Software Stewardship and a Next-Generation Software Stack for the Future of Scientific Computing

Hal Finkel

Acting Director, Computational Science Research and Partnerships Division, Advanced Scientific Computing Research

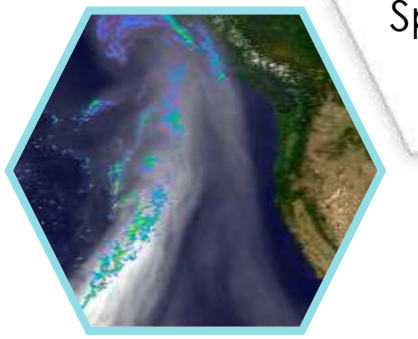
March 12, 2024

U.S. DEPARTMENT OF ENERGY | Office of Science

Frontier is open for science!

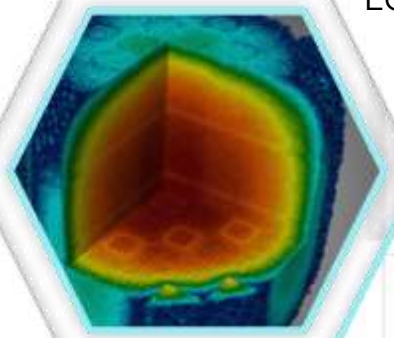
Delivering Exascale computing from system to ecosystem

ALL allocation programs have been enabled since April 2023



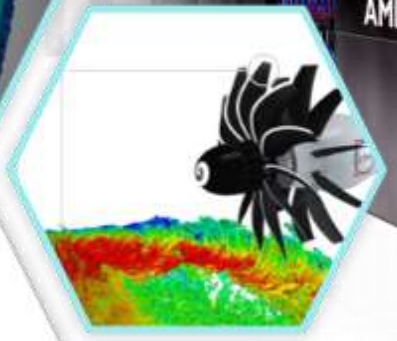
2 Gordon Bell Prize Finalists

1 Gordon Bell Climate Special Prize Finalist



To date 1,580 users & 240 projects active

GE: first-ever 3D Large Eddy Simulations at realistic flight-scale



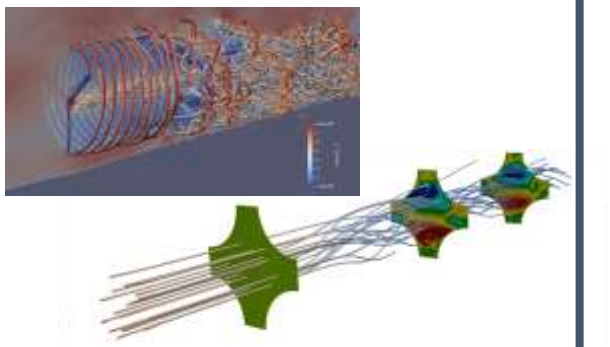
summit PLUS

Now accepting proposals

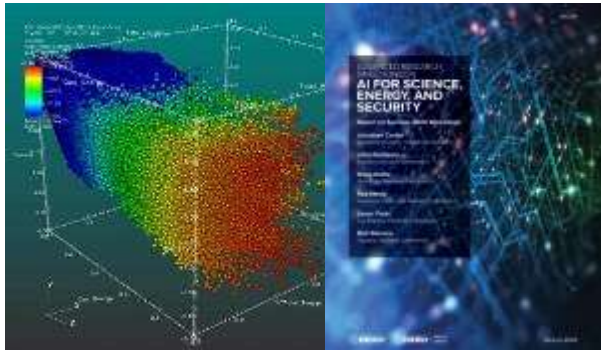
Frontier Supercomputer

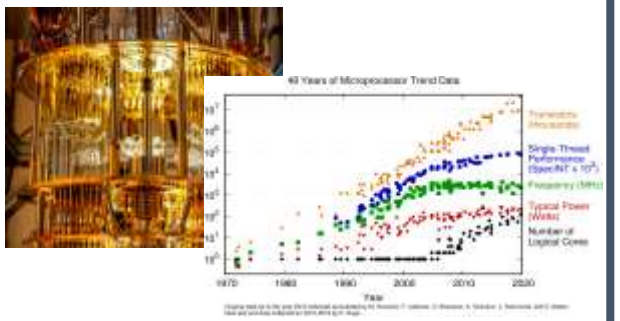
Emerging Technology Trends for Scientific Computing

Advanced Modeling, Simulation, and Visualization



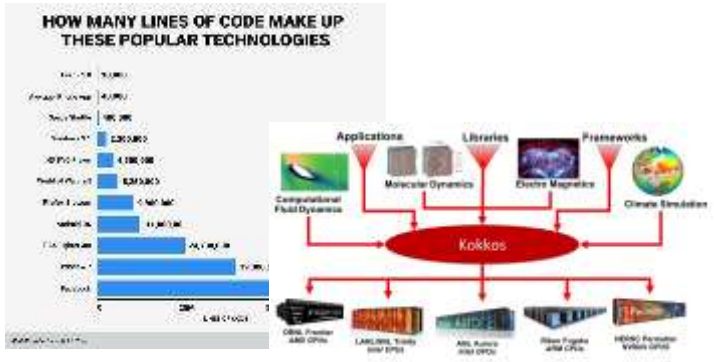
Trustworthy Artificial Intelligence and Data





Heterogeneous, Distributed, Co-Designed, Energy-Efficient Computing and Algorithms

Software Complexity for Increased Versatility



High-Performance Computing and Networking across Experiments, Exascale, and the Edge

Exascale Today Enables the AI of Tomorrow

Long-term investments in applied mathematics and computer science enabled exascale.

MPICH2

R&D 100

2005

TOP500	GREEN500	HPL-MxP
# 1	# 2	# 1

Frontier, #1 on the Top500, **leads the world in computational capability**, and is also **#2 in the world in energy efficiency**, and is **#1 in the world for AI capability**.

The exascale and AI-enabled science era will lead to dramatic capabilities to predict extreme events and their impacts on the electric grid across weather and climate time scales…

and will accelerate the design and deployment of clean-energy technologies to create a better future.

Exascale Computing Project (ECP)

DOE's Exascale Computing Initiative: A partnership between SC and NNSA/ASC to accelerate R&D, acquisition, and deployment to deliver exascale computing capability to DOE national labs by the early- to mid-2020s

6 Core DOE Labs
100 R&D Teams
1000 Researchers

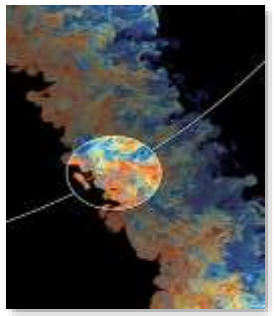
Exascale System deployment Frontier, Aurora, El Capitan

APPLICATION DEVELOPMENT

Develop and enhance the predictive capability of applications critical to DOE

National security

- Stockpile Stewardship
- Reentry-vehicles
- High-energy density physics



Energy security

- Wind farms
- Small Modular Reactors
- Nuclear materials
- Subsurface Science
- Combustion
- Clean fossil fuels
- Biofuel catalysts

Economic security

- Additive manufacturing
- Power grid
- Seismic risk

Scientific discovery

- Astrophysics
- Lattice QCD
- Accelerators
- Materials
- Chemistry
- Fusion
- Standard Model

SOFTWARE TECHNOLOGY

Expanded & vertically integrated software stack for capable exascale computing

Earth system

- Earth system models
- Biomass
- Metagenomics (DOE applications)

HARDWARE AND INTEGRATION

Integrated delivery of ECP products on targeted systems at leading DOE HPC facilities

Health care

- Cancer

On track for CD-4 in FY24

ASCR Software Stewardship

- *Key question*: How should ASCR best support continuing software-technology development and community vibrancy in the exascale science era, across next-generation HPC systems, and advanced scientific computing broadly?
  - Researchers often develop software needed for their research, but research funding is often short term and does not support maintaining and improving the software for use by other researchers, other agencies, and industry.
  - The Exascale Computing Project (ECP) has developed and implemented best practices for software sustainability for the past seven years. However, ECP will wrap up successfully during FY 2024.
- What resources and stewardship models are needed by the community? The community identified needs and responded to ASCR's Request for Information in October 2021:
  - Training on software development and use and on best practices for community development/management
  - Community outreach and networking, legal and administrative support, application engagement and support, and workforce support
  - Infrastructure for common development needs
  - Curation and governance processes and maintaining situational awareness
  - Shared engineering resources and project support (including the incorporation of new capabilities)
  - For additional information, see https://science.osti.gov/-/media/ascr/ascac/pdf/meetings/202203/ASCAC_202203_Finkel-RFI-Codesign-PDES.pdf

ASCR Software Stewardship – Seed Collaborations

- 6 collaborations were selected after peer review from proposals submitted under the DOE SC Open Call and submitted by DOE National Laboratories. Total funding, from ASCR Advanced Computing Technology and Facilities Divisions, is $720 thousand for one year starting in FY 2023.
- Collaborations engaged, and continue to engage, with the HPC and scientific-software communities to gather requirements, build coalitions, and plan for the future.
- After additional review, in FY 2024, additional funding has been provided to some of the teams, and to the SciDAC institutes, to further stewardship objectives and continue innovating on the next-generation software stack.

ASCR Software-Stewardship Request for Information (RFI)

- On October 29th, 2021, ASCR released an RFI on the stewardship of software for scientific and high-performance computing.
- Responses were due by December 13th, 2021.

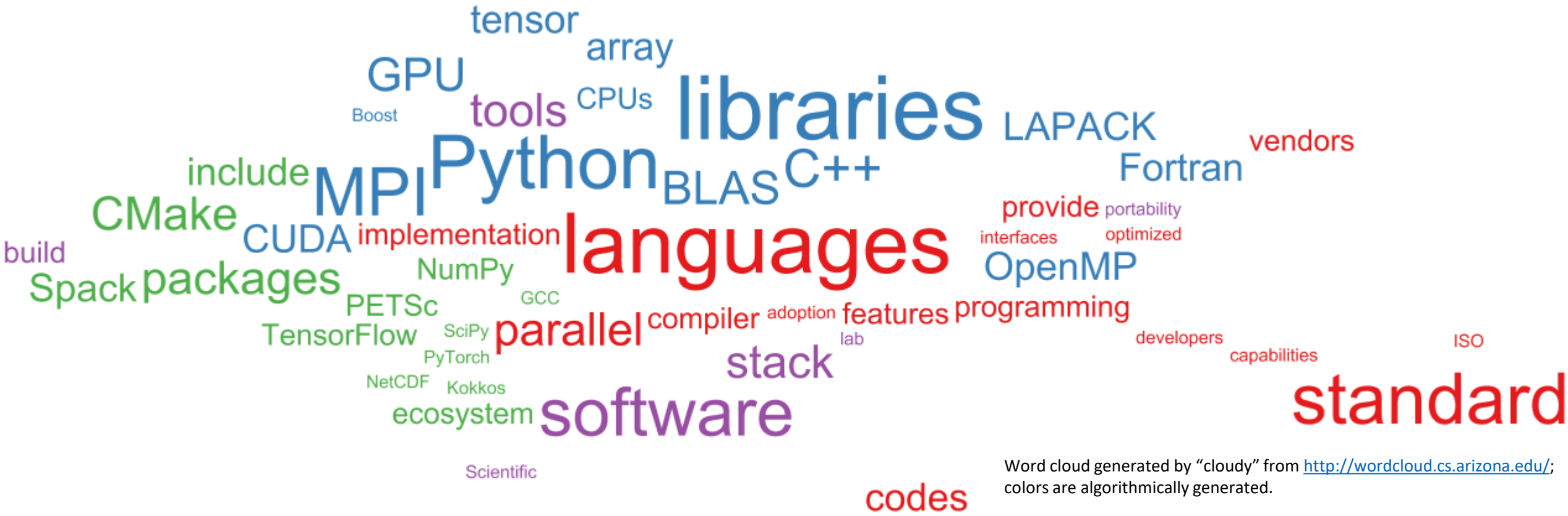
- The RFI details the potential scope of stewardship activities, including but not limited to:
  - Training on software development and use
  - Workforce support
  - Infrastructure for common development needs
  - Curation and governance processes
  - Maintaining situational awareness
  - Shared engineering resources
  - Project support

- ASCR received 37 independent responses*, quality of most was very high
  - ECP responses from the ECP ST leadership team, the ECP task force on broader engagement, NWChemEx Project.
  - 11 responses from DOE national laboratories.
  - Responses from non-profit organizations: HDF5 Group and NumFOCUS.
  - Response from the US Research Software Engineer Association
  - 6 responses from small businesses.
  - Responses from medium/large businesses: CloudBees, HPE, NVIDIA, Google.
- Responses available: https://doi.org/10.2172/1843576 – over 360 pages of text were provided.

(*) Counting the two independently-authored submissions from SNL separately.

ASCR Software-Stewardship RFI: Software Dependencies

Software dependencies and requirements: What software packages and standardized languages or Application Programming Interfaces (APIs) are current or likely future dependencies for your relevant research and development activities?

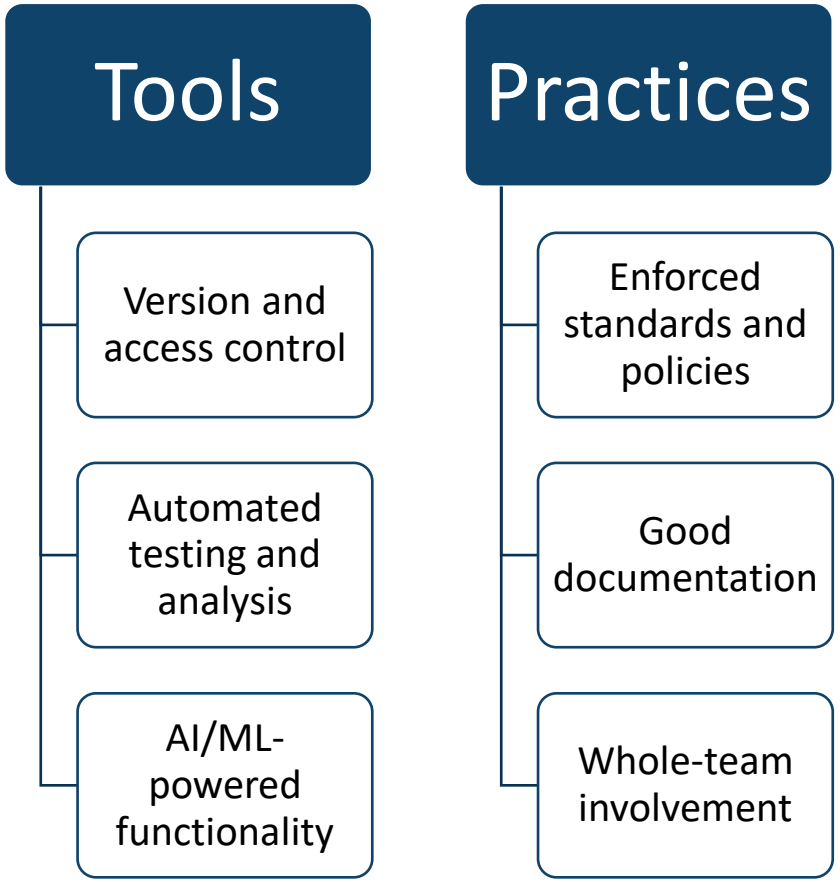


Word cloud generated by "cloudy" from http://wordcloud.cs.arizona.edu/; colors are algorithmically generated.

This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Strategies and Technologies for Software Integrity

Practices related to the security and integrity of software and data: What strategies and technology do you employ, or intend to employ in the foreseeable future, to ensure the security and integrity of your software and its associated provenance metadata?



Community best practices, including SLSA (Supply chain Levels for Software Artifacts) and SBOM (Software Bill of Materials), are actively emerging.

This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Community Software Non-Monetary Impediments

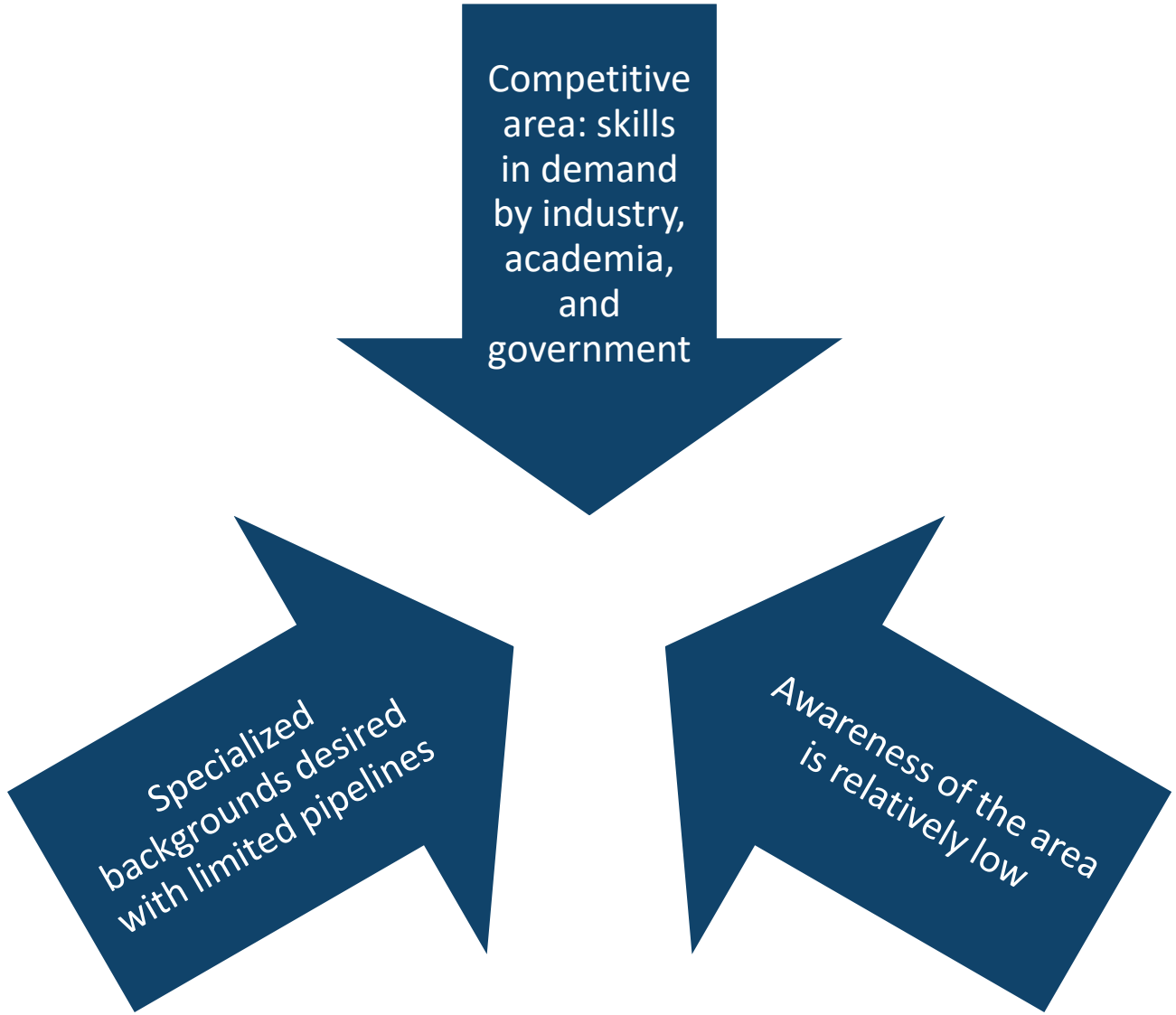
Developing and maintaining community software: What are the largest non-monetary impediments to performing this additional work?

- Maintaining additional infrastructure
- Managing additional software and process complexity
- Understanding external user requirements
- Lack of associated incentives and recognition
- Lack of time and ability to provide support commitments

This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Recruiting and Retention Challenges

Challenges in building a diverse workforce and maintaining an inclusive professional environment: What challenges do you face in recruiting and retaining talented professionals to develop software for scientific and high-performance computing?



This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Successful Strategies for Diversity and Inclusion

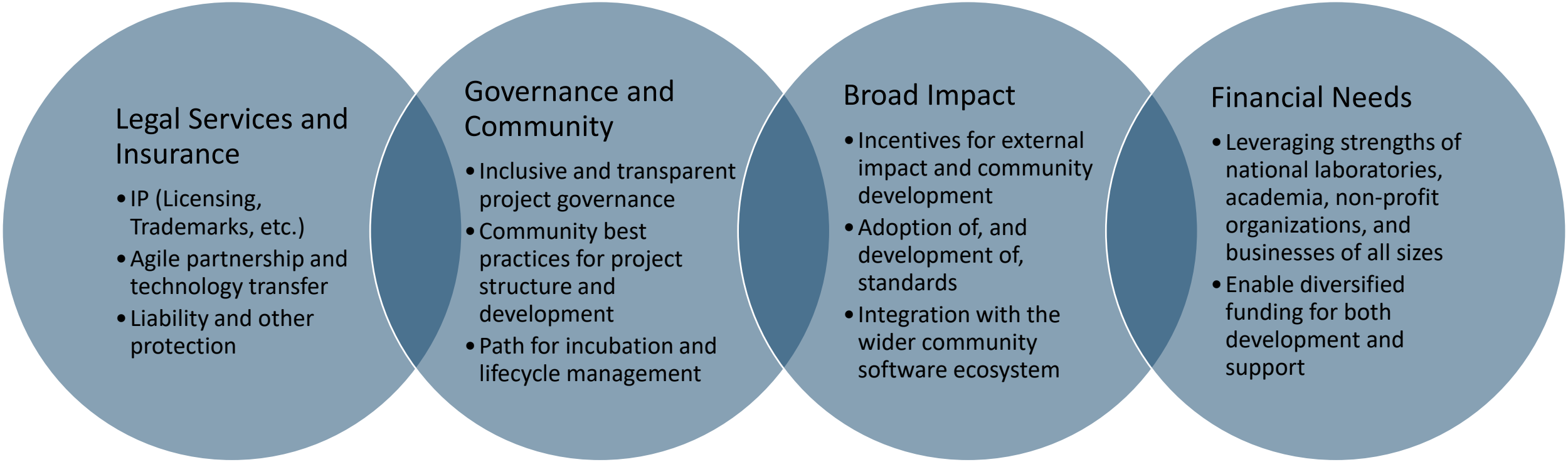
Challenges in building a diverse workforce and maintaining an inclusive professional environment:

- What successful strategies have you employed to help overcome these challenges?
- What opportunities for professional recognition and career advancement exist for those engaged in developing scientific and high-performance computing software?

This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Components of Sustainable Models

Requirements, barriers, and challenges to technology transfer, and building communities around software projects, including forming consortia and other non-profit organizations: How to encourage sustainable, resilient, and diversified funding and development models for the already-successful software within the ecosystem. What are the important characteristics and components of sustainable models for software for scientific and high-performance computing?



This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Additional Scope

Overall scope of the stewardship effort: Are there activities that should be added to, or removed from, this list?

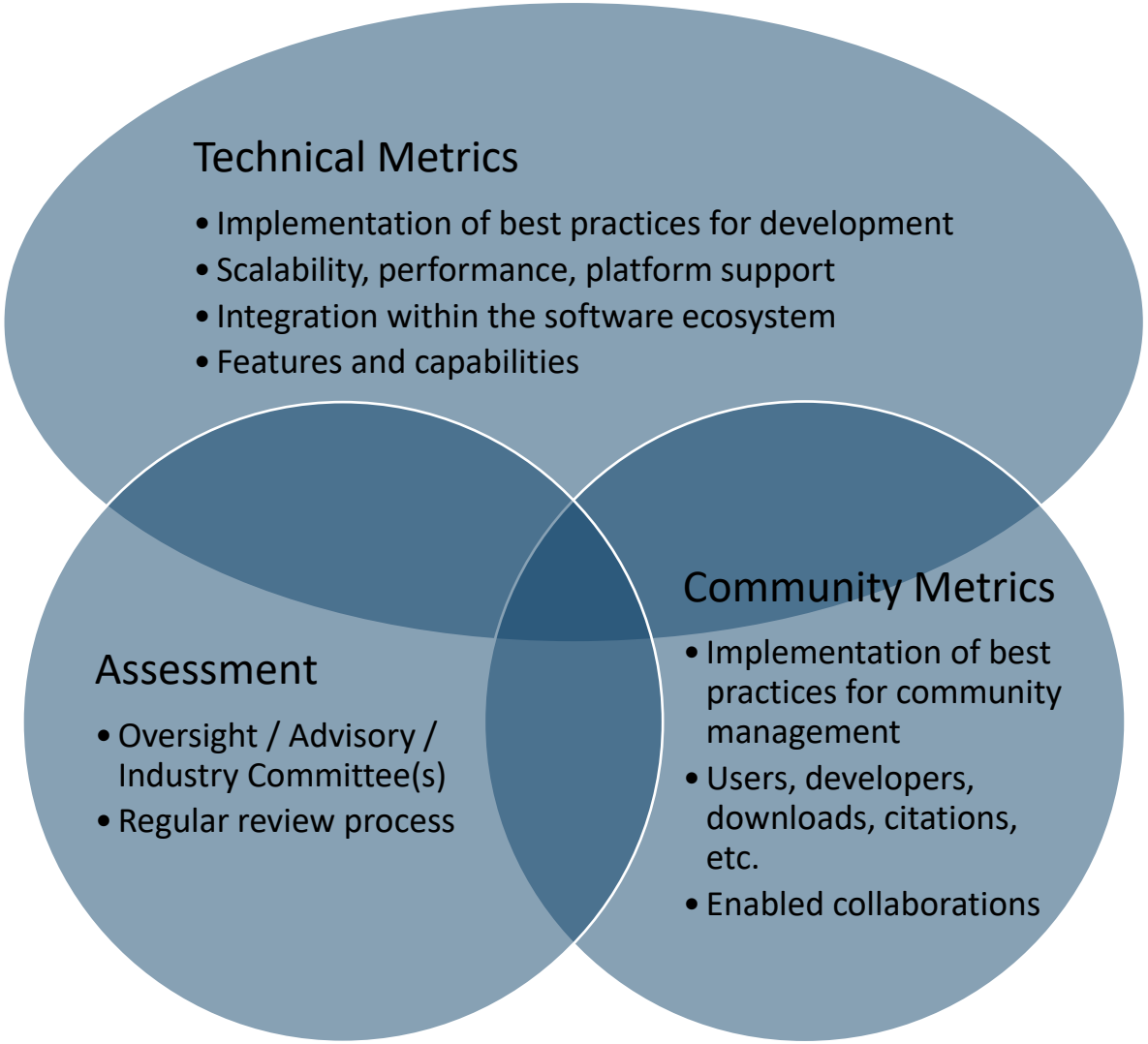
- ✓ Training on software development and use
- ✓ Workforce support
- ✓ Infrastructure for common development needs
- ✓ Curation and governance processes
- ✓ Maintaining situational awareness
- ✓ Shared engineering resources
- ✓ Project support (including the incorporation of new capabilities)

- \+ Application engagement and support
- \+ Training on best practices for community development/management
- \+ Community outreach and networking
- \+ Legal and administrative support

This slide summarizes themes noted in the RFI responses and is ***not*** comprehensive. No endorsement, recommendation, or favoring is intended or implied.

ASCR Software-Stewardship RFI: Success Criteria and Assessment

Assessment and criteria for success for the stewardship effort: What kinds of metrics or criteria would be useful in measuring the success of software stewardship efforts in scientific and high-performance computing and its impact on your scientific fields or industries?



ASCR Software Stewardship – Seed Collaborations

Principal Investigator	Title	Lead Organization	Collaborating Organizations
Dubey, Anshu	COLABS: Collaboration of ORNL, LBNL, and ANL for Better Software	Argonne National Laboratory	Lawrence Berkeley National Laboratory, Oak Ridge National Laboratory
Ferreira da Silva, Rafael	Center for Sustaining Workflows and Application Services	Oak Ridge National Laboratory	Argonne National Laboratory, Brookhaven National Laboratory, Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory
Heroux, Michael	Toward a Post-ECP Software-Sustainability Organization (PESO)	Sandia National Laboratories	Argonne National Laboratory, Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Los Alamos National Laboratory, Oak Ridge National Laboratory, Sandia National Laboratories, University of Oregon
Jones, Terry	STEP: Sustainable Tools Ecosystem Project	Oak Ridge National Laboratory	Advanced Micro Devices, Inc.; Ames National Laboratory; Argonne National Laboratory; Hewlett Packard Enterprise; Intel Corporation; International Business Machines Corporation; Lawrence Berkeley National Laboratory; Lawrence Livermore National Laboratory; Los Alamos National Laboratory; Northeastern University; Pacific Northwest National Laboratory; Rice University; Sandia National Laboratories; University of Oregon; University of Tennessee; University of Wisconsin
(Continued on the next slide)			

ASCR Software Stewardship – Seed Collaborations

Principal Investigator	Title	Lead Organization	Collaborating Organizations
(Continued from the previous slide)			
Teranishi, Keita	S4PST: Sustainability for Node Level Programming Systems and Tools	Oak Ridge National Laboratory	Argonne National Laboratory, Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Louisiana State University, Massachusetts Institute of Technology, University of Delaware, University of Oregon
Watson, Gregory	Open Scientific Software Foundation	Oak Ridge National Laboratory	Eclipse.org Foundation, Inc.; The HDF Group; Kitware, Inc.; Sandia National Laboratories; University of Illinois Urbana-Champaign

ASCR Software Stewardship – Next Steps

- Phase-II proposals on software stewardship for our next-generation software stack have been reviewed, and additional FY 2024 funding is being provided.

- The community is pursuing other partnerships which may be of wide interest, for example:

By Todd Gamblin (LLNL) and Christian Trott (SNL).