Scientific software ecosystems and communities:
Why we need them and how each of us can help them thrive

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
Science and beyond: Applications and discovery in ECP

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

And more:
- 24 applications
- 6 co-design centers
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](https://doi.org/10.1137/17M1143826), U. Rüde, K. Willcox, L.C. McInnes, H. De Sterck, *SIAM Review*, 2018

**SIAM Activity Group on Computational Science and Engineering**
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, …
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

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
Thank you to my collaborators and communities

- 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 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.

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 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.

**Selected program office application development**

**System procurement projects and facilities**

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

**Exascale Computing Project (ECP)**
ECP by the numbers

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

- 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

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

- More than 80 top-notch R&D teams

- Hundreds of consequential milestones delivered on schedule and within budget since project inception
ECP’s holistic approach uses co-design and integration to achieve exascale computing

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

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

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

Perfomrant mission and science applications at scale

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

Andrew Siegel, AD Director
Erik Draeger, AD Deputy Director

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

Katie Antypas, HI Director
Susan Coghlan, HI Deputy Director
### Science and beyond: Applications and discovery in ECP

<table>
<thead>
<tr>
<th>National security</th>
<th>Energy security</th>
<th>Economic security</th>
<th>Scientific discovery</th>
<th>Earth systems</th>
<th>Health care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next-generation, <strong>stockpile stewardship</strong> codes</td>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Additive manufacturing</strong> of qualifiable metal parts</td>
<td><strong>Cosmological probe</strong> of the standard model of particle physics</td>
<td>Accurate regional impact assessments in <strong>Earth system models</strong></td>
<td>Accelerate and translate <strong>cancer research</strong> (partnership with NIH)</td>
</tr>
<tr>
<td>Reentry-vehicle-environment simulation</td>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td>Reliable and efficient planning of the <strong>power grid</strong></td>
<td>Validate fundamental laws of nature</td>
<td>Stress-resistant crop analysis and catalytic conversion of <strong>biomass-derived alcohols</strong></td>
<td>Thank you to Andrew Siegel, Erik Draeger and ECP applications teams</td>
</tr>
<tr>
<td>Multi-physics science simulations of high-energy density <strong>physics</strong> conditions</td>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Seismic</strong> hazard risk assessment</td>
<td><strong>Plasma wakefield accelerator</strong> design</td>
<td><strong>Metagenomics</strong> for analysis of biogeochemical cycles, climate change, environmental remediation</td>
<td></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td>Subsurface use for <strong>carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Light source-enabled analysis of protein and molecular structure</strong> and design</td>
<td><strong>Find, predict, and control materials and properties</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Predict and control magnetically confined fusion plasmas</strong></td>
<td><strong>Predict and control magnetically confined fusion plasmas</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Turbine <strong>wind plant</strong> efficiency</td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Design and commercialization of <strong>SMRs</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
<tr>
<td>Nuclear fission and fusion reactor <strong>materials design</strong></td>
<td><strong>Subsurface use for carbon capture</strong>, petroleum extraction, waste disposal</td>
<td><strong>High-efficiency</strong>, <strong>low-emission combustion engine</strong> and gas turbine design</td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Demyystify origin of chemical elements</strong></td>
<td><strong>Earth systems</strong></td>
</tr>
</tbody>
</table>
DOE HPC Roadmap to Exascale Systems

<table>
<thead>
<tr>
<th>FY 2012</th>
<th>FY 2016</th>
<th>FY 2018</th>
<th>FY 2021</th>
<th>FY 2022</th>
<th>FY 2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Titan</td>
<td>Summit</td>
<td>ORNL</td>
<td>ORNL</td>
<td>ANL</td>
<td>Exascale Systems</td>
</tr>
<tr>
<td>ORNL Cray/AMD/NVIDIA</td>
<td>ORNL IBM/NVIDIA</td>
<td>ORNL HPE/AMD</td>
<td>ANL HPE/AMD/NVIDIA</td>
<td>ANL Intel/HPE</td>
<td></td>
</tr>
<tr>
<td>Mira</td>
<td>Theta</td>
<td>ANL</td>
<td>ANL</td>
<td>LBNL</td>
<td></td>
</tr>
<tr>
<td>ANL IBM BG/Q</td>
<td>Cray/Intel KNL</td>
<td>Cray/Intel KNL</td>
<td>HPE/AMD/NVIDIA</td>
<td>HPE/AMD/NVIDIA</td>
<td></td>
</tr>
<tr>
<td>LBNL</td>
<td>Cori</td>
<td>LBNL</td>
<td>LBNL</td>
<td>LLNL/SNL</td>
<td></td>
</tr>
<tr>
<td>Cray/Intel Xeon/KNL</td>
<td>Cray/Intel Xeon/KNL</td>
<td>HPE/AMD/NVIDIA</td>
<td>Cray/Intel Xeon/KNL</td>
<td>HPE/Intel</td>
<td></td>
</tr>
<tr>
<td>Sequoia</td>
<td>Trinity</td>
<td>Sierra</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLNL IBM BG/Q</td>
<td></td>
<td>LLNL IBM/NVIDIA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trinity</td>
<td>Sierra</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LANL/SNL Cray/Intel Xeon/KNL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLNL HPE/AMD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Version 2.0
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.

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

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

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

Addressing a National Imperative

The Exascale Computing Project is an aggressive research, development, and deployment project focused on delivery of mission-critical applications, an integrated software stack, and exascale hardware technology advances.

Training and Productivity

Lead: Ashley Barker, Oak Ridge National Laboratory

For applications to take full advantage of exascale hardware and software, a robust developer training and productivity program keeps application and software team members, staff, and other stakeholders abreast of emerging technologies and key technologies of importance to ECP. These projects are done in close collaboration between the facilities, vendors, and the ECP community.

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
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

- **WDMAApp: 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
ECP applications require consistency across the software stack.

**Selected ECP Software Technologies**

<table>
<thead>
<tr>
<th>Programming Models and Runtimes</th>
<th>Tools and Technology</th>
<th>Math Libraries (xSDK)</th>
<th>Visualization Analysis and Reduction</th>
<th>Data Mgmt, I/O, Checkpoint Restart</th>
<th>Ecosystem: E4S at large</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPI</td>
<td>TAU</td>
<td>h ypre</td>
<td>ALPINE, Cmfnaw</td>
<td>SCR</td>
<td>Spack</td>
</tr>
<tr>
<td>Kokkos</td>
<td>HPCToolkit</td>
<td>PETSc/TAO</td>
<td>Cinema</td>
<td>MPI-IO</td>
<td></td>
</tr>
<tr>
<td>RAJA</td>
<td>Flux</td>
<td>STRUMPACK</td>
<td>VTK-m</td>
<td>HDF5</td>
<td></td>
</tr>
<tr>
<td>CHAI</td>
<td>Caliper</td>
<td>SuperLU</td>
<td>S2</td>
<td>PnetCDF</td>
<td></td>
</tr>
<tr>
<td>Umpire</td>
<td>PAPI</td>
<td>Trilinos</td>
<td>zip</td>
<td>ADIOS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SUNDIALS</td>
<td>SPOT</td>
<td>UnityFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ArborX</td>
<td>VeloC</td>
<td>VeloC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FFT</td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
<tr>
<td>Compilers and Support</td>
<td></td>
<td></td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
<tr>
<td>LLVM</td>
<td></td>
<td></td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
<tr>
<td>OpenMP</td>
<td></td>
<td></td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BLAS, LAPACK</td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MFEM</td>
<td>… and more</td>
<td>… and more</td>
<td></td>
</tr>
</tbody>
</table>

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.
**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

**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

ECP ST Director: Mike Heroux
ECP ST Deputy Director: L.C. McInnes
<table>
<thead>
<tr>
<th>WBS</th>
<th>WBS Name</th>
<th>CAM/PI</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>Software Technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.1</td>
<td>Programming Models &amp; Runtimes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.1.01</td>
<td>PMR SDK</td>
<td>Thakur, Rajeef</td>
<td></td>
</tr>
<tr>
<td>2.3.1.07</td>
<td>Exascale MPI (MPI5)</td>
<td>Badje, Sameer</td>
<td></td>
</tr>
<tr>
<td>2.3.1.08</td>
<td>Legion</td>
<td>Balaji, Pavan</td>
<td></td>
</tr>
<tr>
<td>2.3.1.09</td>
<td>PaRSEC</td>
<td>McCormick, Pat</td>
<td></td>
</tr>
<tr>
<td>2.3.1.14</td>
<td>Pagoda: UPC++/GASNet for Lightweight Communication and Global Address Space Support</td>
<td>Bosilica, George</td>
<td></td>
</tr>
<tr>
<td>2.3.1.16</td>
<td>SICM</td>
<td>Biskin, Michael</td>
<td></td>
</tr>
<tr>
<td>2.3.1.17</td>
<td>OMPI-X</td>
<td>B overhead, David</td>
<td></td>
</tr>
<tr>
<td>2.3.1.18</td>
<td>RAJA/Kokkos</td>
<td>B overhead, David</td>
<td></td>
</tr>
<tr>
<td>2.3.1.19</td>
<td>Argo: Low-level resource management for the OS and runtime</td>
<td>Shende, Sameer</td>
<td></td>
</tr>
<tr>
<td>2.3.2</td>
<td>Development Tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.2.01</td>
<td>Development Tools Software Development Kit</td>
<td>Vetter, Jeff</td>
<td></td>
</tr>
<tr>
<td>2.3.2.06</td>
<td>Exa-PAPI++: The Exascale Performance Application</td>
<td>Miller, Barton</td>
<td></td>
</tr>
<tr>
<td>2.3.2.08</td>
<td>Extending HPC Toolkit to Measure and Analyze Code</td>
<td>Vetter, Jeff</td>
<td></td>
</tr>
<tr>
<td>2.3.2.10</td>
<td>PROTEAS-TUNE</td>
<td>ASH, Christian Robert</td>
<td></td>
</tr>
<tr>
<td>2.3.2.11</td>
<td>SOLVE: Scaling OpenMP with LLVM for Exascale</td>
<td>ASH, Christian Robert</td>
<td></td>
</tr>
<tr>
<td>2.3.2.12</td>
<td>FLANG</td>
<td>ASH, Christian Robert</td>
<td></td>
</tr>
<tr>
<td>2.3.3</td>
<td>Mathematical Libraries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.3.01</td>
<td>Extreme-scale Scientific xSDK for ECP</td>
<td>Li, Sherry</td>
<td></td>
</tr>
<tr>
<td>2.3.3.06</td>
<td>Preparing PETSc/TAO for Exascale</td>
<td>Meng, Xiaozhu</td>
<td></td>
</tr>
<tr>
<td>2.3.3.07</td>
<td>STRUMPACK/SuperLU/FFTX: sparse direct solvers</td>
<td>Woodward, Carol</td>
<td></td>
</tr>
<tr>
<td>2.3.3.12</td>
<td>Enabling Time Integrators for Exascale Through SU</td>
<td>Woodward, Carol</td>
<td></td>
</tr>
<tr>
<td>2.3.3.13</td>
<td>CLOVER: Computational Libraries Optimized Via Ex</td>
<td>Woodward, Carol</td>
<td></td>
</tr>
<tr>
<td>2.3.3.14</td>
<td>ALExa: Accelerated Libraries for Exascale/ForTrilinos</td>
<td>Vetter, Jeff</td>
<td></td>
</tr>
<tr>
<td>2.3.3.15</td>
<td>Sake: Scalable Algorithms and Kernels for Exascale</td>
<td>Vetter, Jeff</td>
<td></td>
</tr>
<tr>
<td>2.3.4</td>
<td>Data and Visualization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.4.01</td>
<td>Data and Visualization Software Development Kit</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.06</td>
<td>ADIOS Framework for Scientific Data on Exascale</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.10</td>
<td>DataLib: Data Libraries and Services Enabling Exascale</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.13</td>
<td>ECP/VTK-m</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.14</td>
<td>VeLOC: Very Low Overhead Transparent Multilevel</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.15</td>
<td>ExaIO - Delivering Efficient Parallel I/O on Exascale</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.4.16</td>
<td>ALPINE: Algorithms and Infrastructure for In Situ Visualization</td>
<td>Ahrens, James</td>
<td></td>
</tr>
<tr>
<td>2.3.5</td>
<td>Software Ecosystem and Delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.5.01</td>
<td>Software Ecosystem and Delivery Software Development Kit</td>
<td>Willenbring, James M</td>
<td></td>
</tr>
<tr>
<td>2.3.5.09</td>
<td>SW Packaging Technologies</td>
<td>Willenbring, James M</td>
<td></td>
</tr>
<tr>
<td>2.3.5.10</td>
<td>ExaWorks</td>
<td>Willenbring, James M</td>
<td></td>
</tr>
<tr>
<td>2.3.6</td>
<td>NNSA ST</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.6.01</td>
<td>LANL ATDM</td>
<td>Mike Lang</td>
<td></td>
</tr>
<tr>
<td>2.3.6.02</td>
<td>LLNL ATDM</td>
<td>Becky Springmeyer</td>
<td></td>
</tr>
<tr>
<td>2.3.6.03</td>
<td>SNL ATDM</td>
<td>Jim Stewart</td>
<td></td>
</tr>
</tbody>
</table>

**ECP ST Stats**

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

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

As motivated and validated by the needs of ECP applications:

- Extreme strong scalability
- Optimization, UQ, solvers, discretizations
- Advanced, coupled multiphysics, multiscale
- Performance on new node architectures
- Interoperability, complementarity: xSDK

Timeline:
- xSDK release 0.7.0 (Nov 2021)
- 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

- **Group similar products**
- **Make interoperable**
- **Assure policy compliant**
- **Include external products**

### Product

- **E4S**
  - Source: ECP E4S team; Non-ECP Products (all dependencies)
  - Delivery: spack install e4s; containers; CI Testing

- **SDKs**
  - Source: SDK teams; Non-ECP teams (policy compliant, spackified)
  - Delivery: Apps directly; spack install sdk; future: vendor/facility

- **ST Products**
  - Source: ECP L4 teams; Non-ECP Developers; Standards Groups
  - Delivery: Apps directly; spack; vendor stack; facility stack

### Source and Delivery

- **ECP ST Open Product Integration Architecture**
- **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

https://e4s.io

E4S lead: Sameer Shende (U Oregon)

Also includes other products, e.g.,
AI: PyTorch, TensorFlow, Horovod
Co-Design: AMReX, Cabana, MFEM
E4S Community Policies: A commitment to quality improvement

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**

- **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)

---

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
# E4S summary

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

2. **Incrementally and iteratively improve software practices**
   - Determine high-priority topics for improvement and track progress
   - *Productivity and Sustainability Improvement Planning (PSIP)*

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

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](https://bssw.io))
IDEAS-ECP team

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 (LLNL)
- 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 

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
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)

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)

Mailing list to follow IDEAS-led events (webinars, panels, BOFs, etc.): http://eepurl.com/cQCyJ5
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
    — Ref: A Look at Software-Focused Topics at SIAM CSE21, March 2021
  • Thematic poster sessions
  • 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:
  — 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

Panel Series: Performance Portability & ECP

• Lead: Anshu Dubey (2020 series). Refs:
  — 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

Panel Series: Performance Portability & ECP

• Lead: Anshu Dubey (2020 series). Refs:
  — 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

Panel Series: Performance Portability & ECP

• Lead: Anshu Dubey (2020 series). Refs:
  — 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
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

Recent articles

• The Contributions of Scientific Software 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

• 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
BSSw Fellowship: Meet the 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.

2018 - 2021

Meet Our Fellows

Community Growth

2018 Class

- Hai Ah Nam
- Jennifer Kukula
- Haoyu Cai
- Ronghong Xue
- Erin Conner
- Dan Jurkin
- Ehsan Bagherzadeh
- Marilise Groot
- Maria Souto

2019 Class

- Matthew Weimer
- Ryan McQuiggan
- John Crocker
- Venkatasen
- Alexander Fiedler
- Tao Yang
- Tianyu Qin
- Yi-Song Wang
- Shuai Li

2020 Class

- Naveen Dukkipati
- Andrew Moore
- Brian Marchetto
- Ian Allen
- Yoni Wiener
- Kaushik Ghose
- Benji Rizzo
- Bongki Min
- Saurabh Parashar

2021 Class

- Marisol García-Reyes
- Mary Ann Leung
- Chase Million
- Amy Roberts

Fellows

- Farallon Institute
- Sustainable Horizons Institute
- Million Concepts
- University of Colorado Boulder

Increasing accessibility of data & cloud technologies
- Increasing developer productivity and innovation through diversity
- Project management best practices for research software
- Enabling collaboration through version control user stories

Honorable Mentions

- Keith Beattie
- Sarah Lowndes
- Jonathan Madsen
- Addi Thakur

Lawrence Berkeley National Laboratory
- National Center for Ecological Analysis and Synthesis (NCEAS)
- Lawrence Berkeley National Laboratory
- Oak Ridge National Laboratory

Supporting scientific computing innovation
- Supporting science and engineering through computational software infrastructure
- Supporting the development of new computational tools and methods
- Supporting the development of new computational tools and methods
- Supporting the development of new computational tools and methods

BSSw Fellowship Coordinator: Hai Ah Nam

2020 BSSw Fellows: Projects and Perspectives

https://bssw.io/fellowship
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](https://www.exascaleproject.org/hpc-workforce)
- Partnership with Mary Ann Leung, founder & president of Sustainable Horizons Institute

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


---

**Sustainable Research Pathways: Basic Idea**

- Build relationships centered on research collaborations
- Recruit
  - Faculty working with underrepresented students
  - Students from underrepresented groups
- Provide opportunities for staff scientists
- Research collaborations
- Learn/contribute to diversity and inclusion efforts
- Supplement existing D&I Laboratory programs

---

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

---

Challenges and Lessons

Learned in Expanding Participation in Computational Science and Engineering

Advanced Scientific Computing Advisory Committee

Thursday, July 29, 2021

Mary Ann Leung, Ph.D.

Sustainable Horizons Institute

[https://science.osti.gov/-/media/ascr/asrac/pdf/meetings/202107/ASCAC_meeting_202107_Challenges_Lessons_Expanding_CSE.pdf](https://science.osti.gov/-/media/ascr/asrac/pdf/meetings/202107/ASCAC_meeting_202107_Challenges_Lessons_Expanding_CSE.pdf)
SRP-HPC Internship Program
Sustainable Research Pathways for HPC
Broadening participation of underrepresented groups

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.

Application deadline: December 31, 2021
See: https://www.exascaleproject.org/hpc-workforce
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

- **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

- **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

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

<table>
<thead>
<tr>
<th>Day/Time</th>
<th>Event Type</th>
<th>Event Title (see linked program page for full details)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wednesday, Nov. 17 1:30pm-3:00pm CST</td>
<td>Panel</td>
<td>Performance and Correctness Tools for Extreme-Scale Computing</td>
</tr>
<tr>
<td>Wednesday, Nov. 17 5:15pm-6:45pm CST</td>
<td>BOF</td>
<td>HPC Carpentry: Introducing New Users to HPC</td>
</tr>
<tr>
<td>Thursday, Nov. 18 12:15pm-1:15pm CST</td>
<td>BOF</td>
<td>Ethics in HPC</td>
</tr>
<tr>
<td>Thursday, Nov. 18 12:15pm-1:15pm CST</td>
<td>BOF</td>
<td>Spack Community BoF</td>
</tr>
<tr>
<td>Thursday, Nov. 18 12:15pm-1:15pm CST</td>
<td>BOF</td>
<td>Towards FAIR for Machine Learning (ML) models</td>
</tr>
<tr>
<td>Thursday, Nov. 18 3:30pm-5:00pm CST</td>
<td>Panel</td>
<td>Strategies for Working Remotely: Sustainable Hybrid Approaches for HPC</td>
</tr>
<tr>
<td>Friday, Nov. 19 8:30am-12:00pm CST</td>
<td>Workshop</td>
<td>Correctness 2021: 5th International Workshop on Software Correctness for HPC Applications</td>
</tr>
<tr>
<td>Friday, Nov. 19 10:30am-12:00pm CST</td>
<td>Panel</td>
<td>Reproducibility in HPC: Passing Fad or a Work in Progress?</td>
</tr>
</tbody>
</table>
References
More info about the impact of ECP software technologies

**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

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

| IEEE Special Issue on Innovative R&D Toward the Exascale Era | p. 781-785 |
| Online Power Management for Workloads in a Hierarchical Learning Based Approach | p. 761-766 |
| New Data Center Exchanger: Towards a novel Fluid Cooler | p. 765-770 |
| Simple Assembly Configuration/Reduction of SIMI Instructions | p. 776-781 |
| CODEC: A Framework for High-Performance Scalable Parallel Inference | p. 790-805 |
| OpenMP 4.6 | p. 806-811 |
| Keras 2.0: Programming Language for the Exascale Era | p. 812-817 |
| Design and Performance Characterization of Xilinx FPGAs for Leadership Class Patronys | p. 818-823 |
| LKROM: A Nozomi-Based Acceleration Library for Multi-CPU Applications | p. 824-829 |
| The Pulkett المجتمع للعلوم التكنولوجية الدولية | p. 850-852 |

Related conferences:

- 24 papers by the international community

PASC22 Challenge: Computing and Data … for all Humankind

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

Hybrid: SIAM Conference on Parallel Processing for Scientific Computing (PP22)
February 23 - 26, 2022
Hyatt Regency Seattle | Seattle, Washington, U.S.

#SIAMPP22

The PASC22 Conference

June 27 to 29, 2022

Read more

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](http://e4s.io).

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.