



# ALCF Site Update ACM SIGHPC SYSPROS22 Workshop @SC22

**Susan Coghlan**  
Argonne Leadership Computing Facility  
ALCF-3 Project Director

November 15, 2022

# DOE SC Advanced Scientific Computing Research User Facilities

The Advanced Scientific Computing Research (ASCR) program leads the nation and the world in supercomputing, high-end computational science, and advanced networking for science.

**ALCF and OLCF make up the  
DOE Leadership Computing Facility**

Argonne  
Leadership  
Computing  
Facility  
(ALCF)

Oak Ridge  
Leadership  
Computing  
Facility  
(OLCF)

National Energy  
Research Scientific  
Computing Center  
(NERSC)

Energy Sciences  
Network (ESnet)



# 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



# Broad Engagement in HPC

## We enable and support science campaigns

- ALCF computational scientists assist science teams to ready their codes to efficiently use our resources
- Researchers are supported by performance engineers, user support staff, and data analysis and visualization services
- Each year, ALCF-supported research results in hundreds of refereed publications, in journals such as Proceedings of the National Academy of Sciences, Nature, and Physical Review Letters

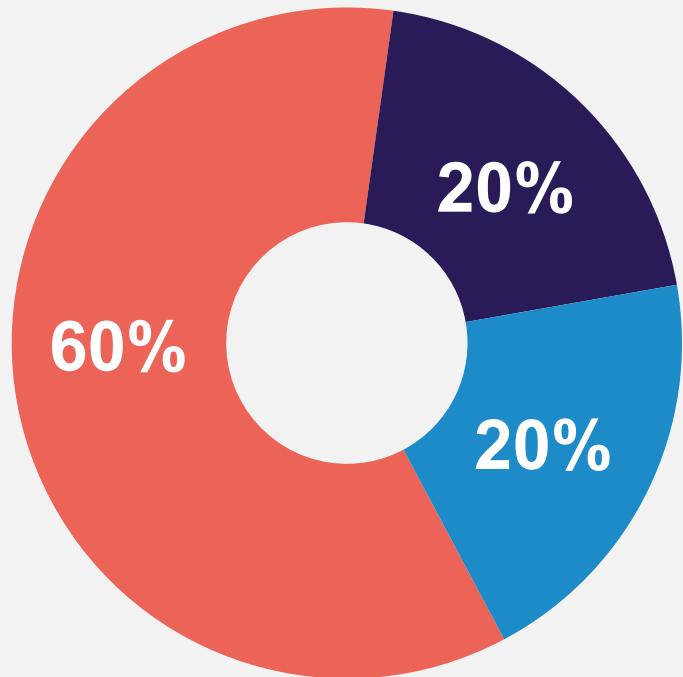
## We deliver cycles to computational scientists

- Delivers millions of node hours of compute time
- Scheduled availability for the resource exceeds 99%

## We partner with community on R&D in hardware and software



# 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

## DD: Director's Discretionary Program

- Rapid allocations for project prep and immediate needs
  - Early Science Program (ESP)
  - Exascale Computing Project (ECP)
  - ALCF Data Science Program (ADSP)
  - Proprietary Projects

# Preparing Users for Exascale

## Early Science Program (ESP)

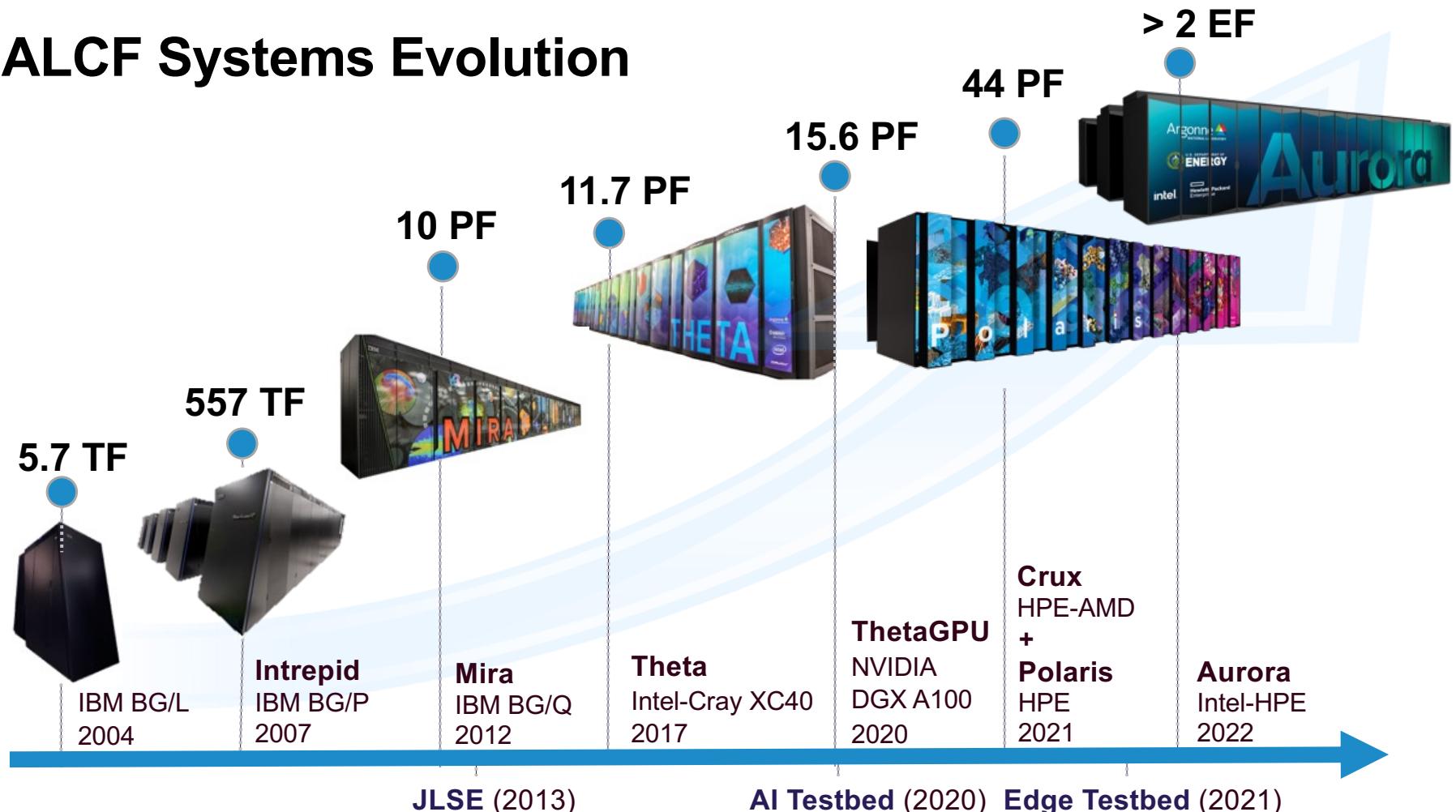
- ALCF conducts ESP to ensure the facility's next-generation systems are ready for science on day one
- Provides research teams with critical pre-production computing time and resources
  - prepares applications for the architecture and scale of a new supercomputer
  - solidifies libraries and infrastructure for other production applications to run on the system
- ALCF hosts hundreds of researchers for a multitude of training opportunities to prepare for Aurora
  - **Aurora Workshops** are held for a multiple days for ESP project members
  - **Aurora Hackathons** are project deep-dives hosted by ALCF/Intel Center of Excellence
  - **Aurora Webinars and On-Demand Videos** are open to all researches who are preparing their code

## 2021 Training Activities

- 13 Workshops
- 16 Webinars
- 8 Hackathons



# ALCF Systems Evolution





PEAK PERFORMANCE

**44 Petaflop DP**

NVIDIA GPU

**A100**

AMD EPYC PROCESSOR

**Milan**

PLATFORM

**HPE Apollo Gen10+**

**Compute Node**

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

**GPU Architecture**

NVIDIA A100 GPU; HBM stack

**Processor Interconnects**

CPU-GPU: PCIe  
GPU-GPU: NVLink

**System Interconnect**

HPE Slingshot 10\*; Dragonfly  
topology with adaptive routing

**Network Switch**

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

**Programming Models**

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

**Node Performance**

78 TF

**Aggregate Memory**

368 TB

**System Size**

560 nodes, 1.78 MW

\*Initial technology to be upgraded later

## Polaris

Polaris provides a platform utilizing several of the Aurora technologies and similar architectures to provide ALCF staff and users a platform for early scaling and testing purposes.

Production August 2022



# 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

# Aurora

Argonne's upcoming exascale supercomputer will leverage several technological innovations to support machine learning and data science workloads alongside traditional modeling and simulation runs.

PEAK PERFORMANCE  
**≥2 Exaflop DP**

INTEL® XEON® ARCHITECTURE-BASED GPU  
**Data Center GPU Max Series**  
INTEL® XEON® SCALABLE PROCESSOR  
**CPU Max Series**  
PLATFORM  
**HPE Cray EX**

## Compute Node

2 Intel® Xeon® CPU Max Series processors; 6 Intel® Data Center GPU Max Series GPUs; Unified Memory Architecture; 8 fabric endpoints; RAMBO

## GPU Architecture

Intel® Data Center GPU Max Series; Tile-based chiplets, HBM stack, Foveros 3D integration, 7nm

## CPU-GPU Interconnect

CPU-GPU: PCIe  
GPU-GPU: X<sup>e</sup> Link

## System Interconnect

HPE Slingshot; Dragonfly topology with adaptive routing

## Network Switch

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

**High-Performance Storage**  
≥230 PB, ≥25 TB/s (DAOS)

## Programming Models

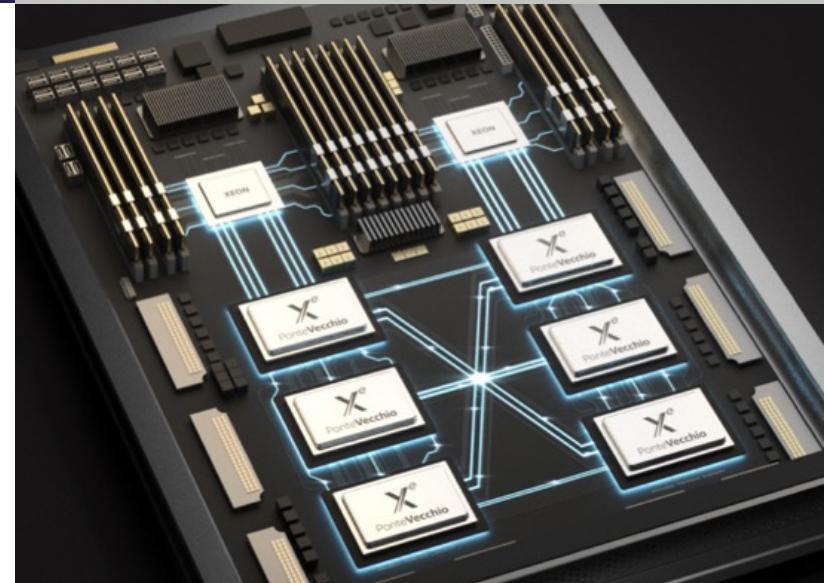
Intel oneAPI, MPI, OpenMP, C/C++, Fortran, SYCL/DPC++

## Node Performance

>130 TF

## System Size

>10,000 nodes



# POWER AND COOLING INFRASTRUCTURE



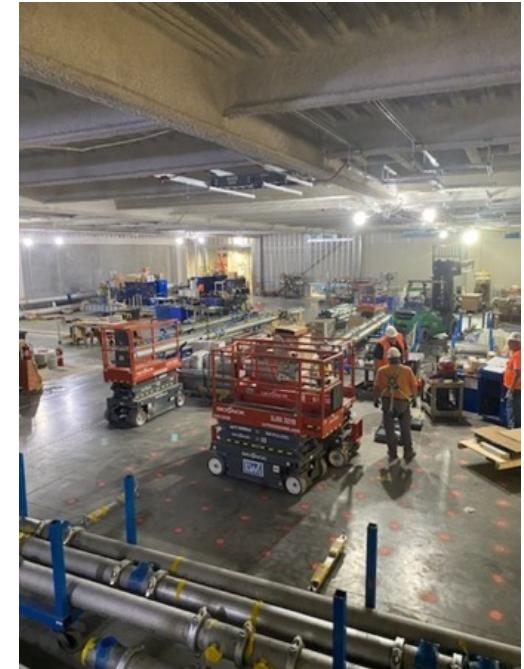
Substations and  
transformers on 2<sup>nd</sup> floor



Switch gear on 2<sup>nd</sup> floor

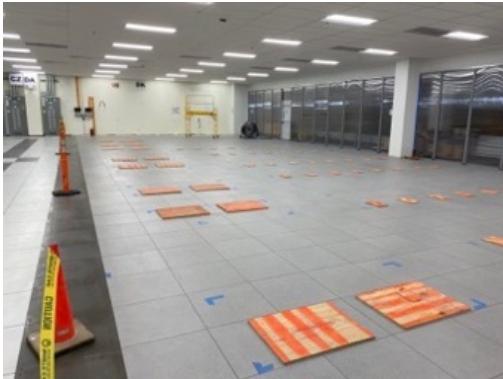


Trench where wall used to be, large header  
pipe put in to provide water for Aurora and  
future systems to South and North of  
previous external wall location



Laying branch pipes  
Floor marking was location  
for raised floor stations

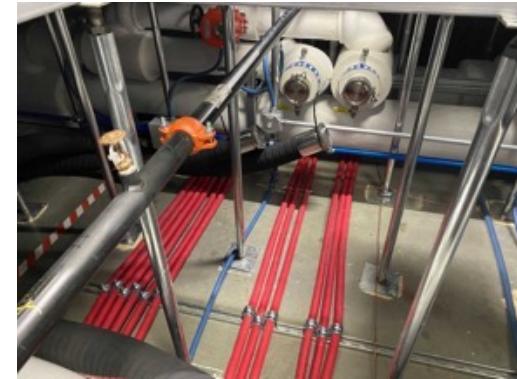
# PREPARATIONS FOR INSTALLATION



Infrastructure rack area after floor cuts were made  
Blocks were installed so cut tiles could safely be left in the data center

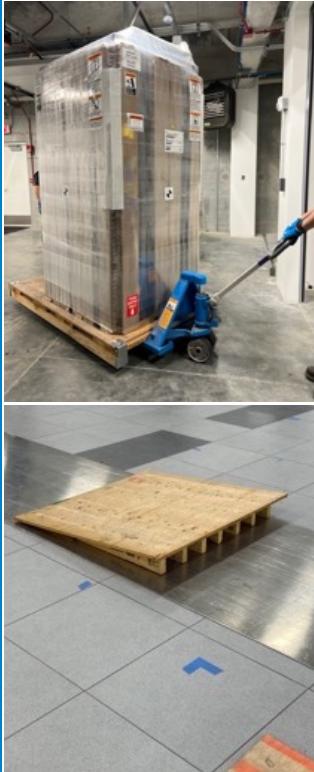


Power Cords from PDU's to infrastructure racks were run out and mounted to unistrut

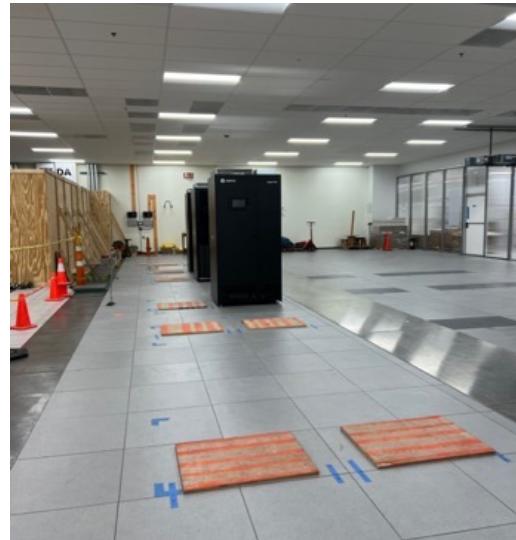


Underfloor piping connections for compute racks

# ON FLOOR PDU INSTALLATION



First on floor equipment



Re-designed ramp for PDUs



Setting PDUs for Infrastructure Racks

Blocks for tiles with cutouts  
Metal path for all deliveries

# INFRASTRUCTURE RACKS



Management  
Logins  
Network  
DAOS Storage



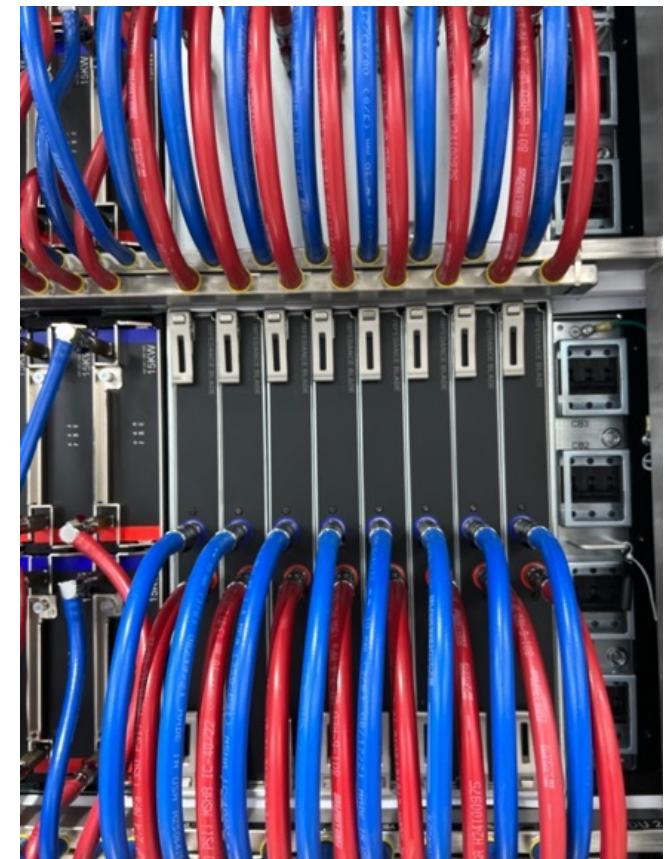
Argonne  
NATIONAL LABORATORY

# MANAGING THROUGH SUPPLY CHAIN CONSTRAINTS

- Aurora has been delivered and built on the data center floor
- Innovations to allow progress to continue
  - Adaptability and flexibility required
  - Plan reviewed and adjusted daily/weekly based on what is available

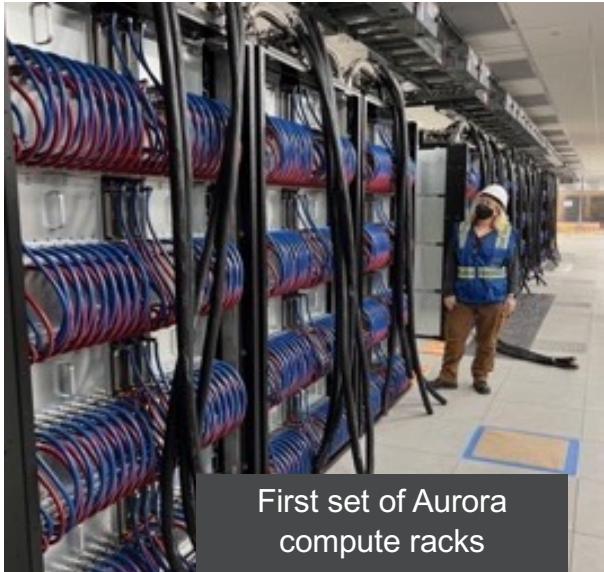


Compute racks prior to compute blade installation

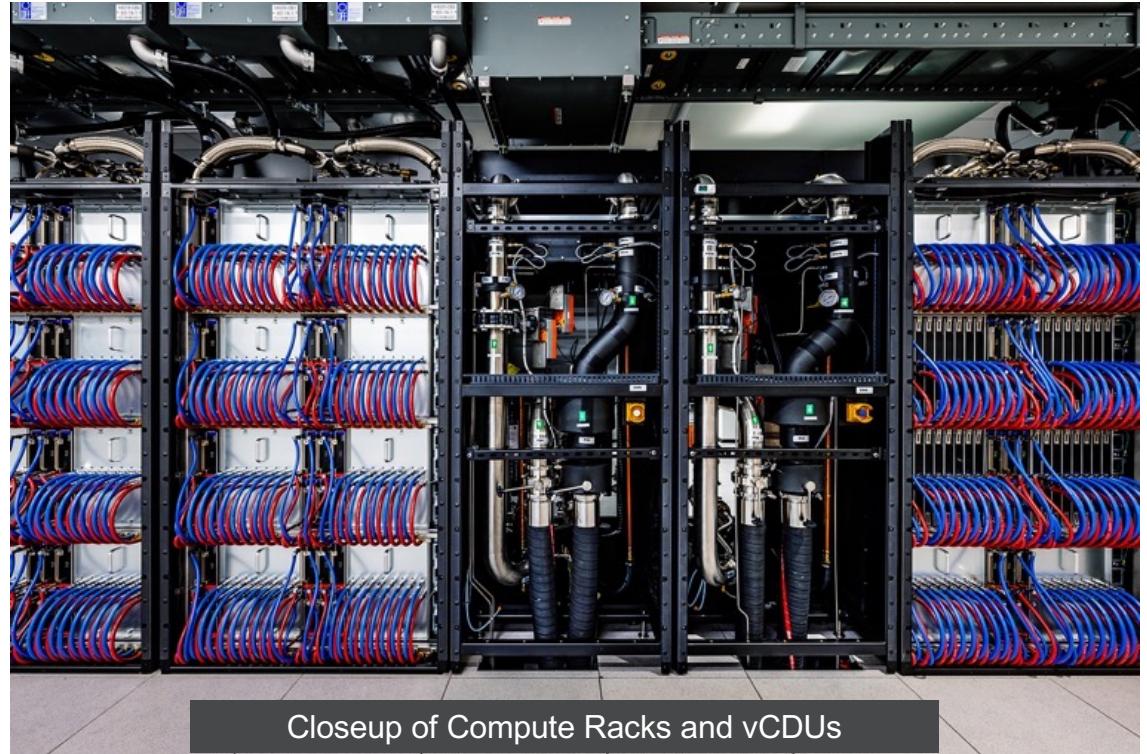


Argonne  
NATIONAL LABORATORY

# AURORA COMPUTE RACKS



First set of Aurora  
compute racks



Closeup of Compute Racks and vCDUs

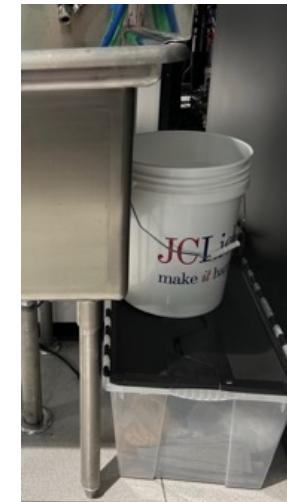
# AURORA SPECIAL COOLING SETUP

- Aurora only Cray Ex system with virtual Cooling Distribution Units (vCDUs)
  - Historical reasons
  - One water loop (secondary loop) through all compute racks
  - Cooled water comes from Argonne area 200 chilled water plant (primary loop) with chillers and cooling towers
  - Hx, pumps, filters, make-up water in building mechanical rooms



# MANAGING WATER COOLING

- Water cooling comes with some challenges
  - Balancing flows
  - Managing pressure
  - Managing water quality
- Aurora scale increases challenges
  - Filter management
  - Mechanical equipment management



Argonne  
NATIONAL LABORATORY

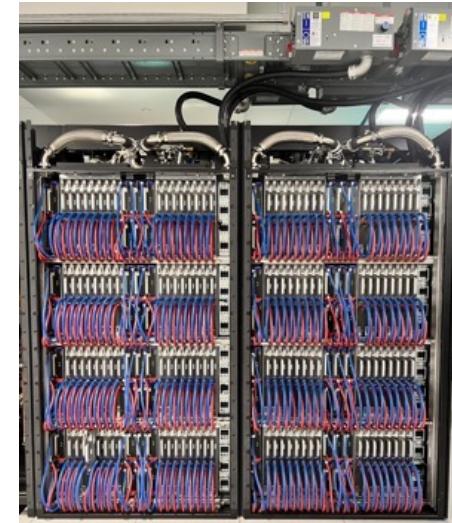
# AURORA AND SUNSPOT STATUS

- Sunspot is fully built and in testing
  - Sunspot is a two rack test and development system
  - Early users will be given access soon
- Aurora is built with exception of compute blades
  - Installation of Exascale Compute Blades (ECBs) has begun

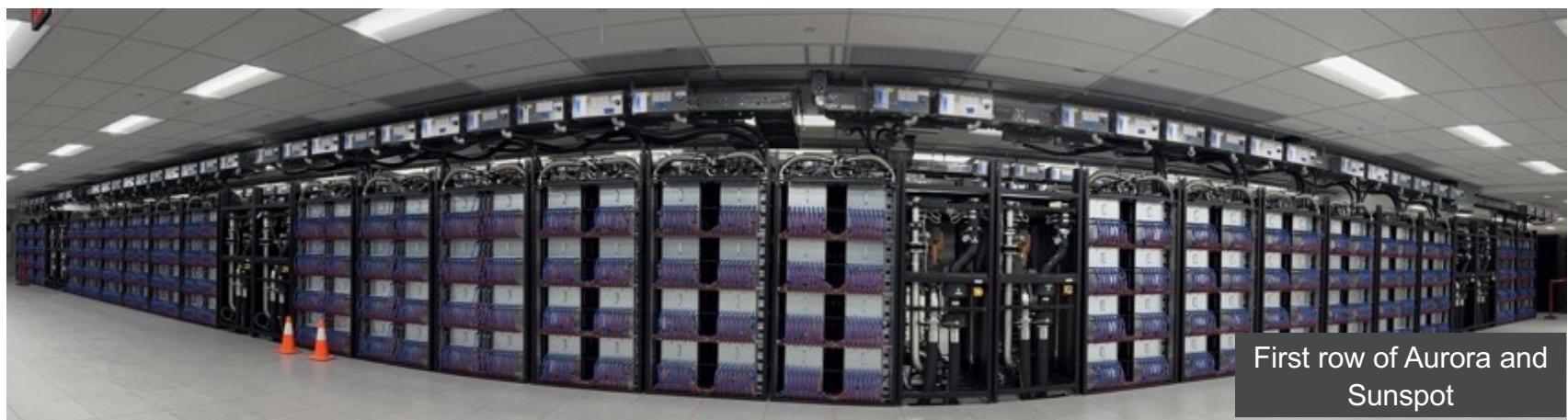
Aurora  
exascale  
compute  
blade



Aurora compute rack



Sunspot compute racks  



# THANK YOU