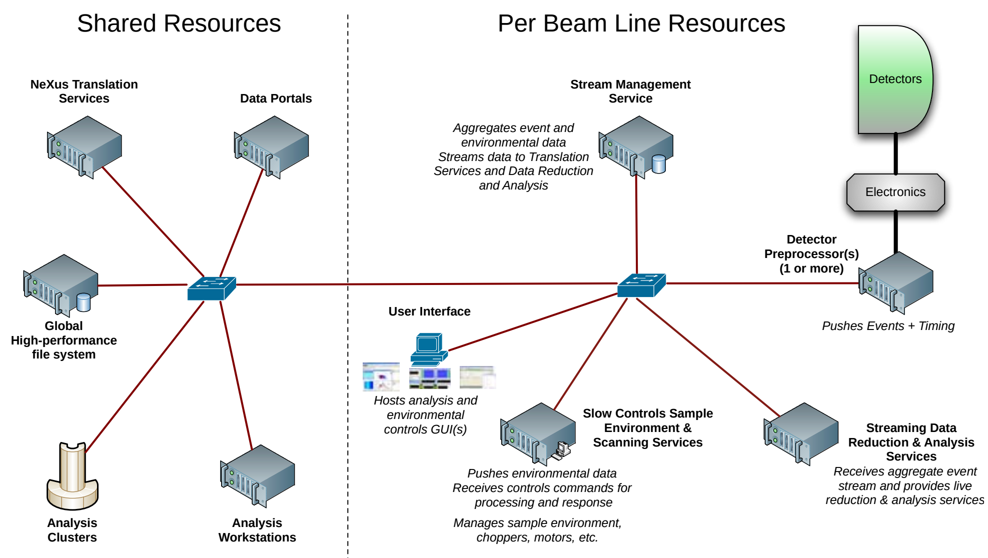


Data Capture and Transformation

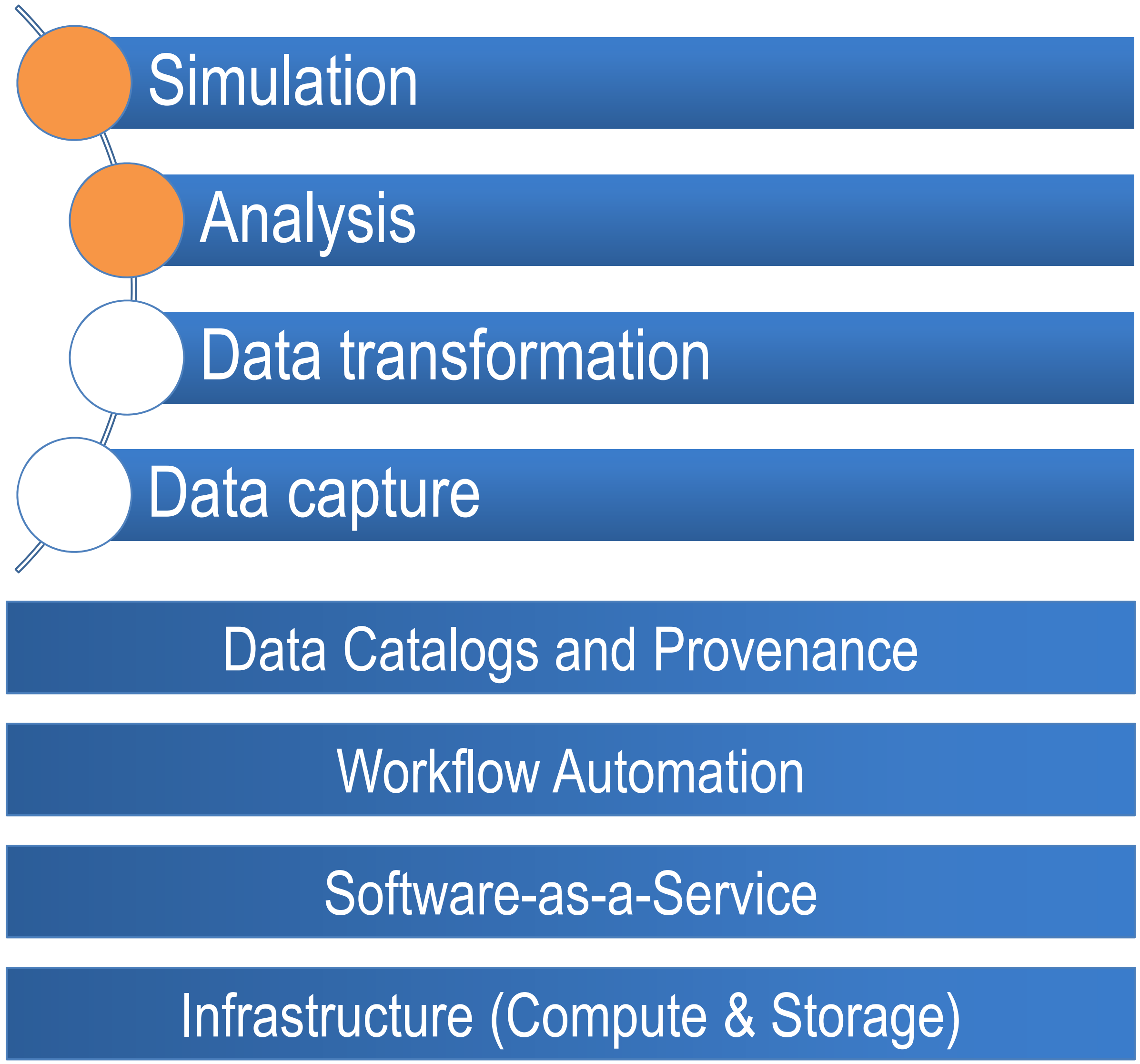
Accelerating Data Acquisition, Reduction, and Analysis at the SNS

- Accelerating Data Acquisition, Reduction, and Analysis at SNS
- We stream data (neutron and SE) from the DAS to a publish subscribe system
  - *Stream Management Service (SMS)*
- We modify MANTID (data reduction) to read from the data stream live from SMS
  - *Streaming Reduction Service (SRS)*
- We re-configure the data translation (file creation) to read the data stream from SMS and create the files while the run is taking place… end of run = close file [file appears "instantly"]
  - *Streaming Translation Service (STS)*
- Files are created on an HPC infrastructure for subsequent parallel analysis and data reduction



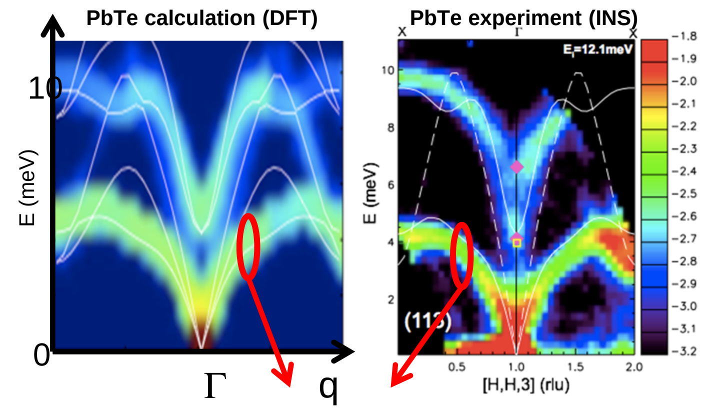


A Knowledge Discovery Architecture for Neutron Science



Data Analysis and Simulation
- Center for Accelerating Materials Modeling

- The CAMM will integrate materials modeling/simulation (MD/DFT) directly into the chain for neutron scattering data analysis, **offline** and **online** (in near real time)
- Developing workflows for refinement, integration of MD codes, **neutron scattering corrections** ..
- The CAMM is working with ORNL's Materials Science and Technology Division to study coarse grained MD simulations of polymers PEO-AA (CNMS), *ab-initio* MD simulations for ferroelectrics/thermoelectrics



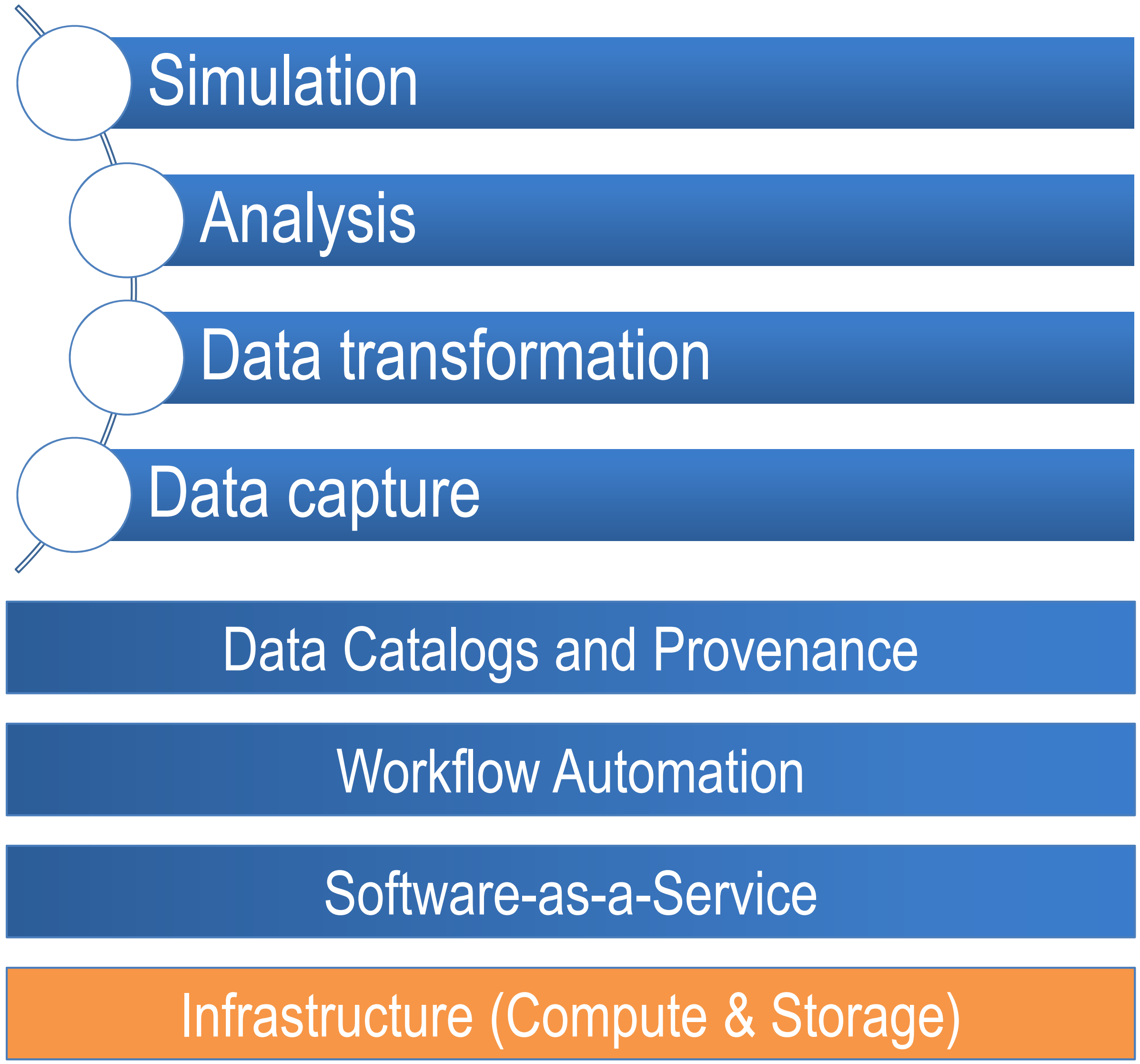
Focus on *width* of dispersions:
- measured line widths need corrections for instrument resolution effects.
- will support development of reliable line width predictions.

Delaire et al., *Nature Materials* (2011).

The Center for Accelerating Materials Modeling (CAMM)

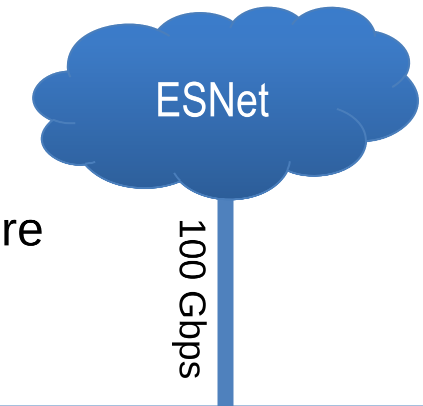
- *Partnership between ORNL's Neutron Sciences, Physical Sciences and Computing and Computational Sciences Directorates*
- *Studying force field refinement from quasi-elastic and inelastic neutron scattering data*
- *CAMM formed in response to BES proposal call for* Predictive Theory and Modeling

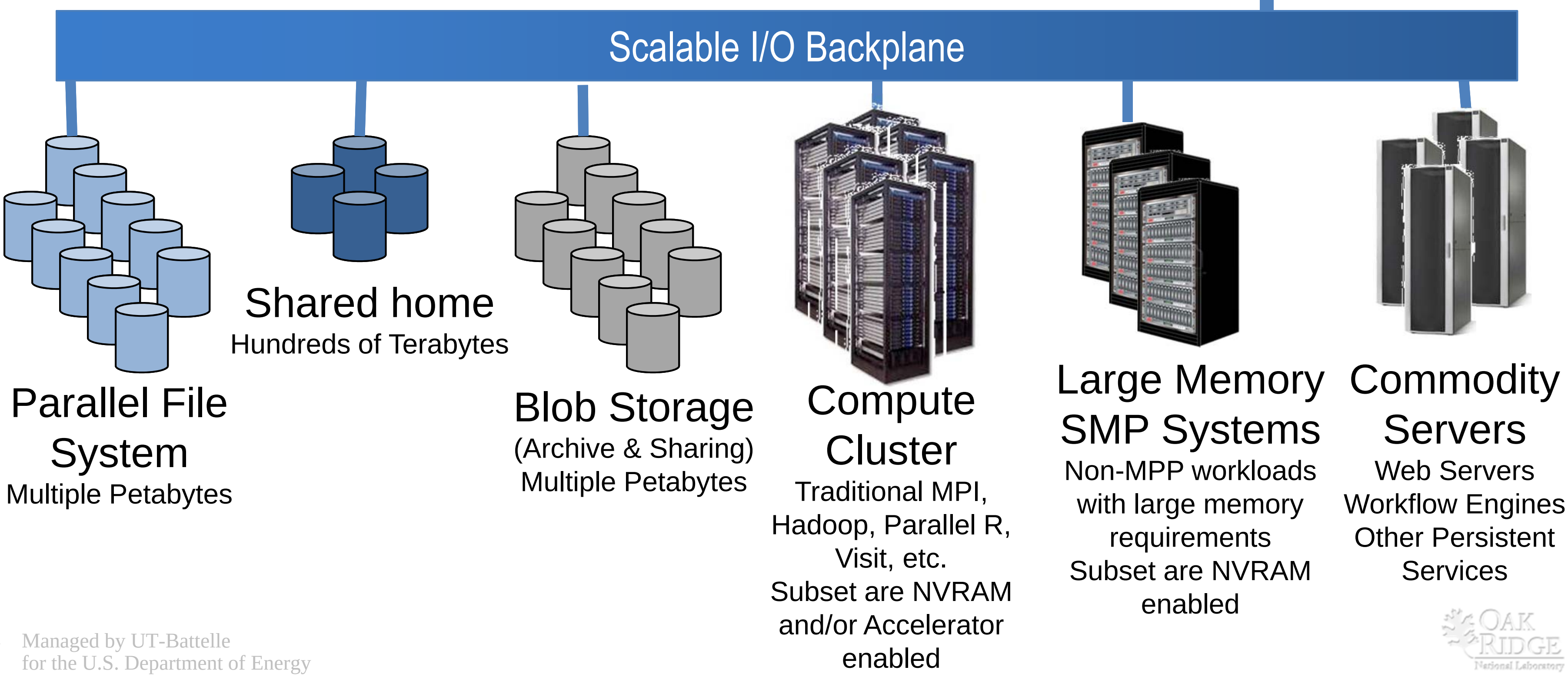
A Knowledge Discovery Architecture for Neutron Science



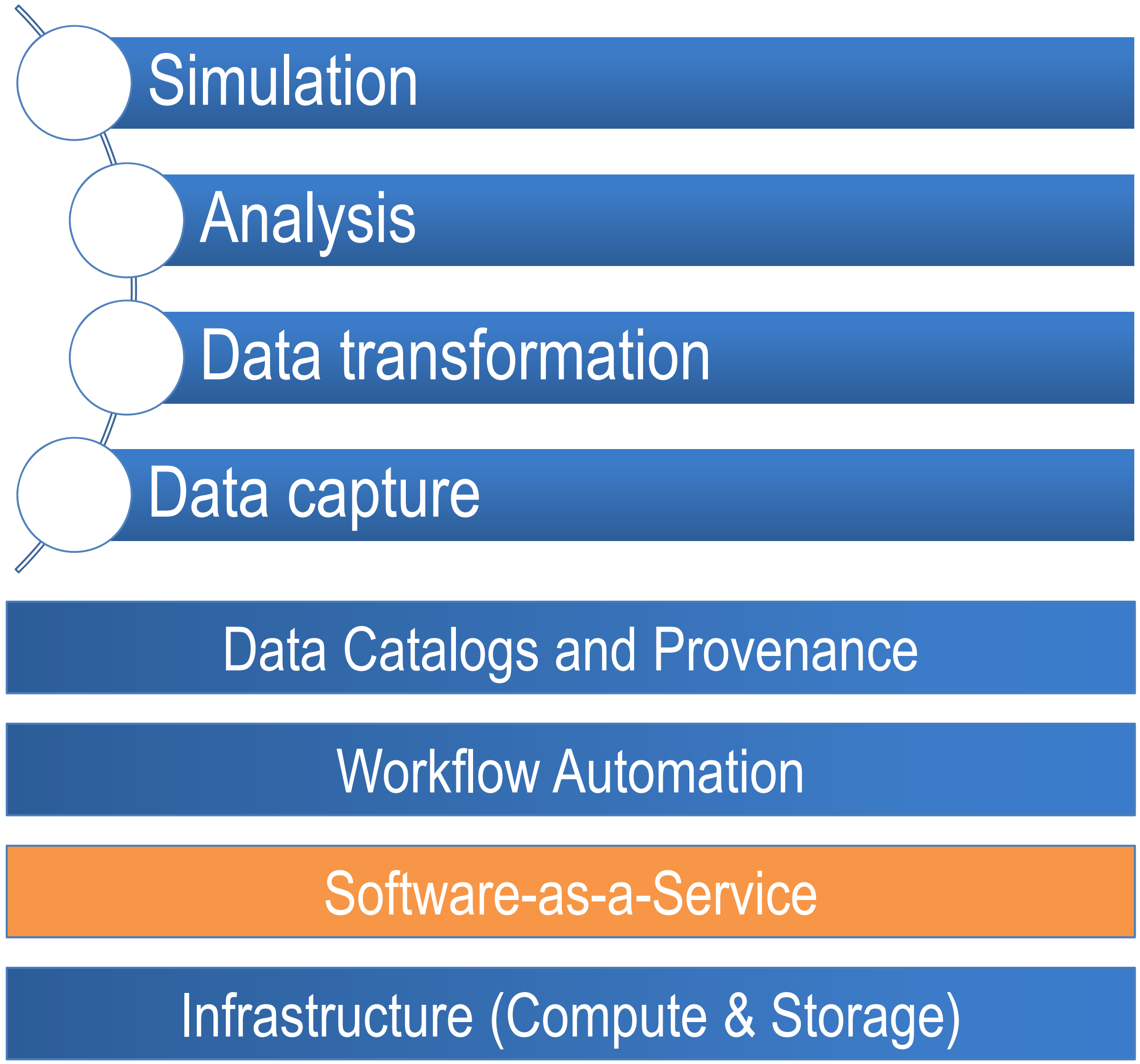
Compute and Data Infrastructure

- Flexible: tailored software stacks coupled with workload appropriate compute
- Elastic: meet point-in-time storage and compute requirements
- Ability to provision persistent services for data capture services, portals, databases, etc.



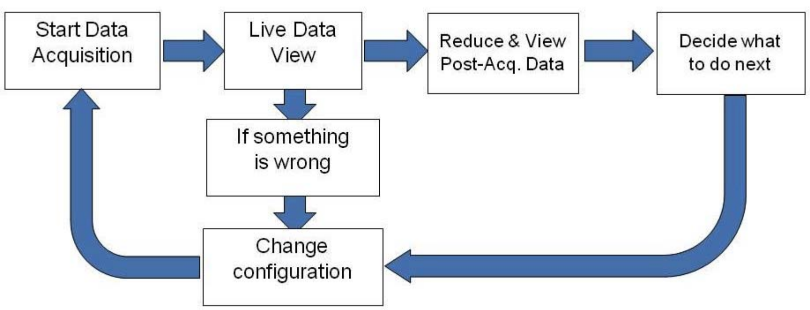


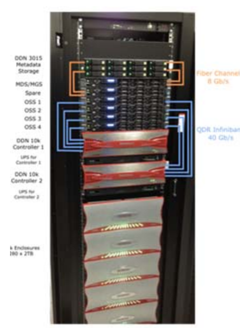
A Knowledge Discovery Architecture for Neutron Science



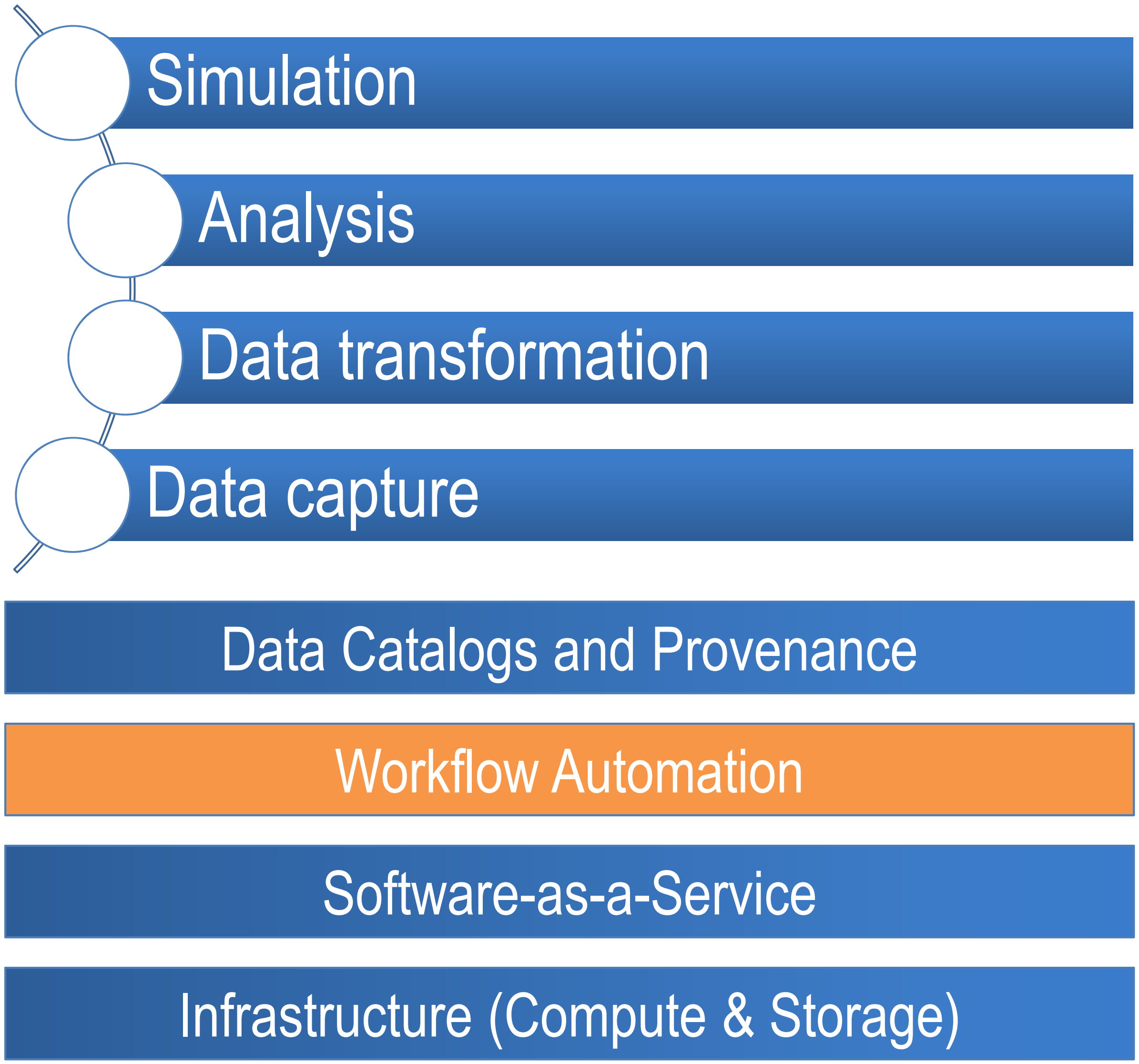
Software-as-a-Service Platform

- A hybrid HPC cluster and software-as-a-service “cloud”
- Full HPC software stack and parallel file system environment
- Unlike traditional HPC, users are insulated from this environment
- Parallel reduction algorithms are configured and ready to run
  - Leverages Mantid in a “Framework” mode and runs this data parallel
- These reductions are available for execution from anywhere via MOAB web-services
- Users continue to interact with Mantid for reduction as they always have on their workstation or via thin-client
  - Mantid executes reduction remotely through MOAB web-services
- Other software services will be required - modeling and simulation



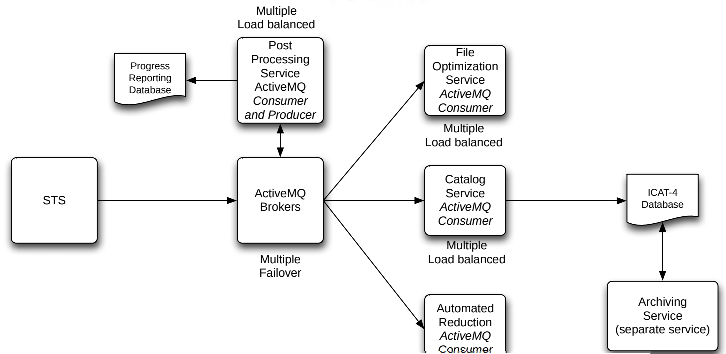


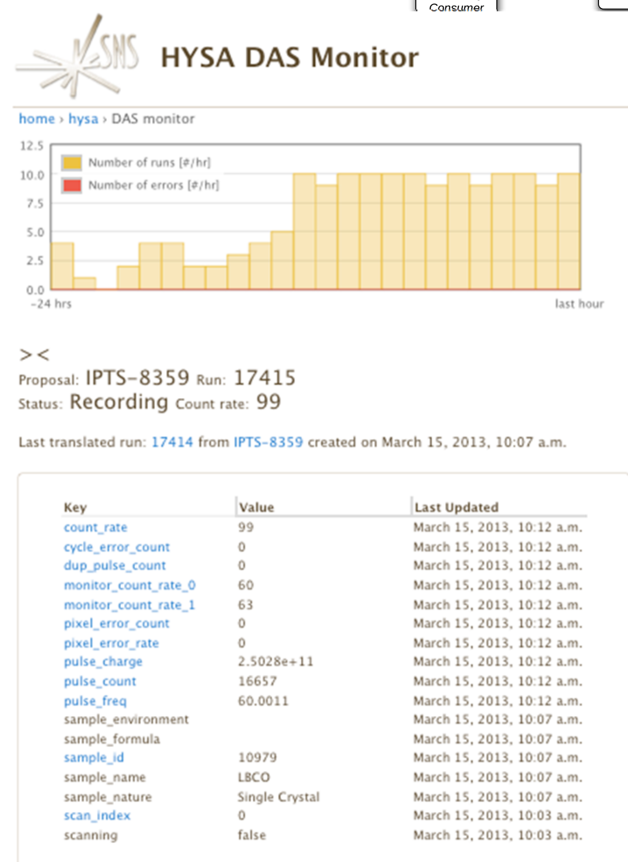
A Knowledge Discovery Architecture for Neutron Science



Workflow Automation

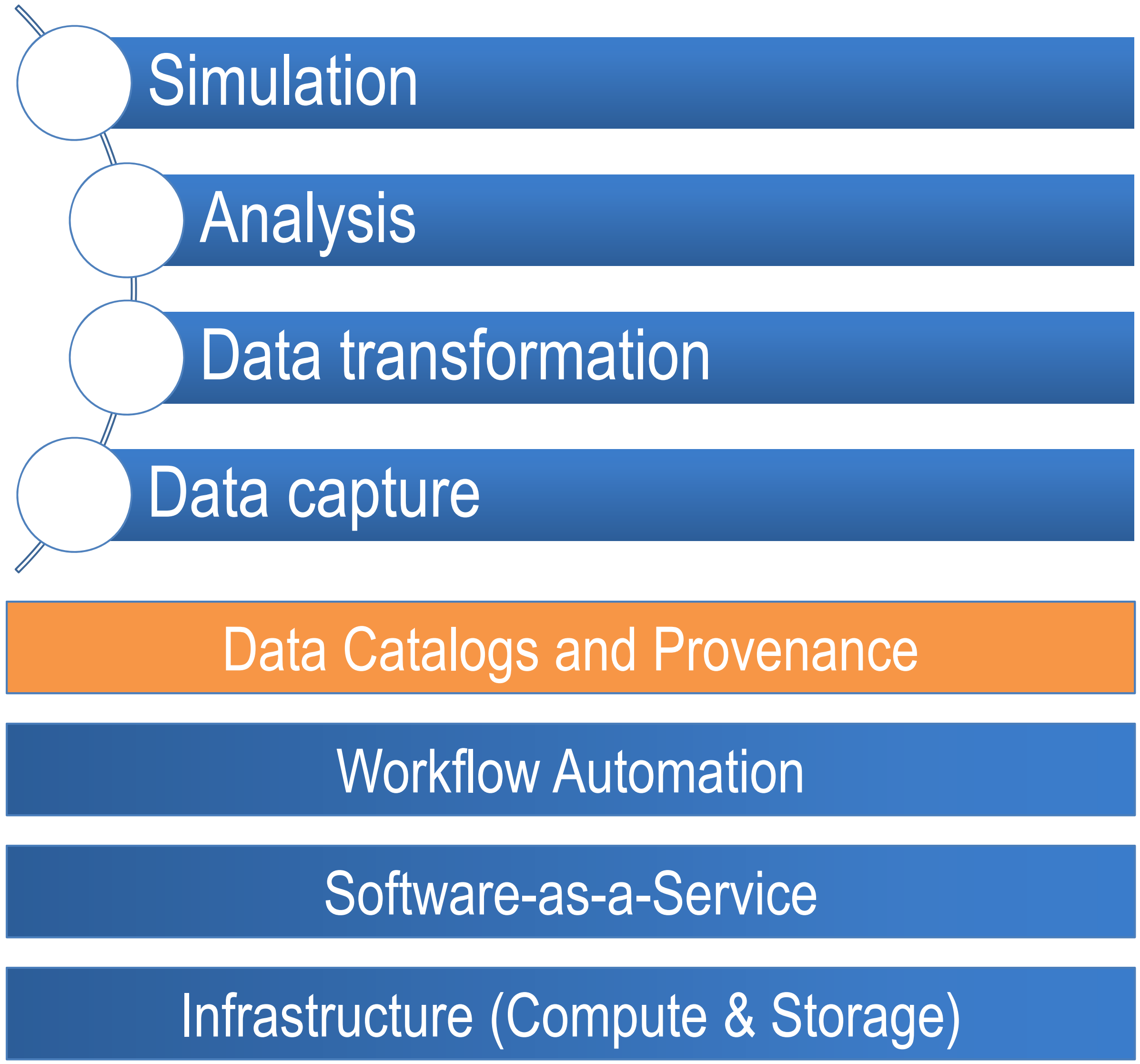
- The Workflow Manager handles data cataloging, automated reduction, and archiving of datasets post acquisition
- Built on Apache ActiveMQ allowing loosely coupled architecture
  - Resilience through durable messages and clustered message brokers
  - Load balancing of tasks across resources through message queues
- Tasks within the workflow are handled by independent processes running on one or more systems
- Each task type maps to an ActiveMQ queue, to activate a task you simply send a message to that queue
- A director process manages the workflow sending messages to these queues
- Workflows are defined and stored within the director's redundant MySQL system





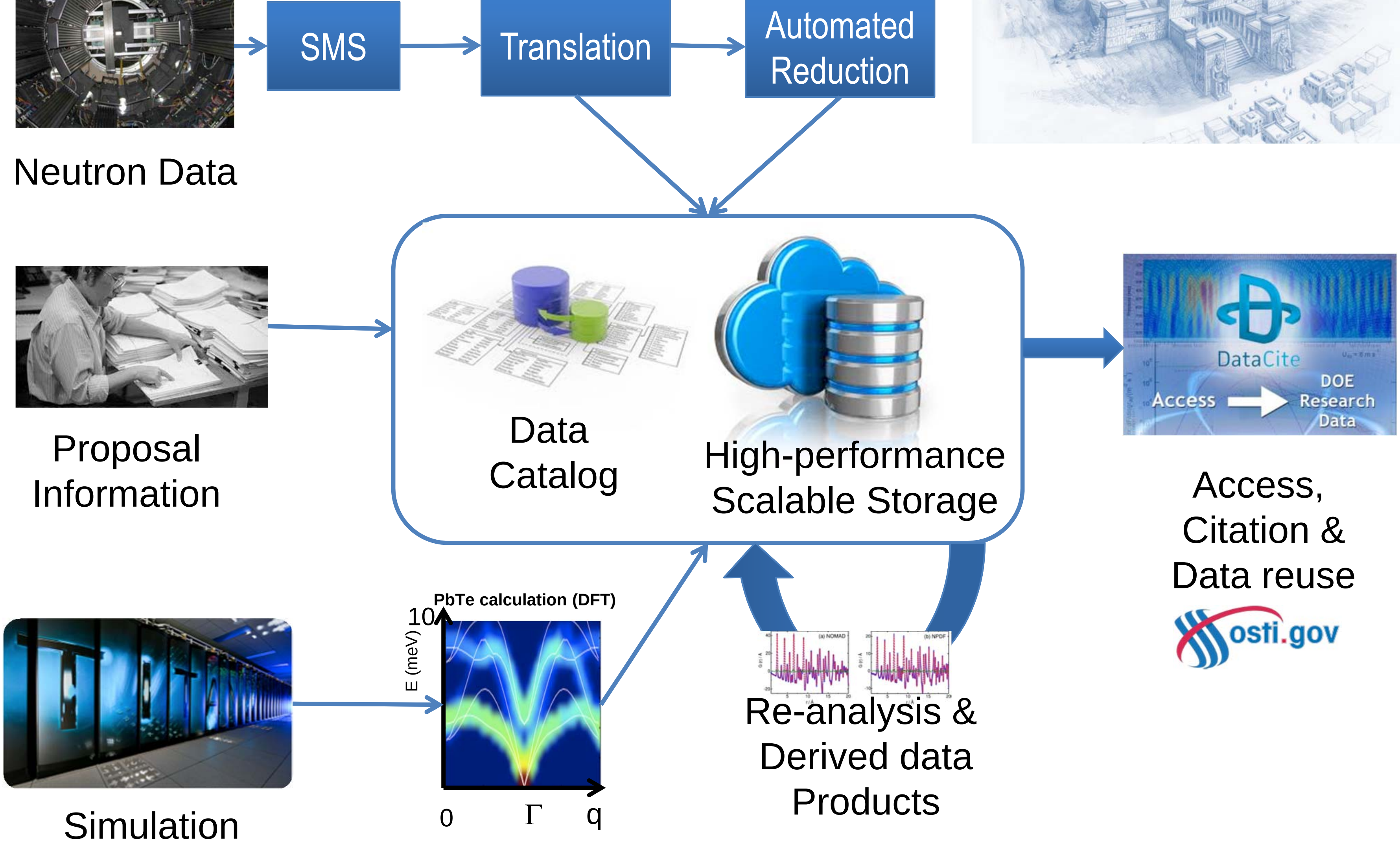
Key	Value	Last Updated
count_rate	99	March 15, 2013, 10:12 a.m.
cycle_error_count	0	March 15, 2013, 10:12 a.m.
dup_pulse_count	0	March 15, 2013, 10:12 a.m.
monitor_count_rate_0	60	March 15, 2013, 10:12 a.m.
monitor_count_rate_1	63	March 15, 2013, 10:12 a.m.
pixel_error_count	0	March 15, 2013, 10:12 a.m.
pixel_error_rate	0	March 15, 2013, 10:12 a.m.
pulse_charge	2.5028e+11	March 15, 2013, 10:12 a.m.
pulse_count	16657	March 15, 2013, 10:12 a.m.
pulse_freq	60.0011	March 15, 2013, 10:12 a.m.
sample_environment		March 15, 2013, 10:07 a.m.
sample_formula		March 15, 2013, 10:07 a.m.
sample_id	10979	March 15, 2013, 10:07 a.m.
sample_name	LBCO	March 15, 2013, 10:07 a.m.
sample_nature	Single Crystal	March 15, 2013, 10:07 a.m.
scan_index	0	March 15, 2013, 10:03 a.m.
scanning	false	March 15, 2013, 10:03 a.m.

A Knowledge Discovery Architecture for Neutron Science



OAK RIDGE National Laboratory

Data Catalogs and Provenance



Galen Shipman
gshipman@ornl.gov

OAK RIDGE
National Laboratory