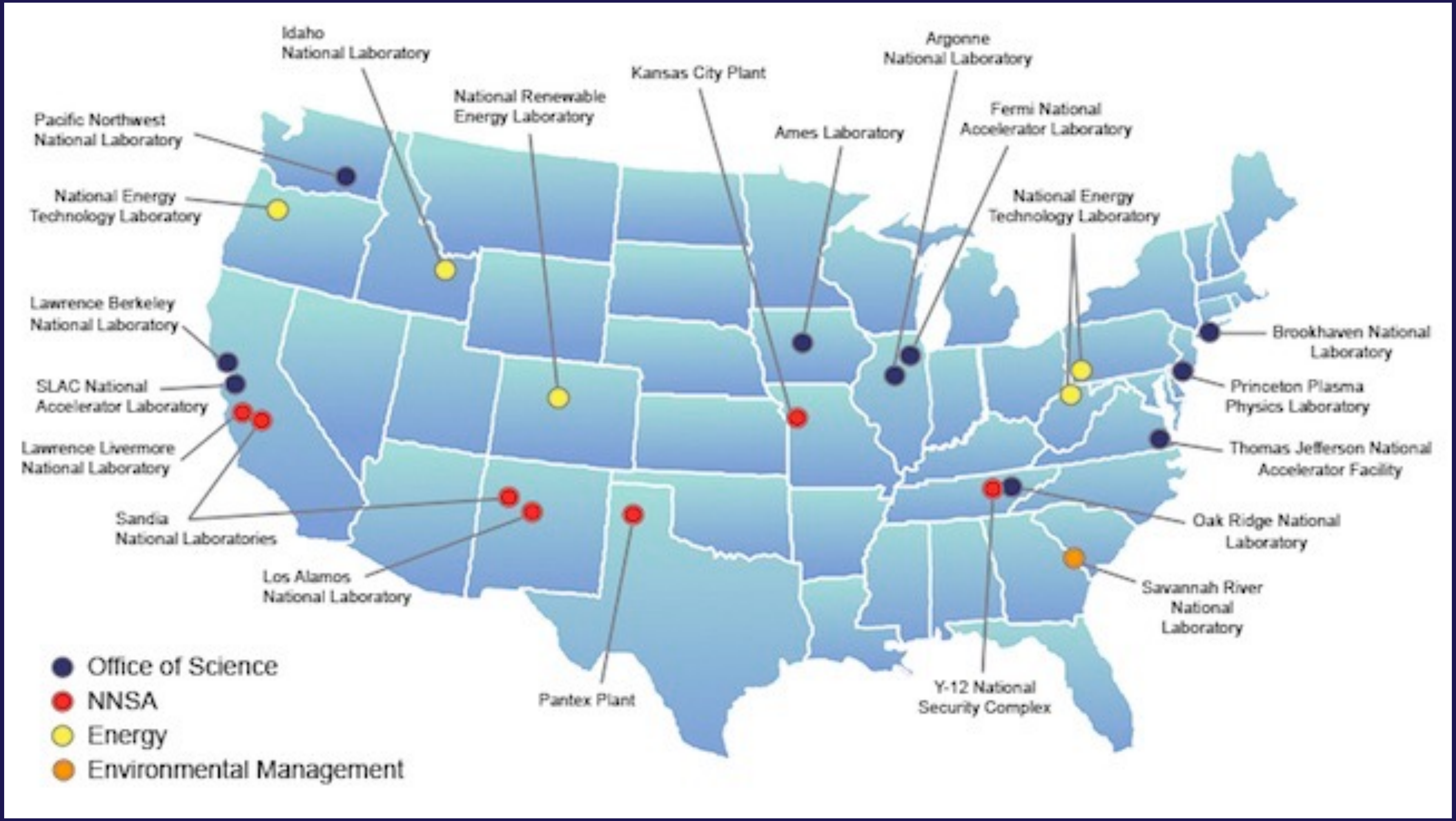
High Performance Computing – DOE Facilities, Direction and Applications

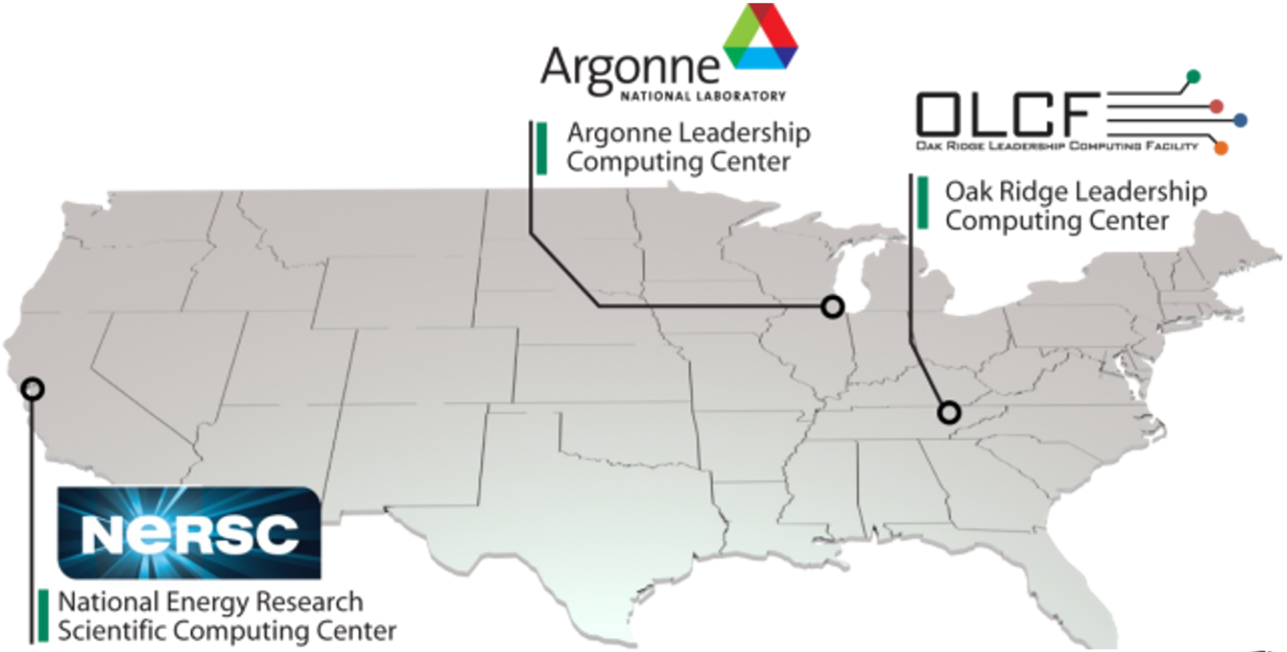
David E. Martin
Argonne Leadership Computing Facility
dem@alcf.anl.gov

August 9, 2022

DOE National Lab Complex



DOE’s Office of Science, Advanced Scientific Computing Research (ASCR) has 3 *Computation User Facilities* with some of the most powerful supercomputers in the world



NERSC
Perlmutter is 70 PF

ALCF
Polaris is 44 PF

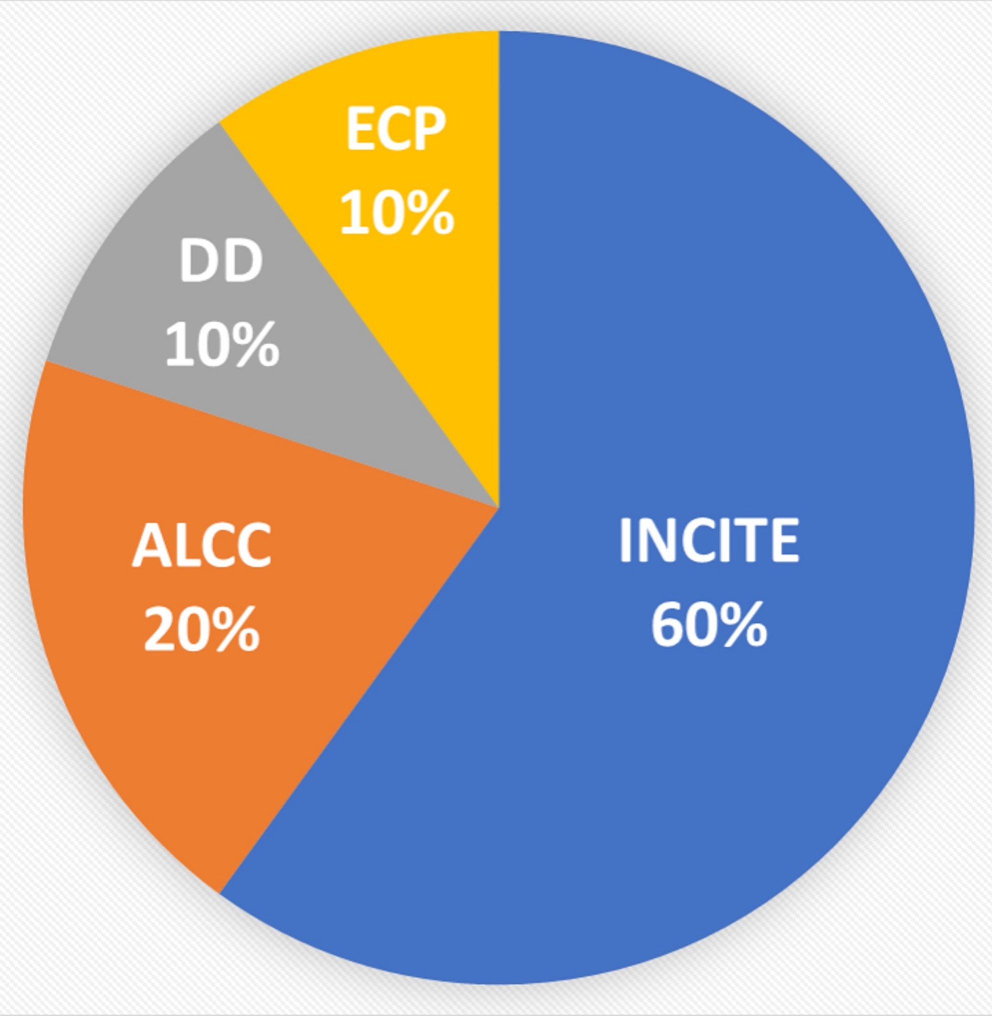
OLCF
Summit is 200 PF

- **Oak Ridge Leadership Computing Facility** (OLCF): *DOE Leadership Computing Facility*
- **Argonne Leadership Computing Facility** (ALCF): *DOE Leadership Computing Facility*
- **National Energy Research Scientific Computing Center (NERSC):** *A scalable parallel computing facility for Office of Science research needs*

COMMON RULES FOR DOE NATIONAL SCIENTIFIC USER FACILITIES

- Open to all
  - No restriction on organization, funding source, nationality, or research area
- Access through peer-reviewed proposal process
  - Project must enable breakthrough science
  - Rapid access for scale-up and smaller projects available
- Two ways to “pay”
  - Publish significant scientific results
  - Pay cost recovery to keep *everything* proprietary
- Expert support
  - Dedicated staff help to users utilize unique resources
  - Collaborative work with domain experts

ALCF and OLCF Allocation Programs



INCITE – INNOVATIVE AND NOVEL COMPUTATIONAL IMPACT ON THEORY AND EXPERIMENT

- Yearly call with computational readiness and peer reviews
- Open to all domains and user communities

ALCC – ASCR LEADERSHIP COMPUTING CHALLENGE

- Yearly call with peer reviews
- Focused on DOE priority

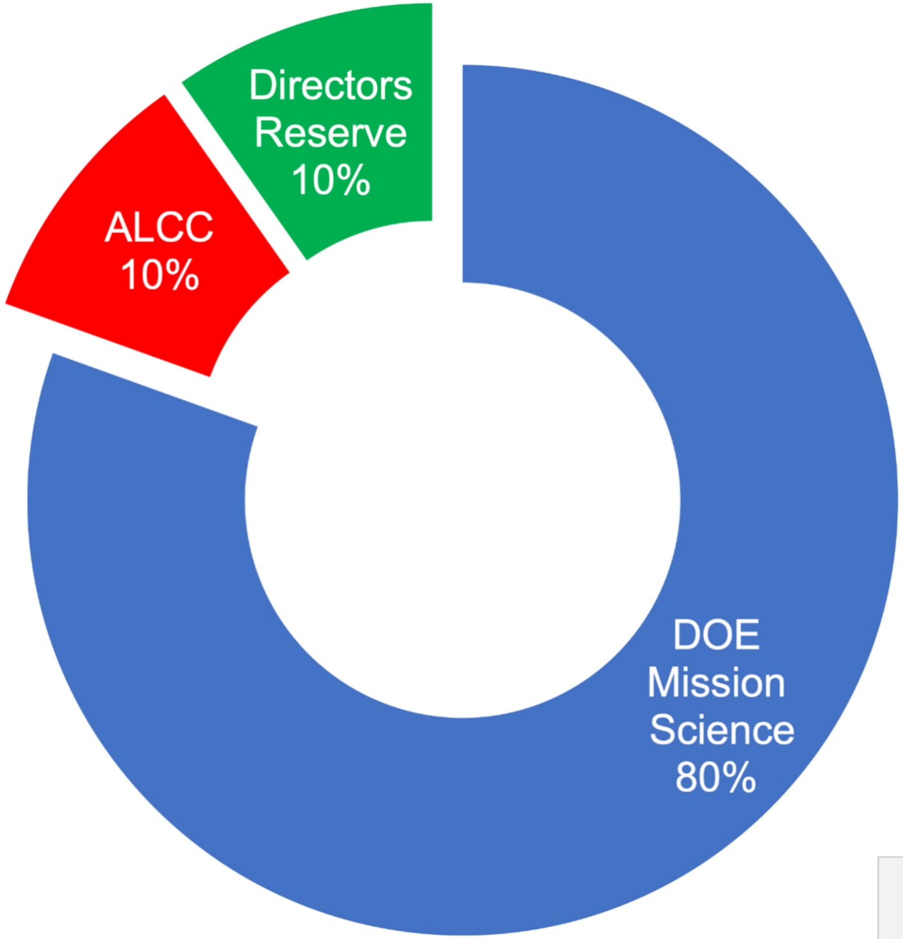
ECP - EXASCALE COMPUTING PROJECT

- Available for ECP efforts to achieve key performance parameters

DIRECTOR'S DISCRETIONARY PROGRAM

- Rapid allocations for project prep and immediate needs, apply any time
- Early science programs

NERSC Allocation Programs



Awarded by DOE Office of Science program managers to support science of interest to their program. Based on proposals submitted by researchers.

Competitive process run by DOE Advanced Scientific Computing Research Office

Strategic awards from NERSC

e.g., COVID research, ECP, QIS@Perlmutter program, Industrial Partners, Education, CSGF fellows, LDRD innovative research

Anyone conducting Office of Science relevant research can apply to use NERSC, but you must demonstrate a need to use the unique capabilities offered by the center.

Argonne Leadership Computing Facility

A DOE Office of Science National User Facility

Supercomputing resources that enable breakthroughs in science and engineering

Argonne NATIONAL LABORATORY

ALCF FOCI

LEADERSHIP COMPUTING

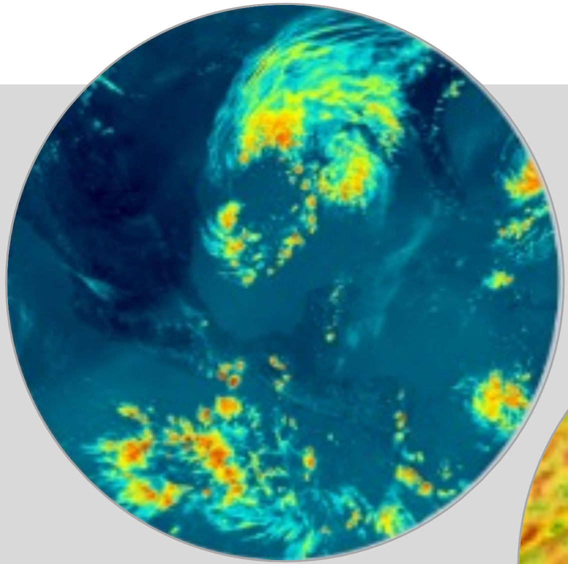
- Maintain a High-Performance Computing system for the largest, most complex modeling and simulations
- Maintain storage, networking and software

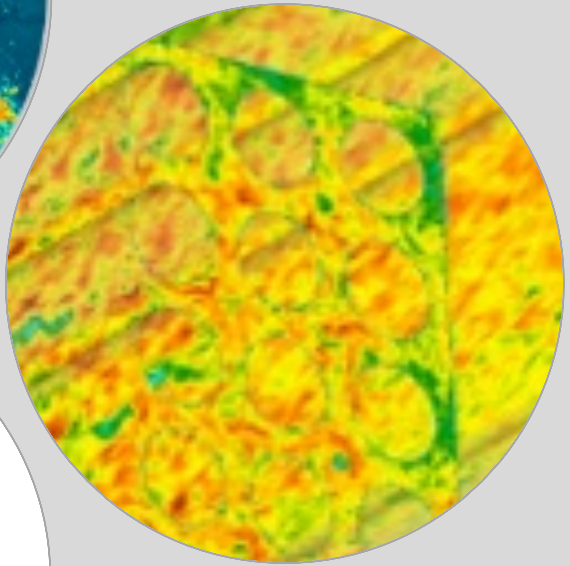
EXPERT SUPPORT

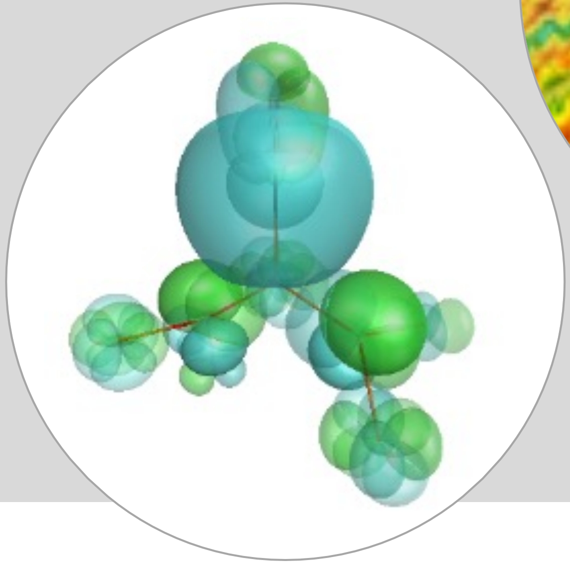
- Computational scientists that are domain scientists who translate problems to computational representations
- Performance engineers port and optimize code for massively parallel machines

NEXT GENERATION MACHINES AND SOFTWARE

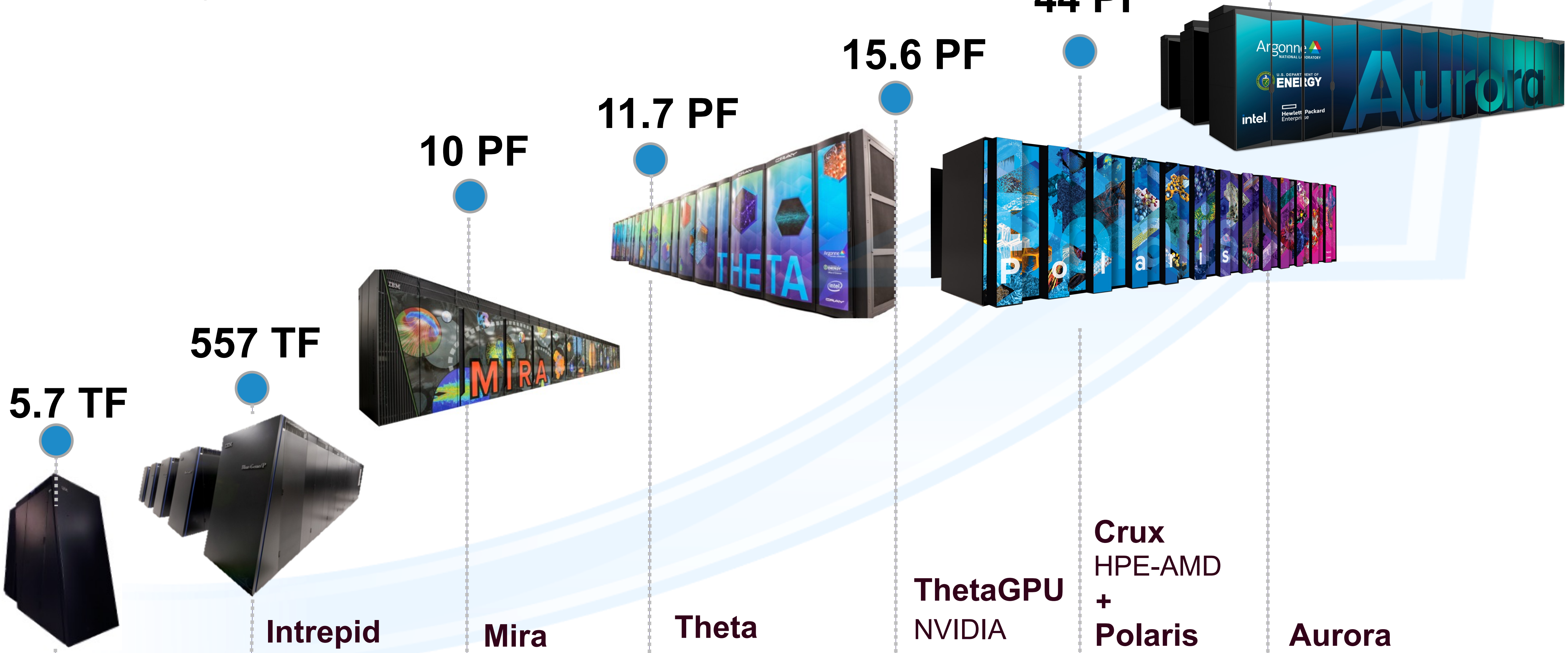
- Work with vendors to develop the next generation of HPC
- Design, procure and install cutting edge computing







ALCF Systems Evolution



5.7 TF — IBM BG/L, 2004

557 TF — **Intrepid**, IBM BG/P, 2007

10 PF — **Mira**, IBM BG/Q, 2012

11.7 PF — **Theta**, Intel-Cray XC40, 2017

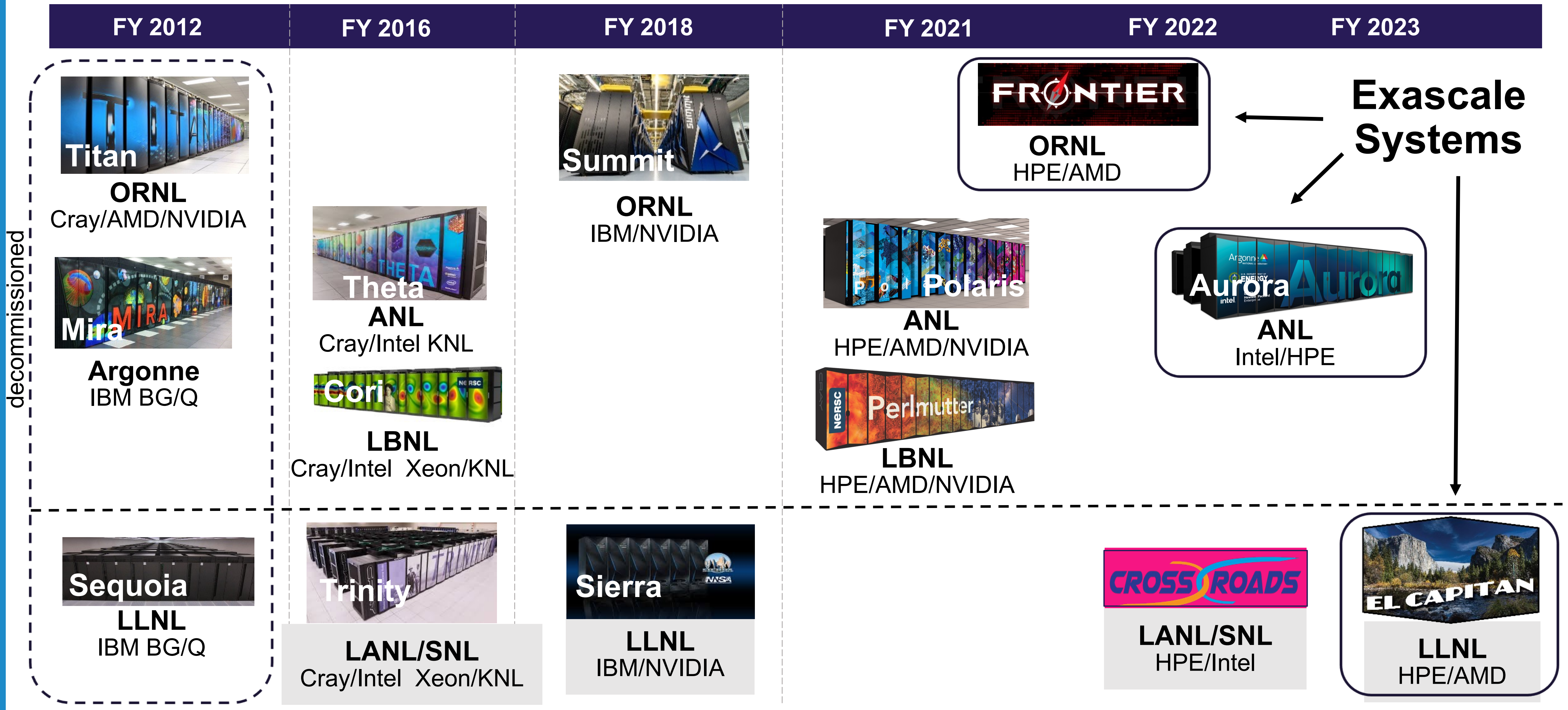
15.6 PF — **ThetaGPU**, NVIDIA DGX A100, 2020

44 PF — **Crux** HPE-AMD + **Polaris** HPE, 2021

> 2 EF — **Aurora**, Intel-HPE, 2022

JLSE (2013) **AI Testbed** (2020) **Edge Testbed** (2021)

DOE HPC Road to Exascale Systems



Aurora Cabinets Installed at Argonne

Aurora

Leadership Computing Facility
Exascale Supercomputer

PEAK PERFORMANCE

≧ 2 Exaflops DP

Intel GPU

Ponte Vecchio

Intel Xeon PROCESSOR

Sapphire Rapids with HBM

PLATFORM

HPE Cray-Ex

Compute Node
2 SPR+HBM processor;
6 PVC; Unified Memory Architecture; 8 fabric endpoints;

GPU Architecture
Xe arch-based “Ponte Vecchio” GPU
Tile-based chiplets
HBM stack
Foveros 3D integration

System Interconnect
HPE Slingshot 11; Dragonfly topology with adaptive routing

Network Switch
25.6 Tb/s per switch, from 64–200 Gb/s ports (25 GB/s per direction)

Node Performance
>130 TF

System Size
>9,000 nodes

Aggregate System Memory
>10 PB aggregate System Memory

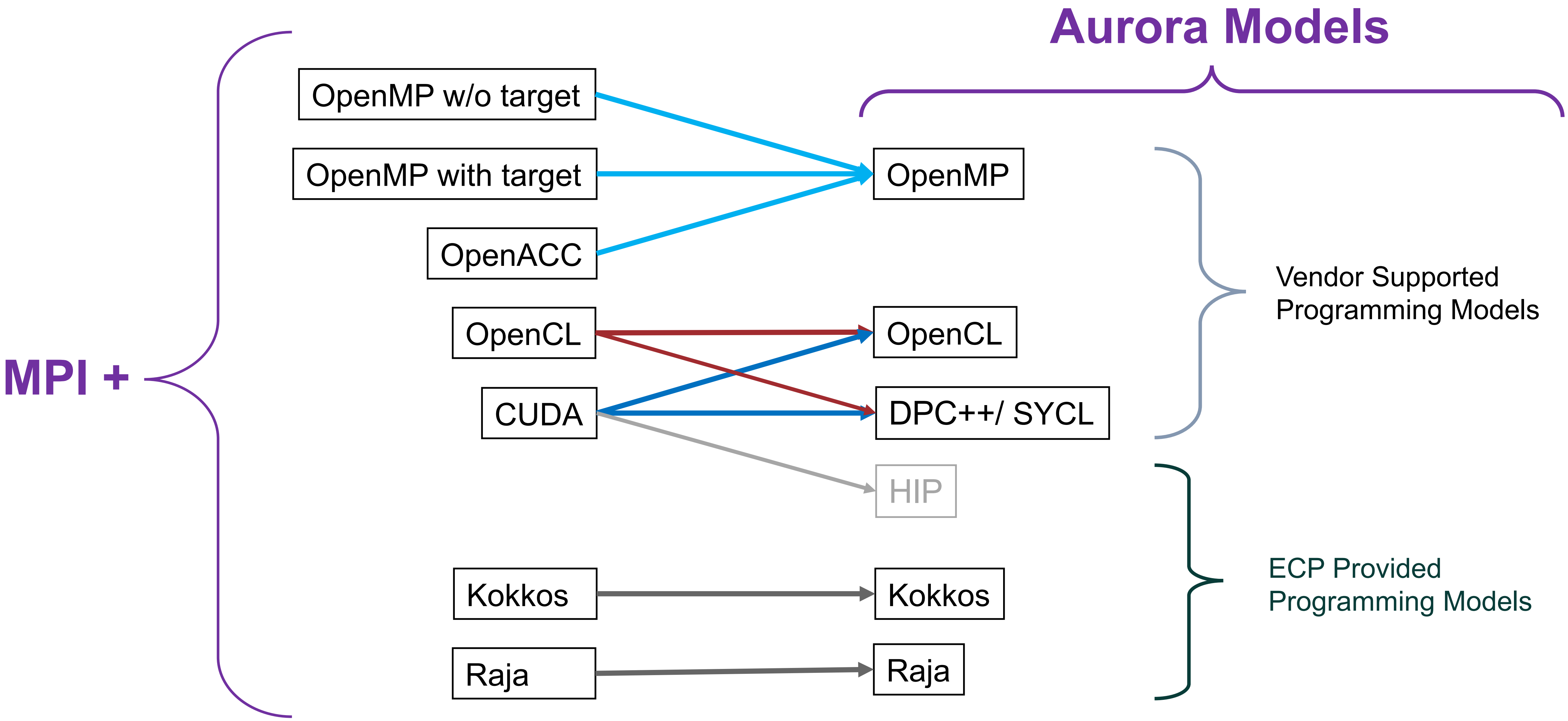
High-Performance Storage
220 PB @ EC16+2**, ≧25 TB/s DAOS**

Programming Models
oneAPI, MPI, OpenMP, C/C++, Fortran, SYCL/DPC++
Python-based environments
Machine learning and Deep learning frameworks

Expert Support

- computational scientists
- computer scientists
- system administrators
- performance engineers
- visualization experts

Mapping of Existing Programming Models to Aurora



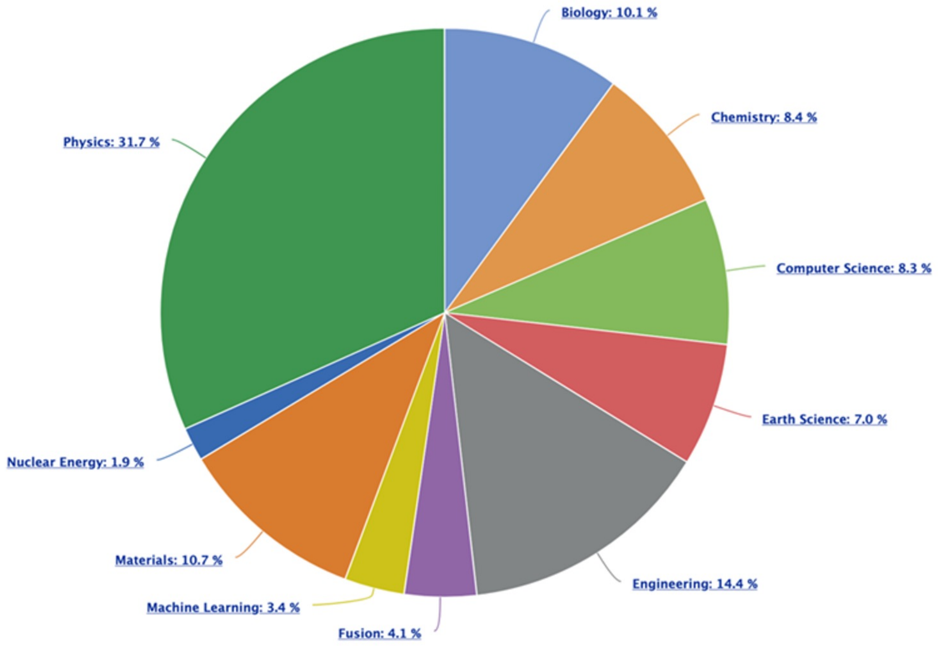
Best Practices

- Consider GPU acceleration generically
  - —May need code refactoring, different algorithms
  - —Do this work on present-day GPU systems
- Consider high-level portability layers
  - —Canned (Kokkos, e.g.)
  - —Roll your own (DSL, e.g.)
- For future systems such as Aurora
  - —Work with closest previous-generation GPU accelerators
  - —Work with beta developer software
- Work with the portability-layer developers
- Do not optimize prematurely
  - —Beta software and previous-generation hardware may optimize differently than final targets
  - —When you do start profiling, use a physically realistic problem

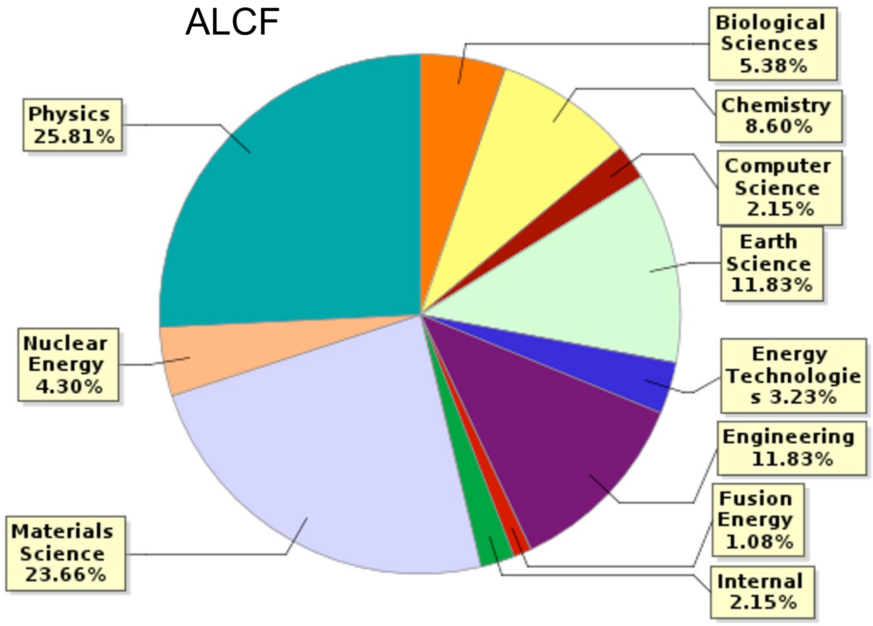
> It is easy to add platform-specific code, such as memory allocators or warp sizes.
> Build and test on other platforms to identify these sooner rather than later.

Projects by Science Domain (2021)

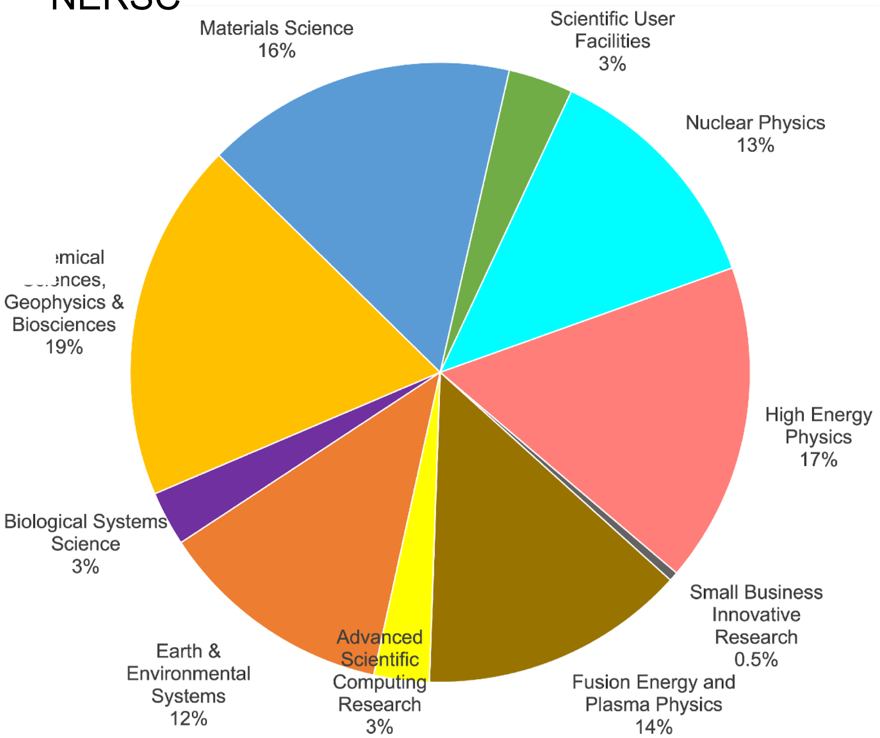
OLCF



ALCF



NERSC



NEW ALCF INITIATIVES

Simulation is ALCF's bread and butter, but the needs of users are expanding. ALCF must

- Provide data storage and management
  - Long term storage and sharing of experimental and computational data
- Enable artificial intelligence and data analytics
  - Machine learning to accelerate simulations and data analytics to understand experimental and sensor data
  - AI Testbed to provide dedicated hardware and software for data analytics and machine learning
- Connect to experimental and observational facilities (at Argonne and beyond)
  - Simulations and data analysis to guide running experiment
  - Digital twin for predictive maintenance and operations

Integrating Different Types of HPC

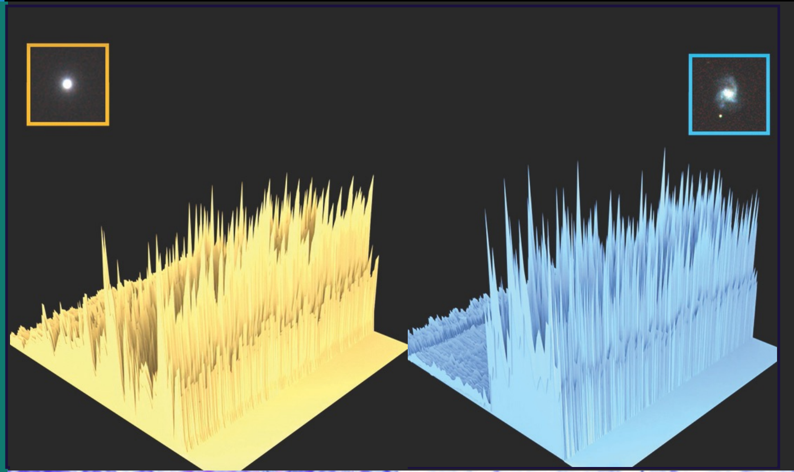
ALCF is driving a new paradigm in high performance computing for science and engineering

Modeling & Simulation

Used to study things that are too big, too small, or too dangerous to study in a laboratory setting.

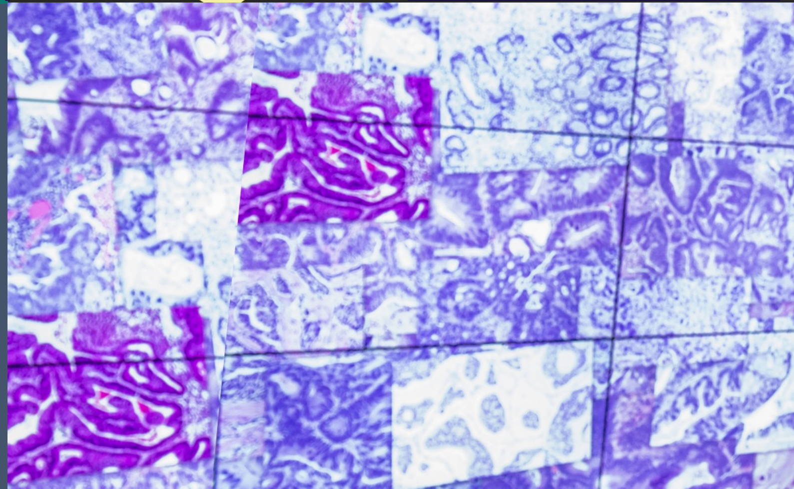
Data Science

Researchers can glean insights from very large datasets produced by experimental, simulation, or observational methods.



Machine Learning

A type of artificial intelligence that trains computers to discover hidden patterns in data to make novel predictions without being explicitly programmed.



COMMUNITY DATA SHARING WITH EAGLE

- **A global filesystem deployed to bring larger and more capable production-level file sharing to facility users**
- **A space for broader distribution of reassembled data acquired from various experiments**
  - Data originating at the ALCF
  - Greater scientific community
- **Science community can access uploaded data, and ALCF users are able to directly access the data for analysis**
- **Designed to foster experimentation**
  - Analysts are able to write new algorithms to attempt analyses that have never been performed

HPE ClusterStor E1000

100 petabytes of usable capacity

8,480 disk drives

Lustre filesystem

160 Object Storage Targets

40 Metadata Targets

HDR Infiniband network

650 GB/s rate on data transfers

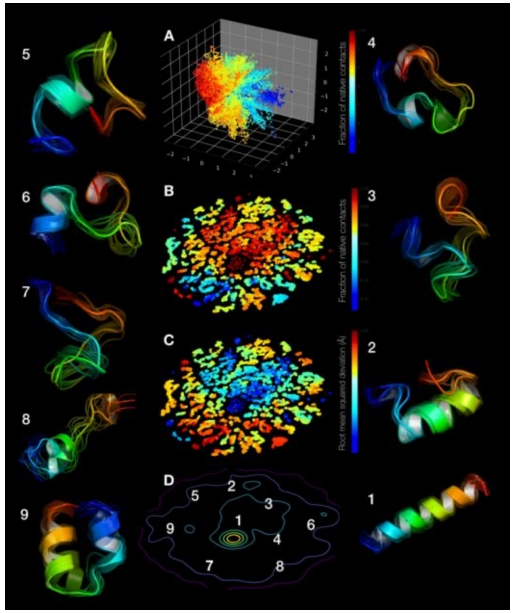
ALCF COMMUNITY DATA CO-OP

- Production system for available the the ALCF user community
- Data portals for discovery and access
  - Researches can post their data for others to browse and analyze
- Based on Django Globus Portal Framework
  - Allows customized portals for each data collections
- Some current projects
  - APS High-Energy Diffraction Microscopy
  - APS X-Ray Photon Correlations Spectroscopy
  - DLHub – a multi-tenant repository of data science models
  - Materials Data Facility

SURGE OF SCIENTIFIC MACHINE LEARNING

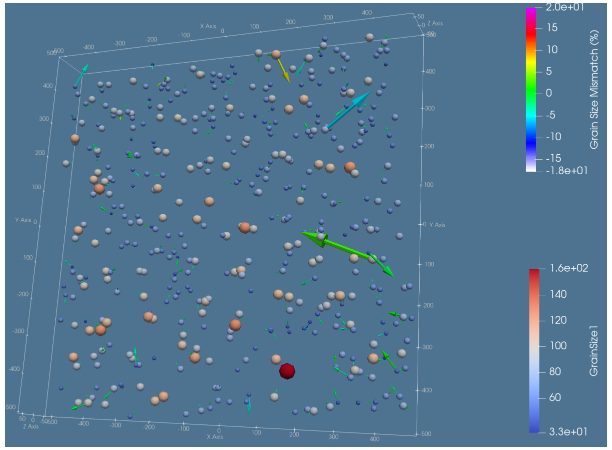
- Simulations/surrogate models
  - Replace, in part, or guide simulations with AI-driven surrogate models
- Data-driven models
  - Use data to build models without simulations
- Co-design of experiments
  - AI-driven experiments

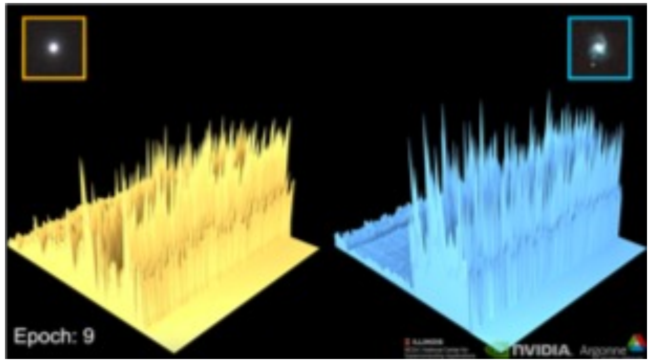
Design infrastructure to facilitate and accelerate AI for Science applications



Protein-folding

Braggs Peak





Galaxy Classification

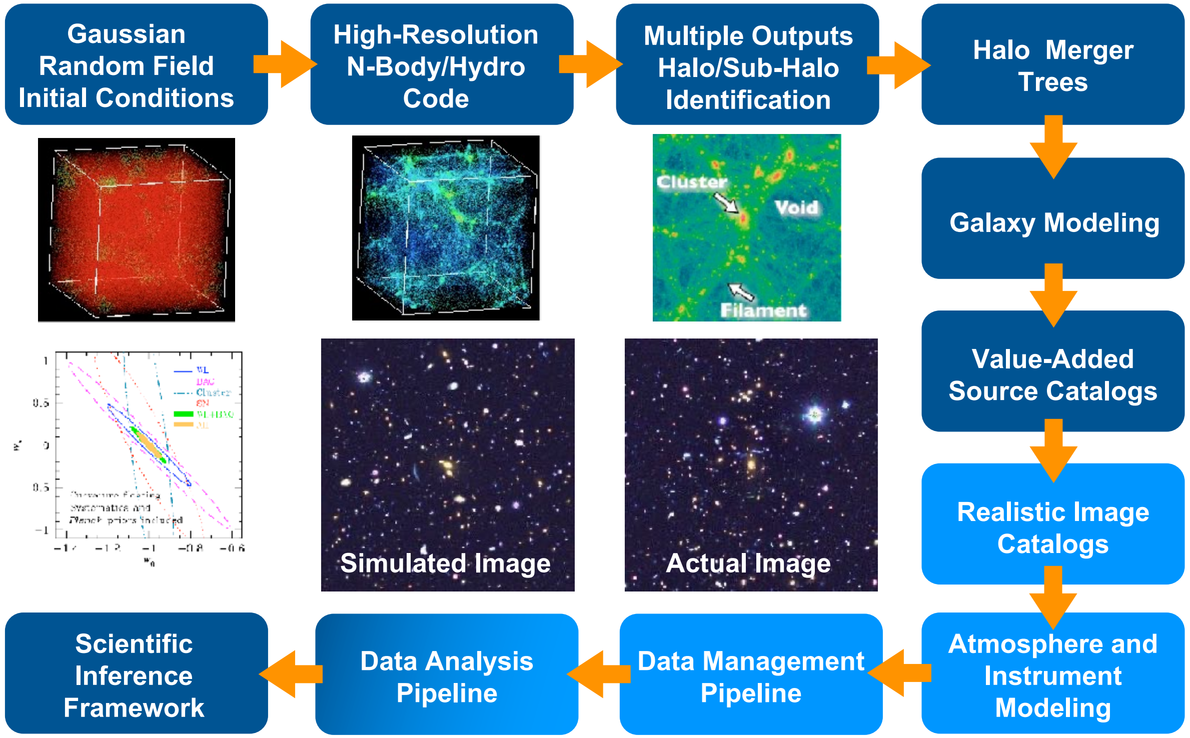
High Energy Physics

Energy/Intensity Frontier:

- Search for Beyond the Standard Model (BSM) physics through AI-driven anomaly detection
- AI-reduced uncertainties to enable precision electroweak measurements for BSM clues
- Generative Adversarial Networks (GANs) for large-scale Large Hadron Collider detector simulation

Cosmic Frontier – AI in end-to-end application:

- Precision Cosmic Microwave Background emulation – AI simulation speed-up of a factor of 1000
- Search for strong lensing of galactic sources for precision cosmology measurements using AI classification, regression, and GANs for image generations
- AI-based Photometric Redshift Estimation
- Combination of AI methods to enable searches for hidden space variables



AI applications in an "end-to-end" Cosmic Frontier application: 1) GANs for image emulation, 2) GP and DL-based emulators for summary statistics, 3) CNN-based image classification, 4) AI-based photometric reshift estimation, 5) Likelihood-free methods for inference *[Work performed under the Argonne-led SciDAC-4 project: "Inference and Machine Learning at Extreme Scales"]*

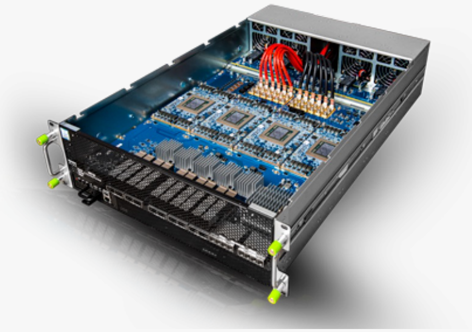
ALCF AI Testbeds

https://www.alcf.anl.gov/alcf-ai-testbed

Cerebras (CS-2)

SambaNova

Graphcore



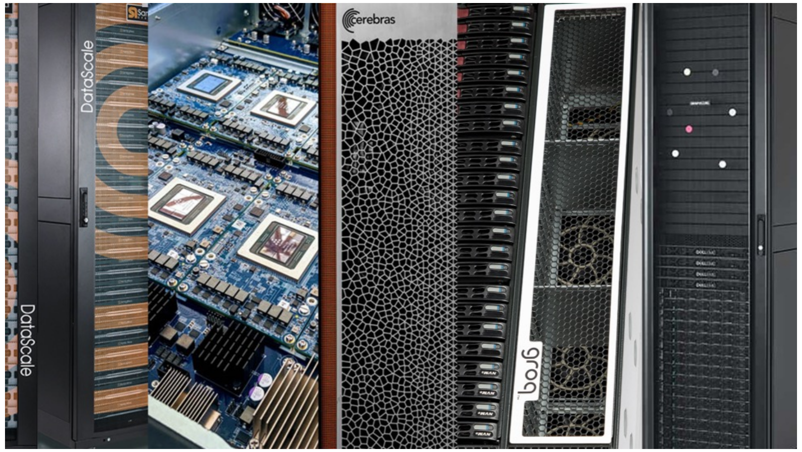
Habana

Groq

- AI Accelerators have unique architectures focused on a particular AI problem space
- ALCF works with companies, often under NDA, to provide testing and feedback
- ALCF has a stable of experimental machines for internal and external use.

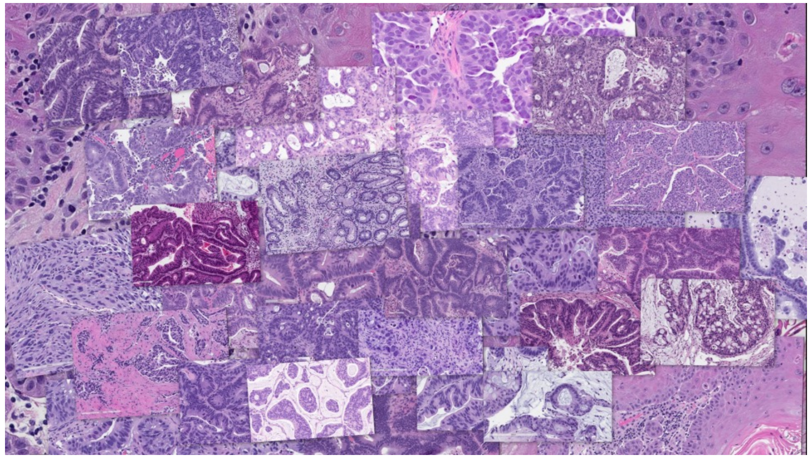
	Cerebras CS2	SambaNova Cardinal SN10	Groq GroqCard	GraphCore GC200 IPU	Habana Gaudi1	NVIDIA A100
Compute Units	850,000 Cores	640 PCUs	5120 vector ALUs	1472 IPUs	8 TPC + GEMM engine	6912 Cuda Cores
On-Chip Memory	40 GB	>300MB	230MB	900MB	24 MB	192KB L1 40MB L2
Process	7nm	7nm	14nm	7nm	7nm	7nm
System Size	2 Nodes	2 nodes (8 cards per node)	4 nodes (8 cards per node)	1 node (8 cards per node)	2 nodes (8 cards per node)	Several systems
Estimated Performance of a card (TFlops)	>5780 (FP16)	>300 (BF16)	>188 (FP16)	>250 (FP16)	>150 (FP16)	312 (FP16), 156 (FP32)
Software Stack Support	Tensorflow, Pytorch	SambaFlow, Pytorch	GroqAPI, ONNX	Tensorflow, Pytorch, PopArt	Synapse AI, TensorFlow and PyTorch	Tensorflow, Pytorch, etc
Interconnect	Ethernet-based	Infiniband	RealScale TM	IPU Link	Ethernet-based	NVLink

AI TESTBED GOALS

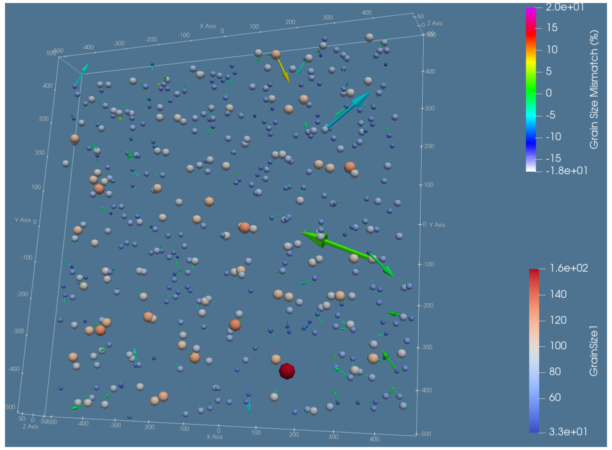


- Learn strengths and weaknesses of different approaches to AI acceleration
- Provide a testbed for labs, academia and industry to try out new AI algorithms on cutting edge hardware
- Work with vendors to help them understand and meet the needs of science and engineering
- Integrate AI accelerators into the ALCF ecosystem of computing, storage and networking
- Make AI accelerators available to the ALCF user community

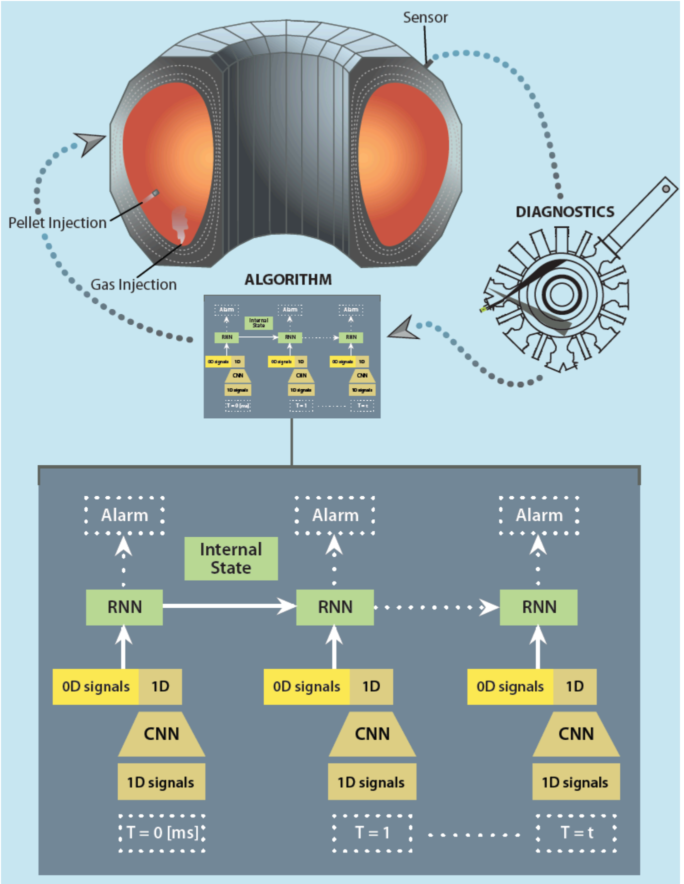
AI FOR SCIENCE APPLICATIONS ON AI TESTBED



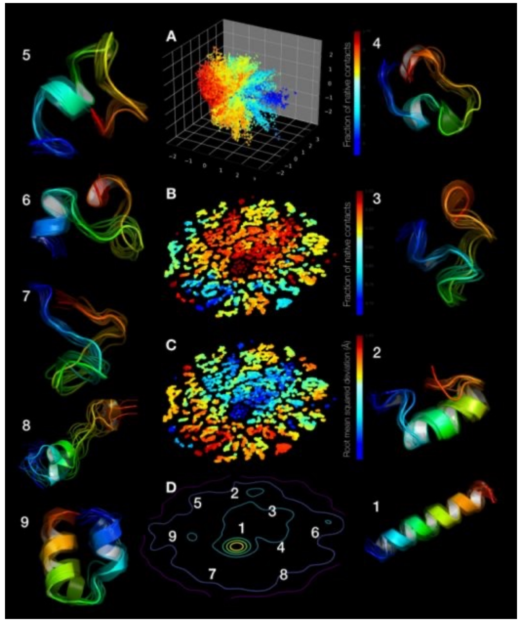
Cancer drug response prediction



Imaging Sciences-Braggs Peak



Tokomak Fusion Reactor operations



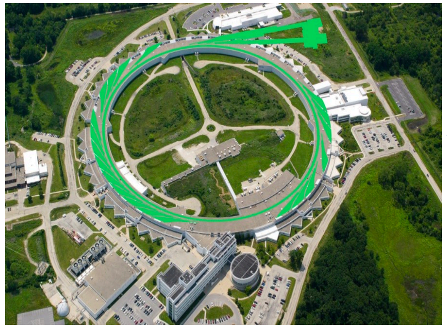
Protein-folding(Image: NCI)

and more..

Argonne NATIONAL LABORATORY

PUTTING IT ALL TOGETHER: INTEGRATION ACROSS FACILITIES

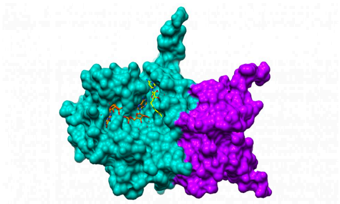
Experimental Facility

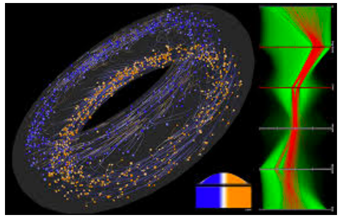


AI-Edge accelerator

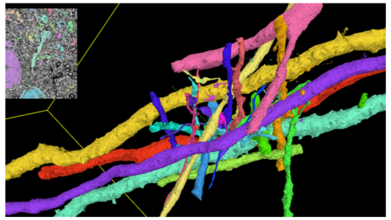
Computing Facility

AI-Accelerators





Simulations



Data-driven Models


Argonne
NATIONAL LABORATORY