

LSSw Meeting 13

June 15, 2023

Heads up for LSSw Meeting 14: July 20, 2020

Topic: Software Foundations

LSSw Meeting 13: June 15, 2023

Topic: Post-ECP Seed Project Outreach Updates: A panel discussion

Description: The Post-ECP Software-Ecosystem Seed Projects have conducted workshops and otherwise reached out to the scientific community to gather input and formulate working plans. During this panel, these projects will provide an outbriefing of these discussions. In opening remarks, the panelists will briefly describe their outcomes. Then we will open up the discussion for Q&A.

Panelists, their affiliations, and seed project titles are:

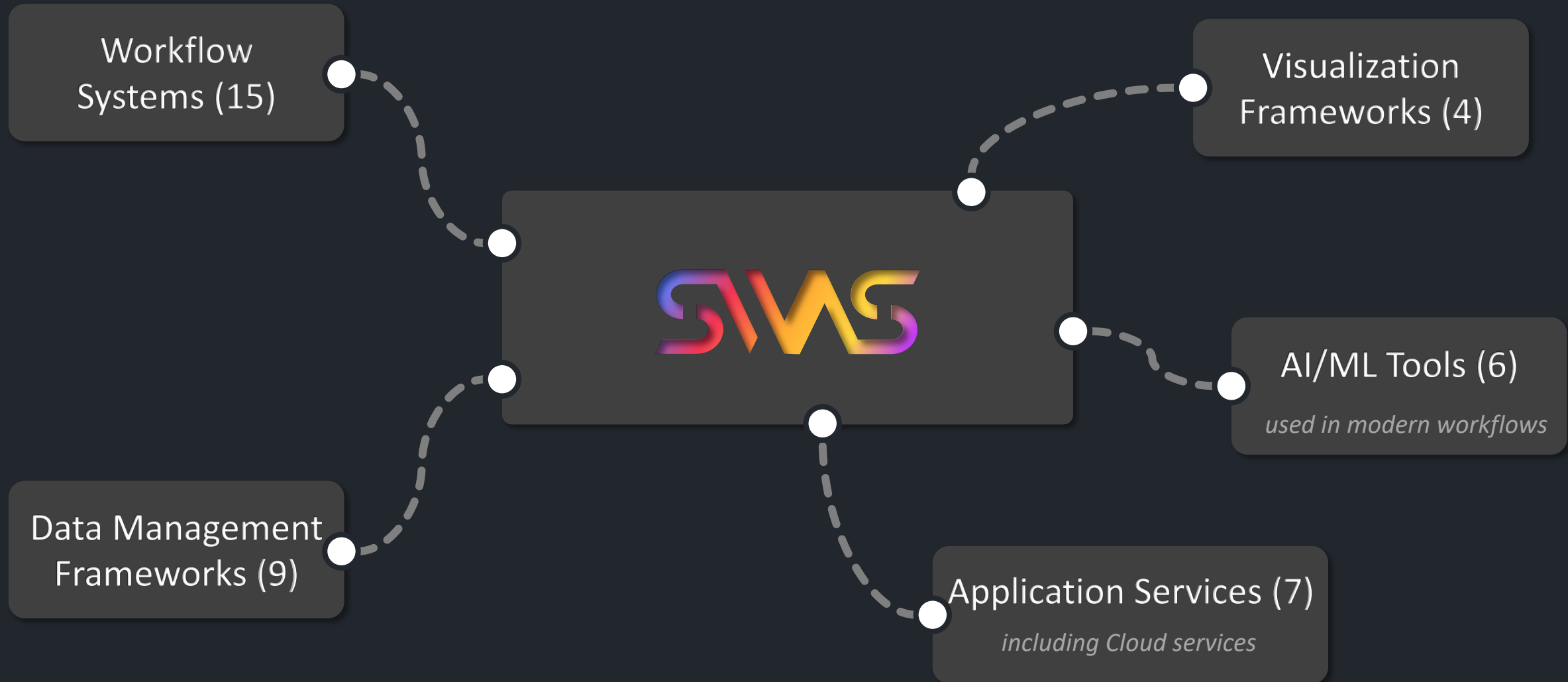
- Anshu Dubey, Argonne National Lab, COLABS: Collaboration of Oak Ridge National Lab, LBNL, and ANL for Better Software
- Rafael Ferreira da Silva, Oak Ridge National Lab, SWAS: Center for Sustaining Workflows and Application Services
- Michael Heroux, Sandia National Labs, PESO: Toward a Post-ECP Software-Sustainability Organization
- Terry Jones, Oak Ridge National Lab, STEP: Sustainable Tools Ecosystem Project
- Keita Teranishi, Oak Ridge National Lab, S4PST: Sustainability for Node Level Programming Systems and Tools
- Gregory Watson, Oak Ridge National Lab, OSSF: Open Scientific Software Foundation



sustaining workflows
& application services

<https://swas.center>

Target Software



What We Offer

SWAS fosters a **collaborative ecosystem** to improve project sustainability in the workflows and application services community through **knowledge sharing and connections**



Get Involved

The SWAS seedling effort is planning to organize an in-person workshop in **early August 2023**

contact@swas.center

Leadership Team



Rafael
Ferreira
da Silva
(ORNL)



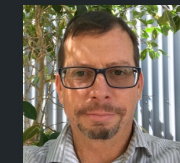
Kyle Chard
(ANL)



Lavanya
Ramakrishnan
(LBL)



Shantenu
Jha
(BNL)



Dan Laney
(LLNL)



sustaining workflows
& application services

<https://swas.center>

Thank you!
Questions?

Toward a Post-ECP Software Sustainability Organization (PESO) Hub and Spoke Model, Role of Sustainability, Funding, Q&A

- Michael Heroux (Sandia National Laboratories; PI)
- James Ahrens (Los Alamos National Laboratory)
- Todd Gamblin (Lawrence Livermore National Laboratory)
- Timothy Germann (Los Alamos National Laboratory)
- Xiaoye Sherry Li (Lawrence Berkeley National Laboratory)
- Lois Curfman McInnes (Argonne National Laboratory)
- Kathryn Mohror (Lawrence Livermore National Laboratory)
- Todd Munson (Argonne National Laboratory)
- Sameer Shende (University of Oregon)
- Rajeev Thakur (Argonne National Laboratory)
- Jeffrey Vetter (Oak Ridge National Laboratory)
- James Willenbring (Sandia National Laboratories)

Argonne National Laboratory

June 8 – 9, 2023



PESO Vision

100X

ECP investments enabled a 100X improvement in capabilities for target apps/libs

- 7 years building an accelerated, cloud-ready software ecosystem
- Positioned to utilize low-cost accelerators from multiple vendors that others cannot
- Emphasized software quality: testing, documentation, design, and more
- Prioritized community engagement: Webinars, BOFs, tutorials, and more
- DOE portability layers are the only credible way to
 - build codes that are sustainable across multiple GPUs and
 - avoid vendor lock-in or
 - deal with growing divergence and hand tuning in your code base
- DOE libraries and tools can lower costs and increase performance for accelerated platforms
- Industry has not caught up outside of AI
 - DOE enables an entirely different class of applications and capabilities to use accelerated nodes
 - In addition to AI

PESO Project Description

Progress on defining PESO

Note: PESO is a work in progress!

Your contributions are essential to define and shape it!

PESO != ECP

- ECP

- Hierarchy
- Centralized finance org
- Fixed set of apps, scope
- Heavyweight reporting
- Justified by size and design

- PESO

- Peer collaboration – Hub
- Leverage institutional finance orgs
- Dynamic, adaptive scope targets
- Tunable reporting strategy
- Lighter weight approach

PESO > Spack+E4S

- Spack+E4S
 - Major PESO deliverable
 - Important product delivery conduits
 - Platforms for agency, industry collaboration
 - Keys for testing on new and diverse platforms & software environments
- PESO = Spack+E4S+More:
 - Impact on science via 100X efforts
 - Collaborative planning, executing tracking and reporting
 - HPC community engagement: apps, facilities, vendors, agencies
 - Cross-cutting training, community engagement, sustainability efforts
 - More

PESO Project Brief Description

PESO will

- **Serve as a hub** for software-ecosystem sustainment efforts for DOE's open-source libraries and tools for advanced scientific computing
- **Work with software project teams** to coordinate development activities for long-term sustainability and benefit to stakeholders
- Work with
 - **software product communities (SPCs) and**
 - **communities of practice (COPs)**
- To provide cross-cutting services and support that are broadly needed by developers, users, and stakeholders
- Realize the full potential of DOE investments in the scientific libraries and tools ecosystem:
 - By taking a broad, strategic view
 - Through project growth, improved software quality and availability, and sustainable delivery, deployment, and support.
 - Realizing the 100X potential enabled by ECP investments

PESO Community Input Form

Ongoing collection of
community input

Please contribute



[PESO Planning Input Google Form](#)

Highlights from PESO Request for Input (Initial 40 Responses)

How to improve software sustainability

- **Testing and Quality Assurance:** There's a strong emphasis on enhancing the testing infrastructure, including unit testing, continuous integration (CI), and deployment testing. This also involves establishing testing environments similar to user environments and using diverse application sets for testing. Also mentioned is the creation of a comprehensive test suite and regular code refactoring. Formalization and automation of development practices and procedures are seen as crucial.
- **Portability and Maintenance:** Several responses highlighted the importance of making the software compatible with newer languages and architectures, as well as improving portability on future systems. This includes maintenance activities for tools such as repositories, issue tracking, and CI, as well as humans for code reviews, bug fixing, training, and documentation. Maintenance of the project on major systems for testing and support was also mentioned.
- **Documentation and Training:** This includes improving and developing user and developer documentation, online short courses for users, tutorials, and other training materials. Some suggested that a software sustainability organization could provide a venue for sharing and highlighting documentation and tutorial development.
- **Community Engagement and User Support:** Responses suggested expanding the developer community, using modern community communication tools, and organizing training and outreach events. Efforts should also be made to increase the software's visibility, encourage more registrations, and identify users' potential needs. A suggestion was made to foster collaborations through a software sustainability organization to smoothen the user experience.
- **Integration and Interoperability:** This involves improving modularity of various features, integrating the software into larger communities, and providing interfaces to other packages. Better architectural documentation and support for application developers were also considered important.
- **Funding and Support:** Dedicated funding for software maintenance, quality assurance, documentation development, porting, and testing was recommended. If sustainability organizations provide personnel, the preference is for the personnel to work with the project team.
- **Forward Planning and Innovation:** Respondents mentioned feasibility studies for emerging technology and software engineering practices, as well as the need to innovate to address potential roadblocks such as language sustainability. Early access to upcoming technologies was also viewed as important.
- **Standards and Automation:** Some respondents advocated for driving standards and automating more aspects of the developer workflow, like code formatting and updates.
- **Development of Advanced Features:** The need for more efficient implementations for advanced hardware support, more fine-grained power monitoring, and support for modern workflows consisting of AI, Big Data, and Data Science workloads were also pointed out.

PESO, Software Product Communities, and Communities of Practice can positively impact most of these

What a software sustainability organization should do to be a successful steward of DOE software

- **Funding and Stability:** Ensure a steady level of funding to support and sustain software products over the long term. It will help in retaining the talented team members and attracting new ones to participate and contribute.
- **Visibility and Engagement:** Increase visibility of software products and carefully choose engagement levels to avoid overstressing resources. Facilitate interactions between different groups developing software and help developers connect with users to understand their needs. Reach out to application developers and users actively.
- **Sustainability and Long-term Planning:** Acknowledge the fear of products not being sustained in the HPC community. Address this by explicitly sustaining products and assuring users that products will be maintained in the long run.
- **User Support and Usability:** Strive for a balance between user support and technological advancement. Ensure software is robust, has repeatable build and install instructions, and provides a clear description of usage expectations. Prioritize excellent user support in developing research libraries or tools.
- **Documentation and Best Practices:** Enforce good documentation practices, CI/CD practices, and circulate best practices across software packages. Advocate for best practices and provide conventions or standards for a uniform experience for application developers.
- **Innovation and Avoiding Monoculture:** Encourage innovation and avoid a situation where only a single dominant package/product is supported in any given category. Promote open-source software and ensure flexibility.
- **Adaptation to Technological Changes:** Adapt the software to changes in operating systems and GPU vendor software stacks. Improve software reliability, scalability, and performance based on user needs.
- **Outreach and Collaboration:** Foster community development, facilitate outreach to grow user base beyond traditional DOE HPC users, and establish mechanisms for meaningful collaborations.
- **Portability and Composability:** Ensure portability on future systems (including emerging architectures and programming models/languages), ensure composability with third-party tools.
- **Integrated Approach:** The organization should encompass various libraries and tools in an integrated manner so that end users and system administrators can deploy the stack easier.
- **Future Orientation:** Have a clear vision of the scientific communities' needs in the next decade and invest in highly performant sustainable software ecosystems.
- **Education and Awareness:** Raise awareness of the need for dedicated support and guidelines for best practices. The organization could also drive sustainability and adoption at the university education level.
- **Quality Assurance and Simplified Installation:** Provide as much quality assurance as possible and maintain simplified installation methods.

Takeaways from Input Responses

- Strong community awareness of importance of sustainability
- Strong common requirements and strategies
- Conjecture: An organization like PESO can be defined to provide:
 - Better – What we produce will be better than with PESO
 - Faster – We will get our work done more quickly
 - Cheaper – We can focus our efforts on other important things

PESO Community Workshop



Argonne National Laboratory

June 8 – 9, 2023

- 80 attendees
- Broad participation
- Workshop report coming



<https://bit.ly/peso-workshop-2023-resources>

Thursday, June 8, 2023

Time (CT)	Topic	Leads	Comments
7:30 am	Registration / Working Breakfast		
8:30 am	Plenary Kickoff	Mike Heroux (PESO PI)	Give overall workshop charge
9:00 am	Applications Perspectives	Andrew Siegel (ECP Applications Development Director)	Challenges and opportunities for increased impact of libraries and tools on application success
9:30 am	Industry Perspectives	Jeff Larkin (NVIDIA), Berk Geveci (Kitware)	Challenges and opportunities for increased impact of libraries and tools in collaboration with industry
10:00 am	Break		
10:30 am	Advanced Computing Facilities Perspectives	Katherine Riley (ALCF), Balint Joo (OLCF), Richard Gerber (NERSC)	Challenges and opportunities for increased impact of libraries and tools in collaboration with computing facilities
11:15 am	Set up for breakouts		Describe charge questions, take Q&A, locate breakout rooms
12:00 pm	Lunch		
12:30 pm	Lunchtime Talk	Ulrike Yang (ECP xSDK Project PI)	How software product communities can enhance the productivity of teams
1:00 pm	Breakout session		
2:45 pm	Break		
3:15 pm	Breakouts resume		
4:00 pm	Report out from breakouts		
5:00 pm	Adjourn		Dinner on your own

Agenda



Friday, June 9, 2023

Time (CT)	Topic	Leads	Comments
7:30 am	Working Breakfast		
8:30 am	Challenges and Opportunities for Computing	Rick Stevens (ANL, Assoc Lab Director, Computing, Environment and Life Sciences)	Roles of sustainable software ecosystems in addressing next-generation computing challenges
8:45 am	Software Foundations	Todd Gamblin (PESO co-PI)	How we can leverage software foundations for DOE software sustainability
9:00 am	Workforce Development	Lois Curfman McInnes (PESO co-PI), Mary Ann Leung (Sustainable Horizons Institute)	Challenges and opportunities for broadening participation in the HPC workforce
9:15 am	Breakout session		
10:15 am	Break		
10:45 am	Breakouts resume		
12:00 pm	Working Lunch		
12:30 pm	Report out from breakouts, discussion		
2:00 pm	Adjourn		

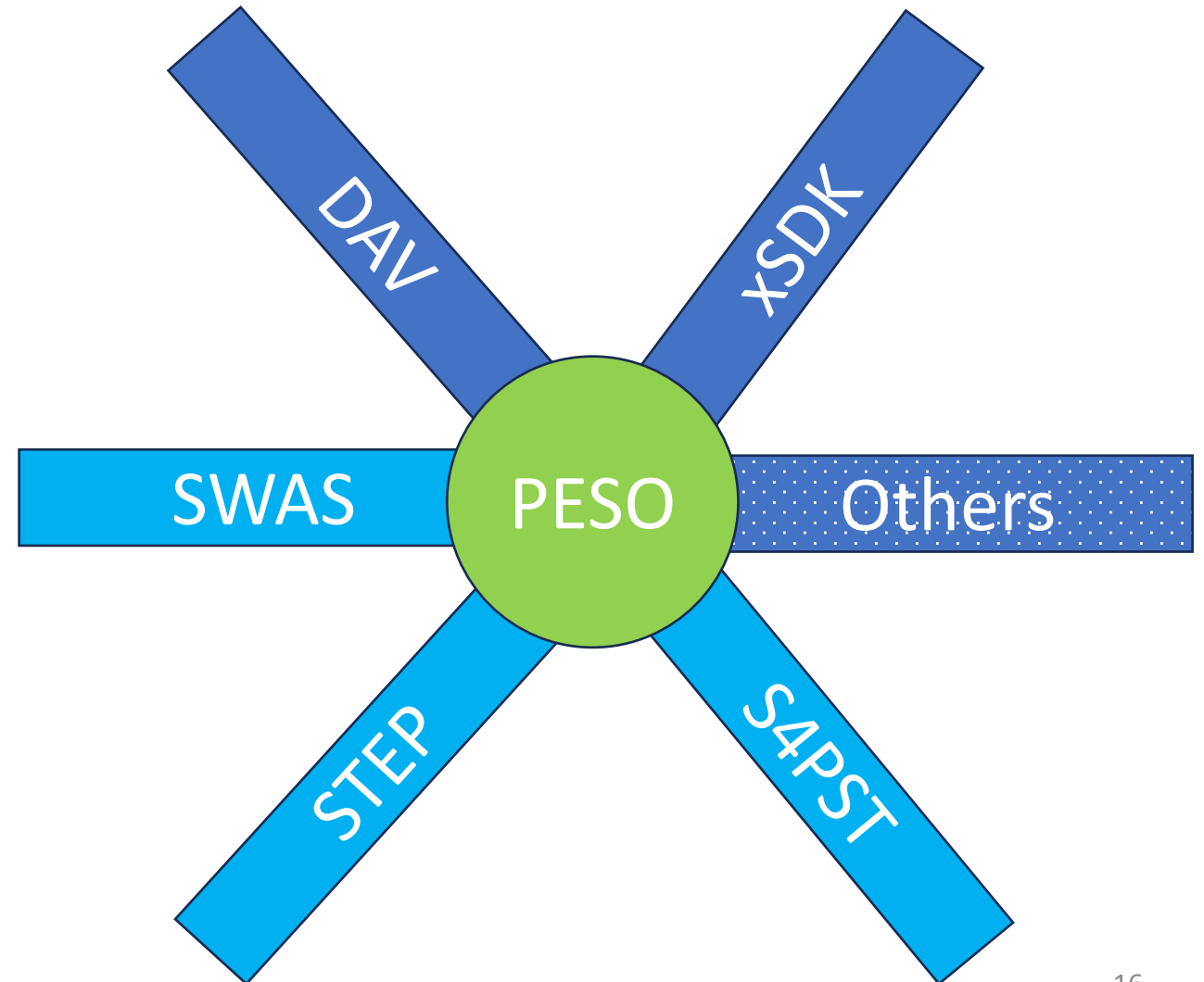
PESO Hub and Spoke Approach



- Software Product Communities (SPCs), (aka, SDKs, Spokes)
 - PESO intends to be a hub that aggregates with communities composed of like products
 - We anticipate SPCs will self-organize and have community-specific governance
 - We anticipate SPCs to include DOE-sponsored and commercial/community software
 - SPC value proposition includes – Shared design space exploration, coordination, more
- Communities of Practice (COPs)
 - PESO anticipates engaging with community leaders in important cross-cutting efforts
 - Examples include:
 - Scientific software developers: (e.g., IDEAS, HPC Best Practices webinars)
 - Community outreach (e.g., Center for Scientific Collaboration and Community Engagement (CSCCE))
 - Software foundations (e.g., NumFOCUS, Linux Foundation)
 - Workforce development (e.g., US RSE, BSSw Fellows, and Sustainable Research Pathways)
- Funding: Transparent criteria and process, with block funding

PESO Hub-and-Spoke Status

- **PESO (Hub)** – Funded seed for cross-community capabilities, engagements, services that are common to most or all software product communities
- **SWAS, STEP, S4PST** – Funded seed projects for workflows, tools, programming systems
- **DAV, xSDK** – Existing ECP SDKs that fit the PESO approach
- **Others** – Some products are missing – need to ID



Hub and Spoke Status

- PESO – SWAS, STEP, S4PST
 - Collaborating and evolving coordination within the hub-and-spoke model
- PESO – xSDK, DAV
 - Working closely under ECP already
 - Includes ECP ST and Co-Design portfolios
- PESO – SciML
 - Initial discussions to propose spoke for the scientific ML product stack
- PESO – Other products
 - What is missing that should be included?

If you are not part of ECP or the SWAS, STEP, or S4PST effort and want to engage with PESO, please contact any PESO team member e.g., Mike Heroux maherou@sandia.gov

Next steps

- Watch for draft PESO workshop report
 - We will send pointer to all workshop participants for review
- PESO Meeting 1: 3 – 4:30 pm ET, June 20
 - <https://lssw.io/PESOMeeting1>
 - ***Topic: Science Impact via Leadership Software Investment: A Holistic, Collaborative Approach from the PESO Project***

Science Impact via Leadership Software Investment: A Holistic, Collaborative Approach from the PESO Project

Description: The US Exascale Computing Project (ECP) elevated the degree of collaboration across DOE labs, universities, US industry, other US agencies and beyond, to address the challenge of producing a new generation of applications, libraries, tools, and processes that unlocked the performance and scientific potential of DOE Exascale computing systems.

The legacy of ECP must include efforts to increase the impact of ECP investments on many more scientific problems within DOE and outside. A critical element to this success is robustly building on the collaborative efforts that established the libraries and tools that were essential to ECP success. The ecosystem composed from these libraries and tools must be further supported, expanded, and leveraged to realize the scientific impact of Exascale systems and their successors.

The PESO team will give an overview of our plan for post-ECP software-ecosystem sustainment, an update on current activities, and welcome questions and comments from the community. While our efforts are informed by the PESO team ECP experiences, the PESO approach is notably different from ECP, emphasizing peer collaboration vs. hierarchy, leveraging institutional finance and process infrastructure vs. a parallel and separate project-specific approach, a dynamic application engagement strategy vs. a focus on pre-defined apps, broader engagement with industry and other agencies, and more. All of our efforts will be conducted upon a foundation of a diverse and inclusive community that provides compelling and stable career paths for all community members.

All community members are welcome to join this event. Please help us create a plan that leads to post-ECP software-ecosystem sustainment success.

Panelists: The PESO Leadership Team

S4PST Sustainability for Programming Systems and Tools

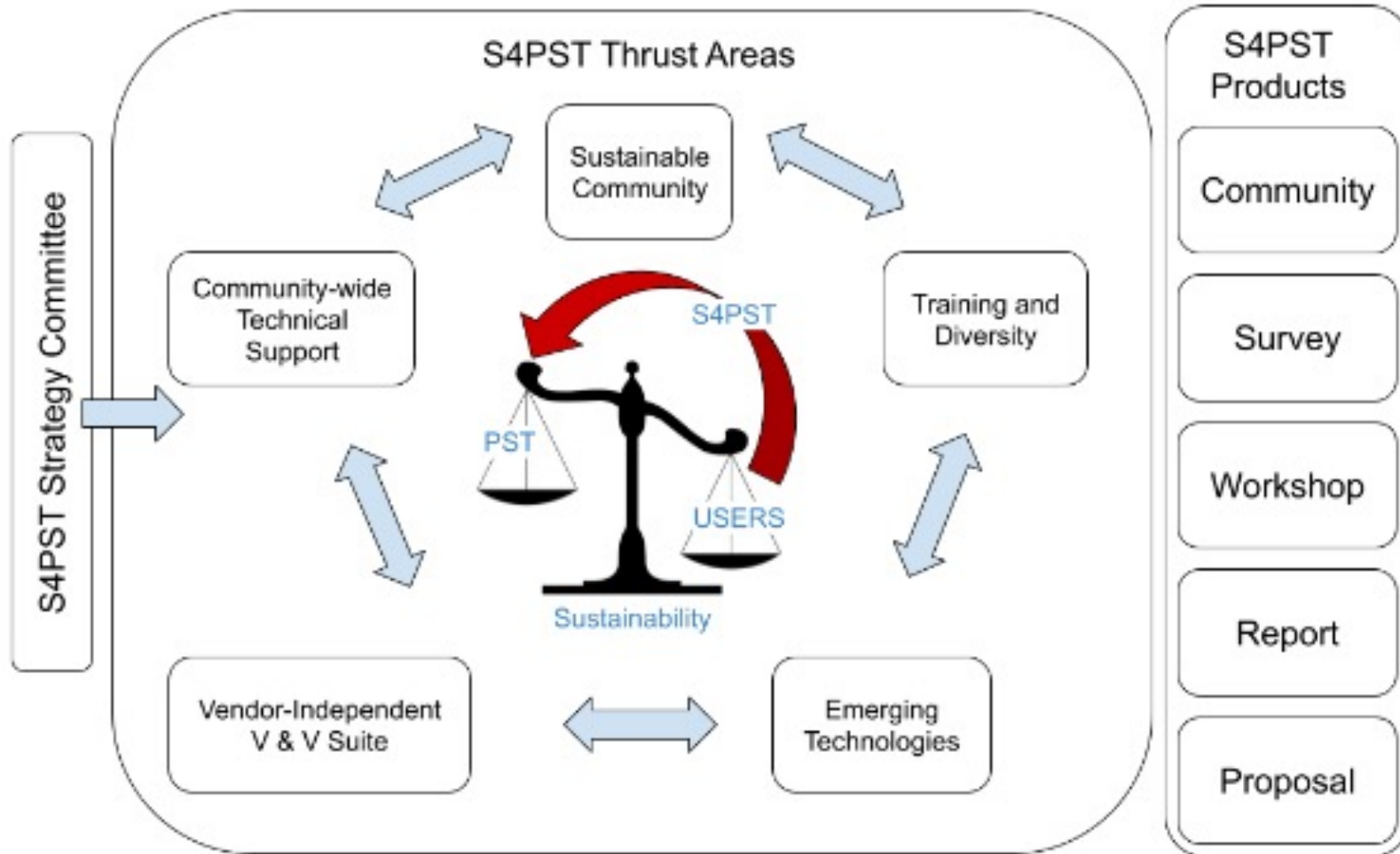
Keita Teranishi
Group Leader, Programming Systems

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

S4PST: Motivation

- ECP has created software ecosystems for scientific HPC community.
- These ecosystems can be fragile without **robust programming systems and compilers.**
 - Offloading ecosystem to third party or non-HPC.
 - Patchworked ecosystem with external dependencies
 - Emerging languages have seamless ecosystem model
 - Reactive approach is costly
- **The sustainability efforts should proactively address:**
 - **Social aspects (community, training, inclusiveness)**
 - **Economic aspects (time, total cost, amortization)**
 - **Technical (interoperability, capability, software/libraries)**

S4PST



Team

Keita Teranishi (PI, ORNL)

Pedro Valero Lara (co-PI, ORNL)

William Godoy (co-PI, ORNL)

Hartmut Kaiser (LSU)

Patrick Diehl (LSU)

Alan Edelman (MIT)

Sameer Shende (U Oregon)

Sunita Chandrasekaran (U

Delaware)

- Michel Schanen (ANL)
- Damian Rouson (LBL)
- Johannes Blaschke (LBL)
- Ignacio Laguna (LLNL)
- Todd Gamblin (LLNL)
- Siva Rajamanickam (SNL)
- Roberto Gioiosa (PNNL)
- Yunfei Guo (ANL)
- Hartwig Rizzi (UTK)
- Christian Trott (SNL)
- Damien Lebrun-Grandie (ORNL)
- Pat McCormick (LANL)
- Johannes Doerfert (LANL)
- Philippe Pebay (NGA)
- Francesco Rizzi (NGA)
- Valentine Churavy (MIT)
- Suzanne Parete-Koon (ORNL)
- Howard Pritchard (LANL)
- DK Panda (OSU)
- Hari Subramoni (OSU)

S4PST Main Activities

1. **Sustainable Community**
Guidelines, Governance, Incentives for good ecosystem practices
2. **Community-Wide Technical Support**
CI/CD, Issue Tracking, Information sharing
3. **Training and Diversity**
Training materials creation for range of skill levels
Engaging with non-DOE communities and non-DOE/ECP projects
4. **Vendor-Independent Verification and Validation**
Engaging with vendors and non-DOE entities
5. **Emerging Technologies**
Incubation and exploration of new practices and new technologies
6. **Assessment of individual products for steering the funding**
Committee (representatives from individual institutions)
Decision criteria

Current Activities

- Community Building
 - Governance model
 - Software Life Cycles (risk management and mitigations, transitions, and third-party software products)
- Survey (William Godoy, ORNL and Johannes Blaschke, LBL)
 - Users, Facilities, Developers
- White Paper (Early July)
 - Community and its governance
 - Engagement with vendors and third-parties (LLVM community, MPI-Forum, OpenMP, OpenACC, C++, Fortran standardization committee, programming-language specific conventions)
 - Training Programs
 - DEI (Education and events at MSIs)
 - Assessment of Programming Systems Projects (approx. 40) from the perspectives of users, facilities and developers
 - Survey Results
- **Online Workshop 7/31 and 8/1**

Current Activities

- Collaboration with other seedling projects
 - PESO
 - Distributed Programming Systems (MPI, UPC++)
 - Math libraries (low level kernels needing specifications for runtime interactions)
 - STEP
 - Tight integration with programming systems and tools
 - Vendor and third-party engagement
 - SWAS
 - Application services
 - V&V ideas
 - COLABS – Provide training from novice to expert (for workforce retention)
 - Introductory materials
 - Focused training (OpenMP tasking, OpenMP offload, Modern Fortran, Kokkos, code contribution policies)
 - OSSF
 - Involvement with non-DOE, legal advice, economic engagements

Software4Science

(Open Scientific Software Foundation Seedling)

Creating a sustainable future for scientific
software

<https://software4science.org>

Outreach

We have been meeting with and obtaining feedback from:

- Seedling teams
- Other labs and projects
- AD projects
- Vendors

Based on feedback we received, we have been continually refining the foundation model

What is the mission of Software4Science?

Provide an equitable and resilient center of gravity that enables and supports a thriving and sustainable ecosystem of open source scientific software projects

What Software4Science is not

- Tied to or aligned with any particular sector, organization, or group
- Trying to control how projects or programs organize or operate
- Creating a one-size-fits-all approach
- Replacing other software foundations

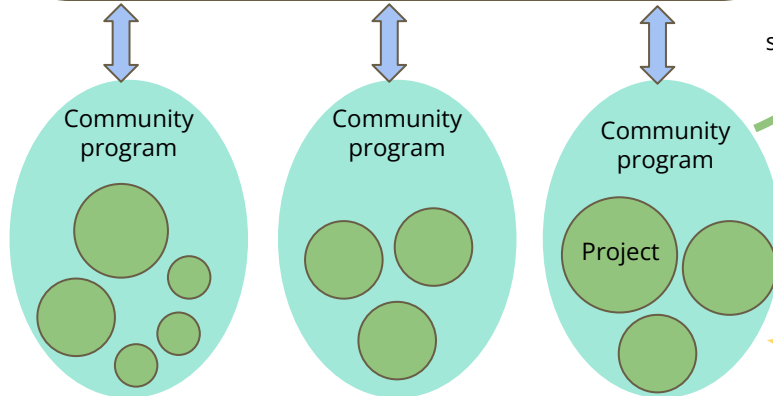
Software4Science

Sustainable Scientific Software (S³) Hub

A neutral and independent hub where projects can collaborate effectively, fostering excellence and innovation within the scientific software community

Services

1. Organize communities; user meetings
2. Software sustainability scorecards
3. Education and training services
4. Software CI and testing
5. Sustainability consulting services



All scientific software projects welcome (+ECP + DOE)

- ★ Membership fees
- ★ Grants and partnerships
- ★ Donations
- ★ Events

- ★ Grants
- ★ Donations
- ★ Fees

S³ Foundation (subsidiary)

Funding the sustainability efforts of scientific software projects

Services

1. Lifecycle management
2. Funding sustainability grants

Funds

Energy
(DOE)

...

General

Project scorecards and rankings

Project funding

Advantages of Software4Science

- Enables existing communities to continue to work in the way they know best
- Broader membership that builds sustainability of the organization
- Open to a wider variety of projects, including ECP AD and other users, which grows contributor base and opens up more opportunities to engage with vendors and sponsors
- Foundations that support open source software have proven to be a highly successful model
 - Linux, Eclipse, Apache, NumFOCUS
 - Others at <https://opensource.com/resources/organizations>

Why not an existing foundation?

- There is no reason why a project couldn't join an existing foundation
- Individual projects may not want the overhead of joining and fundraising
- Other foundations are not focussed on the scientific computing community or scientific software ecosystem
- Some foundations have more restrictive requirements for projects than others

What value would we bring?

- Representative organization that can advocate on behalf of the member projects
- Strong sense of affinity and attachment to other member projects, of “belonging”, of shared purpose
- Host meetings and conferences to bring people together to share ideas, particularly across scientific domains and user communities
- Access to common infrastructure and services
- Training and education relating to scientific software development
- Reduce the costs to DOE by creating a mechanism to bring together resources from multiple funding agencies, organizations, and individuals; for long term sustainability of the ecosystem

Thank you!

Please check out <https://software4science.org>
and join our mailing list for updates!

Contact: watsongr@ornl.gov

Greg Watson,
Oak Ridge
National
Laboratory



**Addi Thakur
Malviya,** Oak
Ridge National
Laboratory



Daniel S. Katz,
University of
Illinois Urbana
Champaign



**Dana
Robinson,** The
HDF Group



**Elaine
Raybourn,**
Sandia National
Laboratories



Bill Hoffman,
Kitware, Inc.



**John
Kellerman,** The
Eclipse
Foundation, Inc.



Clark Roundy,
The Eclipse
Foundation, Inc.

