Argonne Leadership Computing Facility

Accelerating Discovery and Innovation

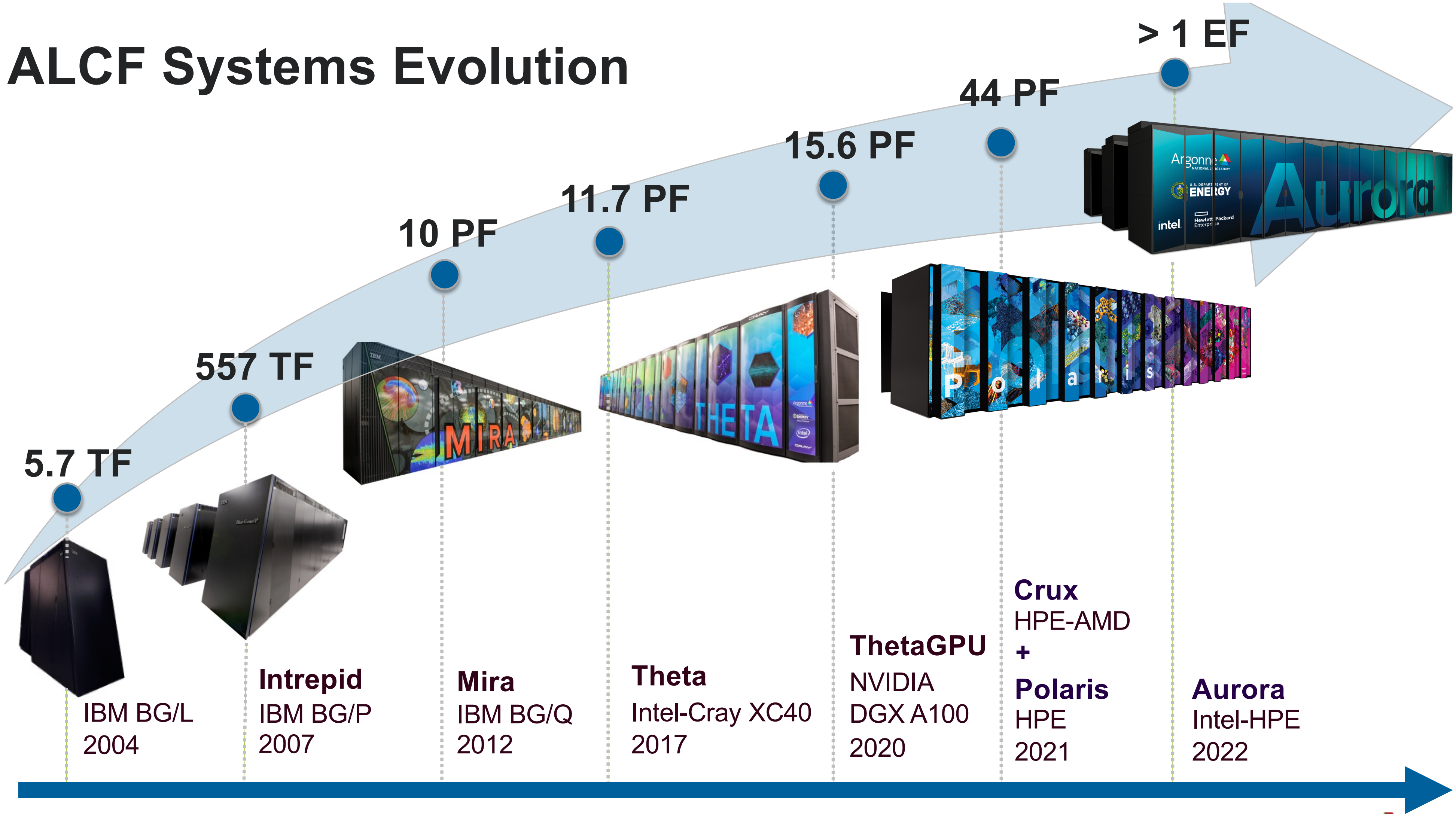
Katherine M Riley
Director of Science, Argonne Leadership Computing Facility

Supercomputing Resources

As a National User Facility we are focused on open science that cannot be readily tackled elsewhere.

Our supercomputers or allocations are 10 to 100 times more powerful than systems typically used for scientific research.

ALCF Systems Evolution



Argonne NATIONAL LABORATORY

System Description	Theta
Total # Compute Racks	24
Total # Compute Nodes	4,392
Total # Cores	281,088
Total # Hardware Threads	1,124,352
Total System Memory	72 TiB MCDRAM 843 TiB DDR4 561 TiB SSD
Interconnect / Topology	Cray Aries / Dragonfly
File System Performance	200 PB (Grand+Eagle, Lustre)
Data Store (PB) (raw)	18 PB, 240 GB/s
Total Peak DP FLOPS (PF)	11.7 PF
Total Floor Space	1,000 sq. ft.

Node Specs	Theta
Node	1.3 GHz Intel “Knights Landing” (KNL) Xeon Phi 7230
Core per node	64
Hardware threads per node	256
Memory per node	16 GiB MCDRAM 192 GiB DDR4 128 GiB SSD

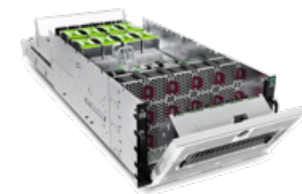
	Peak	Average
Power	2.7 MW	1.7 MW

https://www.alcf.anl.gov/support-center/theta/theta-thetagpu-overview

Argonne NATIONAL LABORATORY

System Configuration	Polaris
# of River Compute Racks	40
# of Apollo Gen10+ Chassis	280
# of Nodes	560
# of AMD EPYC 7532 CPUs	560
# of NVIDIA A100 GPUs	2240
Total GPU HBM2 Memory	87.5 TB
Total CPU DDR4 Memory	280 TB
Total NVMe SSD Capacity	1.75 PB
Interconnect	HPE Slingshot
# of Cassini NICs	1120
# of Rosetta Switches	80
Total Injection BW (w/ Cassini)	28 TB/s
Total GPU DP Tensor Core Flops	44 PF
Total Power	1.8 MW

Single Node Configuration	Polaris
# of AMD EPYC 7532 CPUs	1
# of NVIDIA A100 GPUs	4
Total HBM2 Memory	160 GB
HBM2 Memory BW per GPU	1.6 TB/s
Total DDR4 Memory	512 GB
DDR4 Memory BW	204.8 GB/s
# of NVMe SSDs	2
Total NVMe SSD Capacity	3.2 TB
# of Cassini NICs	2
Total Injection BW (w/ Cassini)	50 GB/s
PCIe Gen4 BW	64 GB/s
NVLink BW	600 GB/s
Total GPU DP Tensor Core Flops	78 TF



Apollo 6500 Gen10+

NVIDIA HGX A100 4-GPU

Argonne NATIONAL LABORATORY

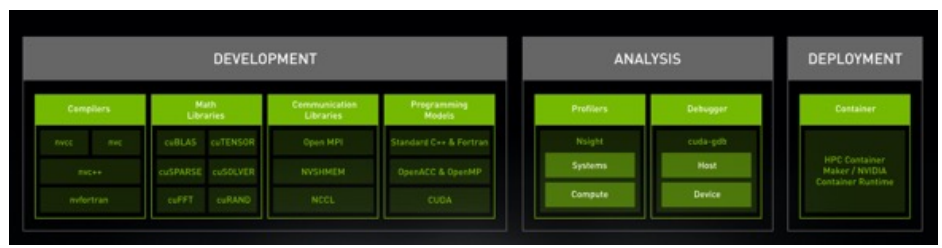
ALCF Resources - Polaris

Software Overview

- Provides an excellent platform for preparing application codes for Aurora
  - All programming models available on Aurora can be tested
  - Features HPE Cray (PE) Programming Environment
  - Built with HPE HPCM system software
- Provides excellent capabilities in simulation, data and learning using Nvidia’s existing HPC SDK
- Support for HPE Cray MPI and MPICH via libfabric using Slingshot provider
  - Initial SS10 feature support
  - later full SS11 feature support for testing all MPI features available on Aurora

Programming Environment

- HPE Cray PE for Polaris
  - HPE Cray MPI support for PGI offload to A100 for Multi-NIC and Multi-GPU support
  - Full Rome and Milan support
- SYCL/Data Parallel C++ provided via
  - CodePlay computecpp compiler with Nvidia support
  - LLVM via Intel DPC++ branch which supports offload to Nvidia GPUs as well as Intel GPUs
- Next NVIDIA HPC SDK will provide primary support for programming A100



Bridge to Aurora

- Polaris will provide a platform for preparation for Aurora
- Polaris and Aurora will have many similarities at the system and user level

Component	Polaris	Aurora
System Software	**HPCM**	HPCM
Programming Models	**MPI, OpenMP, DPC++, Kokkos, RAJA, HIP**, CUDA, OpenACC	**MPI, OpenMP, DPC++, Kokkos, RAJA, HIP**
Tools	**PAT, gdb, ATP,** NVIDIA Nsight, cuda-gdb	**PAT, gdb, ATP,** Intel Vtune
MPI	**HPE Cray MPI, MPICH**	**HPE Cray MPI, MPICH,** Intel MPI
Multi-GPU	1 CPU : 4 GPU	2 CPU : 6 GPU
Data and Learning	**DL frameworks, Cray AI stack, Python/Numba, Spark, Containers,** Rapids	**DL frameworks, Cray AI stack, Python/Numba, Spark, Containers,** oneDAL
Math Libraries	cu* from CUDA	oneAPI

Argonne NATIONAL LABORATORY

Exascale Architecture Designed for Simulation, Data & Learning

- Sustained performance of ≥1EF
- >10 PB of aggregate memory
- Compute nodes with 2 Intel Xeon CPUs and 6 X^e GPUs with unified shared memory
- Cray Slingshot™ fabric in Dragonfly topology with 8 endpoints per node
- ≥230 PB of Distributed Asynchronous Object Storage (DAOS) with ≥ 25 TB/s throughput
- Global file systems with over 200 PB of storage and ~1.3 TB/s throughput
- Cray Shasta™ software stack, Intel software, and Data & Learning frameworks
    - Optimized PyTorch, TensorFlow, Python-based ML and analytics such as OneDAL
- Programming models include OpenMP 5, SYCL/DPC++, OpenCL, Kokkos, and Raja

https://aurora.alcf.anl.gov

Argonne NATIONAL LABORATORY

ALCF Resources - Aurora

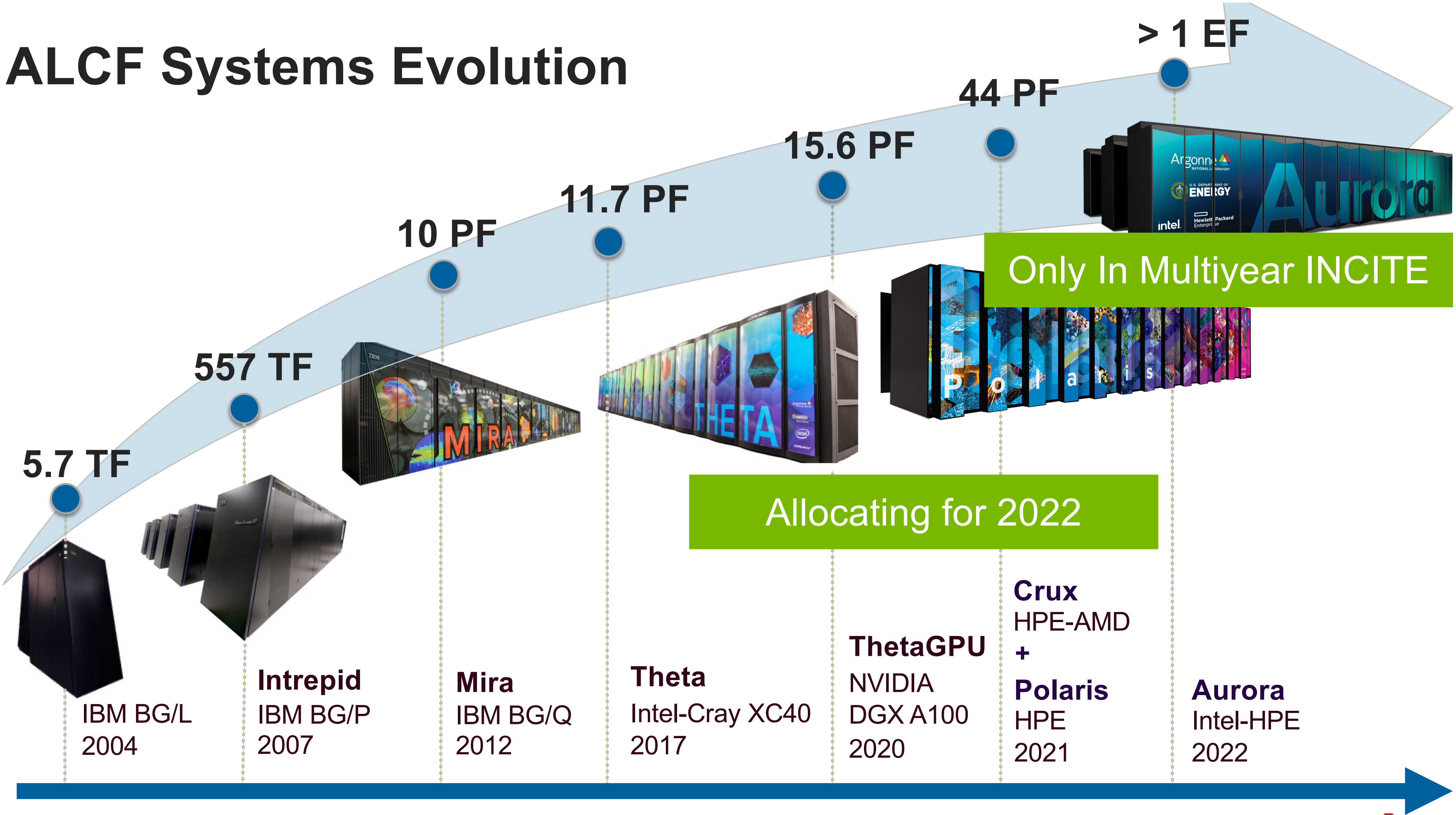
System Configuration	Aurora
Peak Performance	≥1 EF DP sustained
Node	2 Intel Xeon scalable processors (Sapphire Rapids); 6 Intel X^e arch based GPU (Ponte Vecchio)
GPU Architecture	X^e arch based GPU (Ponte Vecchio); Tile based, chiplets, HBM stack, Foveros 3d integration
CPU-GPU Interconnect	PCIe
Aggregate System Memory	>10 PB
Interconnect	HPE Slingshot 11 Dragonfly topology with adaptive routing
Network Switch	25.6 TB/s per switch, from 64-200 GB ports (25GB/s per direction)
High-Performance Storage	≥230 PB, ≥25 TB/s (DAOS)

Configuration	Aurora
Node Performance	> 130 TF
Compute nodes	> 9,000
Cabinets	> 100
Node Memory Architecture	Unified memory architecture, RAMBO

https://aurora.alcf.anl.gov

Argonne NATIONAL LABORATORY

ALCF Systems Evolution



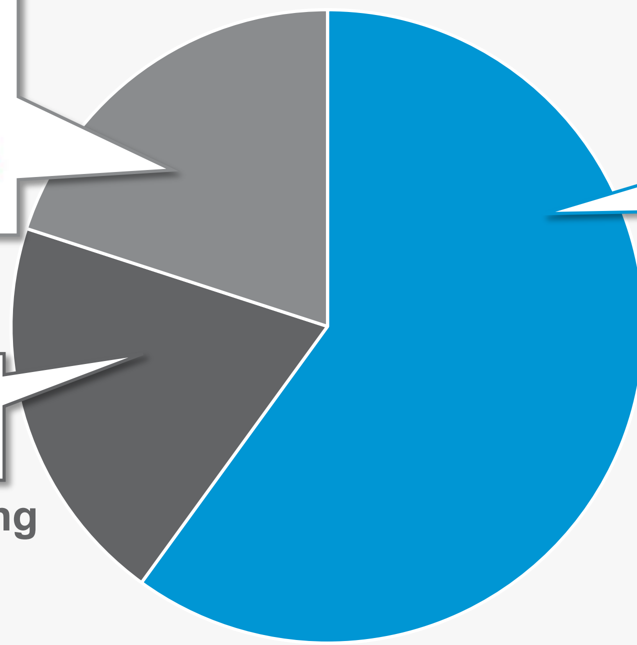
How Do Researchers Gain Access to ALCF?

Primary Allocation Programs for Access to the LCFs

Current distribution of allocatable hours

20% Director's Discretionary

- LCF targeted programs, e.g. ADSP, ESP
- Proposal Prep
- ECP



60% INCITE

Leadership-class computing

20% ASCR Leadership Computing Challenge

DOE/SC capability computing

LCF Allocation Programs	INCITE 60%		ALCC 20%		Director's Discretionary 20%	
Mission	High-risk, high-payoff science that requires LCF-scale resources*		High-risk, high-payoff science aligned with DOE mission		50% Strategic LCF goals 50% ECP	
Call	1x/year – Opens in April, Closes June		1x/year – Opens in November, Closes February		Rolling	
Duration	1-3 years, yearly renewal		1 year		3m,6m,1 year	
Typical Size	**10-15 projects**	**1-3M node-hours**	**5-15 projects**	**0.5-2M node-hours**	**~100 of projects**	**<0.5M node-hours**
Total Hours	**~17.8M Theta node-hours ~2.0M Polaris node-hours**		**~6M Theta node-hours**		**~6M Theta node-hours**	
Review Process	Scientific Peer-Review	Computational Readiness	Scientific Peer-Review	Computational Readiness	Strategic impact and feasibility	
Managed By	INCITE management committee (ALCF & OLCF)		DOE Office of Science		LCF management	
Readiness	**High**		**Medium to High**		**Low to High**	
Availability	Open to all scientific researchers and organizations ***Capability > 20% of resource***					

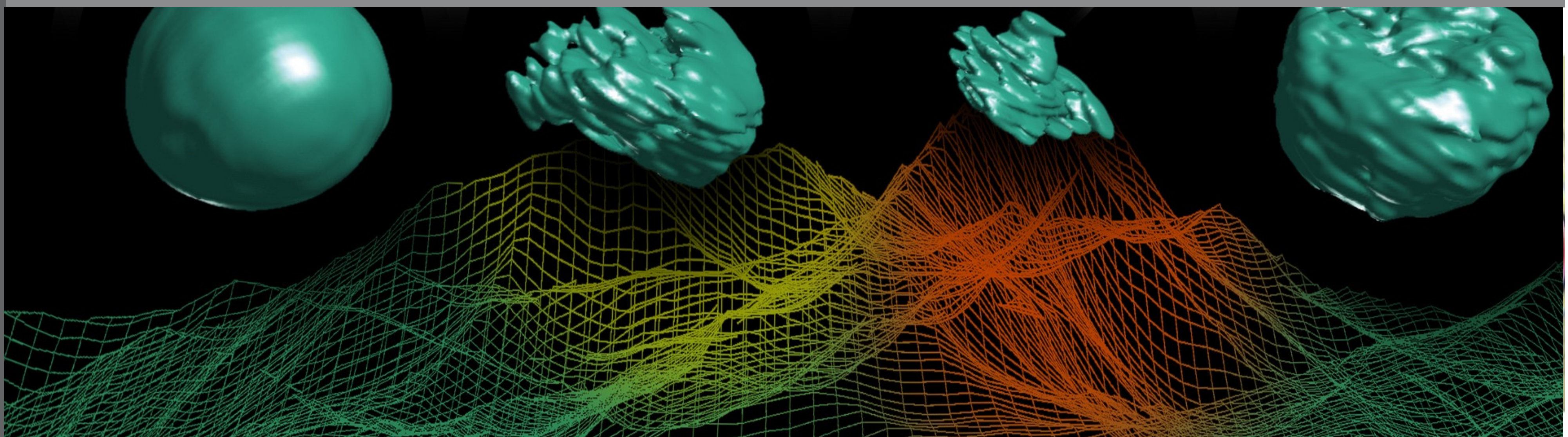
LCF Allocation Programs	INCITE 60%	ALCC 20%	Director's Discretionary 20%
Mission	High-risk, high-payoff science that requires LCF-scale resources*	High-risk, high-payoff science aligned with DOE mission	50% Strategic LCF goals 50% ECP
Call	1x/year – Opens in April, Closes June	1x/year – Opens in November, Closes February	Rolling
Duration	1-3 year	1 year	3m,6m,1 year
Typical Size	**10-15 pro**	**5-15 projects** **0.5-2M node-hours**	**~100 of projects** **<0.5M node-hours**
Total Hours	**~17.8M Theta node-hours** **~2.0M Polaris node-hours**	**~6M Theta node-hours**	**~6M Theta node-hours**
Review Process	Scientific Peer-Review / Computational Readiness	Scientific Peer-Review / Computational Readiness	Strategic impact and feasibility
Managed By	INCITE management committee (ALCF & OLCF)	DOE Office of Science	LCF management
Readiness	**High**	**Medium to High**	**Low to High**
Availability	Open to all scientific researchers and organizations ***Capability > 20% of resource***		

New Early Career Track

Getting Started (DD)

Our Director's Discretionary (DD) allocation program provides researchers with small awards of computing time to "get started" on our computing resources while pursuing real scientific goals.

The DD allocation program allows users to prep their code so that it can take advantage of our massively parallel systems.



DD
Director’s Discretionary

Purpose: A “first step” for projects working toward a major allocation

Eligibility: Available to all researchers in academia, industry, and other research institutions

Review Process: Projects must demonstrate a need for high-performance computing resources; reviewed by ALCF

Award Size: Low 10 thousand of node-hours

Award Duration: 3-6 months, renewable

Total percent of ALCF resources allocated: 20%

Award Cycle

Ongoing (available year-round)

ADSP
ALCF Data Science Program

Targeted at big data science problems, ADSP aims to explore and improve a variety of computational methods that will help enable data-driven discoveries across all scientific disciplines.

Eligibility: Available to researchers in academia, industry, and other research institutions

Review process: Applications undergo a review process to evaluate potential impact, data scale readiness, diversity of science domains and algorithms, and other criteria

Award size: ~Low hundred of thousand of node-hours

Award duration: 2 years

ESP
Early Science Program

As part of the process of bringing a new supercomputer into production, the ALCF hosts the Early Science Program (ESP) to ensure its next-generation systems are ready to hit the ground running.

The intent of the ESP is to use the critical pre-production time period to prepare key applications for the architecture and scale of a new supercomputer, and to solidify libraries and infrastructure to pave the way for other production applications to run on the system.

In addition to fostering application readiness, the ESP allows researchers to pursue innovative computational science projects not possible on today's leadership-class supercomputers.

Award Cycle

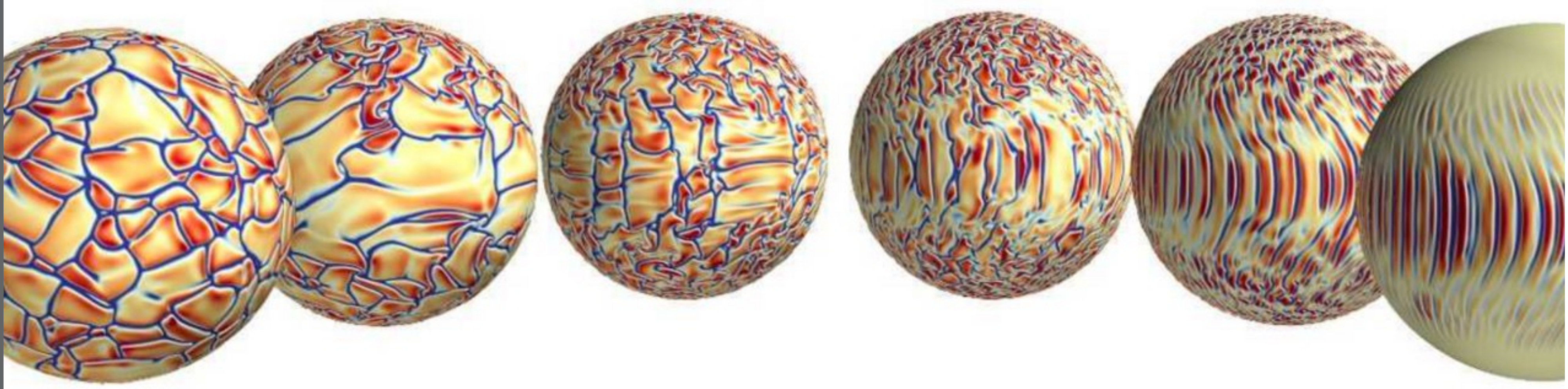
Determined by production timeline

Argonne NATIONAL LABORATORY

Major Awards (INCITE, ALCC)

Our major allocations provides users with computationally intensive, large-scale research projects time on our machines.

The programs conduct a two-part review of all proposals: a peer review by a panel of experts and a computational readiness review.



ALCC
ASCR Leadership Computing Challenge

The DOE's ALCC program allocates resources to projects directly related to the DOE's energy mission, as well as national emergencies, and for broadening the community of researchers capable of using leadership computing resources.

Eligibility: Available to researchers in academia, industry, and other research institutions

Review process: DOE peer reviews all proposals for scientific/technical merit; appropriateness of approach; and adequacy of personnel and proposed resources

Award size: ~1M node-hours

Award duration: 1 year

Total percent of ALCF resources allocated: 20%

Award Cycle
July 1 to June 30

Call open in fall.
Often need LOIs.

INCITE

Innovative & Novel Computational Impact on Theory and Experiment

The DOE's INCITE program provides allocations to computationally intensive, large-scale research projects that aim to address "grand challenges" in science and engineering.

Eligibility: Available to researchers in academia, industry, and other research institutions

Review process: INCITE program conducts a two-part review of all proposals including a peer review by an international panel of experts, and a computational-readiness review

Award size: ~1.0-2.5M node-hours

Award duration: 1-3 years, renewable

Total percent of ALCF resources allocated: 60%

Award Cycle
January 1 to December 31

Call Opens in April

INCITE criteria

Access on a competitive, merit-reviewed basis*

1	**Merit criterion**
	Research campaign with the potential for significant domain and/or community impact
2	**Computational leadership criterion**
	Computationally demanding runs that cannot be done anywhere else: capability, architectural needs
3	**Eligibility criterion**
	• Grant allocations regardless of funding source* • Non-US-based researchers are welcome to apply

*DOE High-End Computing Revitalization Act of 2004: Public Law 108-423

Argonne NATIONAL LABORATORY

Twofold review process

	New proposal assessment	Renewal assessment
1 Peer review: INCITE panels	• Scientific and/or technical merit • Appropriateness of proposal method, milestones given • Team qualifications • Reasonableness of requested resources	• Change in scope • Met milestones • On track to meet future milestones • Scientific and/or technical merit
2 Computational readiness review: LCF centers	• Technical readiness • Appropriateness for requested resources	• Met technical/ computational milestones • On track to meet future milestones
Award Decisions	• INCITE Awards Committee comprised of LCF directors, INCITE program manager, LCF directors of science, sr. management	

Recent Trends in INCITE

Data, Learning and Nontraditional Uses of the Architecture

- In addition to traditional computationally intensive simulation campaigns, INCITE encourages Data and/or Learning projects with unique data requirements (e.g. large scale data analytics) or workflow needs that can only be enabled by the LCFs.
  - A “Learning” panel evaluated proposals that had significant machine / deep learning component to their campaign
  - When appropriate, these proposals were also assessed by their scientific discipline peers as well

2020 and 2021 award statistics, by system

2020	Summit	Theta
Number of projects*	39	14
Average Project	482 K	1.41 M
Median Project	500 K	1.50 M
Total Awards (node-hrs in CY2020)	18.8 M	19.7 M

* Total of 47 INCITE projects (6 projects received time on both Theta and Summit)
* All reported in node-hours native to each resource.

2021	Summit	Theta
Number of projects*	41	16
Average Project	468 K	1.2 M
Median Project	470 K	1.2 M
Total Awards (node-hrs in CY2020)	19.2 M	18.5 M

* Total of 51 INCITE projects (6 projects received time on both Theta and Summit)
* All reported in node-hours native to each resource.

Early Career Track in INCITE

For the INCITE 2022 Call for Proposals, INCITE is committing 10% of allocatable time to an Early Career Track in INCITE.

The goal of the early career track is to encourage the next generation of high-performance computing researchers.

Researchers within 10 years from earning their PhD (on or after December 31st 2011) may choose to apply.

Projects will go through the regular INCITE Computational Readiness and Peer Review process, but the INCITE Management Committee will consider meritorious projects in the Early Career Track separately.

LCF Allocation Programs	INCITE 60%		ALCC 20%		Director's Discretionary 20%	
Mission	High-risk, high-payoff science that requires LCF-scale resources*		High-risk, high-payoff science aligned with DOE mission		50% Strategic LCF goals 50% ECP	
Call	1x/year – Opens in April, Closes June		1x/year – Opens in November, Closes February		Rolling	
Duration	1-3 years, yearly renewal		1 year		3m,6m,1 year	
Typical Size	**10-15 projects**	**1-3M node-hours**	**5-15 projects**	**0.5-2M node-hours**	**~100 of projects**	**<0.5M node-hours**
Total Hours	**~17.8M Theta node-hours ~2.0M Polaris node-hours**		**~6M Theta node-hours**		**~6M Theta node-hours**	
Review Process	Scientific Peer-Review	Computational Readiness	Scientific Peer-Review	Computational Readiness	Strategic impact and feasibility	
Managed By	INCITE management committee (ALCF & OLCF)		DOE Office of Science		LCF management	
Readiness	**High**		**Medium to High**		**Low to High**	
Availability	Open to all scientific researchers and organizations ***Capability > 20% of resource***					

Thank You!

Learn more at: alcf.anl.gov