Applications Readiness Plan

Briefing to ASCAC, November 21, 2014

Straatsma: ALCF, NERSC, OLCF Application Readiness Collaboration
Antypas: Challenges and Strategies for Portability
Williams: Planning of Joint Activities

Argonne NATIONAL LABORATORY | Leadership Computing Facility
NERSC
OLCF OAK RIDGE LEADERSHIP COMPUTING FACILITY

Joint ASCR Facilities Application Readiness and Performance Portability

ALCF, NERSC and OLCF Application Readiness Collaboration

ALCF, NERSC, OLCF Application Readiness Collaboration

The three ASCR sites feel that the close collaboration on application readiness and performance portability activities that has been initiated at the September 2014 joint meeting will result in better programming activities, higher value for the ultimate users of the ported applications, and a natural way to avoid duplication of effort by leveraging each other's efforts.

The outlined collaboration on application readiness and portability does not simply address the next-generation systems that will be coming to ALCF, NERSC, and OLCF, but is also on the vendors' paths to exascale architectures. Using appropriate abstractions to get portability and performance on these pre-exascale systems provides a path to be continued toward exascale.

Two Tracks for Future Large Systems

Tianhe-2 (NUDT): TH-IVB-FEP
Intel Xeon E5-2692 12 C 2.2 GHz
TH Express-2
Intel Xeon Phi 31S1P

Titan (Cray): Cray XK7
AMD Opteron 6274 16C 2.2 GHz
Cray Gemini
NVIDIA K20x

Sequoia (IBM): BlueGene/Q
Power BQC 16C 1.6 GHz

K computer (Fujitsu)
SPARC64 VIIIfx 2.0 GHz
Tofu

Mira (IBM): BlueGene/Q
PowerPC A2 16C 1.6 GHz

Piz Daint (Cray): Cray XC30
Intel Xeon E5-2670 8C 2.6 GHz
Cray Aries
NVIDIA K20x

Edison (Cray): Cray XC30
Intel Xeon E%-2695v2 12C 2.4 GHz
Aries

Many Core

- 10's of thousands of nodes with millions of cores
- Homogeneous cores
- Multiple levels of memory – on package, DDR, and non-volatile
- Unlike prior generations, future products are likely to be self hosted

Hybrid Multi-Core

- CPU / GPU Hybrid systems
- Likely to have multiple CPUs and GPUs per node
- Small number of very fat nodes
- Expect data movement issues to be much easier than previous systems – coherent shared memory within a node
- Multiple levels of memory – on package, DDR, and non-volatile

Cori at NERSC

- Self-hosted many-core system
- Intel/Cray
- 9300 single-socket nodes
- Intel® Xeon Phi™ Knights Landing (KNL)
- 16GB HBM, 64-128 GB DDR4
- Cray Aries Interconnect
- 28 PB Lustre file system @ 430 GB/s
- Target delivery date: June, 2016

Summit at OLCF

- Hybrid CPU/GPU system
- IBM/NVIDIA
- 3400 multi-socket nodes
- POWER9/Volta
- More than 512 GB coherent memory per node
- Mellanox EDR Interconnect
- Target delivery date: 2017

ALCF-3 at ALCF

- TBA
- Target delivery date: 2017-18

Synergy between Application Readiness Efforts

- Application Developer Team involvement
  - Knowledge of the application
  - Work on application in development “moving target”
  - Optimizations included in application release
- Early Science Project
  - Demonstration of application on real problems at scale
  - Shake-down on the new system hardware and software
  - Large-scale science project is strong incentive to participate
- Vendor support is crucial
  - Programming environment often not mature
  - Best source of information on new hardware features
- Access to multiple resources, including early hardware
- Joint training activities

- Portability is a critical concern
- Experience benefits other developers and users
  - Coverage of scientific domains
  - Coverage of algorithmic methods and programming models
- Persistent culture of application readiness
  - More computational ready applications available
  - Experience of science liaisons and catalysts for user programs
  - Synergy with libraries and tools projects

Getting Ready: Application Readiness Programs

NESAP at NERSC

NERSC Exascale Science Application Program

- Call for Proposals – June 2014
- 20 Projects selected
- Partner with NERSC Application Readiness Team
- 8 Postdoctoral Fellows

Criteria

- An application's computing usage within the DOE Office of Science
- Representation among all 6 Offices of Science
- Ability for application to produce scientific advancements
- Ability for code development and optimizations to be transferred to the broader community through libraries, algorithms, kernels or community codes
- Resources available from the application team to match NERSC/Vendor resources

CAAR at OLCF

Center for Accelerated Application Readiness

- Call for Proposals – November 2014
- 8 Projects to be selected
- Partner with OLCF Scientific Computing group and IBM/NVIDIA Center of Excellence
- 8 Postdoctoral Associates

Criteria

- Anticipated impact on the science and engineering fields
- Importance to the user programs of the OLCF
- Feasibility to achieve scalable performance on Summit
- Anticipated opportunity to achieve performance portability for other architectures
- Algorithmic and scientific diversity of the suite of CAAR applications.
- Optimizations incorporated in master repository
- Size of the application's user base

ESP at ALCF

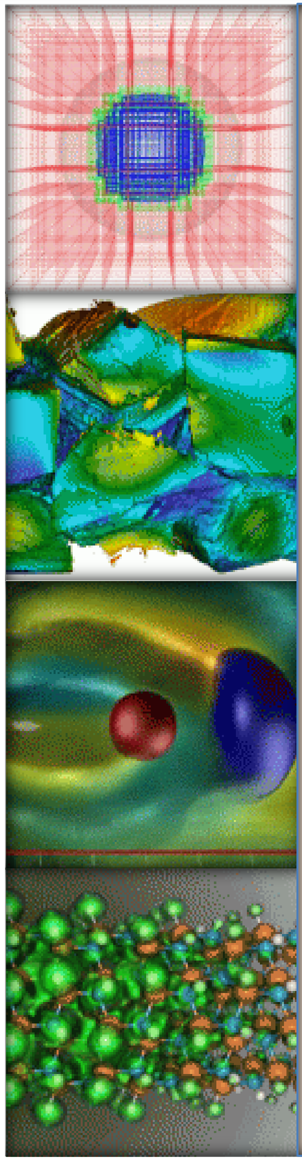
Early Science Program

- Call for Proposals
- 10 Projects to be selected
- Partner with ALCF Catalyst group and ALCF Vendor Center of Excellence
- Postdoctoral Appointee per project

Criteria

- Science Impact
- Computational Readiness
  - Proposed science problem of appropriate scale to exercise capability of new machine
  - Confidence code will be ready in time
  - Project code team appropriate
    - Willing partner with ALCF & vendor
- Diversity of science and numerical methods
  - Samples spectrum of ALCF production apps

NERSC Exascale Science Applications Program (NESAP)



Advanced Scientific Computing Research

Almgren (LBNL) – **BoxLib AMR Framework**

Trebotich (LBNL) – **Chombo-crunch**

High Energy Physics

Vay (LBNL) – **WARP & IMPACT**

Toussaint(Arizona) – **MILC**

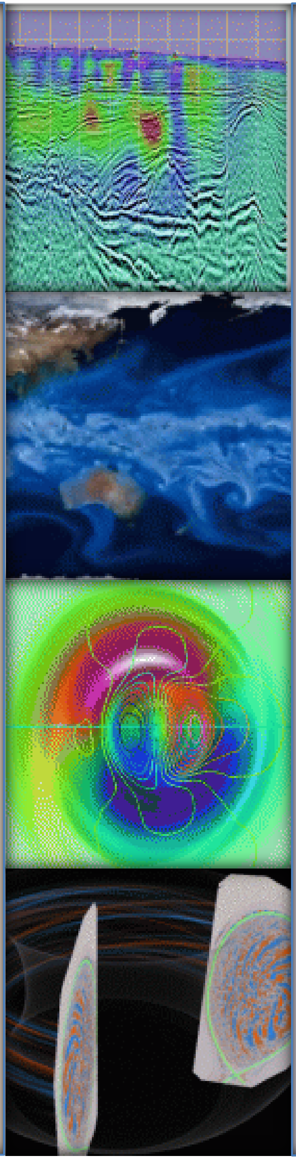
Habib (ANL) – **HACC**

Nuclear Physics

Maris (Iowa St.) – **MFDn**

Joo (JLAB) – **Chroma**

Christ/Karsch (Columbia/BNL) – **DWF/HISQ**



Basic Energy Sciences

Kent (ORNL) – **Quantum Espresso**

Deslippe (NERSC) – **BerkeleyGW**

Chelikowsky (UT) – **PARSEC**

Bylaska (PNNL) – **NWChem**

Newman (LBNL) – **EMGeo**

Biological and Environmental Research

Smith (ORNL) – **Gromacs**

Yelick (LBNL) – **Meraculous**

Ringler (LANL) – **MPAS-O**

Johansen (LBNL) – **ACME**

Dennis (NCAR) – **CESM**

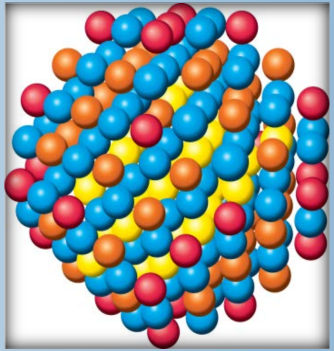
Fusion Energy Sciences

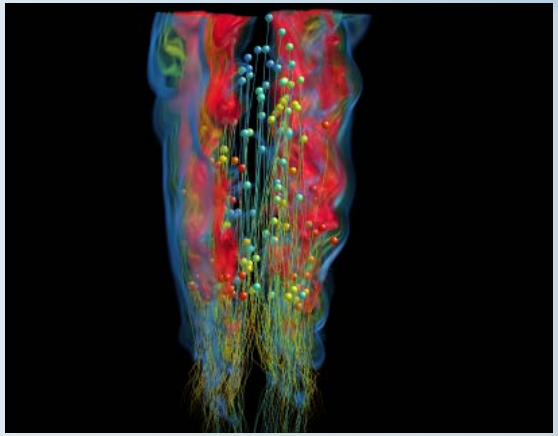
Jardin (PPPL) – **M3D**

Chang (PPPL) – **XGC1**

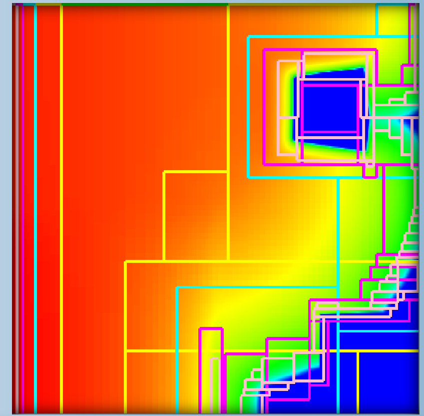
OLCF-3 Center for Accelerated Application Readiness (CAAR)

WL-LSMS
Illuminating the role of material disorder, statistics, and fluctuations in nanoscale materials and systems.

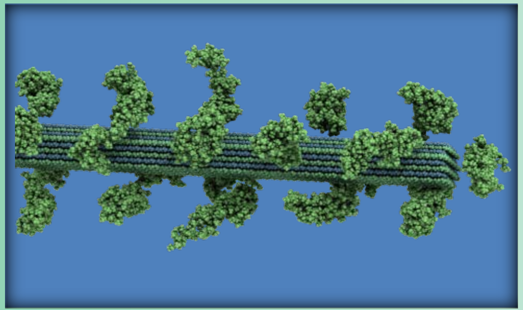




S3D
Understanding turbulent combustion through direct numerical simulation with complex chemistry.

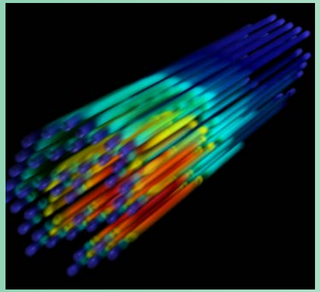


NRDF
Radiation transport – important in astrophysics, laser fusion, combustion, atmospheric dynamics, and medical imaging – computed on AMR grids.



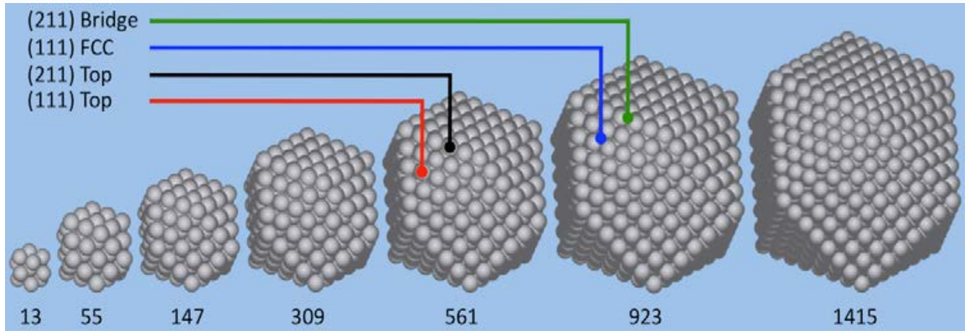
LAMMPS
A molecular description of membrane fusion, one of the most common ways for molecules to enter or exit living cells.

CAM-SE
Answering questions about specific climate change adaptation and mitigation scenarios; realistically represent features like precipitation patterns / statistics and tropical storms.



Denovo
Discrete ordinates radiation transport calculations that can be used in a variety of nuclear energy and technology applications.

ALCF-2 ESP (Mira)

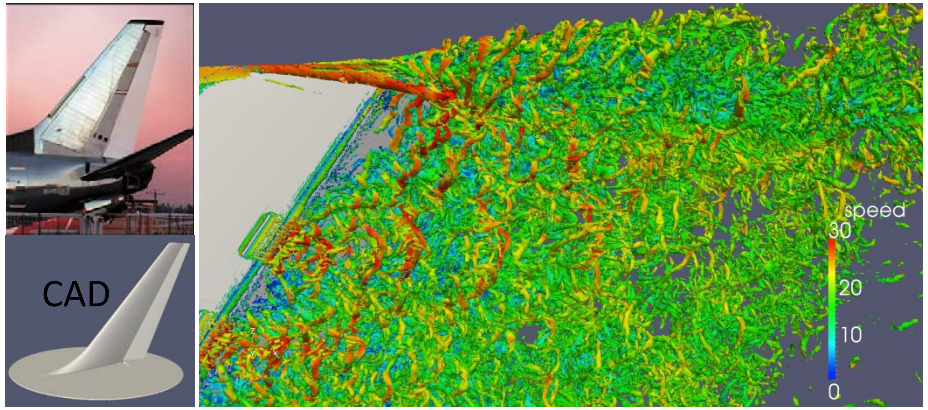


Energy Storage Materials and Catalysis

Quantum Monte Carlo simulations of platinum metal nanoparticles as catalysts for key reactions (QMCPACK).

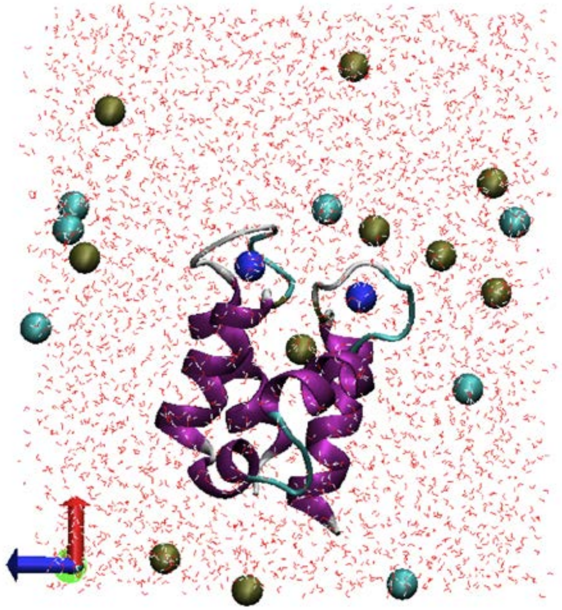
Active Aerodynamic Flow Control

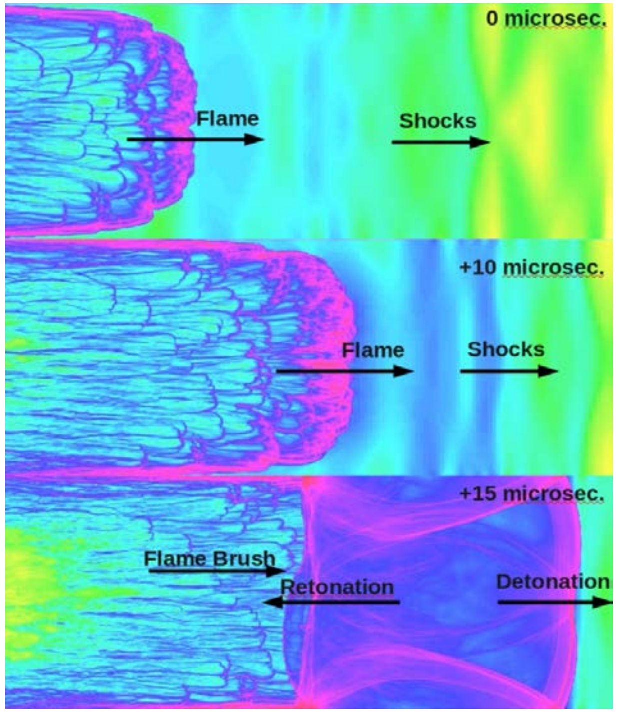
Tiny synthetic jet actuators dramatically improve effectiveness of aerodynamic control surfaces such as rudders. (PHASTA).



Biomolecular Science

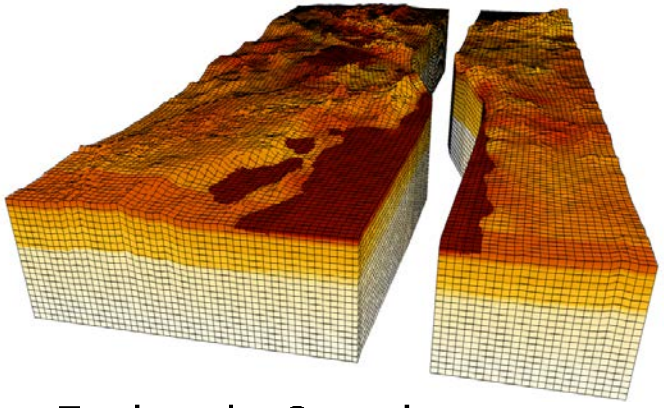
Highly accurate microscopic model of proteins and complexes—polarizable force field (NAMD).





High Speed Combustion and Detonation

Direct numerical simulation of shock tube experiments (ALLA/FTT).

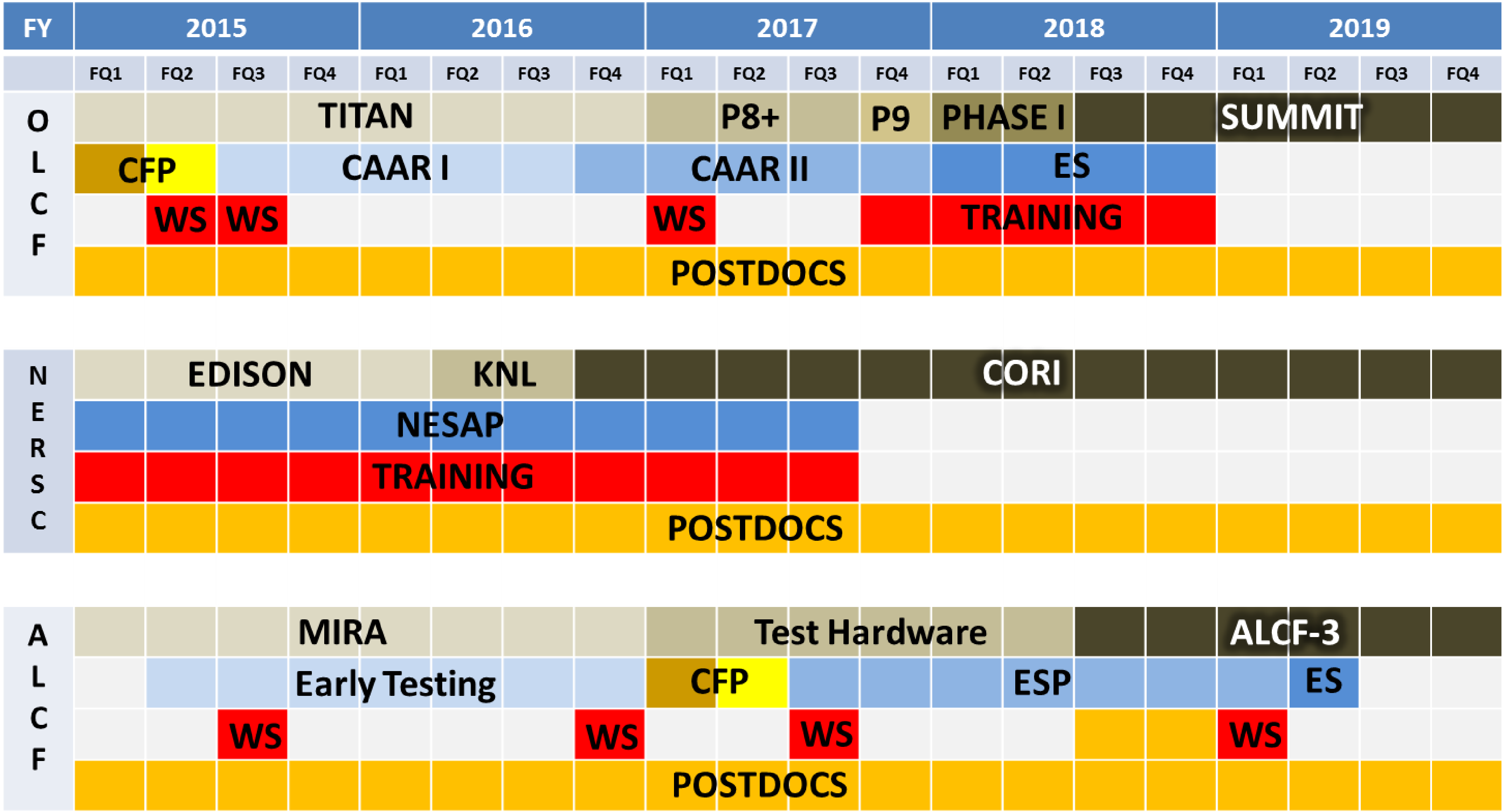


Earthquake Genesis

Realistic 3D fault rupture simulation (SORD).

Argonne National Laboratory | Leadership Computing Facility | NERSC | OLCF Oak Ridge Leadership Computing Facility

Application Readiness Programs Tentative Timeline



Argonne NATIONAL LABORATORY Leadership Computing Facility NERSC OLCF Oak Ridge Leadership Computing Facility

Drivers for Portability

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

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

Primary task is exposing parallelism and data locality

Challenge is to find the right abstraction:

- MPI + X (X=OpenMP, OpenACC)
- PGAS + X
- DSL
- …

ALCF, NERSC, OLCF Collaboration

- March 2014: Meeting to discuss common applications, POCs
- September 2014: ALCF, OLCF participated in NESAP proposal reviews
- September 2014: Meeting to discuss application readiness and architectural and performance portability:
  1. How will we coordinate our application readiness efforts, particularly when more than one center chooses the same application for early access and readiness?
  2. What guidance and tools can we provide users to encourage application development that will be portable across different architectures?
  3. What mechanisms and allocations can we provide to all of our early science teams so that they can test and run their applications on different architectures?

Postdoctoral Associate Programs

The ASCR facilities are hosting Distinguished Postdoctoral Associates programs with the overarching objective of training the next generation of computational scientists. To achieve this goal, the postdoctoral associates programs have the specific goals of providing:

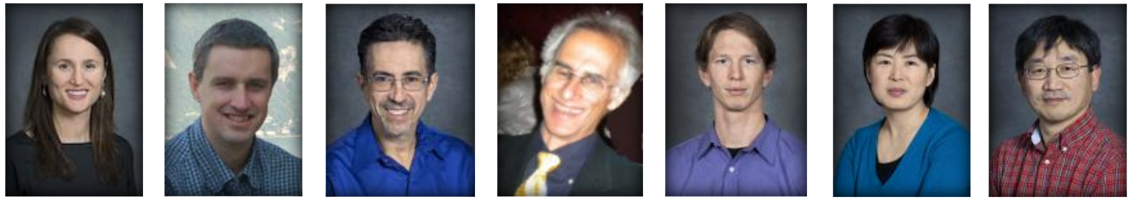
1. Challenging scientific campaigns in critical science mission areas
2. Experience in using ASCR computing capabilities
3. Training in software development and engineering practices for current and future massively parallel computer architectures

Central to achieving these goals is access to leadership computing resources, availability of computational domain scientists to provide adequate mentoring and guidance, and facilities' association with universities with strong computational and computer science programs.

NERSC: https://www.nersc.gov/users/announcements/featured-announcements/nersc-exascale-science-postdoctoral-fellowships/
OLCF: https://www.olcf.ornl.gov/summit/olcf-distinguished-postdoctoral-associates-program/
ALCF: https://www.alcf.anl.gov/about/careers

Application Readiness Staff

NERSC Application Readiness Team



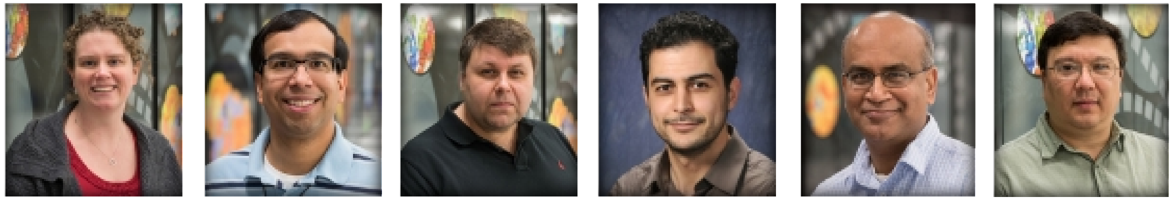
Katie Antypas | Nick Wright | Richard Gerber | Harvey Wasserman | Brian Austin | Zhengji Zhao | Woo-Sun Yang

Rebecca Hartman-Baker | Jon Rood (IPCC postdoc) | Scott French | Jack Deslippe | Helen He | Matt Cordery

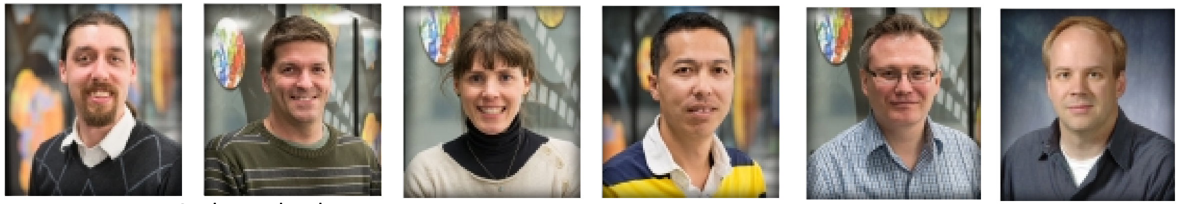
OLCF Application Readiness Team

Left to right: Ramanan Sankaran, Mike Matheson, George Ostrouchov, Duane Rosenberg, Valentine Anantharaj, Bronson Messer, Mark Berrill, Matt Norman, Ed D'Azevedo, Norbert Podhorski, Wayne Joubert, JJ Chai (postdoc, now in CSM), Judy Hill, Mark Fahey, Hai Ah Nam, Jamison Daniel, Dmitry Liakh, Supada Loasooksathit (postdoc), Markus Eisenbach, Arnold Tharrington, Ying Wai Li, Mingyang Chen (postdoc), Peyton Ticknor, Tjerk Straatsma, Dave Pugmire and Jan-Michael Carrillo (postdoc, now at SNS).

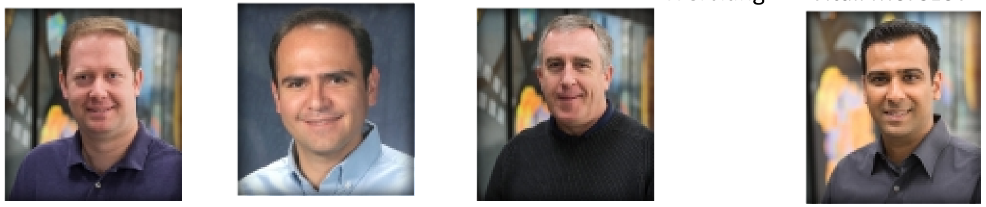
ALCF Application Readiness Team



Katherine Riley | Nichols Romero | Yuri Alexeev | Anouar Benali | Kalyan Kumaran | Ray Loy



Hal Finkel | Graham Fletcher | Marta Garcia Martinez | Wei Jiang | Vitali Morozov | Scott Parker



James Osborn | Alvaro Vazquez-Mayagoitia | Timothy Williams | Venkatram Vishwanath

Not Pictured: Ramesh Balakrishnan, Christopher J. Knight, Adrian Pope, William Scullin, and Emily R. Shemon

Portability Challenges and Strategies

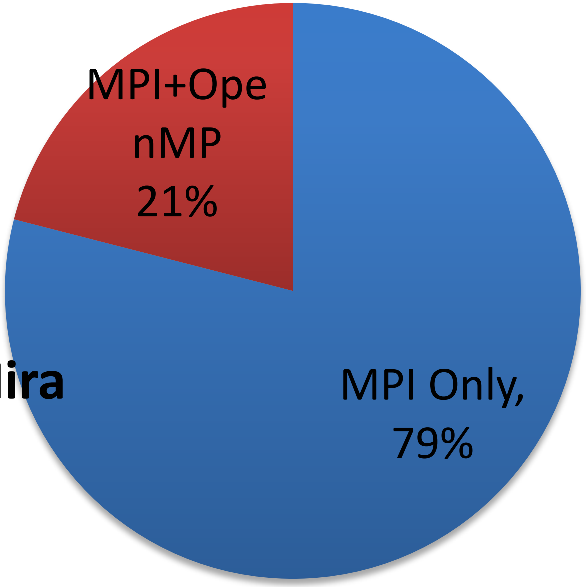
How should we define performance portability?

- Community is just coming to terms with the language and definition.
- From a recent FASTMath meeting two definitions of performance portable emerged:
  - *Same piece of code (from the user perspective) runs on different architectures with 'good' performance*
  - *A relatively small amount of effort is needed to make a change to get good performance within advertised (algorithmic or performance) tolerances across both current and future architectures*

The dominant programming model at centers is MPI, sometimes +X

- MPI only
- MPI+OpenMP
- MPI+OpenACC/CUDA

Programming Model on Hopper



Programming Model GPU usage on Titan

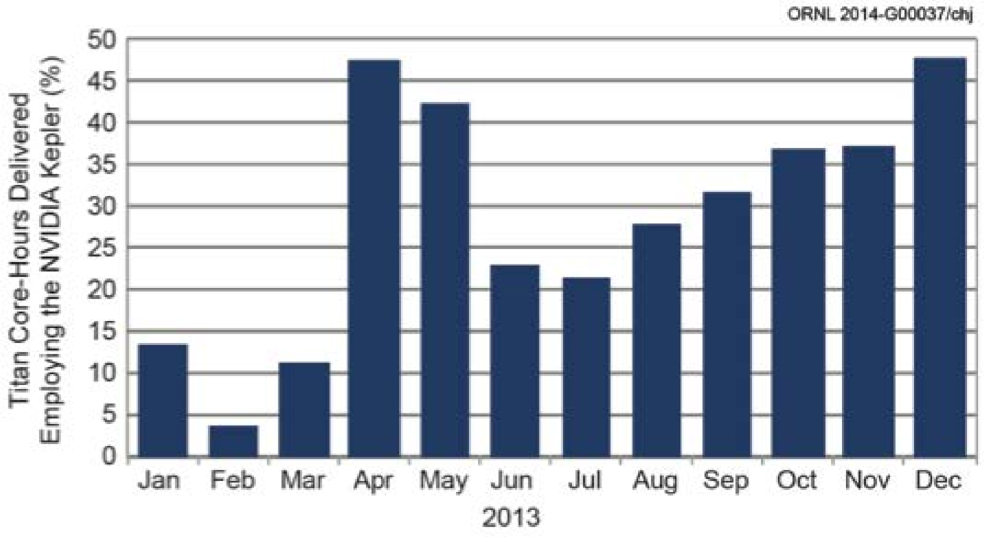
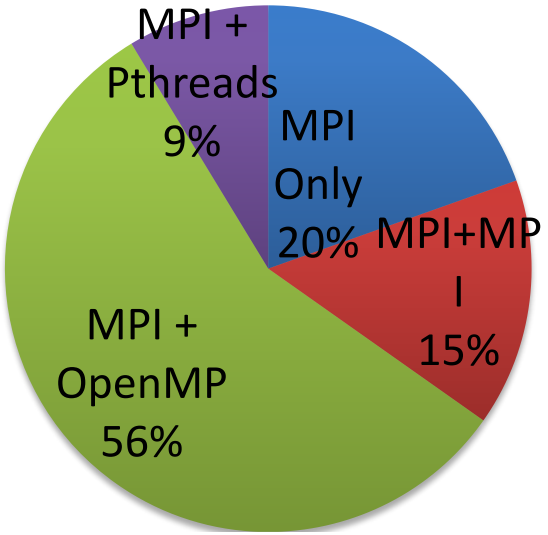


Figure 2.7. Tracking GPU Usage on Titan.

Programming Model on Mira

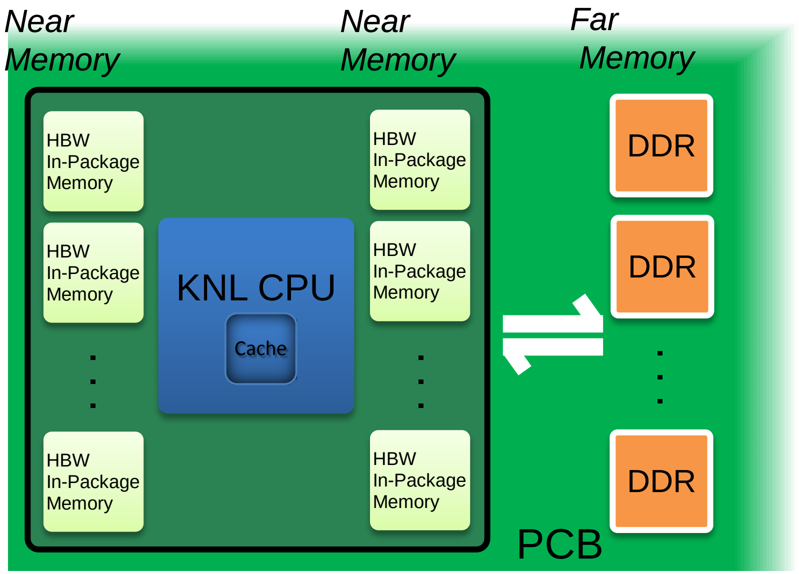


Argonne National Laboratory | Leadership Computing Facility | NERSC | OLCF Oak Ridge Leadership Computing Facility

Portability challenge: Deepening memory hierarchies

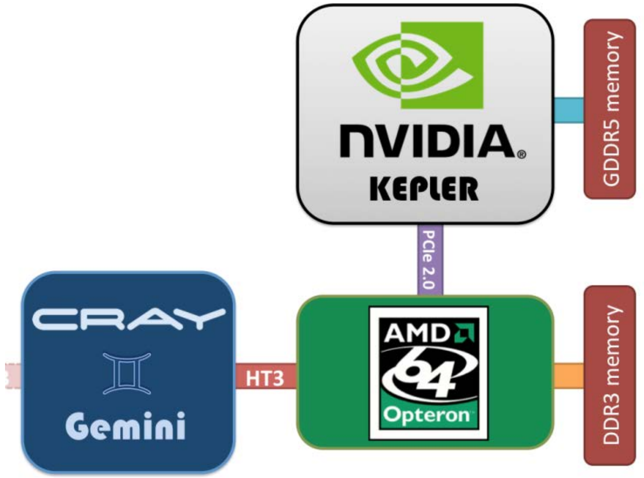
- ASCR facility architectures will have memory hierarchies, each with “faster” and “slower” memory that differ by:
  – Ratios of fast and slow memory bandwidths
  – Sizes of fast and slow memories
  – Attributes of fast and slow memories
  – APIs into fast and slow memories
  – Number of NUMA domains

Knight’s Landing Architecture



Source:Intel

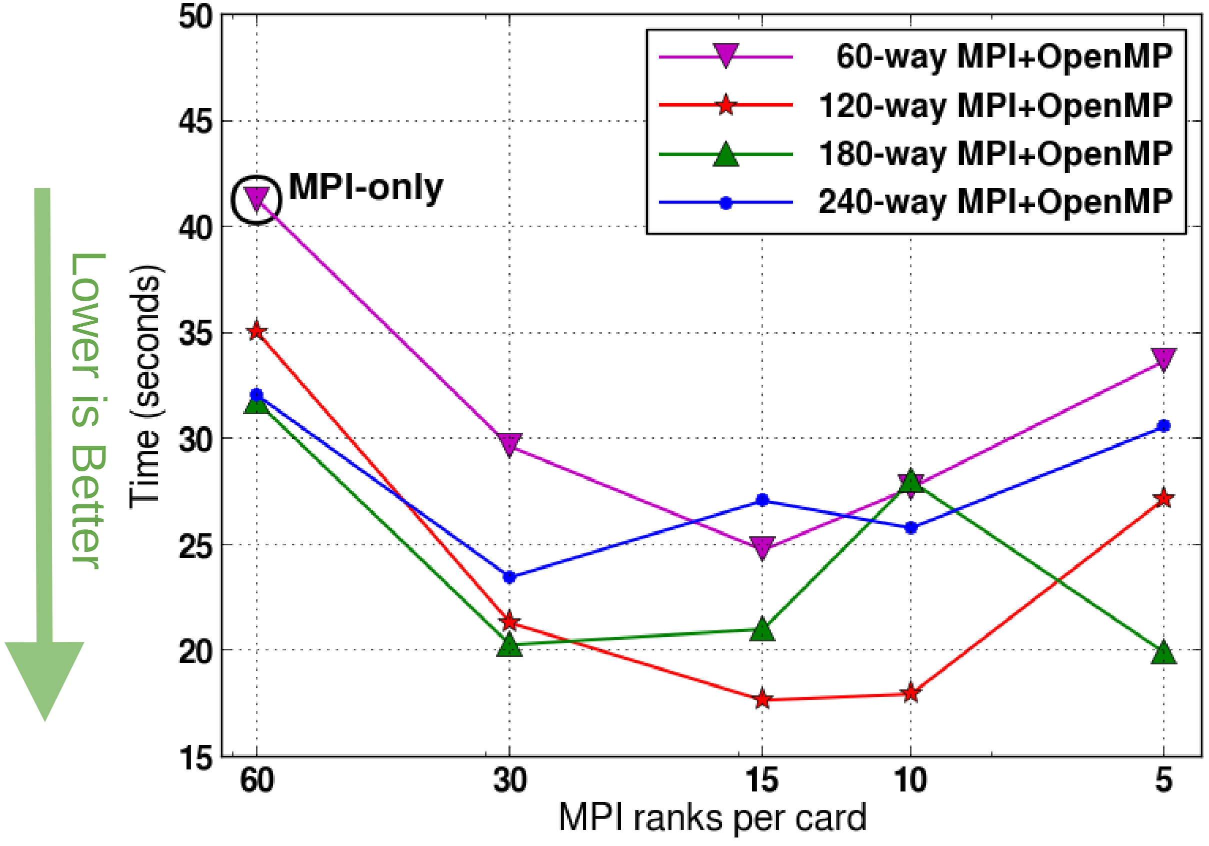
Titan Node Architecture



Portability challenges: Thread Scaling and Management

- Increasing number of threads
- Challenges scaling threaded performance
- Different numbers of threads on each architecture
- Different scheduling policies
- Different levels of coarse and fine parallelism

FLASH Code thread scaling study
Single Knight's Corner Card

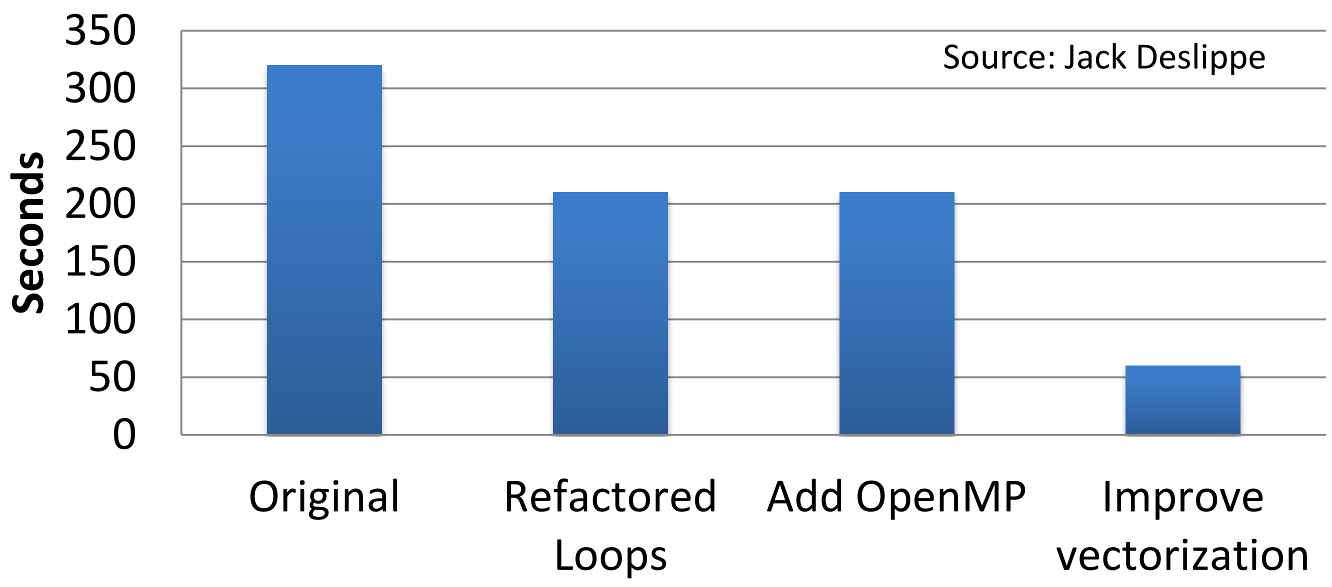


Source: Chris Daley

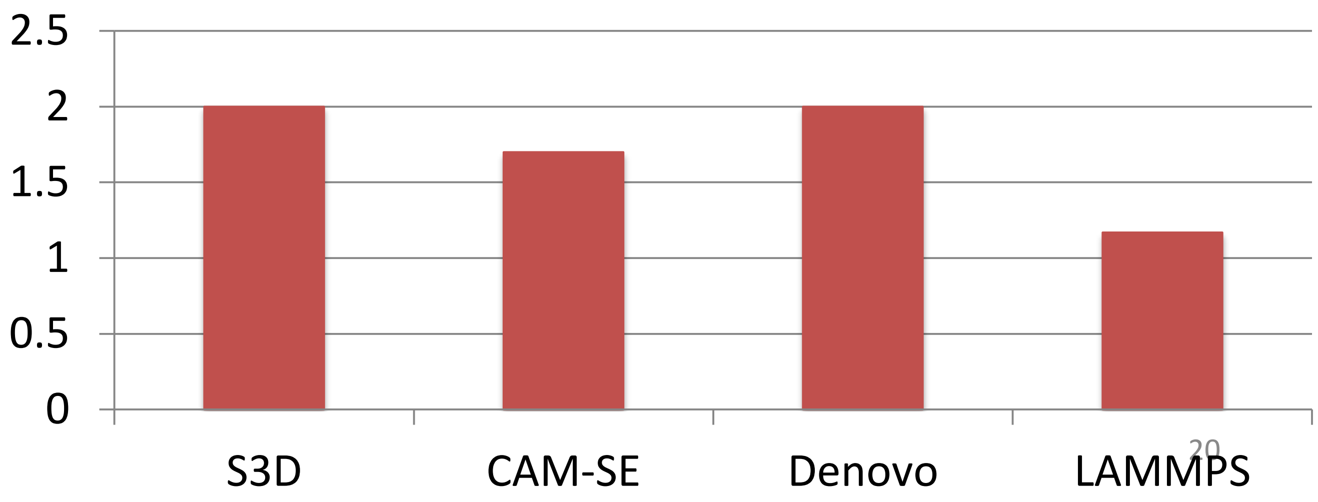
Strategy: Improve data locality and thread parallelism

- Preparing codes for GPUs or manycore architectures will improve performance on all architectures
- Codes exposing finer grained parallelism will more easily transition between architectures
- Applications designed with data locality in mind will also see improved portability

BerkeleyGW Sigma.cplx.x Kernel Improvements – CPU only

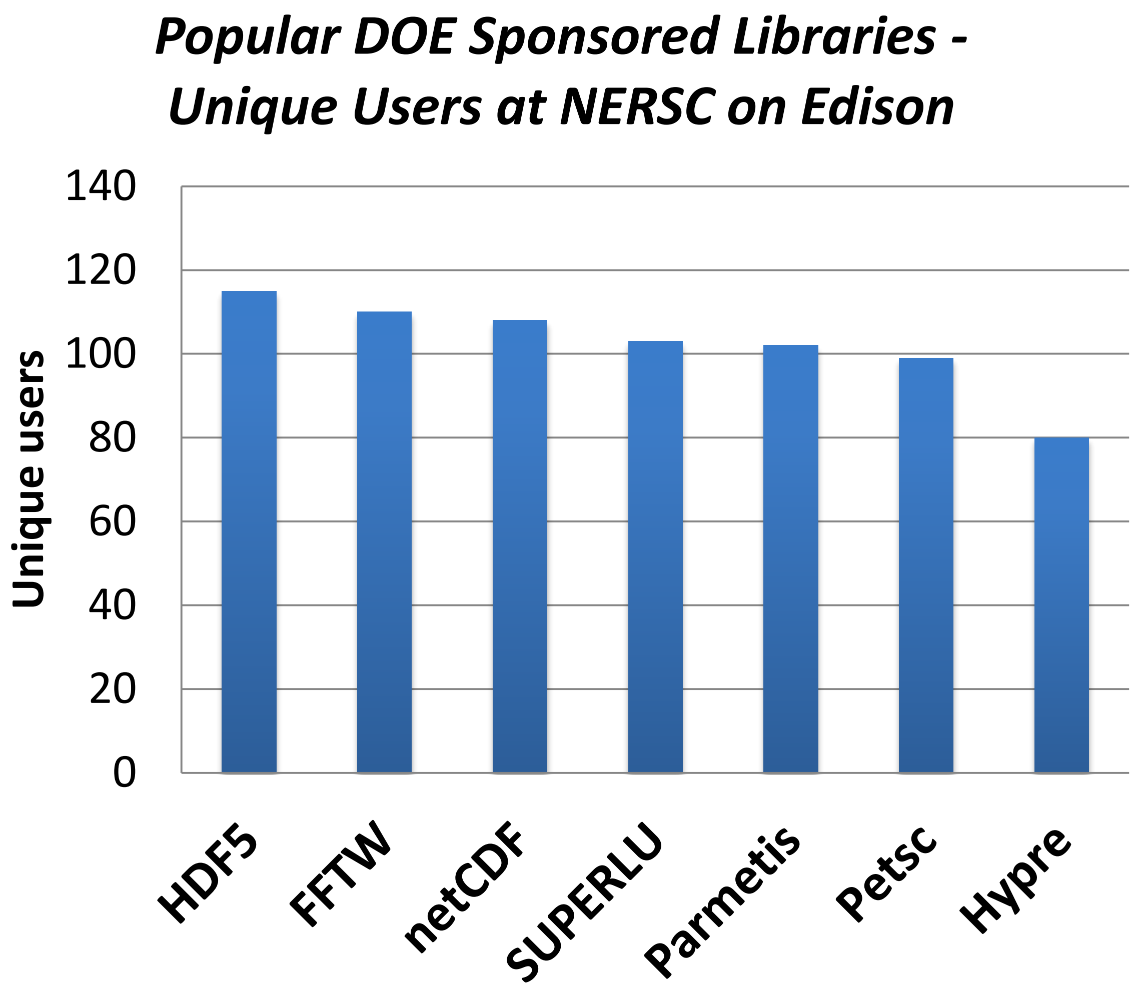


Speedup on CPU due to Refactoring for GPU on Titan



Strategy: Use portable libraries

- Many applications spend significant amounts of time in libraries
- Assumption is that library developers will handle portability challenges
- Differences in ASCR architectures could push users to libraries rather than rolling their own -- ***If libraries are supported and high performing***



Strategy: MPI + OpenMP 4.0 could emerge as a common programming model

- OpenMP 4.0 has new capabilities that make it more attractive for use on accelerators
  - "Target" construct allows offloading of data and computation to accelerators
  - SIMD construct supports more portable vectorization
  - Improved task affinity
- Significant work still necessary
  - For high performance on accelerator architectures
  - For explicitly managing on-package memory
- All 3 centers joining the OpenMP standards committee

Strategy: Abstract portability challenges from users

- A number of frameworks that abstract architecture details away from the user can provide portability
  - KOKKOS
  - TIDA
  - Gridtools
  - Dash
  - hStreams
  - Domain specific languages (DSLs)
- It is important that we work closely with the research community to address performance portability challenges

Strategy: Encourage portable and flexible programming

- Use open and portable programming models where possible
- Try to avoid architecture specific models like:
  - Intel Thread Building Blocks
  - NVIDIA CUDA
  - Where necessary, encapsulate vendor specific code into library or swappable code module
- Good coding practices
  - Use parameters for the amount of threading and placements of threads
  - Allow data structures can be flexibly allocated to different memory spaces
  - Allow task level flexibility so work can be allocated on different computing elements (GPU & CPU)

Planning of Joint Activities

SC Centers Roles

ADVOCATE
portability

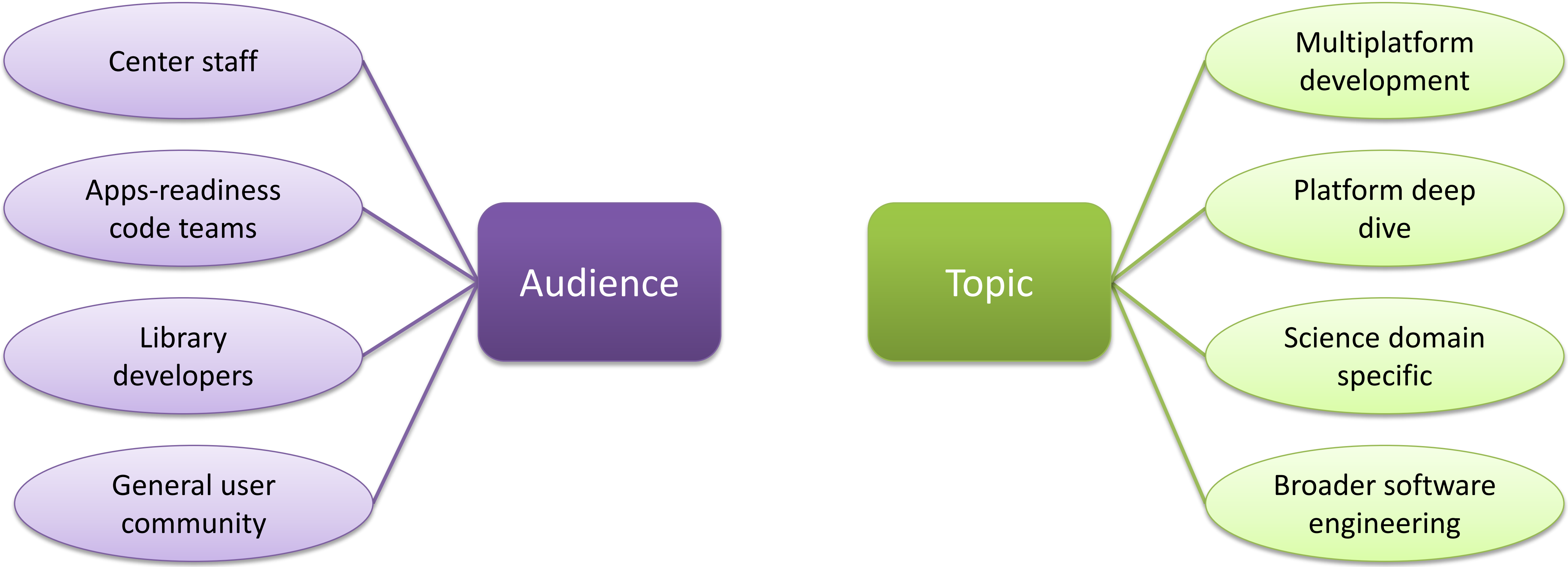
FOCUS
apps efforts

SHARE
best practices

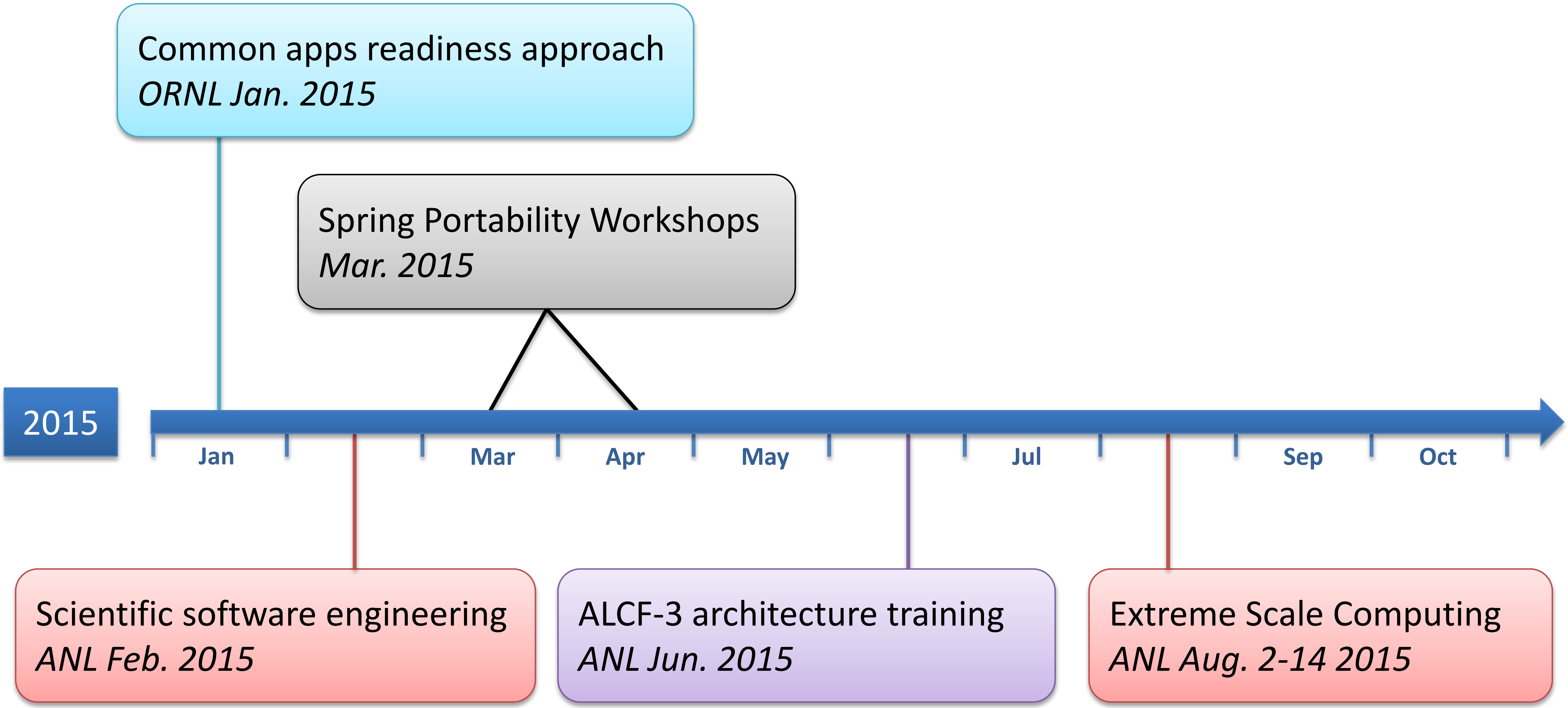
TRAIN
developers

GRANT
hardware access

TRAIN Developers



Upcoming Training



Argonne NATIONAL LABORATORY | Leadership Computing Facility
NERSC
OLCF Oak Ridge Leadership Computing Facility

More Training

- Intel Dungeon sessions (NERSC)
- Vendor workshops (ORNL)
- NESAP training (NERSC)
- ESP hands-on workshops (ANL)
- Hackathons (ORNL)
- Ongoing scientific software engineering workshops (ANL)
- OpenMP 4.0 training

Where appropriate, shared training events

In all training, advocate and instruct on portability

FOCUS Apps Efforts: Application Readiness Program Coordination

- Joint participation in project review, selection
  - Representatives from other centers including CORAL and APEX (NERSC & Trinity) partners
    - NERSC did this in its NESAP proposal review
  - ESP and CAAR proposal forms ask about proposals to other centers' programs
- Annual meetings of ESP, NESAP, CAAR project teams
  - Share best practices
  - Template for presentations
- Tools and Libraries readiness
  - Companions to and components of apps readiness

ESP

CAAR

NESAP

Tools and Libraries

- Companions to/components of ESP, NESAP, CAAR
  - Example: Argonne Early Performance project *(K. Kumaran)*

Kokkos	Trilinos	HPC Toolkit	HDF
Chombo	PETSc	Scalasca	NetCDF
Charm++	LAPACK	TAU	ADIOS
Global Arrays	SCALAPACK	PAPI	
	FFTW	Open|Speedshop	
	P3DFFT		

GRANT Hardware Access

- For center staff: Director's Discretionary or equivalent allocation at other centers
  - For learning, testing
- ESP, NESAP, CAAR projects
  - Next-generation hardware simulators
  - Access to next-generation hardware as early as possible (centers/vendors)
  - Large allocation of pre-production time for early science runs (ALCF-3, Cori, Summit)
  - Time on current systems for interim development

ALCC proposal

Management and Planning of SC Centers Efforts

- Semiannual meetings of cross-labs applications readiness staff
- Tools and libraries working group (W. Joubert)
- Cross-lab training committee (F. Foerttner)
  - Shared calendar
  - Shared training events
- Manage nondisclosure, export control challenges
  - CORAL partners
  - APEX partners (NERSC & Trinity)

March 2014 Meeting
- Apps readiness coordination
- ~15 representatives

September 2014 Meeting
- Apps Portability
- Apps readiness coordination
- ~25 representatives

ALCF, NERSC, OLCF Application Readiness Collaboration

The three ASCR sites feel that the close collaboration on application readiness and performance portability activities that has been initiated at the September 2014 joint meeting will result in better programming activities, higher value for the ultimate users of the ported applications, and a natural way to avoid duplication of effort by leveraging each other's efforts.

The outlined collaboration on application readiness and portability does not simply address the next-generation systems that will be coming to ALCF, NERSC, and OLCF, but is also on the vendors' paths to exascale architectures. Using appropriate abstractions to get portability and performance on these pre-exascale systems provides a path to be continued toward exascale.