Oak Ridge National Laboratory

HPC and AI Convergence in Edge-to-Exascale Science Infrastructures

Arjun Shankar, Ph.D.

Section Head, Advanced Technologies

National Center for Computational Sciences (NCCS)/
Oak Ridge Leadership Computing Facility (OLCF)

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

U.S. DEPARTMENT OF ENERGY

Acknowledgments: This work was sponsored by and used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility at the Oak Ridge National Laboratory supported by the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Agenda

- HPC and AI Convergence Design Patterns
- ORNL Deployment Vignette
- Integrated Research Infrastructures (IRI) – An Emerging DOE Activity

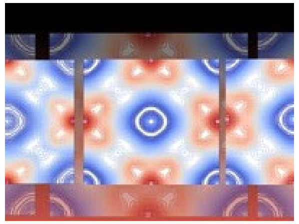
Center for Accelerated Application Readiness (CAAR) Projects

Code: Cholla (Astrophysics)
PI: Evan Schneider, University of Pittsburgh

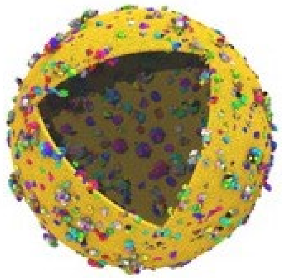
Code: CoMet (Combinatorial Metrics)
PI: Daniel Jacobson, Oak Ridge National Laboratory

Code: GESTS (GPUs for Extreme-Scale Turbulence Simulations)
PI: P. K. Yeung, Georgia Institute of Technology

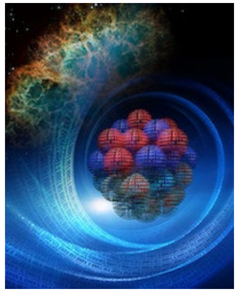
Code: LBPM (Lattice Boltzmann Methods for Porous Media)
PI: James Edward McClure, Virginia Polytechnic



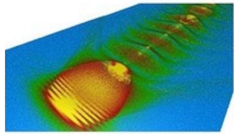
Code: LSMS(Locally-Selfconsistent Multiple Scattering)
PI: Markus Eisenbach, Oak Ridge National Laboratory



Code: NAMD (Nanoscale Molecular Dynamics)
PI: Emad Tajkhorshid, University of Illinois at Urbana-Champaign

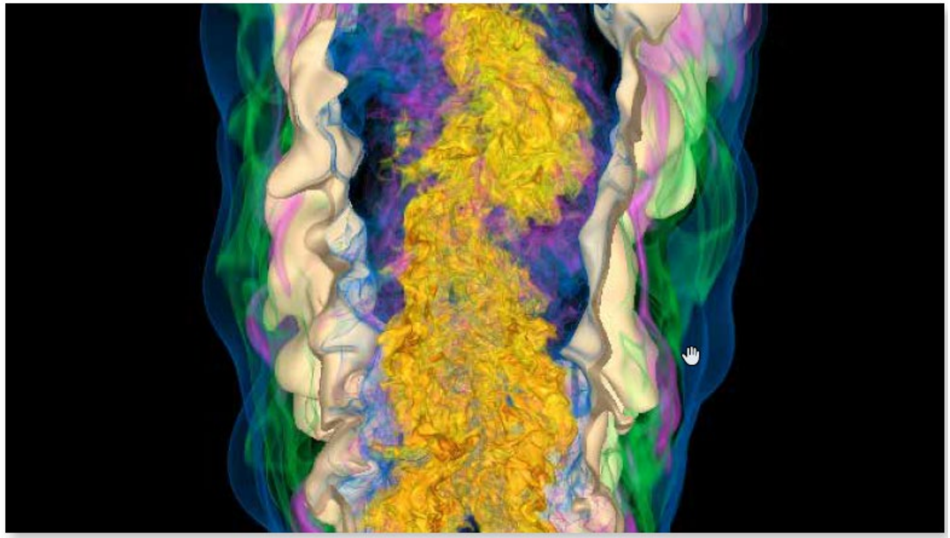


Code: NuCCOR (Nuclear Coupled-Cluster Oak Ridge)
PI: Morten Hjorth-Jensen, Michigan State University

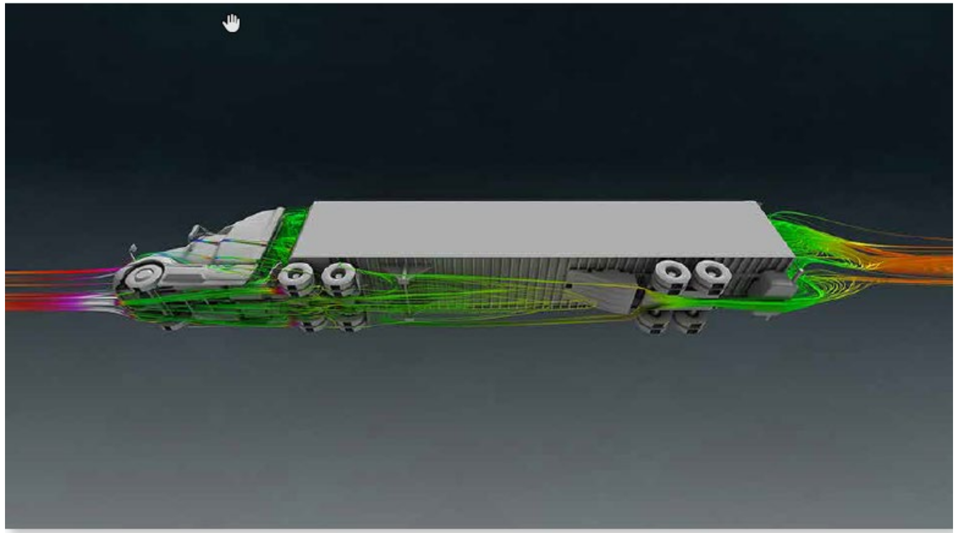


Code: PIConGPU (Particle-in-cell on Graphics Processing Units)
PI: Sunita Chandrasekaran, University of Delaware

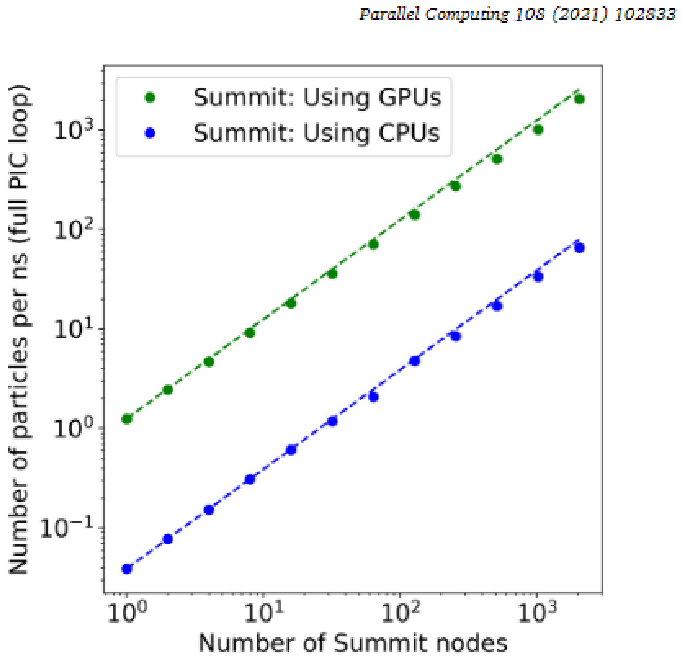
Application Scaling using Accelerators



A visualization of the mixture fraction isosurface (gold) and HO_2, which shows autoignition occurring in fuel-lean mixtures at regions with high temperatures, low mixing rates, and short ignition delay times. Image credit: Hongfeng Yu, University of Nebraska; and Kwan-Liu Ma, University of California, Davis



Supercomputing simulations at ORNL enabled SmartTruck engineers to develop the UnderTray System, some components of which are shown here. The system dramatically reduces drag—and increases fuel mileage—in long-haul trucks. Image credit: Michael Matheson, ORNL

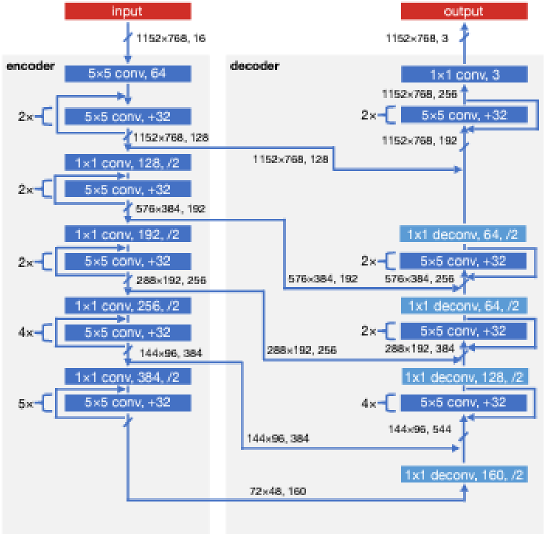


Porting WarpX to GPU-accelerated platforms, A. Myers, et al., https://doi.org/10.1016/j.parco.2021.102833

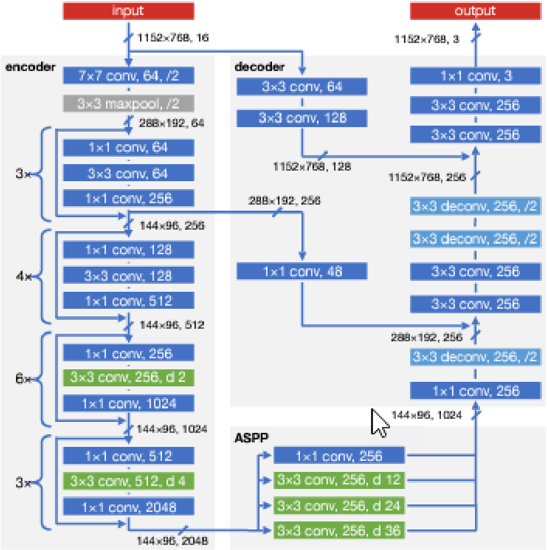
Also, see 25 years of OLCF: https://www.youtube.com/watch?v=CDfANp9ZE9k

Advent of Machine Learning and Deep Learning

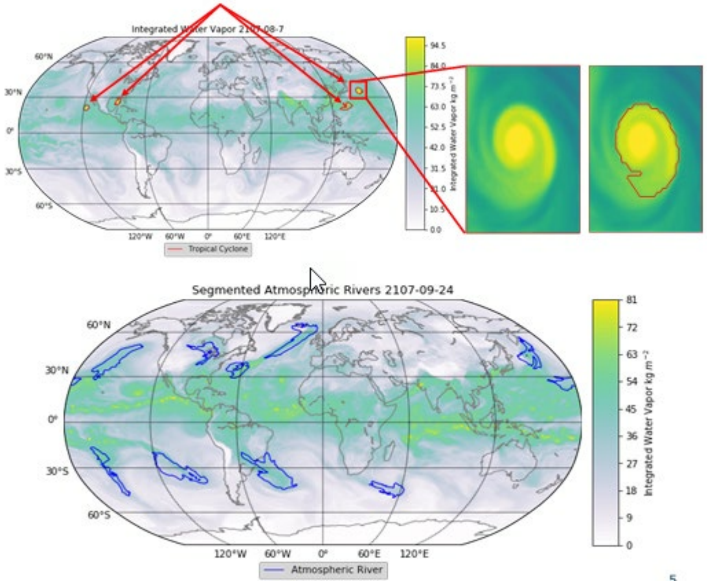
Deep Learning Models for Extreme Weather Segmentation



Tiramisu, 35 layers,
7.8M parameters, 4.2 TF/sample



DeepLabv3+, 66 layers,
43.7M parameters, 14.4 TF/sample

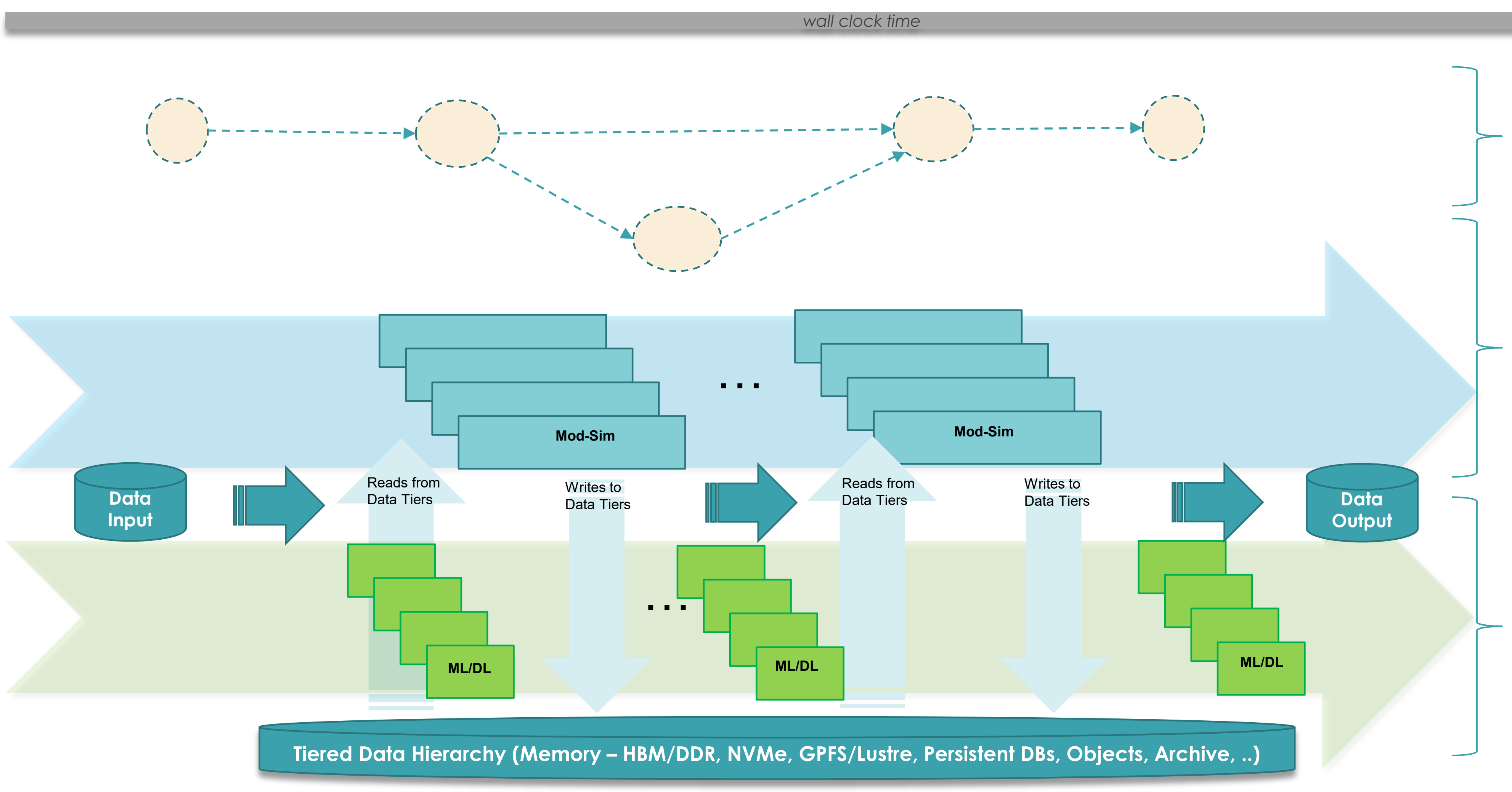


Dataset Size	Required BW (27K GPUs)	GPFS/LUSTRE	BurstBuffer	NVM/e or DRAM
20 TB (~63K samples)	3.8 TB/s	~400 GB/s	~2 TB/s	~26 TB/s

Exascale Deep Learning for Climate Analytics, T. Kurth, et al., Supercomputing 2018, Gordon Bell prize winner;
Table from J. Romero, SOS 2019
Peak performance: 1.13 ExaOps (mixed precision)

Design Pattern for Interleaving ML/DL/AI and Simulation

Design Pattern: Interleaved Mod-Sim + ML/DL/AI at Scale



Workflow: Overlay of Intra/Inter-Facility System workflows help connect various phases of the data lifecycle.

Forward Problems: HPC Simulations, sparse, dense solvers, partitioners, mixed precision, etc.

Analysis and Inverse Problems: ML/DL – proxy/surrogates, un/semi/supervised, steering, in-situ state change and viz, repartitioning, mixed precision, etc.

Example: Surrogates in an HPC Simulation Application
MC-based Exploration of High-Entropy Alloy System (MoNbTaW)

- Probability of N atoms in configuration X at temperature T follows Boltzmann's distribution $exp(-E(X)/k_B T)$ where E is the total configuration energy and k_B is the Boltzmann constant.

- Replica of the alloy systems at various T are simulated via replica exchange Monte Carlo simulations with transition probability between replica m and n:

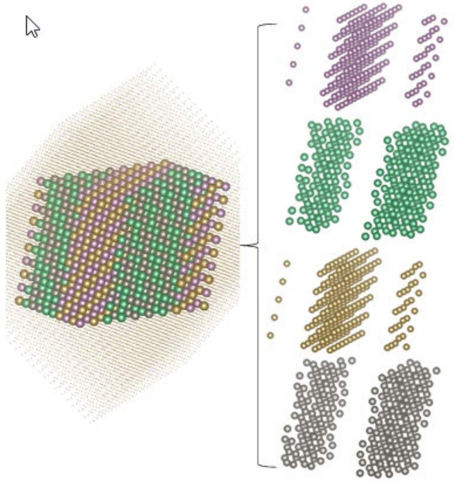
$$W(\{X_m, T_m\}|\{X_n, T_n\}) = \min[1, \exp(-\triangle)],$$

where

$$\triangle = (1/k_B T_n - 1/k_B T_m)(E(X_m) - E(X_n))$$

- Atoms (i,j) are exchanged with acceptance probability $P_{i,j}$ proportional to:

$$\min[1, \exp(-(E(\{x_i, x_j\}) - E(\{x_j, x_i\}))/k_B T)]$$



https://doi.org/10.1038/s43588-021-00139-3

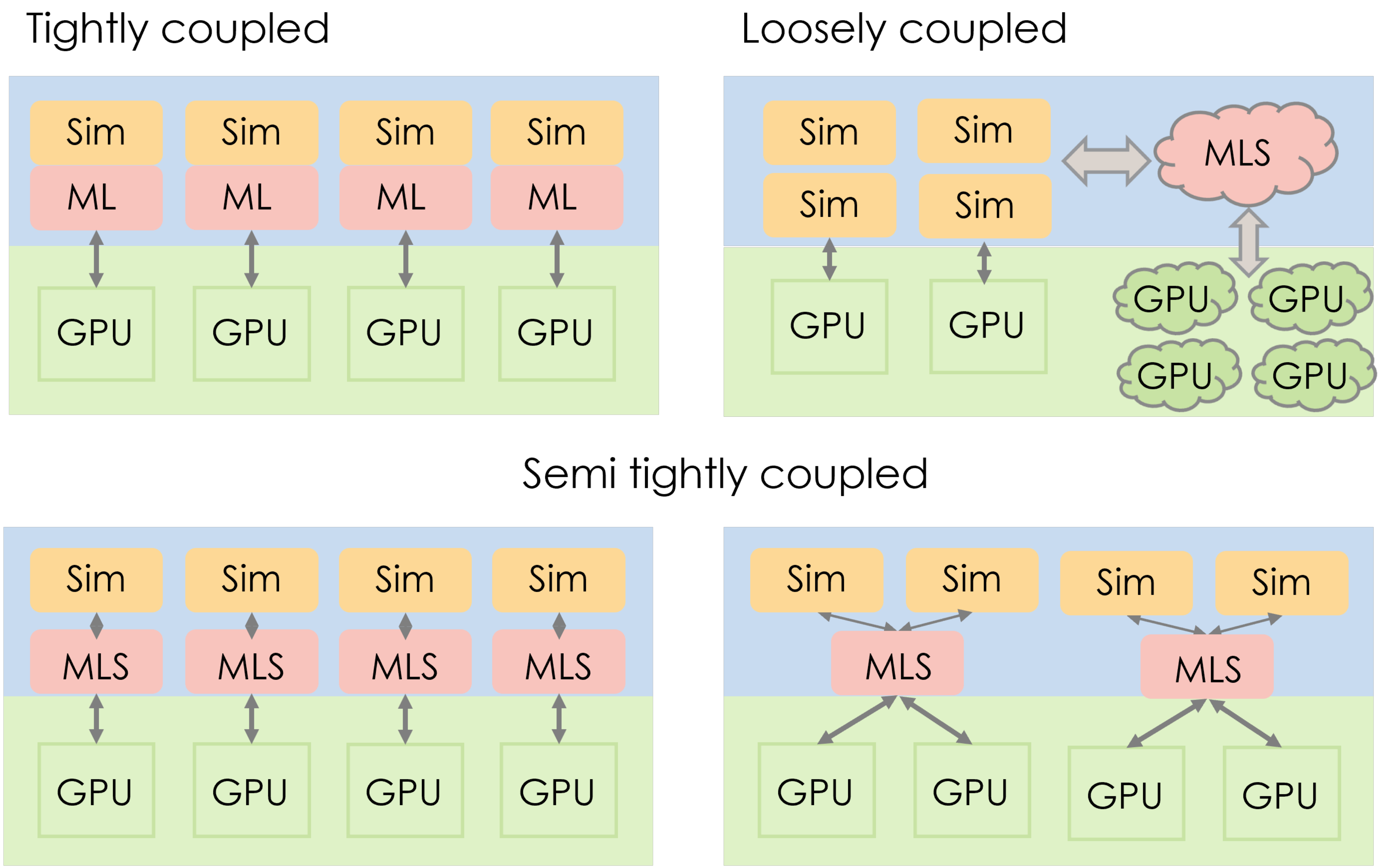
DEEP LEARNING SURROGATE MODELS FOR THE ENERGY EVALUATION OF MONBTAW ALLOY. THE ARCHITECTURE IS GIVEN BY THE NUMBER OF NODES IN EACH HIDDEN LAYER, AND IT IS THE SAME FOR EACH OF THE FOUR ELEMENT. THE R^2 SCORE IS THE AVERAGED MODEL PERFORMANCE OF ALL ELEMENTS.

name	architecture	# parameters	R^2 score FP32	R^2 score mixed
tiny	24-24	5,257	0.991	0.988
small	128x2-64x2 -24x2	55,817	0.991	0.992
mdedium	200x11	440,801	0.994	0.992
large	512x6-256x6 -128x6-64x6-32	2,019,009	0.994	0.993

J. Yin, F. Wang, A. Shankar, Strategies for Integrating Deep Learning Surrogate Models with HPC Simulation Applications, ExSAIS 2022.

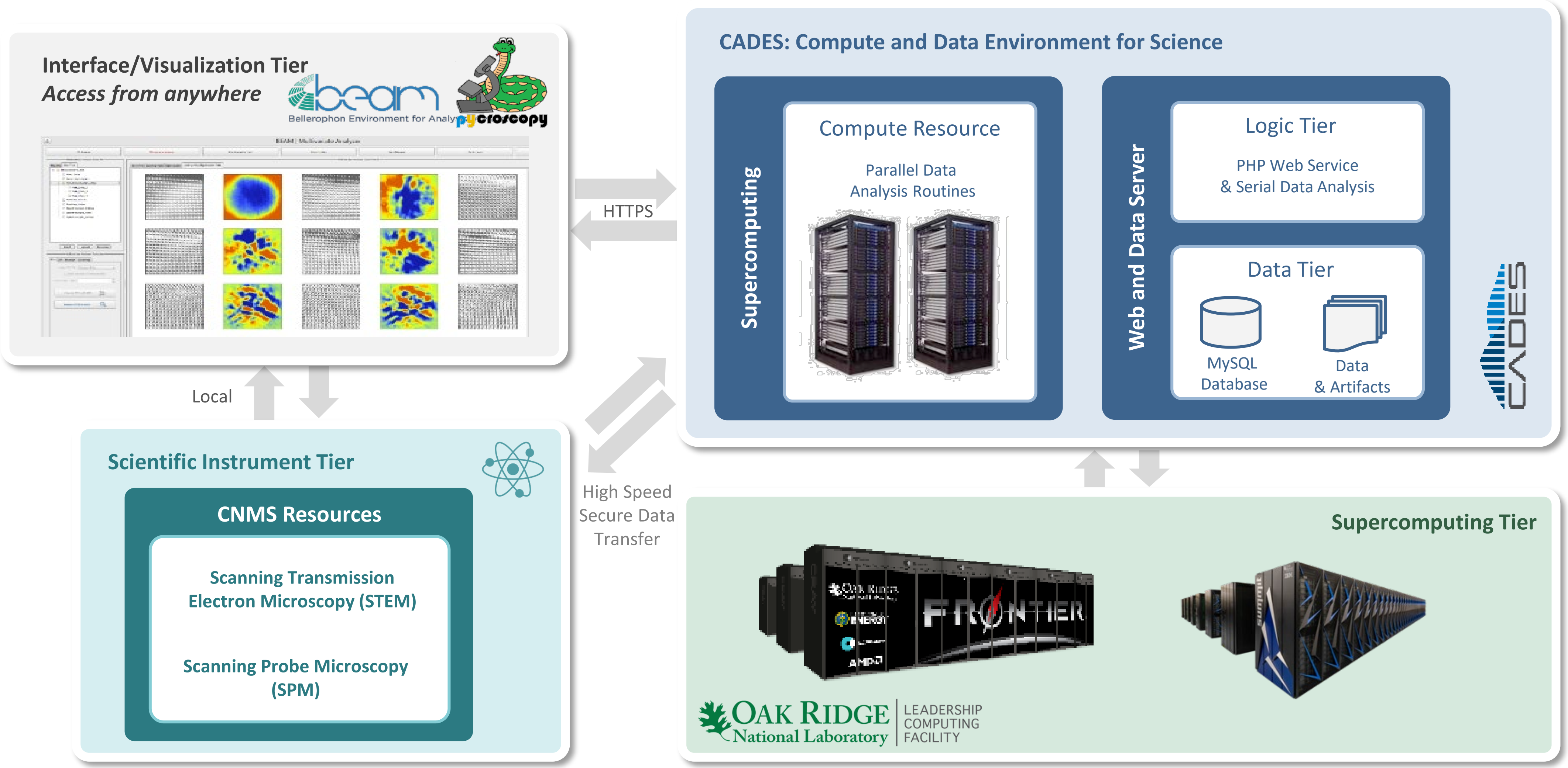
OAK RIDGE National Laboratory

Deployment Patterns of Converged HPC + AI



ORNL Facility Vignettes of Workflows Enabling HPC-AI Convergence

Cross-Facility (ex *situ*) Workflows and Data Science *Cloud + HPC*



Lingerfelt, E., et al., *Procedia Computer Science* **80**, 2276-2280 (2016).

Example going Cross-Facility: Forward/Design of Experiment

Theory

- Model Formulation

Implementation

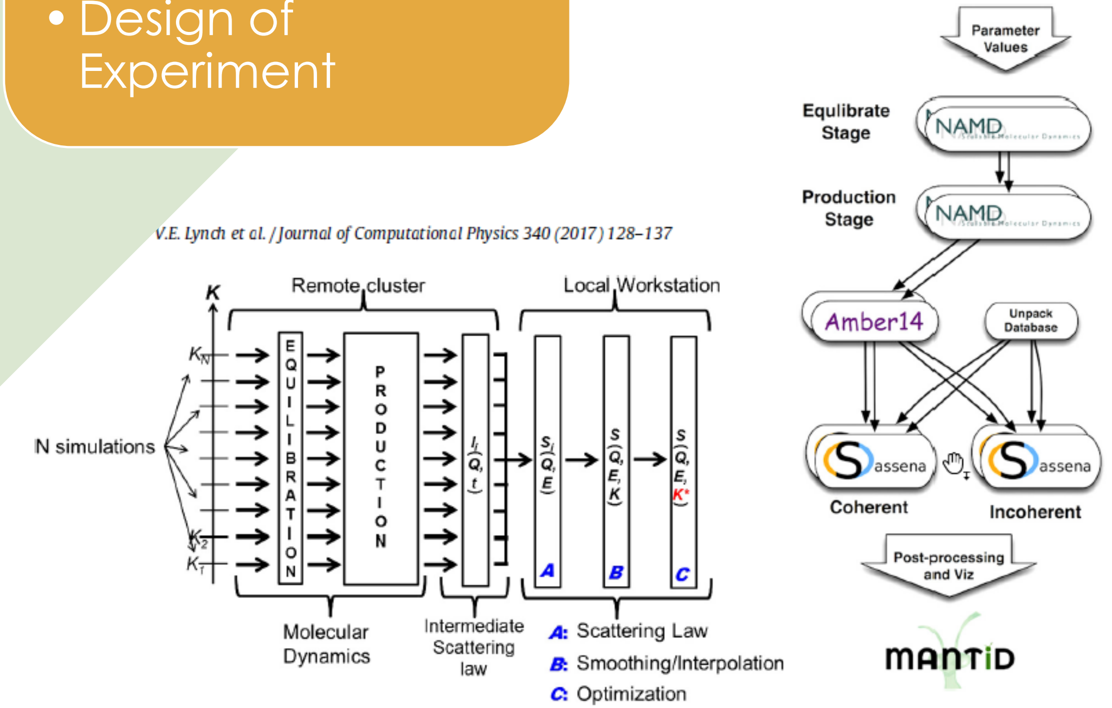
- Algorithm and Software

Execution

- Large amounts of Generated Data
- Design of Experiment

$$I(Q,t) = \left\langle \frac{1}{N_{at}} \sum_{i=1}^{N_{at}} (b_i^{inc})^2 e^{i\vec{Q}[\vec{r}_i(t_0+t)-\vec{r}_i(t_0)]} \right\rangle_{\Omega_{\vec{Q}},t_0}$$

$$S_{sim}(Q,E) = e^{\left(\frac{E}{2k_bT}\right)} \frac{1}{\hbar} \int_{-\infty}^{\infty} e^{-i\frac{Et}{\hbar}} I(Q,t)dt$$



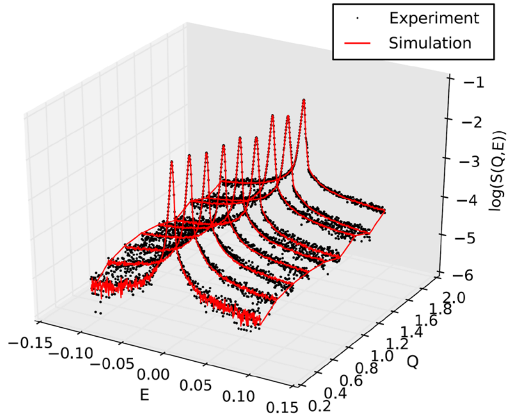
Lynch et al., Journal of Computational Physics 340 (2017) 128–137
Dhindsa, Bhowmik, et al., J. of Physical Chemistry B 2016, 100059-10068

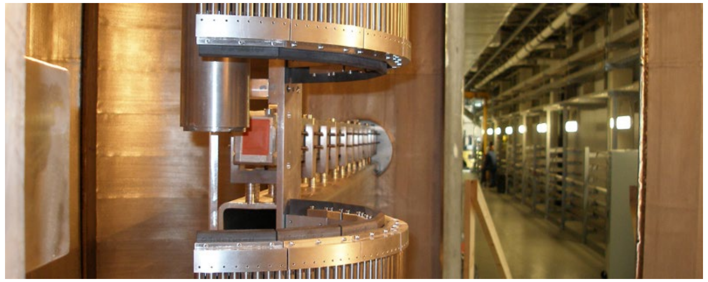
Example going Cross-Facility: ML/DL Problems

Validate/Disprove and Calibrate
- Data-Analytics/ML
- Fit

Align, Reduce, Transfer
- Workflows

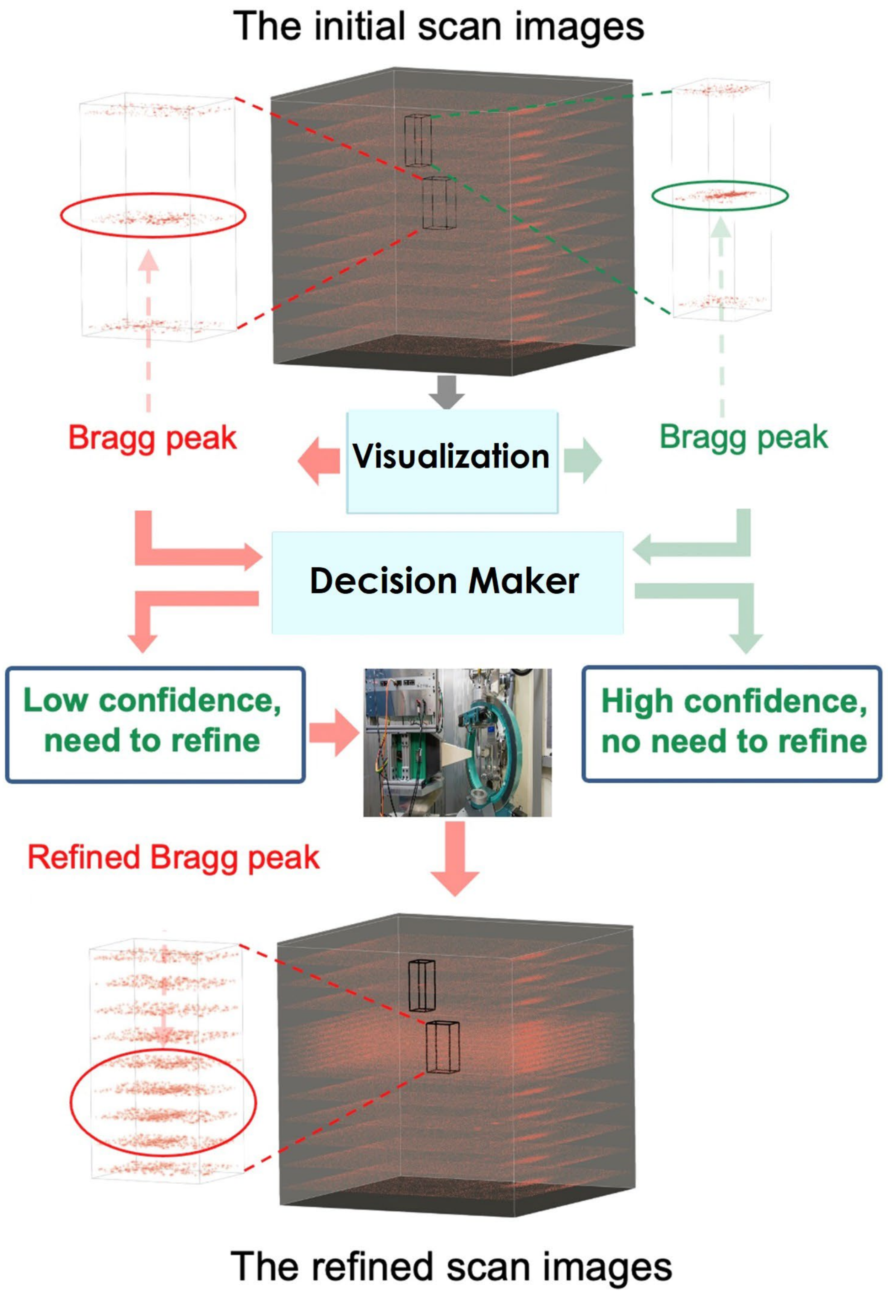
Design of/and Experiment
- Collect Large amounts of Data

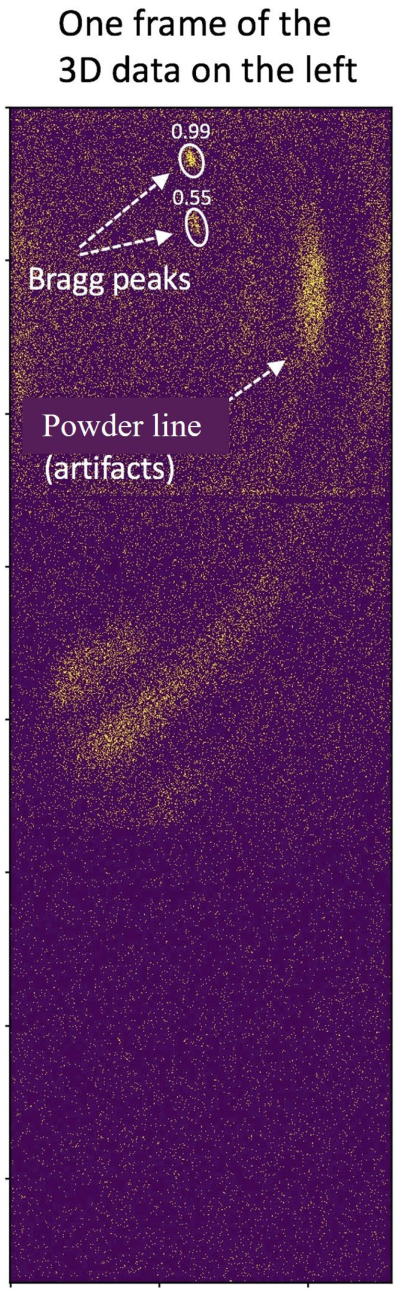




OAK RIDGE National Laboratory

Cross-Facility – SNS to OLCF: Bragg Peak Detection Workflow

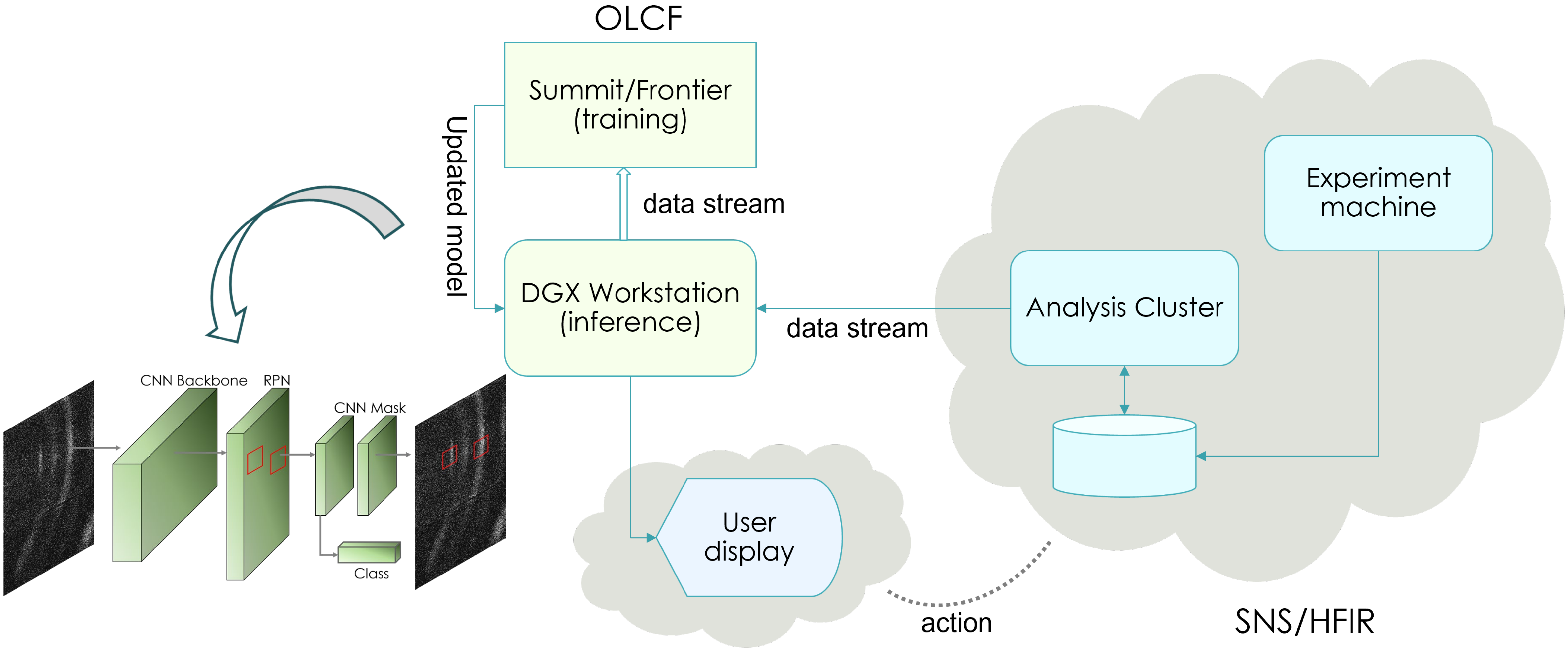




Toward an Autonomous Workflow for Single Crystal Neutron Diffraction
Submitted, May 2022

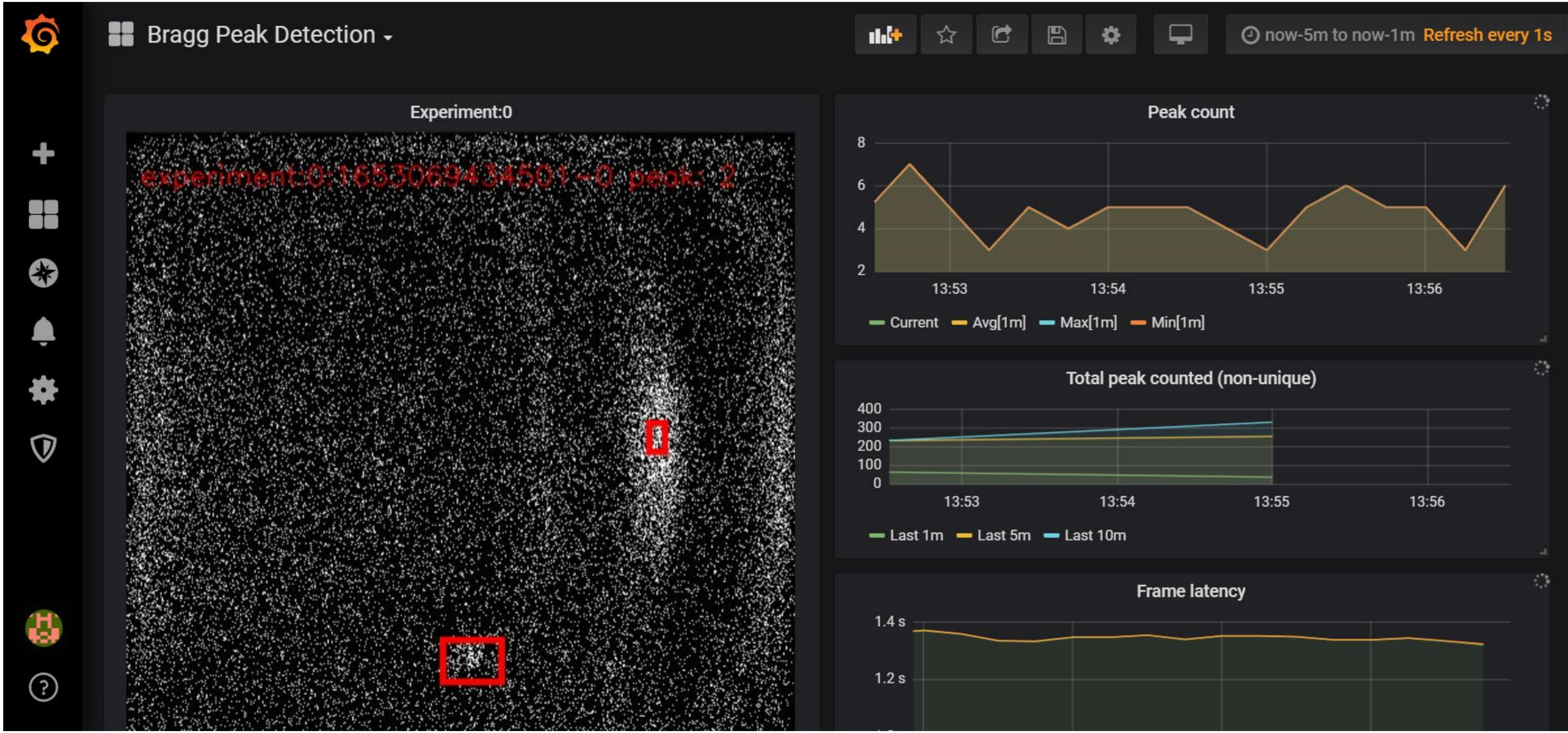
- Junqi Yin,
- Guannan Zhang,
- Huibo Cao,
- Sajal Dash,
- Bryan C. Chakoumakos,
- Feiyi Wang

Using Edge-to-Exascale/Converged AI Pipelines



Toward an Autonomous Workflow for Single Crystal Neutron Diffraction, *Submitted, May 2022, J. Yin, G. Zhang, H. Cao, B. Chakoumakos, and F. Wang*

Scientist Display



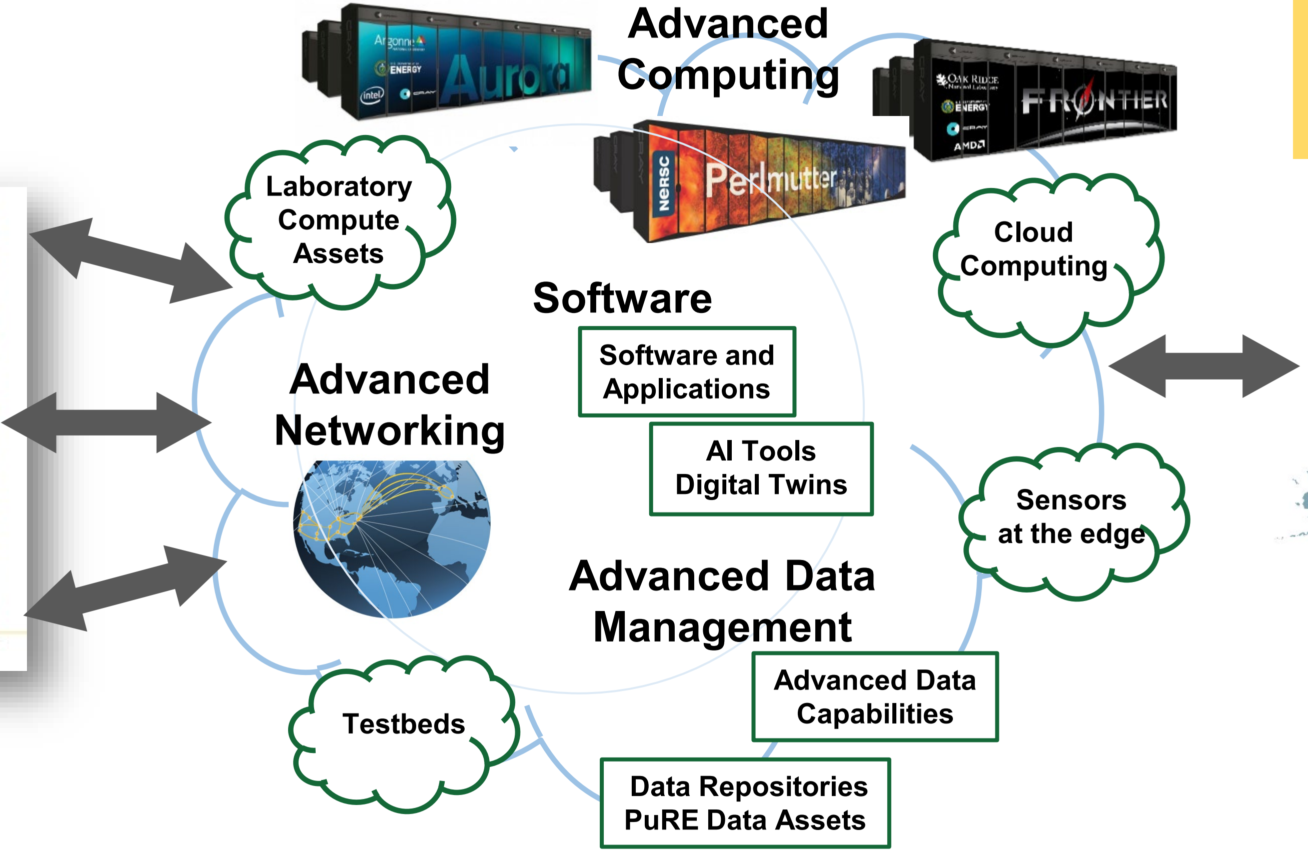
Toward an Autonomous Workflow for Single Crystal Neutron Diffraction, *Submitted, May 2022, J. Yin, G. Zhang, H. Cao, B. Chakoumakos, and F. Wang*

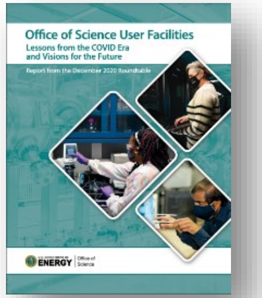
Oak Ridge National Laboratory

DOE Initiative for Integrated Research Infrastructures (IRI)

The vision: A DOE/SC integrated research ecosystem that transforms science via seamless interoperability

Strategic goal: Broadening participation





New modes of integrated science

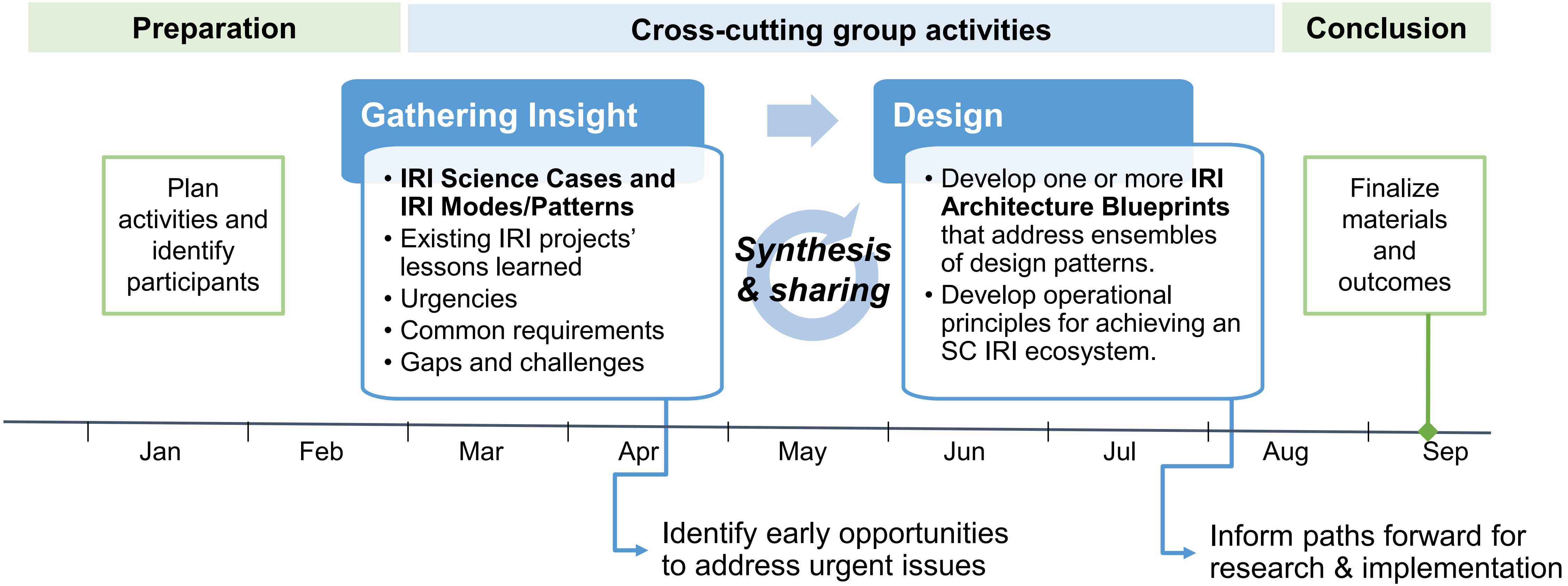
Rapid data analysis and steering of experiments	Novel models for multi-facility allocation/utilization	AI-enabled insight from dynamic, vast multi-modal data	Seamless user interconnectivity via federated IDs

U.S. DEPARTMENT OF ENERGY | Office of Science

Slide credit: IRI-ABA Activity: Design Phase Kickoff Slide Presentation, 6/14/2022, Ben Brown and Bill Miller

SC IRI Architecture Blueprint Activity, FY 2022

Goal: Produce the **reference conceptual foundations** to inform a coordinated "whole-of-SC" strategy for an integrative research ecosystem.



Common recurring sentiments across the user interviews

Data Management

- Users are overwhelmed with large and growing amounts of data to manage, reduce, analyze
  - Users need to move data across facilities and use different systems at different steps of data processing chain
  - Users need bespoke data movement and workflow solutions, and long-duration support for data/metadata.

Automation/AI

- Users need reliable automation & seamless access and try to compensate via human effort.
  - Users need automation, and anticipate AI, but struggle with skills and application of these novel technologies

Heterogeneity

- Users face mismatches between resources, tools and needs
  - Users need heterogeneity in scale and type of resources but have platform fatigue learning many different platforms
  - Users need workflows to be at the center but need software APIs and standardization/uniformity
  - Users have a spectrum of computing needs from elastic computing (matching need to available resources) to urgent computing (near real-time/just-in-time, on-demand)

Ease of Use

- Users find infrastructure hard to use
  - Users encounter a lack of transparency about workflow tools and resources, and many different use policies and cybersecurity barriers
  - Users need infrastructure to be easier to use and be more coordinated across resources and facilities

Workforce Skills Gap

- Users and teams struggle with workforce and training needs
  - Users (and their organizations) struggle with lack of skills, oversubscribed staff, recruiting, and retention
  - Users experience gaps between their working knowledge and skills and those of infrastructure experts.
  - Users need support and expertise in data science.

Slide credit: IRI-ABA Activity: Design Phase Kickoff Slide Presentation, 6/14/2022, Ben Brown and Bill Miller

Forging the Future: Large-Scale Scientific Ecosystems

We will evolve to deploying middleware and workflow systems to provide **abstraction** and **automation** for describing complex computational applications that require efficient and robust management of **large volumes of data** on **high-performance compute** resources.

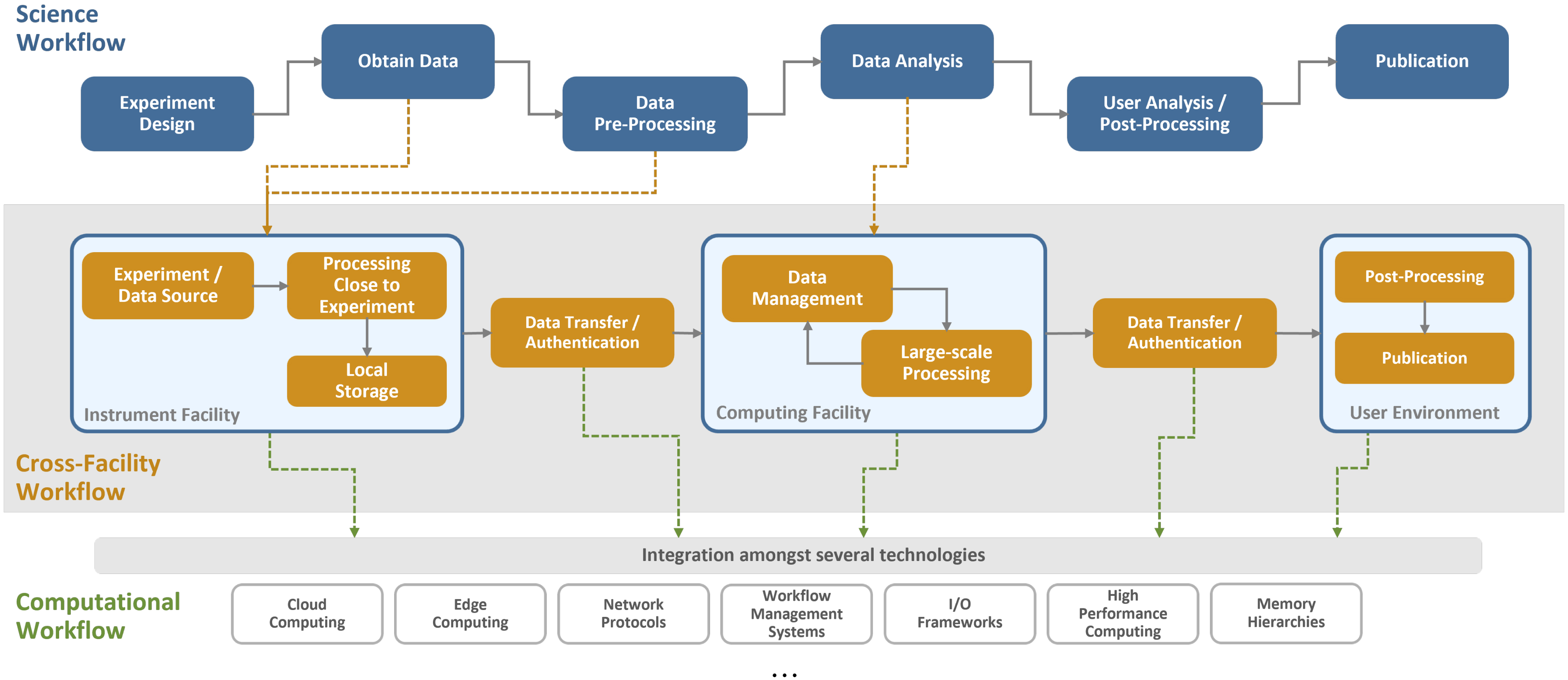


Diagram credit: Rafael Ferreira da Silva

Thank you!

Questions?