Exascale Computing Project Update

Douglas B. Kothe
Oak Ridge National Laboratory (ORNL)
Director, ECP

HPC User Forum
September 7, 2021
Video Conference

National Nuclear Security Administration

U.S. Department of Energy | Office of Science

ECP's Technical Focus Areas

Providing the necessary components to meet national goals

Performant mission and science applications at scale

Aggressive RD&D project	Mission apps; integrated S/W stack	Deployment to DOE HPC Facilities	Hardware technology advances

Application Development (AD)

Develop and enhance the predictive capability of applications critical to DOE

24 applications
National security, energy, Earth systems, economic security, materials, data

6 Co-Design Centers
Machine learning, graph analytics, mesh refinement, PDE discretization, particles, online data analytics

Software Technology (ST)

Deliver expanded and vertically integrated software stack to achieve full potential of exascale computing

71 unique software products
spanning programming models and run times,
math libraries,
data and visualization

Hardware and Integration (HI)

Integrated delivery of ECP products on targeted systems at leading DOE HPC facilities

6 US HPC vendors
focused on exascale node and system design; application integration and software deployment to Facilities

EXASCALE COMPUTING PROJECT

DOE HPC Roadmap to Exascale Systems

FY 2012 | FY 2016 | FY 2018 | FY 2021 | FY 2022 | FY 2023

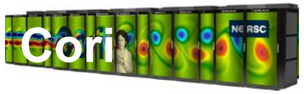
decommissioned

ORNL
Cray/AMD/NVIDIA

ANL
IBM BG/Q

LLNL
IBM BG/Q

ANL
Cray/Intel KNL



LBNL
Cray/Intel Xeon/KNL

LANL/SNL
Cray/Intel Xeon/KNL

ORNL
IBM/NVIDIA

LLNL
IBM/NVIDIA

FRONTIER

ORNL
HPE/AMD

ANL
Intel/HPE

LBNL
HPE/AMD/NVIDIA

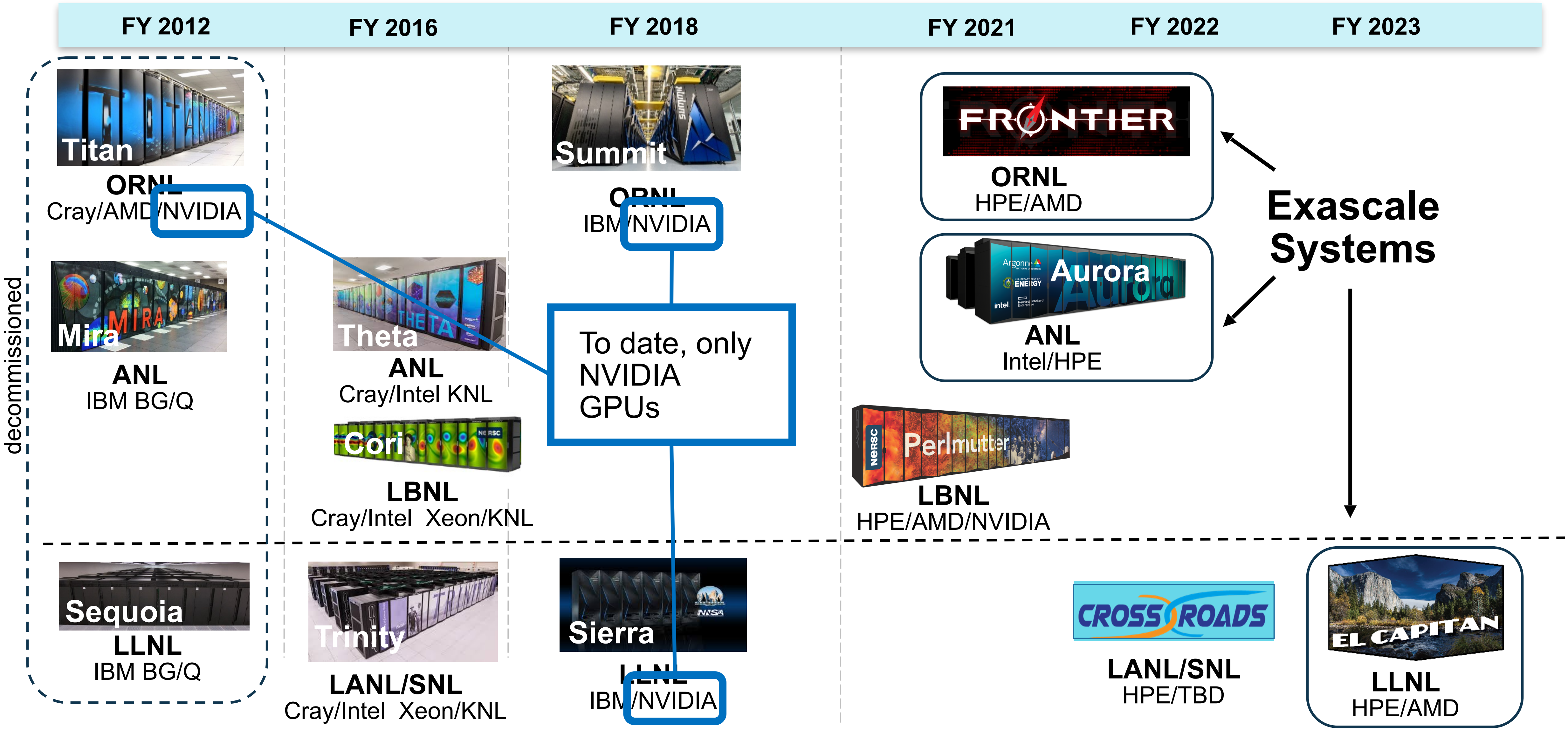
CROSSROADS

LANL/SNL
HPE/TBD

LLNL
HPE/AMD

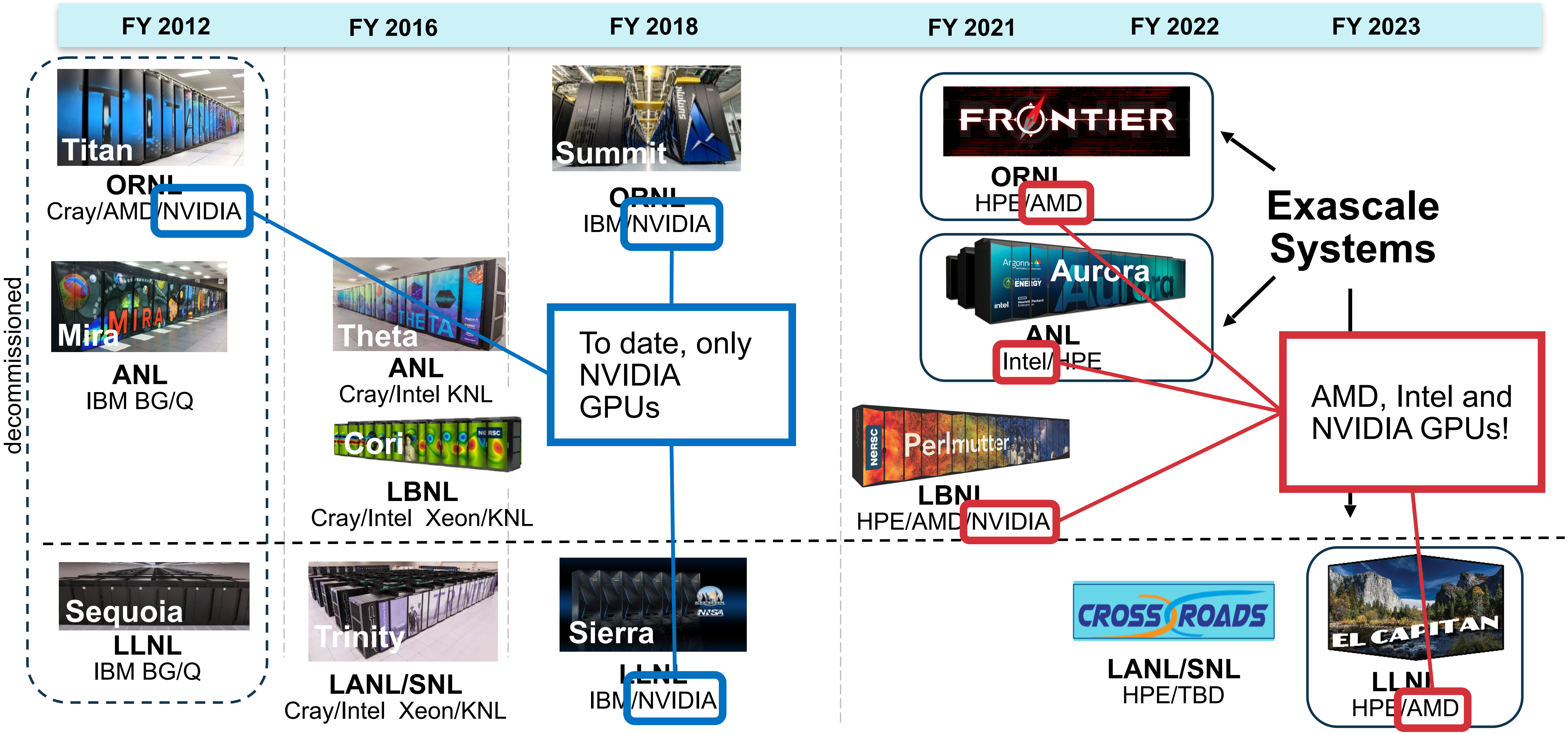
Exascale Systems

DOE HPC Roadmap to Exascale Systems



EXASCALE COMPUTING PROJECT

DOE HPC Roadmap to Exascale Systems



EXASCALE COMPUTING PROJECT

Specs of the Aurora & Frontier Exascale Systems

Details of these impressive systems are becoming firm and real

Frontier: https://www.olcf.ornl.gov/frontier/
Aurora: https://www.alcf.anl.gov/aurora

Specs of the Aurora & Frontier Exascale Systems

Details of these impressive systems are becoming firm and real

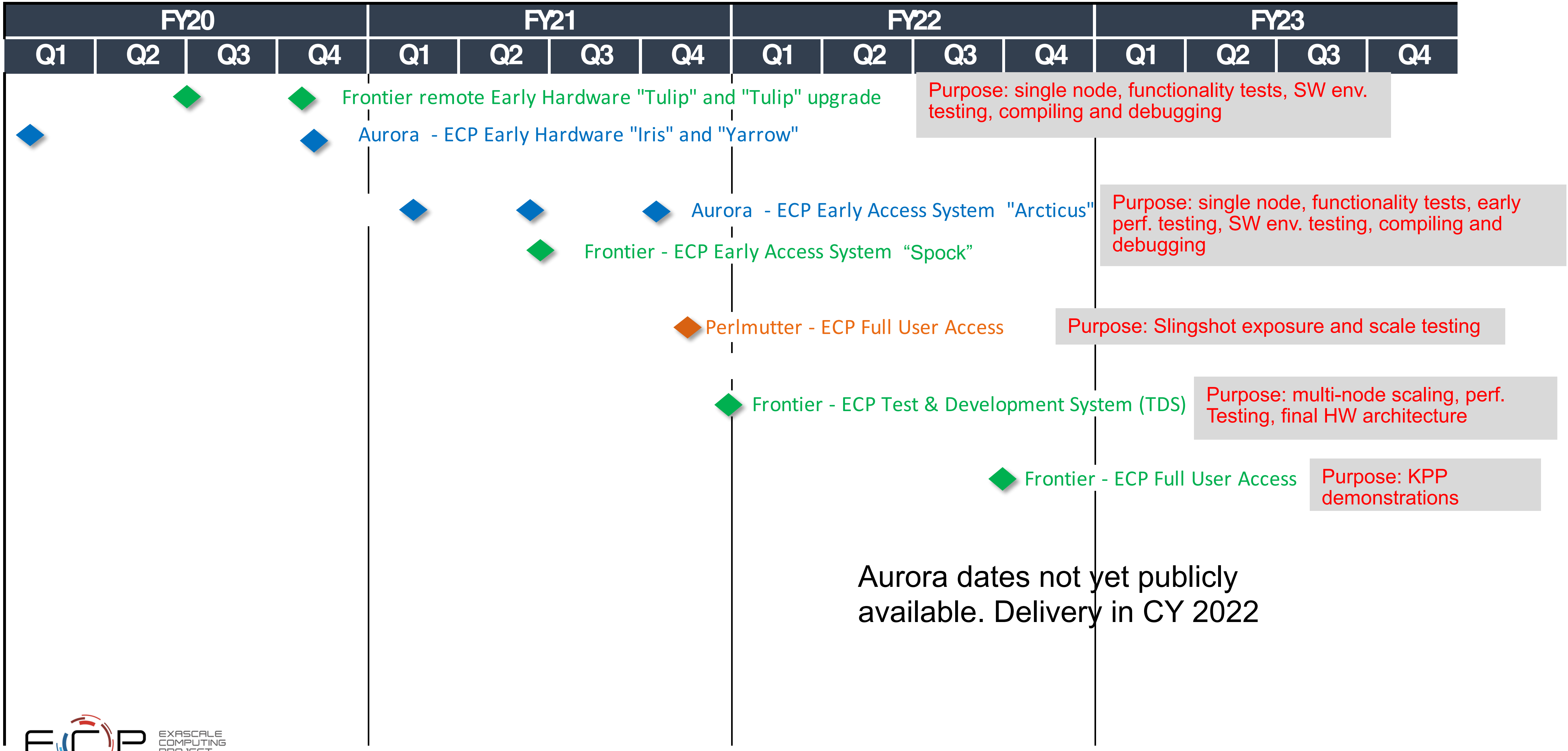
System attributes	ALCF Now	NERSC Now	OLCF Now	NERSC Pre-Exascale	ALCF Pre-Exascale	OLCF Exascale	ALCF Exascale
Name (Planned) Installation	Theta 2016	Cori 2016	Summit 2017-2018	Perlmutter (2020-2021)	Polaris (2021)	Frontier (2021-2022)	Aurora (2022-2023)
System peak	> 15.6 PF	> 30 PF	200 PF	> 120PF	35 – 45PF	>1.5 EF	≥ 1 EF DP sustained
Peak Power (MW)	< 2.1	< 3.7	10	6	< 2	29	≤ 60
Total system memory	847 TB DDR4 + 70 TB HBM + 7.5 TB GPU memory	~1 PB DDR4 + High Bandwidth Memory (HBM) + 1.5PB persistent memory	2.4 PB DDR4 + 0.4 PB HBM + 7.4 PB persistent memory	1.92 PB DDR4 + 240TB HBM	> 250 TB	4.6 PB DDR4 +4.6 PB HBM2e + 36 PB persistent memory	> 10 PB
Node performance (TF)	2.7 TF (KNL node) and 166.4 TF (GPU node)	> 3	43	> 70 (GPU) > 4 (CPU)	> 70 TF	TBD	> 130
Node processors	Intel Xeon Phi 7320 64-core CPUs (KNL) and GPU nodes with 8 NVIDIA A100 GPUs coupled with 2 AMD EPYC 64-core CPUs	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	2 IBM Power9 CPUs + 6 Nvidia Volta GPUs	CPU only nodes: AMD EPYC Milan CPUS; CPU-GPU nodes: AMD EPYC Milan with NVIDIA A100 GPUs	1 CPU; 4 GPUs	1 HPC and AI optimized AMD EPYC CPU and 4 AMD Radeon Instinct GPUs	2 Intel Xeon Sapphire Rapids and 6 Xe Ponte Vecchio GPUs
System size (nodes)	4,392 KNL nodes and 24 DGX-A100 nodes	9,300 nodes 1,900 nodes in data partition	4608 nodes	> 1,500(GPU) > 3,000 (CPU)	> 500	> 9,000 nodes	> 9,000 nodes
CPU-GPU Interconnect	NVLINK on GPU nodes	N/A	NVLINK Coherent memory across node	PCIe		AMD Infinity Fabric Coherent memory across the node	Unified memory architecture, RAMBO
Node-to-node interconnect	Aries (KNL nodes) and HDR200 (GPU nodes)	Aries	Dual Rail EDR-IB	HPE Slingshot NIC	HPE Slingshot NIC	HPE Slingshot	HPE Slingshot
File System	200 PB, 1.3 TB/s Lustre 10 PB, 210 GB/s Lustre	28 PB, 744 GB/s Lustre	250 PB, 2.5 TB/s GPFS	35 PB All Flash, Lustre	N/A	695 PB + 10 PB Flash performance tier, Lustre	≥ 230 PB, ≥ 25 TB/s DAOS

U.S. DEPARTMENT OF ENERGY | Office of Science

Frontier: https://www.olcf.ornl.gov/frontier/
Aurora: https://www.alcf.anl.gov/aurora

ASCR Computing Upgrades At-a-Glance
November 24, 2020

ECP Teams are accessing early exascale system hardware & software



EXASCALE COMPUTING PROJECT

Exascale Early Access Systems at DOE Leadership Computing Facilities

Arcticus at ALCF

HW Description –
- 17 nodes with Intel Xe-HP GPU

SW env – SLES + Intel Aurora SDK

Access: Available to ECP members covered under appropriate NDA.

Communication and trouble shooting: Email list / slack channels, Confluence docs

Support: ALCF staff and Intel Center of Excellence staff

Notes: Systems are shared with Argonne Early Science Program. Incomplete feature support, updated frequently.

Spock at OLCF

HW Description – 12 nodes each with:
- 1x 64 core AMD EPYC CPU
- 4x MI100 GPUs, each with 32 GB HBM
- access to OLCF home and project areas

SW env - SW - RHEL, Slurm, Cray Programming Environment (PE), AMD ROCm (HIP),

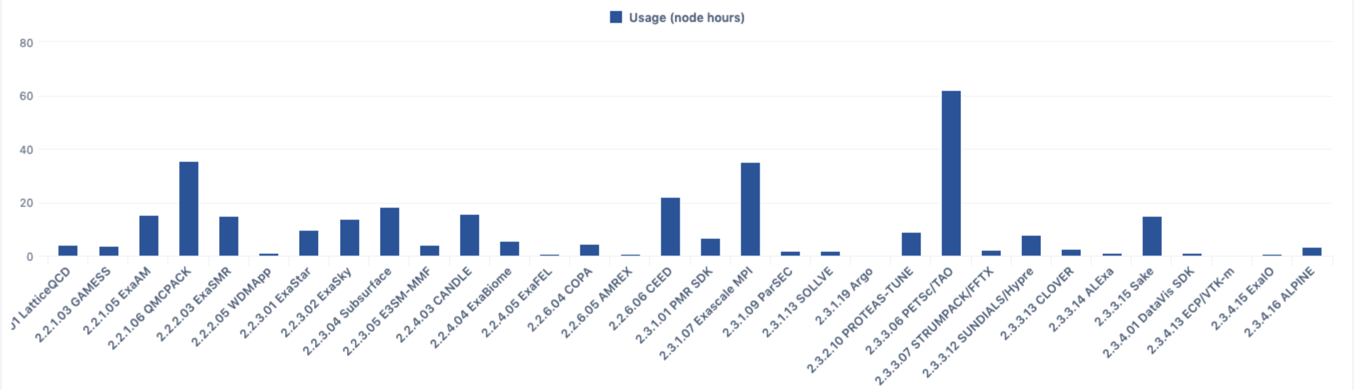
Access: Available to all ECP members

Communication and trouble shooting: Web documentation, help desk, slack channel

Support: OLCF Staff and HPE Center of Excellence staff

Notes: Hardening of programming environment and site specific configurations on-going

Other early hardware includes Iris and Yarrow systems at ALCF, and Birch and Tulip systems for Frontier



EXASCALE COMPUTING PROJECT

ECP Application Portfolio: 24 First-Movers of Strategic Importance to DOE

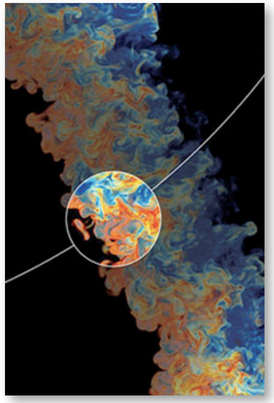
ECP Application Portfolio: 24 First-Movers of Strategic Importance to DOE

National security

Next-generation, **stockpile stewardship** codes

Reentry-vehicle- environment simulation

Multi-physics science simulations **of high-energy density physics** conditions



Energy security

Turbine **wind plant** efficiency

Design and commercialization of **SMRs**

Nuclear fission and fusion reactor **materials design**

Subsurface use for **carbon capture**, petroleum extraction, waste disposal

High-efficiency, low-emission **combustion engine** and gas turbine design

Scale up of **clean fossil fuel** combustion

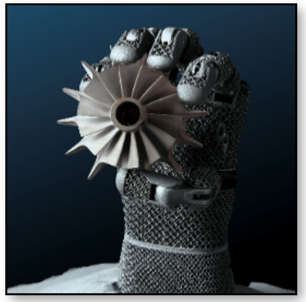
Biofuel catalyst design

Economic security

Additive manufacturing of qualifiable metal parts

Reliable and efficient planning of the **power grid**

Seismic hazard risk assessment



Scientific discovery

Cosmological probe of the standard model of particle physics

Validate fundamental laws of nature

Plasma wakefield accelerator design

Light source-enabled **analysis of protein and molecular structure** and design

Find, predict, and control materials and properties

Predict and control **magnetically confined fusion plasmas**

Demystify **origin of chemical elements**

Earth system

Accurate regional impact assessments in **Earth system models**

Stress-resistant crop analysis and catalytic conversion of **biomass-derived alcohols**

Metagenomics for analysis of biogeochemical cycles, climate change, environmental remediation

Health care

Accelerate and translate **cancer research** (partnership with NIH)

EXASCALE COMPUTING PROJECT

ECP Application Portfolio: 24 First-Movers of Strategic Importance to DOE

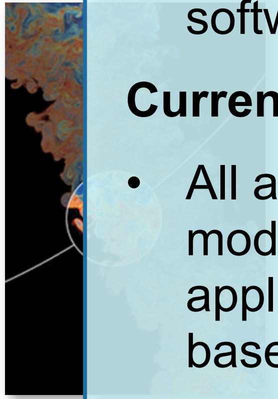
Starting Point

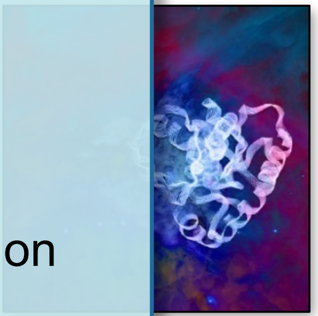
- **24 applications** and **6 co-design** projects
  - Including **78 separate codes**
  - Representing over **10 million lines of code**
  - Many supporting large user communities
  - Covering broad range of mission critical S&E domains
  - Mostly all MPI or MPI+OpenMP on CPUs
- Each envisioned innovative S&E enabled by 100X increase in computing power
- Path to harnessing 100-fold improvement initially unknown likely to have disruptive impact on software unlike anything in last 30 years

Current status

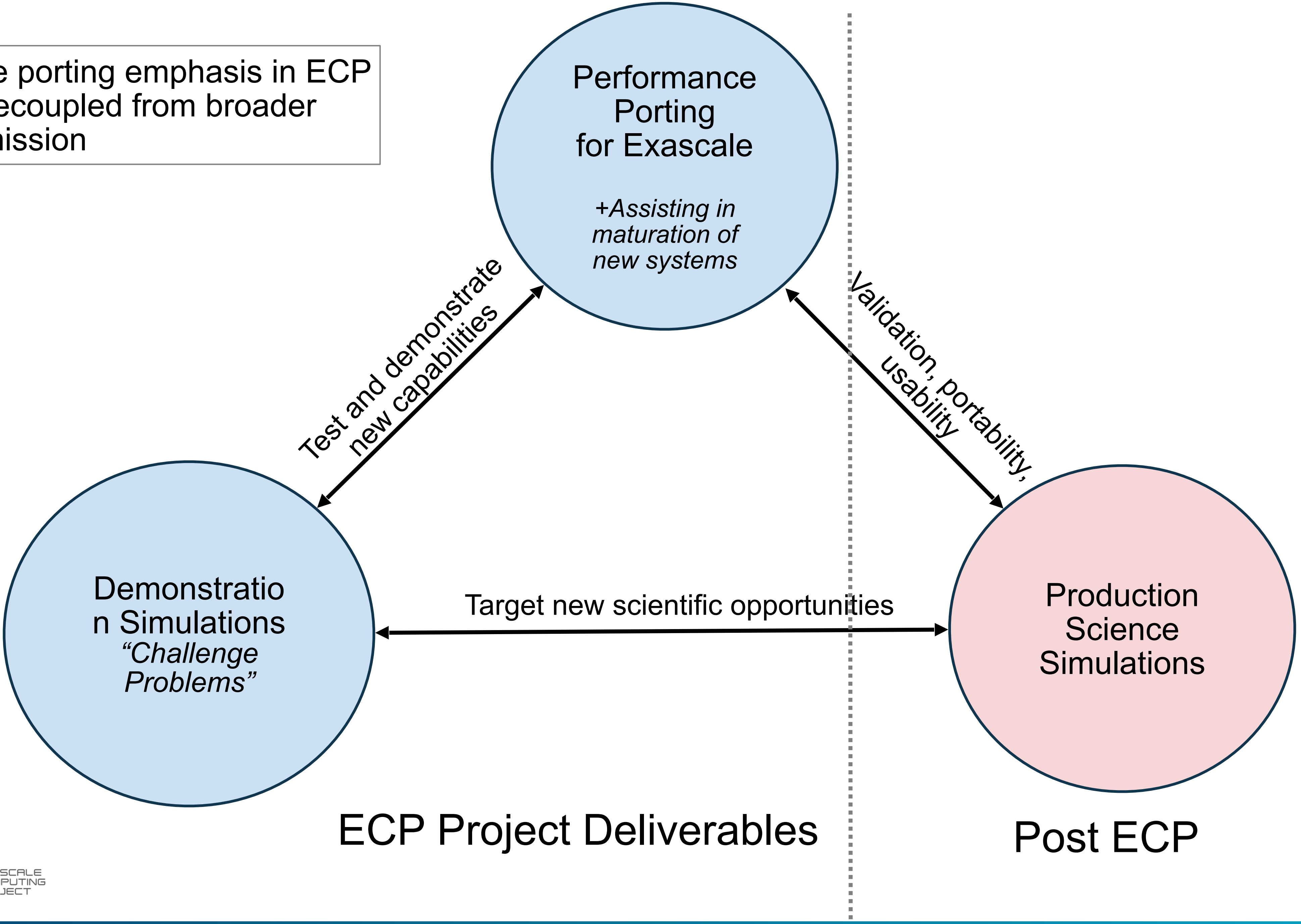
- All applications have, with their own unique development plans, made tremendous progress in model and algorithm development and software architecture redesign / refactor. Most applications have integrated and adopted software abstraction layers or co-designed motif-based components and frameworks to ensure efficient and portable GPU implementations.
- Many application have already realized >50X increase in science work rate performance on the Summit system at ORNL since starting ECP development activities in 2016

→ **Massive software investments**





Performance porting emphasis in ECP cannot be decoupled from broader simulation mission



Efficiently utilizing GPUs goes far beyond typical code porting

Port Code

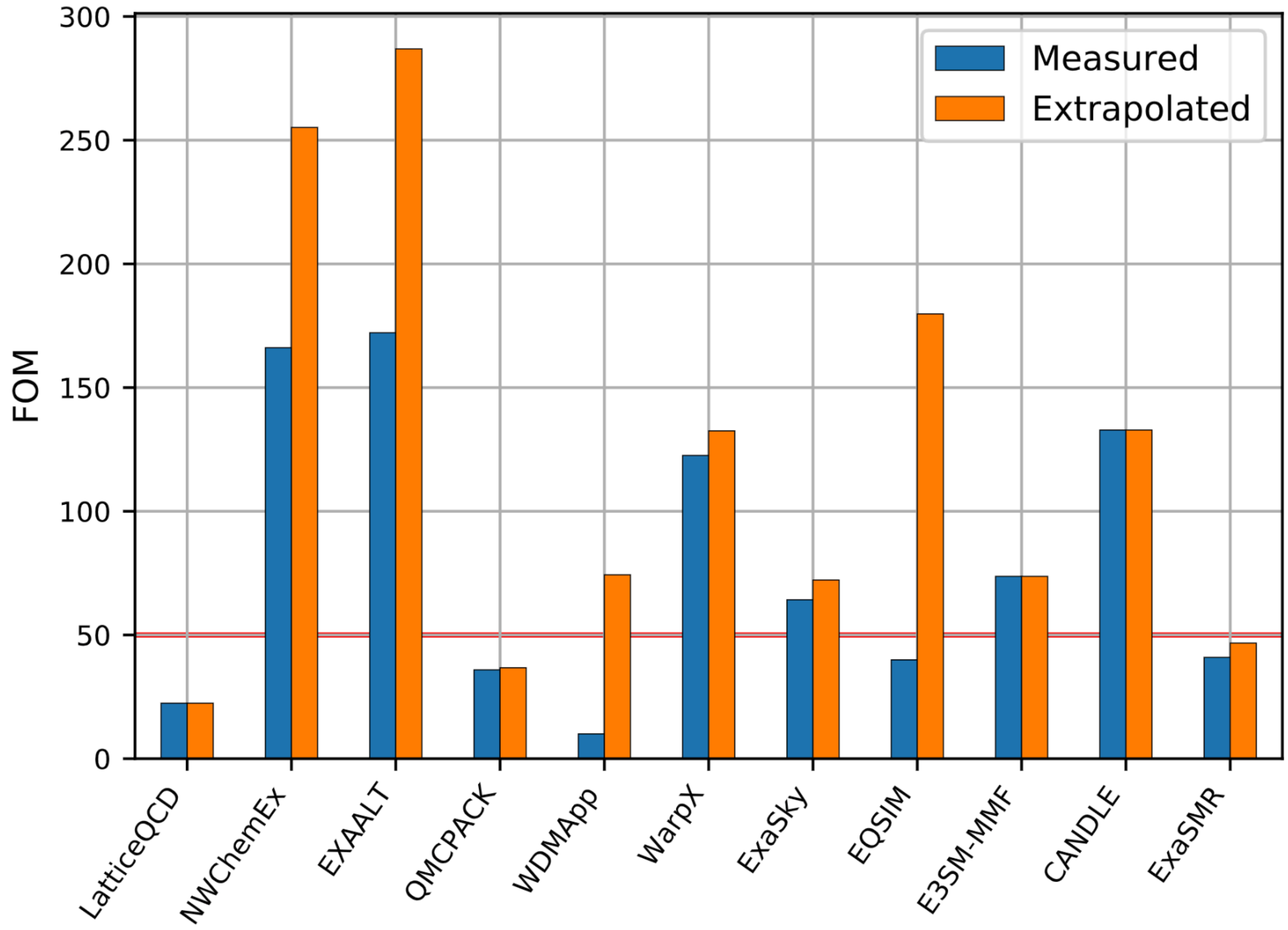
- Rewrite, profile, and optimize
- Memory coalescing
- Loop ordering
- Kernel flattening

Adapt Numerics

- Reduced synchronization
- Reduced precision
- Communication avoiding

Adapt Models

- Mathematical representation
- “On the fly” recomputing vs. lookup tables
- Prioritization of new physical models



EXASCALE COMPUTING PROJECT

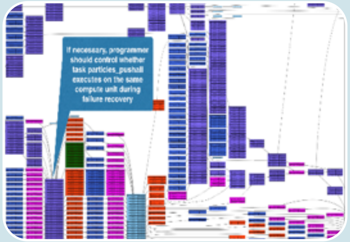
ECP Application Development Highlights

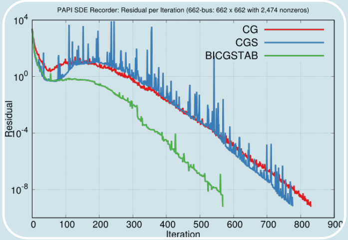
- **ExaSGD** (*Optimizing Stochastic Grid Dynamics at Exascale*) team released version 0.5 of ExaPF, a power flow solver for GPUs written in the Julia programming language
  - Package supports the toolchain for the reduced space optimization method where a reduced Hessian is accumulated on the GPU. ExaPF is based on the recently investigated optimized cuSOLVER_RF linear solver, exploiting the storage of the factors on the GPU for later refactorization
  - Reduced Hessian was computed on various case sizes ranging from 300 to 9241 buses, showing a 10x speedup of the GPU NVIDIA V100 implementation over the CPU implementation using UMFPACK.
- **CoPA** (*Co-Design Center for Particle Applications*) team issue a new version of the Cabana particle library that fully supports running on Intel GPUs with SYCL
  - Nearly feature-complete Kokkos SYCL backend allowing Cabana SYCL runs without any disabled features.
  - Demonstrated with CabanaMD (molecular dynamics) and ExaMPM (solver for the direct numerical simulation of powder-scale melt pool behavior) proxy apps
- **ExaWind** (*Exascale Predictive Wind Plant Flow Physics Modeling*) team demonstrated improved time-to-solution for AMR-Wind (Eulerian background flow solver for domain between turbines)
  - Compared strong- and weak-scaling, as well as time-to-solution, between flow solvers
  - Adopted a preconditioned Krylov solver from AMReX software framework for momentum-equation solves: 1.7x speedup
  - Explored adoption of a mixed CPU/GPU multigrid solver (similar to NekRS) combining geometric multigrid on fine levels using GPUs and algebraic multigrid on coarse levels using CPUs

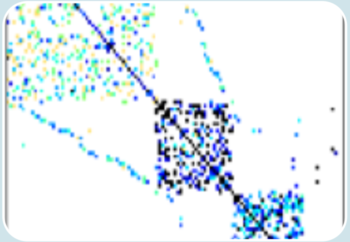
ECP Application Development Highlights

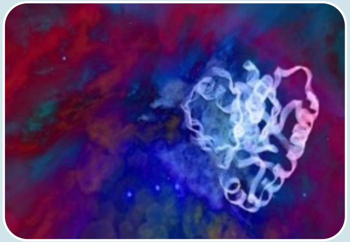
- **ExaFEL** (*Data Analytics at the Exascale for Free Electron Lasers*) team completed a new framework to characterize the robustness of the Multi-Tiered Iterative Phasing (MTIP) algorithm for single particle imaging reconstruction in the presence of realistic noise
  - Instrumental in measuring and understanding the effect of different types of noise, tuning the algorithm for maximum efficiency, and will enable preparation for analysis of Linac Coherent Light Source data in FY22
  - Viable approach to characterize Poisson noise and examine MTIP's tolerance to shot-to-shot beam mis-centering
- **EXAALT** (*Molecular Dynamics at the Exascale – Materials Science*) team is advancing the development of the SNAP machine-learned potential form, resulting in a significant increase in performance using Kokkos and in accuracy with a new Neural Network formulation.
  - On V100 processor the performance has improved by about 30× since the beginning of ECP
- **WarpX** (*Exascale Modeling of Advanced Particle Accelerators*) team developed a novel hybrid nodal-staggered pseudo-spectral Particle-In-Cell (PIC) scheme that combines the advantages of stability of nodal PIC methods and the efficiency of staggered PIC methods
  - Numerical tests were performed with WarpX in a variety of physical scenarios, ranging from the simulation of laser-driven and particle beam-driven plasma wakefield acceleration, showing 2× speedup, to the modeling of electron-positron pair creation in vacuum, with a speedup of at least 10×.

ECP ST has six technical areas









Programming Models & Runtimes

- Enhance and get ready for exascale the widely used MPI and OpenMP programming models (hybrid programming models, deep memory copies)
- Development of performance portability tools (e.g. Kokkos and Raja)
- Support alternate models for potential benefits and risk mitigation: PGAS (UPC++/ GASNet) ,task-based models (Legion, PaRSEC)
- Libraries for deep memory hierarchy and power management

Development Tools

- Continued, multifaceted capabilities in portable, open-source LLVM compiler ecosystem to support expected ECP architectures, including support for F18
- Performance analysis tools that accommodate new architectures, programming models, e.g., PAPI, Tau

Math Libraries

- Linear algebra, iterative linear solvers, direct linear solvers, integrators and nonlinear solvers, optimization, FFTs, etc
- Performance on new node architectures; extreme strong scalability
- Advanced algorithms for multi-physics, multiscale simulation and outer-loop analysis
- Increasing quality, interoperability, complementarity of math libraries

Data and Visualization

- I/O via the HDF5 API
- Insightful, memory-efficient in-situ visualization and analysis – Data reduction via scientific data compression
- Checkpoint restart

Software Ecosystem

- Develop features in Spack necessary to support all ST products in E4S, and the AD projects that adopt it
- Development of Spack stacks for reproducible turnkey deployment of large collections of software
- Optimization and interoperability of containers on HPC systems
- Regular E4S releases of the ST software stack and SDKs with regular integration of new ST products

NNSA ST

- Open source NNSA Software projects
- Projects that have both mission role and open science role
- Major technical areas: New programming abstractions, math libraries, data and viz libraries
- Cover most ST technology areas
- Subject to the same planning, reporting and review processes

ECP's Software Development Kits (SDKs) Span All Technology Areas

ECP's Extreme Scale Scientific Software Stack (E4S) embodies the latest Software Technology products developed in ECP and packaged in SDKs. The latest May 2021 release (https://e4s.io) includes 76 distinct products using the Spack package manager in a full-feature containerized release. E4S also supports AI/ML packages such as TensorFlor, PyTorch, and Horovod. E4S is available for download from Dockerhub, with bare metal and custom containers also supported using the E4S Spack build cache.

PMR Core (17)	Compilers and Support (7)	Tools and Technology (11)	xSDK (16)	Visualization Analysis and Reduction (9)	Data mgmt, I/O Services, Checkpoint restart (12)	Ecosystem/E4S at-large (12)
QUO	openarc	TAU	hypre	ParaView	SCR	mpiFileUtils
Papyrus	Kitsune	HPCToolkit	FleSCI	Catalyst	FAODEL	TriBITS
SICM	LLVM	Dyninst Binary Tools	MFEM	VTK-m	ROMIO	MarFS
Legion	CHiLL autotuning comp	Gotcha	Kokkoskernels	SZ	Mercury (Mochi suite)	GUFI
Kokkos (support)	LLVM openMP comp	Caliper	Trilinos	zfp	HDF5	Intel GEOPM
RAJA	OpenMP V & V	PAPI	SUNDIALS	VisIt	Parallel netCDF	BEE
CHAI	Flang/LLVM Fortran comp	Program Database Toolkit	PETSc/TAO	ASCENT	ADIOS	FSEFI
PaRSEC*		Search (random forests)	libEnsemble	Cinema	Darshan	Kitten Lightweight Kernel
DARMA		Siboka	STRUMPACK	ROVER	UnifyCR	COOLR
GASNet-EX		C2C	SuperLU		VeloC	NRM
Qthreads		Sonar	ForTrilinos		IOSS	ArgoContainers
BOLT			SLATE		HXHIM	Spack
UPC++			MAGMA			
MPICH			DTK			
Open MPI			Tasmanian			
Umpire			TuckerMPI			
AML						

Legend
- PMR
- Tools
- Math Libraries
- Data and Vis
- Ecosystems and delivery

Extreme-scale Scientific Software Stack (E4S)

- E4S: HPC Software Ecosystem – a curated software portfolio
- A **Spack-based** distribution of software tested for interoperability and portability to multiple architectures
- Available from **source**, **containers**, **cloud, binary caches**
- Leverages and enhances SDK interoperability thrust
- Not a commercial product – an open resource for all
- Oct 2018: E4S 0.1 - 24 full, 24 partial release products
- Jan 2019: E4S 0.2 - 37 full, 10 partial release products
- Nov 2019: E4S 1.0 - 50 full, 5 partial release products
- Feb 2020: E4S 1.1 - 61 full release products
- Nov 2020: E4S 1.2 (aka, 20.10) - 67 full release products
- Feb 2021: E4S 21.02 - 67 full release, 4 partial release
- May 2021: E4S 21.05 - 76 full release products

Spack

E4S

https://e4s.io

Lead: Sameer Shende (U Oregon)

Also include other products .e.g.,
AI: PyTorch, TensorFlow, Horovod
Co-Design: AMReX, Cabana, MFEM

Strategic Vision for E4S

- E4S has emerged as a new top-level component in the DOE HPC community, enabling fundamentally new relationships
- E4S has similar potential for new interactions with other US agencies, US industry and international collaborators.
- The E4S portfolio can expand to include new domains (ML/AI), lower—level components (OS), and more.
- E4S can provide better (increased quality), faster (timely delivery of leading edge capabilities) and cheaper (assisting product teams)
- ECP believes E4S is a viable springboard for *the* scientific software stack well past (decades?) ECP. We plan to work with DOE and our colleagues in HPC community in laying out a strategic plan for how E4S must evolve and adapt to coming needs and programs (AI4S, Edge, Quantum, etc.)

ECP Software Technology Highlights

- **SICM** (*Simplified Interface to Complex Memory*) team investigated the potential of guiding data management with Nalu-Wind, a generalized, unstructured, massively parallel, incompressible flow solver for wind turbine and wind farm simulations
  - Analysis shows significant potential for increasing throughput of Nalu-Wind on heterogenous memory systems
- **xSDK4ECP** (*Extreme-scale Scientific xSDK for ECP*) team delivered the second release of xSDK-examples v.0.2.0, a suite of example codes that demonstrate interoperability between select xSDK libraries and provide training for xSDK users on mixed package use
  - Available at https://github.com/xsdk-project/xsdk-examples, with two examples that demonstrate CUDA capabilities
- **SLATE** subteam, part of the **CLOVER** (*Computational Libraries Optimized Via Exascale Research*) team, added ROCm support for AMD GPU platforms to the latest releases of the SLATE and BLAS++ linear algebra libraries
  - BLAS++ now provides portable interfaces across both cuBLAS and rocBLAS backends for GPU BLAS, including Batched BLAS, as well as interfaces to traditional CPU BLAS
  - Support also being developed for BLAS++ for oneMKL for Intel GPU platforms
  - SLATE uses BLAS++ as a portability layer to support both CUDA and ROCm (https://bitbucket.org/icl/slate)

ECP Software Technology Highlights

- **Argo project** (Low-level resource management for the OS and runtime) released UMap version 2.1.0 (https://github.com/LLNL/umap
  - Major enhancements for improved scalability at high node-level including concurrency-aware adaptation of buffer management and scheduling.
  - Used for memory mapping k-mer databases in a production metagenomic classification software (LMAT) used in COVID-19 research. With application-specific optimizations of paging and caching policies, UMap demonstrates an increase in speed of up to 90% over the system mmap service in LMAT for production genome classifications.
- **SOLLVE** (*Scaling OpenMP with LLVm for Exascale*) released BOLT v2.0 (https://bolt-omp.org), a runtime system for OpenMP.
  - Addresses intranode scalability bottlenecks and MPI interoperability issues by using lightweight Argobots threads.
  - Evaluation shows that BOLT successfully improves the performance of applications that contain fine-grained parallelism.
- The **Hypre** team generated a new release with performance enhancements and additional GPU capabilities.
  - HIP support for AMD GPUs for unstructured solvers, including algebraic multigrid, for specific settings.
  - UMPIRE support for memory pooling on GPUs, GPU support for parallel ILU, CUDA-based triangular smoothers, and GPU-enabled examples, and new interfaces that allow setting memory location and execution policy.

ECP Software Technology Highlights

- **MPI Forum**, keepers of the Message Passing Interface (MPI) standard, has formally approved MPI 4.0.
  - Major new features championed and prototyped by the OMPI-X (*Open MPI for Exascale*) team include partitioned communications and sessions, which allows optimization in highly threaded environments and improving support for complex applications such as coupled multiphysics and in situ data analytics
  - Improvements in error handling and recovery and more complex hardware topologies.
- **Sake** (*Solver and Kernels for Exascale*) team released a new version of the Kokkos Kernels library (version 3.4) with support for HIP and SYCL backends targeting Frontier and Aurora systems.
  - Provides path for sparse/dense linear algebra and graph kernels on exascale platforms.
  - Early performance tests of the HIP backend on Spock show that Kokkos Kernels' implementation of sparse matrix-vector multiplication (SpMV) is competitive with AMD's rocSPARSE implementation.
- **ALPINE** (*Algorithms and Infrastructure for In Situ Visualization and Analysis*), **ZFP** (*Compressed Floating-Point Arrays*), and VTK-m teams improved and validated the Lagrangian-based in situ reduction approach for vector field data.
  - Resulted in a 3× speedup using 2048 Summit GPUs
  - Validated using ECP application codes Nyx and SW4, including SW4 runs of 384 GPUs on Summit

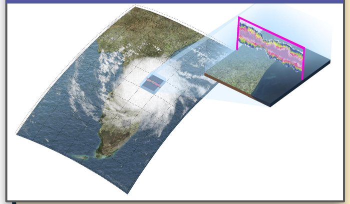
ECP **Hardware and Integration (HI)**

A capable exascale computing ecosystem made possible by “integrating” applications, software, and hardware innovations with training, outreach and allocation management with deep partnerships with DOE Facilities.

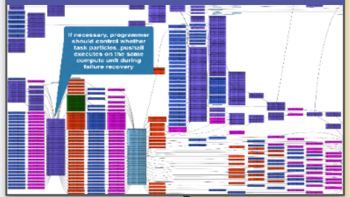
HI enables the “last mile.”

Hardware Innovations and Evaluation

Application Integration



Software Deployment



Training Outreach and Allocations

ANL
Intel/Cray

LLNL
Cray

ORNL
Cray/AMD

HI leadership team : Accomplished technical leaders with Facility experience

Katie Antypas, **HI Director**
15 years experiencing supporting HPC users and deploying HPC systems (LBNL)

Bronis de Supinski, **PathForward**
5 years as the CTO for the Livermore Computing facility (LLNL)

Scott Pakin, **HW Evaluation**
17 years in performance analysis and SW development at the ACES Facility (LANL)

Scott Parker, **Application Integration at Facilities**
13+ years experience working on performance optimization for scientific applications (ALCF)

Susan Coghlan, **HI Deputy Director**
30 years experience acquiring, deploying, managing extreme scale systems at DOE Facilities (Argonne)

Ryan Adamson, **Software Deployment at Facilities**
12 years of systems and security administration OLCF HPC Core Operations Group Lead (ORNL)

Haritha Siddabathuni Som, **Facility Resource Utilization**
14 years in field and manager of the ALCF User Experience Team (ANL)

Ashley Barker, **Training and Productivity**
8 years as a group leader of user assistance and outreach at the OLCF (ORNL)

ECP EXASCALE COMPUTING PROJECT

ECP’s Hardware and Integration Focus Area: Recent Activities

- Early hardware is available for ECP early users at ALCF and OLCF
- Application Integration teams using early hardware to aid in Application Development (AD) teams in optimization and porting
- Software Deployment team has converged around E4S as deployment vehicle and are building and installing ST software on early hardware
- Training and Productivity subproject is partnering closely with facilities to put on joint training events
- Perlmutter at NERSC is identifying bugs/issues with system software and networking and serving as risk mitigation for Frontier and Aurora
- Tracking early hardware usage and plans for a new user program for application demonstration
- Hardware Evaluation is set to complete final studies on memory technologies, analytical modeling and network simulation

ECP's Software Deployment element

Tremendous progress the last 18 months

- Just transitioned leadership to Ryan Adamson (ORNL)
- Opportunities for integrating spack, SDK, and E4S software ecosystems had not yet been explored fully
- Demonstrated early instances of continuous integration infrastructure on production systems
- Limited testing of ST packages on production systems
- Had a strong vision, but lacked a simple, easily understood plan by stakeholders

18 months later

- Broke apart Software Deployment area into Software Integration (testing and deployment) and Continuous Integration
- Made decision to leverage the E4S software stack and packaging ***significantly simplifying*** testing, deployment and continuous integration plan
- E4S build pipelines ***with CI*** deployed at each facility on production systems (including GPU architectures)
- Software testing on early access hardware has begun
- Site-Local CI deployed on all facility managed early hardware delivered

	Q3 FY21
E4S Testsuite Tests	61/67
ST Spack Tests	42/67
ST Spack Tests Targeting Exascale Features	25/67
ST products ready for installation on early hardware	16/67

Application Integration at the Facilities Portfolio

Accelerating application readiness for the exascale architectures

Scott Parker, ALCF

Balint Joo, OLCF

Strategy: Match applications with facility readiness efforts.

Progress Assessment: Progress towards technical execution plans measured quarterly; annual external assessment.

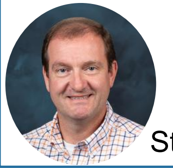
Goal: 21 performant exascale applications that run on Aurora, Frontier

Application Development 2.2

- Chemistry and Materials Applications 2.2.1
- Energy Applications 2.2.2
- Earth and Space Science Applications 2.2.3
- Data Analytics and Optimization Applications 2.2.4
- National Security Applications 2.2.5
- Co-Design 2.2.6

Chris Knight

01 Aurora – 14 applications strongly engaged by ALCF for Aurora; others to follow as resources allow. Best practices are being developed and shared.



Stephen Nichols

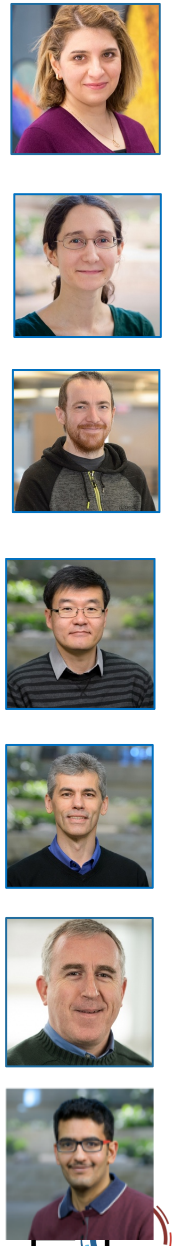
02 Frontier - 12 ECP applications were selected to participate with CAAR program in July, 2019. Applications may transition in and out of the program as progress is made.

Debbie Bard

03 PreExascale – 6 ECP AD applications identified to participate in NESAP for Perlmutter with ECP funding.

Goal: Progress towards exascale readiness develops, and NESAP-ECP apps transition to LCF facilities after NESAP.

AppInt Teams are deeply engaged and critical to ECP application success



Yasaman Ghadar
ALCF
EXXALT

Colleen Bertoni
ALCF
GAMESS

Thomas Applencourt
ALCF
QMCPACK

JaeHyuk Kwack
ALCF
ExaWind

Victor Anisimov
ALCF
NWChemEx

Timothy Williams
ALCF
WDMApp

Murali Emani
ALCF
CANDLE

Christopher Knight
ALCF
EXAALT

James Osborn
ALCF
LatticeQCD

Ray Loy
ALCF
E3SM-MMF

Vitali Morozov
ALCF
ExaSky

Brian Homerding
ALCF
EQSIM

Kris Rowe
ALCF
ExaSMR

Brice Videau
ALCF
ExaStar

Servesh Muralidharan
ALCF
ExaFEL

Abhishek Bagusetty
ALCF
E3SM-MMF
NWChemEx

Dmitry Liakh
OLCF
NWChemEX

Dima Bykov
OLCF
GAMESS

Austin Harris
OLCF
ExaStar

Matt Norman
OLCF
E3SM-MMF

Stephen Nichols
OLCF
ExaAM

Bronson Messer
OLCF
ExaSky

Balint Joo
OLCF
LatticeQCD

Antigoni Georgiadou
OLCF
ExaSky
WDMApp

Phil Roth
OLCF
ExaSGD
ExaBiome

Dhruva Kulkani
NERSC
WDMApp

Muaaz Gul Awan
NERSC
ExaBiome

Kevin Gott
NERSC
WarpX

Rahul Gayatri
NERSC
EXXALT

Paul Lin
NERSC
WDMApp

Ronnie Chatterjee
NERSC
Combustion

Michael Rowan
NERSC
WarpX

Johannes Blaschke
NERSC
ExaFEL

Neil Mehta
NERSC
EXXALT

Felix Wittwer
NERSC
ExaFEL

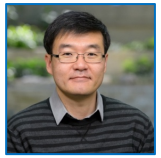
Raphaël Prat
NERSC
Subsurface

AppInt Teams are deeply engaged and critical to ECP application success


Yasaman Ghadar
ALCF
EXXALT


Colleen Bert
ALCF
GAMESS

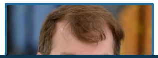

Thomas App
ALCF
QMCPACK


JaeHyuk Kw
ALCF
ExaWind


Victor Anisin
ALCF
NWChemEx


Timothy Will
ALCF
WDMApp


Murali Eman
ALCF
CANDLE


Christopher Knight

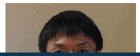

Abhishek

What do Application Engineers do?

- 45 staff members and post-docs engage with vendor experts, ECP application teams, and facility performance staff
- Implement algorithms on new architectures
- Optimize applications for new hardware and software
- Implement new programming models in applications

How are teams managed?

- AppInt Engineers are embedded with facilities staff, and also partner closely with application teams
- Quarterly milestones and work plans


Paul Lin
NERSC
WDMApp

Ronnie Chatterjee
NERSC
Combustion

Michael Rowan
NERSC
WarpX

Johannes Blaschke
NERSC
ExaFEL

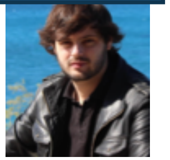
Neil Mehta
NERSC
EXXALT

Felix Wittwer
NERSC
ExaFEL


Servesh Muralidharan
ALCF
ExaFEL


Bronson Messer
OLCF
ExaSky

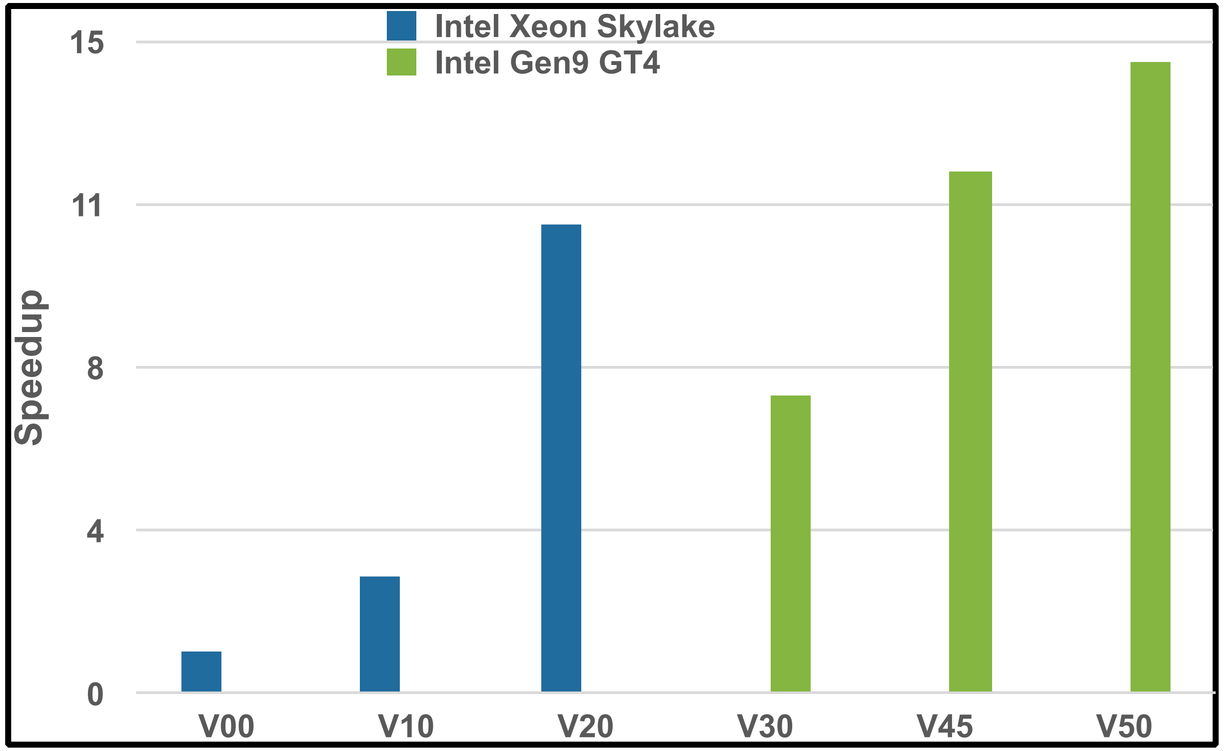

Rahul Gayatri
NERSC
EXXALT


Raphaël Prat
NERSC
Subsurface

GAMESS

Colleen Bertoni, ALCF

- ***Goal:*** Enable quantum chemistry on extremely large systems of interest in catalysis and energy research.
- ***Programming models***: Linear algebra libraries, CUDA, plans for HIP/DPC++ OpenMP
- Key physics module: RI-MP2 electron correlation method kernel
  - MPI/OpenMP threading for CPU with OpenMP offload for part of the RI-MP2 code for GPU



Porting RI-MP2 mini-app to Intel GPUs with OpenMP offload

Series of progressive optimizations, including OpenMP threading (V10), porting to MKL (V20), offloading to GPU (V30), restructuring loops (V45), and enabling concurrent CPU+GPU computation (V50)

GAMESS PI: Mark Gordon, Iowa State University

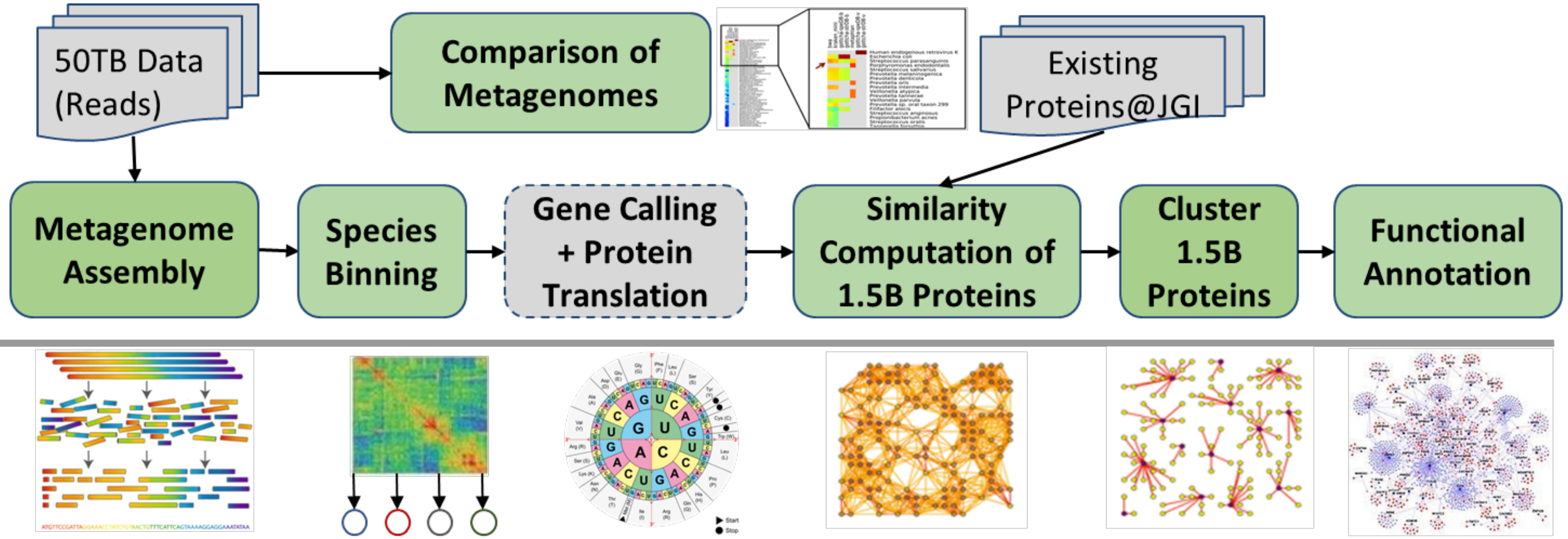
ExaBiome: Exascale Solutions to Microbiome Analysis

How do microbes affect disease and growth of switchgrass for

What happens to microbes after a wildfire?

Challenge Problem: Metagenomic Assembly
- Find species, genes and relative abundance in microbial communities

Stretch Goal: Metagenome Analysis
- Improve understanding of tree of life for microbes; aid in identifying gene function
- Track microbiome over time or space, changes in environment, etc.



Assembly: construct genome from "reads"

Bin contigs by species

Find start/stop codons, map to proteins

Similarity: Compute distance between "all pairs" of proteins

Clustering: find closely related proteins

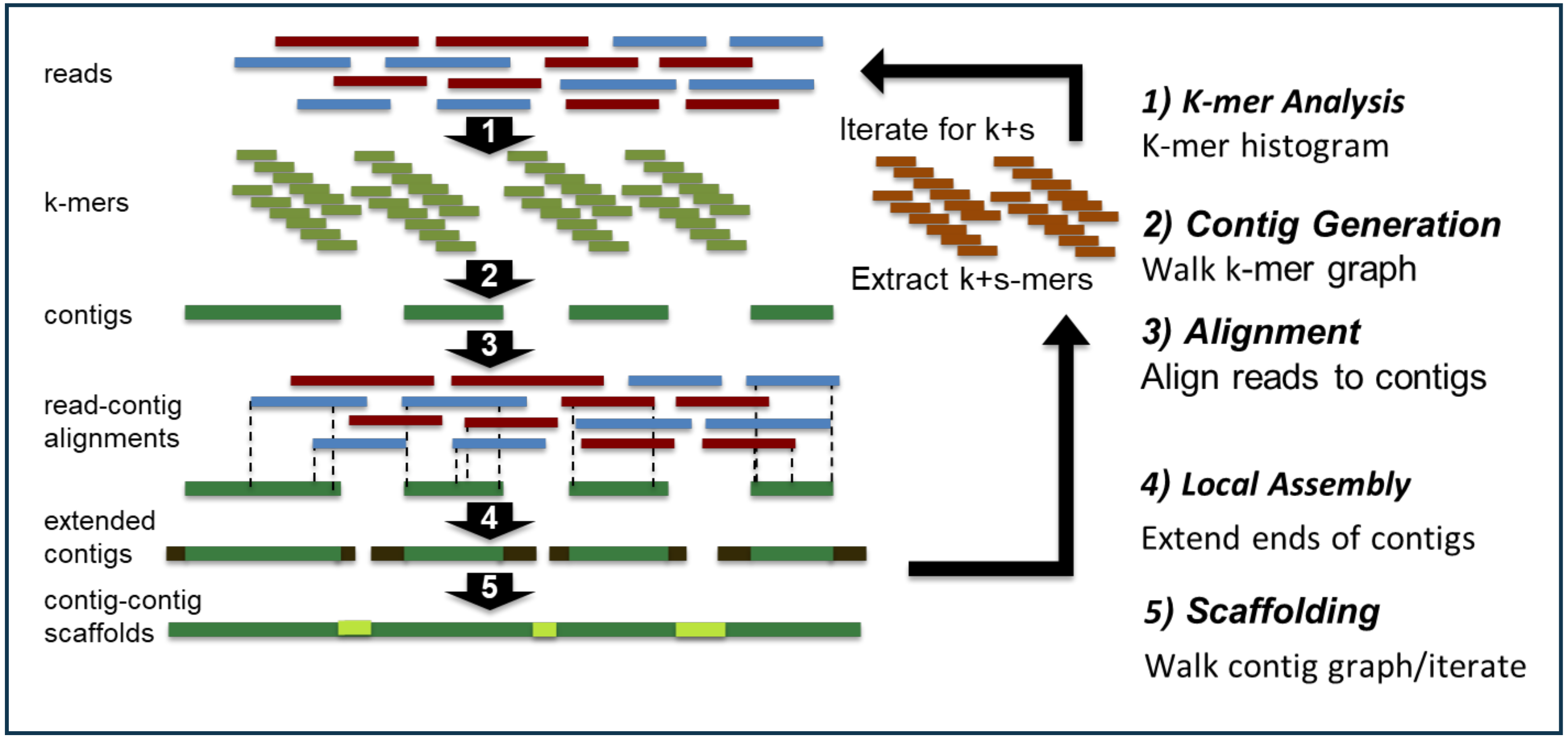
Assign function to proteins

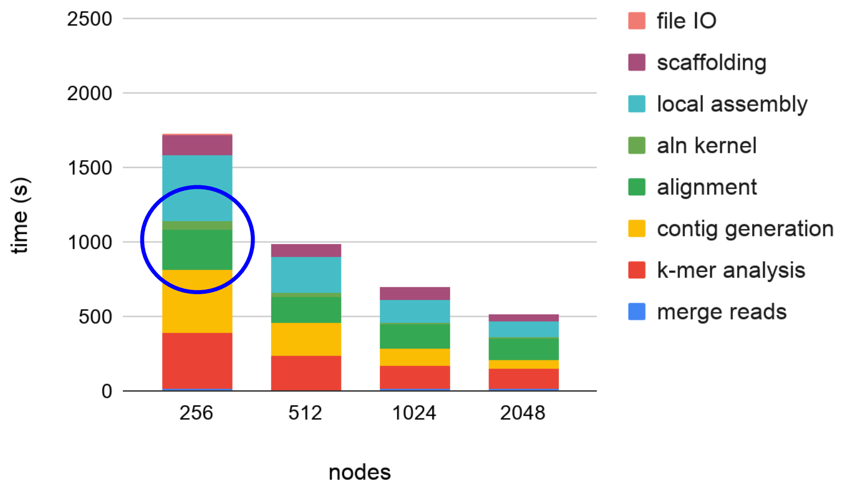
EXASCALE COMPUTING PROJECT

ExaBiome PI: Kathy Yelick, UC Berkeley

ExaBiome Challenge Problem: Metagenomic Assembly

MetaHipMer: metagenomic sequence assembly for short reads





MetaHipMer Approach: distributed hash tables, graph algorithms and other discrete algorithms

GPU Kernels: alignment (implemented), local assembly, k-mer analysis

EXASCALE COMPUTING PROJECT

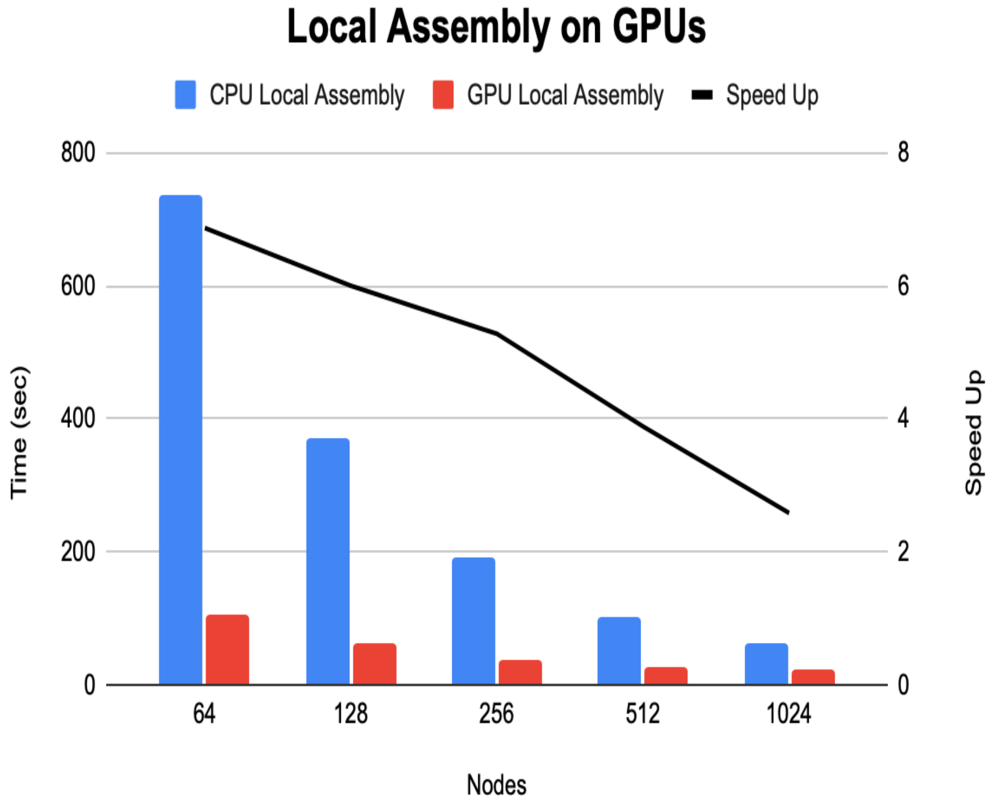
ExaBiome PI: Kathy Yelick, UC Berkeley

ExaBiome


Muaaz Gul Awan, Jonathan Madsen, NERSC

GPU Local Assembly Module for MetaHipMer:

- Graph traversal implemented using hashtables which induce a random memory access pattern that is not suitable for GPUs. Low level CUDA intrinsics used for performant implementation.
- **Local assembly module is 7x faster with GPUs when using 64 Summit nodes.** Performance becomes more communication dominant at higher number of nodes.
- **Integrating this in MetaHipMer pipeline improved performance by up to 42% at 64 nodes**.
- **Local assembly portion has now been reduced from 34% to 6.3% of total MetaHipMer runtime**.
- Accepted to SC21 and **nominated for best paper award**
- Able to process a 16TB metagenome assembly, largest to date



Feedback from ExaBiome:

- Performance improvement on GPUs required low level intrinsics making code somewhat NVIDIA specific.
- Team is working to port code to SYCL and HIP; some differences in intrinsics
- ExaBiome codes rely greatly on Integer performance and for that reason instruction roofline is used frequently for performance deep dives. However, the performance metrics required for constructing instruction roofline are not yet available through Intel and ROCM profilers. We are working closely with vendors on this.

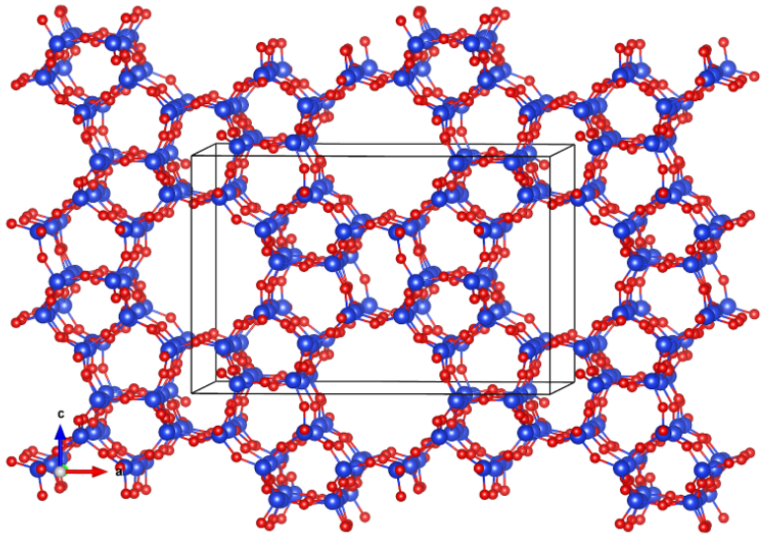
Development of high performance computing models for studying advanced biofuels

– Efficient Production of Biomass

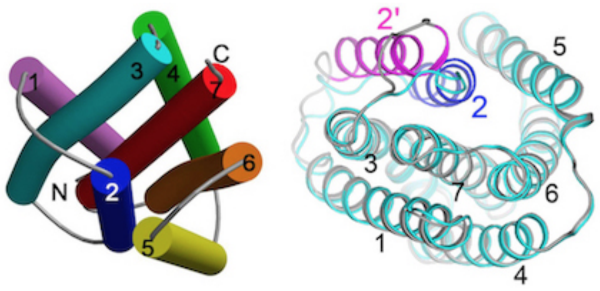
Development of detailed molecular model of ion transport across cellular membranes that control stress responses to aid in the development of stress-resistant crops (Q. Liu, BNL)

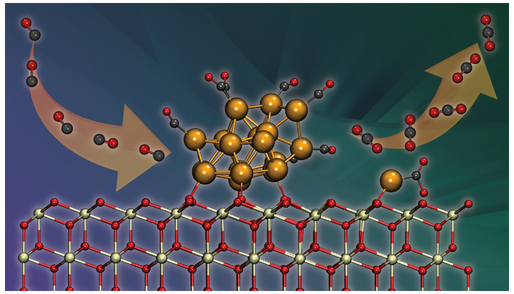
– Efficient Conversion of Biomass to Biofuels

Development of detailed molecular model of catalytic conversion of biomass-derived alcohols to biofuels to aid in the discovery of energy efficient conversion processes (P. Sushko, PNNL)



The zeolite H-ZSM-5 is a promising candidate for the conversion of biomass-derived alcohols to more useful hydrocarbons. The volume marked with black lines corresponds to a single unit cell.





NWChemEx: Project Overview

- **Software Goals**
  - Enable efficient and effective use of pre-exascale and exascale computing technologies for chemical science research and address targeted science challenges
    - Dramatically increase the performance and scalability of NWChemEx
    - **Reduce the dependence of highly accurate electronic structure methods on molecular system size (N_{atoms})**
  - Provide framework to support community-wide effort to develop next-generation molecular modeling capabilities to address broad range of molecular problems critical to DOE's mission
- **Work Plan**
  - Redesign and re-implement NWChem for exascale computing technologies
    - Address limitations of NWChem (**improved scalability, new methodologies and algorithms, enhanced portability**)
    - Update code base to C++ to enable more flexible and modular design
  - Predicting structure and energetics associated with targeted science challenges
    - Hartree-Fock/Density Functional Theory models
    - Coupled Cluster models for high fidelity predictions: CCSD(T)
    - Embedding (DFET) method for modeling complex molecular systems (CC active site plus DFT environment)

NWChemEx

Dmitry Liakh,
and Elvis Maradzike OLCF Post-Doc

NWChemEx is a new redesigned and improved version of the NWChem code capable of performing various quantum-chemistry simulations on heterogeneous exascale HPC platforms

Goal: Larger molecular systems can be calculated **faster** using **coupled-cluster** methods, allowing new quantum chemistry applications

Key challenges:

- Efficient utilization of GPU accelerators
- Optimization of the intra-node data transfers and inter-node communication
- Efficient implementation of the reduced-scaling coupled-cluster algorithms (task granularity challenge)
- Interoperability between all necessary libraries and runtimes

Accomplishments:

- NWChemEx introduces a **portability layer** called Tensor Algebra for Many-body Methods (**TAMM**).
- **TAMM** offloads tensor operations to appropriate processing **backends**.
- **OLCF TAL-SH library** has been integrated with TAMM as a processing backend for NVIDIA and AMD GPU.
- The **CCSD** module of NWChemEx has been fully integrated with the TAL-SH backend on Summit (NVIDIA GPU), demonstrating a **speed-up of 40X** compared to CPU-only NWChem on Titan.

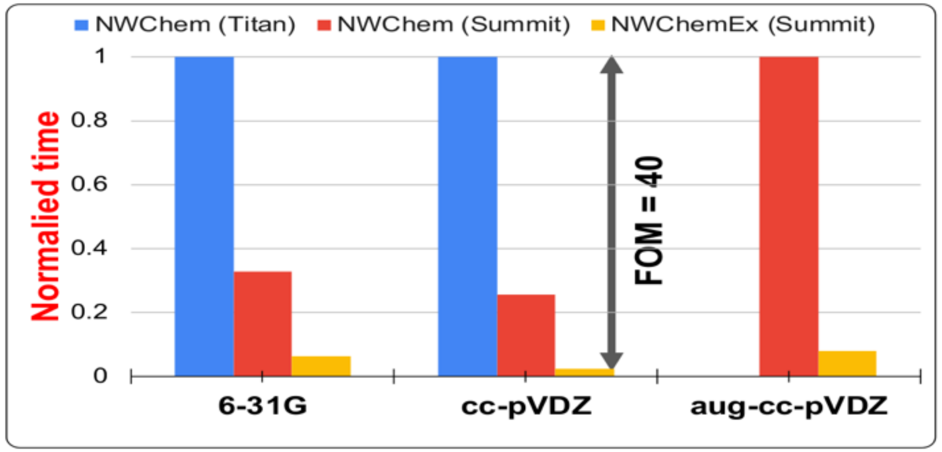
~39,000 basis functions

Basis set	# basis functions
6-31G	424
cc-pVDZ	737
aug-cc-pVDZ	1243

Coupled cluster singles and doubles

Basis	Platform	Nodes	NWChem	NWChemEx
6-31G	Titan Summit	406 100	8.5 min 2.8 min	-- 0.53 min
cc-pVDZ	Titan Summit	890 220	51 min 13 min	-- 1.3 min
aug-cc-pVDZ	Summit	256	74 min	5.8 min

Note: Timing for single iteration



TAMM based coupled-cluster single and doubles (CCSD) achieves **FOM of 40**

EXASCALE COMPUTING PROJECT

NWChemEx PI: Theresa Windus, Ames Lab

ECP is on track for successfully delivering on its objectives

- Application Development (AD)
  - Results on Summit have consistently exceeded expectations.
  - Early access hardware has allowed teams to begin tangible preparation for non-Nvidia GPUs.
  - Teams show keen awareness of what is needed to succeed and are eager to surpass threshold requirements
  - Applications are on track to meet project key parameters for performance and science outcomes
- Software Technology (ST)
  - Established a working testing framework and process in collaboration with HI and DOE HPC Facilities
  - Making excellent progress and are on track with delivery of key milestones and performance parameters
  - A software portfolio (E4S) that provides regular integration, testing, documentation and QA for ECP products
- Hardware & Integration (HI)
  - AD and ST teams are making productive use of Early Access Systems (EAS)
  - Integration of performance engineers on application teams showing to be best practice
  - Training programs are targeting EAS systems, well attended and received by ST and AD teams
  - ST products are being tested with Continuous Integration (CI) on the EAS platforms and NERSC Perlmutter

Questions?