DOE Exascale Computing Project – The Final Update

Lori Diachin, LLNL, Project Director

ASCAC Meeting
Washington, DC
May 29, 2024

The Exascale Computing Project has successfully delivered – on time, under budget and far exceeding expectations!

Maintain **international leadership in HPC**

- ❖ Frontier is the world's first exascale machine – in part due to ECP/ECI investments
- ❖ Aurora is the second exascale system and the fastest AI mixed-precision machine
- ❖ 1000+ researchers trained in GPU computing

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

- ❖ Six vendors funded under PathForward; outcomes realized in exascale systems
- ❖ Accelerator-based computing lowers cost of energy across the board
- ❖ The ECP Industry and Agency Council stimulates consumption of HPC resources

Deliver a long-term, **sustainable software ecosystem** that can be used and maintained for years to come

- ❖ E4S deployed at HPC facilities around the US and the world
- ❖ 76 ECP-funded HPC products available for computing at all scales
- ❖ Performance portability tools developed and widely used

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

- ❖ ECP applications demonstrate outstanding performance and capabilities at exascale
- ❖ Previously unattainable results in real-world challenge problems
- ❖ ECP lessons learned pave the way for many additional applications to leverage accelerator-based computing

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
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

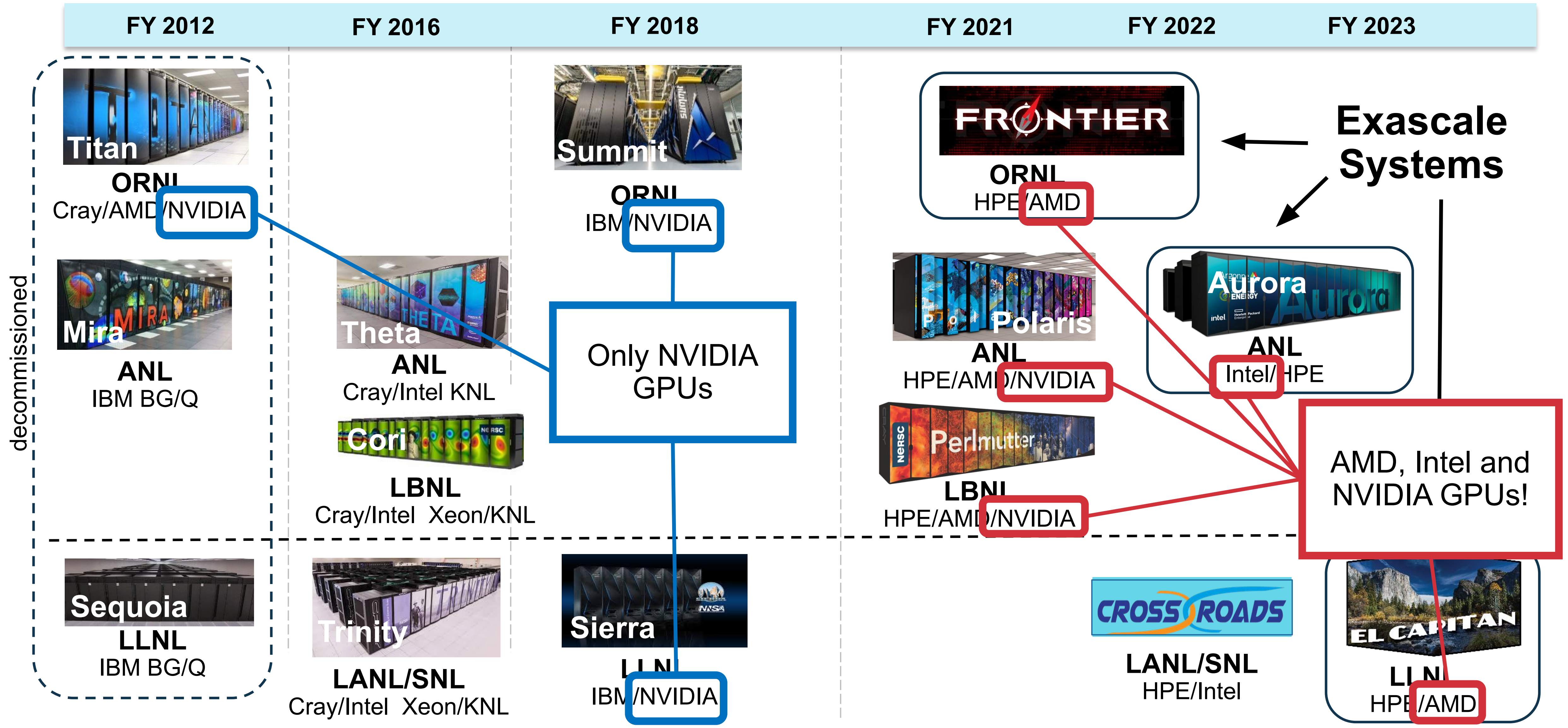

Richard Gerber, HI Director
Susan Coghlan, HI Deputy Director

EXASCALE COMPUTING PROJECT

ECP's success was measured using Key Performance Parameters for applications, software and impact on exascale systems

KPP ID	Description of Scope	Threshold KPP	Objective KPP	Completion Status (02/05/2024)
KPP-1	**11 selected applications** demonstrate performance improvement for mission-critical problems ✔	**6 of 11** applications demonstrate **Figure of Merit improvement ≥50** on their base challenge problem	**All 11** selected applications demonstrate their **stretch challenge problem**	✔ **9 SC apps complete**
KPP-2	**14 selected applications** broaden the reach of exascale science and mission capability ✔	**5 of 10** DOE **Science and Applied Energy** applications ***and*** **2 of 4 NNSA** applications demonstrate their base challenge problem	**All 14** selected applications demonstrate their **stretch challenge problem**	✔ **7 SC apps complete** ✔ **3 ATDM apps complete**
KPP-3	**76 software products** selected to meet an aggregate capability integration score ✔	Software products achieve an aggregate capability integration score of at least **34 out of a possible score of 68 points**	Software products **achieve the maximum** aggregate capability integration **score of 68 points**	✔ **55 integration points complete**
KPP-4	Delivery of **267 vendor baselined milestones** in the PathForward element ✔	Vendors meet **214 out of** the total possible **267** PathForward **milestones** ✔	Vendors meet **all 267** possible PathForward **milestones**	✔ **267 PathForward milestones complete**

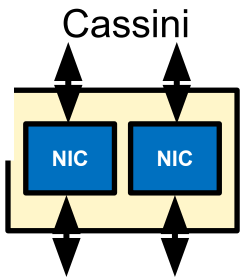
HPC systems have come a long way since ECP’s inception

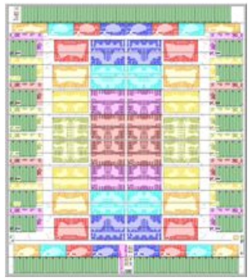


The *Forward programs had several payoffs in exascale systems

Design Forward
DOE

Slingshot Network Architecture





Design Forward 2
DOE
System Design

Slingshot Network Architecture and Implementation

Shasta Implementation

Energy Efficiency

- Frontier TDS #7 on Green 500; Frontier #13
- 7 of top 20 systems based on AMD GPUs

PathForward
DOE

Scalable in-fabric high performance storage

Aurora

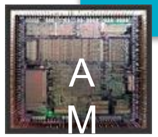
FRONTIER

Fast Forward 1&2
DOE

DAOS Object Store Architecture

AMD Node Architecture

AMD Node Architecture and Implementation



EL CAPITAN
ADVANCED SIMULATION & COMPUTING

EXASCALE COMPUTING PROJECT

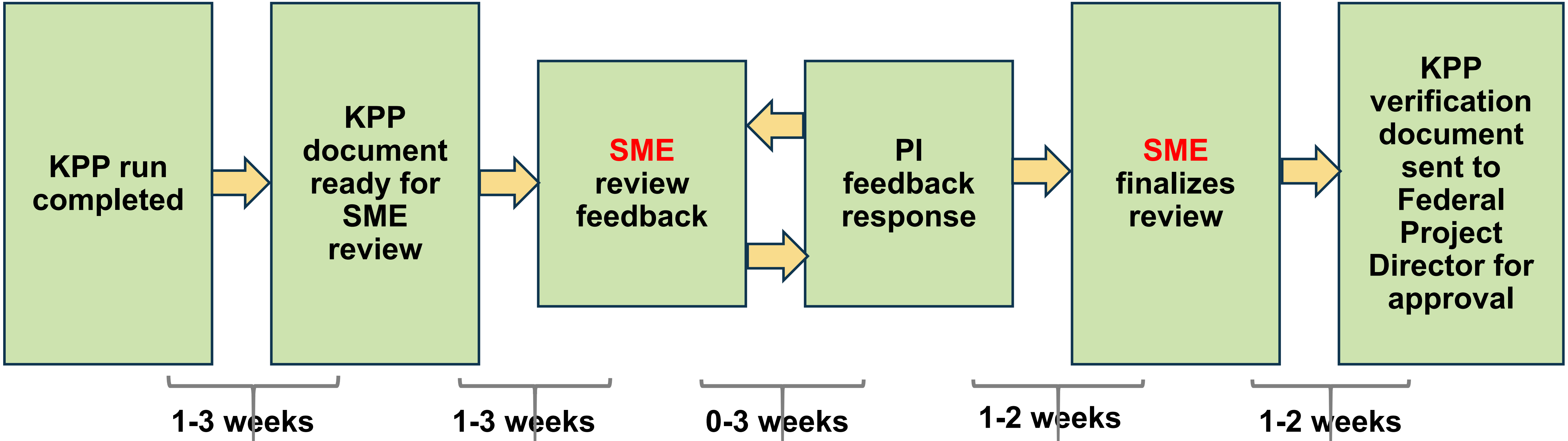
Several exascale capabilities are sited at the DOE facilities and used by ECP teams

- #1 on the Top 500 list since May 2022: Currently 1.19 exaflops
- #2 on HPL-MxP mixed-precision benchmark with 10.2 exaflops of AI performance
- Full system available to all ECP teams since April 2023
- Excellent performance: 7 Gordon Bell & Special Prize Finalists to-date
- Many first-of-a-kind science goals achieved

- Full system installation completed June 2023
- #2 on the Top 500 list (May 2024),. Currently 1.01 exaflops using 87% of the machine
- #1 on HPL-MxP mixed-precision benchmark with 10.6 exaflops of AI performance
- Limited ECP access available July 2023 (ANL personnel); full system access for ECP November 2023

- Hardware installation at LLNL is underway; network, cooling, cabinets all in place – blades being installed now
- 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

KPP verification involved tens of external subject matter experts and was a time intensive, closely tracked process



Example: KPP-1 and KPP-2 verification timeline; similar for KPP-3

KPP completion was closely tracked (example KPP-1)

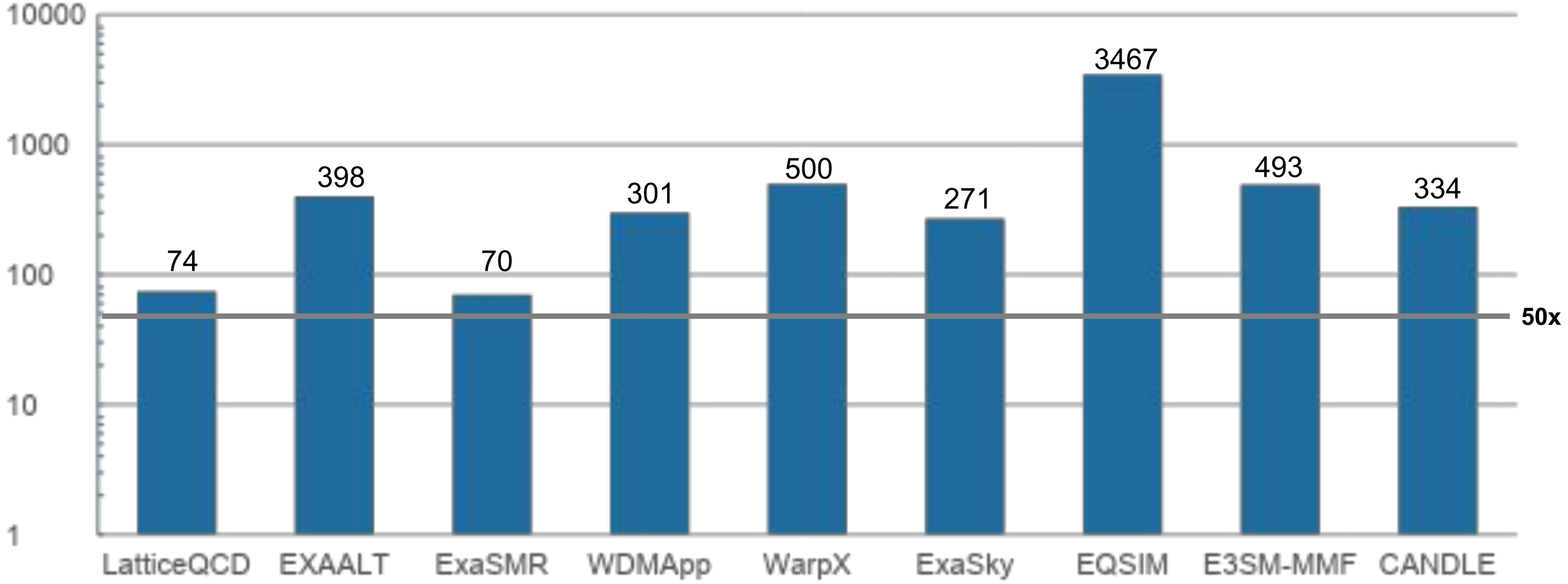
KPPs met: 10/11
Threshold: 6/11

KPP in progress	KPP completed	KPP submitted	KPP signed (SME)	KPP signed (FPD)
LatticeQCD				
NWChemEx				
EXAALT				
QMCPACK				
ExaSMR				
WDMApp				
WarpX				
ExaSky				
EQSIM				
E3SM-MMF				
CANDLE				

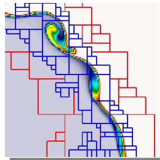
April 2024

ECP application results exceeded expectations!

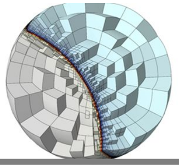
9 out of 11 KPP-1 projects surpassed ambitious 50x performance target



AMReX

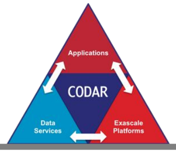


CEED

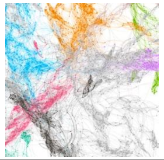


CoPA

CODAR



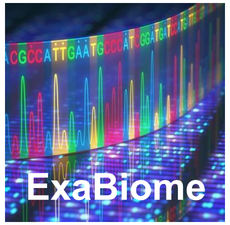
ExaGraph

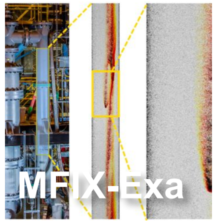


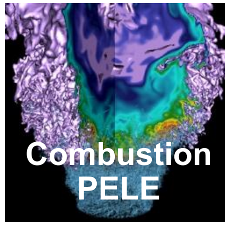
ExaLearn

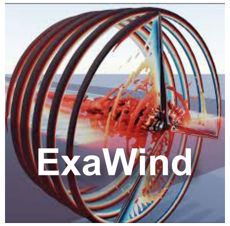
Codesign played a critical role

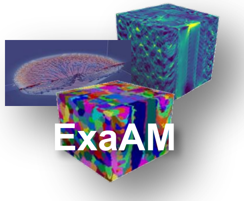
7 out of 10 SC KPP-2 projects demonstrated exascale capability

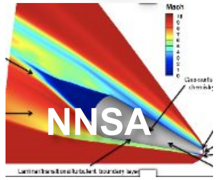












3 out of 4 NNSA KPP-2 projects demonstrated exascale capability

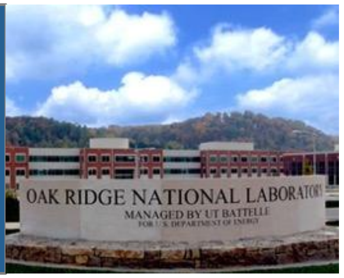
50x

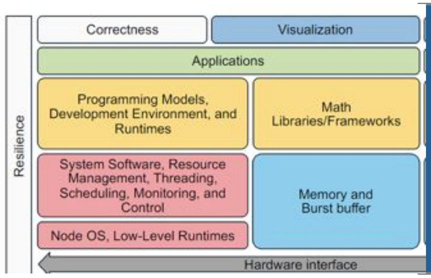
ECP EXASCALE COMPUTING PROJECT

Partnership with ECP facility-embedded performance engineers enabled applications and improved the ecosystem for all

Collaborated via workshops, hackathons, & dungeon sessions to improve application readiness and performance on Aurora, Frontier & Perlmutter.

Collaborated with AD, ST, Facility staff, & vendors through the Facilities' Centers of Excellence on porting codes to exascale programming models.



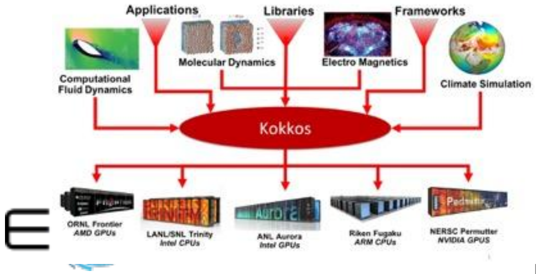


Tested applications with the exascale system software stacks and helped them perform at high performance at scale in each environment.

Created and tested proxy apps used as easy-to-run and test surrogates for key parts of larger codes.

Identified, reported, and tracked through resolution many bugs in the software stack, thus improving the environment for all users.

Created work-arounds for codes that enabled them to run on Frontier, Aurora, and proxy systems, while waiting for bugs or other issues to be addressed.



Helped implement and test portable programming paradigms like Kokkos, Raja, OpenMP.

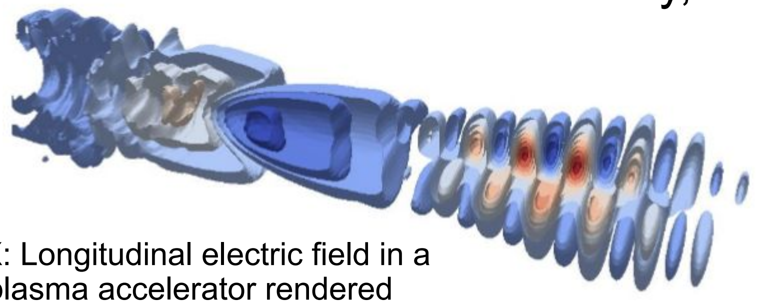
Managed contracts with vendors to implement HIP for Aurora and SYCL on Frontier to enable application portability.

One example of success WarpX: co-design enabled computational design of next-generation plasma-based accelerators

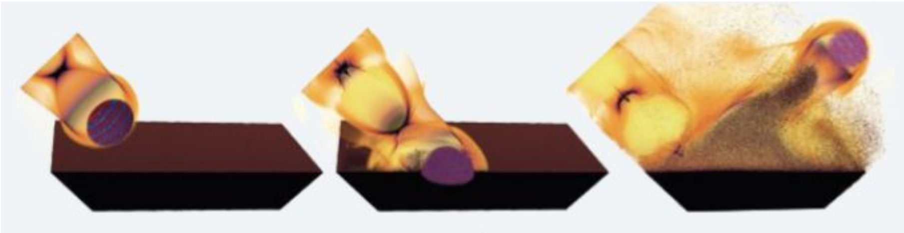
PI: Jean-Luc Vay, LBNL

- **Objective**: Model chains of plasma-based particle accelerators for future high-energy physics colliders

- **ECP accomplishments**
  - New exascale code capability **built on AMReX adaptive mesh refinement library** incorporated full physics and first-principles models in three dimensions
  - Worked with AMReX team to optimize backend code to run on four of fastest supercomputers in the world (Frontier, Fugaku, Summit, Perlmutter); 2022 Gordon Bell winner
  - **Enabled first-of-a-kind “ground truth” 3D simulations** to go from one to tens of consecutive multi-GeV laser-driven plasma stages allowing first computational prototypes of many-stage plasma accelerators
  - Achieved over **500X performance improvement on Frontier** compared to baseline ECP calculation in 2016



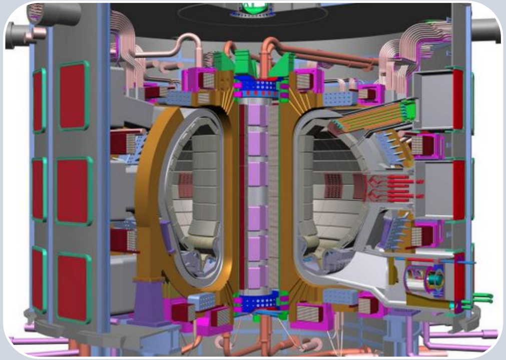
WarpX: Longitudinal electric field in a laser-plasma accelerator rendered with ECP software libraries as the simulation was running

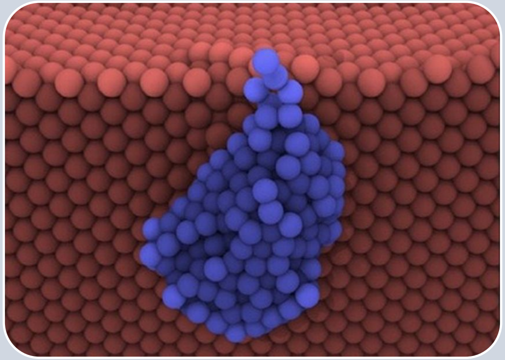


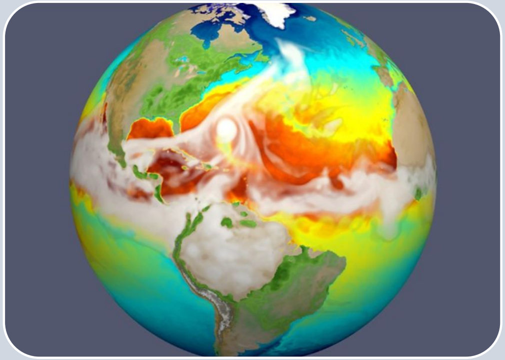
Enables models of novel e- beam injection processes to deliver highly-charged, short duration pulses without damaging other tissues to improve cancer treatment options

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

Four ECP technical teams worked to demonstrated the promise of AI / ML on Frontier through March 31, 2024







ExaWind

Use high resolution models run on Frontier as the basis for **surrogate models for unsteady loads** on wind turbine blades to be deployed in engineering design codes

ExaLearn/Fusion

Use machine learning to enhance the **control of plasmas in tokamaks** by predicting disruptions using fast surrogate models derived from high resolution runs on Frontier. Goal is to test in real time on DIII-D tokamak.

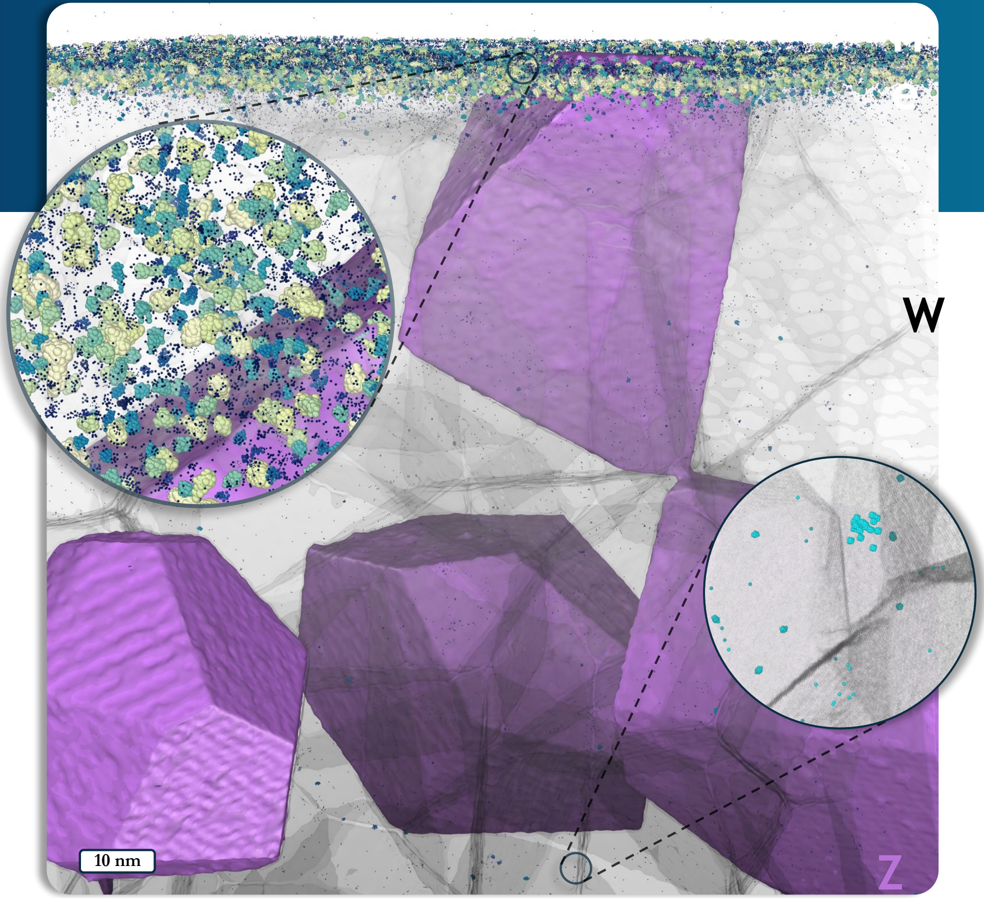
EXAALT

Use machine-learned potentials derived from high precision quantum simulations in molecular dynamics models run on Frontier to **predict survivability of plasma facing materials** proposed for ITER

Climate

Use machine learning for **automatic rapid tuning** of the 5 to 7 most important parameters used in subgrid scale models related to clouds. High resolution runs performed using the SCREAM Gordon Bell winning code.

ITER fusion reactor materials selected using machine learning



To operate safely and efficiently, fusion power plants must be designed with materials that can withstand the extreme heat and particle loads generated by the fusion process.

Large-scale and accurate molecular dynamics simulations of plasma-wall interactions are needed to identify new materials that can survive
under the extreme conditions of fusion reactors.

- **Challenge:** Current approaches to virtual assessment are too slow and expensive, because they rely on human physical intuition, domain expertise, and trial and error.
- **Goal:** Demonstrate fast virtual assessment of new materials, from the training of accurate machine-learned models to large-scale, long-time exascale simulations.

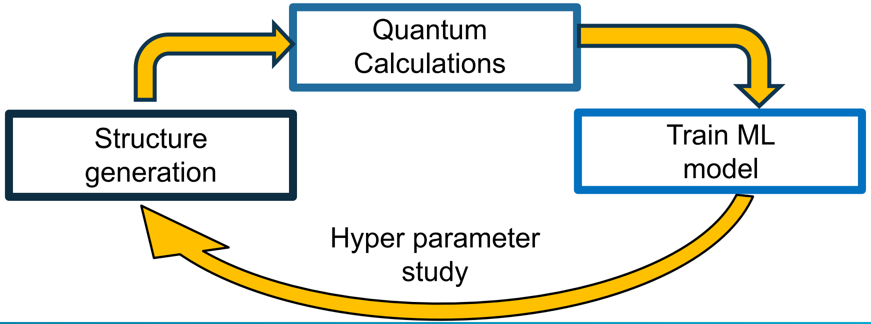
Materials scientists and engineers can use this capability to assess the survivability of candidate plasma-facing materials proposed for the ITER tokamak reactor, and potentially dramatically accelerate the process.

EXASCALE COMPUTING PROJECT

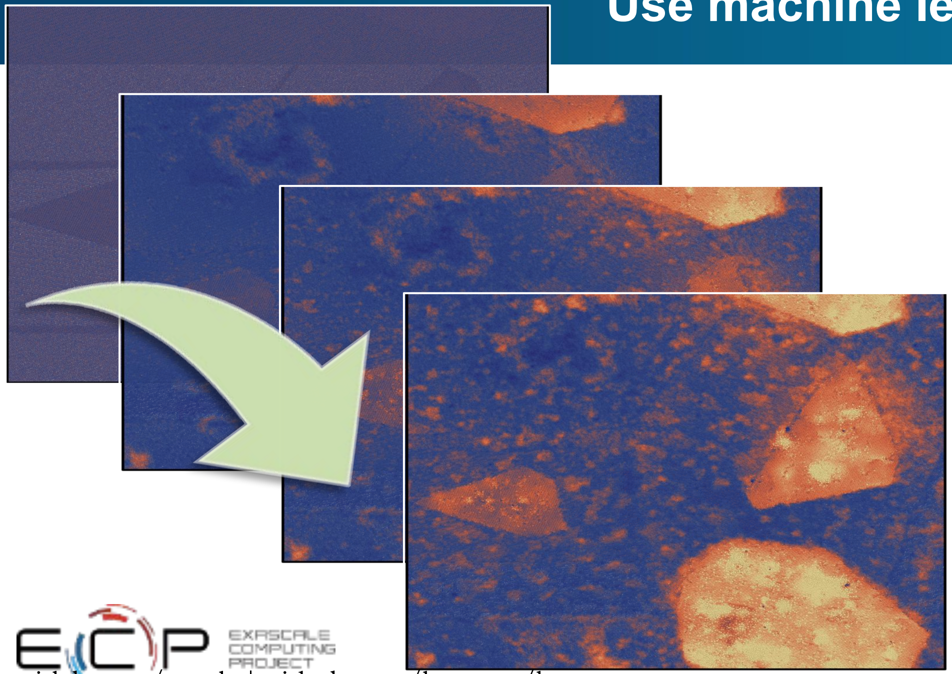
ECP funded a sprint effort to demonstrate use of ML in two ways to identify candidate materials for ITER

Develop autonomous HPC workflow that trains ML models of new materials from scratch

Demonstrated the training of accurate models of tungsten/rhenium and tungsten/osmium alloys *in a single day* on NERSC/Perlmutter. Turned weeks of development into hours. No human intervention required!



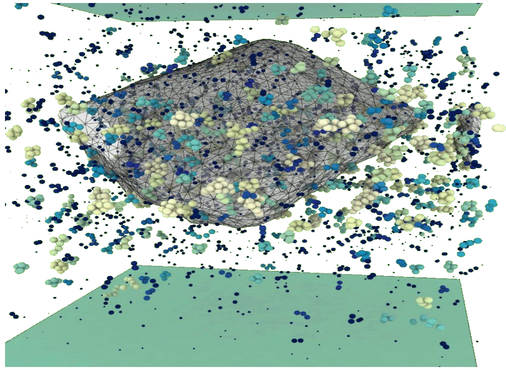
Use machine learned potentials in exascale simulations to evaluate materials



(Left) Simulation codes are now fast enough to study surface erosion from thermal loads and high-energy neutrons.

(Right) He ash creates detrimental porosity, removing the benefit of ZrC additions in tungsten plasma facing components.

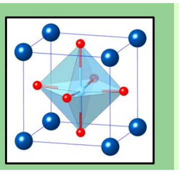
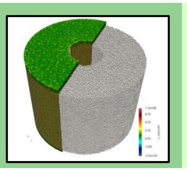
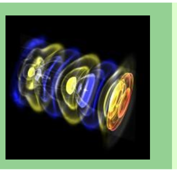
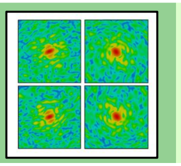
Sprint effort captured these phenomena in a candidate material using a mere four Frontier-days.



EXASCALE COMPUTING PROJECT

gitlab.com/exaalt | github.com/lammps/lammps

ECP applications would not have been successful without close collaboration with software technology developers

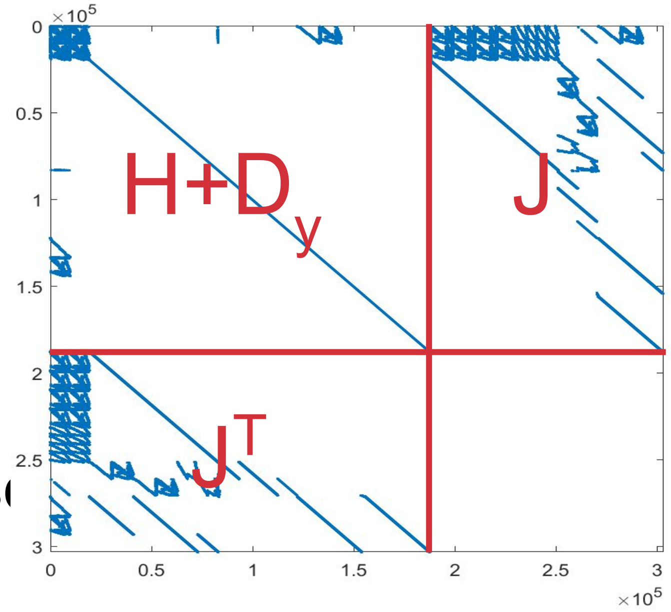
ExaWind: advanced wind farm modeling 	Sparse linear solvers in hypre and Trilinos optimized for strong scaling and GPU performance
QMCPACK: quantum Monte Carlo for materials 	Batched dense linear algebra kernels significantly improved GPU performance
ExaSGD: power grid optimization 	Optimize sparse indefinite 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

ExaSGD helps ensure a more reliable and robust power grid to better leverage renewable energy sources

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 spars systems of equations were required

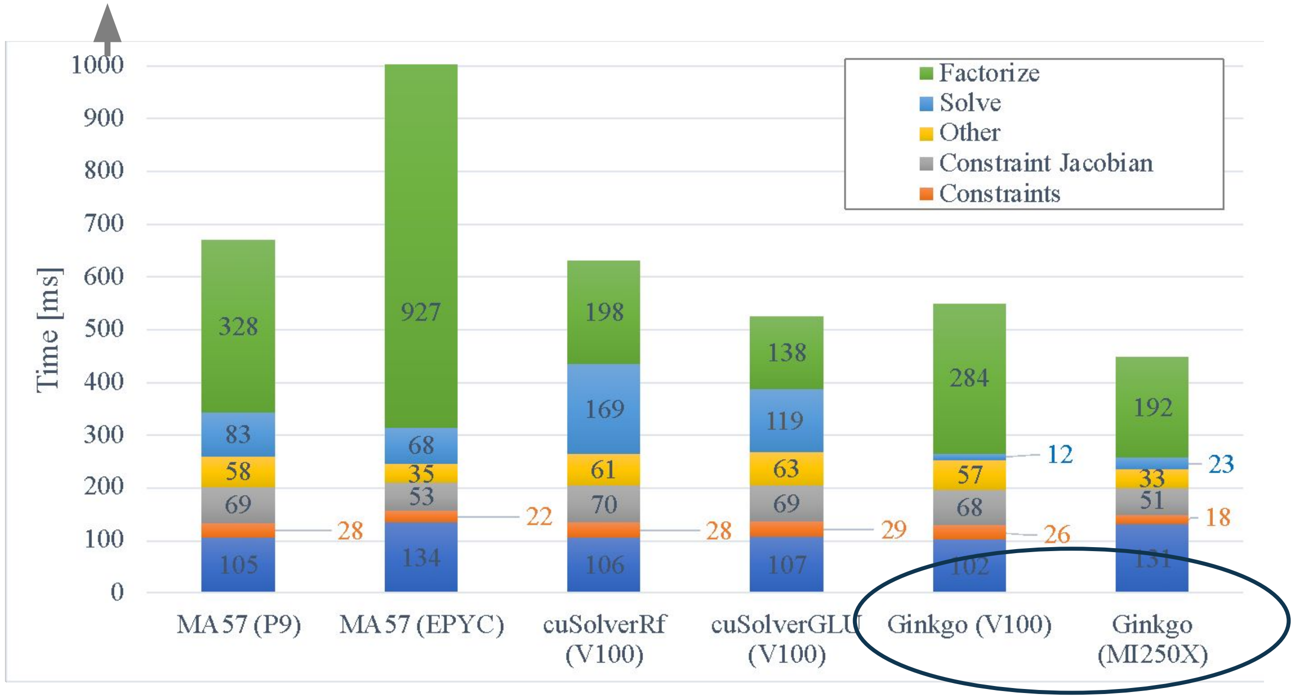
Can analyze Western power grid operations, with multiple scenarios and 1000's of contingencies to help operators plan for and respond to emergencies



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 teams 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

ECP EXASCALE COMPUTING PROJECT

E4S will be one of ECP's lasting legacy; demonstrating the concept of software as a facility

- E4S: HPC software ecosystem – a curated software portfolio built on software development toolkits (SDKs)
- 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: February 2024: E4S 24.02 – 120+ full release products

https://spack.io
Spack lead: Todd Gamblin (LLNL)

https://e4s.io
E4S lead: Sameer Shende (U Oregon)

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	Build caches 10X build time improvement
Turnkey stack A new user experience	https://e4s.io	Post-ECP Strategy LSSw, ASCR Task Force

One of the largest impacts of ECP

- Deployed on Frontier, Aurora, and Perlmutter
- Used on DoD, NOAA, NSF, and many other systems

ECP's software deployment (SD) team collaborated closely with DOE Facilities to ensure success on exascale systems

Deployed ~100 E4S packages across various programming environments.

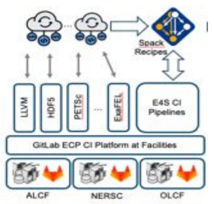
Consistently kept up to date with quarterly releases of E4S on Frontier, Crusher, Summit, Sunspot, Cori, and Perlmutter.

SD iterated extensively with the Facilities, vendors and ST developers to improve E4S software.

Ran more than 10,000 tests of the E4S installations.

Implemented and tested GitLab CI with Jacamar runner to satisfy the Facilities' security and provenance requirements.

The team helped install, configure, debug, & test the Spack software package manager.



Worked with AD and ST teams to implement CI with more than 20,000 hours of CI pipeline runs and 175,000 CI jobs.

Worked closely with ST and the Facilities to develop a "Level 2 (L2) Support Plan" in line with Facilities' SLAs and metrics.

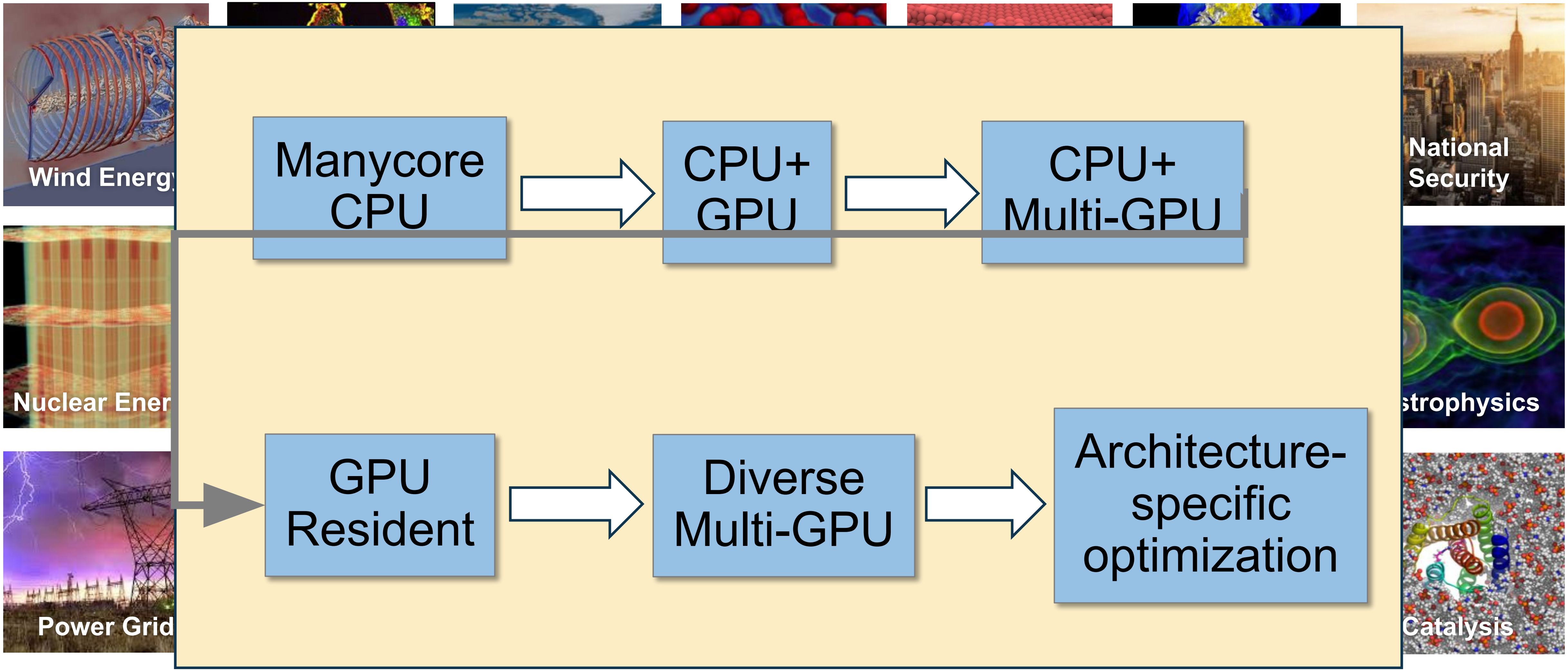
Achieving our KPPs would have been impossible without close collaborations with the ASCR HPC facilities

- Full GPU partition available to all ECP 10/2021
- Provided critical pre-exascale system for developing and testing GPUs implementations
- Helped shake out Slingshot network
- 51 ECP teams (285 users) on Perlmutter: 4.1M GPU node hours
- E4S deployed and widely used

- Full system available to all ECP 4/2023
- All 16 successful AD SC KPPs run on Frontier with excellent performance
- Many ST / Co-design KPP-3s achieved on Frontier
- All ECP teams (361 users) have run on Frontier: ~10.1M node hours (thru 12/31)
- 6 ECP teams were awarded a total of 7M INCITE node hours at OLCF
- 11 OLCF INCITE projects are led or co-led by former ECP team members

- System available to all ECP teams 11/2023
  - Lustre, interconnect, and some hardware issues being worked; Aurora is in the scaling phase
- ~10% of ST KPP-3 integrations achieved on Sunspot; major test for performance portability in final quarter
- 54 ECP teams (142 users) have run on Sunspot: ~ 47K node hours
- 46 ECP teams (83 users) have run on Aurora: 411K node hours

Exascale applications were designed to be flexible and adaptive



ECP applications teams used several different programming models to achieve performance portability

GPU-specific kernels

- Isolate the computationally-intensive parts of the code into CUDA/HIP/SYCL kernels.
- Refactoring the code to work well with the GPU is the majority of effort.

Loop pragma models

- Offload loops to GPU with OpenMP or OpenACC.
- Most common portability strategy for Fortran codes.

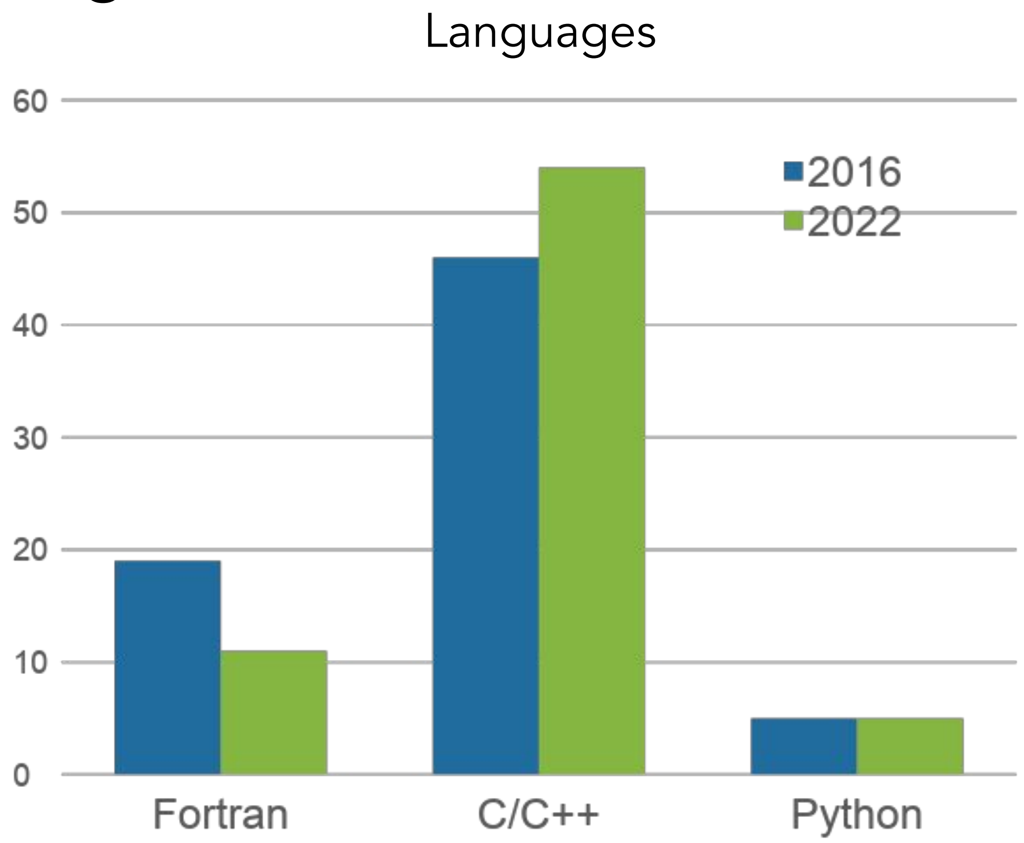
C++ abstractions

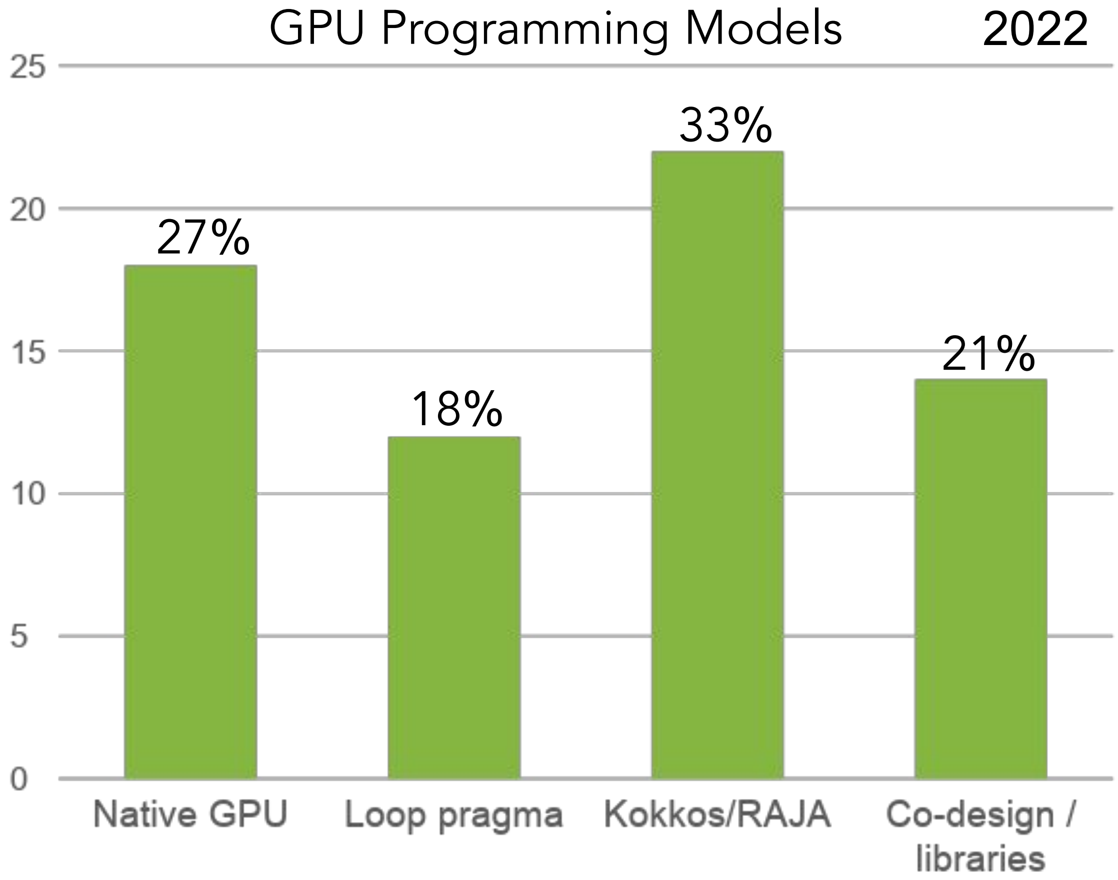
- Fully abstract loop execution and data management using advanced C++ features.
- Kokkos and RAJA developed by NNSA in response to increasing hardware diversity.

Co-design frameworks

- Design application with a specific motif to use common software components
- Depend on co-design code (e.g. CEED, AMReX) to implement key functions on GPU.

Distribution of ECP programming languages and models has changed over time





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.

ECP applications and software technologies were able to quickly make use of the Aurora system

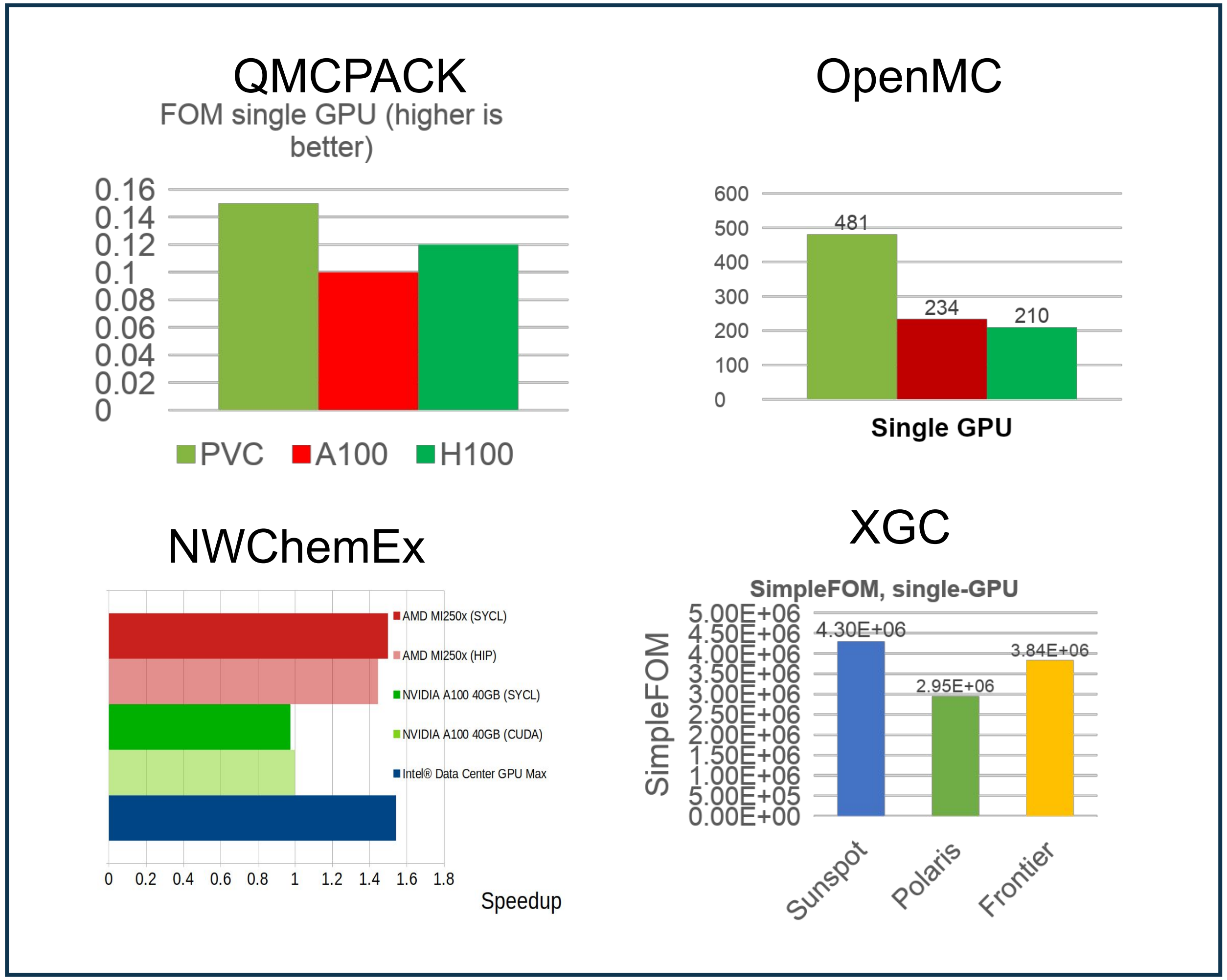
- Every application and software technology team designed their code with performance portability in mind
  - Preparations on Sunspot critical to teams' ability to run on Aurora
  - 13 AD teams are fully running on Aurora; 2 more are partially running
  - ~30 ST Libraries useful to Aurora are being installed from E4S
- Several applications have been able to run on 1000 more Aurora nodes, e.g., EXAALT, GAMESS, ExaSky, MFIX-exa
- Two teams have demonstrated their FOM and challenge problem on Aurora (EXAALT and ExaSky)

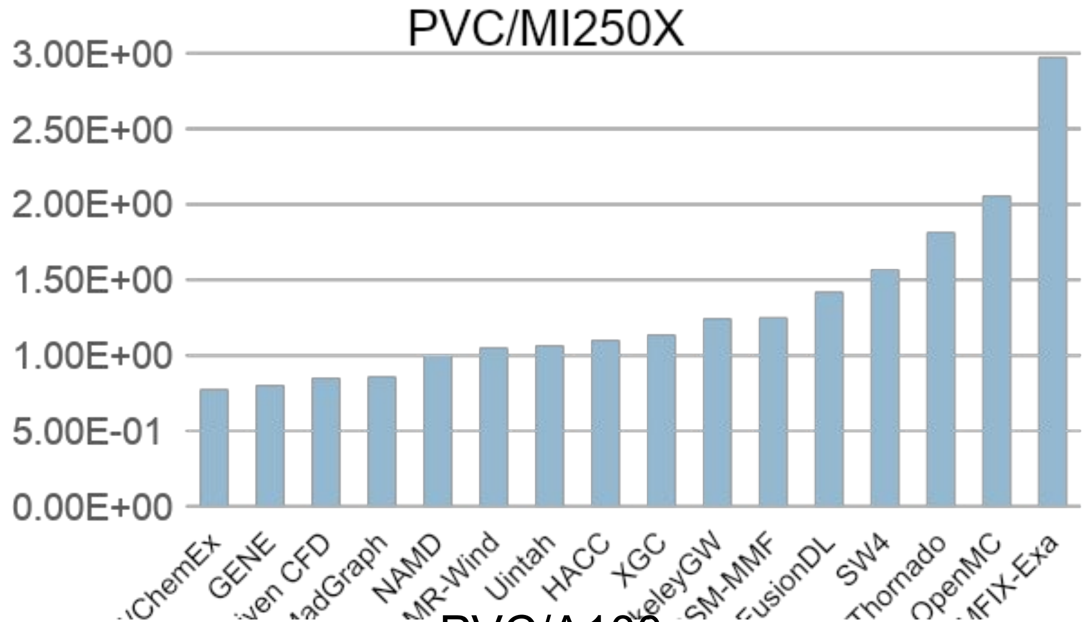
Project	Code	Science	Port	Runs on PVC
LatticeQCD	Grid	Full	Full	Running
	QUDA	Full	Full	Running
	Chroma	Full	Full	Running
	MILC	Full	Full	Running
ExaFEL	spiniFEL	Full	Partial	NT
	cctbx	Full	Full	Running
E3SM	E3SM-MMF	Full	Full	Running
EXAALT	LAMMPS	Full	Full	Running
	LATTE	Full	Partial	Partial
ExaWind	AMR-Wind	Full	Full	Running
	Nalu-Wind	Full	Full	Partial
	OpenFast	Full	N/A	N/A
ExaStar	Flash-X	Full	Full	Running

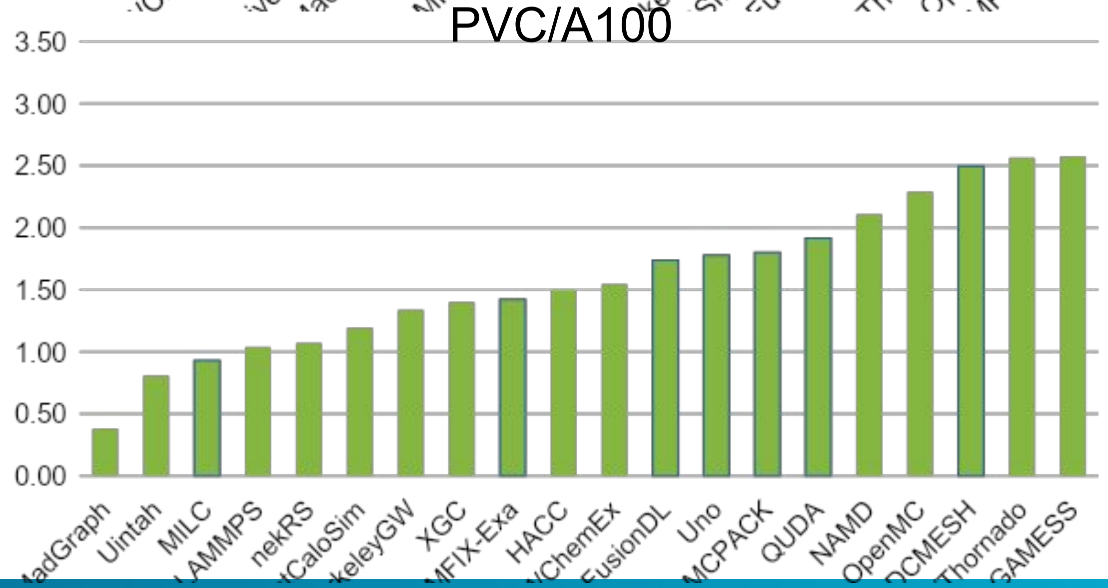
Project	Code	Science	Port	Runs on PVC
EQSim	SW4	Full	Partial	Running
	ESSI	Full	N/A	N/A
	HACC	Full	Full	Running
	NYX	Full	Full	Running
	QMCPACK	Full	Full	Running
	NekRS	Full	Full	Running
	OpenMC	Full	Full	Running
	XGC	Full	Full	Running
	GENE	Full	Full	Running
	GEM	Full	Full	Running
NWChemEx	NWChemEx	Full	Full	Running
	MFIX-Exa	Full	Full	Running
	Uno	Full	Full	Running
	GAMESS	Full	Full	Running

EXASCALE COMPUTING PROJECT

Performance portability layers have allowed several teams to successfully run on multiple GPU architectures







In the last two years ECP engaged in significant outreach and stakeholder activities as we bring the project to a close

- Engaging stakeholders on the new capabilities developed
- Engaging industry and other agencies with outreach and lessons learned to broaden the community of exascale-ready applications and technologies
- Working on a series of communications
  - Shared Fate
  - ECP and Industry
  - ECP and Agencies
- CiSE special issues
- ECP book proposals
- Significant presence at conferences such as SC23, SIAM PP24

AD stakeholder engagement

Office	POC	Briefing Date
HEP	Harriet Kung	June 10, 2022
BES	Linda Horton	June 29, 2022
WETO	Benjamin Hallissy	September 26, 2022
FES	James Van Dam	October 13, 2022
NE	Katie Huff	October 31, 2022
NP	Tim Hallman	May 10, 2023
BER	Gary Geemaert	June 15, 2023
OE	Gil Bindewald	June 27, 2023
BES	Linda Horton	October 31, 2023
FES	Jean Paul Allain	November 20, 2023
HEP	Jeremy Love	December 13, 2023
CESER	Daniel Lagriff	February 22, 2024
BER	Dorothy Koch	March 11, 2024

ECP has been very active with our Industry and Agency Council

NOAA collaboration points

- E4S use on the cloud
- Allowed first successful run of GFLD Earth system model on the cloud
- Exploring Spack-stack builds

DoD collaboration points

- E4S deployment on five DoD systems (Cray EX, HPE SGI 8600)
- Spack tools extensively used
- Deep dive evaluation of 16 E4S packages (Ascent, Exaworks, Flux, Kokkos, RAJA, PETSc, LAMMPS, etc.)

TAE Technologies collaboration points

- Adopted WarpX, AMReX, Adios, Ascent, for fusion framework
- Extending WarpX with new physics models
- Went from ALCF Theta CPUs to Perlmutter GPUs

Date	ECP/IAC Events
April 11, 2023	NASA/AD deep dive
July 18, 2023	NASA/ST deep dive
August 17, 2023	IAC Quarterly Call
October 3, 2023	Continuous Integration Workshop
October 25, 2023	Final in-person meeting of the IAC members
November 16, 2023	Impact of Exascale on Industry Panel at SC23
January 31-Feb 1	Final in-person meeting of the IAC tech reps

ECP leadership is actively engaged in project closeout activities

Jan 2024 | April 2024 | July 2024 | Oct 2024

Q2 | **Q3** | **Q4**

Cost reconciliation with partner labs

Finalize and submit CD-4 documents

Archive project data and decommission project systems

★ Finalize CD-4 documentation as much as possible

★ Capitol Hill Exascale Celebration

★ Project End
Hand off to ORNL Business Services Directorate

★ CD-4 Red Team

★ CD-4 Review

★ ESAAB Review

ECP leadership team funded to work on closeout, outreach, and communications

ECP project office continues to work on archiving, decommissioning systems, and financial close out

ORNL staff complete subcontract closeout and final report

As of April 2024, the ECP has closed 221 of 599 of total funding elements including some of the awards to universities, HPC vendors, and software procurements. 362 are pending closure; 16 remain open. Remaining costs are $4M with a remaining cost variance of $29.9M. Five months of schedule contingency remain.

EXASCALE COMPUTING PROJECT

Successful CD-4 review completed April 17, 2024

- Four Charge Questions all answered with a resounding “Yes!”:
  1. Has the project met all of the Key Performance Parameters as described in the Project Execution Plan?
  2. Will the final project costs be within the project baselines when closeout activities are complete?
  3. Has a draft project closeout report been developed and have the lessons learned from the project been identified, discussed, and captured in a draft report?
  4. Is the ECP ready for approval of CD-4, Project Completion?

- New project documents developed:
  - DOE O413.3B required documents
    - Project Closeout Report
    - Transition to Operations Plan
    - Lessons Learned
  - Other useful documents
    - Financial Closeout Plan
    - Final PEP Milestone Report
    - Document Control and Records Management

- A few key themes from the report out:
  - Kudos for the technical accomplishments of the project and successful tailoring of 413.3b to ensure productive outcomes
  - The idea of software as a facility resonated with reviewers; reiterated the need for software sustainability as a priority
  - The *-Forward programs saw significant impact on exascale systems; encouragement to continue this bridge
  - Continue to expand on and promulgate lessons learned to the broader community
  - ECP created culture change and strong collaborations among applications/software technologies/facilities and NNSA/SC

Performance measurement baseline budget by Fiscal Year

WBS	Description	FY16	FY17	FY18	FY19	FY20	FY21	FY22	FY23	FY24	Total
2	Exascale Computing Project	$69.7M	$153.2M	$255.4M	$321.4M	$313.2M	$230.3M	$214.9M	$177.1M	$47.9M	$1,783.1M
2.1	Project Management	$4.6M	$11.7M	$11.1M	$9.7M	$16.6M	$13.8M	$11.4M	$3.1M	$13.9M	$95.9M
2.2	Application Development	$21.7M	$61.1M	$90.3M	$115.5M	$131.6M	$101.4M	$99.0M	$75.7M	$16.1M	$712.4M
2.3	Software Technology	$23.3M	$55.1M	$75.2M	$81.8M	$64.6M	$71.3M	$73.6M	$71.0M	$11.1M	$527.0M
2.4	Hardware and Integration	$20.1M	$25.3M	$78.8M	$114.4M	$100.4M	$43.8M	$30.9M	$27.3M	$6.8M	$447.8M

- Baseline set June 1, 2019; changes made through a formal Project Change Request system
- Budget at complete is $1,783.1M with contingency of $29.1M for a total project cost of $1,812.3M
- ECP came in on schedule (SPI of 1.00) and under budget (CPI of 1.02)
- Current cost variance is $29.9M
- Remaining funds in contingency and cost variance are more than adequate to cover remaining costs

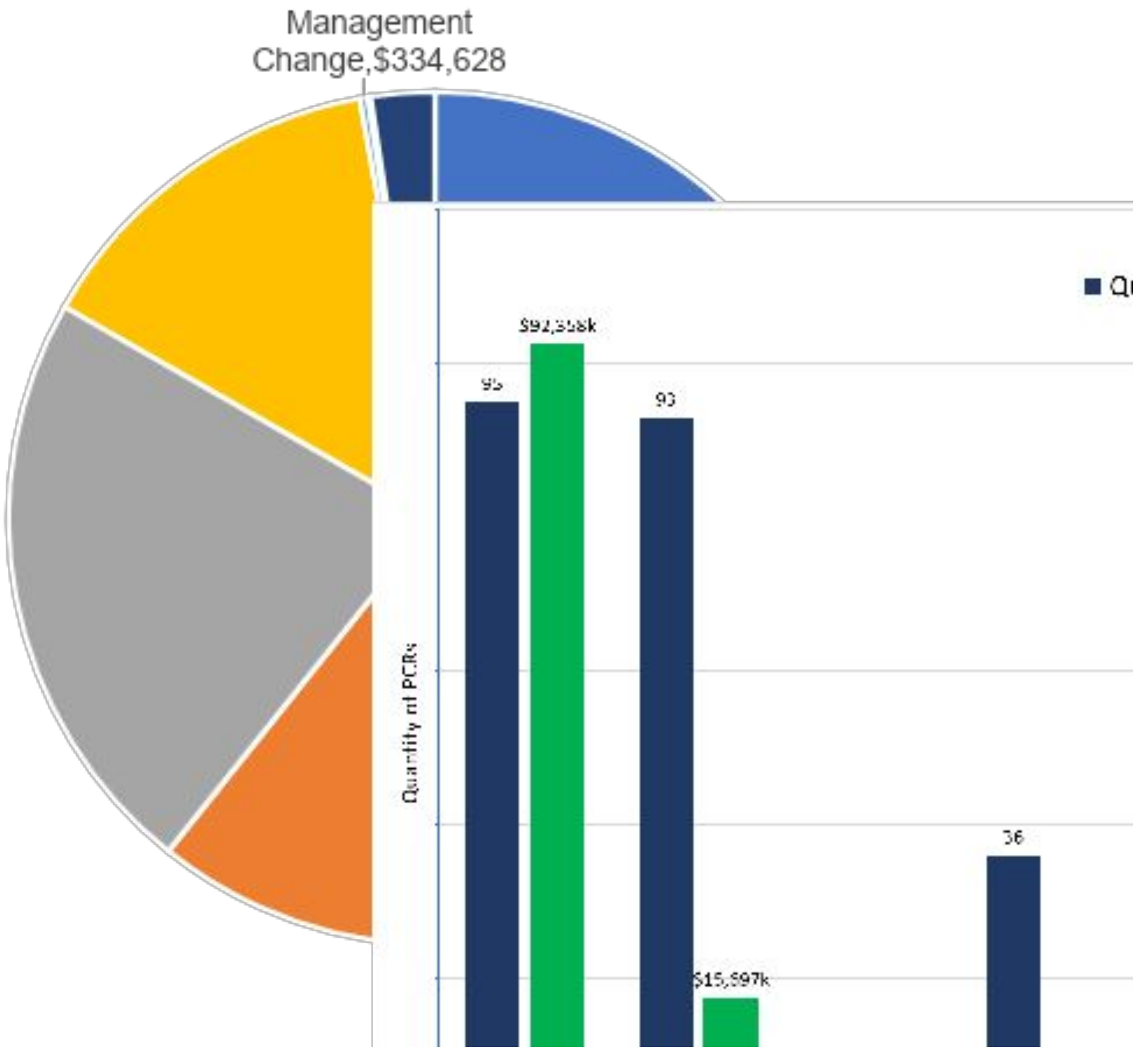
- ECP leadership team flattened the budget profile relative to the funding profile expecting a large push late in the project to run on exascale systems
- Two one-quarter extensions required to mitigate risks associated with delayed access to exascale systems
  - Reflected in budgets in FY23 and FY24

Aggressive use of project change requests was crucial to success

366 PCRs executed since CD-2/3 in 11 broad categories, Total Value $107.3M

Category	Value	PCRs
Risk Mitigation	$35,888,874	89
1 Qtr Project Extension (FY23Q4)*	$32,615,919	3
1 Qtr Project Extension (FY24Q1)*	$25,383,380	4
Scope Adjustment	$15,697,440	93
Management Change	$334,628	36
COVID-19	$126,000	9
Convert Planning Packages	$0	37
Milestone Adjustment	$0	16
Schedule Adjustment	$0	40
WBS Dictionary	$0	5
Resource Alignment	($2,715,832)	34
Grand Total	**$107,330,408**	**366**

*Separated from Risk Mitigation Category



Major PCRs over $1M (since 4/23)

- PCR-800 – 1Q Extension ($25M)
- PCR-823 – PMO FY24 Baseline (4.5M)
- PCR-829 – AI/ML Extension ($1.8M)

We are capturing and sharing our lessons learned in a living document: **Successes**

Integration

- Diverse, multi-disciplinary teams are required to tackle technology inflection points
- Build integration into the project structure and operations, e.g., integration meetings, hierarchical/cross team communications
- Make integration a metric for success for subprojects, e.g., shared fate milestones and KPP-3
- Explicitly manage external dependencies throughout the life of the project, e.g., with Facilities, sponsors, etc.
- Build and use tools that enable integration, e.g., the integration database, Jira, Confluence

Agility

- Use agile project management practices to assess and manage RD&D scope; allow for replanning and reprioritization as knowledge is gained
- Use milestone constructs that are appropriate for a software RD&D project
- Formalize and use project change control processes to adjust scope and schedule where needed
- Rely on technical CAMs for assessing progress and managing challenges for L4 subproject teams
- Address extraordinary events in a timely fashion, e.g., delay of system access due to COVID-19 resulted in two 1Q extensions

Project Management

- Adopt and tailor formal 413.3B project management practices; e.g. performance measurement processes
- Train technical staff and leaders on formal management practices early and partner them with project management specialists
- Trust and empower a strong leadership team with clearly defined roles and responsibilities
- Identify, provide, and require use of centralized tools for project management, e.g., Jira, Confluence, Slack
- Record institutional commitments through MOUs and MOAs

We are capturing and sharing our lessons learned in a living document: **Improvements**

Impact of differences in funding/ staffing/sponsors

- More proactive education of technical staff on formal project management terms (e.g. Budget vs. Funding) would have reduced confusion
- Managing SC construction funds with NNSA operating funds resulted in differing rate overheads and posed challenges for managing program elements that were funded by both , e.g., PathForward
- Dedicate staff resources and tailor reporting for subcontracts funded through partner institutions
- Managing a large number of MPOs in ECP could have benefitted from a dedicated and embedded procurement specialist

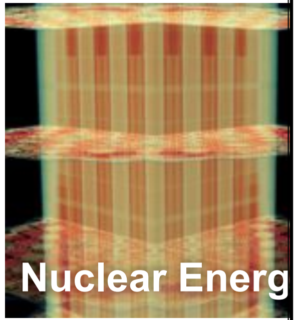
Impact of using 413.3B for an RD&D project

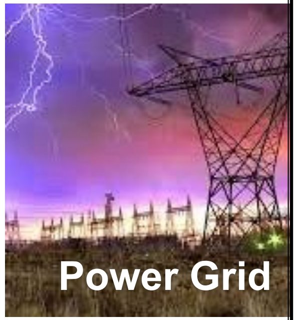
- It took time to develop tailored practices for implementing 413.3B constructs in a software RD&D project; the burden was heavy on project teams at the beginning
- While the ECP KPPs measured success in creating a exascale ecosystem, aspects could be improved in appropriately defining challenge problems and in understanding the review burden a priori
- Plan to support technical PIs with additional centralized and/or local project coordinators to help manage 413.3B requirements
- Absorbing research programs into ECP resulted in a focus on shorter-term goals at the expense of longer-term ones

Project Planning Impact

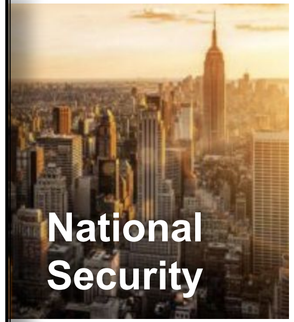
- ECP had several major transitions in the WBS that required significant effort to plan and implement. There were several downstream effects in tracking project metrics that should be carefully understood to minimize disruption.
- Significant effort by ECP leadership was required to help mitigate the distraction and loss of key personnel associated with a lack of a coordinated plan for post-ECP activities

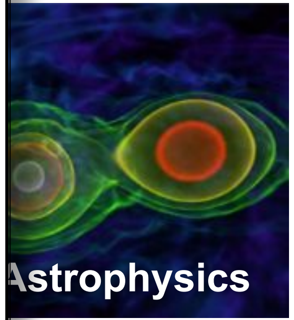
What’s next for ECP applications?

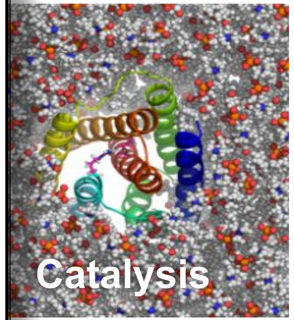




- A mix of funding sources
  - SC application offices base programs (e.g., E3SM, QMCPack, WarpX, ExaBiome)
  - SciDAC (e.g., ExaStar, ExaSky)
  - DOE Earthshots (e.g., ExaWind, Subsurface)
  - NNSA (e.g., ATDM applications, ExaAM)
  - Applied energy offices (e.g., ExaSGD, ExaAM, EQSim, ExaSMR)
  - Laboratory LDRD (e.g., EXAALT)
  - Other (e.g., CANDLE)
- About half of the teams have 50% or more of their ECP level of funding
  - Scope generally focused in different ways
  - Team often disbursed







What's next for ECP software technologies?

The Consortium for the Advancement of Scientific Software funded by DOE

- **PESO:** Stewarding, evolving and integrating a cohesive ecosystem for DOE software
- **RAPIDS & FASTMath:** Stewardship, advancement, integration for math, data/vis, and ML/AI packages.
- **SWAS:** Stewardship and project support for scientific workflow software and its community
- **S4PST:** Stewardship, advancement and engagement for programming systems.
- **STEP:** Stewardship, advancement of software tools for understanding performance and behavior.
- **COLABS:** Training, workforce development, and building the RSE community.
- **CORSA:** Partnering with foundations to provide onboarding paths for DOE-funded software.

Addressing the stewardship needs of the DOE ASCR scientific software ecosystem

Goal: to ensure the long-term viability of the ASCR software ecosystem so that it may continue to serve as the base for future DOE-funded research.

Member organizations have distinct and complementary foci; all have some ECP representation

https://cass.community

The PESO project is focused on the Scientific Software Ecosystem

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

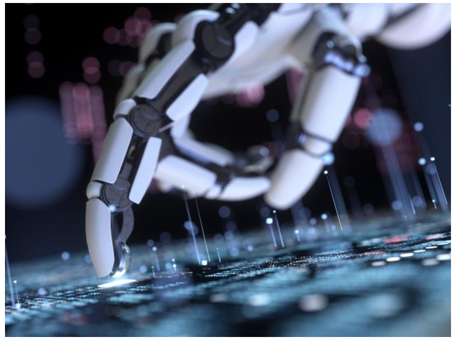
Port & Test Platforms
- Frank test & development system
- Cloud resources
- Documentation, training

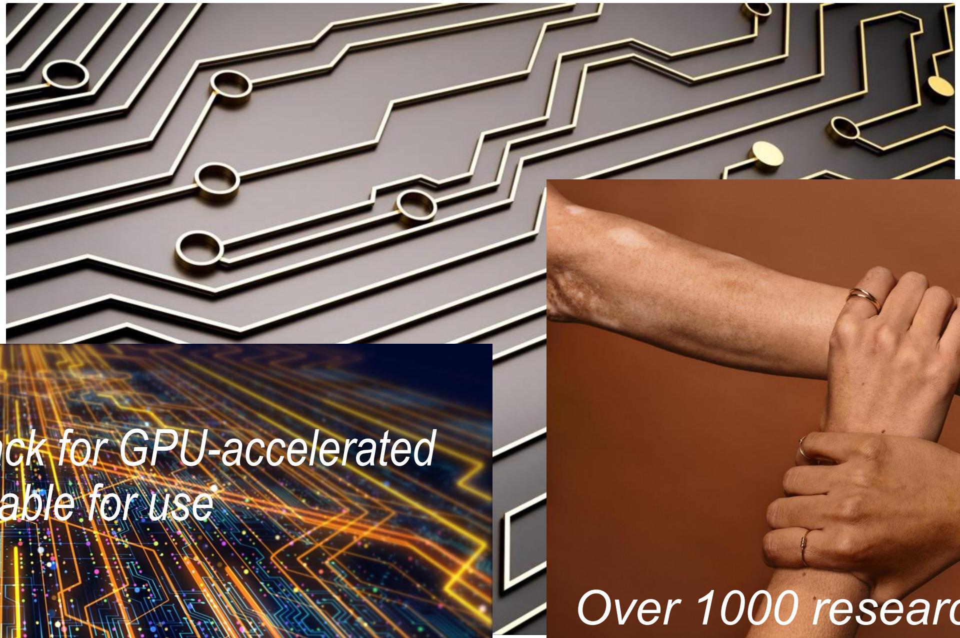
BSSw.io Content (w. COLABS)
- Short articles on topics related to scientific software productivity and sustainability (recruit, write, review, & edit)

ECP’s impact and legacy will be long-lived and far-reaching

A suite of applications that will impact problems of national importance for decades to come

Best practices for running a large-scale software development RD&D 413.3B project





Best practices and lessons learned for thinking about how to program GPUs – moving the nation forward at all levels of computing

Thank you

https://www.exascaleproject.org

This research was supported by the Exascale Computing Project (17-SC-20-SC), a joint project of the U.S. Department of Energy's Office of Science and National Nuclear Security Administration, responsible for delivering a capable exascale ecosystem, including software, applications, and hardware technology, to support the nation's exascale computing imperative.

EXASCALE COMPUTING PROJECT

ECP Director: Lori Diachin
ECP Deputy Director: Ashley Barker

Thank you to all collaborators in the ECP and broader computational science communities. **The work discussed in this presentation represents creative contributions of many people who are passionately working toward next-generation computational science.**

Questions?