EXASCALE AND ARTIFICIAL INTELLIGENCE: A GREAT MARRIAGE

Douglas Kothe, Chief Research Officer

Sandia National Laboratories, U.S.

February 12 – 16, 2024

EXASCALE CAPABILITIES DEPLOYED AT THE DOE FACILITIES

- #1 on the Top 500 list since May 2022: Currently 1.19 Exaflops
- Machine is increasingly stable and performant
- Full system available to all ECP teams since April 2023
- Excellent performance: 6 Gordon Bell & Special Prize Finalists to-date
- Many first-of-a-kind science goals achieved

- Full system installation completed June 2023
  - Currently in the scaling test phase; usual issues seen with serial #1 and large-scale systems
  - Lustre, interconnect, and some hardware issues being worked
- Limited access available July 2023 (ANL personnel); full system access for ECP early science expected November 2023

- Hardware being delivered to LLNL now; installation is underway
- Expected to exceed 2 Exaflops when deployed in mid to late 2024
- Brief period of open science before machine transitions to classified use focused on stockpile stewardship

THE EXASCALE COMPUTING PROJECT

Maintain **international leadership in HPC**

Promote the **health of the US HPC industry**

Deliver a **sustainable software ecosystem** used and maintained for years to come

Ensure that exascale systems can be used to deliver **mission-critical**

7-year, $1.8B

US Department of Energy project funded 1000+ people at national labs, universities, US industries

This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

Application Development

- Develop and enhance the predictive capability of applications, **25 applications, 6 Co-Design Centers**

Software Technology

- Deliver expanded and vertically integrated software stack, **70 unique products**

Hardware and Integration

- Application integration and software deployment to facilities, exascale node and system design, **6 US HPC vendors**

THE EXASCALE COMPUTING PROJECT

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
ML, graph analytics, mesh refinement, PDE discretization, particles, online data analytics


Andrew Siegel, AD Director
Erik Draeger, AD Deputy Director

Software Technology (ST)

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

70 unique software products spanning programming models and runtimes, math libraries, data and visualization, development tools


Mike Heroux, ST Director
Lois Curfman McInnes, ST Deputy Director

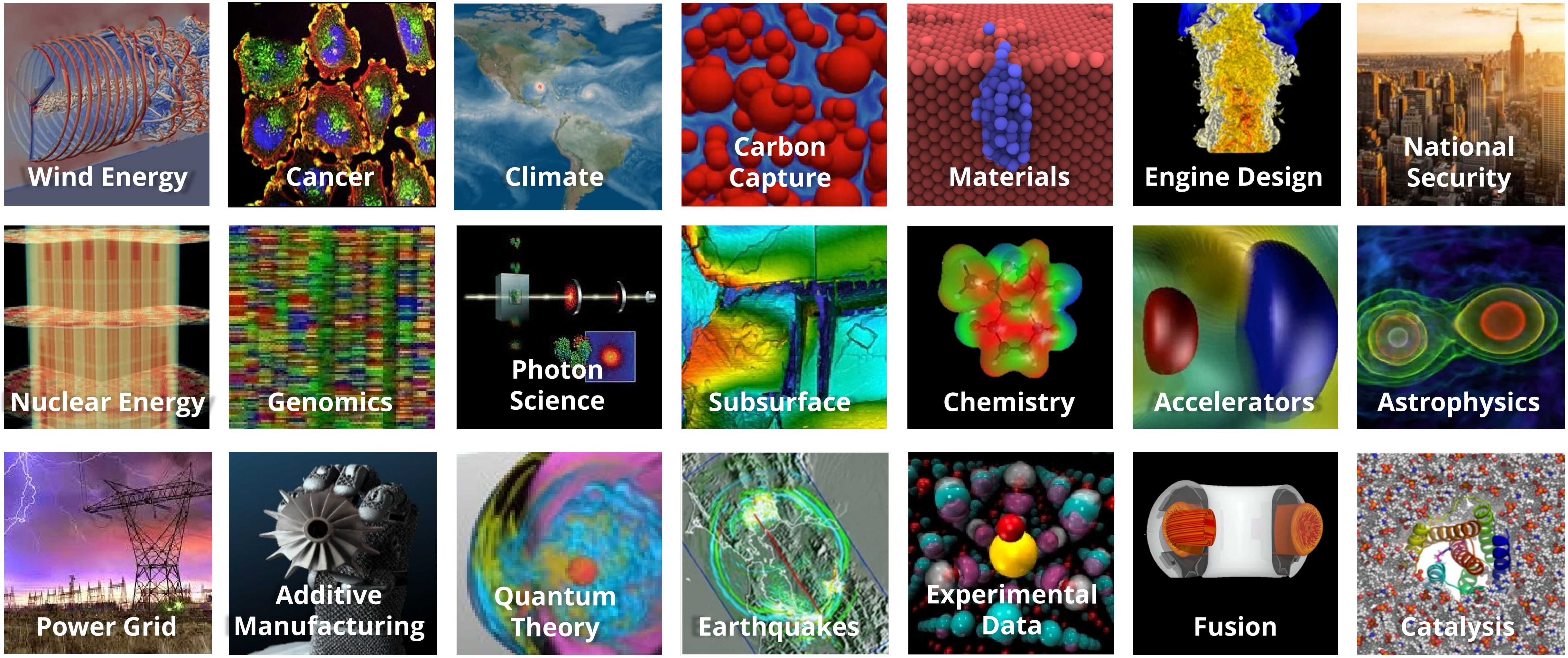
Hardware and Integration (HI)

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


Richard Gerber, HI Director
Susan Coghlan, HI Deputy Director

THE BREADTH OF EXASCALE-INDICATIVE OF A SEA CHANGE IN COMPUTING ABILITIES FOR DOE AND THE NATION



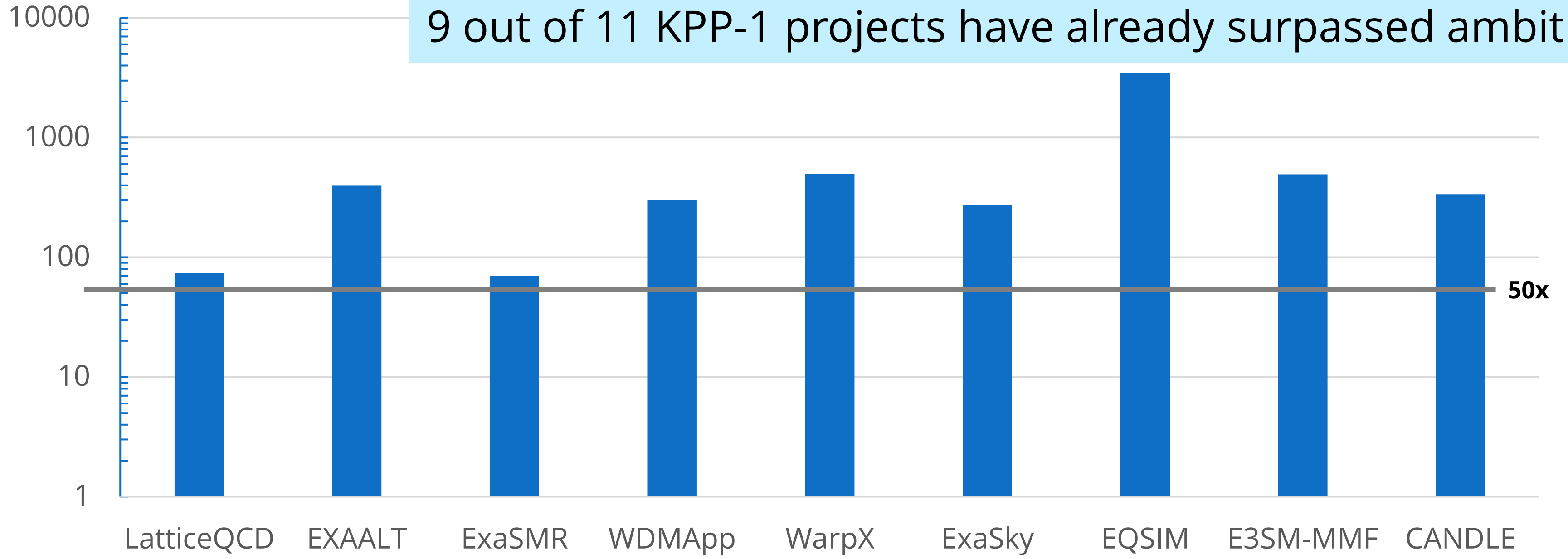
THE BREADTH OF EXASCALE-INDICATIVE OF A SEA CHANGE IN COMPUTING ABILITIES FOR DOE AND THE NATION

24 applications and **6 co-design** projects

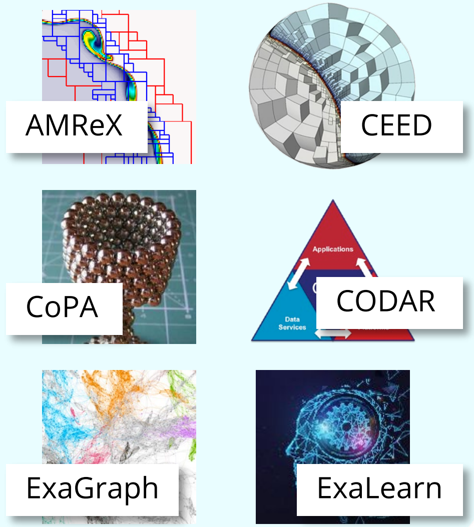
- Including **62 separate codes**
- Representing over **10 million lines of code**
- Many supporting large user communities
- Covering broad range of mission-critical science and engineering domains
- Mostly all MPI or MPI+OpenMP on CPUs at beginning of ECP
- Each project defines a domain-specific challenge problem for final benchmark
- Applications are evaluated in one of two categories
  - Performance – achieve a **50x** performance increase
  - Capability – utilize new architectures for expanded science and engineering

ECP APPLICATION RESULTS ARE EXCEEDING EXPECTATIONS!

9 out of 11 KPP-1 projects have already surpassed ambitious 50x performance target

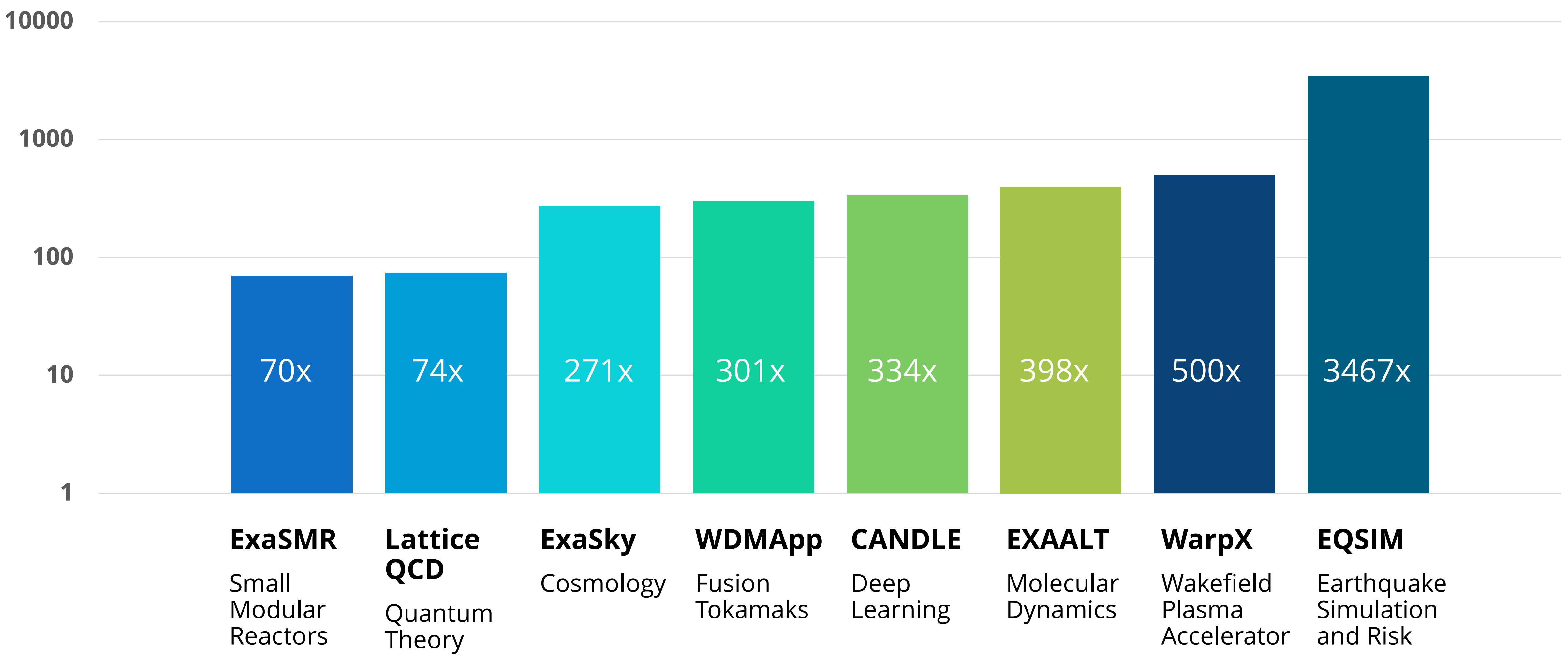


Co-design played a critical role



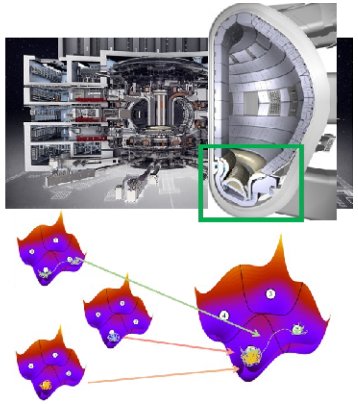
7 out of 10 KPP-2 projects have already completed exascale capability demonstrations

APPLICATION CODE IMPROVEMENTS

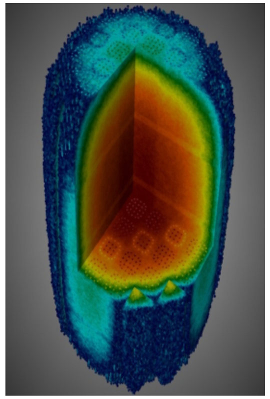


100X DEMONSTRATED: ECP-SPONSORED APPLICATION FOMS

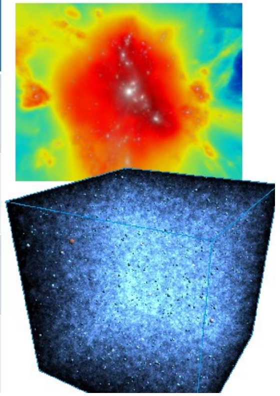
Project/PI	EXAALT: Molecular Dynamics Danny Perez
Challenge Problem	Damaged surface of Tungsten in conditions relevant to plasma facing materials in fusion reactors • 100,000 atoms • T=1200K
FOM Speedup	398.5
Nodes Used	7000
ST/CD Tools	Used in KPP Demo: **Kokkos, CoPa**



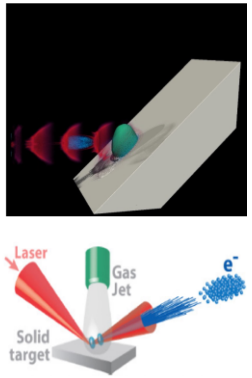
Project/PI	ExaSMR: Small Modular Reactors Steve Hamilton
Challenge Problem	NuScale-style Small Module Reactor (SMR) with depleted fuel and natural circulation • 213,860 Monte Carlo tally cells/6 reactions • 5.12×10^{12} particle histories/cycle, 40 cycles • 1098×10^{6} CFD spatial elements • 376×10^{9} CFD degrees of freedom • 1500 CFD timesteps
FOM Speedup	70
Nodes Used	6400
ST/CD Tools	Used in KPP Demo: **CEED** Additional: Trilinos



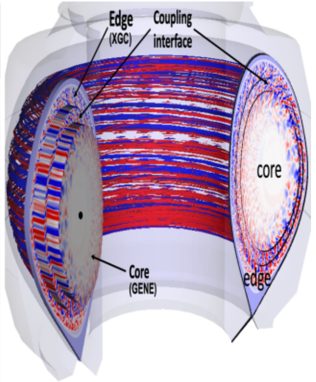
Project/PI	ExaSky: Cosmology Salman Habib
Challenge Problem	Two large cosmology simulations • gravity-only • hydrodynamics
FOM Speedup	271.65
Nodes Used	8192
ST/CD Tools	Used in KPP demo: **none** Additional: CoPa, VTK-m, CINEMA, HDF5.0



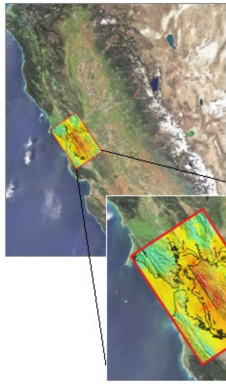
Project/PI	WarpX: Plasma Wakefield Accelerators Jean-Luc Vay
Challenge Problem	Wakefield plasma accelerator with a 1PW laser drive • 6.9×10^{12} grid cells • 14×10^{12} macroparticles • 1000 timesteps/1 stage
FOM Speedup	500
Nodes Used	8576
ST/CD Tools	Used in KPP Demo: **AMReX**, **libEnsemble** Additional: ADIOS, HDF5, VTK-m, ALPINE



Project/PI	WDMApp: Fusion Tokamaks Amitava Bhatacharjee
Challenge Problem	Gyrokinetic simulation of the full ITER plasma to predict the height and width of the edge pedestal
FOM Speedup	150
Nodes Used	6156
ST/CD Tools	Used in KPP Demo: **CODAR**, **CoPA**, **PETSc**, **ADIOS** Additional: VTK-m



Project/PI	EQSIM: Earthquake Modeling and Risk Dave McCallen
Challenge Problem	Impacts of Mag 7 rupture on the Hayward Fault on the bay area.
FOM Speedup	3467
Nodes Used	5088
ST/CD Tools	Used in KPP Demo: **RAJA, HDF5**



EXASCALE COMPUTING WILL HAVE FAR REACHING CONSEQUENCES

Exascale is real and is delivering remarkable science. It took widespread collaboration throughout the HPC community, and we've delivered – but it's only just the start.

Effectively utilizing GPU-CPU architectures has changed the landscape for the better.

Energy consumption per flop can be decreased at all levels – from desktops to petascale systems to extremely large systems. That's a win for everyone.

The **performance improvements** possible in science work-rate are tremendous…50x, 100x, 400x and more. Imagine how this can impact industry, science, national security, etc. as those gains become more widespread.

Applications and software already deployed have large user communities with **tens of thousands of users.**

The software needed to leverage these capabilities in a performance portable way − E4S − is available to anyone and offers tremendous value to systems of all sizes and in the cloud.

KEY INGREDIENTS OF AN ECP APPLICATION DEVELOPMENT PROJECT

Science goal

Algorithmic innovation

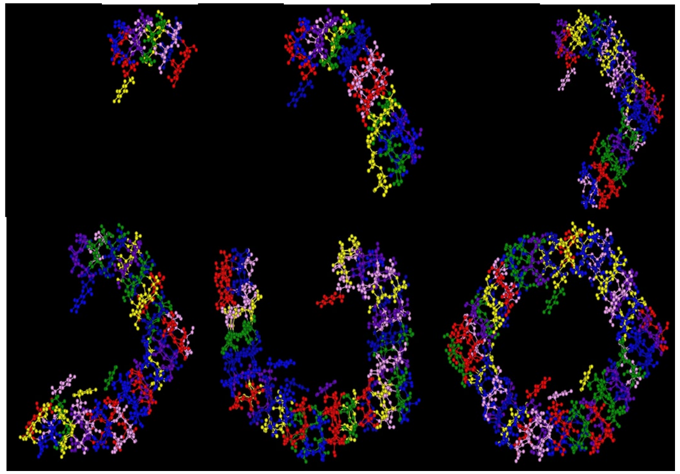
Porting

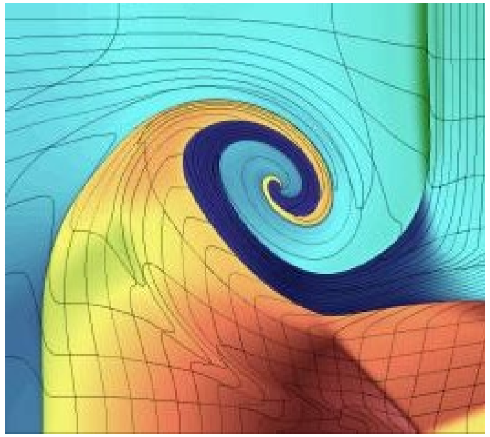
Integration

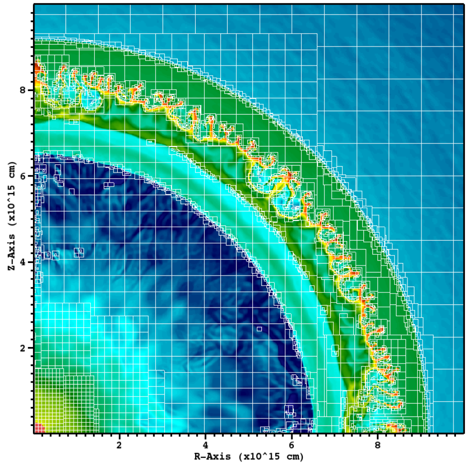
ALGORITHMIC INNOVATION

Domain-driven adaptations critical for making efficient use of exascale systems

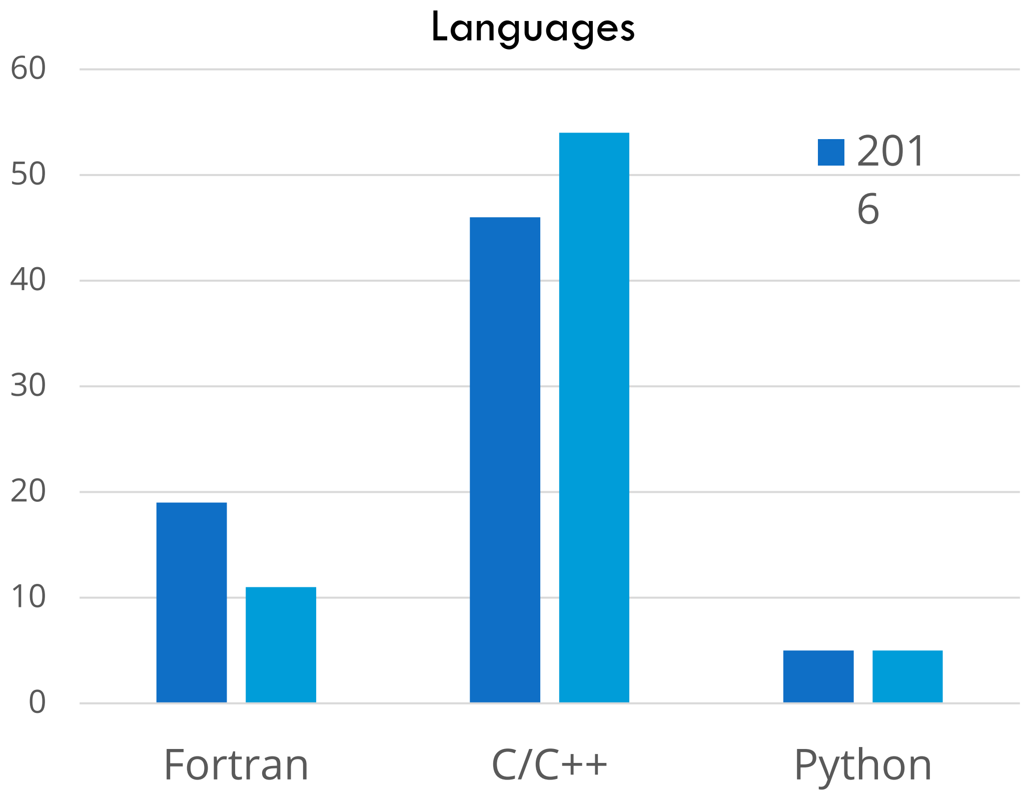
- Inherent strong scaling challenges on GPU-based systems
  - Ensembles vs. time averaging
  - Fluid dynamics, seismology, molecular dynamics, time-stepping
- Increase dimensions of (fine-grained) parallelism to feed GPUs
  - Ray tracing, Markov Chain Monte Carlo, fragmentation methods
- Localized physics models to maximize "free flops”
  - MMF, electron subcycling, enhanced subgrid models, high-order discretizations
- Alternatives to sparse linear systems
  - Higher order methods, Monte Carlo
- Reduced branching
  - Event-based models

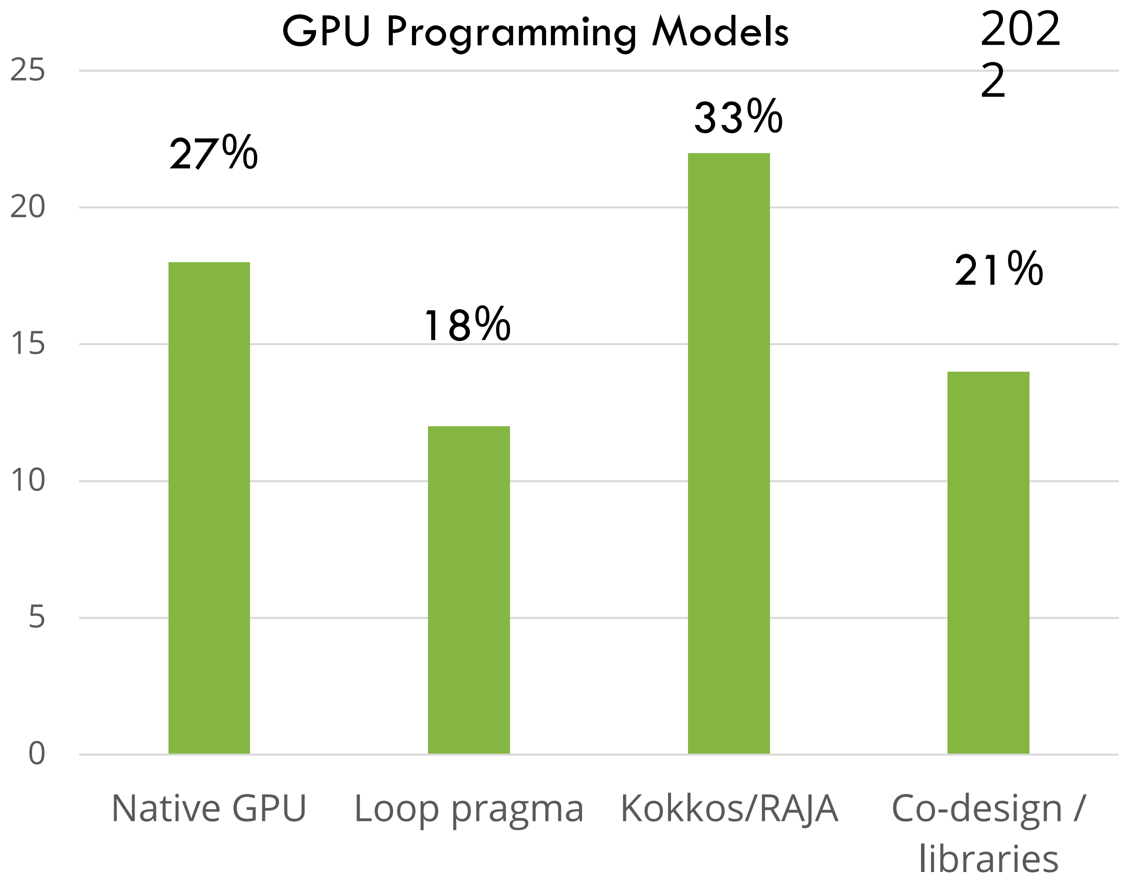






DISTRIBUTION OF ECP PROGRAMMING MODELS





Many ECP applications started out using native GPU and loop pragma models before moving to C++ abstractions and co-design libraries

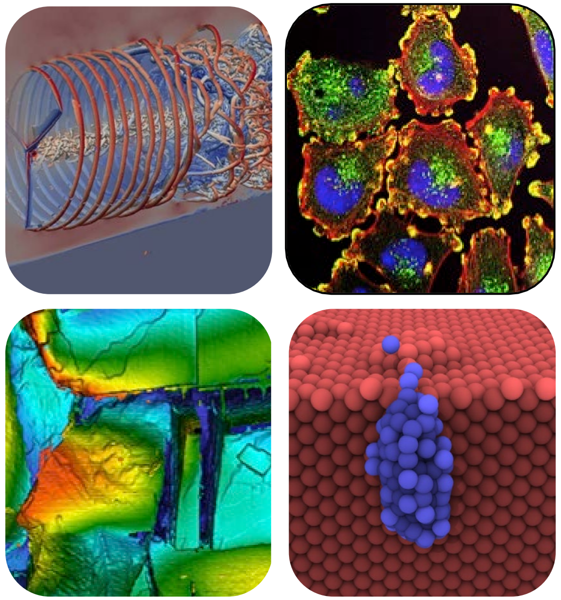
Evans TM, Siegel A, Draeger EW, *et al.* *"A survey of software implementations used by application codes in the Exascale Computing Project."* The International Journal of High Performance Computing Applications. 2022;36(1):5-12.

EXASCALE APPLICATIONS AND SOFTWARE TECHNOLOGY INTEGRATIONS

Programming abstractions

Memory / heterogeneous data management

In situ visualization and analysis



Solvers / preconditioners

Numerical infrastructure / discretizations

Build and deployment

Performance analysis and monitoring

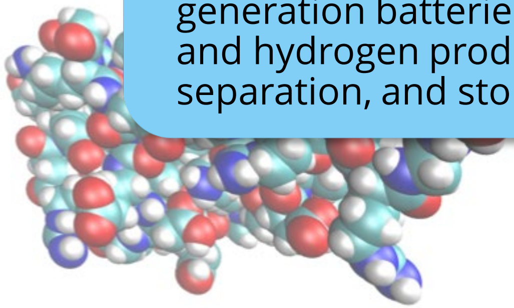
KEY LESSONS IN ACHIEVING EXASCALE IN ECP

1. Code porting exercise
2. “Over the fence” producer-consumer enabling software
3. Do-it-yourself application groups with minimal external dependencies

1. Holistic rethinking and restructuring of implementation, algorithm, and models
2. Software/application tight co-design loop:
   Develop → Deploy → integrate → Accrete
3. Leverage community tools and expertise broadly – E4S, LCFs, Centers of Excellence, teams with highly diverse expertise

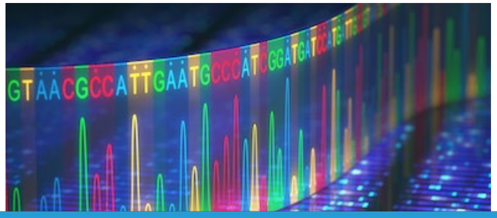
ECP APPLICATIONS ARE ENABLING NEW SCIENCE (1 OF 4)

NWChemEx



Enable predictive investigations of new chemicals and materials for solar energy conversion, next-generation batteries, CO_2 capture, and hydrogen production, separation, and storage

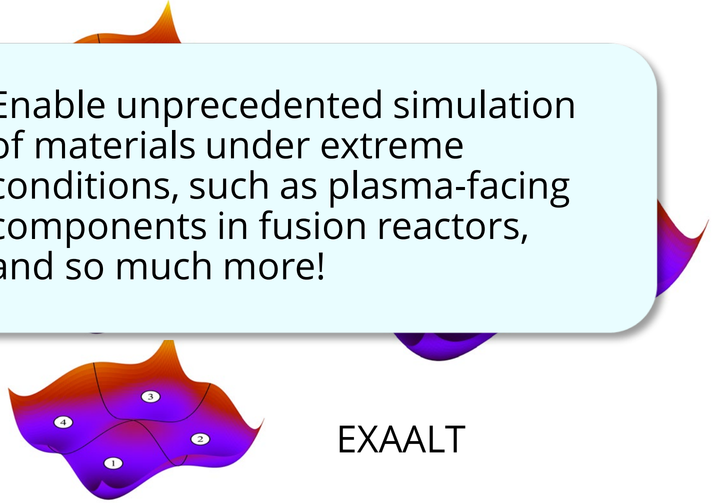
ExaBiome



Promote environmental remediation, improve food production, aid medical research and help understand impact of climate change, including wildfires and algae blooms

Reveal the amount and distribution of heavy elements we can expect to find on planet Earth and deepen our fundamental knowledge of gravity and matter at extreme densities

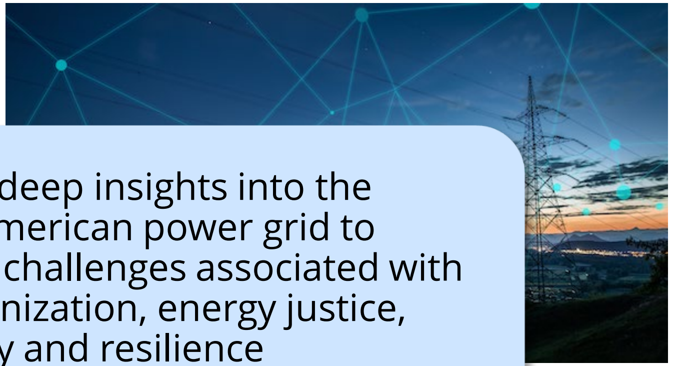
ExaStar



Enable unprecedented simulation of materials under extreme conditions, such as plasma-facing components in fusion reactors, and so much more!

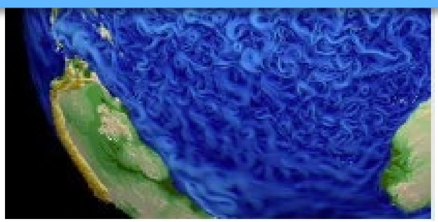
EXAALT

ExaSGD



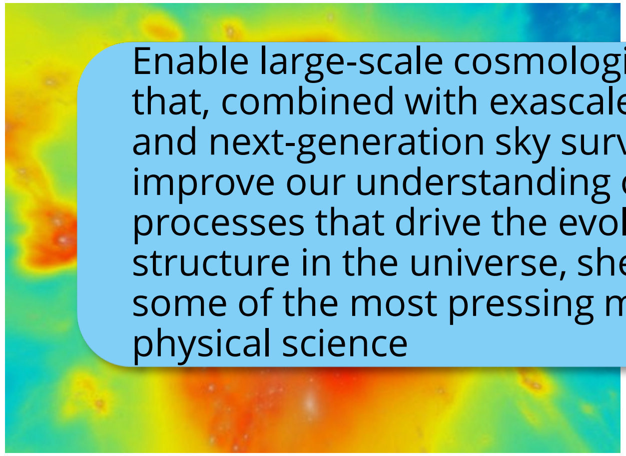
Provide deep insights into the North American power grid to address challenges associated with decarbonization, energy justice, reliability and resilience

Understand the regional impacts of climate change on the water cycle informs policy makers in multiple sectors of the US and global economies including agriculture and energy production



E3SM-MMF

ECP APPLICATIONS ARE ENABLING NEW SCIENCE (2 OF 4)

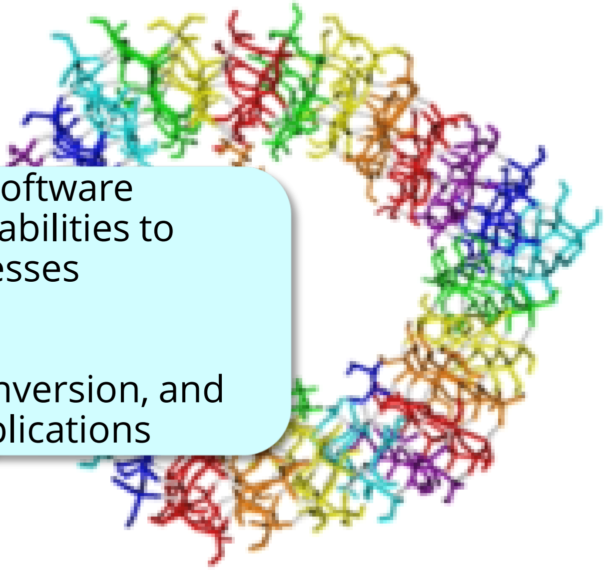


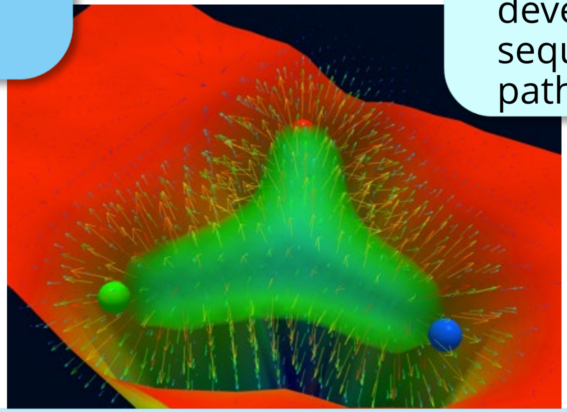
Enable large-scale cosmological simulations that, combined with exascale computing and next-generation sky surveys, will improve our understanding of the physical processes that drive the evolution of structure in the universe, shedding light on some of the most pressing mysteries in physical science

ExaSky

GAMESS

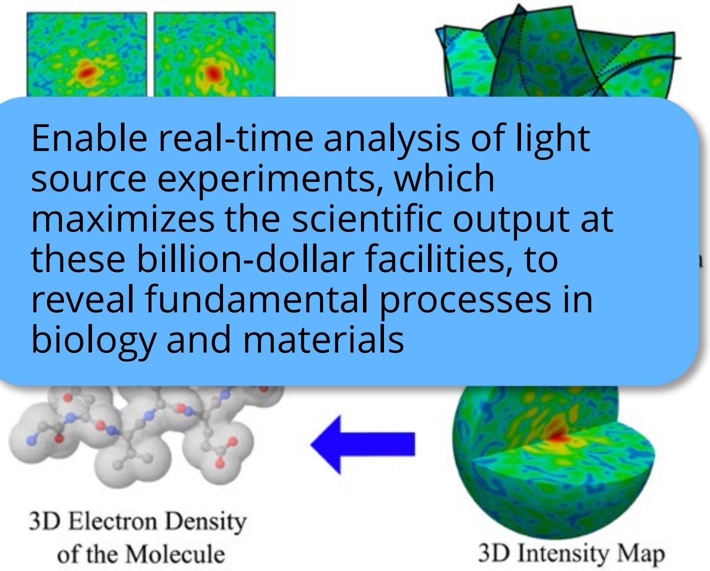
New algorithms and software leapfrog previous capabilities to model chemical processes important for catalyst development, carbon sequestration and conversion, and pathogen sensing applications





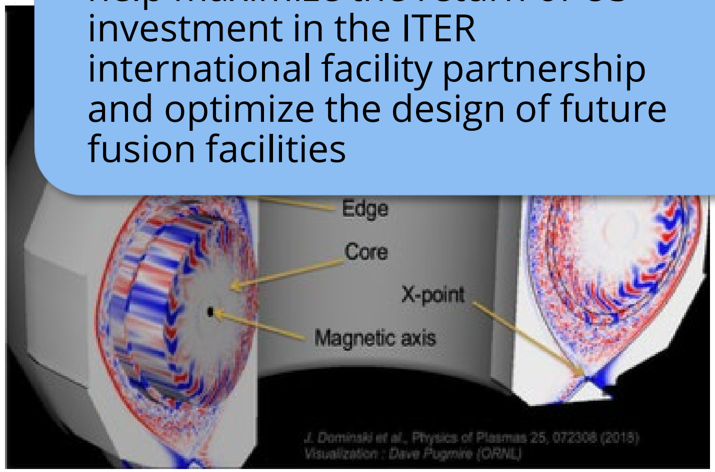
Enables detailed physics investigations at the subatomic level advancing the understanding of interactions of matter including matter/antimatter asymmetries, the Higgs Boson, and neutrino oscillations

LatticeQCD



Enable real-time analysis of light source experiments, which maximizes the scientific output at these billion-dollar facilities, to reveal fundamental processes in biology and materials

ExaFEL

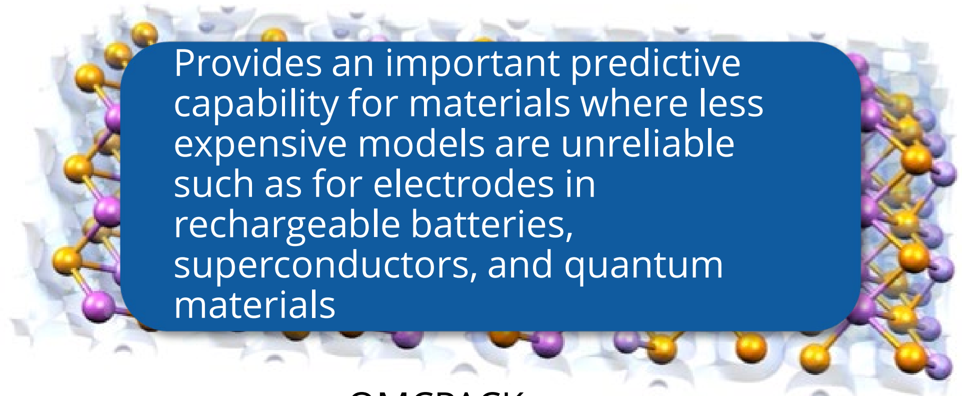


Exascale fusion simulations will help maximize the return of US investment in the ITER international facility partnership and optimize the design of future fusion facilities

WDMApp

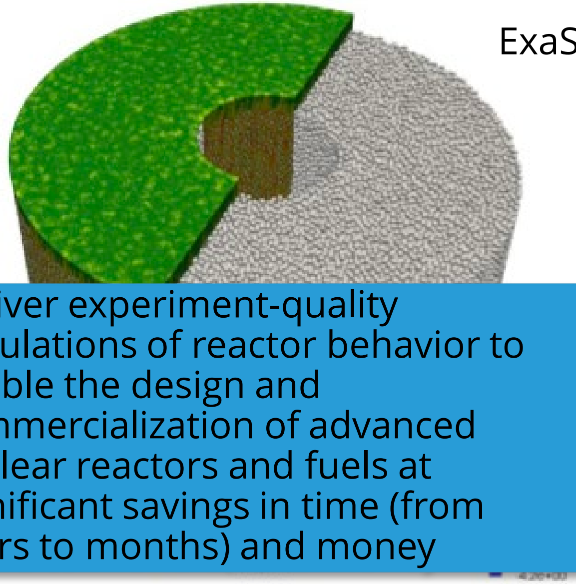
ECP APPLICATIONS ARE ENABLING NEW SCIENCE (3 OF 4)

Provides an important predictive capability for materials where less expensive models are unreliable such as for electrodes in rechargeable batteries, superconductors, and quantum materials

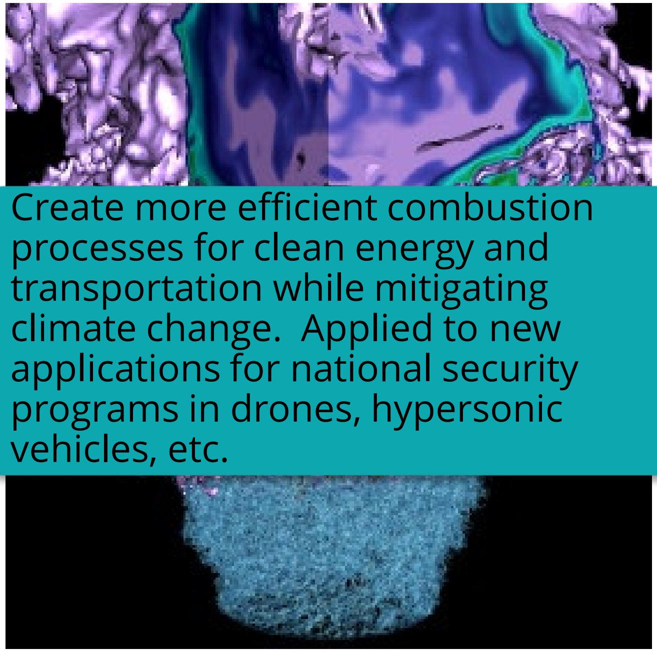


QMCPACK

ExaSMR



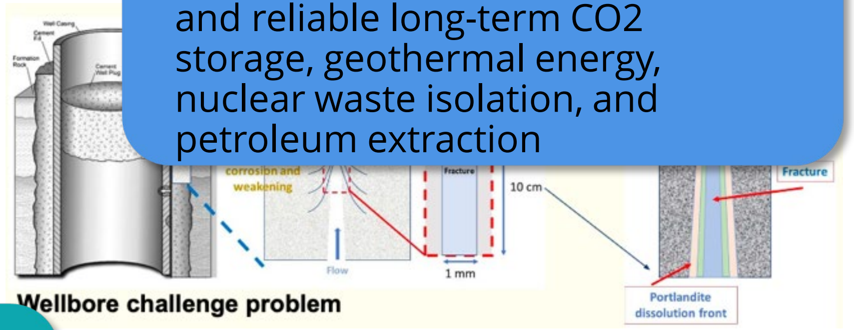
Deliver experiment-quality simulations of reactor behavior to enable the design and commercialization of advanced nuclear reactors and fuels at significant savings in time (from years to months) and money



Create more efficient combustion processes for clean energy and transportation while mitigating climate change. Applied to new applications for national security programs in drones, hypersonic vehicles, etc.

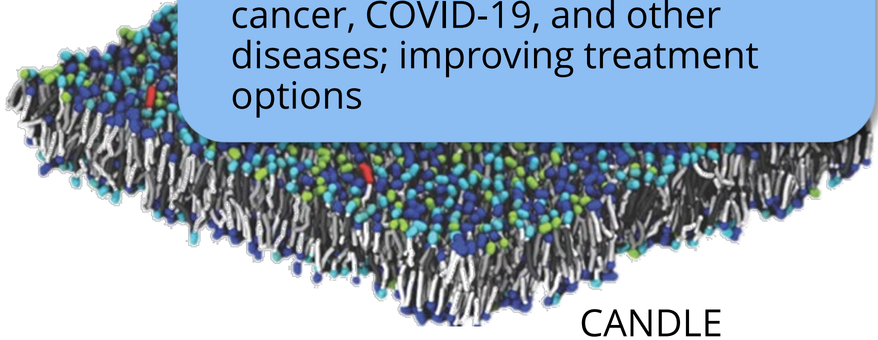
Combustion-PELE

Understanding subsurface processes, especially fractures, is critical for the development of safe and reliable long-term CO2 storage, geothermal energy, nuclear waste isolation, and petroleum extraction



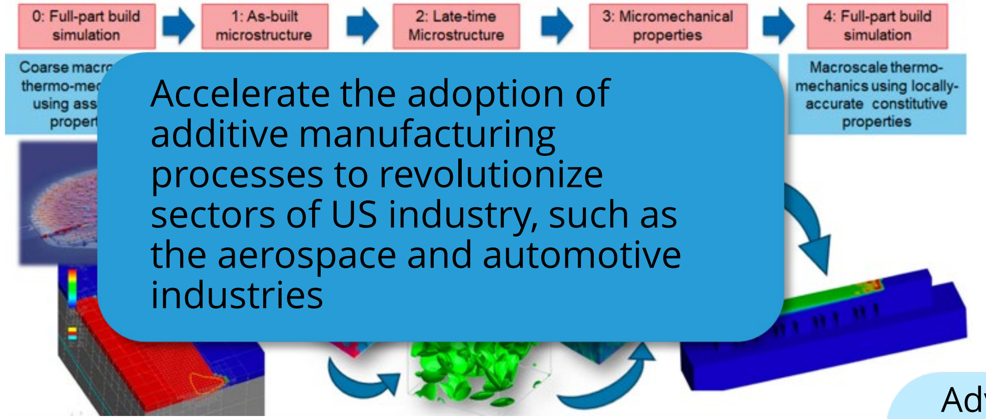
Subsurface

Applying high performance computing and machine learning has led to new insights into cancer, COVID-19, and other diseases; improving treatment options



CANDLE

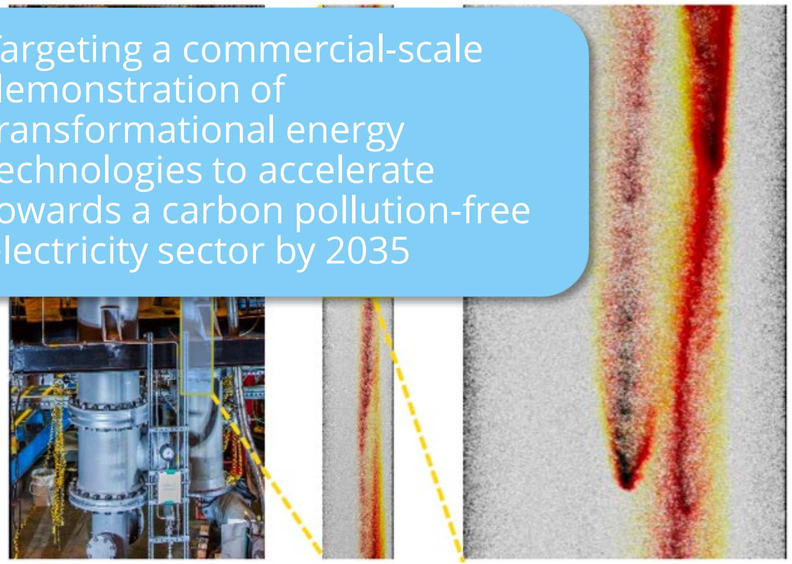
ECP APPLICATIONS ARE ENABLING NEW SCIENCE (4 OF 4)



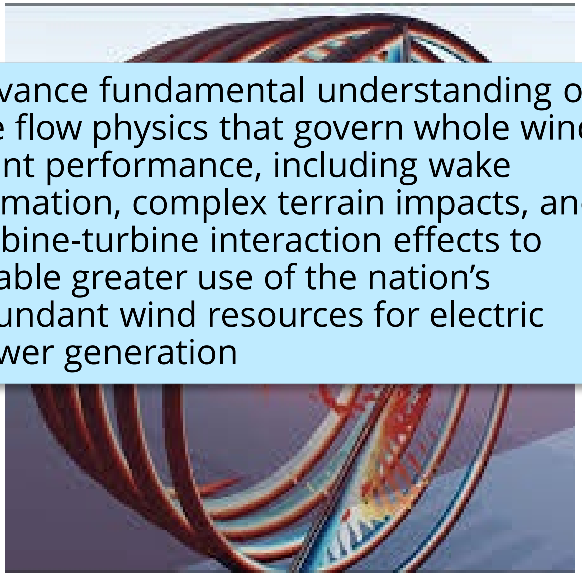
Accelerate the adoption of additive manufacturing processes to revolutionize sectors of US industry, such as the aerospace and automotive industries

ExaAM

MFIX-Exa

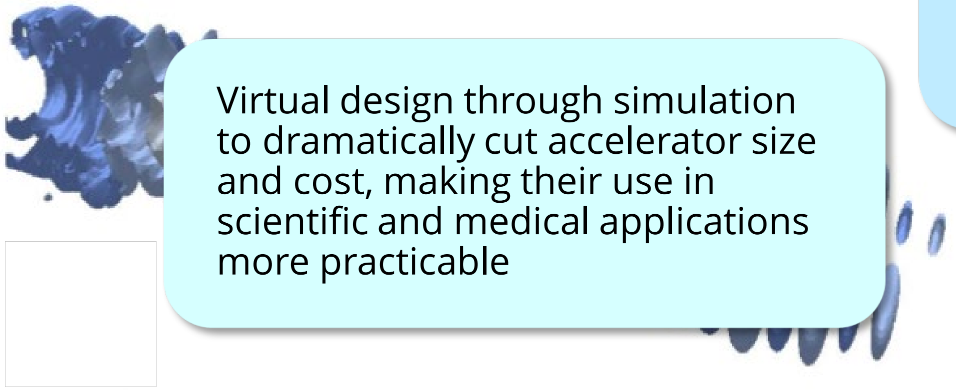


Targeting a commercial-scale demonstration of transformational energy technologies to accelerate towards a carbon pollution-free electricity sector by 2035



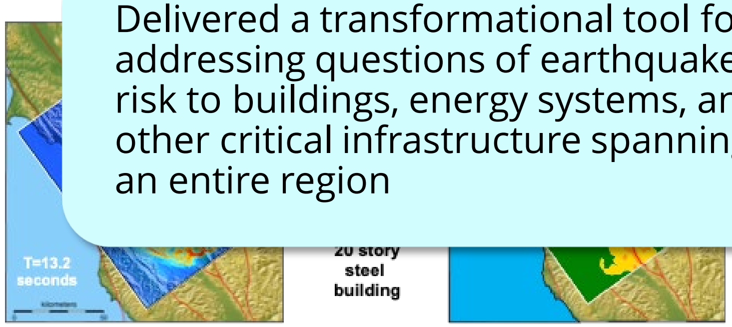
Advance fundamental understanding of the flow physics that govern whole wind plant performance, including wake formation, complex terrain impacts, and turbine-turbine interaction effects to enable greater use of the nation's abundant wind resources for electric power generation

ExaWind



Virtual design through simulation to dramatically cut accelerator size and cost, making their use in scientific and medical applications more practicable

WarpX



Delivered a transformational tool for addressing questions of earthquake risk to buildings, energy systems, and other critical infrastructure spanning an entire region

EQSIM

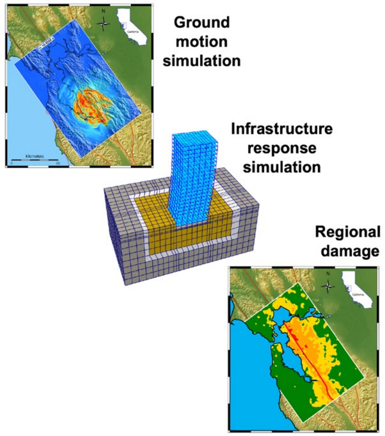
ExaSMR developed first-of-a-kind simulations of advanced nuclear reactors such as small modular reactors

- **Objective**: Help DOE meet its goal of an operational small modular reactor (SMR) in 10 years, a key part of the Department's goal to develop safe, clean, and affordable nuclear power options
- Accomplishments:
  - First fully coupled, fully resolved simulation of nuclear reactor core with coupled Monte Carlo neutron transport and CFD; achieved nearly 100X overall performance improvements in the science work rate
  - Allows study of evolution of nuclear fuel for the first time
  - **Gordon Bell Finalist – SC23!**
- Deliver experiment-quality simulations of reactor behavior to enable the design and commercialization of advanced nuclear reactors and fuels at significant savings in time (from years to months) and money

Small modular reactor. Image reproduced with permission of NuScale

EQSIM developed a framework that advances the ability to predict and use site-specific ground motions in earthquake risk assessment

- **Objective**: Create high-performance simulation tools that establish a coupled assessment of earthquake hazard (ground motion) and earthquake risk (infrastructure demands) at regional scale
- Accomplishments:
  - Increased the resolution of earthquake ground motion frequencies from 2Hz to 10Hz
  - Simulations now include strong coupling between regional geophysics and local soil/building models
  - Achieved more than 1000X improvement in computational performance comparec to all previous San Francisco Bay Area simulations



- This is a transformational tool for addressing questions of earthquake risk to buildings, energy systems, and other critical infrastructure spanning an entire region

LESSONS ON APPLICATION-SOFTWARE INTEGRATION FROM ECP

It's not optional any more

Many applications have historically tried to minimize dependencies to maintain control. For exascale, staying ahead of fast-moving hardware trends requires delegating key math and CS capabilities to others.

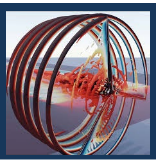
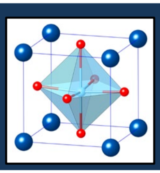
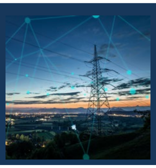
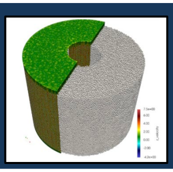
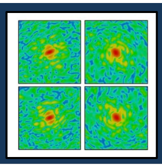
"Off the shelf" is rarely sufficient

We found that deep engagement between applications and math/CS libraries was often needed to fully understand problem space and create coupled implementations that perform well on exascale hardware. Standalone APIs/libraries developed in isolation can miss out on performance.

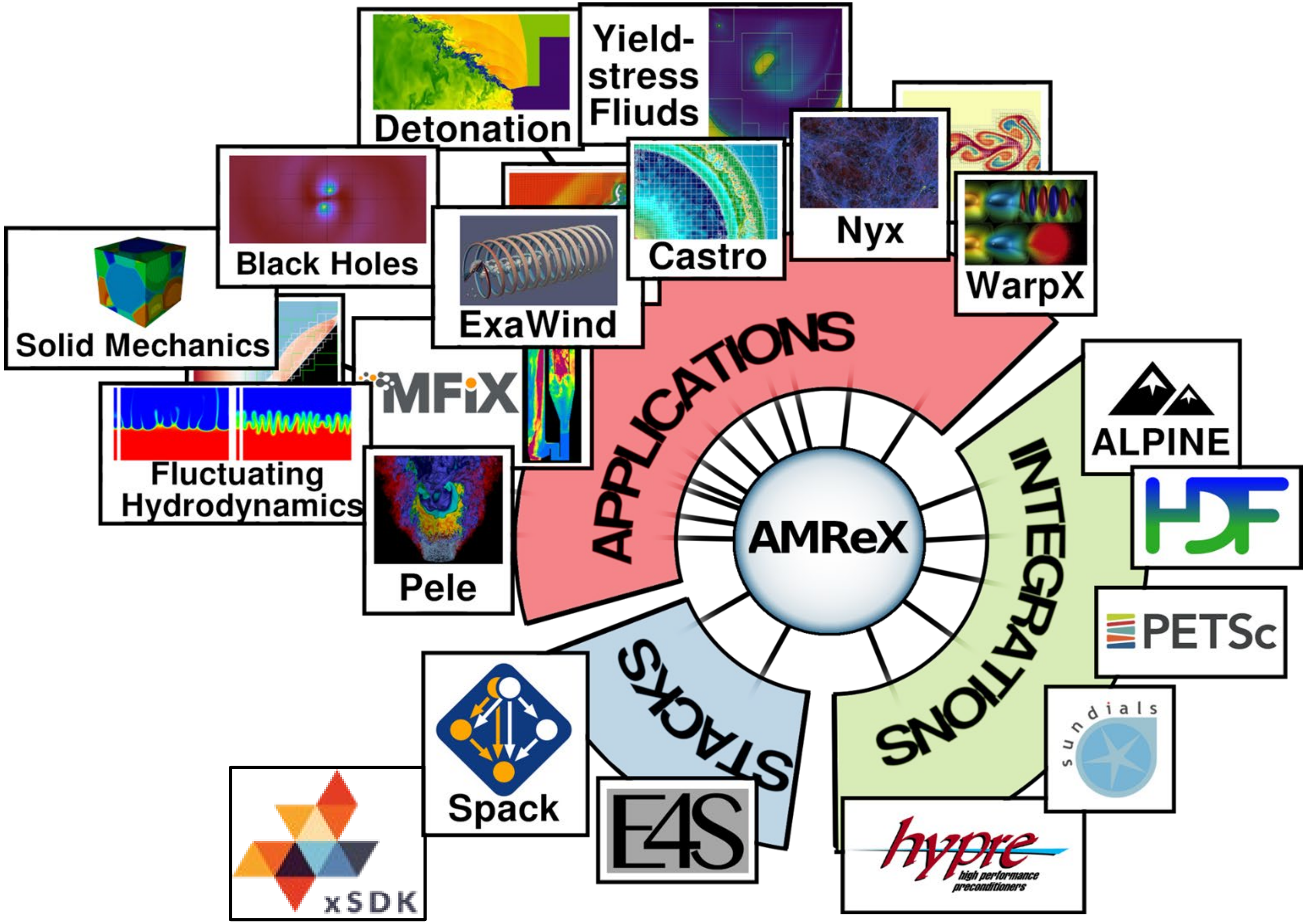
Having multiple competing implementations is ideal

There is rarely any ground truth for performance of complex problems, having different packages (ideally with compatible APIs) drives improvements.

SMALL SAMPLING OF ECP INTEGRATION EXAMPLES

ExaWind: advanced wind farm modeling		Sparse linear solvers in hypre and Trilinos optimized for strong scaling and GPU performance
QMCPACK: quantum Monte Carlo for materials		Batched dense linear algebra kernels significantly improved GPU performance
ExaSGD: power grid optimization		Optimize sparse indefinite solvers developed and optimized for large-scale grid problems
ExaSMR: small modular reactor modeling		Custom discretization designed and tuned for specific reactor assembly
WarpX: plasma wakefield accelerator design		Adaptive mesh data structures and solvers highly optimized for GPU performance
ExaFEL: real-time light source analysis and reconstruction		Non-uniform FFTs designed to minimize data motion

AMReX IS PART OF A LARGER SOFTWARE ECOSYSTEM



Software Integrations:

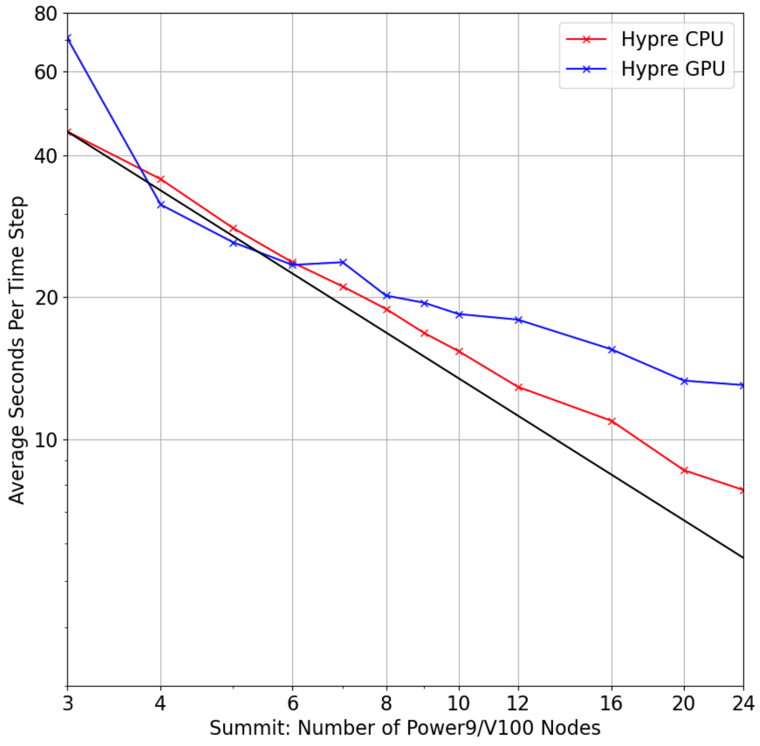
- SUNDIALS – Chemical reactions, time integrators
- Hypre, PETSc – Linear solvers on mesh data
- In-situ: Ascent, Sensei
- Offline visualization: VisIt, Paraview, yt
- IO: HDF5, ADIOS

Software Stacks:

- Spack
  - Smoke test – CUDA, AMD HIP
- xSDK
- E4S

EXAWIND – STRONG SCALING SPARSE LINEAR SOLVERS

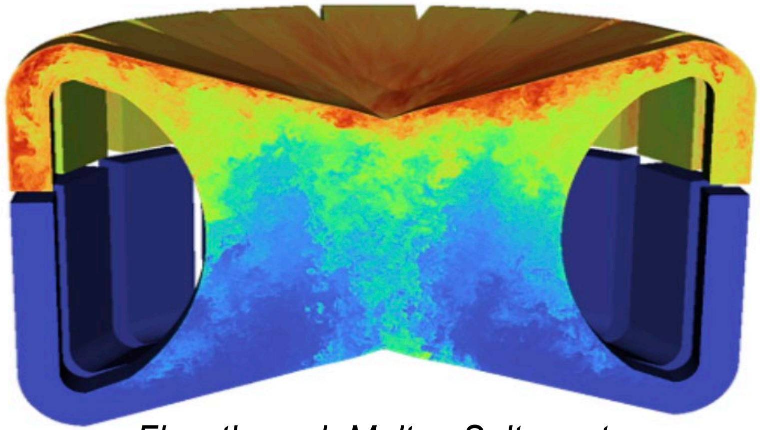
- **ExaWind**'s challenge based on strong-scaling perfor
- **Hypre**: novel single-reduction GMRES algorithms inc GPUs
  - >2x faster than baseline Hypre GPU implementa
  - Flat communication time and retains numerical
  - Adopted by Trilinos PEEKS, PETSc (ANL), CEED pr
- Applied low-synch ideas to create low-rank approxin
  - Extremely fast set-up and solve time (20x faster)
- Low-synch GMRES multi-MPI multi-GPU
  - Extended strong scaling roll-off by 2x for ExaWin >2x speedup.
  - Low-synchronization Gram-Schmidt and GMRES Swirydowicz, J. Langou, S. Ananthan, U. Yang, S. Thomas



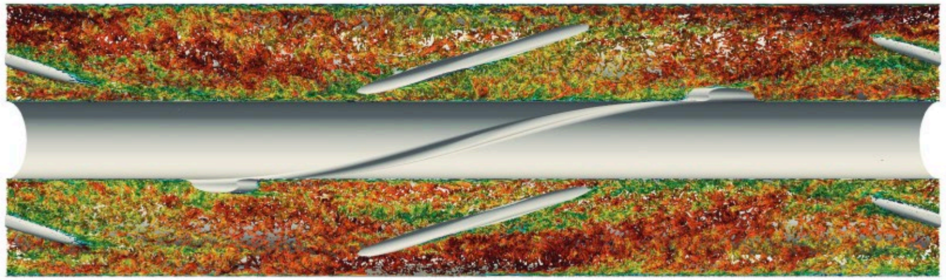
DEVELOPING NEW ADVANCED REACTOR MODELS

ExaSMR project advanced predictive modeling and simulation for nuclear reactors by coupling Monte Carlo neutron transport with computational fluid dynamics

- Advanced reactor designs lack extensive experimental and operational data of current nuclear fleet
- High-fidelity simulations can complement and reduce need for costly physical experiments



Flow through Molten Salt reactor



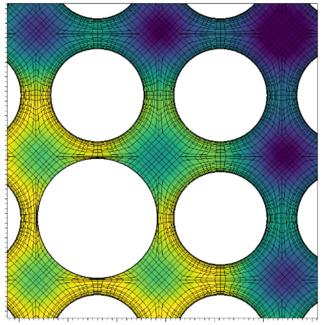
Flow structures in a sodium fast reactor

CEED codesign project delivered efficient tools for discretization and numerical solution of PDEs

- Enhanced performance portability thought automated performance tuning of computational routines
- Improved time to solution through mixed precision algebraic multigrid

ExaSMR+CEED enabled nuclear simulations of unprecedented size and complexity

- Already being used to benchmark and improve predictions of advanced nuclear systems
- NekRS code jointly developed by CEED and ExaSMR
- Multiphysics reactor simulation of small modular reactor on Frontier was finalist for 2023 Gordon Bell Prize



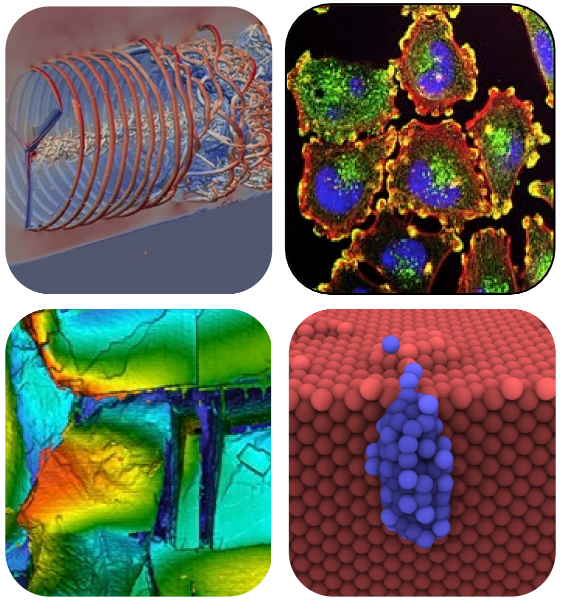
Coolant temperature distribution near guide tube in SMR core

EXASCALE APPLICATIONS AND SOFTWARE TECHNOLOGY INTEGRATIONS

Programming abstractions

Memory / heterogeneous data management

In situ visualization and analysis

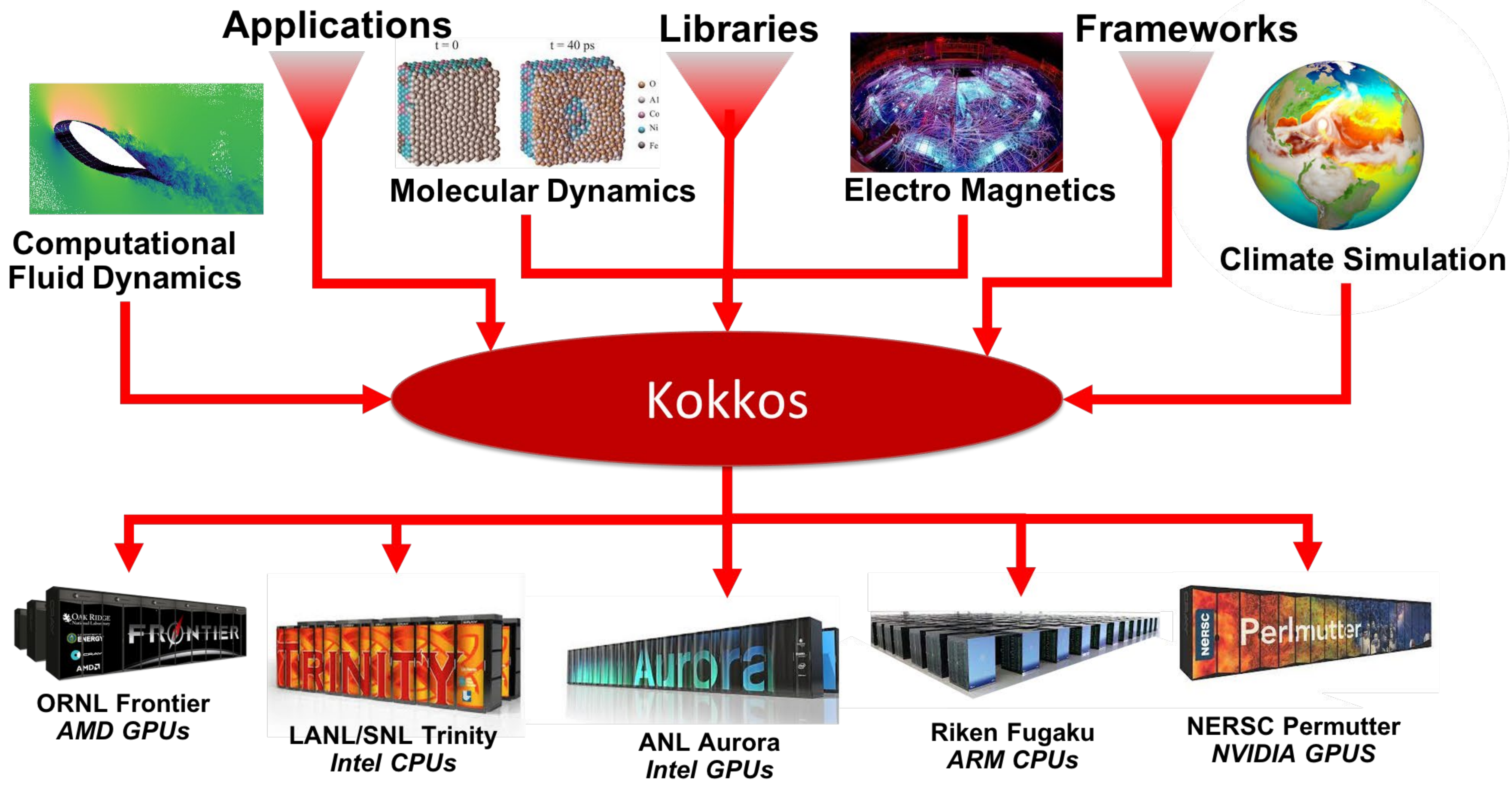


Solvers / preconditioners

Numerical infrastructure / discretizations

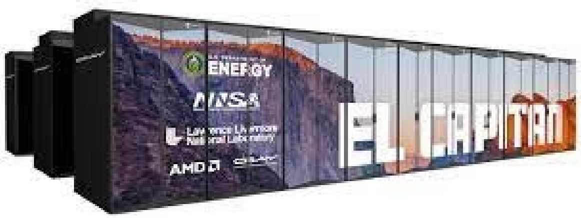
Build and deployment

Performance analysis and monitoring


Applications
Libraries
Frameworks
t = 0
t = 40 ps
O
Al
Co
Ni
Fe
Molecular Dynamics
Electro Magnetics
Computational Fluid Dynamics
Climate Simulation
Kokkos
ORNL Frontier
AMD GPUs
LANL/SNL Trinity
Intel CPUs
ANL Aurora
Intel GPUs
Riken Fugaku
ARM CPUs
NERSC Permutter
NVIDIA GPUS

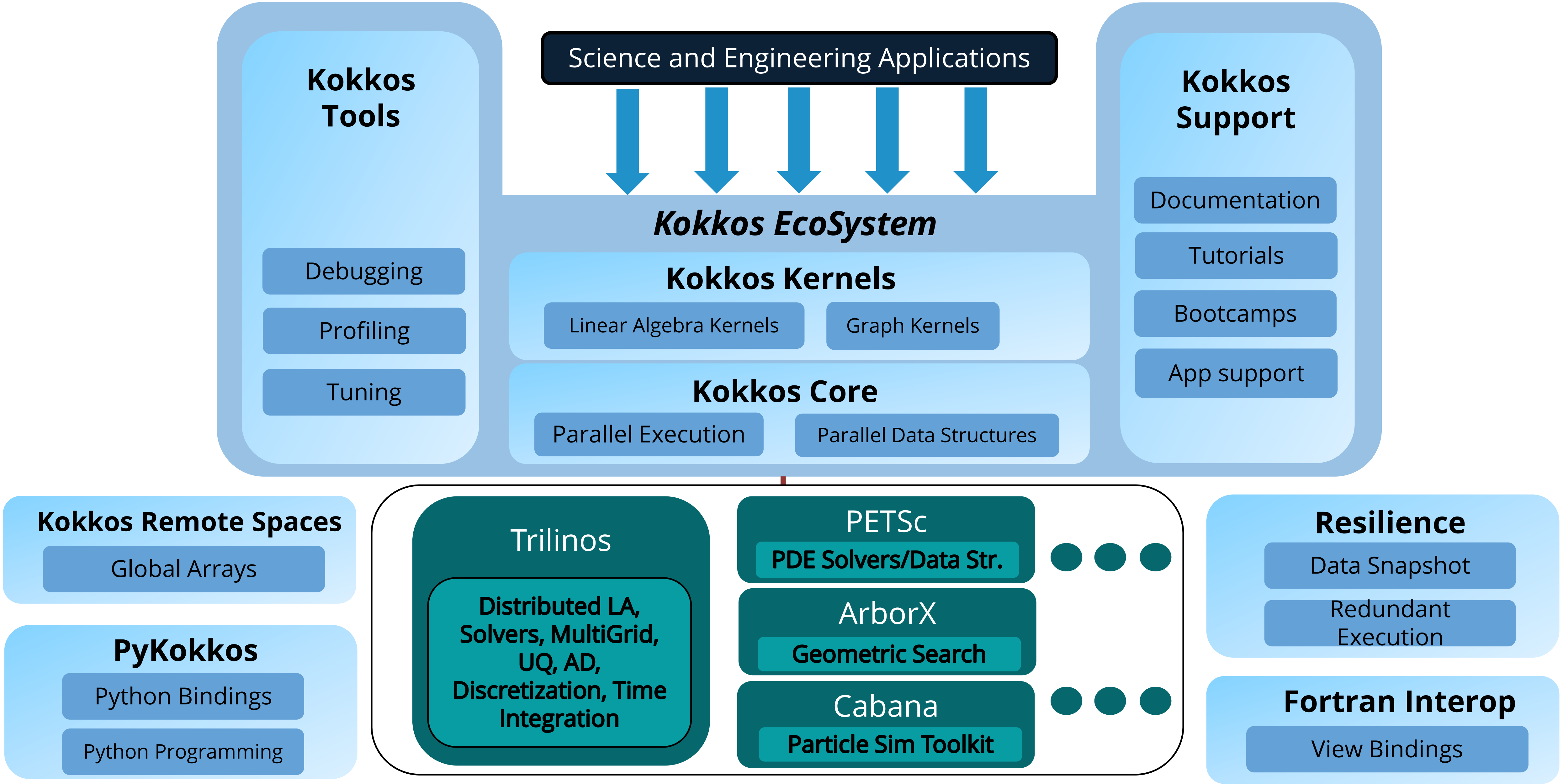
WHAT IS KOKKOS?

- A C++ Programming Model for Performance Portability
  - Implemented as a template library on top of CUDA, OpenMP, HPX, …
  - Aims to be descriptive not prescriptive
  - Aligns with developments in the C++ standard
  - Replaces usage of CUDA, OpenMP, HIP, etc.
- Expanding solution for common needs of modern science/engineering codes
  - Math libraries based on Kokkos
  - Tools which enable insight into Kokkos
- It is Open Source
  - Maintained and developed at https://github.com/kokkos
- It has many users at wide range of institutions.

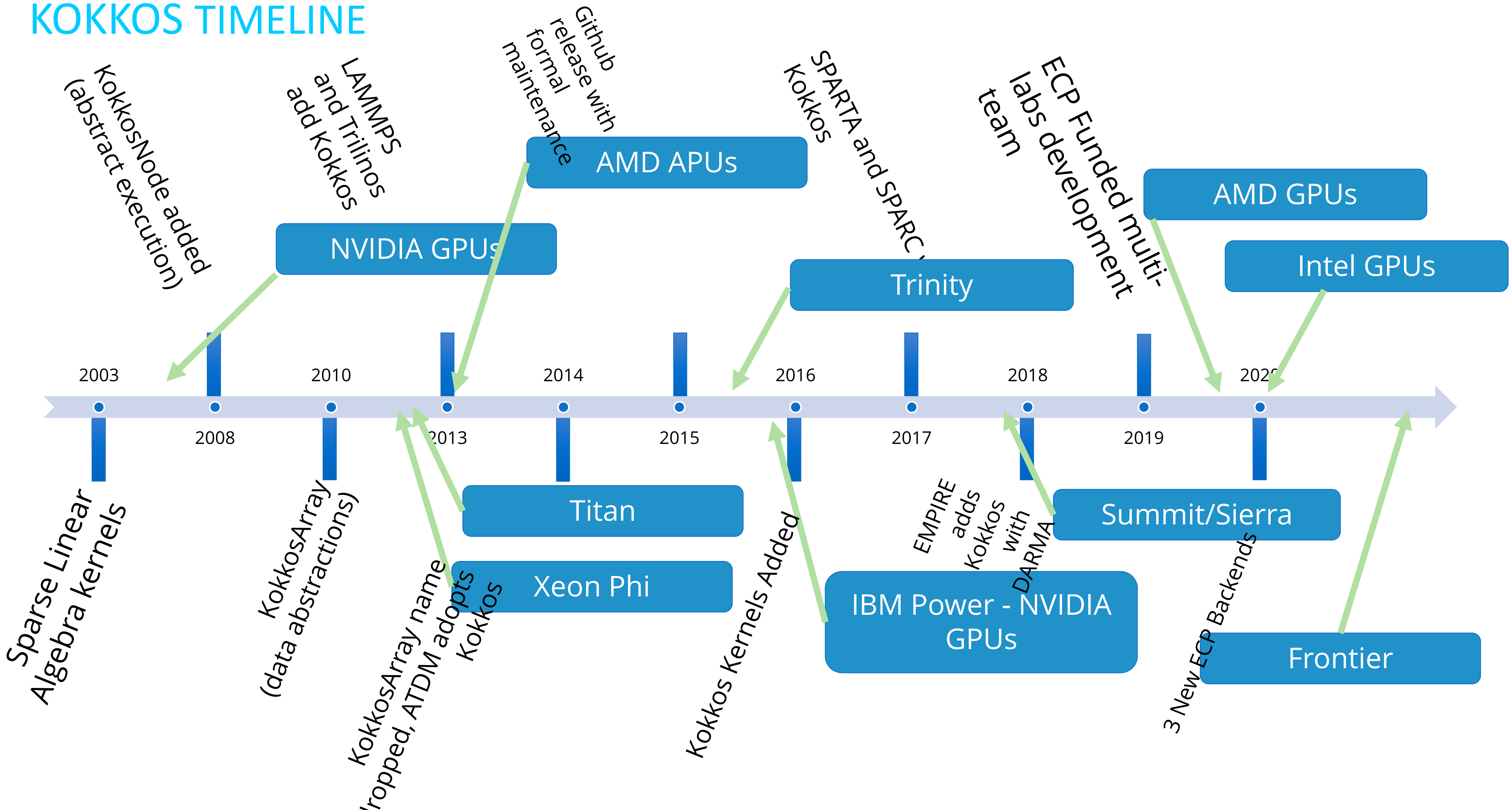


Kokkos is NOT just for GPUs!

THE KOKKOS ECOSYSTEM - TODAY



KOKKOS TIMELINE

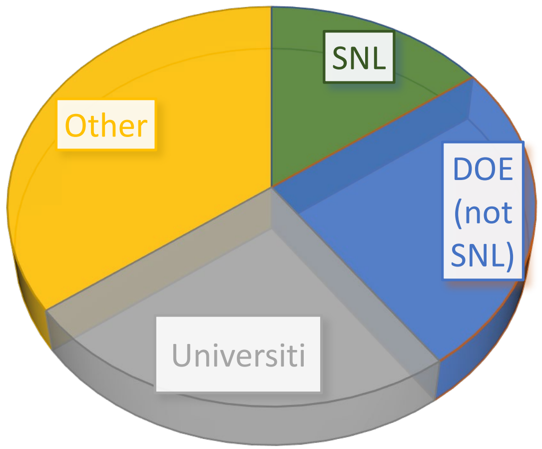


KOKKOS COMMUNITY

Kokkos Slack

https://kokkosteam.slack.com

- >1500 Registered Users
- >150 Institutions
    - Including 34 European



Kokkos Developers

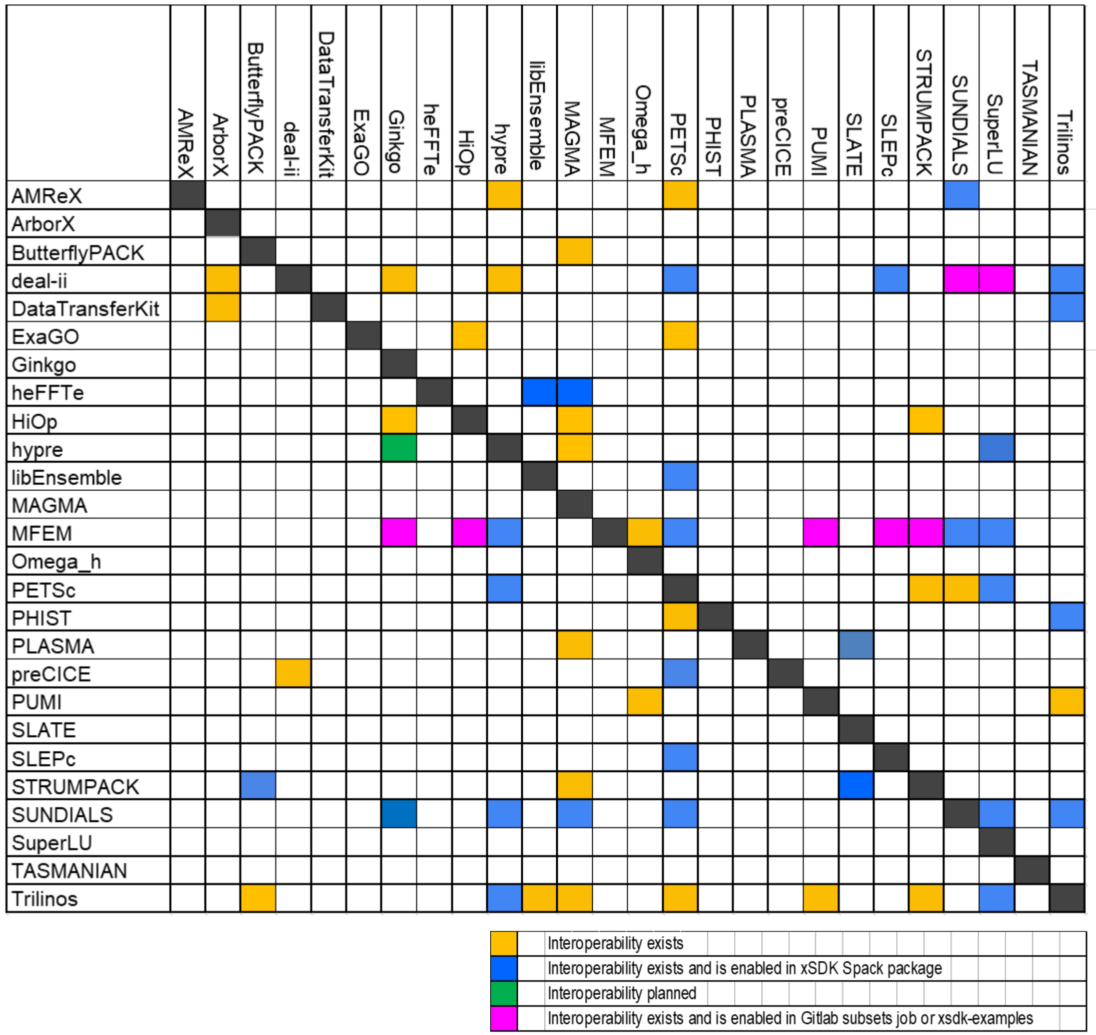
Sandia National Laboratories
Oak Ridge National Laboratory
Argonne National Laboratory
Los Alamos National Laboratory EST. 1943
Berkeley Lab
CSCS

Applications and Libraries

- Estimated 150-300 HPC projects using Kokkos
- On the order of three-dozen apps run science and engineering production runs with Kokkos
    - Many apps use multiple Kokkos based libraries
- Similar distribution as the Slack User

xSDK's MULTI-LAYER TESTING FRAMEWORK

xSDK Interoperability Matrix



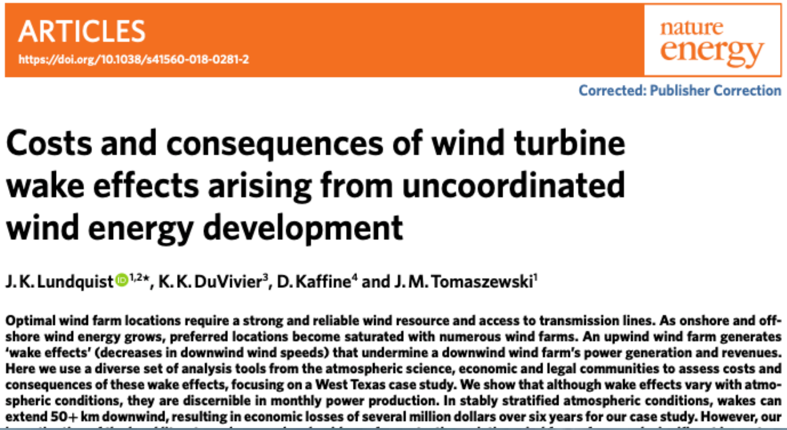
Multi-layered testing

- Testing strategies of the individual xSDK libraries
- Testing of the interfaces between libraries
- Test subsets of various interoperable packages in combination
- Test the whole xSDK (final level)

Because it is tested as a complete set and all works together, users can be confident any subset will compile and work together

THE ExaWIND PROJECT

Wind turbines produce a wake that can affect other turbines downstream of it, reducing their efficiency, increasing noise and reducing the life of the turbines. Can we design a wind farm that minimizes these effects?



ARTICLES

https://doi.org/10.1038/s41560-018-0281-2

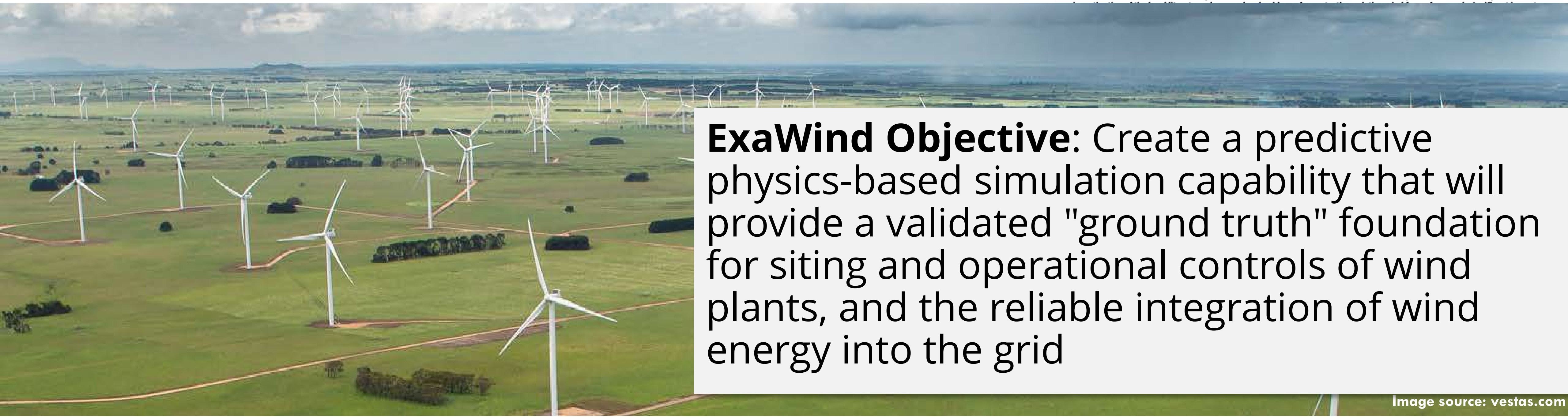
nature energy

Corrected: Publisher Correction

Costs and consequences of wind turbine wake effects arising from uncoordinated wind energy development

J. K. Lundquist[1,2]*, K. K. DuVivier[3], D. Kaffine[4] and J. M. Tomaszewski[1]

Optimal wind farm locations require a strong and reliable wind resource and access to transmission lines. As onshore and offshore wind energy grows, preferred locations become saturated with numerous wind farms. An upwind wind farm generates 'wake effects' (decreases in downwind wind speeds) that undermine a downwind wind farm's power generation and revenues. Here we use a diverse set of analysis tools from the atmospheric science, economic and legal communities to assess costs and consequences of these wake effects, focusing on a West Texas case study. We show that although wake effects vary with atmospheric conditions, they are discernible in monthly power production. In stably stratified atmospheric conditions, wakes can extend 50+ km downwind, resulting in economic losses of several million dollars over six years for our case study. However, our

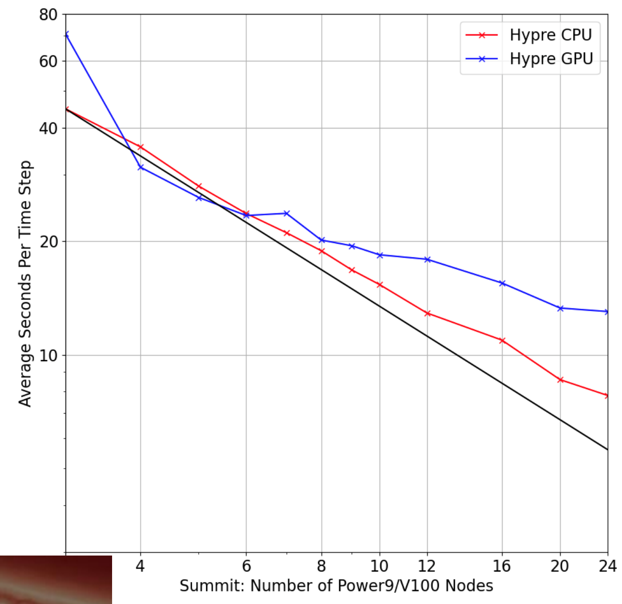


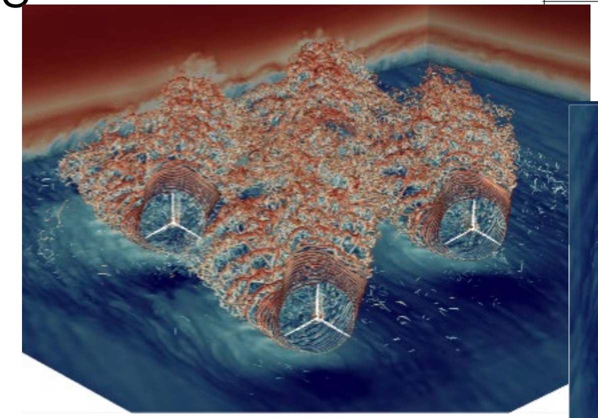
ExaWind Objective: Create a predictive physics-based simulation capability that will provide a validated "ground truth" foundation for siting and operational controls of wind plants, and the reliable integration of wind energy into the grid

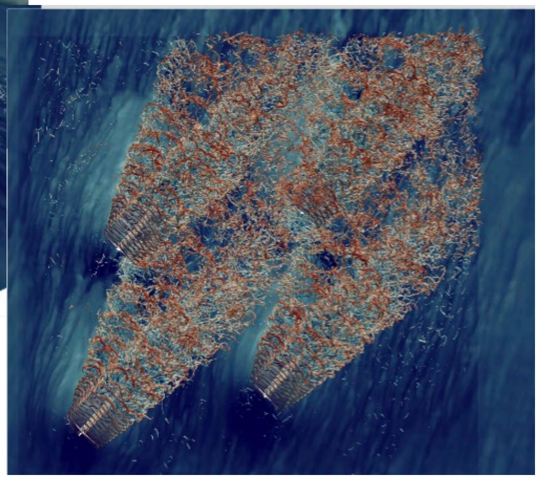
Image source: vestas.com

ExaWind REQUIRED STRONG SCALING FOR SPARSE LINEAR SOLVERS

- ***hypre***: novel single-reduction GMRES algorithms increase strong-scaling limit on GPUs
  - >2x faster than baseline hypre GPU implementation on V100 GPU
  - Flat communication time and retains numerical stability
  - Adopted by Trilinos PEEKS, PETSc (ANL), CEED project, et al.
- Applied low-synchronization ideas to create low-rank approximate AINV smoothers for AMG
  - Extremely fast set-up and solve time (20x faster) on GPU
- Low-synch GMRES multi-MPI multi-GPU
  - Extended strong scaling roll-off by 2x for ExaWind
  - Used low-synchronization Gram-Schmidt and GMRES algorithms, NLAA (2019) K. Swirydowicz, J. Langou, S. Ananthan, U. Yang, S. Thomas
- Enabled ExaWind to reach challenge problem success with hybrid solvers







Hybrid solver simulation of a small wind farm; simulation with 650-millsion grid points (Large-scale Summit run)

Credit to PI: Mike Sprague (NREL)

ExaSGD - A MORE RELIABLE AND ROBUST POWER GRID

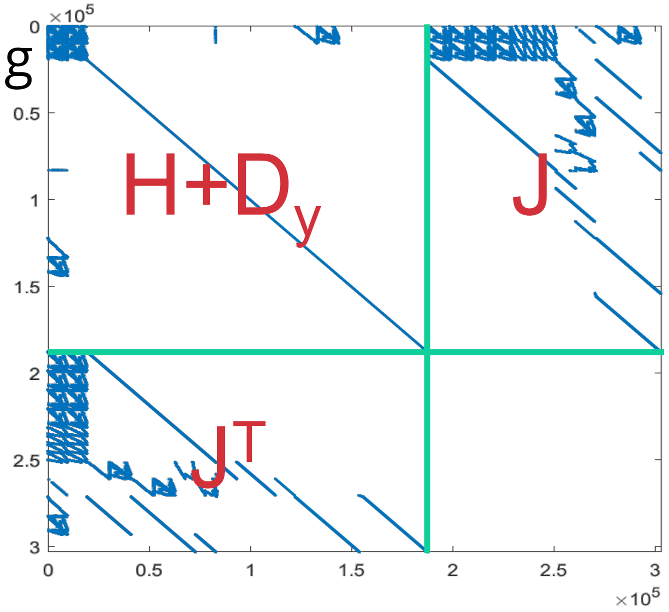
PI: Chris Oehmen, PNNL

- **Objective**: Optimize power grid operation and control to reliably incorporate intermittent energy sources such as wind and solar while minimizing disruptions

- **ECP Scope:**
  - Enable "what if" evaluation of complex damage from extreme weather or cyber attack by optimizing with many weather scenarios and complex disruptions ("contingencies")
  - Allow for flexibility to run on laptops to exascale computers; highly advanced numerical algorithms leverage accelerated computing

- New Solver strategies for GPUs for large sparse systems of equations were required

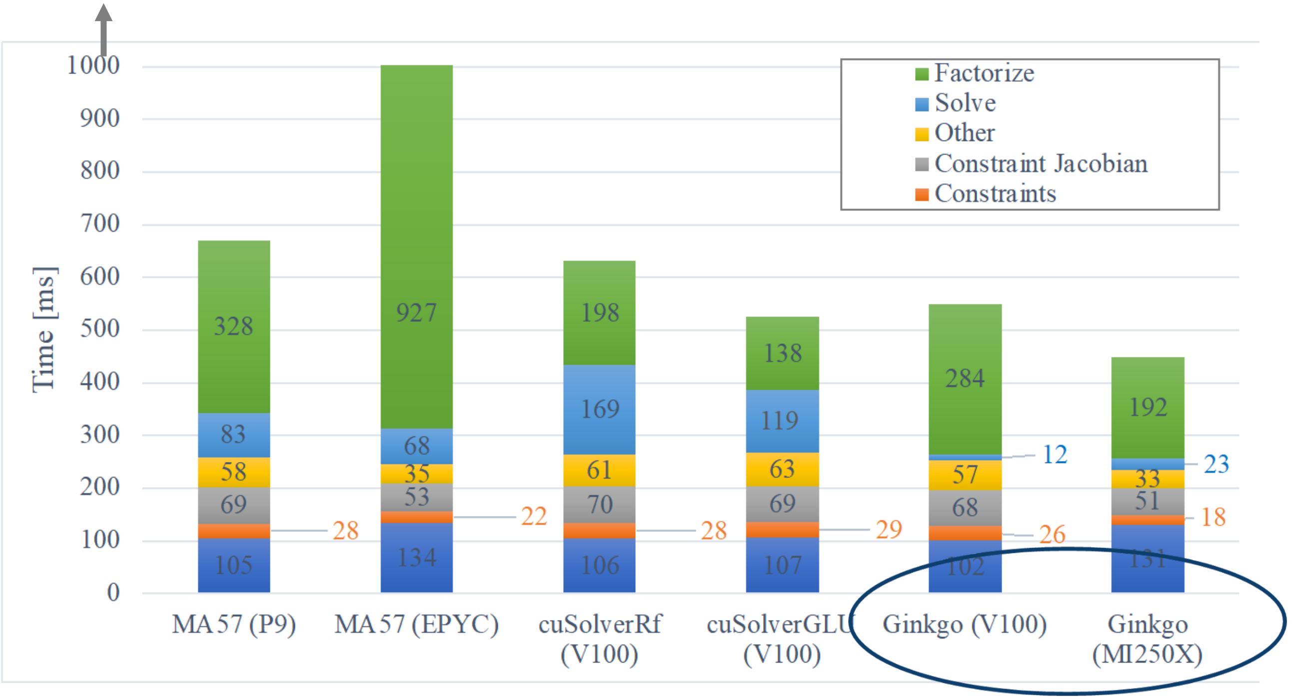
Can analyze Western power grid operations, with multiple scenarios and 1000's of contingencies to help operators plan for and respond to emergencies



Typical sparsity pattern of optimal power flow matrices: No obvious structure that can be used by linear solver. Needed to solve long sequences of such systems

COLLABORATION EFFORT BETWEEN SUPERLU, GINGKO AND ExaSGD PRODUCED NEW DIRECT SPARSE SOLVERS

- Used non-supernodal structures for GPUs
- Extract parallelism for extreme sparsity of the graph; find level sets of coefficients that can be processed simultaneously
- Cholesky for symmetric, LU for non-symmetric
- Results:
  - Each GPU solution outperforms all CPU baselines
  - Ginkgo performance improves on a better GPU
  - Iterative refinement configuration affects linear solver performance and optimization solver convergence



Ginkgo provides the first portable GPU-resident sparse direct linear solver for non-supernodal systems

A MULTI-PRECISION EFFORT WAS CREATED IN FY20 TO BETTER LEVERAGE GPUS

- Goal is to integrate mixed precision technology as production-ready implementation into ECP software products to allow for the smooth integration into ECP applications.
  - Leverage low-precision cores developed for AI/ML
  - Algorithm development and evaluation primarily focused on GPU hardware (AMD, Intel, NVIDIA GPUs)
- Cross-laboratory expert teams focused on:
  - Mixed precision dense direct solvers (MAGMA and SLATE)
  - Mixed precision sparse direct solvers (SuperLU)
  - Mixed precision multigrid (on a theoretical level and in hypre)
  - Mixed precision FFT (heFFTE)
  - Mixed precision preconditioning and Krylov solvers (Ginkgo, Trilinos)
  - Separating the arithmetic precision from the memory precision (Ginkgo)

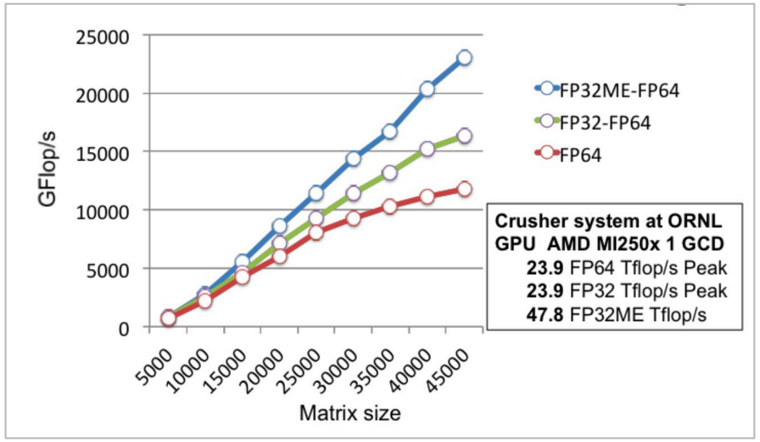
PETSc
Ginkgo
hypre
ECP SLATE ICL
MAGMA ICL UT

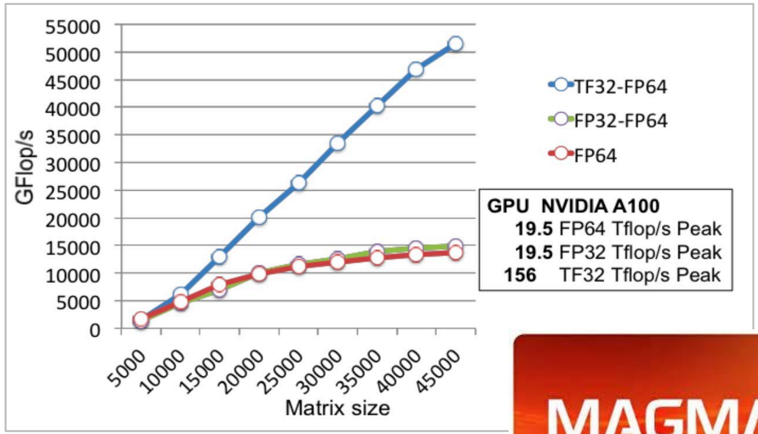
Working in a large collaborative team allowed sharing lessons learned and interoperability of results in the library ecosystem

EXAMPLE OUTCOME

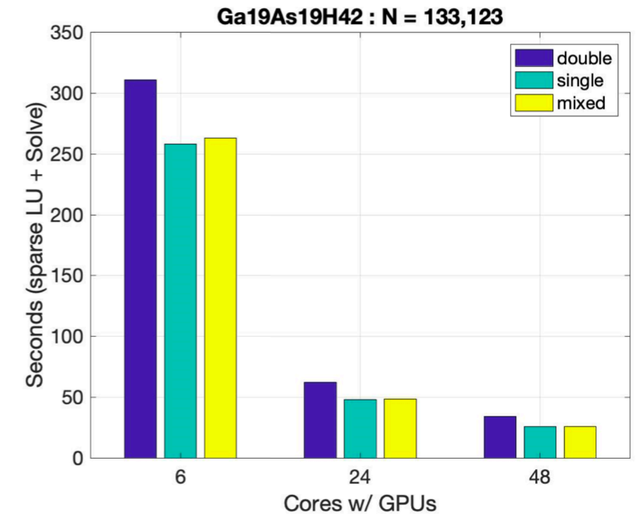
Several teams saw significant improvements using mixed precision in high precision solvers using iterative refinement

Magma dense LU using MPIR





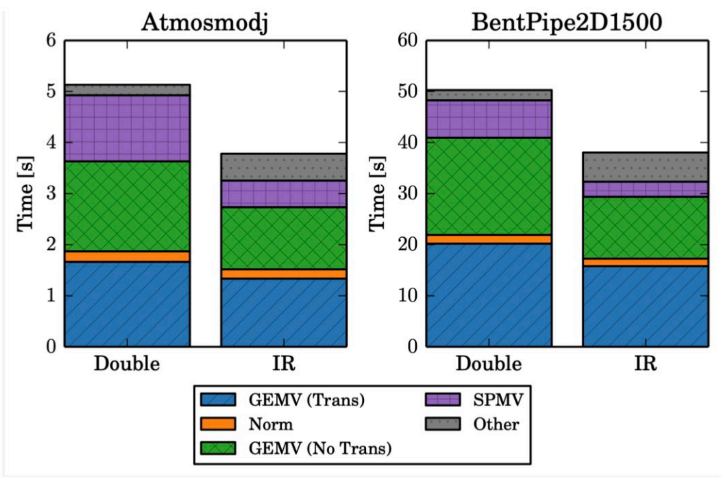
SuperLU sparse LU using MPIR



Mixed precision in SuperLU gives the results of high precision solvers at costs only slightly more than single precision

Trllinos mixed precision GMRES solvers see improvements for nearly al stages of the algorithm

Trilinos GMRES(5) using MPIR



The MAGMA dense LU factorization improved by factors of 2X to 3X using MPIR

DELIVERING AN OPEN, HIERARCHICAL SOFTWARE ECOSYSTEM

ECP ST Open Product Integration Architecture

Levels of Integration	Product	Source and Delivery
• Build all SDKs • Build complete stack • Containerize binaries	E4S	**Source:** ECP E4S team; Non-ECP Products (all dependencies) **Delivery:** spack install e4s; containers; CI Testing
• Group similar products • Make interoperable • Assure policy compliant • Include external products	SDKs	**Source:** SDK teams; Non-ECP teams (policy compliant, spackified) **Delivery:** Apps directly; spack install sdk; future: vendor/facility
• Standard workflow • Existed before ECP	ST Products	**Source:** ECP L4 teams; Non-ECP Developers; Standards Groups **Delivery:** Apps directly; spack; vendor stack; facility stack

ECP ST Individual Products

ACHIEVEMENTS IN NEW APPLICATION SCIENCE WOULD HAVE NOT BEEN POSSIBLE WITHOUT A ROBUST SOFTWARE STACK

ECP's 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
- Growing functionality: Nov 2023: E4S 23.11 – 120+ full release products

Community Policies Commitment to SW quality	DocPortal Single portal to all E4S product info	Portfolio testing Especially leadership platforms
Curated collection The end of dependency hell	Quarterly releases Release 23.11 – November	Build caches 10X build time improvement
Turnkey stack A new user experience	https://e4s.io	Post-ECP Strategy LSSw.io, PESO

E4S

https://e4s.io

E4S lead: Sameer Shende (U Oregon)

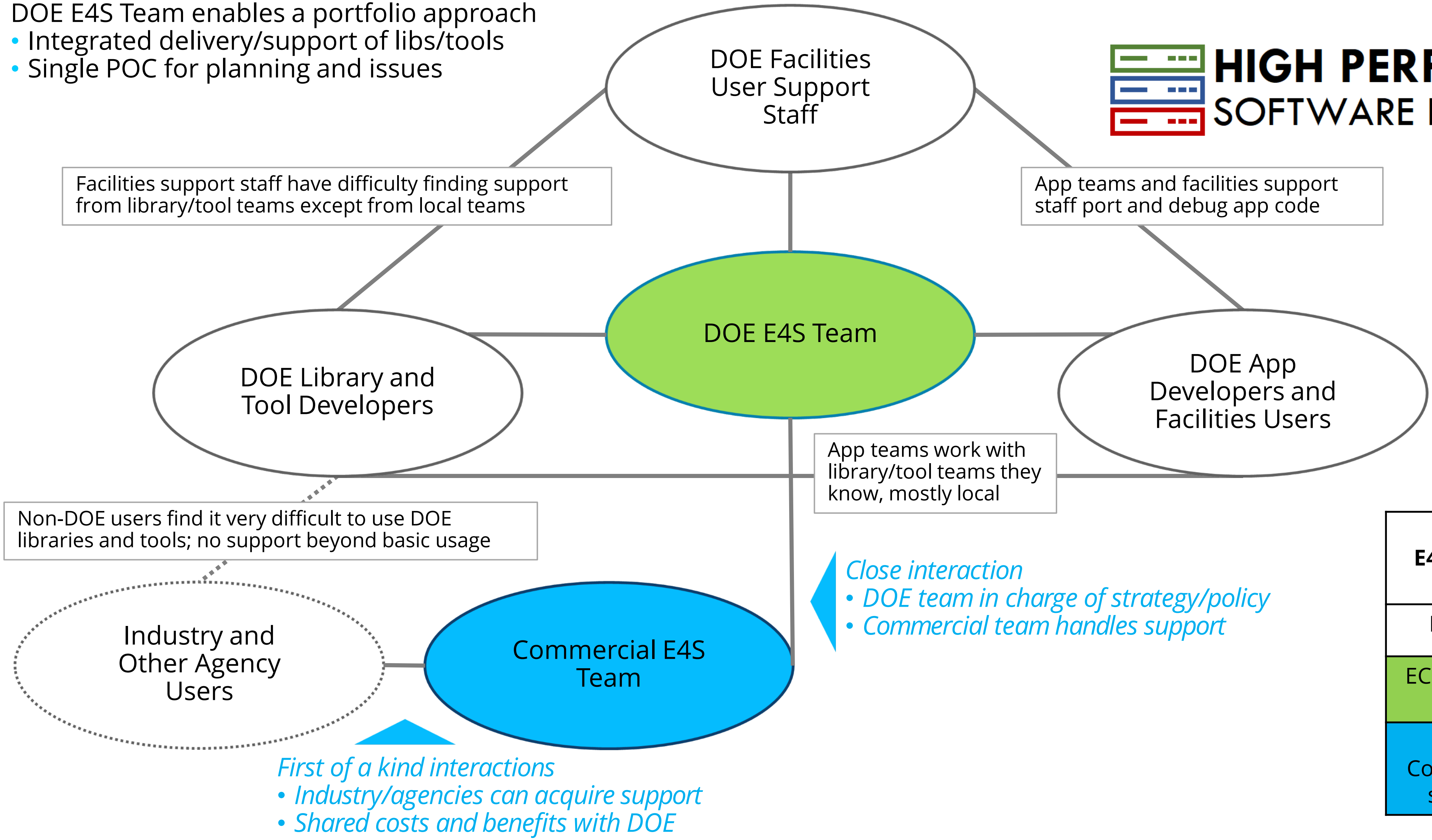
Also includes other products, e.g.,
AI: PyTorch, TensorFlow, Horovod
Co-Design: AMReX, Cabana, MFEM

E4S BUSINESS MODEL: OPTIMIZE COST & BENEFIT SHARING

DOE E4S Team enables a portfolio approach
- Integrated delivery/support of libs/tools
- Single POC for planning and issues

New

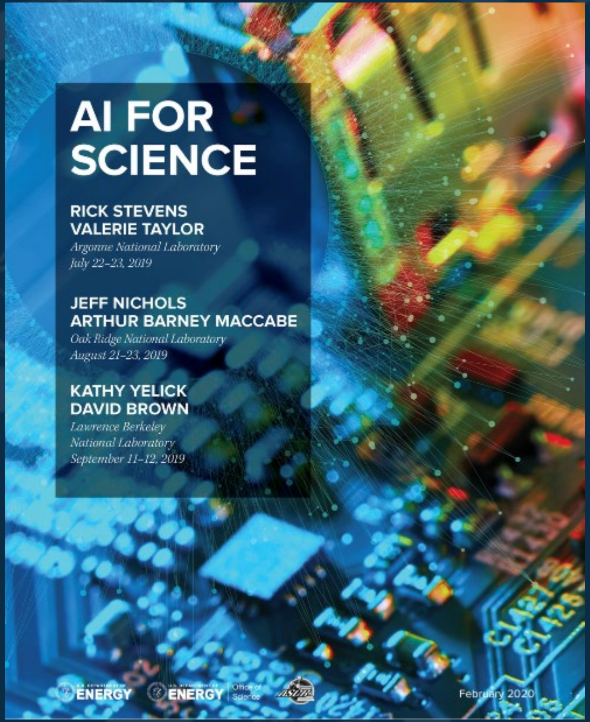
HIGH PERFORMANCE SOFTWARE FOUNDATION



E4S Phase	Cost & Benefit Scope
Pre-E4S	Local Facility
ECP support	DOE complex
+ Commercial support	Universal

DOE HAS BEEN GATHERING WIDE COMMUNITY INPUT (>1300 RESEARCHERS)

2019



What changed in three years?

- Language Models (e.g. ChatGPT) released
- Artificial image generation took off
- AI folded a billion proteins
- AI hints at advancing mathematics
- AI automation of computer programming
- Explosion of new AI hardware
- AI accelerates HPC simulations
- Exascale machines start to arrive

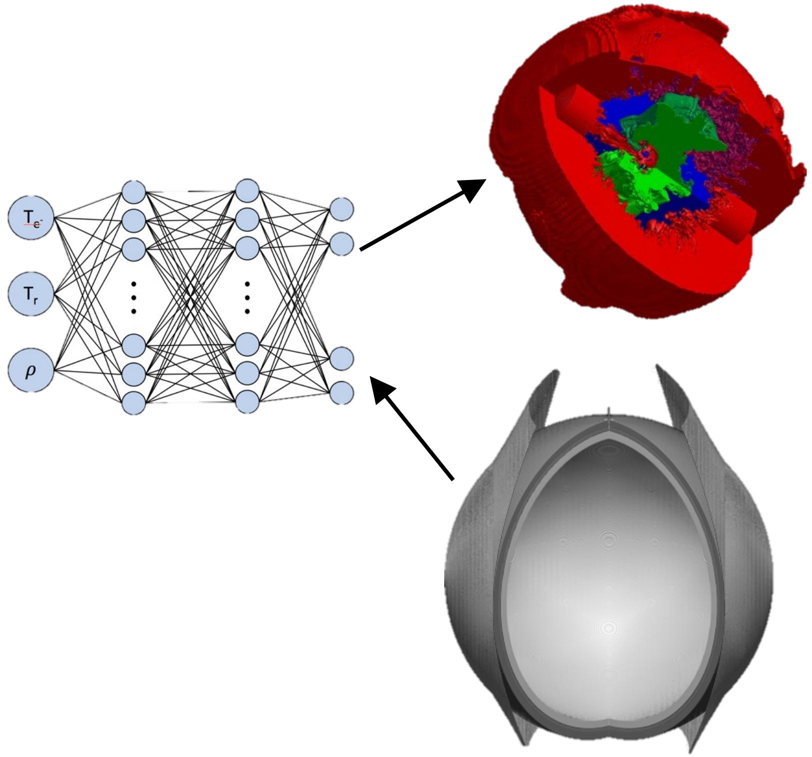
2022

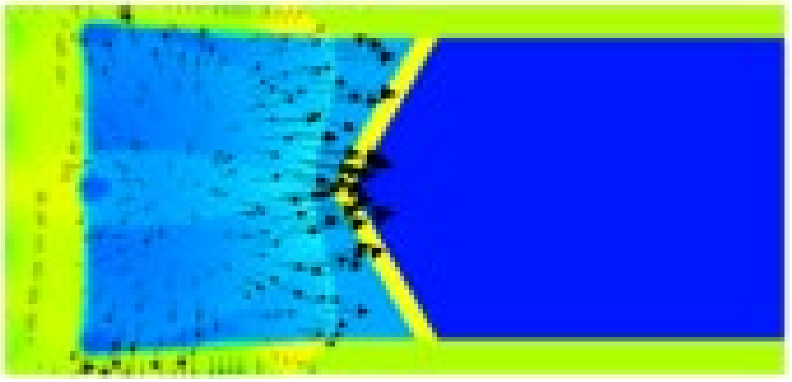
2020 DOE Office of Science ASCR Advisory Committee report recommending major DOE AI4S program

Report posted here:
https://www.anl.gov/ai-for-science-report

U.S. DEPARTMENT OF ENERGY | Office of Science | NNSA National Nuclear Security Administration

FROM THE WORKSHOPS IT WAS CLEAR THAT AI REPRESENTS A POWERFUL NEW FOUNDATION FOR PROGRESS IN SCIENCE AND TECHNOLOGY





AI based surrogates for HPC

Climate Ensembles
Effective Zettascale on Exa

AI and robotics for autonomous discovery

Materials, Chemistry, Biology
Light-Sources, Neutrons

AI for software engineering and programming

Code Translation, Optimization
Quantum Compilation, QAlgs

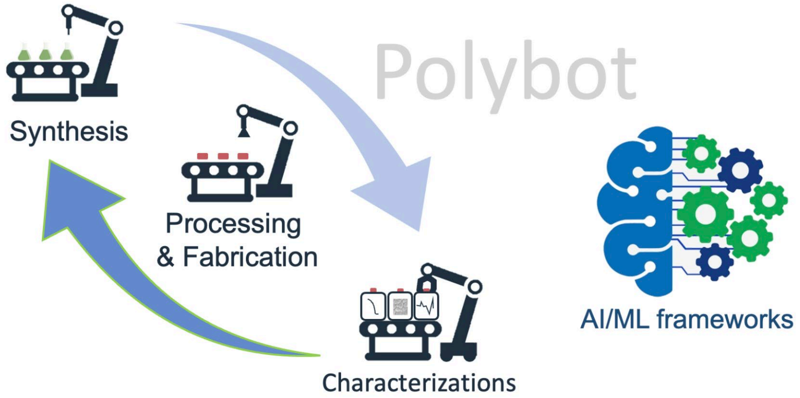
AI for prediction and control of complex engineered systems

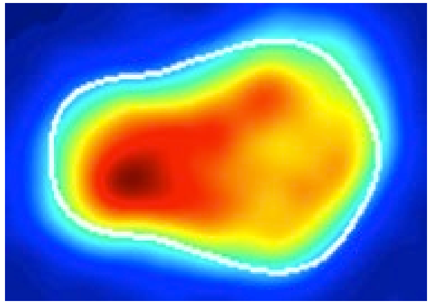
Accelerators, Buildings, Cities,
Power Grid, Networks

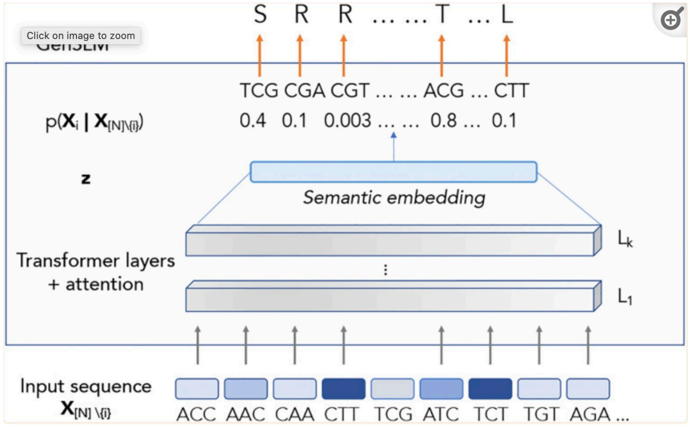
AI for advanced properties inference and inverse design

Energy Storage
Proteins, Polymers

Foundation AI for scientific knowledge

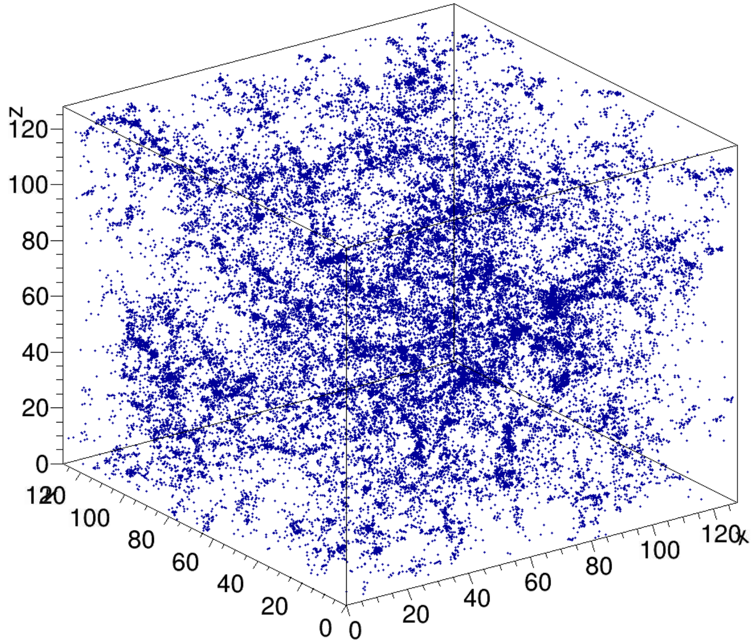






TRAINING DNNS ON VERY LARGE DATA SAMPLES

Solving the inverse problem of determining the cosmological parameters from simulation output.



- Sure, that should be possible as we have large-scale GPU machines!

  → *Not actually true*

- Even a mini-batch with just one sample may require O(10) GB of memory

- e.g., cosmological datasets
  - 512^3, 4 channels, 2 bytes per element → 1 GB per sample
  - A regression model for the 512^3 dataset does not fit to the GPU memory of Sierra
  - Splitting one sample into smaller disjoint samples can workaround the memory pressure, but discards long-range information

Use LBANN for this effort.

EXASCALE COMPUTING PROJECT

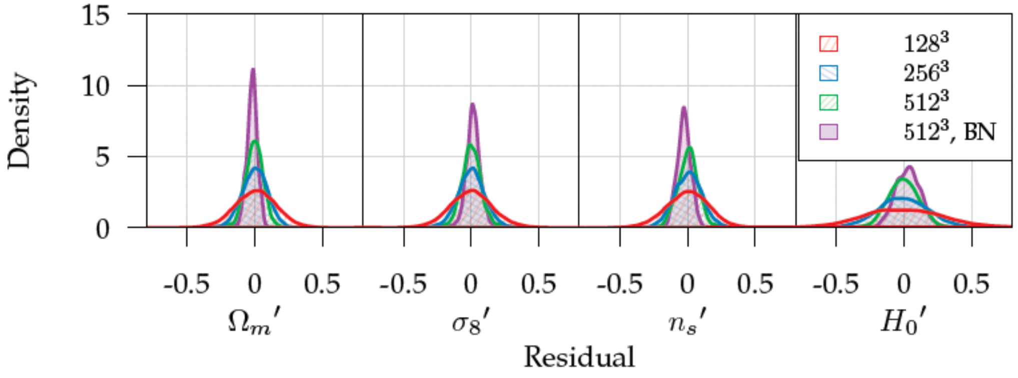
TRAINING DNNS ON VERY LARGE DATA SAMPLES

Solving the inverse problem of determining the cosmological parameters from simulation output.

We solve these challenges by extensively applying hybrid parallelism throughout the end-to-end training pipeline, including both computations and I/O with LBANN.

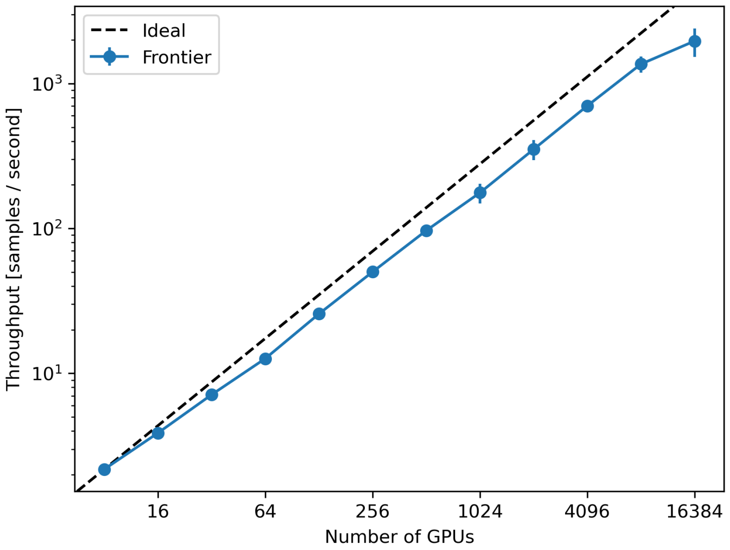
Our hybrid-parallel algorithm extends the standard data parallelism with spatial parallelism, which partitions a single sample in the spatial domain, realizing strong scaling beyond the mini-batch dimension with a larger aggregated memory capacity.

We were able to enable training of CosmoFlow with much larger samples than previously possible, realizing an order-of-magnitude improvement in prediction accuracy.



True/predicted cosmological parameters (normalized to [-1; 1]) from four different configurations. Our analysis with the 512^3 samples was ~10X better.

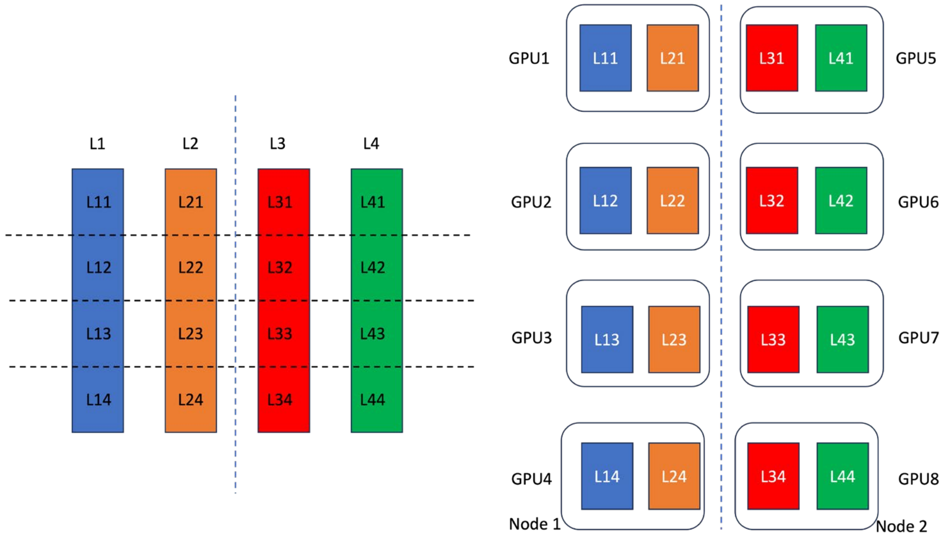
The scaling on Frontier was nearly ideal up to 16k GPUs



EXASCALE COMPUTING PROJECT

LLMs ON FRONTIER: TRAINING A TRILLION PARAMETER GPT MODEL

- 3D parallelism by porting Megatron-DeepSpeed
  - Model parallel
  - Pipeline parallelism
  - Data parallelism
- Trade-off between memory, compute, and communication
- Hyperparameter tuning to find the best distribution strategy
- First tune a 175 B model, then grow it to 1T model

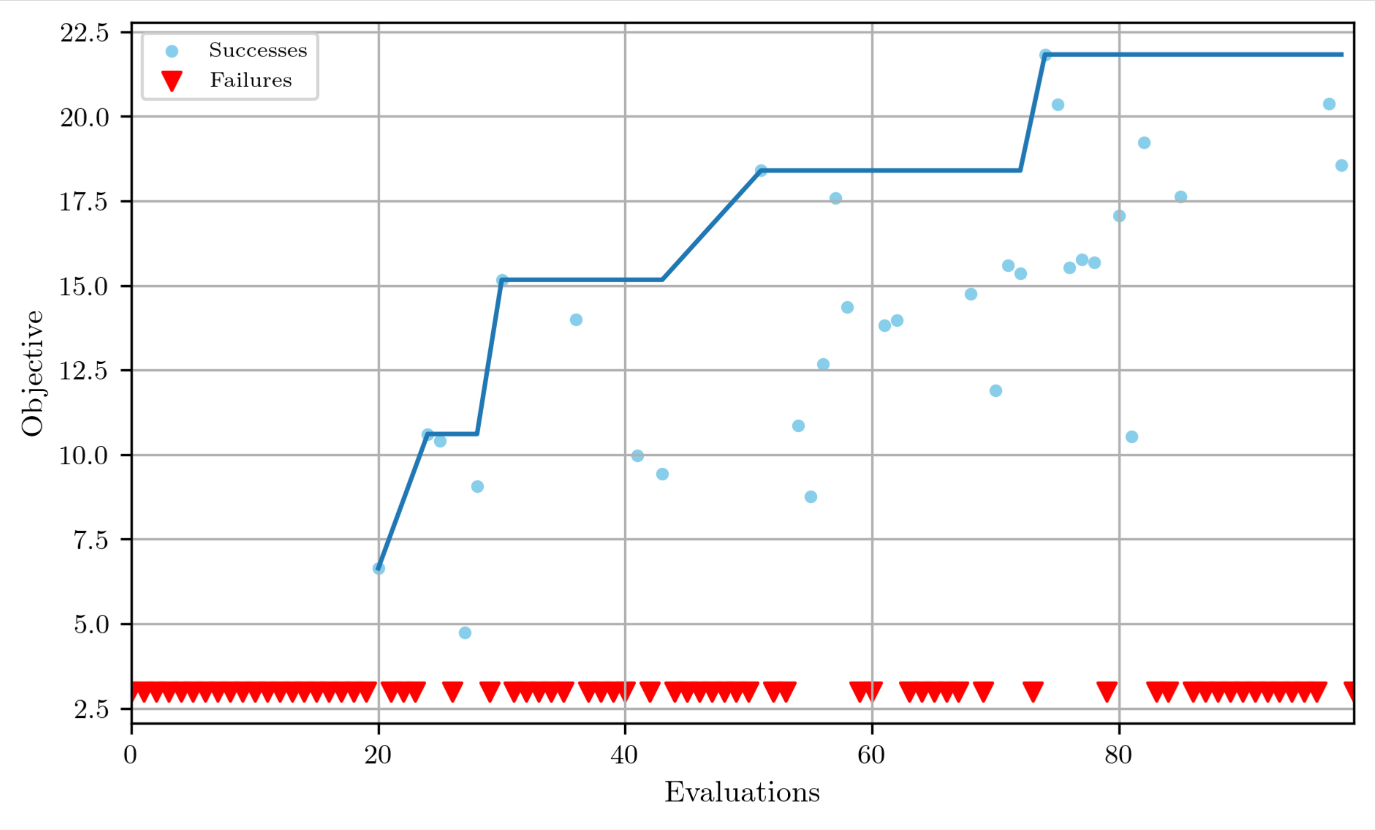


S. Dash, I. Lyngaas, J. Yin, X. Wang, R. Egele, G. Cong, F. Wang, and P. Balaprakash. ***Optimizing Distributed Training on Frontier for Large Language Models***. arXiv preprint arXiv:2312.12705 (2023).

LLMs ON FRONTIER: TUNING 175 BILLION PARAMETER MODEL

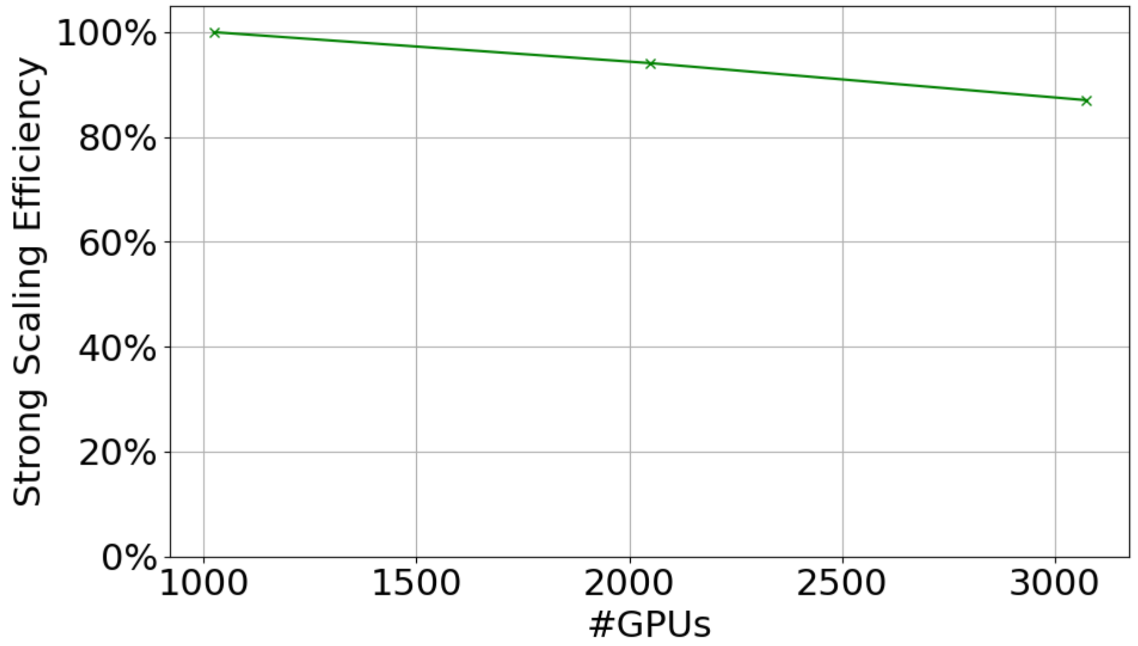
Distribution Parameter	Search-Space
Pipeline-parallel-size (PP)	$PP \in \{1, 2, 4, 8, 12, 16\}$
Tensor-parallel-size (TP)	$TP \in \{1, 2, 4, 8\}$
Micro-batch-size (MBS)	$MBS \in [4, 20]$
Gradient accumulation steps (GAS)	$GAS \in \{5, 10\}$
ZeRO-1 Optimizer	$ZeRO-1 \in \{True, False\}$
Number of Nodes (NNODES)	$NNODES \in \{12, 16\}$

Hyperparameter tuning using DeepHyper



S. Dash, I. Lyngaas, J. Yin, X. Wang, R. Egele, G. Cong, F. Wang, and P. Balaprakash. ***Optimizing Distributed Training on Frontier for Large Language Models***. arXiv preprint arXiv:2312.12705 (2023).

LLMs ON FRONTIER: FROM A 175B PARAMETER MODEL TO 1T!



Frontier trained a ChatGPT-sized large language model with only 3,000 of its 37,888 Radeon GPUs — the world's fastest supercomputer blasts through one trillion parameter model with only 8 percent of its MI250X GPUs

News By Matthew Connatser published January 07, 2024

Now you're playing with AI power!

Comments (19)

(Image credit: ORNL)

Researchers at Oak Ridge National Laboratory trained a large language model (LLM) the size of ChatGPT on the Frontier supercomputer and only needed 3,072 of its 37,888 GPUs to do it. The team published a research paper that details

S. Dash, I. Lyngaas, J. Yin, X. Wang, R. Egele, G. Cong, F. Wang, and P. Balaprakash. ***Optimizing Distributed Training on Frontier for Large Language Models***. arXiv preprint arXiv:2312.12705 (2023).

ECP Summary

- A $1.8B project started in 2016
  - 15 DOE Laboratories
  - Dozens of university and industrial partners
  - 1000 participants
  - Three technical focus areas and 81 research projects
- ECP met all the key performance parameters
- ECP technical work completed Dec 31, 2023
- The ECP leadership team is now focused on project closeout, communications, and outreach

ECP impact and legacy

- A suite of applications that will impact problems of national importance for decades to come
- Integrated software stack for GPU-accelerated computing widely available for use
- Best practices and lessons learned for thinking about how to program GPUs – moving the nation forward
- Over 1000 researchers trained and ready for accelerator-based computing
- Best practices for running a large-scale software development RD&D 413.3b project

Lessons Learned

- Diverse, multi-disciplinary teams with stable funding can achieve great things
- Collaborative solutions can't be dictated but they can be incentivized
- Build integration into project structure and measures of success
- High quality software is the foundation for collaboration in scientific computing

Next Steps

- Formally close the 413.3b project
- Long term sustainability efforts have begun in the Software Stewardship projects (e.g. PESO)
- Would like to keep releasing xSDK and E4S broadly to DOE, other agencies, industry, academia
- Help enable everyone to achieve the 100X potential of GPU accelerated computing

ECP-FUNDED TECHNOLOGIES & PARTICIPANTS ARE RECEIVING RECOGNITION

R&D 100 AWARDS

CANDLE (2023)
ZFP (2023)
Variorum (2023)
Flash-X (2022)
Mochi (2021)
SZ (2021)
Flux (2021)
Legion (2020)
Spack (2019)
UCX (2019)
Darshan (2018)
Swift/T (2018)
Nek5000 (2016)

HPCwire People to Watch

Jeff Vetter (2021)
Bronis de Supinski (2021)
Lori Diachin (2019)
Doug Kothe (2018)

acm

Gordon Bell Finalist | ExaSMR (2023)
Gordon Bell | WarpX (2022)
Gordon Bell Special Prize | CANDLE (2022)
A. M. Turing Award | Jack Dongarra (2021)
ACM Fellow | Bronis de Supinski (2022), Rob Ross (2021), Rick Stevens (2020)
IEEE CS Ken Kennedy Award | Ian Foster (2022)
IEEE CS Sidney Fernbach Award | Salman Habib (2020)

...and many more notables, including

CRA Distinguished Service Award | Paul Messina (2018)
Best Open Source Contribution (For IPDPS 2023 paper) | ExaIO (2023)
SIAM/ACM prize in Computational Science and Engineering | SUNDIALS (2023)
The Ernest Orlando Lawrence Award | ExaStar and DataLib (2021)
SIAM Fellow | Mike Heroux (2019); IEEE Fellow | Rajeev Thakur (2022)

Supplemental Material

THE xSDK EFFORT

xSDK lead: Ulrike Meier Yang (LLNL)
xSDK release lead: Satish Balay (ANL)


xSDK release 1.0.0 (Nov 2023)

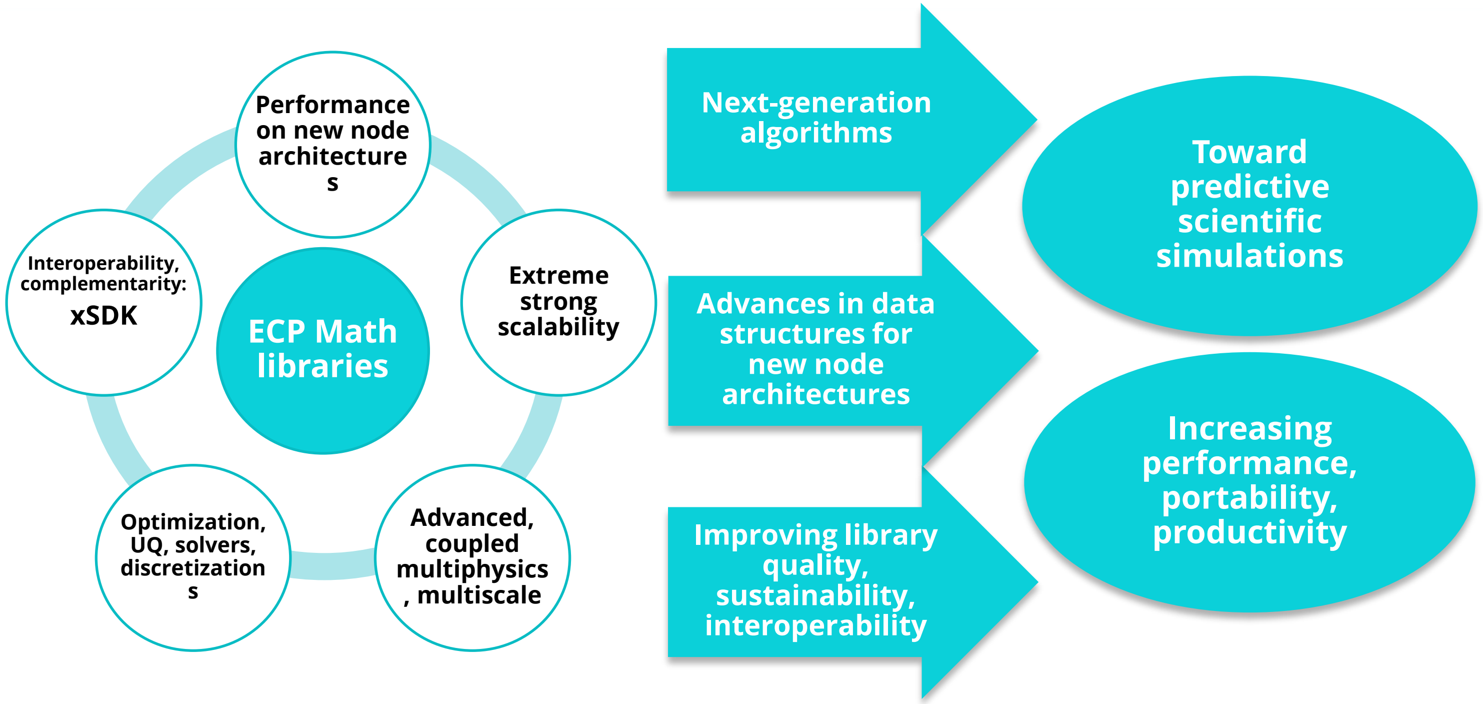
- hypre
- PETSc/TAO
- SuperLU
- Trilinos

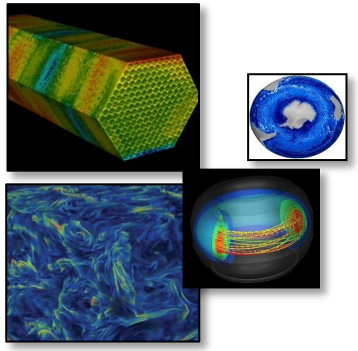
} Began in 2016 with 4 libraries, IDEAS-Classic project

- AMReX
- ArborX
- ButterflyPACK
- DTK
- ExaGO
- Ginkgo
- heFFTe
- HiOp
- libEnsemble
- MAGMA
- MFEM
- Omega_h
- PLASMA
- PUMI
- SLATE
- Tasmanian
- SUNDIALS
- Strumpack
- Alquimia
- PFLOTRAN
- deal.II
- preCICE
- PHIST
- SLEPc

} from the broader community

- Facilitates the combined use of independently developed software packages for ECP applications
- Provides community policies for better software quality and sustainability
- Provides deeper levels of library interoperability

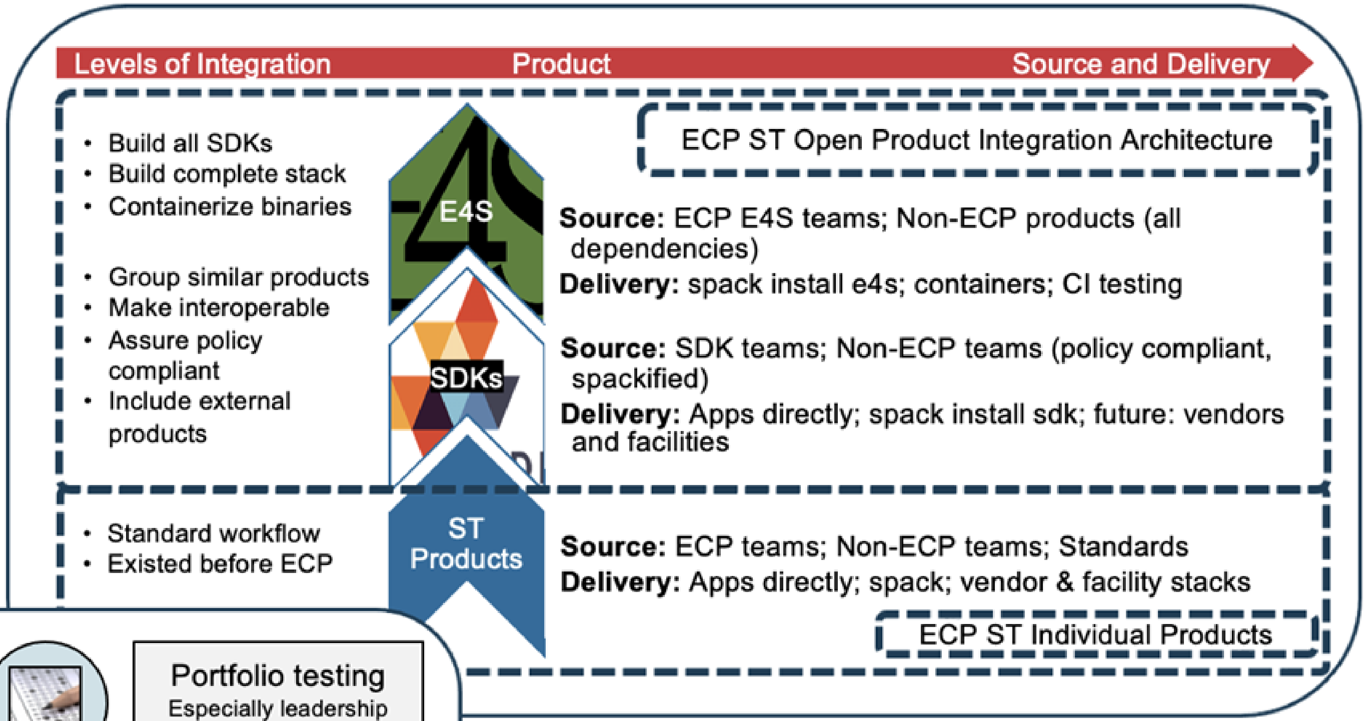


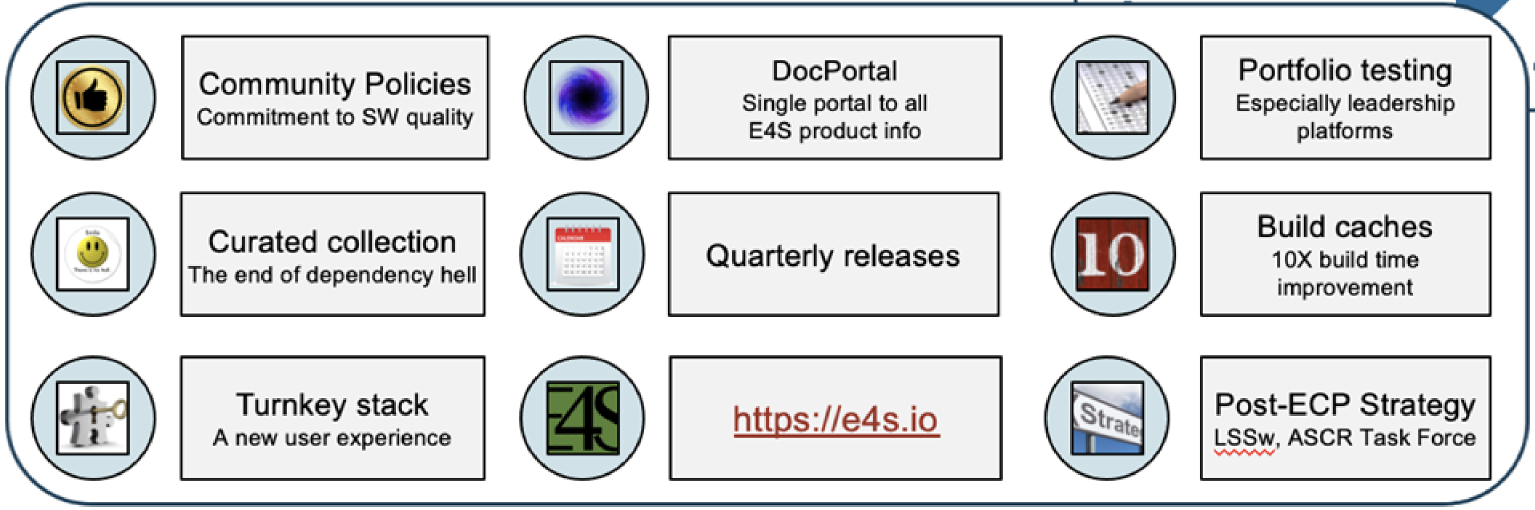


Ref: xSDK: Building an Ecosystem of Highly Efficient Math Libraries for Exascale, **SIAM News**, Jan 2021

SDKS WERE BROADLY DEPLOYED

- ECP Software Technology lead: Mike Heroux (SNL)
- 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**
- Not a commercial product – an open resource for all
- Supported by DOE and commercial entities (Paratools)
- Growing functionality: November 2023: E4S 23.11 – 100+ full release products





https://e4s.io

E4S lead: Sameer Shende (U Oregon)

Spack https://spack.io

Spack lead: Todd Gamblin (LLNL)

ECP APPLICATIONS AND MATHEMATICAL LIBRARIES

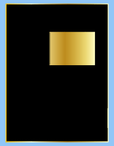
Vision	Provide high-quality, sustainable extreme-scale math libraries that are constantly improved by a robust research and development effort and support exascale needs of the ECP community	
Challenges	Need advances in algorithms and data structures to exploit emerging exascale architectures (high concurrency, limited memory bandwidth, heterogeneity); need new functionality to support predictive simulation and analysis	
Portfolio Goals	**Advanced algorithms**	• Advanced, coupled multiphysics and multiscale algorithms (discretizations, preconditioners & Krylov solvers, nonlinear & timestepping solvers, coupling) • Toward predictive simulation & analysis (optimization, sensitivities, UQ, ensembles)
	Performance	• Performance on new node architectures • Extreme strong scalability

	Improving library sustainability & complementarity	• Math library interoperability and complementarity through the xSDK
		  ◦ Improving package usability, quality, sustainability
		  ◦ Community coordination and collaboration while retaining package autonomy

ECP SOFTWARE TECHNOLOGY – PRESENT AND FUTURE

Key themes:
- Focus: GPU node architectures and advanced memory & storage technologies
- Create: New high-concurrency, latency tolerant algorithms
- Develop: New portable (Nvidia, Intel, AMD GPUs) software product
- Enable: Access and use via standard APIs

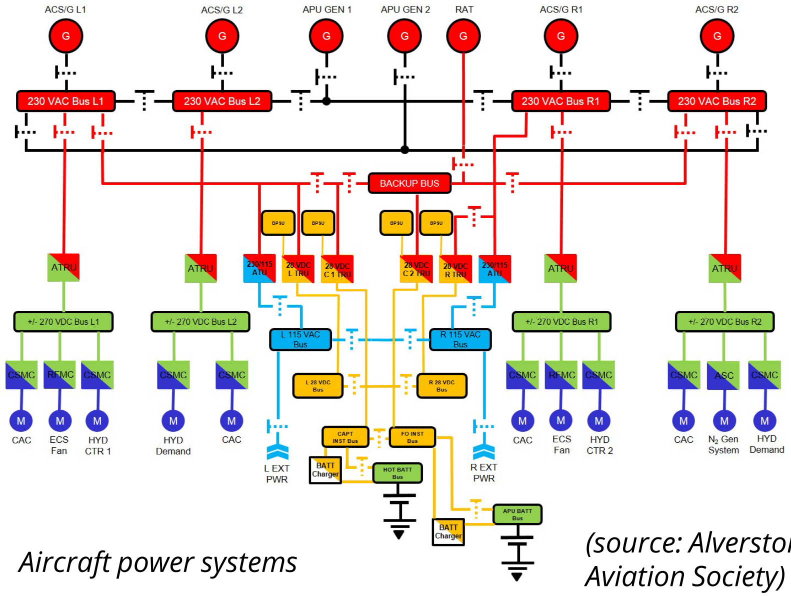
Software categories:
- **Next generation established products:** Widely used HPC products (e.g., MPICH, OpenMPI, PETSc)
- **Robust ~~emerging~~ products:** Address key new requirements (e.g., Kokkos, RAJA, Spack)
- **New products:** Enable exploration of emerging HPC requirements (e.g., zfp, Variorum)



Legacy: A stack that enables performance portable application development on leadership platforms.

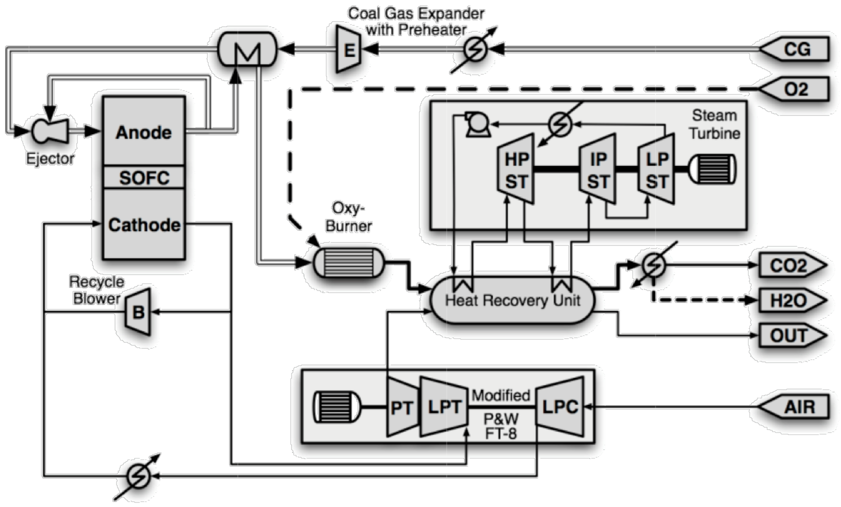
Example Products	Engagement
MPI – Backbone of HPC apps	Explore/develop MPICH and OpenMPI new features & standards
OpenMP/OpenACC –On-node parallelism	Explore/develop new features and standards
Performance Portability Libraries	Lightweight APIs for compile-time polymorphisms
LLVM/Vendor compilers	Injecting HPC features, testing/feedback to vendors
Perf Tools - PAPI, TAU, HPCToolkit	Explore/develop new features
Math Libraries: BLAS, sparse solvers, etc.	Scalable algorithms and software, critical enabling technologies
IO: HDF5, MPI-IO, ADIOS	Standard and next-gen IO, leveraging non-volatile storage
Viz/Data Analysis	ParaView-related product development, node concurrency

SYSTEMS ENGINEERING DOMAIN



Aircraft power systems

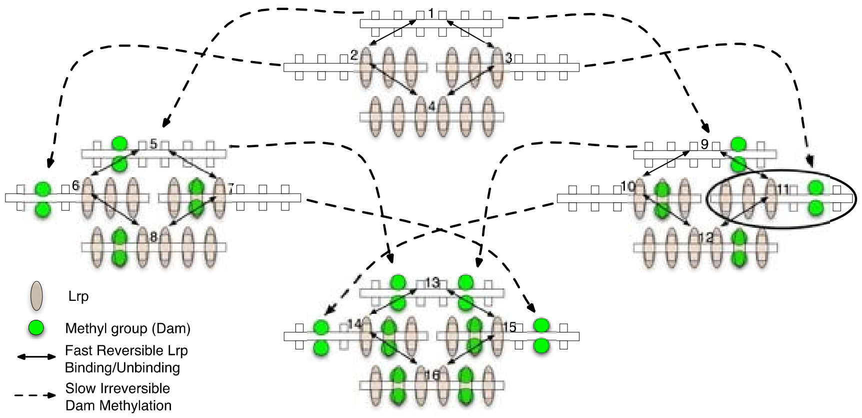
(source: Alverstone Aviation Society)



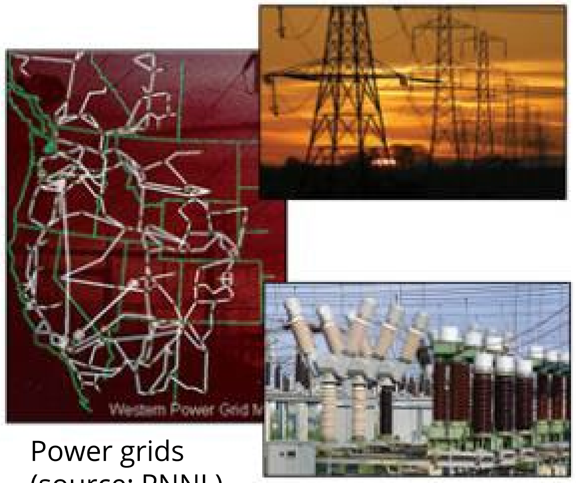
Solid oxide fuel cell plant (source: Kameswaran at al. 2010)

- ExaSGD addresses systems engineering problems
- Produced new direct sparse solvers using non-supernodal structures, for GPUs
- Cholesky for symmetric, LU for non-symmetric
- Joint effort between SuperLU, Ginkgo, ExaSGD teams

Buildings (source: EEB Hub, B661 2014)



Gene regulatory networks (source: Peles et al. 2006)



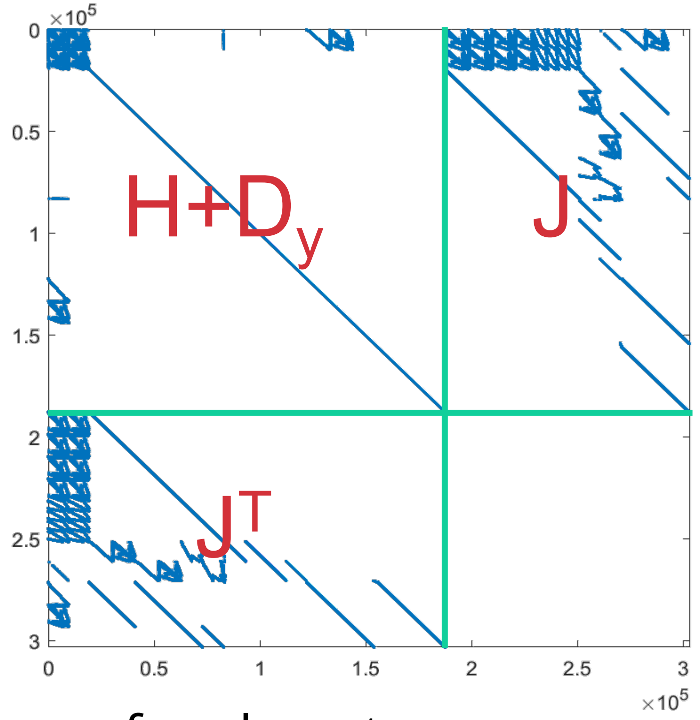
Power grids
(source: PNNL)

UNDERLYING KKT LINEAR SYSTEM PROPERTIES

- Security constrained optimal power flow analysis
- The interior method strategy leads to symmetric indefinite linear systems

$$\overbrace{\begin{bmatrix} H + D_y & J \\ J^T & 0 \end{bmatrix}}^{K_k} \overbrace{\begin{bmatrix} \Delta y \\ \Delta \lambda \end{bmatrix}}^{\Delta x_k} = \overbrace{\begin{bmatrix} r_y \\ r_\lambda \end{bmatrix}}^{r_k},$$

- H – sparse Hessian,
- D_y – arises from log-barrier function



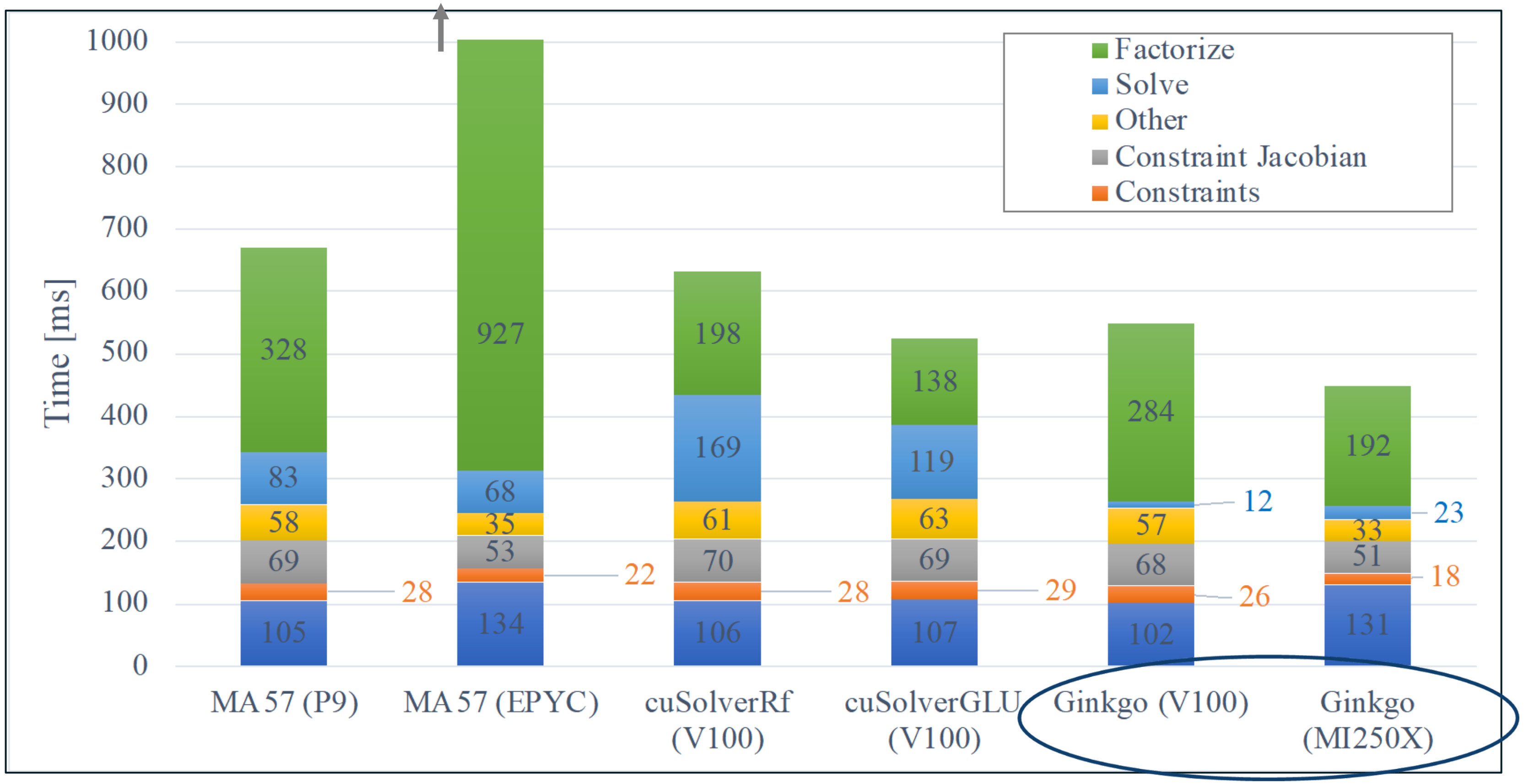
Typical sparsity pattern of optimal power flow matrices: No obvious structure that can be used by linear solver.

- The challenge: we need to solve a long sequences of such systems

LINEAR SOLVER PERFORMANCE WITHIN OPTIMIZATION ALGORITHM

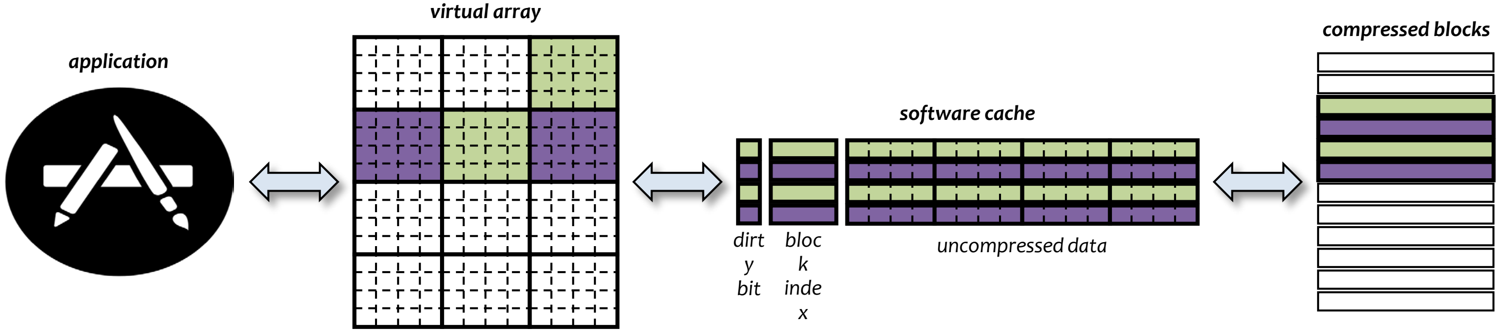
Average per iteration times (including first iteration on CPU)

- Each GPU solution outperforms all CPU baselines
- Ginkgo performance improves on a better GPU
- Iterative refinement configuration affects linear solver performance and optimization solver convergence



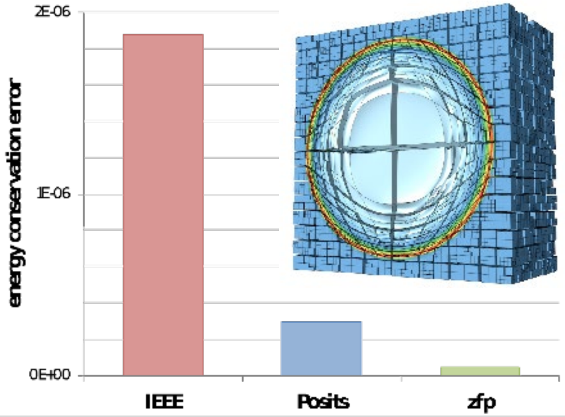
Ginkgo provides the first portable GPU-resident sparse direct linear solver for non-supernodal systems

ADDRESSING GROWING GAP OF OPS, BW, & MEMORY: ZFP



- Fixed-length compressed blocks enable fine-grained read & write **random access**
  - C++ compressed-array classes hide complexity of compression & caching from user
  - User specifies per-array storage footprint in bits/value
- Absolute and relative **error tolerances** supported for offline storage, sequential access
- Fast, hardware friendly, and parallelizable: **150 GB/s throughput** on NVIDIA Volta
- **HPC tool support**:

Inline compression boosts solution accuracy by 40x over IEEE in CEED code



Prepared by LLNL under Contract DE-AC52-07NA27344.

ECP generated a:

- Collection of portable GPU-capable libraries and tools for AMD, Intel, and NVIDIA devices
- Designed for future adaptation to next-generation highly-concurrent node architectures
- Foundation for others who will make the transition from CPU to GPU and beyond

ECP DID LEADERSHIP SCIENCE ON LEADERSHIP SYSTEMS

- AI/ML
  - E4S already builds AI/ML products: PyTorch, TensorFlow, Horovod
  - Opportunity: Curate and support additional stacks
    - Many scientific teams rely on their own ad hoc fragile stack, often generations behind latest
    - DOE teams are working on their own AI/ML capabilities, need integration and support
  - The “Frank” system sponsored by DOE includes key AI target devices
  - Bottom line:
    - **Extension of ecosystem efforts to AI should require modest changes to our approach**
    - Certainly, better than establishing a different stack
    - For science, M&S and AI/ML software are used in combination – **a single stack makes sense**
- Cloud
  - **E4S is already available in containers, on AWS, and Google Cloud**
  - **We use these resources for testing, and so do the cloud providers (to assure their SW works with ours)**
  - Provide a common test and evaluation setting when working with non-DOE users
- Quantum
  - Most people I know in this field are physicists
  - We don’t know enough to say what is needed
  - Even so, these devices will be hosted - a lot of what we know about HPC software can apply

ECP INVESTMENTS ENABLED A 100X IMPROVEMENT IN CAPABILITIES

- **7 years** building an **accelerated**, **cloud-ready** software ecosystem
- Positioned to utilize **accelerators from multiple vendors** that others cannot
- **Emphasized software quality**: testing, documentation, design, and more
- Prioritized **community engagement**: Webinars, BOFs, tutorials, and more
- DOE **portability layers** are the credible way to
  - Build codes that are sustainable **across multiple GPUs** and
  - **Avoid vendor lock-in**
  - **Avoid growing divergence** and hand tuning in your code base
- ECP software can **lower costs** and **increase performance** for **accelerated** platforms
- Outside of AI, industry has not caught up
  - DOE enables an entirely different class of applications and capabilities to use accelerated nodes
  - In addition to AI
- **ECP legacy: A path and software ecosystem for others to leverage**

MORE THAN ONE WAY TO LEVERAGE 100X

- 100X can be realized as exciting new science capabilities at the high end
  - Fundamental new science on leadership platform
  - New opportunities on affordable machines that fit in current data centers
- **But can also reduce costs**
- Migration to accelerated platforms can be used to
  - Migrate a problem from an HPC cluster to a deskside or laptop systems
  - Lower your AWS monthly charges – E4S is available for container/cloud
  - Keep energy costs in check while still growing computing capabilities
- **Biggest ECP impact will be accelerating GPU transition – at all levels**
- **Transitioning software stacks to GPUs is essential**
  - **CPU-based HPC system realize only modest energy efficiency improvements**
  - **Migrating to GPUs is key to improving HPC environmental impact**

100X RECIPE

- **Ingredients**
  - A compelling science impact story
  - $$$ - $$$$
  - Staff
  - Computing resources, training
  - The deliverables and experience from DOE/ECP
  - Delivered via post-ECP organizations like PESO
  - And more…
- **Steps**
  - Translate science story to strategy and plan – leverage experience from ECP, others
  - ID node-level parallelization strategy – CUDA, HIP, DPC++, Kokkos, RAJA, OpenMP, others
  - Survey existing libraries and tools – Vendors, E4S, others
  - Explore available platforms – DOE Facilities, cloud, others
  - Leverage existing software ecosystem – containers, Spack, others
  - Leverage software communities – Product communities, communities of practice, others
  - Construct new codes within the broader ecosystem
  - Produce new science results

PESO: Partnering for Scientific Software Ecosystem Stewardship Opportunities

PESO Partnerships

Stakeholder Engagement and Consortium Partnerships

Consortium
- Spack support
- Products in E4S
- SW practices
- PIER activities

Applications Community
- Spack support
- E4S support

Computing Facilities
- Spack support
- E4S integration
- E4S user support

Commercial HPC Companies
- Engage in joint activities to advance the scientific software ecosystem
- Develop business models to further partnerships
- Increase release coordination activities

US Agencies
- Engage in business model discussions & plans for use of E4S
- Work with commercial providers to establish a support model
- Explore opportunities for joint product development

Industrial Users
- Engage in business model discussions & plans for use of E4S
- Work with commercial providers to establish a support model
- Explore models for mixed open-proprietary software stacks

Community Development

Broadening Participation of Underrepresented Groups in DOE Computing Sciences
- Coordinate consortium crosscutting-layer PIER planning
- Seek support for Sustainable Research Pathways Program
- Lead HPC Workforce Development and Retention Action Group

Better Scientific Software (BSSw) Fellowship Program
- Coordinate BSSw Fellowship Program – which gives recognition and funding to leaders and advocates of high-quality scientific software
- Seek sustainable support for BSSw Fellows, Honorable Mentions, travel, and program management for 2025 and beyond

PESO Services

Integration Partnerships

Provide resources and support for portfolio build, integration, and testing capabilities
- Spack integration
- CI testing
- Portfolio support & management

in collaboration with & co-funded by SSOs
- On-node & inter-node programming systems (w. S4PST)
- Math libraries, Data & viz, ML/AI (w. OASIS)
- Tools (w. STEP), Workflows (w. SWAS)
- NNSA software (funded by NNSA)

SQA & Security

Provide infrastructure to support and leverage product team SQA efforts
- Supply chain, Product quality
- Testing, Documentation

PESO Products

E4S
- Support for product integration
- E4S website
- Documentation, Training

Spack
- Features for consortium products
- Documentation, Training

Port & Test Platforms
- Frank test & development system
- Cloud resources
- Documentation, training

BSSw.io Content (w. COLABS)
- Short articles on topics related to scientific software productivity and sustainability (recruit, write, review, & edit)

PESO

Stakeholders:
Applications Community
Commercial HPC Companies
Industrial Users
US Agencies

DOE Computing Facilities:
ALCF NERSC
OLCF

CRLC: Computational Research Leadership Council: ANL, BNL, LBNL, LLNL, LANL, ORNL, PNNL, SNL

PESO Advisory Board
Reps from ANL, LBNL, LLNL, LANL, ORNL, SNL

S3C Consortium
PESO, COLABS, CORSA, OASIS, STEP, SWAS, S4PST

DOE Program Managers
ASCR: Hal Finkel, Ben Brown, Saswata Hier-Majumder, Robinson Pino, Bill Spotz, David Rabson
NNSA: Si Hammond

PESO: Partnering for Scientific Software Ecosystem Stewardship Opportunities

Mike Heroux, SNL - PI **Lois Curfman McInnes, ANL - Co-PI**

PESO Partnerships

Stakeholder Engagement
(Mike Heroux, SNL)

Partnerships Coordinator
(Terece Turton, LANL)

Strategic engagement with consortium partners, applications, facilities, industry and agencies (in collaboration with and co-funded by SSOs)

- William Godoy, ORNL, On-node programming systems (w. S4PST)
- Rajeev Thakur, ANL, Inter-node programming systems (w. S4PST)
- Sameer Shende, Univ of Oregon, Tools (w. STEP)
- Sherry Li, LBNL, Math libraries (w. OASIS)
- Berk Geveci, Kitware, Data and viz (w. OASIS)
- Lavanya Ramakrishnan, LBNL, Workflows (w. SWAS)
- Mahantesh Halappanavar, PNNL, AI/ML (w. OASIS)

Unfunded partners: Strategic engagement with NNSA, communities of practice, applications, facilities, industry, agencies

- David Bernholdt, ORNL, RSE engagement (funded by COLABS)
- Addi Malviya-Thakur, ORNL, Foundation engagement (funded by CORSA)
- Elaine Raybourn, SNL, Consortium-wide community development (funded by CORSA)
- Ulrike Yang, LLNL, NNSA software (funded by NNSA)
- Partners at ALCF, NERSC, OLCF (funded by facilities, SW integration)

Community Development
(Lois Curfman McInnes, ANL)

Broadening Participation Initiative

- Mary Ann Leung, Sustainable Horizons Institute, PIER planning, lead of Sustainable Research Pathways (SRP)
- Daniel Martin, LBNL, lab lead of Sustainable Research Pathways
- Suzanne Parete-Koon, ORNL, lead of HPC Workforce Development and Retention Action Group

Better Scientific Software (BSSw) Fellowship Program

- Elsa Gonsiorowski, LLNL, Coordinator of BSSw Fellowship Program
- Erik Palmer, LBNL, Deputy Coordinator of BSSw Fellowship Program

PESO Services

Integration Coordinator
(Jim Willenbring, SNL)

Software portfolio management and integration (in collaboration with and co-funded by SSOs)

- Damien Lebrun-Grandie, ORNL, On-node prog systems (w. S4PST)
- Hui Zhou, ANL, Inter-node programming systems (w. S4PST)
- Bill Hoffman, Kitware, Tools (w. STEP)
- Satish Balay, ANL, Math libs (w. OASIS)
- Patrick O'Leary, Kitware, Data & viz (w. OASIS)
- Matteo Turilli, BNL, Workflows (w. SWAS)
- Sam Browne, SNL, NNSA software (funded by NNSA)

SQA & Security
(David Bernholdt, ORNL)

- Ross Bartlett (SNL)
- Berk Geveci (Kitware)
- Jim Willenbring (SNL)

PESO Products

E4S
(Sameer Shende, U Oregon)

- Luke Peyralans, Erik Keever, Wyatt Spear, Jordi Rodriguez

Spack (Todd Gamblin, LLNL)

- Greg Becker, LLNL
- Tammy Dahlgren, LLNL

Port & Test Platforms
(Gamblin & Shende)

- In partnership with Univ of Oregon, Cloud, etc.

BSSw.io Content (w. COLABS)

- Ross Bartlett, SNL
- Keith Beattie, LBNL
- Patricia Grubel, LANL
- Mark Miller, LLNL

Strategy & Integration – Members are part of other SSO teams, NNSA, for tight collaboration

PESO

SWOT ANALYSIS GOING FORWARD

Strengths

- Successfully completed $1.8B, 7-year US Exascale Computing Project
- Rich, growing software stack in E4S and Spack from ECP investments
- Progress due to a hierarchical, process-driven org structure
- Evolving to a resilient and adaptable horizontal, community-driven constellation of orgs
- De-risked a path for many others to pursue 100X

Weaknesses

- Successfully **completed** $1.8B, 7-year US Exascale Computing Project
- Rich, growing software stack in E4S and Spack **that needs investment after ECP**
- Progress **needs to continue without** a hierarchical, process-driven org structure
- **Evolving** to a resilient and adaptable horizontal, community-driven constellation of orgs
- De-risked a path for many others to pursue 100X **but realizing it still very challenging**

Threats

- **The end of ECP signals that Exascale was achieved. Mission accomplished.**
- The **transition** from hierarchical (ECP) org model to collaborative (post-ECP) model **won't happen quickly enough.** With **little margin for error,** we won't acquire community support needed for ecosystem efforts after ECP ends

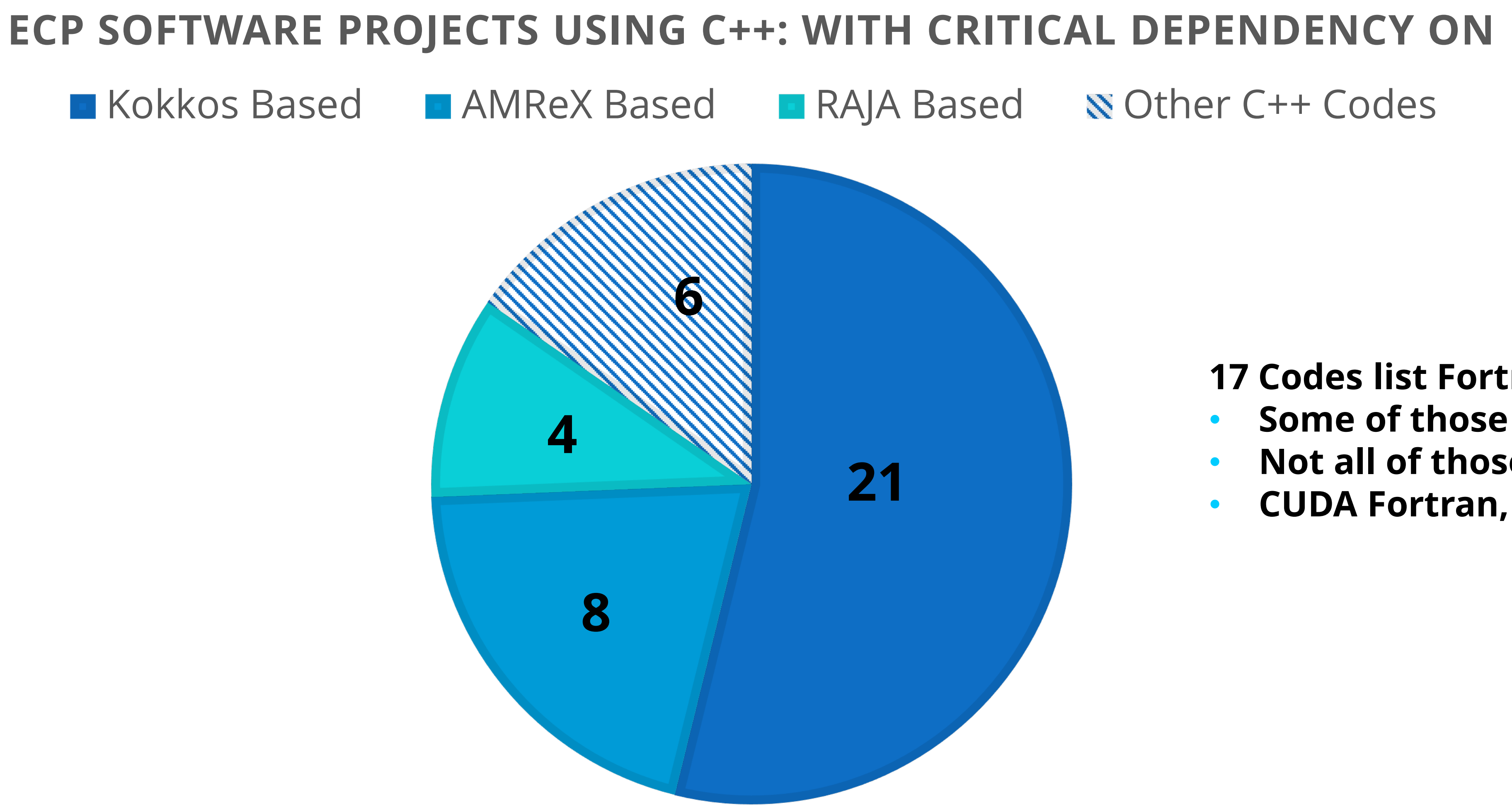
Opportunities

- See next slide

POST-ECP

- DOE/ECP has **learned a lot about producing software contributions** to the HPC community:
  - Improved planning, executing, tracking, assessing, integrating, and delivering
  - Improved interactions with the broader HPC software and hardware community
  - Direct engagement with industry, US agencies, and international collaborators
- In post-ECP efforts **we propose to continue and expand these efforts:**
  - Further engage with commercial partners to provide a rich, robust software ecosystem
  - Evolve a stable, sustainable business model for engaging with agencies and industry
  - Engage with cloud providers, software foundations, and others to optimize cost & benefit sharing
  - Further the ECP strategy for direct industry and agency engagement
- We intend to **realize the potential of the ECP legacy across the HPC community:**
  - **Realize the "100X" potential** by transferring scientific computations to accelerated architectures
  - **Increase the trustworthiness, sustainability, and cost effectiveness** of our software in the future
- **We want to work with the HPC community to realize the legacy of ECP, and beyond**
  - We have many new means to interact
  - Many new opportunities to pursue

DOE EXASCALE COMPUTING PROJECT

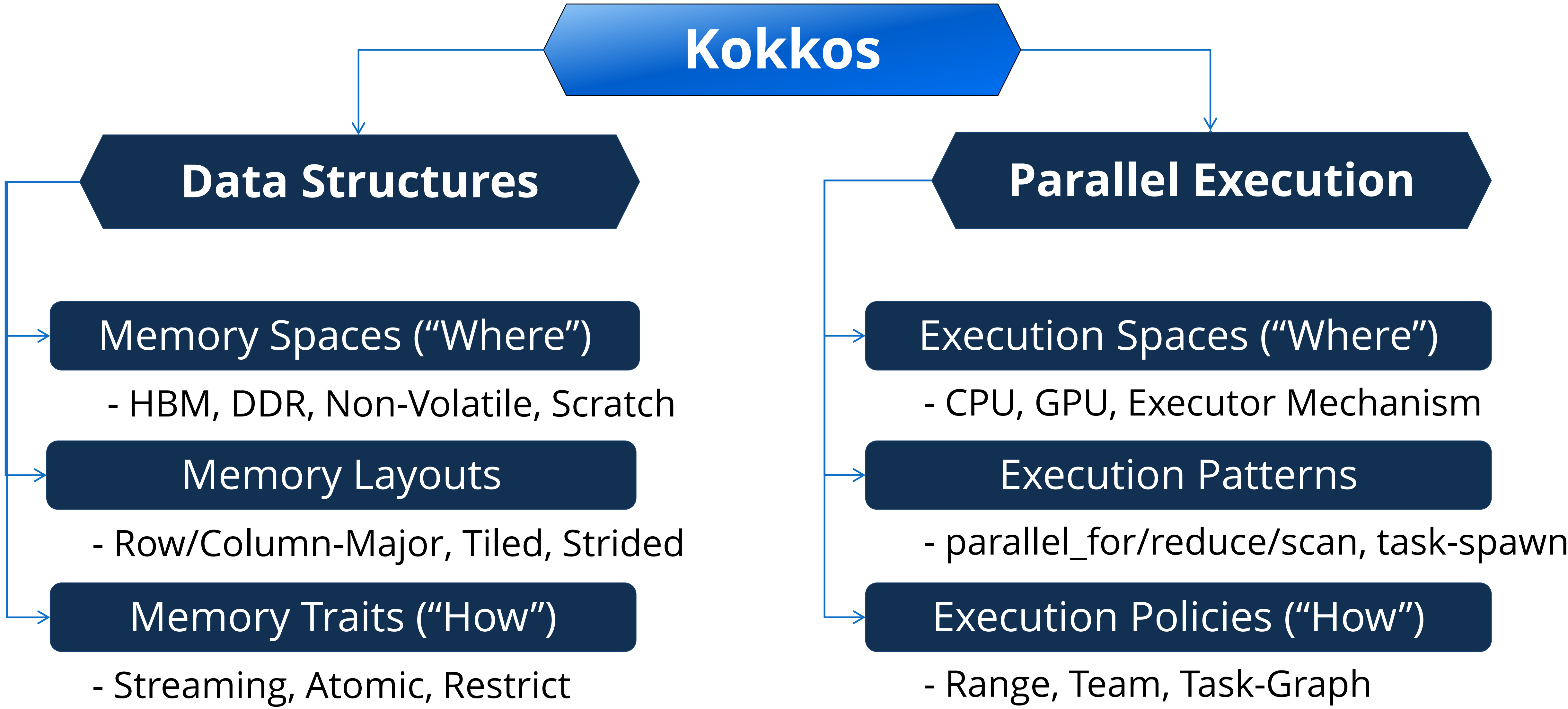


17 Codes list Fortran as critical

- **Some of those actually moved on**
- **Not all of those run on GPUs**
- **CUDA Fortran, OpenMP, OpenACC**

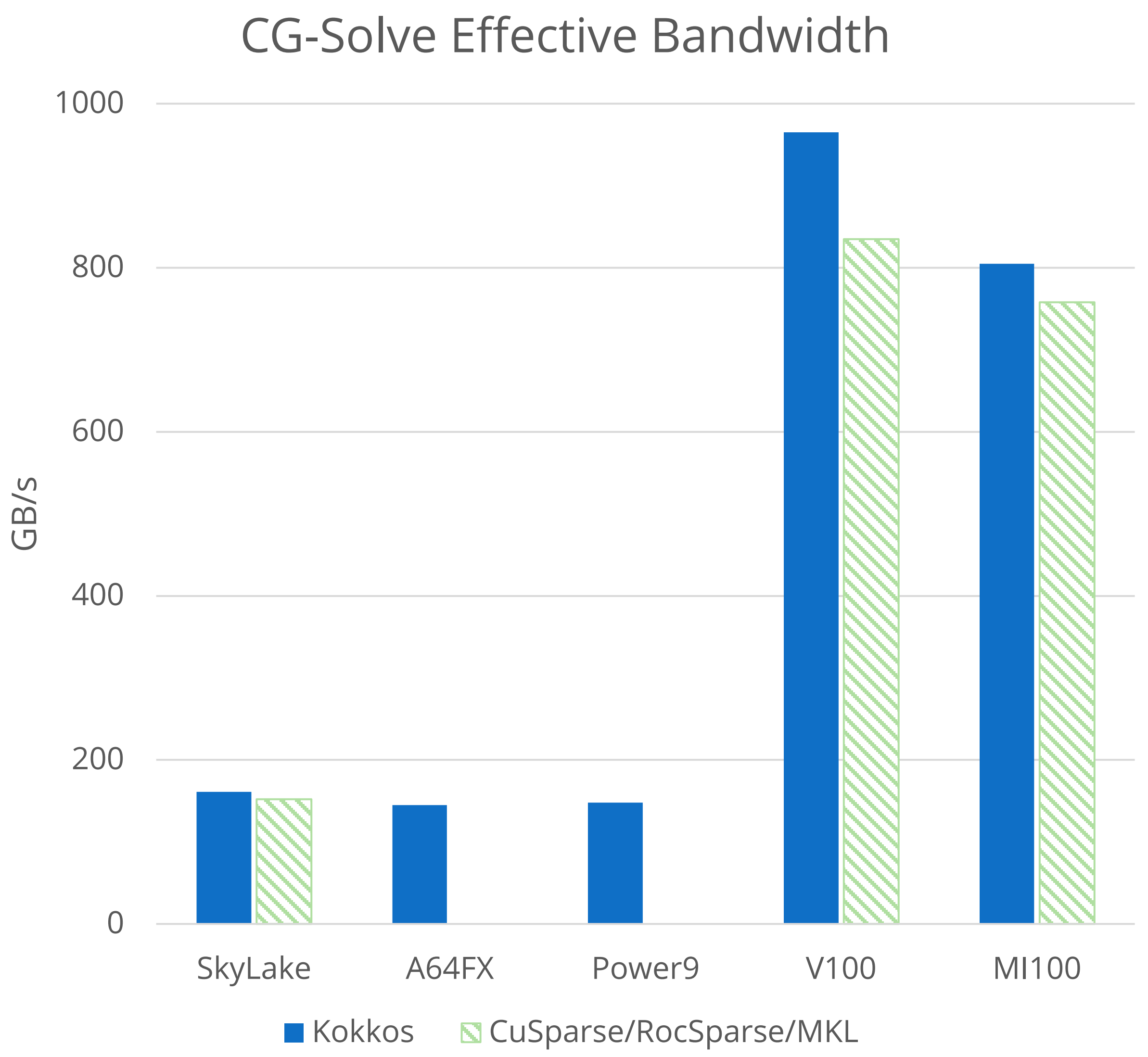
Large Majority of C++ Codes Chose Abstraction over Vendor Models!

KOKKOS CORE ABSTRACTIONS



CG SOLVE PERFORMANCE 2022

- CG-Solve as discussed above
- Also try replacing SPMV with TPL
- Running 100x100x100 heat conduction problem
  - "MiniFE" Proxyapp setup
- Measure effective Bandwidth
  - Algorithmical memory ops per time
- Why is this beating vendor libs?
  - Its complicated, but a real effect



KOKKOS SUPPORT

- The Kokkos Lectures
  - 8 lectures covering most aspects of Kokkos
  - 15 hours of recordings
  - > 500 slides
  - >20 exercises
- Extensive Wiki
  - API Reference
  - Programming Guide
- Slack as primary direct support

- Module 1: Introduction
  - Introduction, Basic Parallelism, Build System
- Module 2: Views and Spaces
  - Execution and Memory Spaces, Data Layout
- Module 3: Data Structures and MDRangePolicy
  - Tightly Nested Loops, Subviews, ScatterView,…
- Module 4: Hierarchical Parallelism
  - Nested Parallelism, Scratch Pads, Unique Token
- Module 5: Advanced Optimizations
  - Streams, Tasking and SIMD
- Module 6: Language Interoperability
  - Fortran, Python, MPI and PGAS
- Module 7: Tools
  - Profiling, Tuning , Debugging, Static Analysis
- Module 8: Kokkos Kernels
  - Dense LA, Sparse LA, Solvers, Graph Kernels

https://kokkos.link/the-lectures

KOKKOS KERNELS

- BLAS, Sparse and Graph Kernels on top of Kokkos and its View abstraction
  - Scalar type agnostic, e.g. works for any types with math operators
  - Layout and Memory Space aware
- Can call vendor libraries when available
- Views contain size and stride information => Interface is simpler

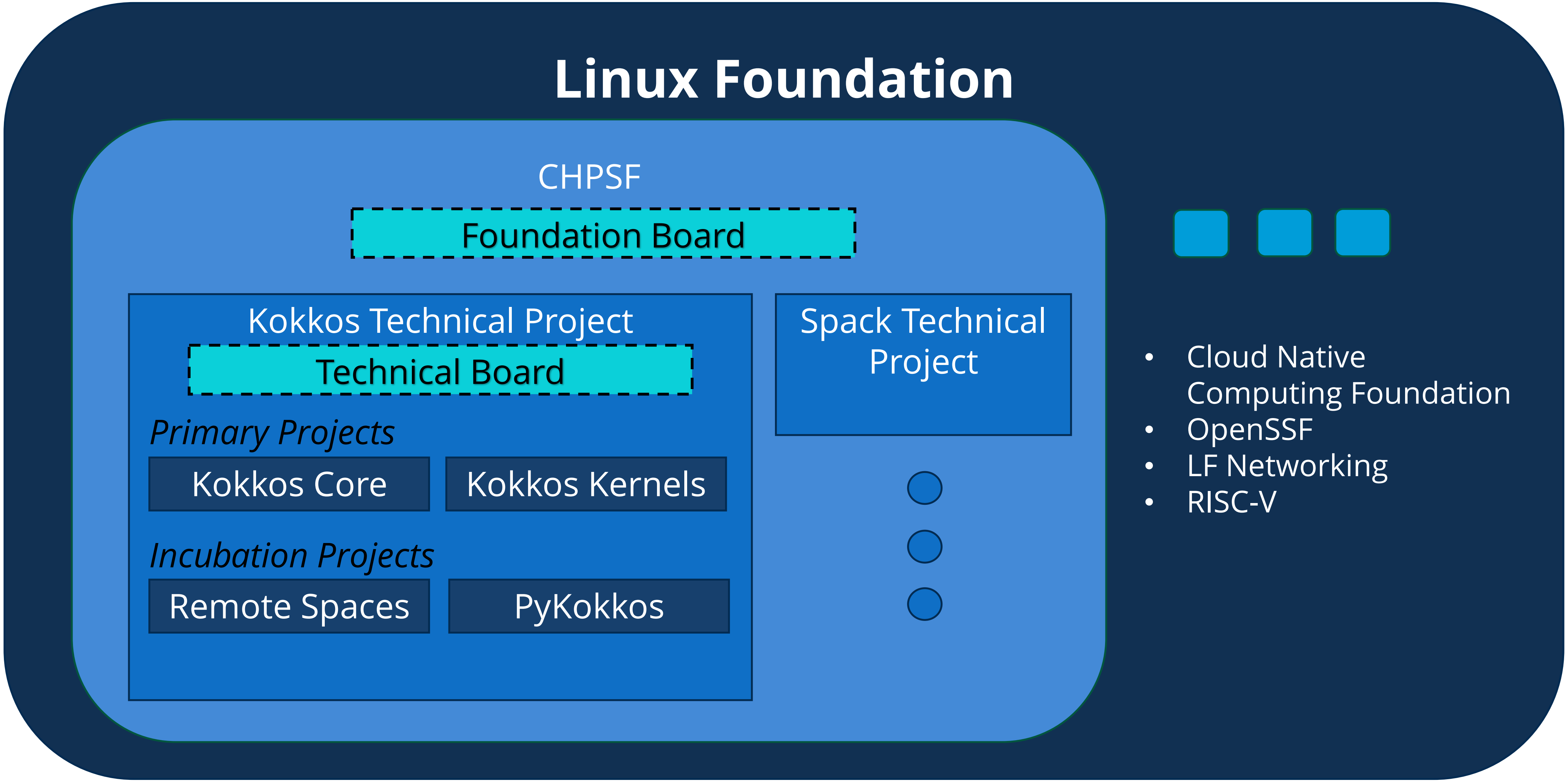
```
// BLAS
int M, N, K, LDA, LDB; double alpha, beta; double *A, *B, *C;
dgemm('N','N', M, N, K, alpha, A, LDA, B, LDB, beta, C, LDC);
```

vs.

```
// Kokkos Kernels
double alpha, beta; View<double**> A,B,C;
gemm('N','N', alpha, A, B, beta, C);
```

CORE HIGH PERFORMANCE SOFTWARE FOUNDATION

- **Primary Goal:** *Enable true partnerships on Kokkos via open governance.*



MORE FOUNDATION DETAILS

- Primary Organizers: Todd Gamblin (Spack) and Christian Trott (Kokkos)
  - ***Sponsors:***
    - *US Labs:* SNL, LLNL, ORNL, LBL, LANL and ANL (
    - *Tentative Industry:* HPE, Amazon, NVIDIA, AMD and Intel
    - *Others:* CEA, maybe Rieken, CSCS, Julich? Open to other interested parties
  - ***Scope:***
    - Help sustain critical software technologies for HPC and related areas
  - ***What does it do:***
    - Guarantees Open Governances (e.g. github.com/kokkos will be owned by LF)
    - Help organize and pay for meetings and workshops
    - Help finance and organize project infrastructure (e.g. Slack)

FOUNDATION TIMELINE

- *November 2023:* Announcement for intend to launch foundation
    - Describes goals and initial partners
    - Invitation for interested partners to participate in finalization of foundation charter
    - Potentially add more initial technical projects
- *By April 2024:* Spack, Kokkos etc. become ”Unfunded technical projects” in Linux Foundation
    - Sets up Open Governance
    - Transfer assets (github organization, trademark, URLs)
    - Copyright stays with contributing organizations!
    - No change in license
- *May 2024 (ISC):* Foundation is up and running
    - Enables donations to go reach this effort

IDEAS FOR FUTURE DIRECTIONS OF KOKKOS

- **Edge computing / Embedded Support**
  - Many of the same concerns as HPC – resource constraint, performance critical
  - Many different devices including FPGAs
- **Programming Language Safety**
  - More concern about cyber security – how do we write safer code?
  - Kokkos data abstractions (View/mdspan/mdarray) enable safer encapsulation – could make it almost impossible to have out of bounds memory access
  - Combined with static analysis could be significant step to enable C++ codes which are memory safe by design
- **More Asynchronous Dependency Based Software design**
  - Incorporate ISO C++ Senders/Receivers interface
- **Better integration with distributed computing**
  - Remote spaces
  - MPI interface taking Kokkos data structures

THE BOTTOM LINE

- Exascale systems successfully delivered
- Applications improved by factors of hundreds to thousands relative to 2016 performance
- Dramatically reduced energy use per flop
- Department of Energy (DOE) investments significantly de-risked the broader adoption of these technologies by other agencies and industry
- Thousands of users already benefiting from exascale science and software

SUSTAINABLE CAPABILITY INTEGRATION

- **Capability:** Any significant product functionality, including existing features adapted to the pre-exascale and exascale environments, that can be integrated into a client environment
  - Approximately 1 FTE-year cost
  - Single capability intended to be broadly useful

- **Integration:** Complete, sustainable integration of a significant product capability into a client environment
  - Integration typically includes availability via E4S
  - Example: New PETSc solver integrated into WDMApp on Frontier – 1 integration
  - Target integrations include AD/ST codes, tool usage, facility deployment, community standards, vendor deployment

- **Sustainable:** Integrated into the intended home for testing, integration, deployment, support

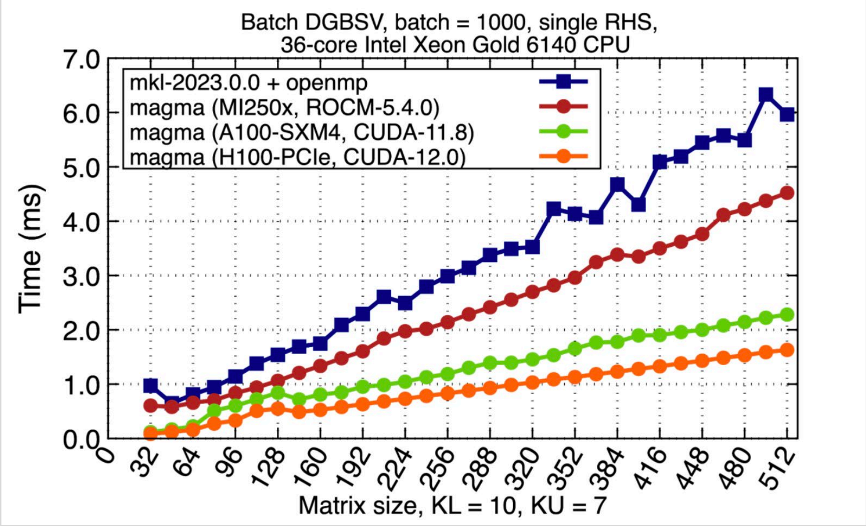
BATCHED SPARSE LINEAR SYSTEM SOLUTION

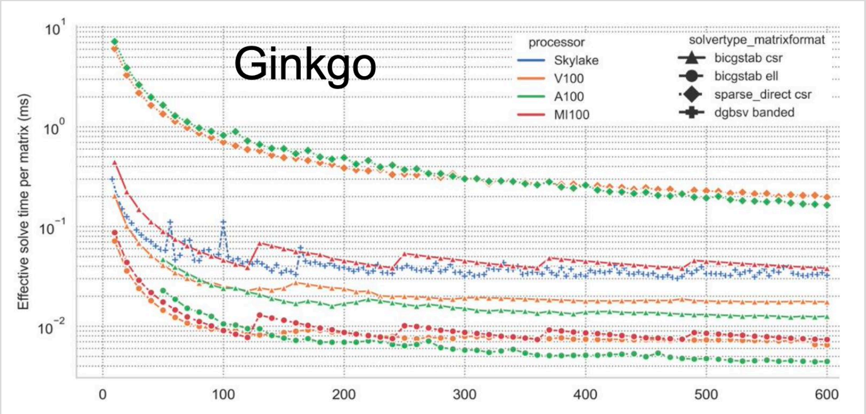
- **Challenges**:
  - Many applications require simultaneous solution of many small linear systems of equations that are structurally sparse (e.g. reactions in combustion applications)
  - Difficult to achieve high GPU utilization in non-batched form
  - Need to design appropriate interfaces to be both functionally efficient and performance portable

- **ECP Scope**:
  - Interface design for accessing batched solvers
  - Develop sparse batched kernel codes for solvers and preconditioners
  - Combine with mixed precision efforts to improve performance
  - Support GPU accelerators from AMD, Intel and NVIDIA; vendor relationships critical to a successful outcome

- **Libraries**: MAGMA, Gingko, hypre, KokkosKernels, PETSc, SuperLU

Batch Band Factorization in Magma on three different GPUs



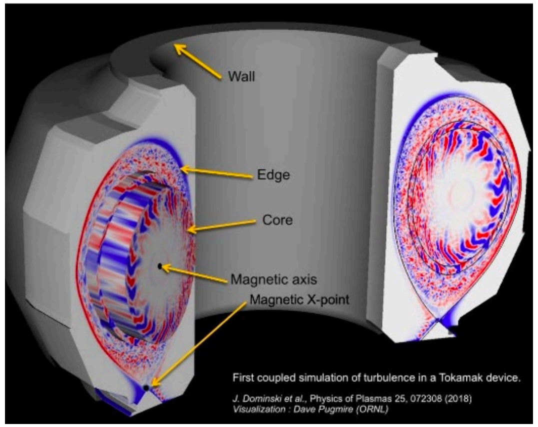


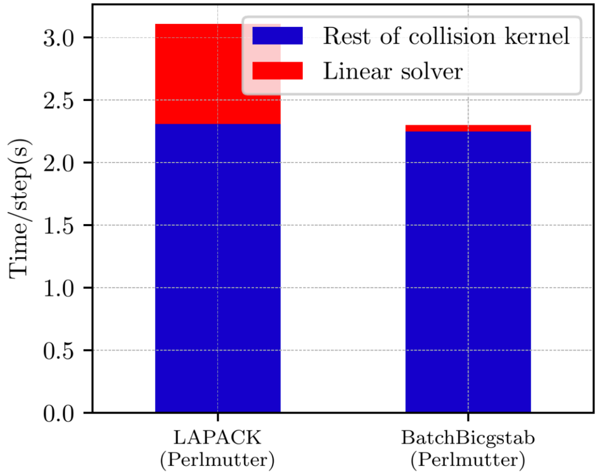
BATCHED SOLUTION TECHNIQUES SHOW SIGNIFICANT IMPROVEMENTS

FUSION

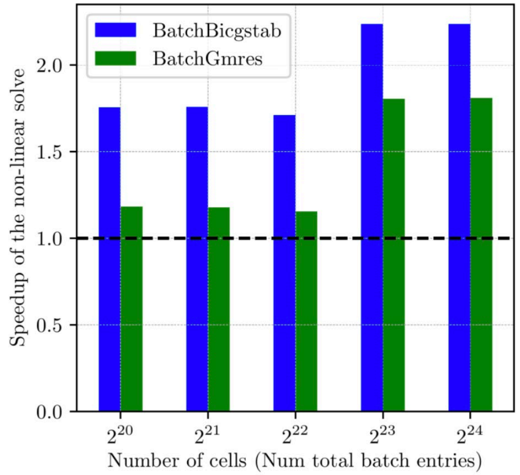
Fast, fully conservative Landau collision operator is developed in structure-preserving methods for fusion plasma. It is used in the recent 10 species models optimized for WDMApp milestone of ITER simulations with tungsten impurities.

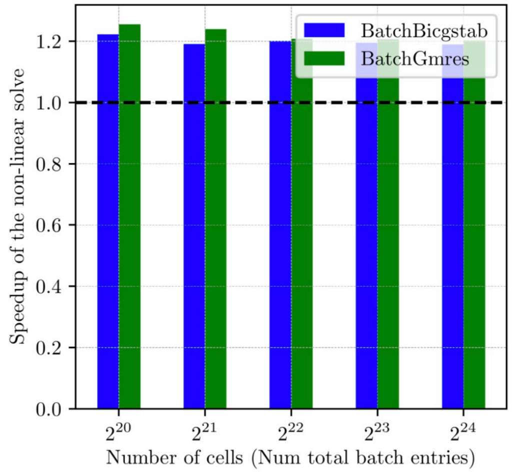
Time/step of batched **Gingko GPU solvers** integrated into XCG fusion application reduce linear solver time compared to LAPACK on the multicore CPU (AMD EPYC)





COMBUSTION





Performance studies with the Pele combustion application using **Gingko batched solvers and SUNDIALS** for small problem sizes. Enabled speedups of almost 2x for some reaction mechanisms on AMD MI250X GPUs (Frontier).