Scientific software ecosystems and communities:

Why we need them and how each of us can help them thrive

EXASCALE COMPUTING PROJECT

Lois Curfman McInnes
Argonne National Laboratory

December 8, 2021

In the webinar series: *Best Practices for HPC Software Developers*

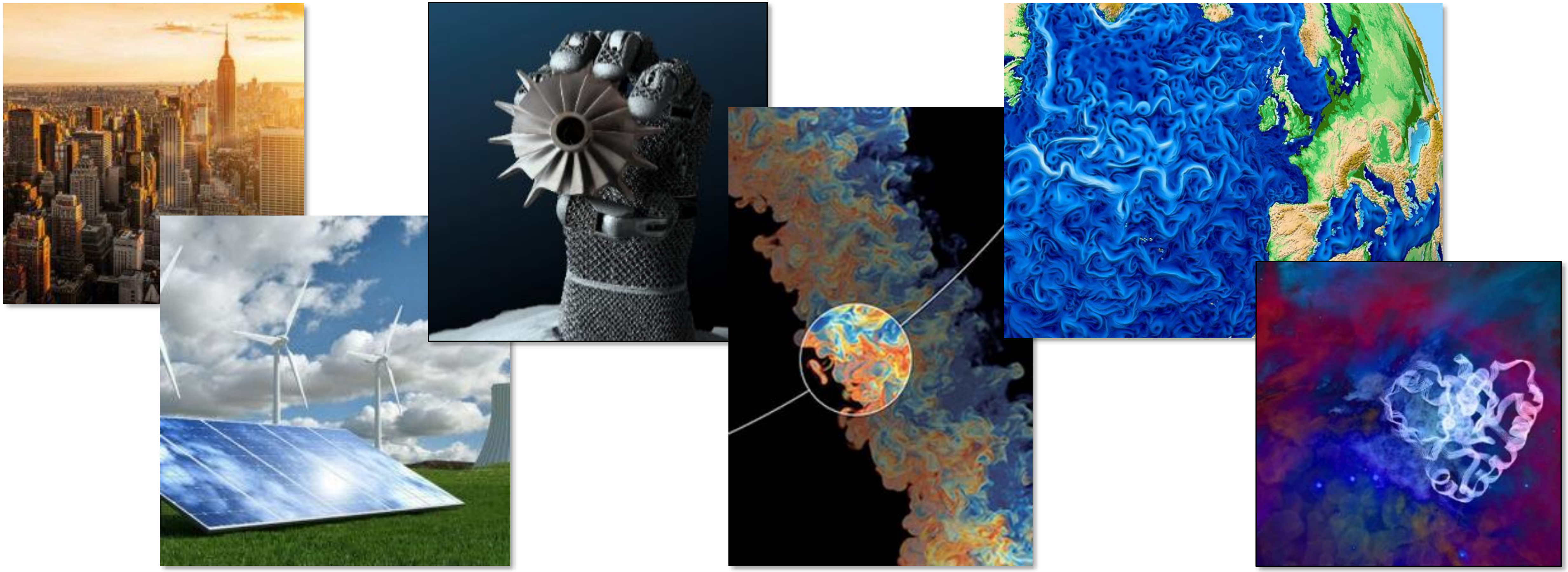
Based on a presentation at SC21

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

Science and beyond: Applications and discovery in ECP

And more:
- 24 applications
- 6 co-design centers



National security | Energy security | Economic security | Scientific discovery | Earth systems | Health care

Thank you to ECP applications teams and their collaborators in software technologies

ECP EXASCALE COMPUTING PROJECT

Software quality is a critical component of quality science.

Focus: The fundamental roles of scientific software ecosystems and communities ...
… and how each of us can help them thrive

Ann Pachett essay: *Fact vs Fiction*

“Every story you read, you realize that the writer has made a decision for what to include, and what to leave out. It doesn’t mean that he or she isn’t telling the truth, it means events can’t be recorded exactly. They can only be interpreted.

Even a photograph reveals only part of the picture. The frame is defined by its own 4 edges. Whom to you choose to leave out of the portrait? Whom do you choose to include?

You realize that one answer is not enough and that you have to look at as many sources as are available to you so that you can piece together a larger picture.

Everyone adds a chip of color to the mosaic and from there some kind of larger portrait begins to take shape.

We each tell 1 version of a complicated story.”

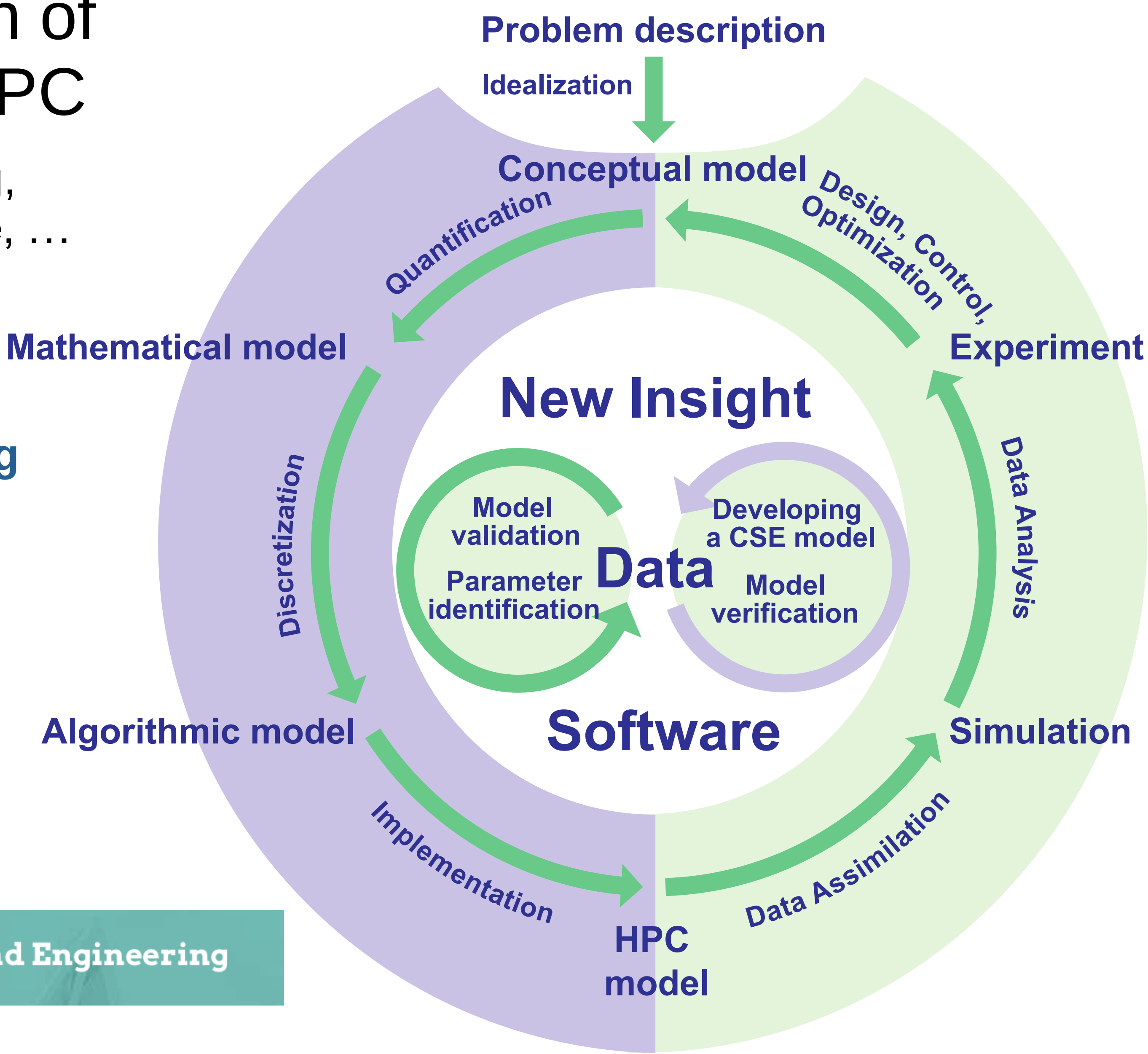
Software is the foundation of sustained collaboration in HPC

- computational science and engineering, data science, learning/AI, infrastructure, …

HPC = High-performance computing

CSE = Computational science and engineering

Ref: Research and Education in Computational Science and Engineering, U. Rüde, K. Willcox, L.C. McInnes, H. De Sterck, **SIAM Review**, 2018



SIAM Activity Group on Computational Science and Engineering

EXASCALE COMPUTING PROJECT

Audience query: HPC software and YOU

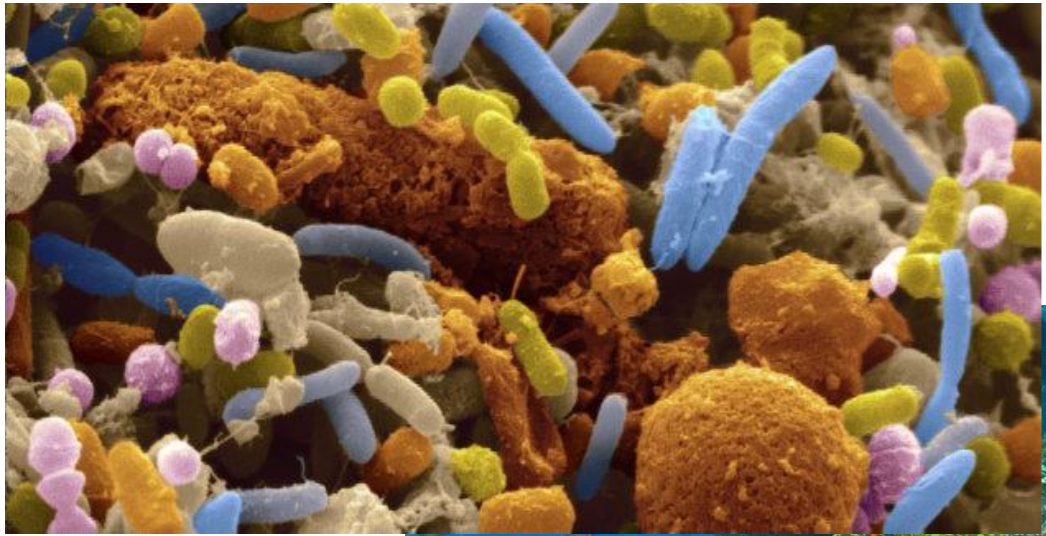
- **Do you develop HPC software?**
  - **That you use yourself**
  - **That you provide to others**
    - In your research group
    - In the broader community
- **Do you use HPC software developed by others?**
- **Do you contribute to teams who develop and use HPC software?**
  - Strategy, planning, logistics, raising funds, …
- **Do you lead projects or organizations where teams develop and use HPC software?**
- **Are you a stakeholder or supporter of projects that develop and use HPC software?**
- **Are you a fan of HPC software?**

☺

HPC software, including computational science and engineering, data science, learning/AI, infrastructure, …

EXASCALE COMPUTING PROJECT

Ecosystem: A group of independent but interrelated elements comprising a unified whole



Diversity is essential for an ecosystem to thrive.

- No element functions in isolation.
- Each element fulfills unique roles.
- The whole is greater than the sum of its parts.

We must explicitly consider **community software ecosystem perspectives** for next-generation computational science

- **Complex, intertwined challenges**
  - Technical and sociological
- **Need community efforts … and each of us!**
  - Improve software sustainability
  - Change research culture
- **Building an ECP software ecosystem**
  - While advancing scientific productivity through better scientific software
- **Get involved!**

Why? Reduce technical risk

Provide a firmer foundation for science at exascale and beyond

nature computational science

Comment | Published: 22 February 2021

How community software ecosystems can unlock the potential of exascale computing

Lois Curfman McInnes, Michael A. Heroux, Erik W. Draeger, Andrew Siegel, Susan Coghlan & Katie Antypas

Nature Computational Science **1**, 92–94(2021) | Cite this article

Metrics

Emerging exascale architectures and systems will provide a sizable increase in raw computing power for science. To ensure the full potential of these new and diverse architectures, as well as the longevity and sustainability of science applications, we need to embrace software ecosystems as first-class citizens.

https://dx.doi.org/10.1038/s43588-021-00033-y

EXASCALE COMPUTING PROJECT

Thank you to my collaborators and communities

SIAG/CSE

- DOE Exascale Computing Project
- DOE Advanced Scientific Computing Research – Applied Math Program
- Argonne National Lab (MCS Division)
- SIAM Activity Group on CSE
- Developers of E4S and xSDK
- IDEAS Productivity Project
- Better Scientific Software (BSSw) community
- PETSc / TAO developers and users
- Sustainable Horizons Institute

Thank you

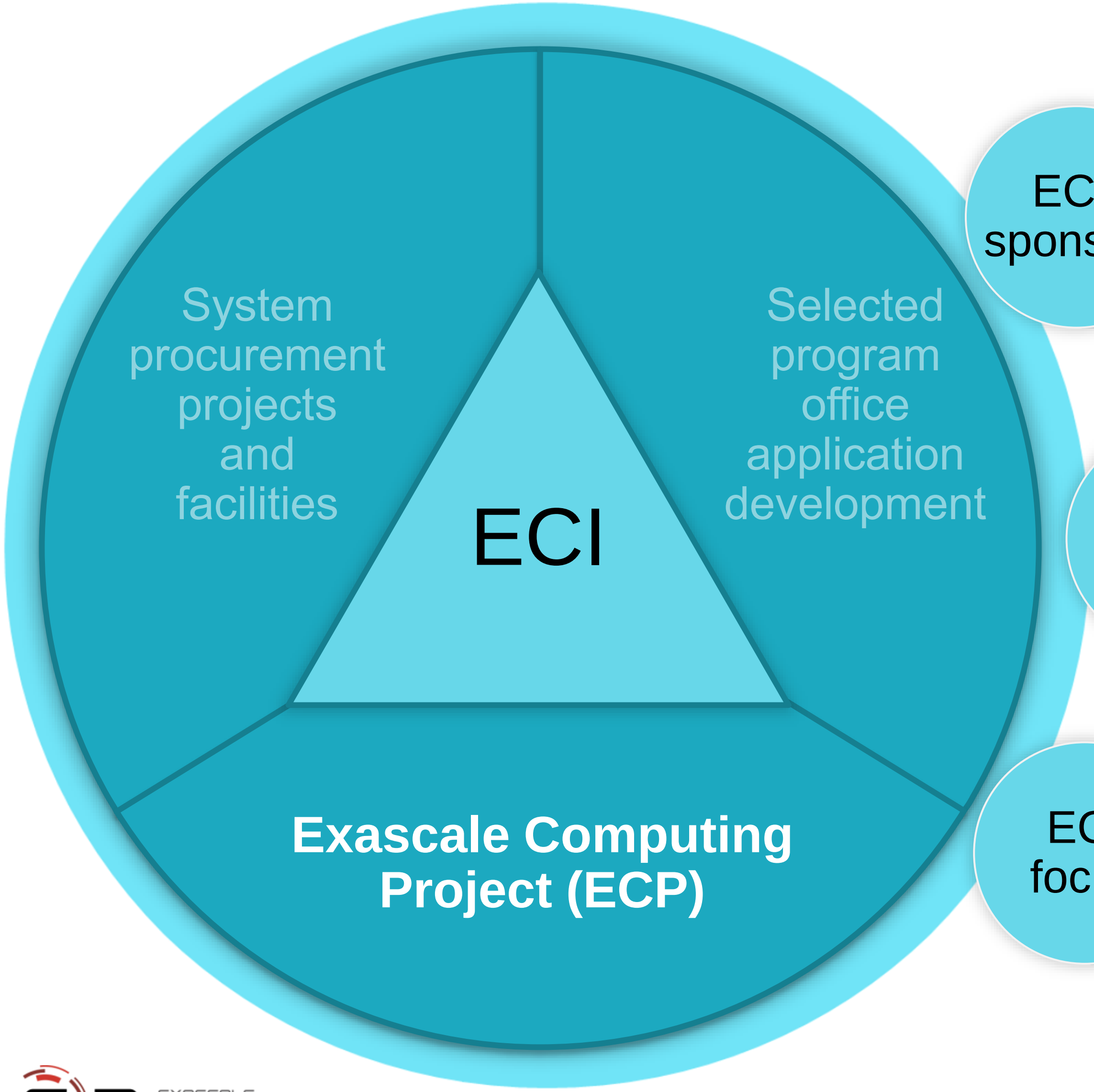
https://www.exascaleproject.org

<u>Thank you</u> 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.**

ECP Director: Doug Kothe
ECP Deputy Director: Lori Diachin

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.

DOE Exascale Computing Initiative (ECI)



ECI sponsors: DOE Office of Science (SC);
National Nuclear Security Administration (NNSA)

ECI mission: Accelerate R&D, acquisition, and deployment to deliver exascale computing capability to DOE national labs by the early to mid-2020s

ECI focus: Delivery of an enduring and capable exascale computing capability for use by a wide range of applications of importance to DOE and the US

ECP by the numbers

7
YEARS

$1.8B

A seven-year, $1.8B R&D effort that launched in 2016

6
CORE DOE
LABS

Six core DOE National Laboratories: Argonne, Lawrence Berkeley, Lawrence Livermore, Oak Ridge, Sandia, Los Alamos

- Staff from most of the 17 DOE national laboratories take part in the project

3
FOCUS
AREAS

Three technical focus areas: Hardware and Integration, Software Technology, Application Development supported by a Project Management Office

80+
R&D TEAMS

1000
RESEARCHERS

More than 80 top-notch R&D teams

Hundreds of consequential milestones delivered on schedule and within budget since project inception

Oak Ridge National Laboratory
Argonne National Laboratory
Berkeley Lab Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Sandia National Laboratories

ECP EXASCALE COMPUTING PROJECT

ECP’s holistic approach uses co-design and integration to achieve exascale computing

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

Katie Antypas, HI Director
Susan Coghlan, HI Deputy Director

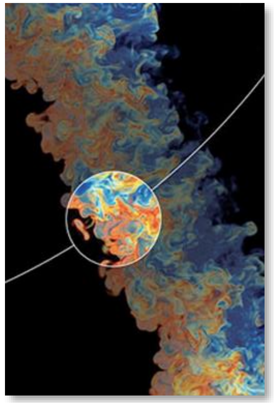
Science and beyond: Applications and discovery in ECP

National security

Next-generation, **stockpile stewardship** codes

Reentry-vehicle-environment simulation

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



Energy security

Turbine **wind plant** efficiency

Design and commercialization of **SMRs**

Nuclear fission and fusion reactor **materials design**

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

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

Scale up of **clean fossil fuel** combustion

Biofuel catalyst design

Economic security

Additive manufacturing of qualifiable metal parts

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

Seismic hazard risk assessment

Scientific discovery

Cosmological probe of the standard model of particle physics

Validate fundamental laws of nature

Plasma wakefield accelerator design

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

Find, predict, and control materials and properties

Predict and control **magnetically confined fusion plasmas**

Demystify **origin of chemical elements**

Earth systems

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

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

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

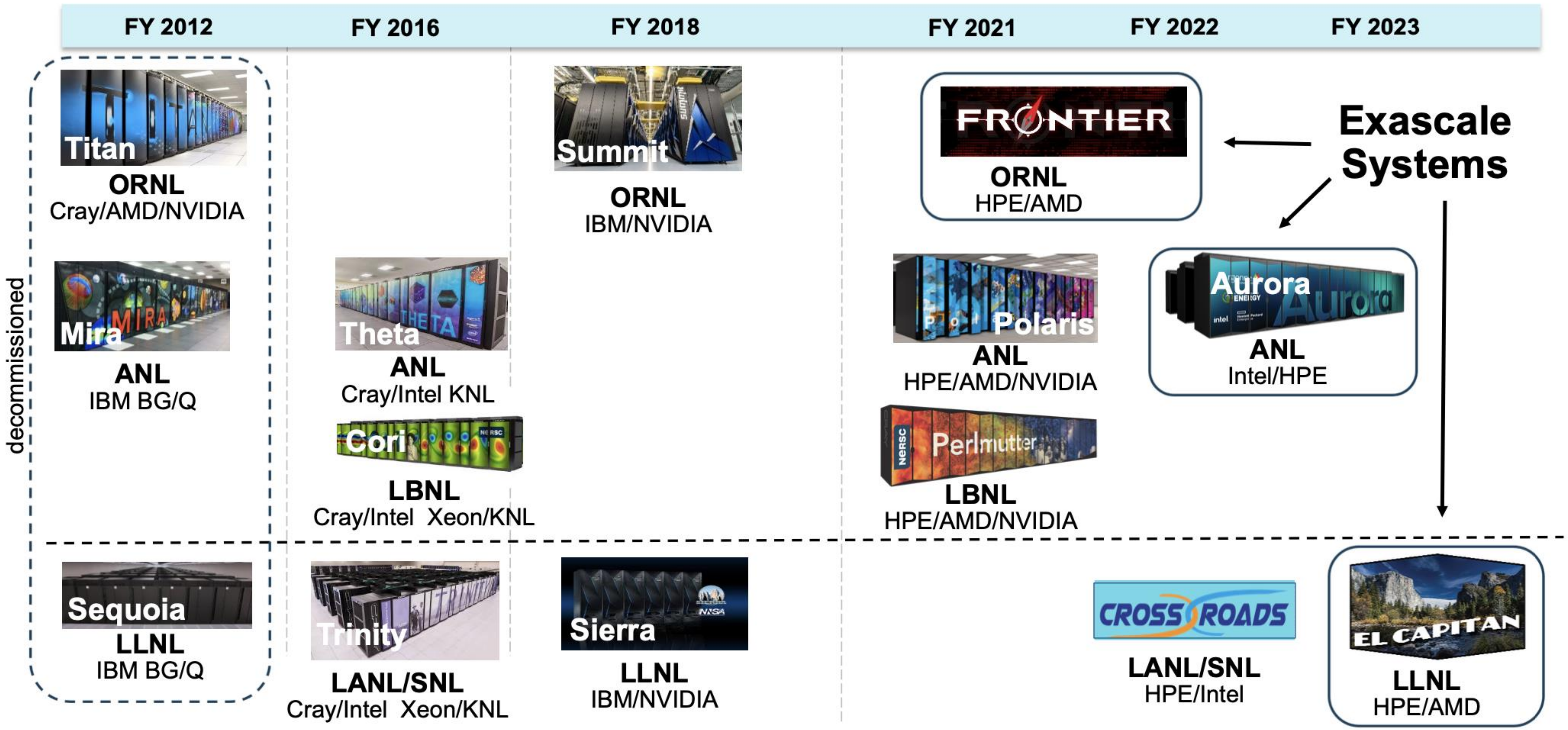
Health care

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

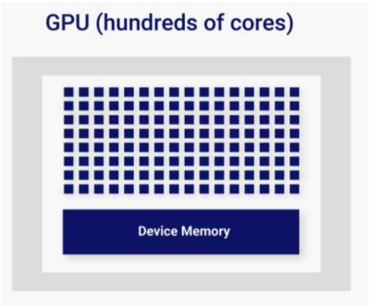
Thank you to Andrew Siegel, Erik Draeger and ECP applications teams

EXASCALE COMPUTING PROJECT

DOE HPC Roadmap to Exascale Systems



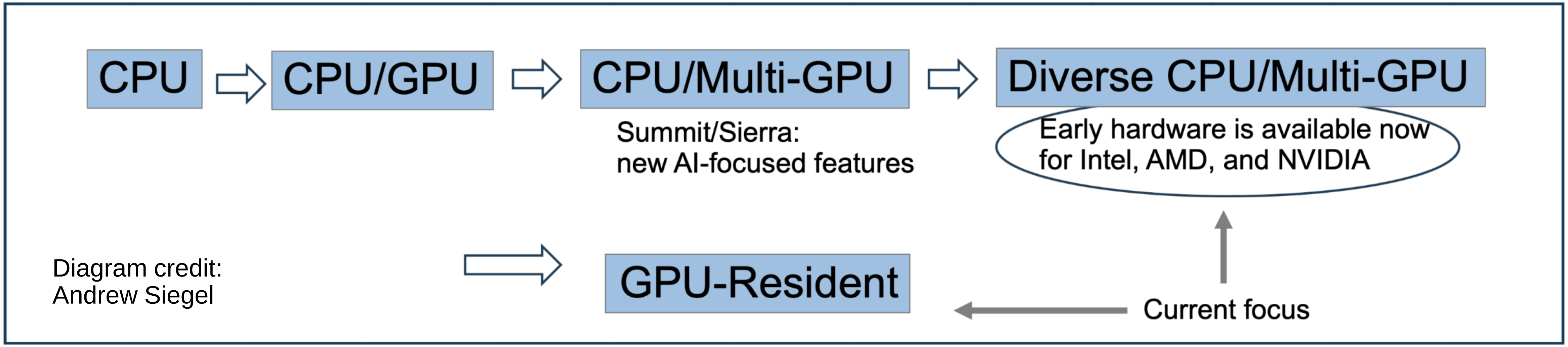
Heterogeneous accelerated-node computing



Accelerated node computing: Designing, implementing, delivering, & deploying agile software that effectively exploits heterogeneous node hardware

- Execute on the largest systems … AND on today and tomorrow's laptops, desktops, clusters, …
- We view *accelerators* as any compute hardware specifically designed to accelerate certain mathematical operations (typically with floating point numbers) that are typical outcomes of popular and commonly used algorithms. We often use the term GPUs synonymously with accelerators.

Text credit: Doug Kothe



Ref: A Gentle Introduction to GPU Programming, Michele Rosso and Andrew Myers, May 2021

New science opportunities at extreme scale

Beyond interpretive simulations … working toward predictive science

- Multirate, multiscale, multicomponent, multiphysics simulations
- Simulations involving stochastic quantities, sensitivities, UQ, optimization
- Coupling of simulations, data analytics and learning … HPC / AI
- Complex workflows among DOE facilities (compute and scientific / observational)
  - Ref: Ben Brown, DOE ASCAC Meeting, Sept 2021

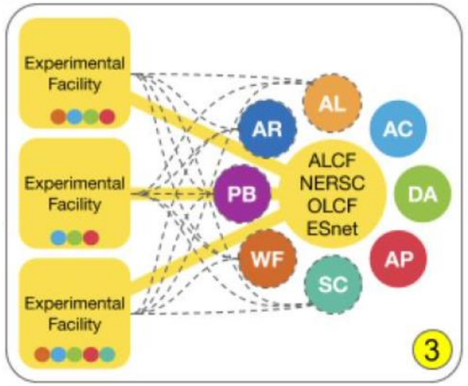
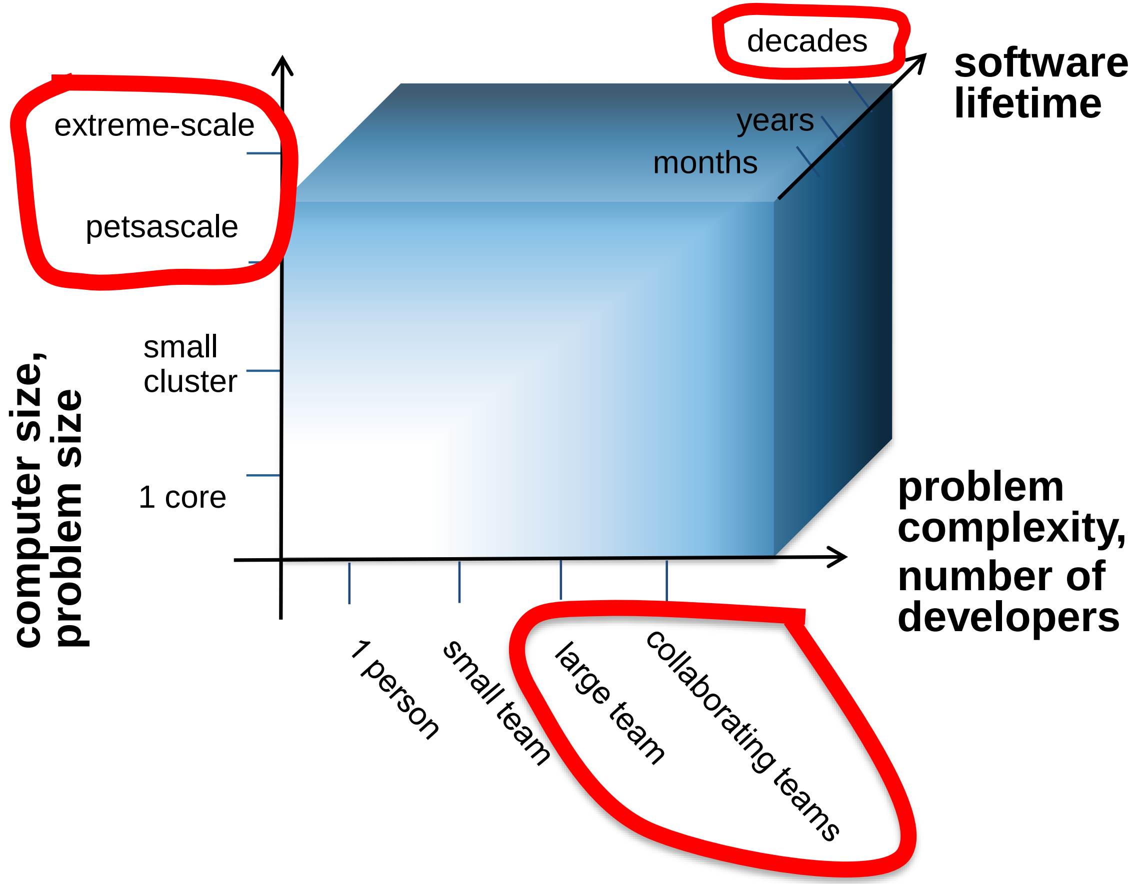
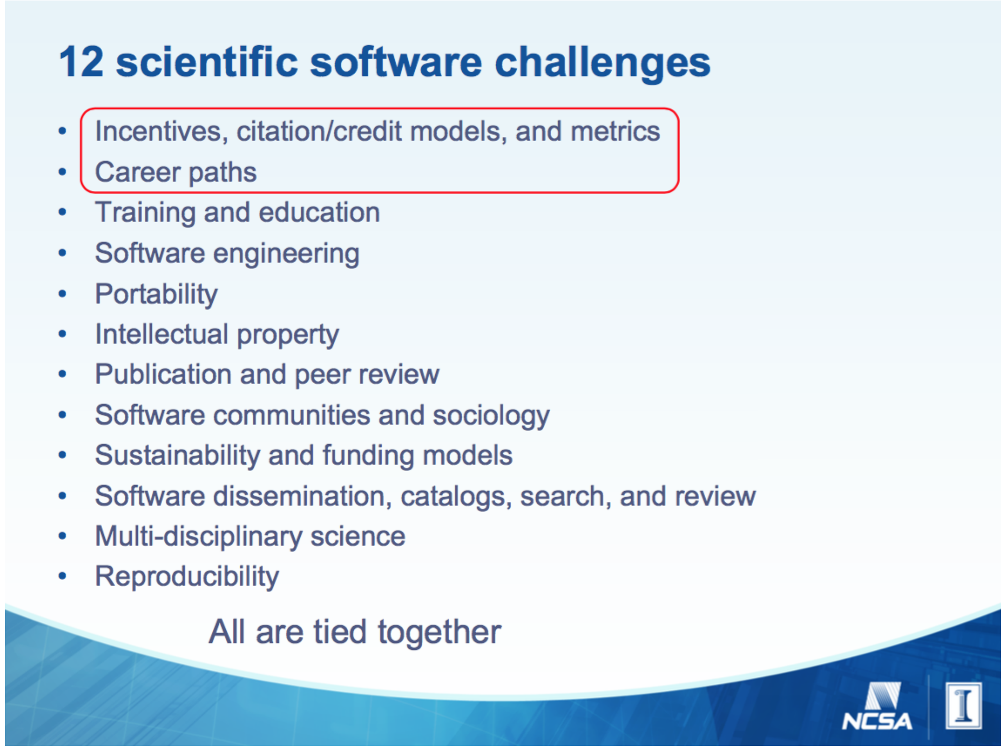


Diagram credit: ASCR Integrated Research Infrastructure Task Force, C. Adams et al., March 2021

Community software ecosystems require high-quality software

Complex, intertwined challenges





Ref: D.S. Katz, *Software in Research: Underappreciated and Underrewarded*, 2017 eResearch Australasia conference, https://doi.org/10.6084/m9.figshare.5518933

Challenges of scientific software

Technical

- All parts of the cycle can be under research
- Requirements change throughout the lifecycle as knowledge grows
- Importance of reproducibility
- Verification complicated by floating point representation
- Real world is messy, so is the software

Sociological

- Competing priorities and incentives
- Limited resources
- Perception of overhead with deferred benefit
- Need for interdisciplinary interactions

Technical debt: The implied cost of additional rework caused by choosing an easy (limited) solution now instead of using a better approach that would take longer.

- Wikipedia

EXASCALE COMPUTING PROJECT

Community organizations: Resources and opportunities to get involved

- Research Software Alliance: https://www.researchsoft.org
- Software Sustainability Institute: https://www.software.ac.uk
- US Research Software Sustainability Institute: https://urssi.us/
- NumFOCUS: https://www.numfocus.org
- WSSSPE: http://wssspe.researchcomputing.org.uk/
- Software Carpentry: https://software-carpentry.org
- Research Software Engineering (RSE) movement:

  https://society-rse.org   https://us-rse.org
- IDEAS Productivity: https://ideas-productivity.org
- Better Scientific Software: https://bssw.io
- And more …

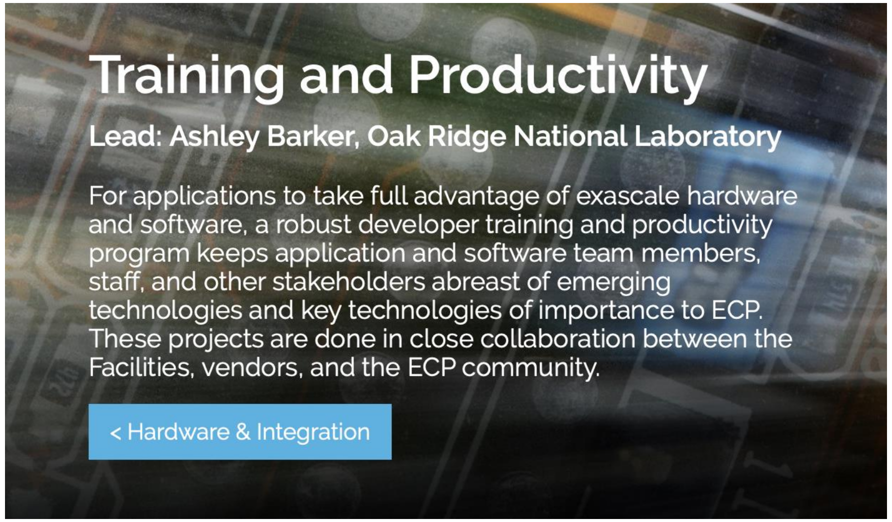
Ref: Community Organizations: Changing the Culture in Which Research Software Is Developed and Sustained, D.S. Katz, L.C. McInnes, et al, **IEEE CiSE**, 2019

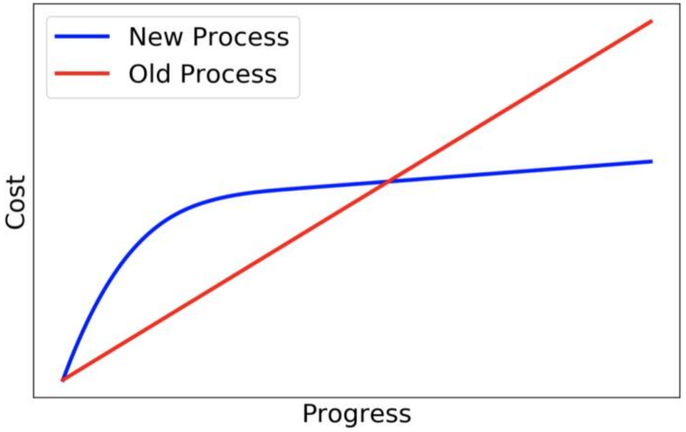
Building an ECP software ecosystem …

… While advancing scientific productivity through better scientific software

Advancing scientific productivity through better scientific software

Reducing technical risk by building a firmer foundation for computational science





Improving developer productivity and software sustainability: nurturing a culture of continual improvement in software practices

Recognizing that change requires investment but pays off over time

Impact: Helping ECP teams to achieve:

- **Better:** Science, portability, robustness, composability
- **Faster:** Execution, development, dissemination
- **Cheaper:** Fewer staff hours and lines of code

IDEAS productivity
better scientific software

ECP EXASCALE COMPUTING PROJECT

ECP applications rely on ST products across all technical areas

24 ECP applications: National security, energy, Earth systems, economic security, materials, data
6 co-design centers: machine learning, graph analytics, mesh refinement, PDE discretization, particles, online data analytics

Consider ECP software technologies needed by 5 ECP applications:

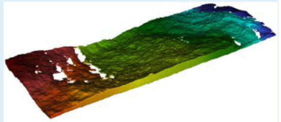
ExaWind: Turbine Wind Plant Efficiency

Harden wind plant design and layout against energy loss susceptibility; higher penetration of wind energy

Lead: NREL
DOE EERE

Subsurface: Carbon Capture, Fossil Fuel Extraction, Waste Disposal

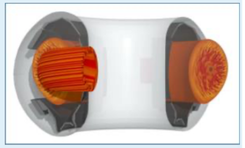
Reliably guide safe long-term consequential decisions about storage, sequestration, and exploration



Lead: LBNL
DOE BES, EERE, FE, NE

WDMApp: High-Fidelity Whole Device Modeling of Magnetically Confined Fusion Plasmas

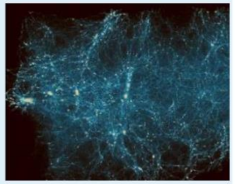
Prepare for ITER experiments and increase ROI of validation data and understanding; prepare for beyond-ITER devices



Lead: PPPL
DOE FES

ExaSky: Cosmological Probe of the Standard Model of Particle Physics

Unravel key unknowns in the dynamics of the Universe: dark energy, dark matter, and inflation

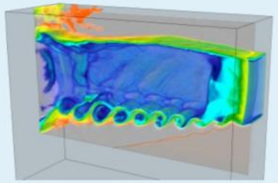


Lead: ANL
DOE HEP

The MARBL Multi-physics Code

Multi-physics simulations of high energy-density physics and focused experiments driven by high-explosive, magnetic or laser based energy sources

- Magneto-radiation-hydrodynamics at the exascale
- Next-generation pulsed power / ICF modeling
- High-order numerical methods



Lead: LLNL

EXASCALE COMPUTING PROJECT

ECP applications require consistency across the software stack

ECP Applications: C **Cosmology** (ExaSky) · F **Fusion Energy** (WDMApp) · N **National Security** (MAPP) · S **Subsurface Flow** · W **Wind Farm** (ExaWind) · **… and more**

24 apps, 6 co-design centers

Shown are 36 ST products (used or being considered by the 5 apps above)

ST overall has 70 unique software products used by 24 apps and 6 co-design centers

Selected ECP Software Technologies

Programming Models and Runtimes

MPI	C F N S W
Kokkos	F W
RAJA	N S
CHAI	N S
Umpire	N S

… and more

Tools and Technology

TAU	F W
HPCToolkit	C F N S
Flux	N
Caliper	N S
PAPI	C F S

Compilers and Support

LLVM	C F N S W
OpenMP	C F N W

… and more

Math Libraries (xSDK)

hypre	N S W
PETSc/TAO	F N S
STRUMPACK	F
SuperLU	F S W
Trilinos	F S W
SUNDIALS	N S
ArborX	C W
FFT	C W
BLAS, LAPACK	F W
MFEM	N

… and more

Visualization Analysis and Reduction

ALPINE	C N S W
Cinema	C N
VTK-m	C F N S W
SZ	C
zfp	F N
SPOT	N

… and more

Data Mgmt, I/O, Checkpoint Restart

SCR	N
MPI-IO	F
HDF5	C F N S W
PnetCDF	W
ADIOS	F W
UnifyFS	N
VeloC	C

… and more

Ecosystem: E4S at large

Spack	F N W

… and more

Legend:
- Prog Models & Runtimes
- Tools
- Math Libraries
- Data and Viz
- Ecosystems and Delivery

See E4S.io for more ST products
- AID
- AML
- BEE
- Darshan
- DTK
- Dyninst
- FleCSI
- ForTrilinos
- GASNet
- Ginkgo
- Kokkoskernels
- Legion
- libEnsemble
- MarFS
- NRM
- OpenACC
- Papyrus
- PaRSEC
- PDT
- PowerStack
- ScaLAPACK
- SCR
- SICM
- SLATE
- SWIG
- Tasmanian
- Umap
- UPC++

ECP apps rely on multiple software technologies; some software products contribute to multiple distinctly developed components of a multiphysics app (such as fusion energy modeling) that must run within a single executable.

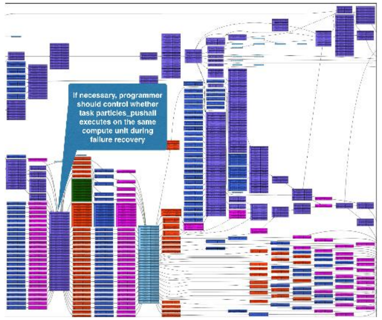
EXASCALE COMPUTING PROJECT

ECP **Software Technology (ST)**

Goal

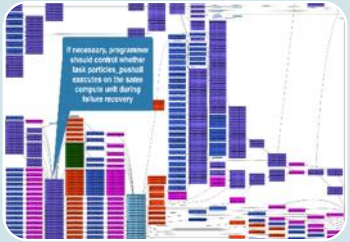
Build a comprehensive, coherent software stack that enables application developers to productively develop highly parallel applications that effectively target diverse exascale architectures

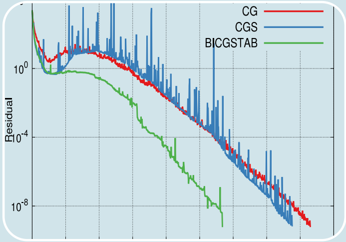
- Prepare SW stack for scalability with massive on-node parallelism
- Extend existing capabilities when possible, develop new when not
- Guide, and complement, and integrate with vendor efforts
- Develop and deliver high-quality and robust software products

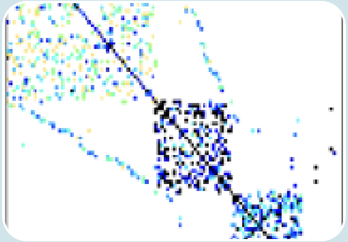


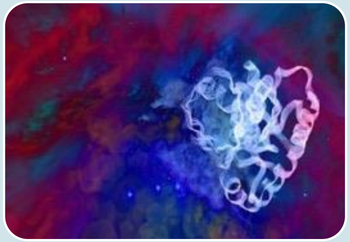
ECP ST has six technical areas

ECP ST Director: Mike Heroux
ECP ST Deputy Director: L.C. McInnes









E4S

NNSA National Nuclear Security Administration

Programming Models & Runtimes

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

Development Tools

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

Math Libraries

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

Data and Visualization

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

Software Ecosystem

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

NNSA ST

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

Area Leads: Rajeev Thakur | Jeff Vetter | Sherry Li | Jim Ahrens | Todd Munson | Kathryn Mohror

ST L4 Teams

- WBS
- Name
- PIs
- PCs - Project Coordinators

WBS	WBS Name	CAM/PI	PC
2.3	**Software Technology**	**Heroux, Mike, McInnes, Lois**	_
2.3.1	**Programming Models & Runtimes**	**Thakur, Rajeev**	_
2.3.1.01	PMR SDK	Shende, Sameer	Shende, Sameer
2.3.1.07	Exascale MPI (MPICH)	Balaji, Pavan	Guo, Yanfei
2.3.1.08	Legion	McCormick, Pat	McCormick, Pat
2.3.1.09	PaRSEC	Bosilica, George	Carr, Earl
2.3.1.14	Pagoda: UPC++/GASNet for Lightweight Communication and Global Address Space Support	Hargrove, Paul	Hargrove, Paul
2.3.1.16	SICM	Lang, Michael	Vigil, Brittney
2.3.1.17	OMPI-X	Bernholdt, David	Grundhoffer, Alicia
2.3.1.18	RAJA/Kokkos	Trott, Christian Robert	Trujillo, Gabrielle
2.3.1.19	Argo: Low-level resource management for the OS and runtime	Beckman, Pete	Gupta, Rinku
2.3.2	**Development Tools**	**Vetter, Jeff**	_
2.3.2.01	Development Tools Software Development Kit	Miller, Barton	Tim Haines
2.3.2.06	Exa-PAPI++: The Exascale Performance Application Programming Interface with Modern C++	Dongarra, Jack	Jagode, Heike
2.3.2.08	Extending HPCToolkit to Measure and Analyze Code Performance on Exascale Platforms	Mellor-Crummey, John	Meng, Xiaozhu
2.3.2.10	PROTEAS-TUNE	Vetter, Jeff	Glassbrook, Dick
2.3.2.11	SOLLVE: Scaling OpenMP with LLVm for Exascale	Chapman, Barbara	Kale, Vivek
2.3.2.12	FLANG	McCormick, Pat	Perry-Holby, Alexis
2.3.3	**Mathematical Libraries**	**Li, Sherry**	_
2.3.3.01	Extreme-scale Scientific xSDK for ECP	Yang, Ulrike	Yang, Ulrike
2.3.3.06	Preparing PETSc/TAO for Exascale	Munson, Todd	Munson, Todd
2.3.3.07	STRUMPACK/SuperLU/FFTX: sparse direct solvers, preconditioners, and FFT libraries	Li, Sherry	Li, Sherry
2.3.3.12	Enabling Time Integrators for Exascale Through SUNDIALS/ Hypre	Woodward, Carol	Woodward, Carol
2.3.3.13	CLOVER: Computational Libraries Optimized Via Exascale Research	Dongarra, Jack	Carr, Earl
2.3.3.14	ALExa: Accelerated Libraries for Exascale/ForTrilinos	Turner, John	Grundhoffer, Alicia
2.3.3.15	Sake: Scalable Algorithms and Kernels for Exascale	Rajamanickam, Siva	Trujillo, Gabrielle
2.3.4	**Data and Visualization**	**Ahrens, James**	_
2.3.4.01	Data and Visualization Software Development Kit	Atkins, Chuck	Bagha, Neelam
2.3.4.09	ADIOS Framework for Scientific Data on Exascale Systems	Klasky, Scott	Grundhoffer, Alicia
2.3.4.10	DataLib: Data Libraries and Services Enabling Exascale Science	Ross, Rob	Ross, Rob
2.3.4.13	ECP/VTK-m	Moreland, Kenneth	Moreland, Kenneth
2.3.4.14	VeloC: Very Low Overhead Transparent Multilevel Checkpoint/Restart/Sz	Cappello, Franck	Ehling, Scott
2.3.4.15	ExaIO - Delivering Efficient Parallel I/O on Exascale Computing Systems with HDF5 and Unify	Byna, Suren	Bagha, Neelam
2.3.4.16	ALPINE: Algorithms and Infrastructure for In Situ Visualization and Analysis/ZFP	Ahrens, James	Turton, Terry
2.3.5	**Software Ecosystem and Delivery**	**Munson, Todd**	_
2.3.5.01	Software Ecosystem and Delivery Software Development Kit	Willenbring, James M	Willenbring, James M
2.3.5.09	SW Packaging Technologies	Gamblin, Todd	Gamblin, Todd
2.3.5.10	ExaWorks	Laney, Dan	Laney, Dan
2.3.6	**NNSA ST**	**Mohror, Kathryn**	_
2.3.6.01	LANL ATDM	Mike Lang	Vandenbusch, Tanya Marie
2.3.6.02	LLNL ATDM	Becky Springmeyer	Gamblin, Todd
2.3.6.03	SNL ATDM	Jim Stewart	Trujillo, Gabrielle

- ~250 staff
- ~70 products
- 34 teams
- ~30 universities
- ~9 DOE labs
- 6 technical areas
- 1 focus area of 3 in ECP

ECP ST Stats

- 35 L4 subprojects
- 11 PI/PC same
- 24 PI/PC different
- ~27% ECP budget

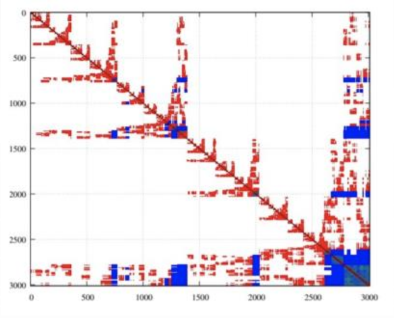
EXASCALE COMPUTING PROJECT

Recent advances in ECP software technologies as driven by needs of ECP apps

Scalable Solvers

Speeding sparse algorithms on CPUs and GPUs

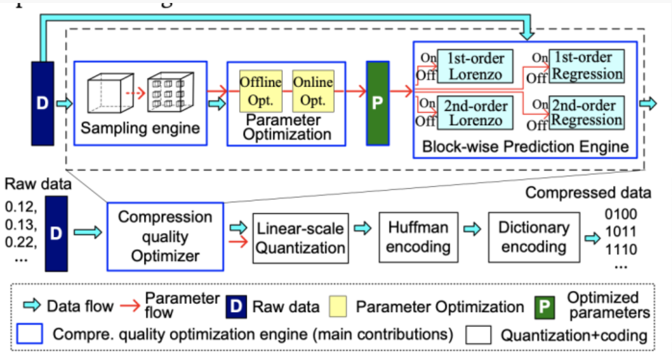
- The STRUMPACK team has developed new capabilities for multifrontal rank-structured preconditioning.
- **Impact:** STRUMPACK provides robust and scalable factorization-based methods for ill-conditioned and indefinite systems that arise in multiscale, multiphysics simulations.
- **More info:** https://www.exascaleproject.org/highlight/strumpack-speeds-sparse-algorithms-on-cpus-and-gpus



Lossy Compression

Optimizing lossy compression methods to manage data volumes

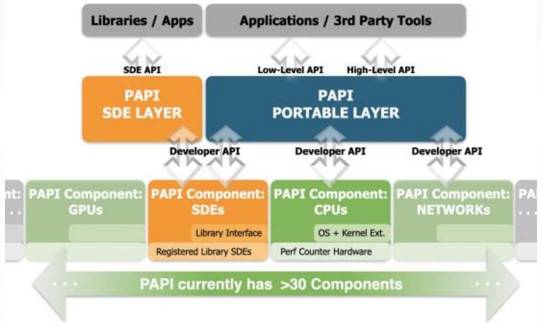
- The VeloC-SZ team has optimized SZ, an error-bounded prediction-based lossy compression model.
- **Impact:** SZ reduces dataset size while meeting users' speed and accuracy needs by storing the most pertinent data during simulation and experiments.
- **More info:** Significantly Improving Lossy Compression for HPC Datasets with Second-Order Prediction and Parameter Optimization, HPDC20, K. Zhao et al.



Performance Monitoring

Advancing performance counter monitoring capabilities for new ECP hardware

- The Exa-PAPI team provides a consistent interface and methodology for the use of low-level performance counter hardware found across the entire system (CPUs, GPUs, on/off-chip memory, interconnects, I/O system, energy/power, etc).
- **Impact:** Exa-PAPI enables users to see, in near real time, relations between software performance and hardware events.
- **More info:** https://icl.utk.edu/exa-papi



A few examples, among many more … all software available via **https://E4S.io**

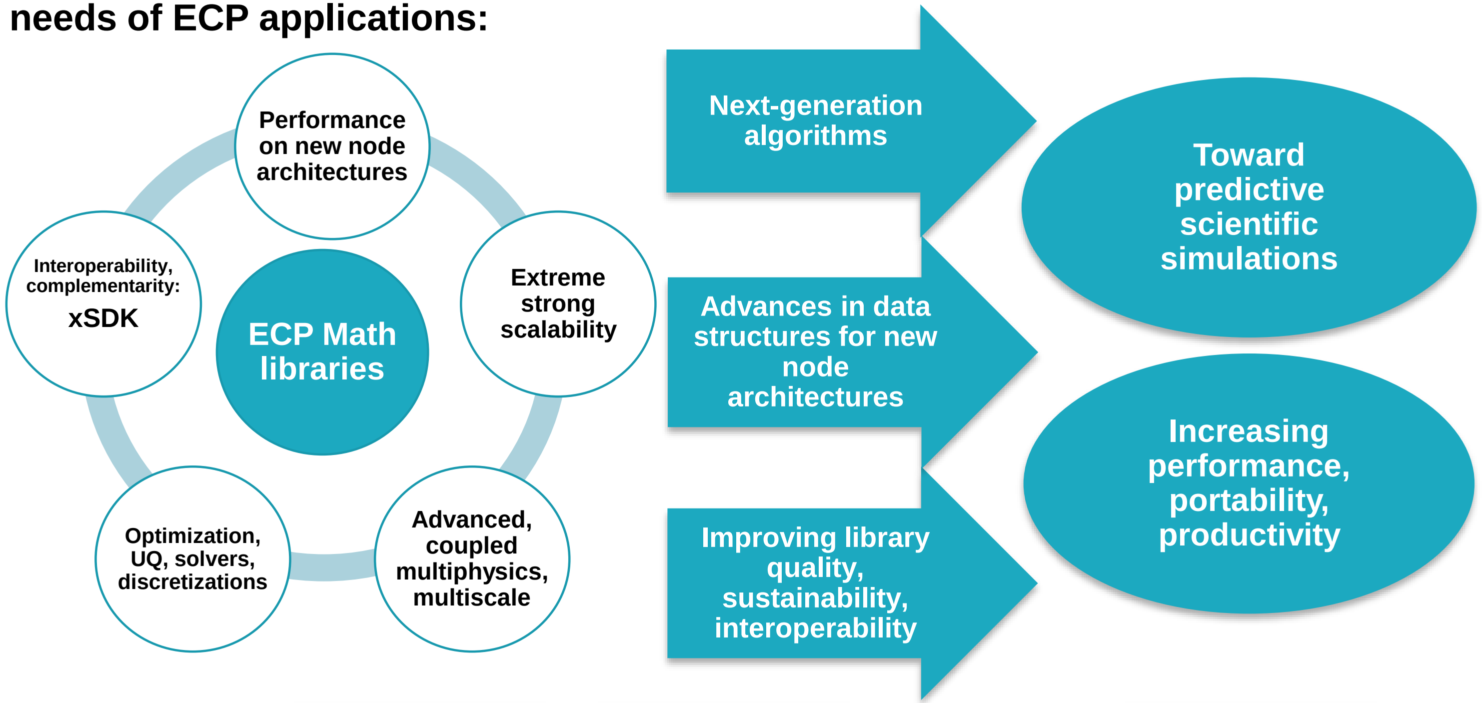
xSDK: Primary delivery mechanism for ECP math libraries' continual advancements

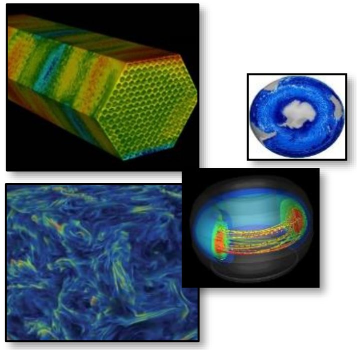

xSDK lead: Ulrike Meier Yang (LLNL)
xSDK release lead: Satish Balay (ANL)

xSDK release 0.7.0 (Nov 2021)

hypre
PETSc/TAO
SuperLU
Trilinos
AMReX
ArborX
ButterflyPACK
DTK
Ginkgo
heFFTe
libEnsemble
MAGMA
MFEM
Omega_h
PLASMA
PUMI
SLATE
Tasmanian
SUNDIALS
Strumpack
Alquimia
PFLOTRAN
deal.II, preCICE, PHIST, SLEPc } from the broader community

As motivated and validated by the needs of ECP applications:





Timeline: xSDK release 1 | xSDK release 2 | ..... | xSDK release n

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

Delivering an open, hierarchical software ecosystem

Levels of Integration | Product | Source and Delivery

ECP ST Open Product Integration Architecture

- 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

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 2021: E4S 21.11 – 91 full release products

Community Policies Commitment to software quality	DocPortal Single portal to all E4S product info	Portfolio testing Especially leadership platforms
Curated collection The end of dependency hell	Quarterly releases Release 1.2 – November	Build caches 10X build time improvement
Turnkey stack A new user experience	https://e4s.io	E4S Strategy Group US agencies, industry, international

Spack

https://spack.io

Spack lead: Todd Gamblin (LLNL)

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

EXASCALE COMPUTING PROJECT

E4S Community Policies: *A commitment to quality improvement*

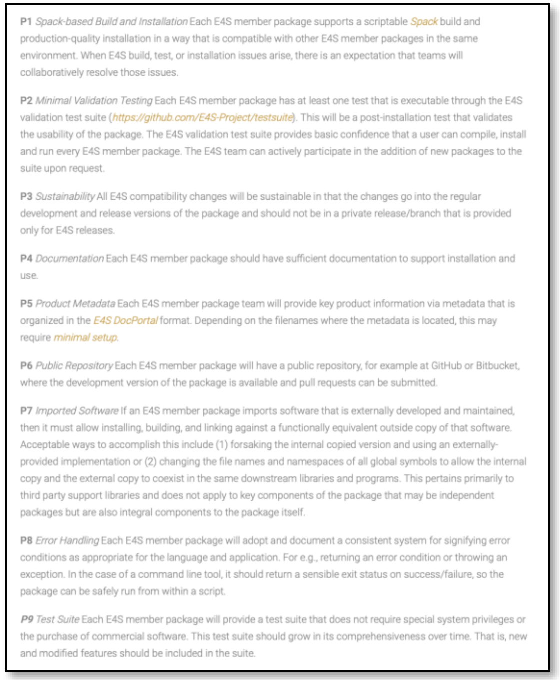
- Purpose: Enhance sustainability and interoperability
- Will serve as membership criteria for E4S
  - Membership is not required for *inclusion* in E4S
  - Also includes forward-looking draft policies
- Modeled after xSDK community policies
- Multi-year effort led by SDK team
  - Included representation from across ST
  - Multiple rounds of feedback incorporated from ST leadership and membership

SDK lead: Jim Willenbring (SNL)

Policies: Version 1

https://e4s-project.github.io/policies.html

- **P1: *Spack-based Build and Installation***
- **P2: *Minimal Validation Testing***
- **P3: *Sustainability***
- **P4: *Documentation***
- **P5: *Product Metadata***
- **P6: *Public Repository***
- **P7: *Imported Software***
- **P8: *Error Handling***
- **P9: *Test Suite***



P1 *Spack-based Build and Installation* Each E4S member package supports a scriptable *Spack* build and production-quality installation in a way that is compatible with other E4S member packages in the same environment. When E4S build, test, or installation issues arise, there is an expectation that teams will collaboratively resolve those issues.

P2 *Minimal Validation Testing* Each E4S member package has at least one test that is executable through the E4S validation test suite (*https://github.com/E4S-Project/testsuite*). This will be a post-installation test that validates the usability of the package. The E4S validation test suite provides basic confidence that a user can compile, install and run every E4S member package. The E4S team can actively participate in the addition of new packages to the suite upon request.

P3 *Sustainability* All E4S compatibility changes will be sustainable in that the changes go into the regular development and release versions of the package and should not be in a private release/branch that is provided only for E4S releases.

P4 *Documentation* Each E4S member package should have sufficient documentation to support installation and use.

P5 *Product Metadata* Each E4S member package team will provide key product information via metadata that is organized in the *E4S DocPortal* format. Depending on the filenames where the metadata is located, this may require *minimal setup*.

P6 *Public Repository* Each E4S member package will have a public repository, for example at GitHub or Bitbucket, where the development version of the package is available and pull requests can be submitted.

P7 *Imported Software* If an E4S member package imports software that is externally developed and maintained, then it must allow installing, building, and linking against a functionally equivalent outside copy of that software. Acceptable ways to accomplish this include (1) forsaking the internal copied version and using an externally-provided implementation or (2) changing the file names and namespaces of all global symbols to allow the internal copy and the external copy to coexist in the same downstream libraries and programs. This pertains primarily to third party support libraries and does not apply to key components of the package that may be independent packages but are also integral components to the package itself.

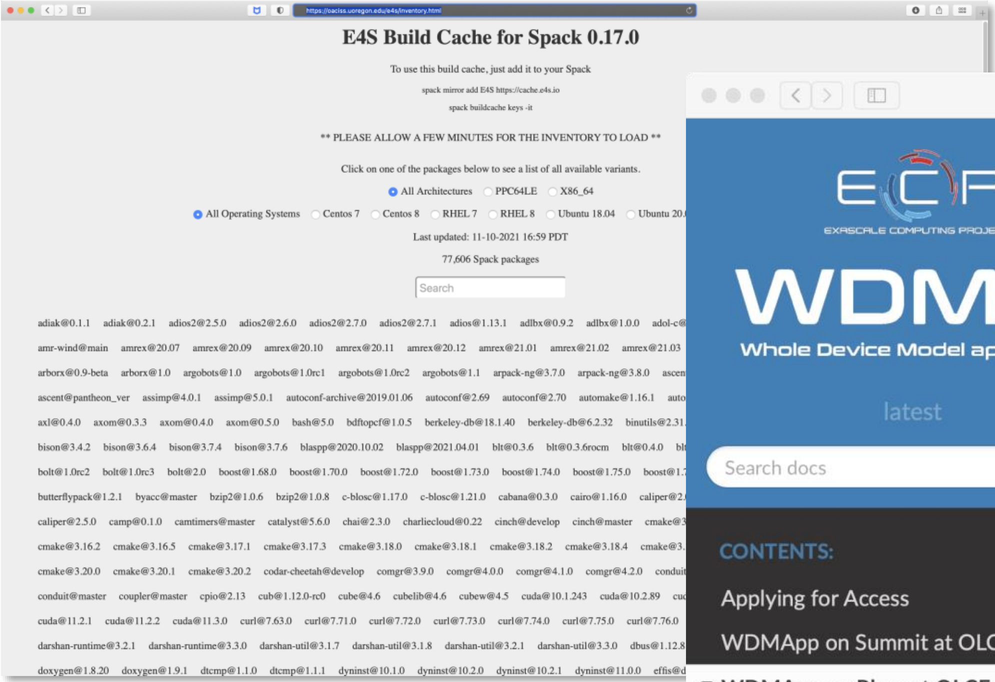
P8 *Error Handling* Each E4S member package will adopt and document a consistent system for signifying error conditions as appropriate for the language and application. For e.g., returning an error condition or throwing an exception. In the case of a command line tool, it should return a sensible exit status on success/failure, so the package can be safely run from within a script.

P9 *Test Suite* Each E4S member package will provide a test suite that does not require special system privileges or the purchase of commercial software. This test suite should grow in its comprehensiveness over time. That is, new and modified features should be included in the suite.

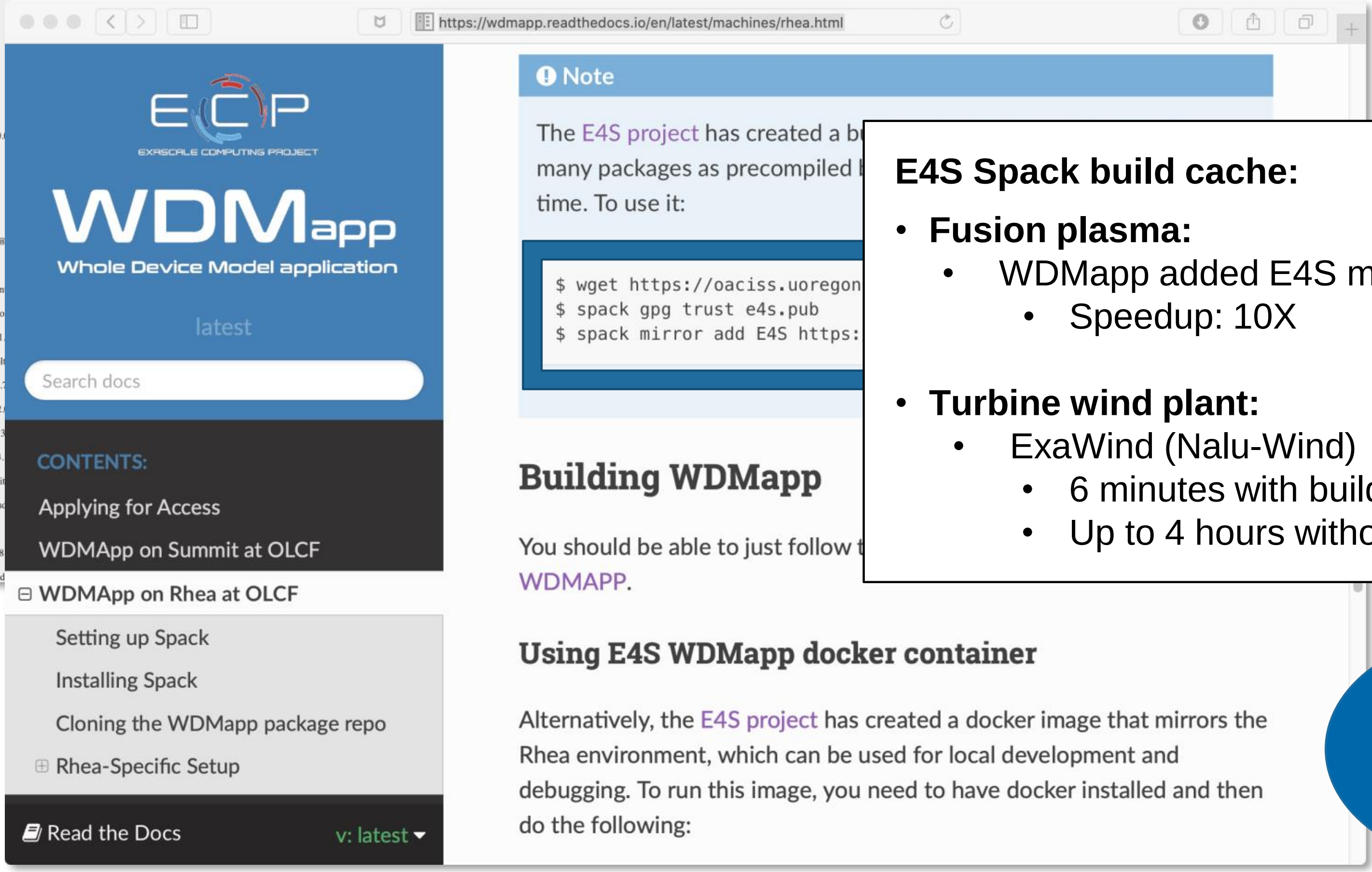
We welcome feedback. What policies make sense for your software?

Speeding up bare-metal installs using the E4S build cache

https://oaciss.uoregon.edu/e4s/inventory.html

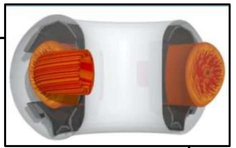


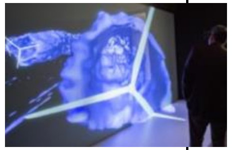
- 75,000+ binaries
- S3 mirror
- No need to build from source code!



E4S Spack build cache:

- **Fusion plasma:**
  - WDMapp added E4S mirror
    - Speedup: 10X
- **Turbine wind plant:**
  - ExaWind (Nalu-Wind)
    - 6 minutes with build cache
    - Up to 4 hours without





Special thanks to Sameer Shende, WDMapp and ExaWind teams

https://wdmapp.readthedocs.io/en/latest/machines/rhea.html

E4S summary

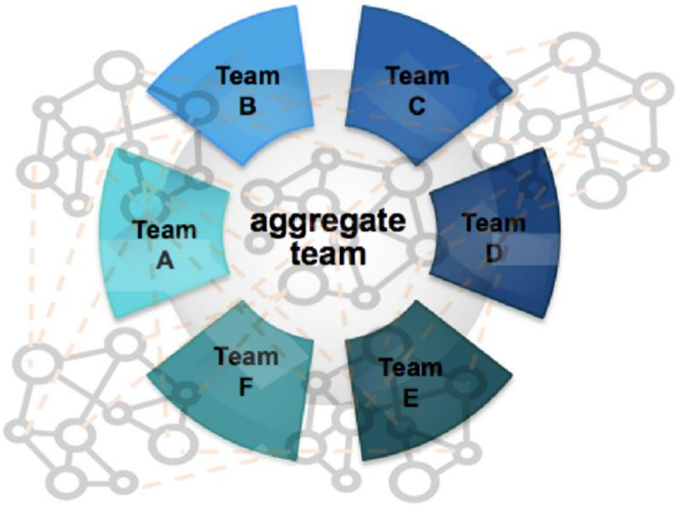
What E4S is not	What E4S is
• A closed system taking contributions only from DOE software development teams.	• Extensible, open architecture software ecosystem accepting contributions from US and international teams. • Framework for collaborative open-source product integration for ECP & beyond, including AI and Quantum.
• A monolithic, take-it-or-leave-it software behemoth.	• Full collection of compatible software capabilities **and** • Manifest of a la carte selectable software capabilities.
• A commercial product.	• Vehicle for delivering high-quality reusable software products in collaboration with others. • New entity in the HPC ecosystem enabling first-of-a-kind relationships with Facilities, vendors, other DOE program offices, other agencies, industry & international partners.
• A simple packaging of existing software.	• Hierarchical software framework to enhance (via SDKs) software interoperability and quality expectations. • Conduit for future leading edge HPC software targeting scalable computing platforms.

Building an ECP software ecosystem …

… While advancing scientific productivity through better scientific software

ECP: A “team of teams”

An aggressive research, development and deployment project, focused on delivery of mission-critical applications, an integrated software stack, and exascale hardware technology advances



Multilayered collaboration across the ECP community

- Ref: Scaling productivity and innovation on the path to exascale with a “team of teams” approach, E. Raybourn et al, 2019

Networked teams, at scale. Multidisciplinary expertise, such as:

Research software engineers (RSEs)

Applications scientists

Computer scientists

Applied mathematicians

Project coordinators

Performance engineers

Computational scientists and engineers

Cognitive and social scientists

Data scientists and engineers

Stakeholders

And more …

A shout-out to 4 terrific RSEs on the forefront of work toward exascale

… among many more terrific RSEs working in research labs, universities and industry

Lisa Childers
Technical Development Lead
Argonne, ALCF
Focus: Workload and resource management (and tracking) on extreme-scale machines; facilitating user interactions with extreme-scale machines to improve productivity in scientific pursuits.

Rinku Gupta
Research Software Specialist
Argonne, MCS
Focus: HPC software design, development, leadership (for resource management, fault tolerance, checkpointing); improving software productivity and sustainability; editor-in-chief of the Better Scientific Software website (BSSw.io); lead of RSE movement at Argonne.

Ken Raffenetti
Principal Software Development Specialist
Argonne, MCS
Focus: Parallel programming models and communication libraries; definition of the Message Passing Interface (MPI) standard and key maintainer of MPICH; member of PMIx Administrative Steering Committee.

Junchao Zhang
Software Engineer
Argonne, MCS
Focus: Developer of PETSc, focusing on software scalability, maintainability, and user support, with emphasis on communication and computation efficiency on heterogeneous architectures with GPUs.

https://ideas-productivity.org

Advancing scientific productivity through better scientific software

Science through computing is only as good as the software that produces it.

1 Customize and curate methodologies

- Target scientific software productivity and sustainability
- Use workflow for best practices content development

3 Establish software communities

- Determine community policies to improve software quality and compatibility
- Create Software Development Kits (SDKs) to facilitate the combined use of complementary libraries and tools

2 Incrementally and iteratively improve software practices

- Determine high-priority topics for improvement and track progress
- *Productivity and Sustainability Improvement Planning (PSIP)*

4 Engage in community outreach

- Broad community partnerships
- Collaboration with computing facilities
- Webinars, tutorials, events
- *WhatIs* and *HowTo* docs
- Better Scientific Software site (https://bssw.io)

IDEAS-ECP team

https://www.ideas-productivity.org


Ann Almgren (LBNL)


Ross Bartlett (SNL)


David E. Bernholdt (ORNL)
Institutional PI


Anshu Dubey (ANL)
Institutional PI


Elsa Gonsiorowski (LLNL)
Institutional PI


Patricia Grubel (LANL)


Rinku Gupta (ANL)


Rebecca Hartman-Baker (LBNL)
Computing Facility Liaison


Michael Heroux (SNL)
Lead Co-PI


Axel Huebl (LBNL)


Lena Lopatina (LANL)
Computing Facility Liaison


Rose Lynch (ANL)


Addi Malviya Thakur (ORNL)


Osni Marques (LBNL)
Institutional PI


Lois Curfman McInnes (ANL)
Lead Co-PI


Reed Milewicz (SNL)

Argonne NATIONAL LABORATORY | BERKELEY LAB | Lawrence Livermore National Laboratory | Los Alamos NATIONAL LABORATORY EST. 1943 | OAK RIDGE National Laboratory | Sandia National Laboratories | UNIVERSITY OF OREGON


Mark Miller (LLNL)


J. David Moulton (LANL)
Institutional PI


Hai Ah Nam (LBNL)


Boyana Norris (U. Oregon)
Institutional PI


Elaine Raybourn (SNL)
Institutional PI


Katherine Riley (ANL)
Computing Facility Liaison


David Rogers (ORNL)
Computing Facility Liaison


Jean Shuler (LLNL)
Computing Facility Liaison


Ben Sims (LANL)


Deborah Stevens (ANL)


Greg Watson (ORNL)


Jim Willenbring (SNL)

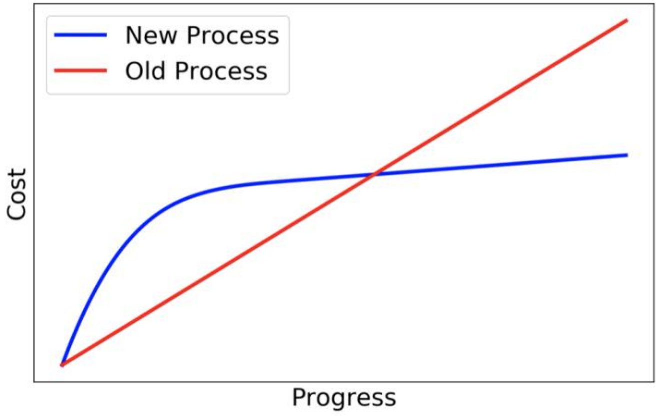
IDEAS-ECP Alumni

- Satish Balay (ANL)
- Lisa Childers (ANL)
- Todd Gamblin (LLNL)
- Judy Hill (ORNL)
- Steve Hudson (ANL)
- Christoph Junghans (LANL)
- Alicia Klinvex (SNL)
- Shannon Lindgren (ANL)
- Jared O'Neal (ANL)
- Michele Rosso (LBNL)
- Barry Smith (ANL)
- Louis Vernon (LANL)
- Paul Wolfenbarger (SNL)

Ref: Research Software Science: A Scientific Approach to Understanding and Improving How We Develop and Use Software for Research, M. Heroux, 2019

Productivity and sustainability improvement planning:
Recent successes with PSIP on HDF5

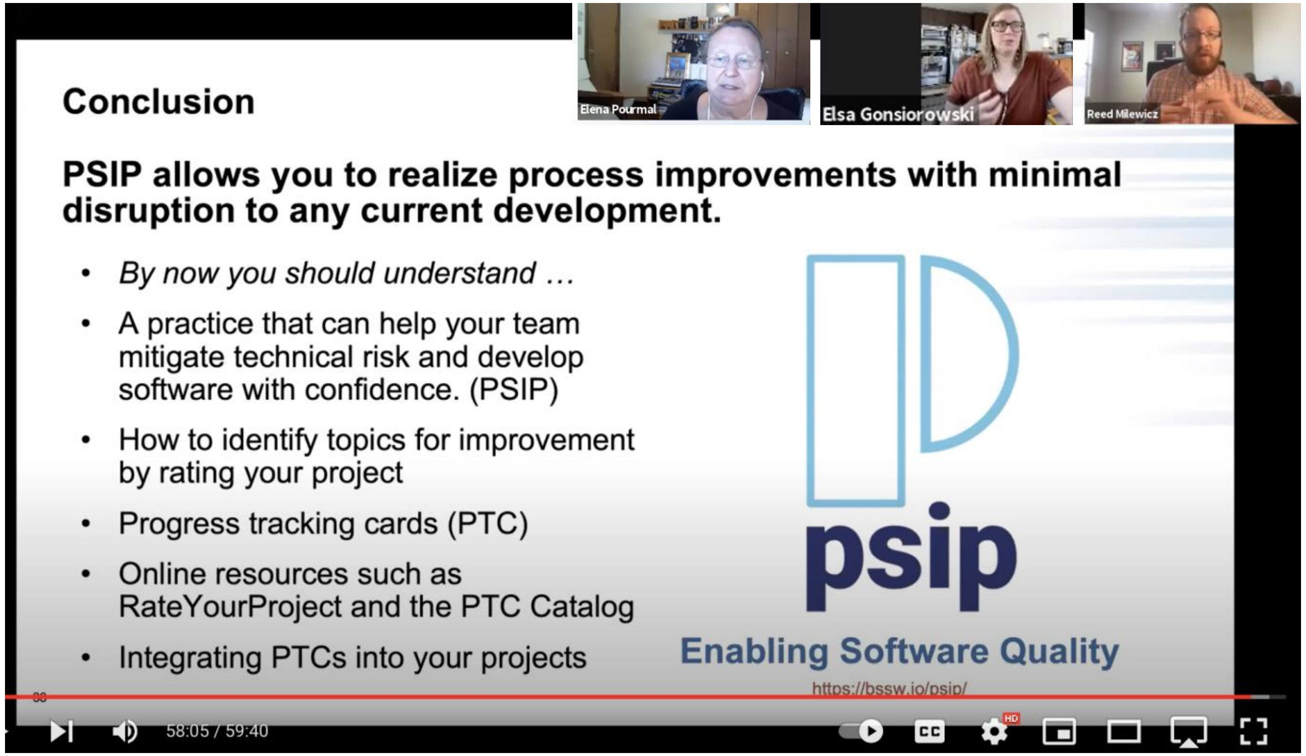
https://bssw.io/psip



HDF5 improvement goals - achieved by using PSIP progress tracking cards (PTC)

- Modernize processes for handling **documentation** (PTC)
- Move HDF5 from a THG managed Bitbucket instance to **GitHub** (PTC)
- Define and adopt a set of consistent **coding standards** (PTC)

"The PSIP project had an immediate impact on our community. With the GitHub move we see increasing amounts of small but very valuable contributions to make HDF5 code and documentation better." **– Elena Pourmal, Director of Engineering, The HDF Group**



Refs:

- Using the PSIP Toolkit to Achieve Your Goals – A Case Study at The HDF Group, E. Pourmal, R. Milewicz, E. Gonsiorowski, webinar, June 2020 [recording / slides]
- Recent successes with PSIP on HDF5, M. Miller, E. Pourmal, E. Gonsiorowski, Nov 2020
- Automating Software Productivity Planning: Lightweight Tools for Upgrading Team Practices, E. Raybourn et al, the International Conference on Software Engineering Research & Practice, SERP'21, July 2021

IDEAS Outreach

Lead: David Bernholdt

Better Scientific Software Tutorials

- Covering issues of developer productivity, software sustainability and reliability, with a special focus on the challenges of complex, large-scale HPC
  - software design, agile methodologies, Git workflows, reproducibility, software testing, continuous integration testing, refactoring, and more
- https://bssw-tutorial.github.io
- Recent venues
  - Supercomputing (2016-2021)
  - SEAS's Improving Scientific Software (2021)
  - ECP Annual Meeting (2017-2021)
  - ISC (2017-2019, 2021), ATPESC (2016-2021)

BSSw Tutorial @ SC21, Nov 15

Mailing list to follow IDEAS-led events (webinars, panels, BOFs, etc.): http://eepurl.com/cQCyJ5

Webinar Series: Best Practices for HPC Software Developers (HPC-BP)

- Covering topics in software development and HPC
- https://ideas-productivity.org/events/hpc-best-practices-webinars
- Lead: Osni Marques
- Presented by the community to the community
- Monthly series, since May 2016 (offered live and archived)
  - Best Practices for HPC Software Developers: The First Five Years of the Webinar Series, O. Marques and D. Bernholdt, Oct 2021

IDEAS Outreach
Lead: David Bernholdt

Technical Meetings and Birds of a Feather Sessions

- Creating opportunities to talk about software development, productivity, and sustainability
- https://ideas-productivity.org/events
- Minisymposia
  - SIAM CSE, SIAM PP (2015-2022), PASC (2018, 2019)
  - Ref: A Look at Software-Focused Topics at SIAM CSE21, March 2021
- Thematic poster sessions
  - SIAM CSE (2017, 2019, 2021)
- BOF sessions
  - Software Engineering and Reuse in Modeling, Simulation and Data Analytics for Science and Engineering
    - Supercomputing (2015-2021), ISC (2019)
- Collegeville Workshop Series on Scientific Software,
  - Ref: Software Team Experiences and Challenges, K. Beattie et al, Oct 2021

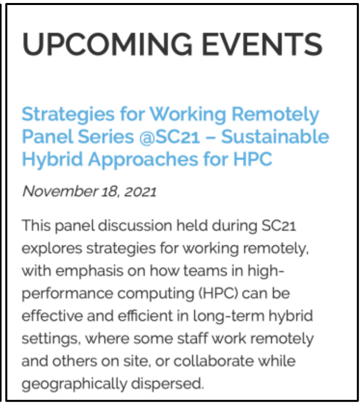
Panel Series: Performance Portability & ECP

- Lead: Anshu Dubey (2020 series). Refs:
  - Performance Portability in the Exascale Computing Project: Exploration Through a Panel Series, A. Dubey et al, IEEE CiSE, Sept 2021
  - SIAM CSE21 minisymposium: https://doi.org/10.6084/m9.figshare.c.5321441
  - Minisymposium accepted for ECCOMAS 2022

Panel Series: Strategies for Working Remotely

- Exploring strategies for working remotely, with emphasis on how HPC teams can be effective and efficient in long-term hybrid settings
- https://www.exascaleproject.org/strategies-for-working-remotely
- Lead: Elaine Raybourn
- Quarterly series, since April 2020 (offered live and archived)
- Ref: Why We Need Strategies for Working Remotely: The ECP Panel Series, E. Raybourn, SC20 State of the Practice, Nov 2020

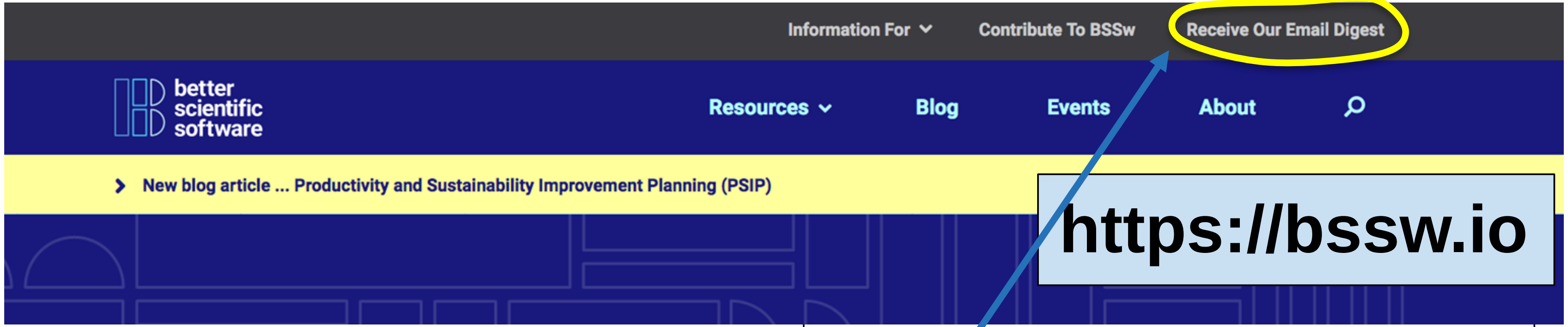
BOF @ SC21 Nov 16	Panel @ SC21 Nov 18



EXASCALE COMPUTING PROJECT

better scientific software

So your code will see the future.



https://bssw.io

What is BSSw?

Community-based hub for sharing information on practices, techniques, and tools to improve developer productivity and software sustainability for computational science.

We want and *need* contributions from the community … Join us!

- **Types of content**
  - Informative articles
  - Curated links
    - Highlight other web-based content
  - Events
  - WhatIs, HowTo docs
  - Blog articles

BSSw.io editor in chief: Rinku Gupta

Receive our email digest

Recent articles

- The Contributions of Scientific Sofware to Scientific Discovery, K. Keahey & R. Gupta
- Software Team Experiences and Challenges, C. Balos, J. Brown. G. Chourdakis et al.
- Performance Portability and the ECP Project, A. Dubey
- Testing Non-Deterministic Research Software, N. Eisty,
- What Does This Line Do? The Challenge of Writing a Well-Documented Code, M. Stoyanov

Better Scientific Software: 2020 Highlights

SHARE

- Unit Testing C++ with Catch, M. Dewing
- The Art of Writing Scientific Software in an Academic Environment, H. Anzt
- FLASH5 Refactoring and PSIP, A. Dubey & J. O'Neal
- Software Sustainability in the Molecular Sciences, T. Windus & T.D. Crawford
- Working Effectively with Legacy Code, R. Bartlett
- Building Community through Software Policies, P. Luszczek & U.M. Yang
- Continuous Technology Refreshment: An Introduction Using Recent Tech Refresh Experiences on VisIt, M. Miller & H. Auten

EXASCALE COMPUTING PROJECT

BSSw Fellowship: Meet the Fellows

better scientific software

https://bssw.io/fellowship

Meet Our Fellows

The BSSw Fellowship program gives recognition and funding to leaders and advocates of high-quality scientific software. Meet the Fellows and Honorable Mentions and learn more about how they impact Better Scientific Software.

Fellowships Overview | Apply | Meet Our Fellows | BSSw Fellowship FAQ

Community Growth

2018 - 2021

2021 Class

Fellows

Marisol García-Reyes
Farallon Institute
Increasing accessibility of data & cloud technologies

Mary Ann Leung
Sustainable Horizons Institute
Increasing developer productivity and innovation through diversity

Chase Million
Million Concepts
Project management best practices for research software

Amy Roberts
University of Colorado Denver
Enabling collaboration through version control user stories

Honorable Mentions

Keith Beattie
Lawrence Berkeley National Laboratory
Computational Research Division, Computer Systems Engineer

Julia Stewart Lowndes
National Center for Ecological Analysis and Synthesis (NCEAS), UC Santa Barbara
Openscapes Director

Jonathan Madsen
Lawrence Berkeley National Laboratory
NERSC, Application Performance Specialist

Addi Thakur Malviya
Oak Ridge National Laboratory
Software Engineering Group, Group Leader

EXASCALE COMPUTING PROJECT

BSSw Fellowship Coordinator: Hai Ah Nam

[2020 BSSw Fellows: Projects and Perspectives]

KRELL institute

U.S. DEPARTMENT OF ENERGY

NSF

Working to broaden participation of underrepresented groups in HPC

ECP Broader Engagement Task Force: Multipronged HPC initiative to complement and leverage existing lab programs in workforce development

- Across DOE labs: Share experiences & best practices, “Intro to HPC” training/outreach, internships w. mentoring/community
- More info: https://www.exascaleproject.org/hpc-workforce
- Partnership with Mary Ann Leung, founder & president of Sustainable Horizons Institute

SUSTAINABLE HORIZONS INSTITUTE

SRP-HPC Internship Program: Expand **Sustainable Research Pathways** program across ECP teams

Nominee in 2021 for the HPCWire Reader's Choice Award in the **Workforce Diversity and Inclusion** category

https://science.osti.gov/-/media/ascr/ascac/pdf/meetings/202107/ASCAC_meeting_202107_Challenges_Lessons_Expanding_CSE.pdf

Ref: Increasing Productivity by Broadening Participation in Scientific Software Communities, M.A. Leung, D. Rouson and L.C. McInnes, 2020

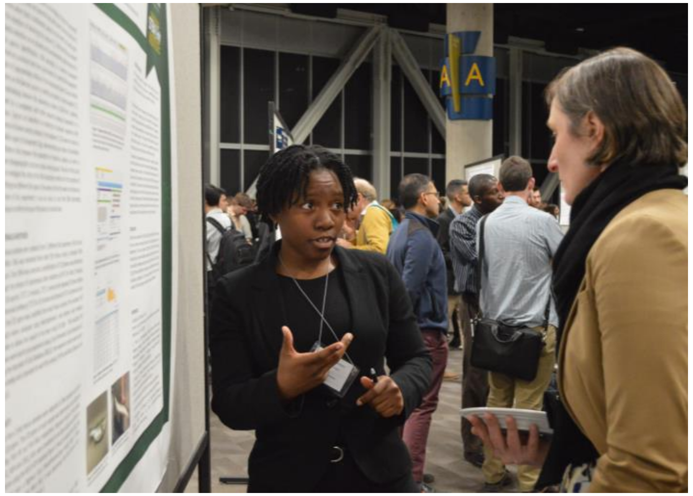
SRP-HPC Internship Program

Sustainable Research Pathways for HPC
Broadening participation of underrepresented groups

Application deadline: December 31, 2021

See: https://www.exascaleproject.org/hpc-workforce

Collaborate with ECP teams



Two tracks*

- Faculty/student teams
- Students on their own

Includes:

- Matching workshop to explore possible research collaborations
- Onboarding & welcome at 2022 Exascale Computing Project (ECP) Annual Meeting, including the start of professional/career development activities
- Summer research experience
- Participation in 2023 ECP Annual Meeting to present research results and engage in the HPC community
- Community building throughout and ongoing

* **Students from (and faculty working with) underrepresented groups** (Black or African American, Hispanic/Latinx, American Indian, Alaska Native, Native Hawaiian, and Pacific Islanders, women, and persons with disabilities) **are strongly encouraged to apply.**

ECP EXASCALE COMPUTING PROJECT

We must explicitly consider **community software ecosystem perspectives** for next-generation computational science

- What software ecosystems do you want to use and be a part of?
- E4S (http://e4s.io) is an extensible, open architecture software ecosystem; contributions and feedback are welcome!

- Software ecosystems require high-quality software, but many complex, intertwined challenges exist.
- Community efforts are working to overcome technical and sociological challenges in scientific software … **Get involved!**

Investment in software quality pays off (better, faster, cheaper).

Software quality is a critical component of quality science.

Call to action for the HPC community:

Each of us must change our expectations and behavior. To consider:

- **Do you develop and use HPC software?**
  - Investigate resources for software improvement
  - Advocate for and lead change in your projects
  - Disseminate insights about software improvement from your own work (blogs, presentations, posters, papers, etc)
  - Check out community activities, such as the Research Software Engineering (RSE) movement
- **Do you lead projects or organizations where teams develop and use HPC software?**
  - Encourage continual software quality improvement
  - Provide clear career paths and mentoring for scientific software professionals, such as RSEs
- **Are you a stakeholder or supporter of projects that develop and use HPC software?**
  - Incorporate expectations of software quality and sustainability, including funding for people to do this important work
  - Incorporate expectations for transparency and reproducibility
- **Everyone**
  - Work toward changes in software citations/credit models, metrics
  - Work toward changes in incentives, training and education

Working toward software sustainability: Join the conversation

Leadership Scientific Software (LSSw) Portal

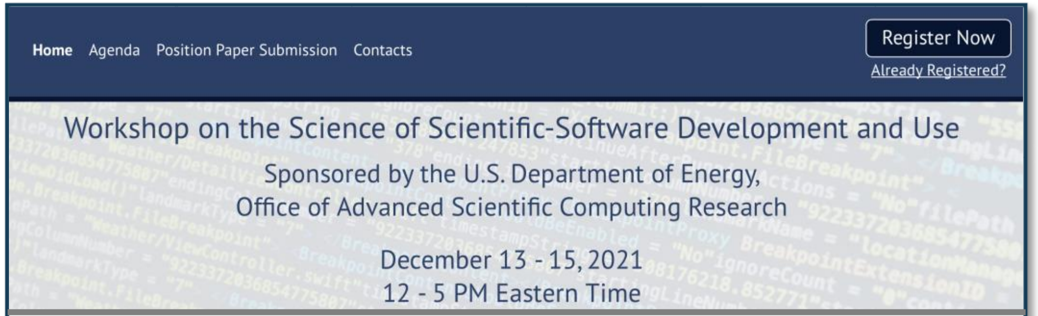
https://lssw.io

The LSSw portal is dedicated to building community and understanding around the development and sustainable delivery of leadership scientific software

- LSSw Town Hall Meetings (ongoing)
  - 3rd Thursday each month, 3 – 4:30 pm Eastern US time
- Slack: Share your ideas interactively
- Whitepapers: Written content for LSSw conversations
  - We need your ideas (2-4 page whitepapers)
  - Submit via GitHub PR or attachment to contribute@lssw.io
- References
  - Help us build a reading list
  - Submit via GitHub PR or email to contribute@lssw.io

Workshop on Research Software Science

Software is an increasingly important component in the pursuit of scientific discovery. Both its development and use are essential activities for many scientific teams. At the same time, very little scientific study has been conducted to understand, characterize, and improve the development and use of software for science.



- Info and registration at: https://www.orau.gov/SSSDU2021
- Whitepaper deadline: Nov 19, 2021

EXASCALE COMPUTING PROJECT

Recent software-related events at SC21 … Get involved!

Day/Time	Event Type	Event Title (see linked program page for full details)	Tags
Sunday, Nov. 14 8:00am-5:00pm CST	Tutorial	Managing HPC Software Complexity with Spack	
Sunday, Nov. 14 9:00am-5:30pm CST	Workshop	P3HPC: 2021 International Workshop on Performance, Portability, and Productivity in HPC	
Monday, Nov. 15 8:00am-5:00pm CST	Tutorial	Better Scientific Software	
Monday, Nov. 15 9:00am-5:30pm CST	Workshop	RSE-HPC-2021: Research Software Engineers in HPC: Creating Community, Building Careers, Addressing Challenges	
Tuesday, Nov. 16 2:15pm-3:00pm CST	Invited Talk	The Importance of Diverse Perspectives in Advancing HPC	HPC Community Collaboration; Workforce; Diversity Equity Inclusion (DEI)
Tuesday, Nov. 16 5:15pm-6:45pm CST	BOF	Software Engineering and Reuse in Modeling, Simulation, and Data Analytics for Science and Engineering	Computational Science; Software Engineering
Tuesday, Nov. 16 5:15pm-6:45pm CST	BOF	Strengthening Reproducibility for SC21 and Beyond	Reproducibility and Transparency
Wednesday, Nov. 17 12:15pm-1:15pm CST	BOF	Words Matter! Promoting Inclusion through Language in Advanced Research Computing	Diversity Equity Inclusion (DEI)
Wednesday, Nov. 17 1:30pm-2:15pm CST	Invited Talk	Powering HPC Discoveries through Scientific Software Ecosystems and Communities	Reproducibility and Transparency; Software Engineering; HPC Community Collaboration; Computational Science

Recent software-related events at SC21

(continued)

Day/Time	Event Type	Event Title (see linked program page for full details)	
Wednesday, Nov. 17 1:30pm-3:00pm CST	Panel	Performance and Correctness Tools for Extreme-Scale Computing	Correctness · Performance · System Software and Runtime Systems
Wednesday, Nov. 17 5:15pm-6:45pm CST	BOF	HPC Carpentry: Introducing New Users to HPC	Data Analytics · Professional Development · Software Engineering
Thursday, Nov. 18 12:15pm-1:15pm CST	BOF	Ethics in HPC	Diversity Equity Inclusion (DEI)
Thursday, Nov. 18 12:15pm-1:15pm CST	BOF	Spack Community BoF	HPC Community Collaboration · Reproducibility and Transparency · Software Engineering
Thursday, Nov. 18 12:15pm-1:15pm CST	BOF	Towards FAIR for Machine Learning (ML) models	Machine Learning and Artificial Intelligence
Thursday, Nov. 18 3:30pm-5:00pm CST	Panel	Strategies for Working Remotely: Sustainable Hybrid Approaches for HPC	Workforce
Friday, Nov. 19 8:30am-12:00pm CST	Workshop	Correctness 2021: 5th International Workshop on Software Correctness for HPC Applications	In cooperation with IEEE computer society TCHPC
Friday, Nov. 19 10:30am-12:00pm CST	Panel	Reproducibility in HPC: Passing Fad or a Work in Progress?	Reproducibility and Transparency

References

More info about the impact of ECP software technologies

https://exascaleproject.org
A few highlights … Check back for the latest ECP news

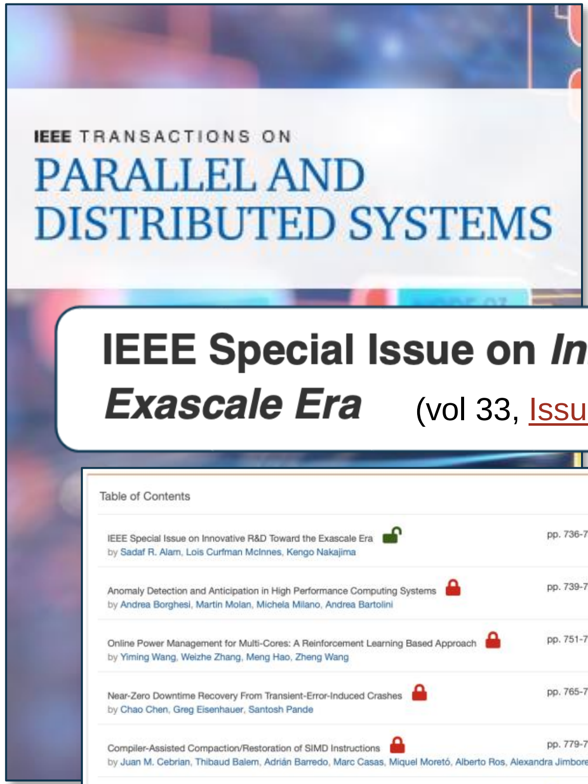
- **ECP News**
  - An Exascale Day Interview with ORNL's Doug Kothe, Director of ECP
  - ECP-funded team investigates NVM techniques to improve data storage & performance speed
  - ECP-funded researchers enable faster time-to-science with novel I/O processing method
  - ECP project optimizes lossy compression methods to manage big science data volumes
  - ALPINE project tests novel algorithm for in situ exascale data analysis
  - Workflow technologies impact SC20 Gordon Bell COVID-19 award winner & two of three finalists
  - The Extreme-Scale Scientific Software Stack (E4S): A new resource for computational and data science research

- **Technical Highlights**
  - LLVM Holds the Keys to Exascale Supercomputing
  - ECP Leads the Way to Cross-Platform Tested and Verified Compilers for HPC and Exascale Architectures
  - RAJA Portability Suite Enables Performance Portable CPU and GPU HPC Codes
  - A New Approach in the HYPRE Library Brings Performant GPU-based Algebraic MultiGrid to Exascale Supercomputers and the General HPC Community
  - Clacc – An Open Source OpenACC Compiler and Source Code Translation Project
  - The ECP SuperLU Library Speeds the Direct Solution of Large Sparse Linear Systems on HPC and Exascale Hardware
  - ECP Provides TAU, a CPU/GPU/MPI Profiler, for All HPC and Exascale Machines
  - HeFFTe – A Widely Applicable, CPU/GPU, Scalable Multidimensional FFT That Can Even Support Exascale Supercomputers

- **Let's Talk Exascale Podcast:**
  - Sunita Chandrasekaran Reflects on Teaching Supercomputing and Leading the ECP SOLLVE Project
  - Supporting Scientific Discovery and Data Analysis in the Exascale Era
  - ECP Leadership Discusses Project Highlights, Challenges, and the Expected Impact of Exascale Computing
  - Flexible Package Manager Automates the Deployment of Software on Supercomputers
  - Accelerating the Adoption of Container Technologies for Exascale Computing
  - Simplifying the Deployment of High-Performance Computing Tools and Libraries
  - Method Enables Collaborative Software Teams to Enhance Effectiveness and Efficiency
  - Tackling the Complex Task of Software Deployment and Continuous Integration at Facilities
  - Optimizing Math Libraries to Prepare Applications for Exascale Computing

EXASCALE COMPUTING PROJECT

Special issue of IEEE TPDS, upcoming conferences



IEEE Special Issue on *Innovative R&D Toward the Exascale Era* (vol 33, Issue: 4, April 2022 -- online Oct 2021)

Table of Contents

https://www.siam.org/conferences/cm/conference/pp22

PASC22 Challenge: Computing and Data … for all Humankind
https://pasc22.pasc-conference.org

ST Capability Assessment Report (CAR)

- ST software products discussed in this presentation are presented with more detail and further citations.
- We classify ECP ST product deployment as broad, moderate, or experimental.
  - Broad and moderate deployment is typically suitable for collaboration.
  - Web links are available for almost all products.
  - All ECP ST products are available as part of the Extreme-scale Scientific Software Stack (E4S) http://e4s.io.


EXASCALE COMPUTING PROJECT

ECP-RPT-ST-0002-2020–Public

ECP Software Technology Capability Assessment Report–Public

Michael A. Heroux, Director ECP ST
Lois Curfman McInnes, Deputy Director ECP ST
Rajeev Thakur, Programming Models & Runtimes Lead
Jeffrey S. Vetter, Development Tools Lead
Sherry Li, Mathematical Libraries Lead
James Ahrens, Data & Visualization Lead
Todd Munson, Software Ecosystem & Delivery Lead
Kathryn Mohror, NNSA ST Lead

November 19, 2020

https://www.exascaleproject.org/wp-content/uploads/2021/01/ECP-ST-CAR-v2.5.pdf

Advancing Scientific Productivity through Better Scientific Software:
Developer Productivity & Software Sustainability Report

Disruptive changes in computer architectures and the complexities of tackling new frontiers in extreme-scale modeling, simulation, and analysis present daunting challenges to software productivity and sustainability.

This report explains the IDEAS approach, outcomes, and impact of work (in partnership with the ECP and broader computational science community).

Target readers are all those who care about the quality and integrity of scientific discoveries based on simulation and analysis. While the difficulties of extreme-scale computing intensify software challenges, issues are relevant across all computing scales, given universal increases in complexity and the need to ensure the trustworthiness of computational results.

https://exascaleproject.org/better-scientific-productivity-through-better-scientific-software-the-ideas-report

Thank you

Abstract

HPC software is a cornerstone of long-term collaboration and scientific progress, but software complexity is increasing due to disruptive changes in computer architectures and the challenges of next-generation science. Thus, the HPC community has the unique opportunity to fundamentally change how scientific software is designed, developed, and sustained—embracing community collaboration toward scientific software ecosystems, while fostering a diverse HPC workforce who embody a broad range of skills and perspectives. This webinar will introduce work in the U.S. Exascale Computing Project, where a varied suite of scientific applications builds on programming models and runtimes, math libraries, data and visualization packages, and development tools that comprise the Extreme-scale Scientific Software Stack (E4S). The webinar will introduce crosscutting strategies that are increasing developer productivity and software sustainability, thereby mitigating technical risks by building a firmer foundation for reproducible, sustainable science. The webinar will also mention complementary community efforts and opportunities for involvement.