ARGONNE LEADERSHIP COMPUTING FACILITY

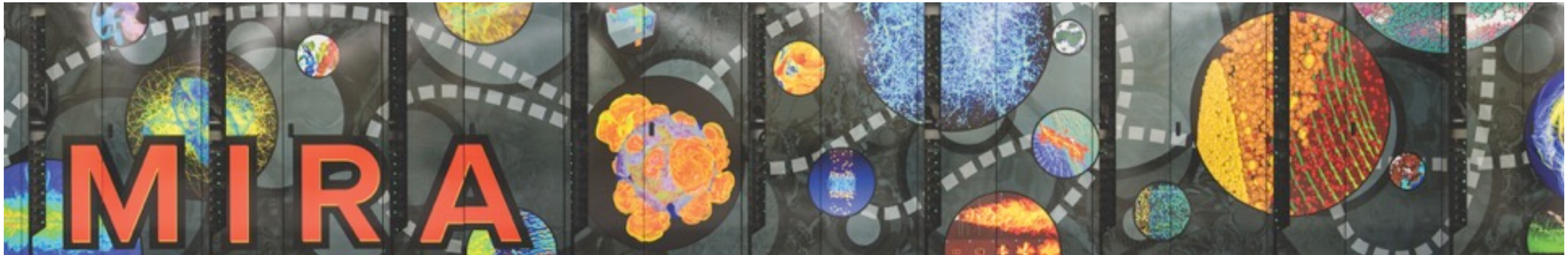
Susan Coghlan
ALCF Project Director

Argonne National Laboratory
22 July 2015

DOE LEADERSHIP COMPUTING FACILITY

Leadership Class Resources

- Dedicated to high impact breakthrough open science
- Supported by DOE's Advanced Scientific Computing Research Program
- Two Centers – ALCF & OLCF
- Two of the world's most powerful supercomputers
- Two diverse architectures



ARGONNE LEADERSHIP COMPUTING FACILITY

Mira supercomputer

Accelerates major scientific discoveries and engineering breakthroughs for humanity

- 10 PetaFlops peak
- 49,152 nodes with 786,432 cores
- 786 TeraBytes of memory
- 26 PB scratch FS, 400 GB/s sustained

Argonne's
MIRA
supercomputer is the fifth-fastest in the world.

ALCF RESOURCES

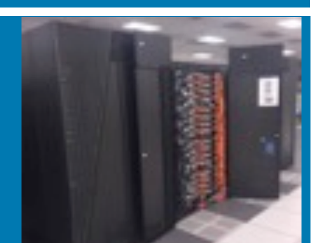
Mira - compute

- 10 PF IBM BG/Q
- 48K nodes/786K cores
- 786 TB memory
- 5D Torus interconnect
- 26 PB GPFS, 400 GB/s

Cooley – data analytics

- 223 TF
- 126 nodes/1512 Xeon cores/126 Tesla K80 GPUs
- 384 TB (CPU)/3 TB (GPU) memory
- FDR InfiniBand interconnect
- Connected to Mira file systems

Cetus – app d&d



- 840 TF IBM BG/Q
- 4K nodes/64K cores
- 64 TB memory
- 5D Torus interconnect
- Connected to Mira file systems

Vesta – system SW d&d

- 420 TF IBM BG/Q
- 2K nodes/32K cores
- 32 TB memory
- 5D Torus interconnect
- 1 PB GPFS,

THREE PRIMARY WAYS TO ACCESS LCF

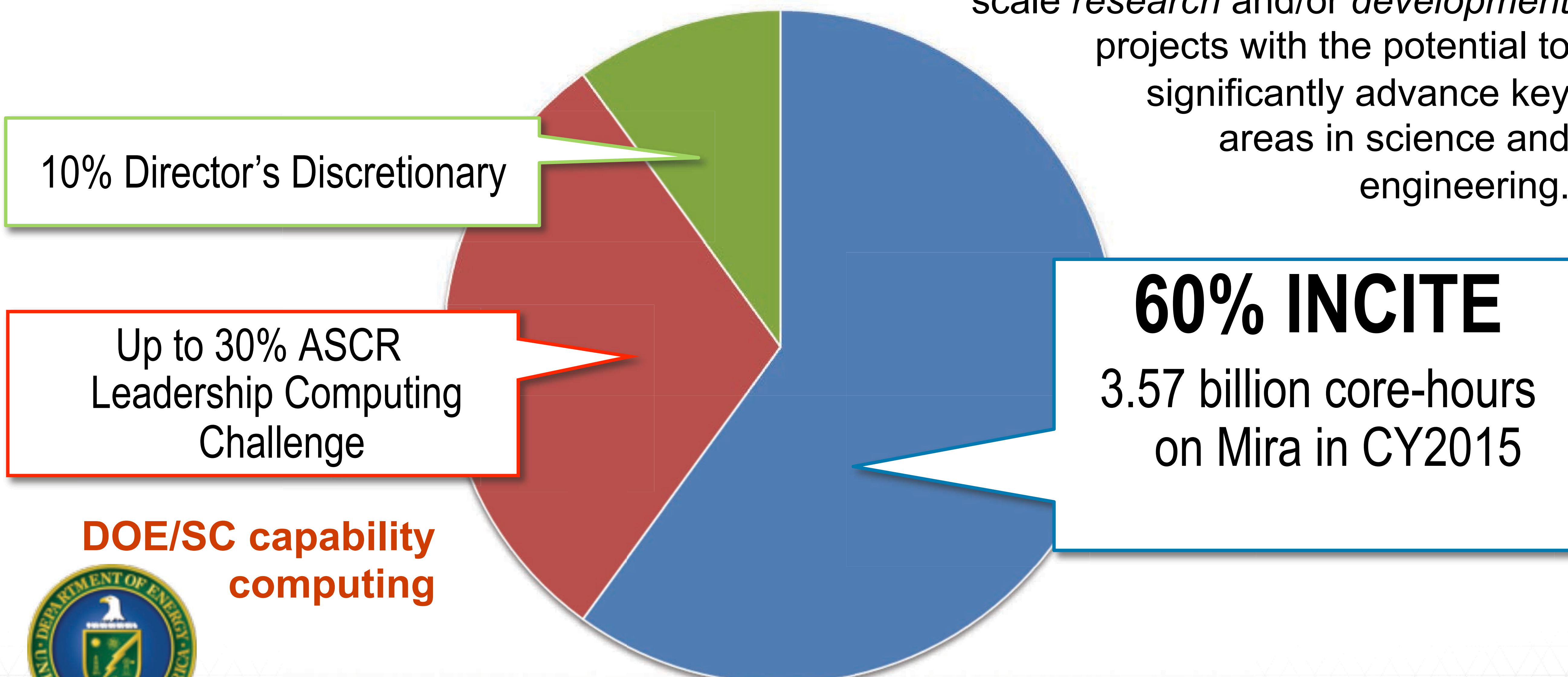
DISTRIBUTION OF ALLOCABLE HOURS

Argonne NATIONAL LABORATORY

OAK RIDGE National Laboratory

Leadership-class computing

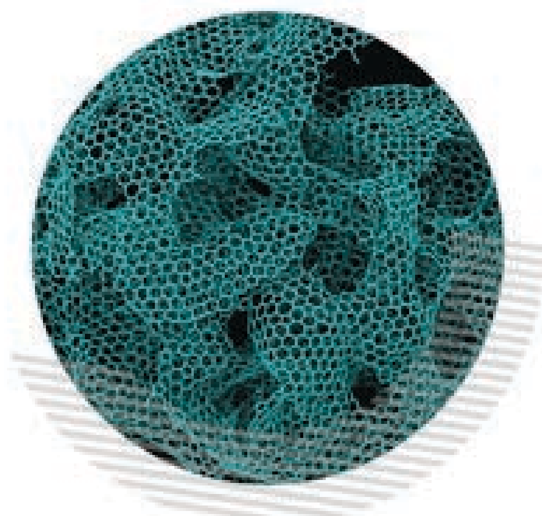
INCITE seeks computationally intensive, large-scale *research* and/or *development* projects with the potential to significantly advance key areas in science and engineering.



DOE/SC capability computing

ALCF SCIENCE HIGHLIGHTS

Materials Science



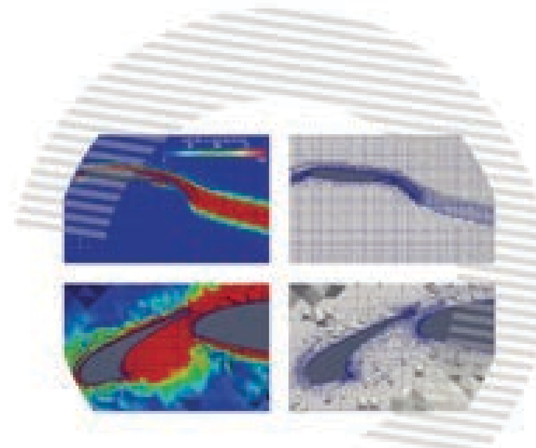
Petascale Simulations of Self-Healing Nanomaterials

INCITE 2015

Rajiv Kalia, University of Southern California

Researchers are performing massive simulations on Mira to help advance the understanding and viability of self-healing nanomaterial systems, which are capable of sensing and repairing damage in materials operating in harsh environments. Ultimately, this work could help enhance the reliability and reduce the cost of components for many energy applications, including high-temperature, turbines and wind and solar energy technologies.

Engineering



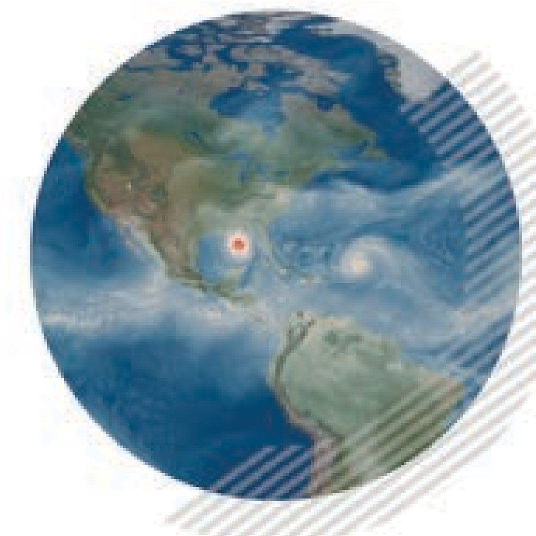
Adaptive Detached Eddy Simulation of a High Lift Wing with Active Flow Control

INCITE 2015

Kenneth Jansen, University of Colorado Boulder

Recent developments in parallel adaptive meshing and parallel solver technology are yielding fundamental insight into the complicated physics of flow control on real aircraft configurations. Using these techniques, project researchers are modeling an array of synthetic jets that have been specifically vectored to improve the aerodynamic performance of wing profiles in aeronautics, which could lead to more-efficient aircraft designs.

Earth Science



Accelerated Climate Modeling for Energy

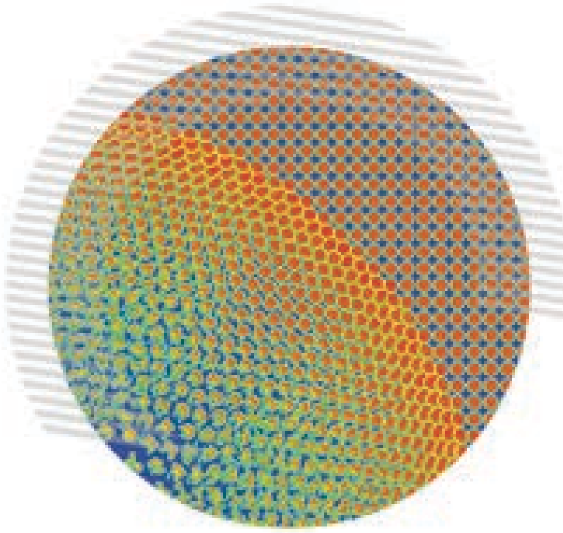
INCITE 2015

Mark Taylor, Sandia National Laboratories

DOE's Accelerated Climate Modeling for Energy (ACME) project aims to develop and apply a computationally advanced climate and Earth system model to investigate the challenges posed by the interactions of climate change and societal energy requirements. This is the only major national modeling project designed to address DOE mission needs to efficiently utilize DOE leadership computing resources now and in the future.

ALCF SCIENCE HIGHLIGHTS

Chemistry



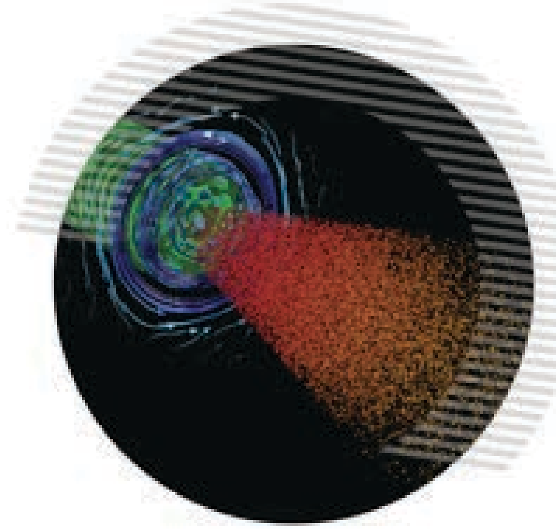
Solving Petascale Public Health and Safety Problems Using Uintah

INCITE 2014

Martin Berzins, University of Utah

Researchers carried out large-scale 3D simulations to study the physical mechanisms that led to an explosion when a semi-truck hauling explosives crashed and caught fire in Utah in 2005. With a better understanding of the deflagration-to-detonation transition (DDT) process and the development of potential mitigation strategies, this project stands to improve the safe storage and transport of explosives.

Physics



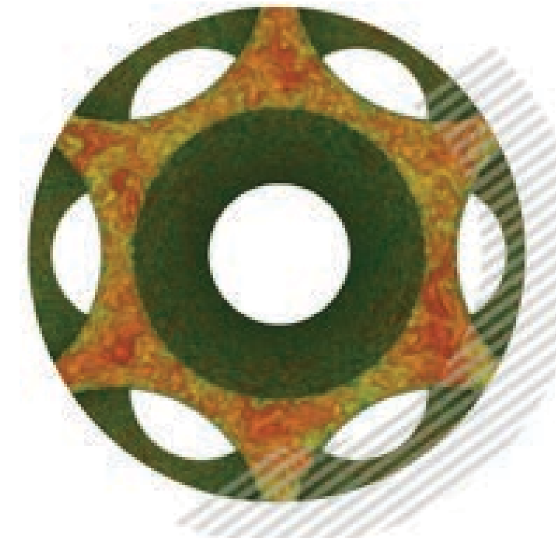
Particle Acceleration in Shocks: From Astrophysics to Laboratory In Silico

INCITE 2015

Frederico Fiuza, Lawrence Livermore National Laboratory

Researchers are using ab initio particle-in-cell simulations to study the physics of shock formation and particle acceleration that is relevant in a number of scenarios, ranging from cosmic-ray acceleration, to the generation of compact ion sources for tumor therapy, to inertial fusion energy. These efforts are expected to help solve some of the central questions in plasma/relativistic phenomena in astrophysics and in the laboratory.

Engineering



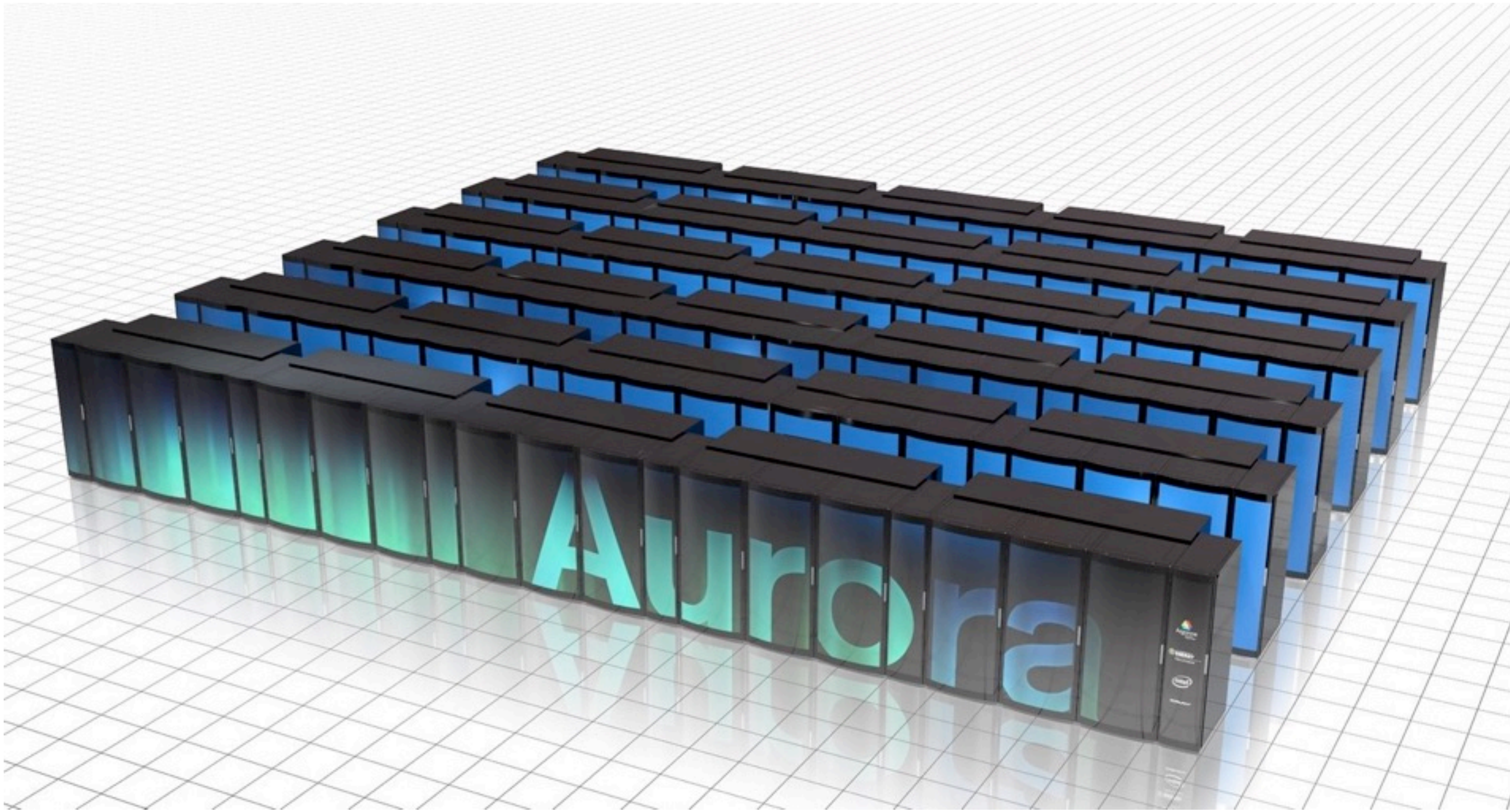
Petascale Simulations in Support of CESAR

ALCC 2014

Elia Merzari, Argonne National Laboratory

This project performs large petascale-level simulations in support of DOE's Center for Exascale Simulation for Advanced Reactors (CESAR), which aims to develop a coupled, next-generation nuclear reactor core simulation tool capable of efficient execution on exascale computing platforms. The resulting simulation code will aid advancements in nuclear engineering and nuclear energy.

THE FUTURE



CORAL

COLLABORATION OF OAK RIDGE, ARGONNE, AND LIVERMORE

- Acquire DOE 2018 – 2023 Leadership Computing Capability
- Three leadership class systems – one each at ALCF, LLNL, OLCF
  - With arch diversity between ALCF and OLCF
- ALCF: Intel (Prime) Cray (Integrator)
- OLCF: IBM (Prime)
- LLNL: IBM (Prime)

TWO NEW ALCF SYSTEMS THETA AND AURORA

- Intel Xeon Phi compute architecture
- Deep memory architecture – very fast memory, slower capacity memory, burst buffers

THE RIGHT PATH FOR OUR USERS

- Many core evolution
- Easy to port codes
- Well-balanced between compute, memory, network, and storage
- Robust and well-known Cray user environment combined with Intel innovations

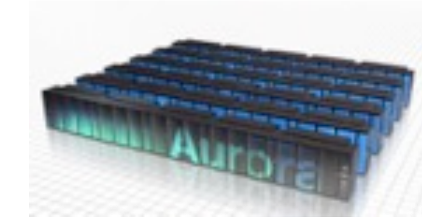
THETA: STEPPING STONE TO AURORA

System Features	Theta Details
Delivery Timeline	2016
Peak System Performance	>8.5 PetaFLOP/s
Compute Node CPU	2nd Generation Intel® Xeon Phi™ processors - **KNL**
Compute Node Count	>2,500 single socket nodes
Compute Platform	Cray XC supercomputing platform
Compute Node Peak Performance	>3 TeraFLOP/s per compute node
Cores Per Node	>60 cores
High Bandwidth On-Package Memory	Up to 16 Gigabytes per compute node
DDR4 Memory	192 Gigabytes per compute node
SSD	128 Gigabytes per compute node
File System	Intel Lustre File System
File System Capacity (Initial)	10 Petabytes
File System throughput (Initial)	200 Gigabytes/s
System Interconnect	Cray Aries Dragonfly topology interconnect
Intel Arch (x86-64) Compatibility	Yes

MIRA -> THETA IMPACT FOR SCIENTISTS

- Same MPI + OpenMP/ pThreads PM
- Dragonfly advantages
  - Much higher connectivity: applications with irregular point-to-point communication will do better
  - Much smaller diameter
- Increased vectorization opportunities
  - Compilers: Intel, Cray, and PGI
  - AVX512: more widely used than QPX, available in other Intel CPUs
  - Wider vector SIMD unit
  - Two independent vector units
- Memory changes
  - Fastest memory is ~12x faster, slower memory is ~2.7x faster
  - Capacity is 13x more per node
- Additional improvements
  - Better single thread performance
  - Larger caches (L1, instruction, L2)
  - Memory per core larger
- Progression of {cores,threads} per node
  - Mira {16,64} → Theta {>60,>240}

AURORA – COMING IN 2018



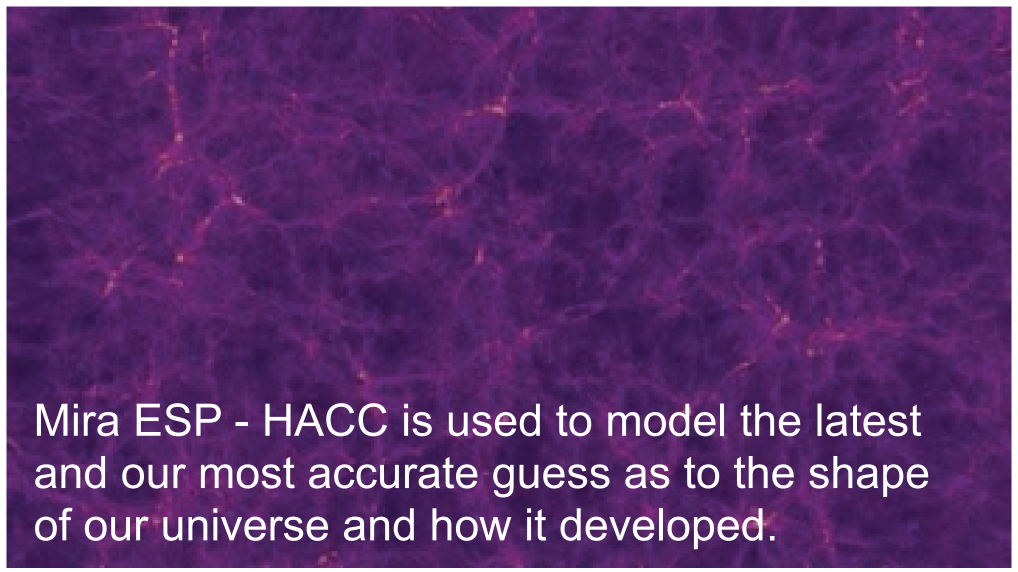
System Feature	Aurora Details
Peak System performance (FLOPs)	180 - 450 PetaFLOPS
Application FOM	13x Mira
Processor	3rd gen Intel Xeon Phi processor - **KNH**
Number of Nodes	>50,000
Compute Platform	Cray Shasta next generation SC platform
HBM, local mem, persistent mem	>7 PetaBytes
High Bandwidth On-Package Memory BW	>30 PB/s
System Interconnect	2nd gen Intel Omni-Path with silicon photonics
Interconnect Aggregate Node Link BW	>2.5 PB/s
Interconnect Bisection BW	>500 TB/s
Interconnect interface	Integrated
Burst Storage Buffer	Intel SSDs
File System	Intel Lustre File System
File System Capacity	>150 PetaBytes
File System Throughput	>1 TeraByte/s
Intel Architecture (x86-64) Compatibility	Yes

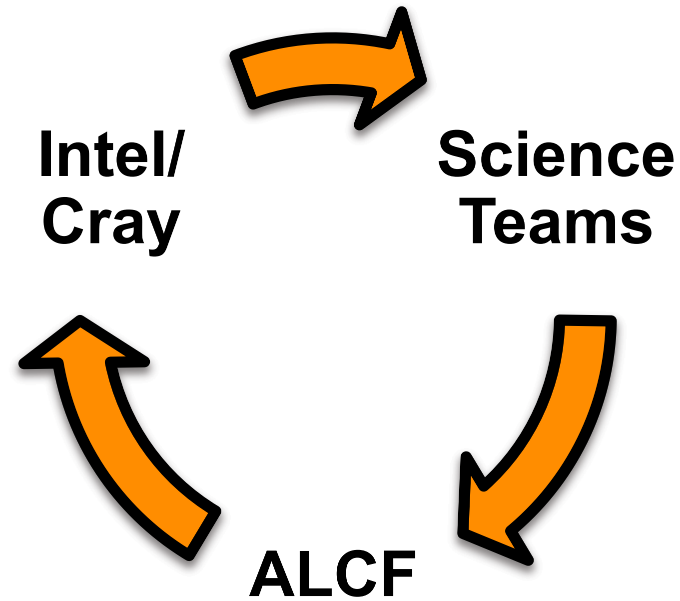
THETA -> AURORA IMPACT FOR SCIENTISTS

- Next gen of Xeon Phi architecture
- Similar number of nodes as Mira
  - Slightly more node concurrency than Theta
- Similar tiered memory and I/O
  - Faster & more HBM, more slower capacity mem
- Interconnect
  - New Intel Omni-Path, but same topology
- Similar Cray software stack
  - +Intel IP improvements
- Same MPI + OpenMP/pThreads PM

ALCF EARLY SCIENCE PROGRAM

GOAL: SCIENCE ON DAY ONE

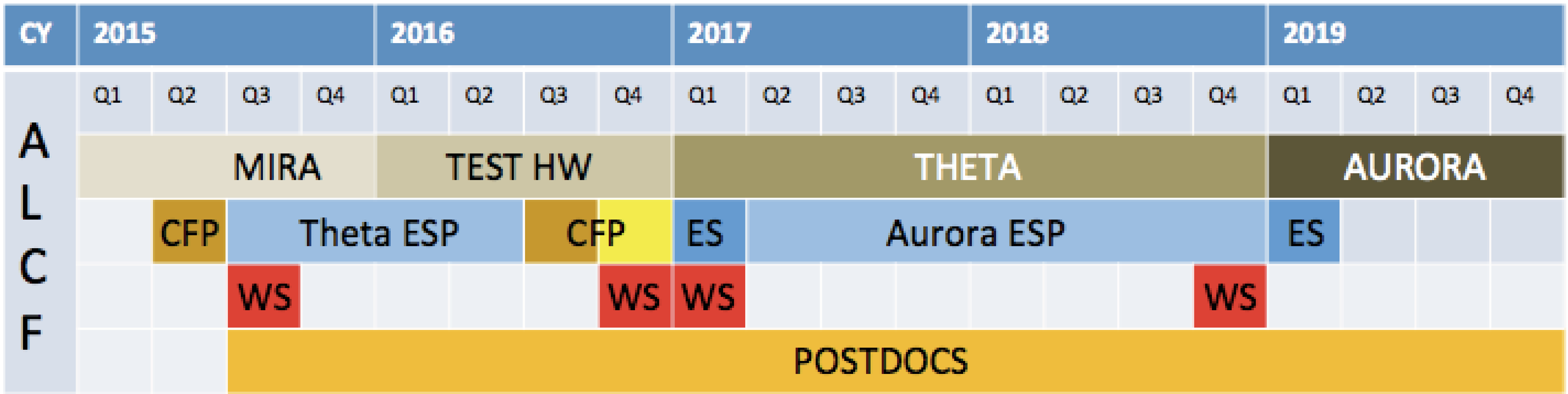




- Based on successful Mira ESP pioneering program
- Theta ESP (2015 – 2017)
- Aurora ESP (2016 – 2019)

Call for Aurora ESP proposals 3Q 2016

ALCF ESP CURRENT TIMELINE



Theta	Aurora
1. Apr 2015: ESP Call	1. Jul 2016: ESP call
2. May 2015: ESP Call closed	2. Sep 2016: ESP call closed
3. Jul 2015: ESP teams selected, work begins	3. Dec 2016: ESP projects selected, work begins
4. Aug 2015: ESP kick-off workshop	4. Jan 2017: ESP kick-off workshop
5. Oct 2016: ESP hands-on workshop	5. Nov 2018: ESP hands-on workshop
6. Jan 2017: Early Science dedicated-access period begins	6. Jan 2019: Early Science dedicated-access period begins
7. Mar 2017: Early Science dedicated-access period ends	7. Mar 2019: Early Science dedicated-access period ends

ARCHITECTURE AND PERFORMANCE PORTABILITY

Application portability among ALCF, OLCF, and NERSC architectures is critical concern of ASCR

- Application developers target wide range of architectures
- Maintaining multiple code versions is difficult
- Porting to different architectures is time-consuming
- Many Principal Investigators have allocations on multiple resources
- Applications far outlive any computer system

Improve data locality and thread parallelism

- Many-core or GPU optimizations improve performance on all architectures
- Exposed fine grain parallelism transitions more easily between architectures
- Data locality optimized code design also improves portability

Use portable libraries

- Library developers deal with portability challenges
- Many libraries are DOE supported

MPI+OpenMP 4.0 could emerge as common programming model

- Significant work is still necessary
- All ASCR centers are on the OpenMP standards committee

Encourage portable and flexible software development

- Use open and portable programming models
- Avoid architecture specific models such as Intel TBB, NVIDIA CUDA
- Use good coding practices: parameterized threading, flexible data structure allocation, task load balancing, etc.

Argonne Leadership Computing Facility

Argonne
NATIONAL LABORATORY