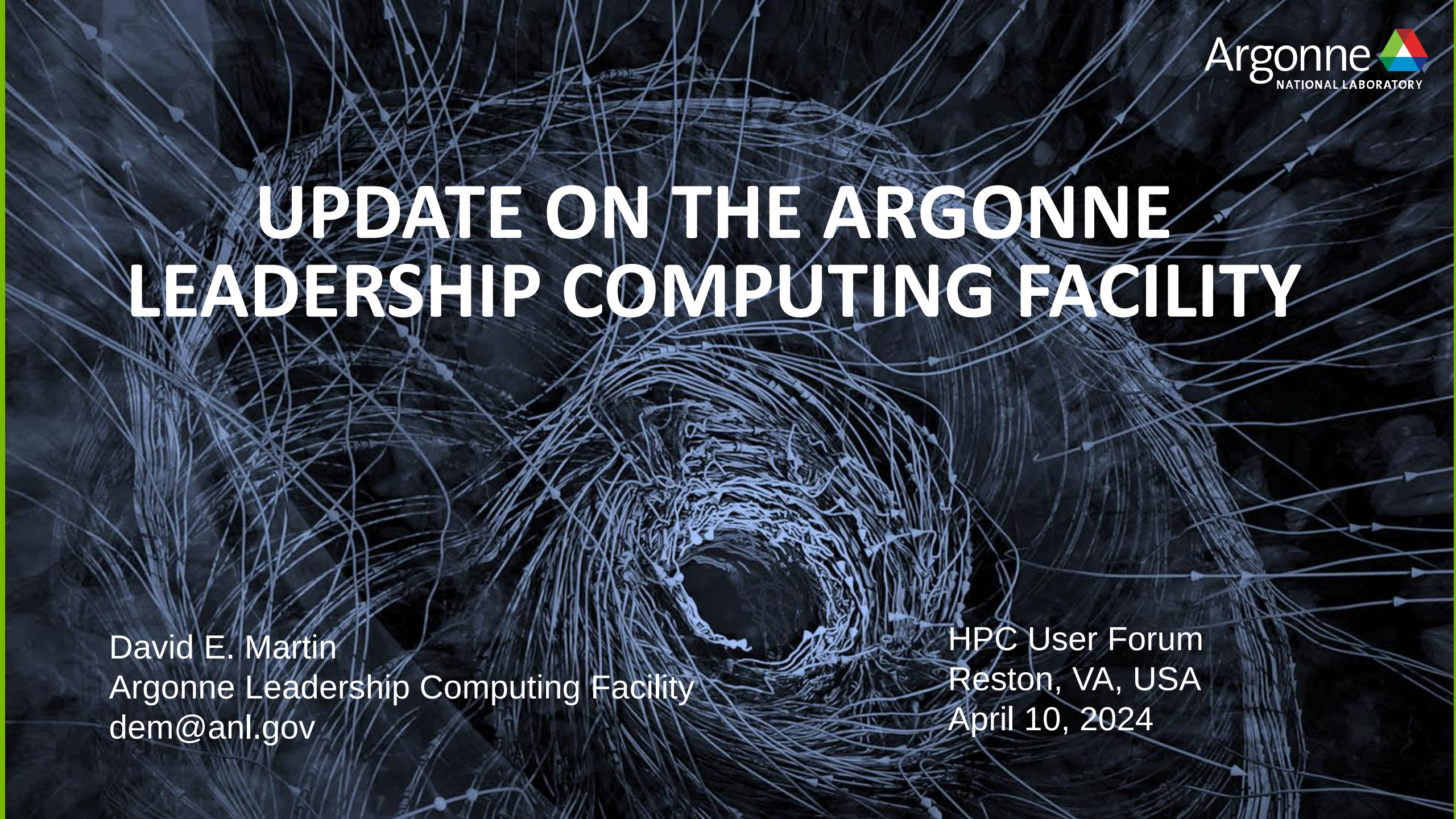


UPDATE ON THE ARGONNE LEADERSHIP COMPUTING FACILITY

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

HPC User Forum
Reston, VA, USA
April 10, 2024

DOE Leadership Computing Facility

- Established in 2004 as a collaborative, multi-lab initiative funded by DOE's ***Advanced Scientific Computing Research*** program
- Operates as **one facility** with two centers, at Argonne and at Oak Ridge National Laboratory
- Deploys and operates at least two advanced architectures that are **10-100 times more powerful** than systems typically available for open scientific research
- **Fully dedicated** to open science to address the ever-growing needs of the scientific community

COMMON RULES FOR 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 discretionary access 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 FOCI

LEADERSHIP COMPUTING

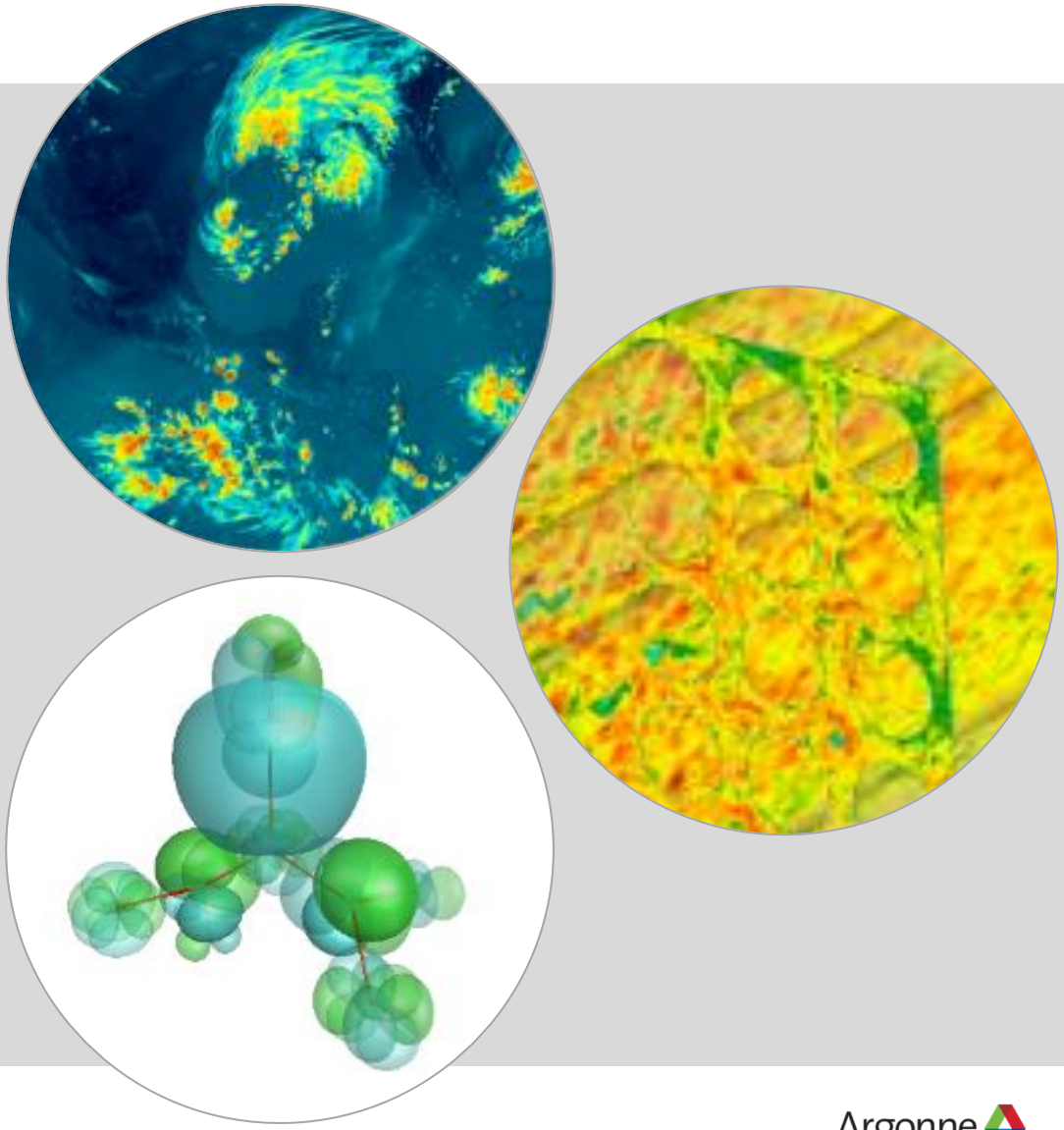
- Maintain a High-Performance Computing system for the largest, most complex modeling and simulations
- Maintain storage, networking and software

EXPERT SUPPORT

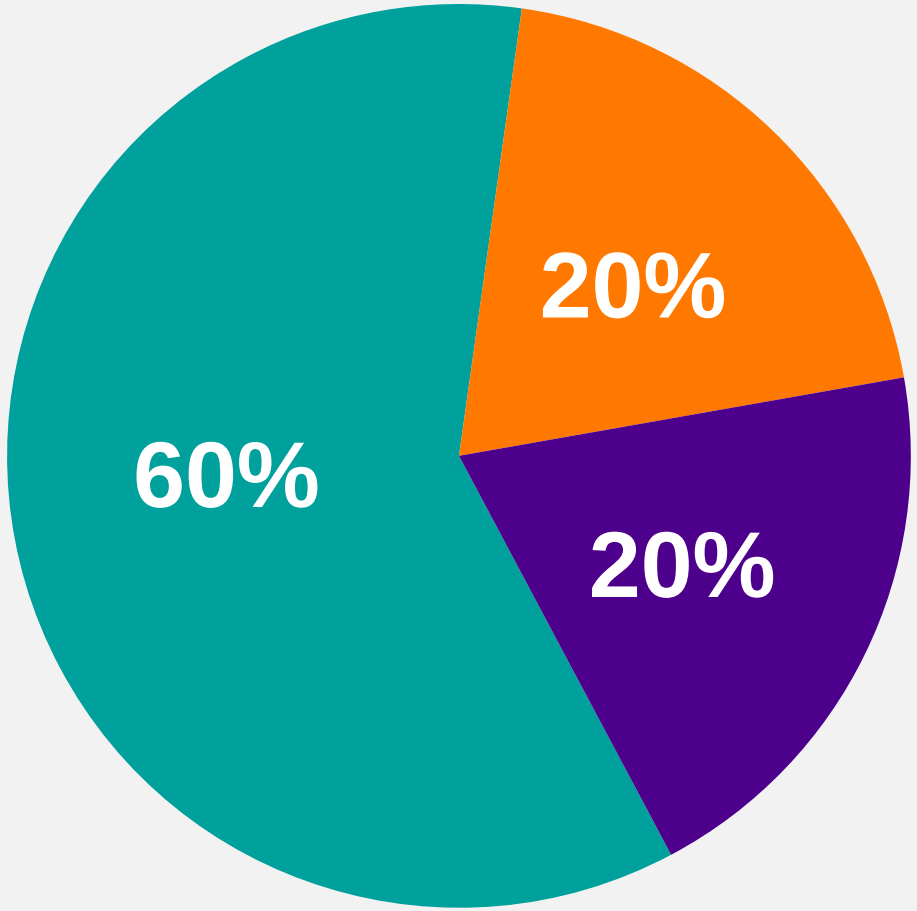
- Computational scientists that are domain scientists who translate problems to computational representations
- Performance engineers port and optimize code for massively parallel machines

NEXT GENERATION MACHINES AND SOFTWARE

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



ALCF 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

DIRECTOR'S DISCRETIONARY PROGRAM

- Rapid allocations for project prep and immediate needs

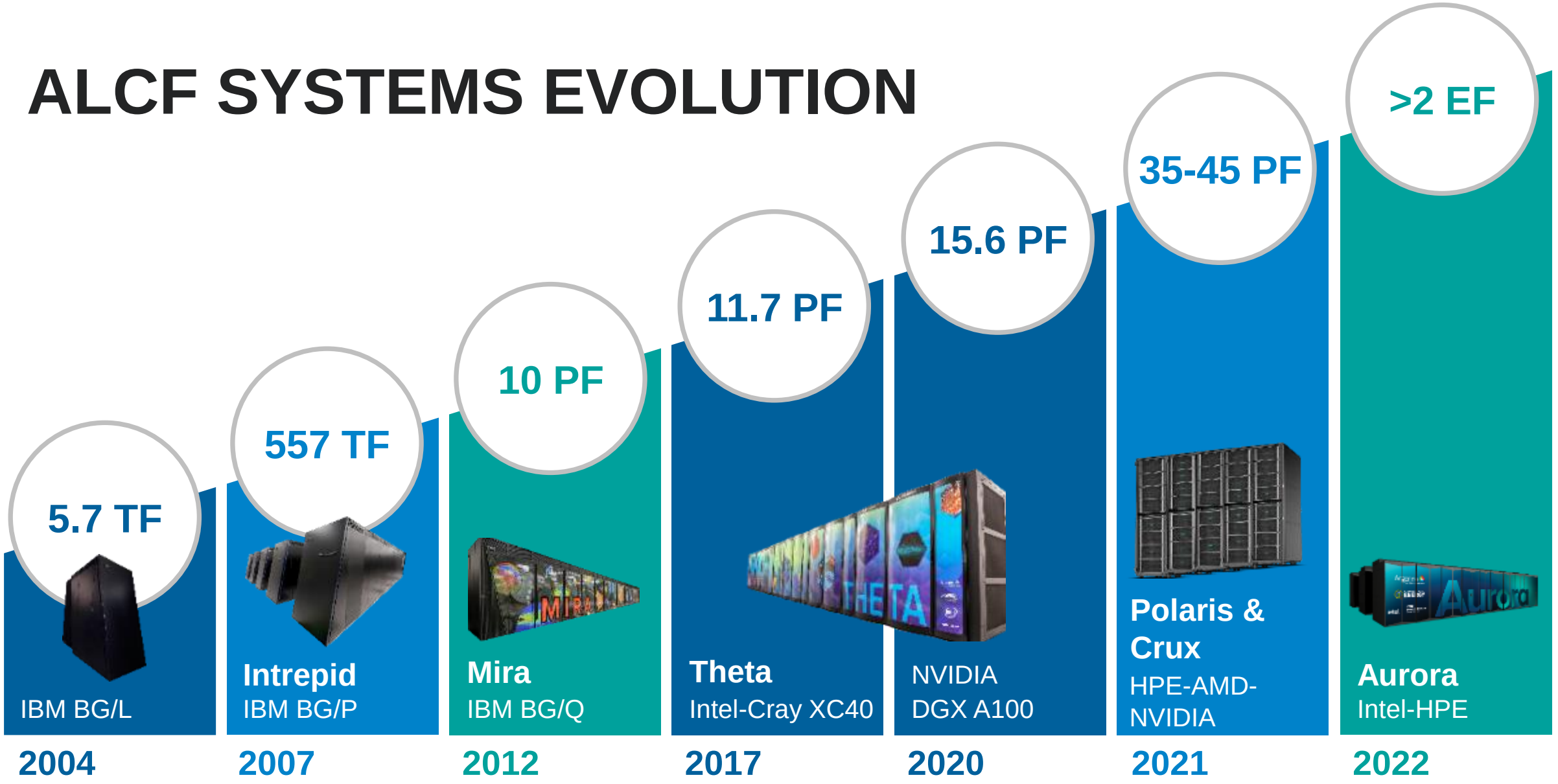
EARLY SCIENCE PROGRAM

EXASCALE COMPUTING PROJECT

ARGONNE DATA SCIENCE PROGRAM

PROPRIETARY PROJECTS

ALCF SYSTEMS EVOLUTION



Polaris

Argonne Leadership Computing Facility
Pre-Exascale Supercomputer

PEAK PERFORMANCE

44 Petaflop DP

NVIDIA GPU

A100

AMD EPYC PROCESSOR

Milan 7543

PLATFORM

HPE Apollo Gen10+

Compute Node
1 AMD Milan processor;
4 NVIDIA A100 GPUs; Unified Memory Architecture; 2 fabric endpoints; 2 NVMe SSDs

GPU Architecture
NVIDIA A100 GPU; HBM stack

CPU-GPU Interconnect
CPU-GPU: PCIe
GPU-GPU: NVLink

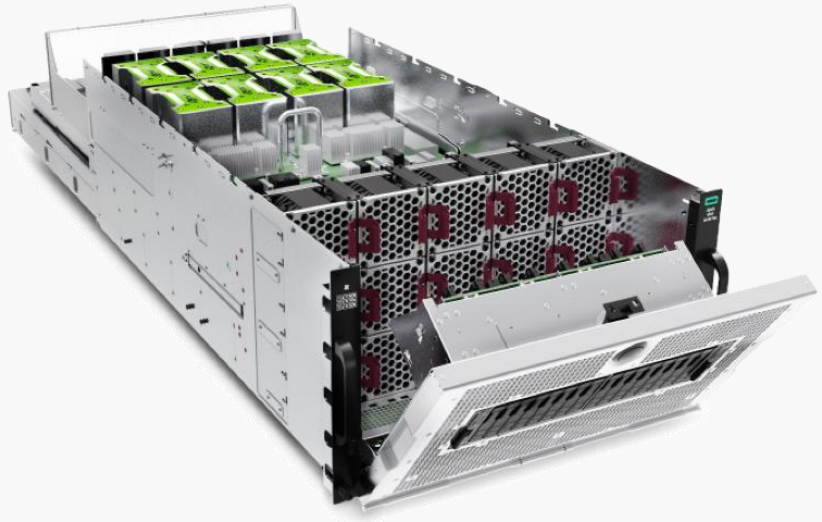
System Interconnect
HPE Slingshot 11; Dragonfly topology with adaptive routing

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

Programming Models
CUDA, MPI, OpenMP, C/C++, Fortran, DPC++

Node Performance
78 TF

System Size
560 nodes



Aurora

Argonne Leadership Computing Facility
Exascale Supercomputer

Peak Performance
≧ 2 Exaflops DP

Intel GPU
Ponte Vecchio (PVC)

Intel Xeon Processor
Sapphire Rapids with High Bandwidth Memory

Platform
HPE Cray-Ex

Compute Node
2 Xeon SPR+HBM processors
6 Ponte Vecchio GPUs
Node Unified Memory Architecture
8 fabric endpoints

GPU Architecture
Intel XeHPC architecture
High Bandwidth Memory Stacks

Node Performance
>130 TF

System Size
>9,000 nodes

Aggregate System Memory
>10 PB aggregate System Memory

System Interconnect
HPE Slingshot 11
Dragonfly topology with adaptive routing

Network Switch
25.6 Tb/s per switch (64 200 Gb/s ports)
Links with 25 GB/s per direction

High-Performance Storage
220 PB
≧25 TB/s DAOS bandwidth

Software Environment
- C/C++
- Fortran
- SYCL/DPC++
- OpenMP offload
- Kokkos
- RAJA
- Intel Performance Tools

Recent Changes in Compute Resources

- Theta decommissioned on 12/31/2023
  - The end of the multicore era at ALCF
- Theta expansion (ThetaGPU)
  - 24 DGX A100 nodes
  - Coming back as a stand-alone system later in the year
- Aurora
  - Significant time for ECP and ESP projects
  - Good performance on a range of applications
  - Full machine now dedicated to stabilization work

Emerging Initiatives at ALCF

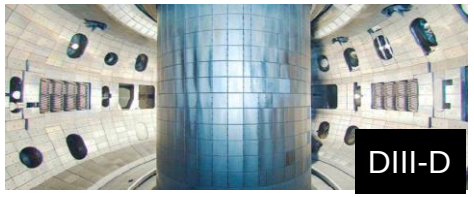
- Integrated Research Infrastructure
  - Nexus at Argonne

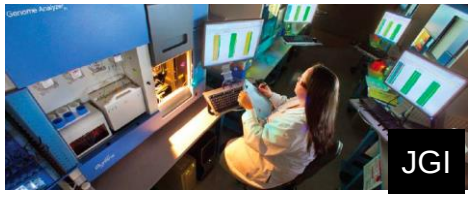
- AI
  - ALCF AI Testbed
  - Trillion Parameter Consortium
  - Aurora GPT
  - National AI Research Resource (NAIRR)

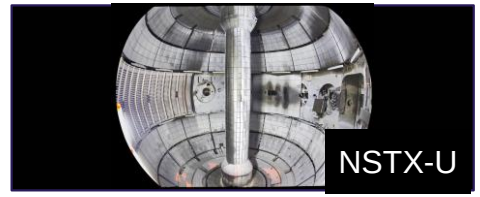
ALCF-4

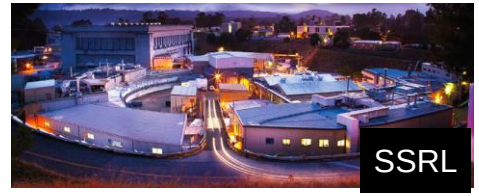
DOE EXPERIMENTAL USER FACILITIES

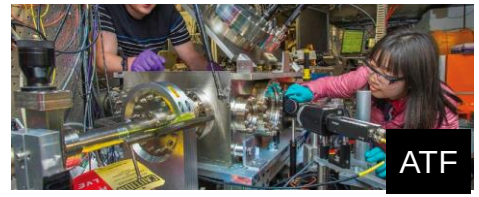
- DOE operates 24 experimental user facilities
- Similar to the computing facilities, some of them are undergoing upgrades
- Their data rates and their computing needs will increase accordingly





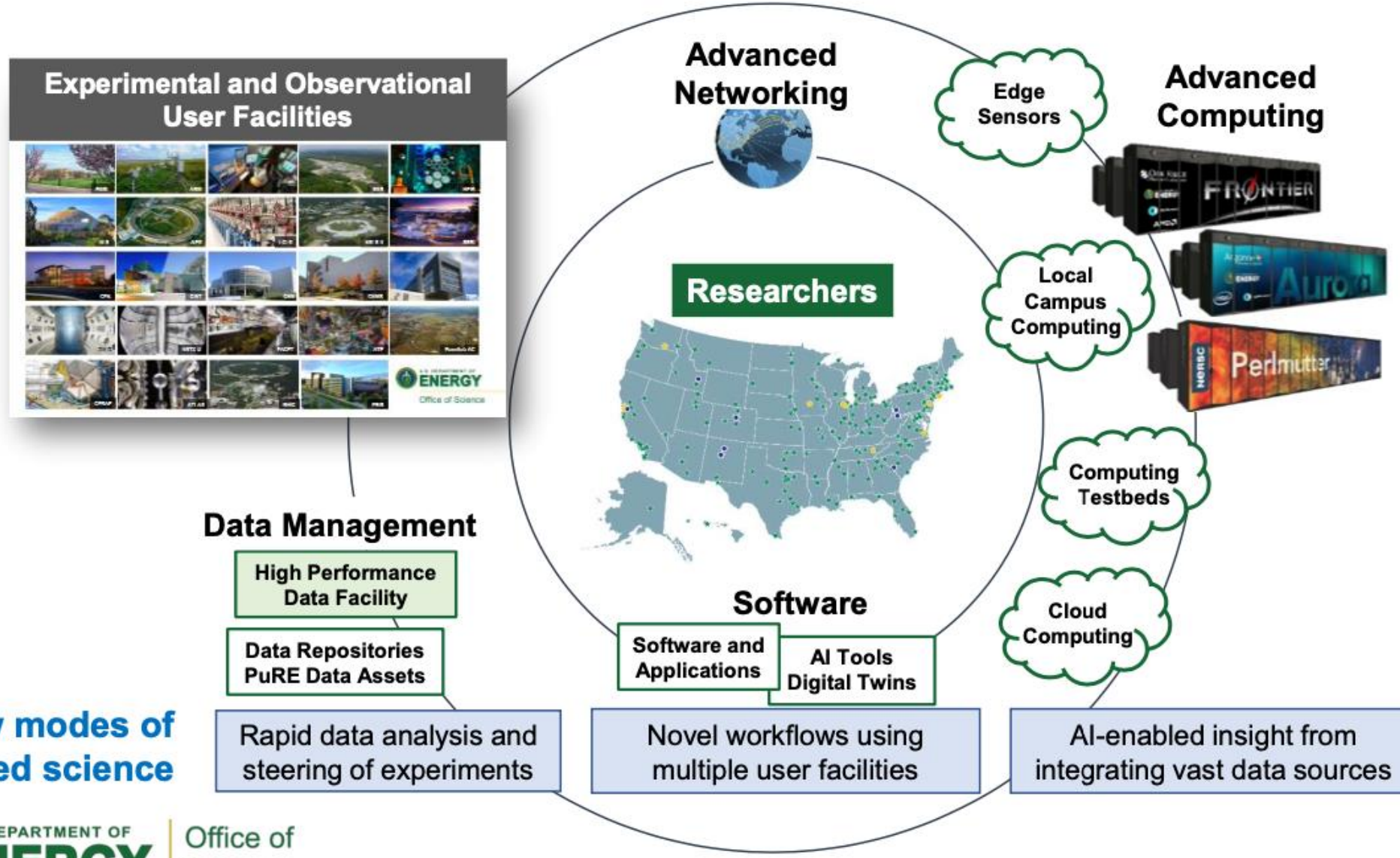






DOE's Integrated Research Infrastructure (IRI) Vision:

To empower researchers to meld DOE's world-class research tools, infrastructure, and user facilities seamlessly and securely in novel ways to radically accelerate discovery and innovation



U.S. DEPARTMENT OF ENERGY | Office of Science

Slide from Ben Brown, DOE, ECP IAC – Jan 2024

The IRI Blueprint Activity created a framework for IRI implementation

IRI Science Patterns (3)

Time-sensitive pattern has *urgency*, requiring real-time or end-to-end performance with high reliability, e.g., for timely decision-making, experiment steering, and virtual proximity.

Data integration-intensive pattern requires combining and analyzing data from multiple sources, e.g., sites, experiments, and/or computational runs.

Long-term campaign pattern requires sustained access to resources over a long period to accomplish a well-defined objective.

IRI Practice Areas (6)

User experience practice will ensure relentless attention to user perspectives and needs through requirements gathering, user-centric (co)-design, continuous feedback, and other means.

Resource co-operations practice is focused on creating new modes of cooperation, collaboration, co-scheduling, and joint planning across facilities and DOE programs.

Cybersecurity and federated access practice is focused on creating novel solutions that enable seamless scientific collaboration within a secure and trusted IRI ecosystem.

Workflows, interfaces, and automation practice is focused on creating novel solutions that facilitate the dynamic assembly of components across facilities into end-to-end IRI pipelines.

Scientific data life cycle practice is focused on ensuring that users can manage their data and metadata across facilities from inception to curation, archiving, dissemination, and publication.

Portable/scalable solutions practice is focused on ensuring that transitions can be made across heterogeneous facilities (portability) and from smaller to larger resources (scalability).

Convened over **150 DOE national laboratory experts** from **all 28 SC user facilities** across **13 national laboratories** to consider the **technological, policy, and sociological challenges** to implementing IRI.

IRI Program launch is a DOE FY24-25 Agency Priority Goal. ASCR is implementing IRI through these four major elements.

1. Invest in IRI foundational infrastructure
2. Stand up the IRI Program governance and FY24 workstreams
3. Bring IRI projects into formal coordination
4. Deploy an IRI Science Testbed across the ASCR Facilities

These are all connected.
These are each essential.

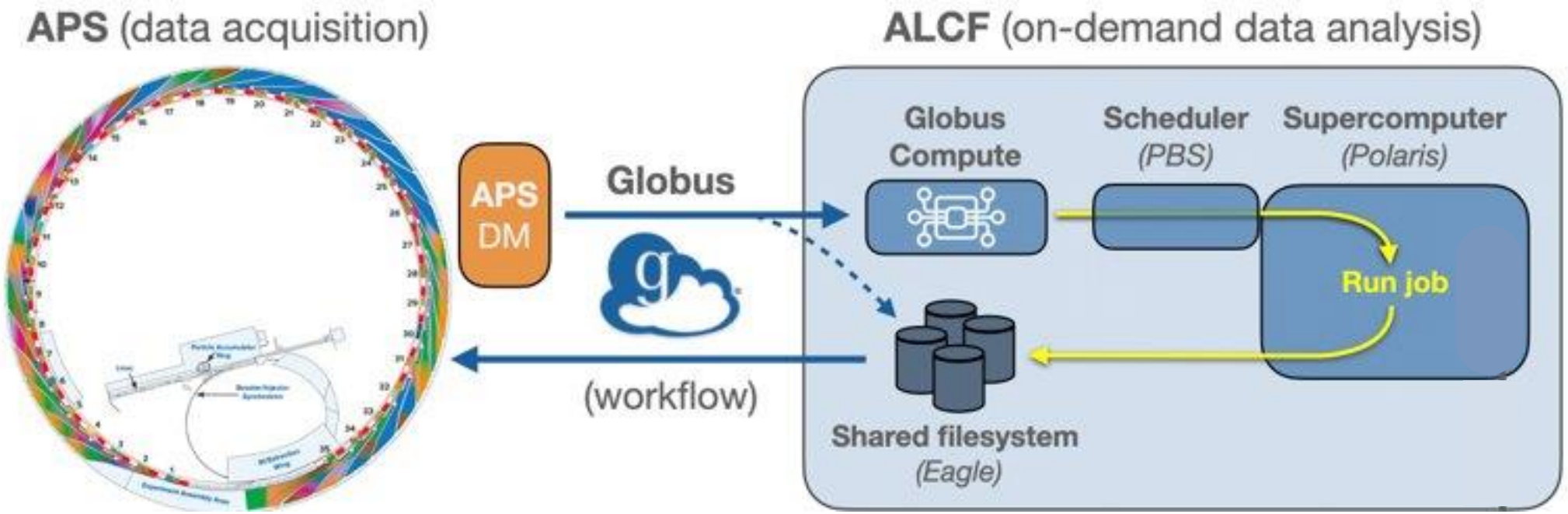
ARGONNE NEXUS

- The **Nexus** initiative at Argonne enables experimental facilities to leverage supercomputing facilities for experiment-time data analysis
- Nexus goals align with the DOE vision for integrated research infrastructure (IRI)
    - Simplify access to DOE computing resources (accounts, job submission)
    - Automate data transfer and workflows triggered by experiment events
    - Provide robust, generalizable solutions, not tailored to a single experiment/facility
    - Support interactive inspection and provenance of data and derived products
    - Publish data for community access, citation, and archiving

ARGONNE NEXUS SERVICES

- Demand Queue
  - Deployed now on Polaris: reduces queue wait time for experiment-time analysis; backfilled by preemptable jobs
- Service Accounts
  - Provides experiment-specific accounts for running analyses in a controlled environment
- Eagle Data Sharing (100PB filesystem)
  - Users can define collections of data to share publicly or with designated Globus users, without an ALCF account
  - Leveraged in ALCF-HEP cosmology data sharing portal
- Globus Infrastructure
  - Distributed compute, transfer, and web-based monitoring
- ALCF Community Data Cooperative (ACDC)
  - Layer atop Eagle data sharing to enable more sophisticated metadata-based navigation and search
  - Metadata extraction and capture services
- Dedicated Web Applications
  - Foundation for user-driven access to data and analysis
  - Developing in collaborations with APS and Argonne-HEP for extension to other science areas
  - Data catalogs resident on Eagle filesystem, searchable via user-provided metadata
  - Community analysis of data supported via back-end integration of workflows on Polaris
  - Reusable web components can be deployed and customized for future science domains

ARGONNE NEXUS: LIGHTSOURCE AUTOMATION



- Integration with the data management (DM) system at APS **allows the workflow to begin as soon as data is obtained**
- Workflow moves data from the APS beamline to ALCF and submits job to demand queue on Polaris
- Results are written to Eagle, where they're reachable via Jupyter, and also returned to APS for evaluation

time intensive
data intensive
long term

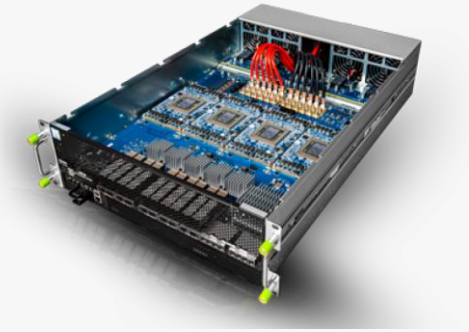
ALCF AI Testbeds

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

Cerebras (CS-2)

SambaNova

Graphcore



Habana

Groq

- Infrastructure of next-generation machines with hardware accelerators customized for artificial intelligence (AI) applications.
- Provide a platform to evaluate usability and performance of machine learning based HPC applications running on these accelerators.
- The goal is to better understand how to integrate AI accelerators with ALCF's existing and upcoming supercomputers to accelerate science insights

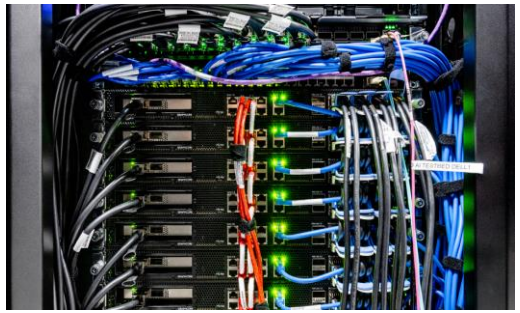
Recent ALCF AI Testbed Updates

ALCF AI Testbed Systems are in production and available for allocations to the research community

https://www.alcf.anl.gov/science/directors-discretionary-allocation-program

SambaNova upgraded to latest 2nd generation SN30 accelerators and scaled to 8 nodes with 64 AI accelerators

SambaNova SN30



Graphcore upgraded to latest Bow generation accelerators and scaled to a Pod-64 configuration with 64 accelerators

Graphcore BowPod64

Cerebras CS-2 upgraded to an appliance mode to include Memory-X and Swarm-X technologies to enable larger models and scaled to two CS-2 engines

Cerebras CS-2

Groq system has been upgraded to a GroqRack with nine nodes, each consisting of eight GroqChip Tensor streaming processors accelerators

GroqRack

https://nairrpilot.org

Director’s Discretionary (DD) awards support various project objectives from scaling code to preparing for future computing competition to production scientific computing in support of strategic partnerships.

Getting Started on ALCF AI Testbed:

Apply for a Director’s Discretionary (DD) Allocation Award

Cerebras CS-2, SambaNova Datascale SN30, Graphcore Bow Pod64, and GroqRack are available for allocations

Allocation Request Form

AI Testbed User Guide

TRILLION PARAMETER CONSORTIUM

Global consortium of scientists to address the challenges of building large-scale artificial intelligence (AI) systems and advancing trustworthy and reliable AI for scientific discovery

GOALS

- **Build an open community of researchers** interested in creating state-of-the-art large-scale generative AI models aimed broadly at advancing progress on scientific and engineering problems by sharing methods, approaches, tools, insights, and workflows.
- **Incubate, launch, and coordinate projects** voluntarily to avoid duplication of effort and to maximize the impact of the projects in the broader AI and scientific community.
- **Create a global network of resources and expertise** to facilitate the next generation of AI and bring together researchers interested in developing and using large-scale AI for science and engineering.

TPC WORK AREAS

- **Identifying and preparing high-quality training data**, with teams organized around the unique complexities of various scientific domains and data sources.
- **Designing and evaluating** model architectures, performance, training, and downstream applications.
- **Developing crosscutting and foundational capabilities** such as innovations in model evaluation strategies with respect to bias, trustworthiness, and goal alignment, among others.

TPC will not hold intellectual property, code or money

AuroraGPT

- Large Language Model for Science
- Interactivity of ChatGPT
- Training from science publications and data (graphs, datasets, etc.)
  - Goal of one trillion parameters
  - Training on Aurora, optimized for Intel PVC
- Looking into multi-modal queries and results
- Argonne-led with external collaborators

National Artificial Intelligence Research Resource Pilot

- NAIRR is a vision for a shared national research infrastructure for responsible discovery and innovation in AI
- The NAIRR pilot will run for two years, beginning January 24, 2024
  - First call for proposals closed in March 2024. More to come.
- ALCF's Contribution to NAIRR
  - Time on AI Testbed
  - Training opportunities

NAIRR FOCUS AREAS

NAIRR Open

This focus area, led by NSF, will support open AI research by providing access to diverse AI resources via the NAIRR Pilot Portal and coordinated allocations.

NAIRR Secure

This focus area, co-led by the National Institutes of Health and the Department of Energy, will support AI research requiring privacy and security-preserving resources and assemble exemplar privacy-preserving resources.

NAIRR Software

This focus area, led by NSF, will facilitate and investigate interoperable use of AI software, platforms, tools and services for NAIRR pilot resources.

NAIRR Classroom

This focus area, led by NSF, will reach new communities through education, training, user support and outreach.

ALCF-4

- Starting the cycle for next ALCF machine
- CD-0 – Mission Need
  - Granted by DOE in 2023
- Now working on technical requirements document
  - Will be posted later this year for comment
  - Trying to be consistent with OLCF requirements document
- Formal RPF
  - Based on technical requirements
  - Looking for a broad set of responses with a variety of approaches

Thanks to Tom Uram, Ben Brown, Venkat Vishwanath and Katherine Riley

Argonne National Laboratory

The U.S. Department of Energy's Argonne National Laboratory delivers world-class research, technologies, and new knowledge that aim to make an impact — from the atomic to the human to the global scale.

About Argonne

Argonne is a multidisciplinary science and engineering research center located outside Chicago.

- Born out of the University of Chicago's work on the Manhattan Project in the 1940s.
- Managed by UChicago Argonne, LLC, for the U.S. Department of Energy's Office of Science.
- Works with universities, industry, and other national labs on questions and experiments too large for any one institution to do by itself.

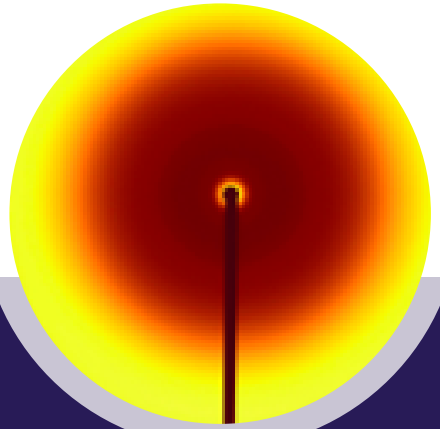
Argonne’s Research Directorates

Computing, Environment and Life Sciences

Couples computing-related activities with science domains whose futures are closely tied to progress in computing

Energy and Global Security

Conducts applied R&D, creates tools that enable scientific and technological breakthroughs, and translates discoveries through engineering to the marketplace



Photon Sciences

Provides the brightest X-ray beams in the Western Hemisphere and provides discoveries in nearly every scientific discipline

Physical Sciences and Engineering

Creates new materials and chemistries and advances accelerator physics

Our one-of-a-kind facilities enable science from the nanoscale to the exascale

Argonne's five flagship facilities support one of the largest user communities in the U.S. Department of Energy complex.

Advanced Photon Source

Argonne Tandem Linear Accelerator System

Argonne Leadership Computing Facility

Center for Nanoscale Materials

Atmospheric Radiation Measurement – The Southern Great Plains


Argonne
NATIONAL LABORATORY