STEP: Sustainable Tools Ecosystem Project

Funding Opportunity Announcement Number: DE-FOA-0002844
DOE/SC Program Office: Advanced Scientific Computing Research (ASCR)
DOE/SC Program Office Technical Contact: Dr. Hal Finkel and Dr. William Spotz

Applicant/Institution:
Oak Ridge National Laboratory
1 Bethel Valley Road
Post Office Box 2008
Oak Ridge, Tennessee 37831

Lead Principal Investigator:
Name: Terry Jones
Telephone: 865-804-7078
Email: trj@ornl.gov
-

Administrative Point of Contact:
Name: Bailie Roy
Telephone: 865-574-5538
Email: roybc@ornl.gov

Proposal Investigator Table		
Role	**Name**	**Institution**
Lead-PI	Terry Jones	Oak Ridge National Lab
Co-Investigator	James Brandt	Sandia National Laboratories
Co-Investigator	Philip Carns	Argonne National Lab
Co-Investigator	Kshitij Doshi	Intel Corporation
Co-Investigator	Jonathan Gallmeier	Hewlett Packard Enterprise
Co-Investigator	Ann Gentile	Sandia National Laboraties
Co-Investigator	Kevin Harms	Argonne National Lab
Co-Investigator	Lucho Ionkov	Los Alamos National Lab
Co-Investigator	Heike Jagode	University of Tennessee
Co-Investigator	Mike Jantz	University of Tennessee
Co-Investigator	Matthew Legendre	Lawrence Livermore Natl Lab
Co-Investigator	Keith Lowery	Advanced Micro Devices
Co-Investigator	John Mellor-Crummey	Rice University
Co-Investigator	Barton Miller	University of Wisconsin
Co-Investigator	José Moreira	IBM
Co-Investigator	Erdal Mutlu	Pacific Northwest Natl Lab
Co-Investigator	Phil Roth	Oak Ridge National Lab
Co-Investigator	Sameer Shende	University Oregon
Co-Investigator	Shane Snyder	Argonne National Lab
Co-Investigator	Devesh Tiwari	Northeastern University
Co-Investigator	Theresa Windus	Ames National Lab

Budget and Summary

Budget Table

Application Type: ✅ Subawards ❎ Collaborative

Role	Name	Institution	Year 1 Budget	Total Budget
Lead-PI	Terry Jones	Oak Ridge National Laboratory	$125K	$125K
		Project Totals	$125K	$125K

* Budget to be used as follows: $115K to be used for three Town Hall Meetings, and $10K to be used for Diversity/Equity/Inclusion (possible Town Hall travel stipends for under-represented groups).

Leadership Structure & Facilities

The *Sustainable Tools Ecosystem Project (STEP)* proposal aims to develop a plan for sustaining a tools ecosystem for High Performance Computing (HPC) over the long term. STEP brings together experts from leadership-class facilities and stakeholders at multiple national laboratories, premier research universities and multiple vendors. Scientific discovery based on extreme-scale machines at DOE's leadership facilities has a long track-record that emphasizes the breadth and importance of the HPC performance tools community. Our project structure therefore seeks an open model which easily accommodates a wide range of stakeholders. ORNL will provide overall project management, and will be the main point of contact to DOE in reporting on progress and results. This team brings together expertise in key tool areas and will utilize an open, inclusive and community-driven process.

Institutional Contacts Table

Role	Name	Email	Institution
Lead-PI	Terry Jones	trj@ornl.gov	Oak Ridge National Laboratory
Co-PI	Philip Carns	carns@mcs.anl.gov	Argonne National Laboratory
Co-PI	Kshitij Doshi	kshitij.a.doshi@intel.com	Intel Corporation
Co-PI	Jonathan Gallmeier	jonathan.gallmeier@hpe.com	Hewlett Packard Enterprise
Co-PI	Ann Gentile	gentile@sandia.gov	Sandia
Co-PI	Mike Jantz	mrjantz@utk.edu	University of Tennessee
Co-PI	Lucho Ionkov	lionkov@lanl.gov	Los Alamos National Laboratory
Co-PI	Matthew Legendre	legendre1@llnl.gov	Lawrence Livermore National Laboratory
Co-PI	Keith Lowery	Keith.Lowery@amd.com	Advanced Micro Devices
Co-PI	John Mellor-Crummey	johnmc@rice.edu	Rice University
Co-PI	Barton Miller	bart@cs.wisc.edu	University of Wisconsin
Co-PI	Jose Moreira	jmoreira@us.ibm.com	IBM
Co-PI	Erdal Mutlu	erdal.mutlu@pnnl.gov	Pacific Northwest National Laboratory
Co-PI	Sameer Shende	sameer@cs.uoregon.edu	University Oregon
Co-PI	Devesh Tiwari	d.tiwari@northeastern.edu	Northeastern University
Co-PI	Theresa Windus	twindus@iastate.edu	Ames National Lab

ABSTRACT

This proposal describes our team's Sustainable Tools Ecosystem Project (STEP). STEP will bring together a diverse community of High Performance Computing (HPC) tools developers and stakeholders to develop plans for the sustainability of the HPC tools ecosystem. We define *HPC tools* as the collection of tools and utilities that can be applied to both understanding performance bottlenecks and optimizing performance-and-resource efficiency. We define the *HPC tools ecosystem* as the broader ecosystem encompassing the collection of stakeholders, platform dependencies, and interactions that influence those tools. By its nature, HPC tool development requires extensive hardware and application interaction and understanding and must be responsive to evolving technology. The performance gains made possible by tools are able to *dramatically* increase the effectiveness of supercomputers for scientific discovery. Through the use of HPC tools, Auxel Huebl explained during the 2022 Gordon Bell prize presentation that their team was able to start with its well-honed ECP version of WarpX (6 years in the making) and increase that version's performance by **a gamechanging 3.3x** through careful architectural mapping to the Fugaku supercomputer [4]. When HPC tools enable performance optimization through insight into how the compiled code is utilizing hardware, they throw off the shackles impeding scientific discovery.

While sustainability in software is a pervasive need, HPC tools have several urgent and domain-specific challenges which further complicate sustainability:

- **Exploding hardware complexity:** The rapid pace of increasing hardware complexity and heterogeneity greatly expands tools' targets and forces HPC tool developers to respond in a reactive manner.
- **Exploding use cases:** New and emerging application paradigms, including AI/ML, edge, and embedded instrumentation are shifting the usages that tools need to support. Additionally, there are new opportunities for tools in traditional HPC areas, such as feedback-driven dynamic resource management.
- **The coordination challenge:** Tools themselves are uniquely and closely tied to design decisions across different layers of the execution stack, including: hardware, system software, middleware, and applications.
- **The management challenge:** Building a sustainable tools ecosystem will require plans for organizing, operating budgets, community standards, technology tracking, workforce development with particular attention to promoting inclusive and equitable research (PIER).

STEP will perform vital community engagement, planning, and preparation steps to address these challenges. Our concept advocates for an open and broad collaborative approach to *identifying and implementing the best way to build a sustainable HPC tools ecosystem.* As part of this proposal, we will therefore establish a *series of open, iterative, and inclusive Town Hall meetings* where we will bring together a cross-cutting team of experts to identify requirements and strategies to support a sustainable and thriving HPC tools ecosystem. In addition, STEP will formulate an *initial plan for a sustainable tools center* which is meant to provide a larger and long-term engagement. This new community will welcome all who wish to contribute to the HPC tools space.

Contents

Executive Summary
This proposal aims to address HPC tool sustainability. As ECP sunsets, essential tools that the community relies on to achieve performance goals are at risk. Further, increasing complexity in computing architectures and trends toward new application usages and programming paradigms exacerbate the need for reliable and effective tools. This proposal will bring together leaders and stakeholders in the tools community to develop plans & processes for creating a sustainable tools ecosystem. If awarded, it will use the $125K grant to support three town hall meetings and DEI stipends.

1 Motivating Overview

The objective of the Sustainable Tools Ecosystem Project (STEP) is to establish processes that enable a sustainable tools ecosystem for the High Performance Computing (HPC) community. We define *HPC tools* to mean the collection of software that can be used to both understand performance bottlenecks and optimize performance and resource efficiency. In addition to their role in enabling supercomputer performance (a decisive determinant of scientific discovery), these tools provide essential feedback to users, operations staff, and system and application software developers. We define the *HPC tools ecosystem* as the broader encompassing collection of stakeholders, platform dependencies, and interactions that influence those tools. Given their various purposes, tools must satisfy varying levels of intrusion, interactivity, performance impact, fidelity, and other characteristics. Tools functionalities naturally have dependencies on evolving hardware, software, and policies outside of their direct control.

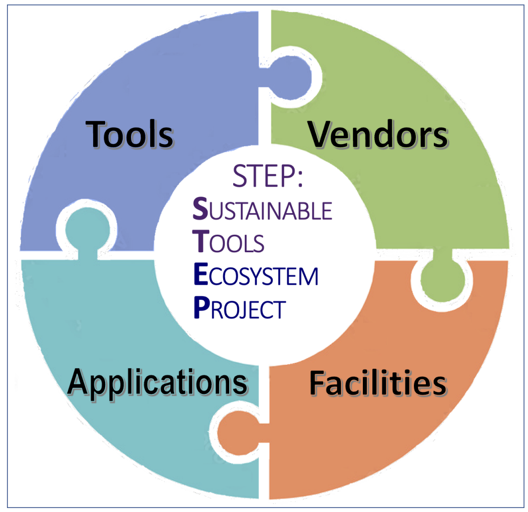

Figure 1: STEP targets 4 HPC communities.

HPC tools enable performance gains that *dramatically* increase the effectiveness of supercomputers for scientific discovery. For example, Auxel Huebl explained during the 2022 Gordon Bell prize presentation that their team was able to start with its well-honed ECP version of WarpX (6 years in the making) and increase its performance by **a gamechanging 3.3x** through careful architectural mapping to the Fugaku supercomputer [4]. When HPC tools enable performance optimization through insight into the interaction between software and hardware, they break down barriers impeding scientific discovery.

Revolutions in computing and applications thus increase the need and space for tools. However, ecosystem dependencies and current development practices threaten the ability of the tool community to meet these increased demands. *Sustainable processes* for the design, development, and maintenance of HPC tools must be *developed, agreed upon, and adopted by the community* to meet current and future needs.

STEP will bring together multiple sub-communities from the HPC domain including **1)** tool developers; **2)** application team members, library team members and other tool users; **3)** facilities staff and those who deploy HPC systems; and **4)** vendors of HPC products (see Figure 1). STEP establishes the means for these sub-communities to develop plans for the sustainability of the HPC tools ecosystem in a bottom-up, community-driven process.

1.1 Background

As computers have increased in complexity and scale, using them effectively has become much more difficult. Today's most powerful systems feature heterogeneous technologies for both computation and storage. Invariably, applications don't take full advantage of machine resources without considerable tuning. Determining how best to tailor an application/workflow for a target machine is an indispensable task that requires specialized tools and expertise. However, the path forward for useful tools is precarious: HPC vendors have never been able to provide the complement of tools needed by application developers and performance engineers, and third-party tools R&D teams lack resources for sustaining essential tools beyond the end of their current funding stream (for many today, the Exascale Computing Project – ECP [3]).

Moreover, tools are closely bound to architectures and system software in ways that other types of software, such as libraries and scientific applications, are not. For example, a tool that tracks how an application uses computing resources must be able to measure low-level architectural events and metrics and relate them to program progress and source code. As a result, tools must closely track changes to applications, architectures, and their software stacks. To complicate matters, the need for tools is most acute for understanding code performance on systems that push the boundaries of technology and scale, but these systems' novelty makes them extremely difficult for tool developers to support when first deployed. The advent of exascale systems, increases in architectural diversity, and additional complexity driven by heterogenity, all make providing effective tools for use by scientists and engineers essential to avoid impeding scientific discovery.

Thus, the tools community provides critical infrastructure for a significant portion of HPC applications and platforms, and is essential for applications to achieve high performance at scale. However, an evolving computing landscape, along with the imminent sunset of the ECP, are threatening to make many of the tools the community relies on of declining value. Hence, our project seeks to establish a community-driven framework for ensuring the long-term sustainability of HPC tools.

1.2 Developing a Plan of Action for Sustaining the Tools Ecosystem

Although HPC tools have proven invaluable in scientific computing, they face a number of shared challenges, both technical and organizational, that put their efficacy at risk if we do not establish a coordinated path to sustainability for the HPC tool community as a whole. We therefore plan to establish a strong foundation for coordinating multiple facets of software sustainability across the HPC tools ecosystem. To achieve this goal, we will need to contend with several trends and obstacles that threaten to limit the sustainability of the tools ecosystem. In particular, this proposal will establish a framework that will enable the community to address *four main long-term challenges:* (1) **exploding hardware complexity –** the rapid pace of increasing hardware complexity and heterogeneity greatly expands tools' targets and forces HPC tool developers to respond in a reactive manner, (2) **exploding use cases –** new and emerging application paradigms, including AI/ML, edge, and embedded instrumentation are shifting the usages that tools need to support; additionally, there are new opportunities for tools in traditional HPC areas, such as feedback-driven dynamic resource management, (3) **the coordination challenge –** tools themselves are uniquely and closely tied to design decisions across different layers of the execution stack, including: hardware, system software, middleware, and applications, and (4) **the management challenge –** building a sustainable tools ecosystem will require plans for organizing, operating budgets, community standards, technology tracking, workforce development with particular attention to promoting inclusive and equitable research (PIER).

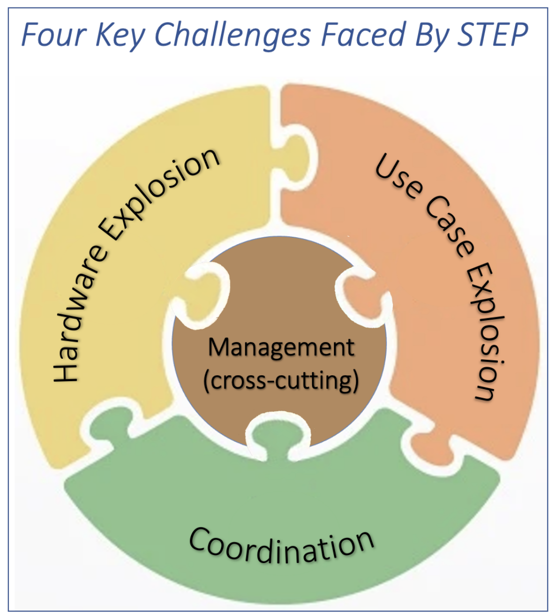

Figure 2: Obstacles to sustainable tools for exascale performance.

To lay the groundwork for a sustainable tools ecosystem, STEP will establish and hold a series of *Town Hall* meetings with the express purpose of defining effective approaches that address these challenges. The meetings will bring together experts in tool design and development as well as key stakeholders including application developers, vendors, and facility operators. All who wish to contribute to the HPC tools community will be welcome. By building and fostering collaborations among stakeholders from a range of backgrounds, we can resolve challenges in coordinating dependencies across the tools ecosystem. Further, a larger community creates an environment of innovation where new ideas to sustain and refine HPC tooling can be developed. Moving from a small set of experts in this area to an open community-driven ecosystem will greatly improve innovation [5].

To build the the coordinated community-driven long-term development directions needed to ensure the quality, breadth, and duration of tools required to have a substantial impact on HPC, we will additionally build on the Town Hall outcomes, to create high-level, initial plans for a future *STEP center*. This planning work will be executed by our distinguished team of experts in the tools field, managers from HPC vendors, liaisons for DOE facilities, and leads from scientific applications (see *Project Members table below*)

2 The Current Approach to Sustaining HPC Tools

Today's HPC tools have been developed mostly by applications developers, researchers, and users in an ad-hoc manner as they struggle to understand application performance. System administrators have also contributed tools in response to the need to identify and mitigate system problems or validate that the system is operating as designed. These HPC tool development efforts take place across all HPC centers. Despite this fact, multi-institutional collaborations are rare unless institutions share compute resources or cross-pollinate activity as a side effect of staff relocation.

A large-scale, long-running project can take the time to define functionalities and interfaces and develop a maintainable set of software tools. It can also fund maintainers to ensure all components remain compliant and interoperable. Current ad-hoc development efforts, on the other hand, have no such luxury or even incentives for compatibility with externally developed components. Thus, every HPC center will either utilize whatever tools they can find that are both open and do the job they want, or they will write their own which they may themselves put out on GitHub for public use. Since few, if any, of these software components are designed to evolve with technology (i.e., they have been written to address technologies currently deployed at some HPC center), each new HPC platform requires substantial tool re-design. The exceptional tools that have succeeded in long-term production use to date have done so by piecing together a variety of funding sources and resources over time, a situation that is not conducive to the confident development of robust, proactive capabilities.

The Exascale Computing Project (ECP) was a notable exception to the typical funding and sustainability model for software tools. However, as funding tails off at the end of the program, so does much of the collaboration. It is also important to note that, while the ECP program was crucial, it was not a panacea for HPC tool sustainability. ECP's focused approach was successful in delivering production-quality solutions for exascale platforms, but in doing so, it de-emphasized similar capabilities for institution- or department-scale systems. Similarly, the ECP program's hierarchical management approach was instrumental in driving progress, but also made it challenging to pursue new, community-driven agile approaches to emerging problems. The experience gained from this program highlights a need for sustainability plans that encompass the full range of DOE computing capabilities and can rapidly adapt to feedback from stakeholders with diverse perspectives. A case study comparing and contrasting ECP style approach and a STEP style approach is provided in Section 4.1.

The community's experience with the current state of the practice in HPC tool sustainability has also brought into focus four urgent challenges that are now facing the HPC tools community at large:

- **Exploding Hardware Complexity** Computer speeds have been increasing at around 1000x every 12 to 14 years in recent decades. With this dramatic increase in speed there has been a corresponding rapid change in computer architectures, expanding the targets for tools and causing universities and national laboratories to make tool improvements in a ***reactive*** manner as a result. Reactive development is undesirable because tools lag the availability of hardware, because retrofitting software to new interfaces and or uses can be labor intensive, and because there may be missed opportunities from insufficient communication of tool experts and machine architects. HPC vendors themselves are not incentivized to support a broad range of tools when faced with the realities of technological and market pressure.

- **Exploding Use Cases** The application landscape is expanding in multiple dimensions, and performance tools must evolve to support them effectively. This calls for the development of new and diverse data gathering and analysis mechanisms. Emerging AI/ML applications, integration of AI/ML within traditional HPC programming models, increasingly sophisticated workflow patterns, dynamic resource provisioning and management, and the expansion of HPC into more problem domains all present opportunities for increasing scientific computing productivity if the community is equipped with HPC tool capabilities that accomodate those use cases.

- **Coordination** The current siloed nature of HPC tool development precludes large community coordination, which in turn limits the breadth of domain expertise that can be brought to bear. The ultimate impact of this coordination challenge is that there is limited interoperability between key tools and piecemeal response to technology trends. At present there is no venue facilitating communication between the HPC tool community at large and key stakeholders such as vendors or facility operators. There is also no ongoing forum for promoting interoperability.

- **Management** Management is a broader challenge that touches upon each of the three previously outlined challenges. The lack of a focused, community-driven management strategy presents an obstacle for objective and inclusive organizing and feedback, long-term budgeting, tactical response to imminent technology evolution, and insuring diversity, equity, and inclusion. The current state of the practice is for each performance tool team to develop its own management best practices to the best of its ability within the scope of its own team. A "big picture" perspective is needed in order to more effectively balance these considerations across the DOE complex.

3 Community Building and Execution Plan

In contrast to current practice (see Section 2), we propose a *community-organized, community-driven* coordination of HPC tools design and development activities in the context of the wider HPC ecosystem. In a series of Town Halls, we will bring together the larger community of HPC tools developers and stakeholders, shown in Figure 3, to explore and address the four identified challenges. Through participation in these Town Halls, this community will develop guiding principles along with associated methodologies and approaches that, if followed, can ensure that the HPC tool needs, both current and future, are met through sustainable development, deployment, and maintenance processes. The Town Hall interactions will encourage continued community interaction and application of new guidelines going forward. Additionally, we will address how these guiding principles can be rapidly implemented to form a STEP center to more sustainably and coherently coordinate activities in the long-term.

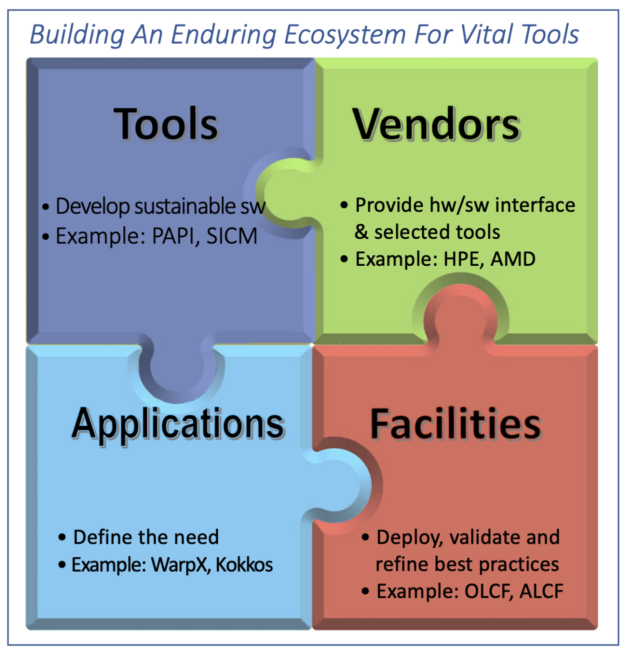

Figure 3: Tools and tools-dependent communities can collaborate to address sustainability

Section 3.1 describes STEP's Town Hall activities. Section 3.2 describes how the Town Hall outcomes can inform a blueprint for a STEP Center. Section 3.3 describes how the STEP team composition positions it for success. Section 3.4 presents information on organization of STEP and at a high level how that maps onto Town Hall activities. Section 3.5 provides information on initial STEP project members and tasks to be performed in Calendar Year (CY) 2023. Finally, Section 3.6 provides information on how the STEP project will measure success in CY2023.

3.1 STEP Town Hall meetings

STEP will establish a series of three Town Hall meetings, held in the Spring-Summer of 2023. Each Town Hall meeting is expected to have between 60 and 120 experts representing the HPC tools developers, vendors, HPC facilities and application teams communities. These disparate communities have not typically or regularly interacted as a group, but have significant inter-dependencies (Figure 3). Participants will collaboratively explore the current HPC tools space and develop solutions to the sustainability challenges identified above.

Town halls will consist of a variety of sessions that seek to clarify the important challenges and urgent gaps in the HPC tools ecosystem and generate concrete actions to move this area forward in the near and long-term future. The sessions will be will be organized into survey and sustainability sessions:

- **Tool survey** sessions aim to establish current and future directions of HPC tools and identify opportunities for increasing sustainability. These sessions will include discussion of current tools and their characteristics such as their purposes, usages, requirements, development processes, community support, and existing challenges. These sessions will also include identification of potential new tools and directions for emerging use cases, architectures and code reuse. They will also explore how new interactivity among existing tools might provide more holistic insights and expand tool impact. Broader understanding of the HPC tools space will seed brainstorming of opportunities in community issues such as standardization, common API's, code reuse.
- **Tool sustainability** sessions will explore critical challenges and opportunities in building a sustainable tools ecosystem. **Table 1 (left)** lists each of the challenges to be addressed along with examples of how they currently, and will in the future, affect sustainability. These examples will facilitate discussion of approaches to overcome the challenges in these and similar cases. Potential outcome objectives for each challenge area are listed in **Table 1 (right)**. These will seed brainstorming of how such objectives could be realized.

Note that one of the aspects of the management challenge is workforce development. Wider industry opportunities (e.g., social media and search companies) have resulted in competition for students who would have previously explored HPC careers. Lack of established education and training in production-quality code development results in variability in code quality and style making sustainability more difficult. The STEP team includes a number of university partners who can influence educational directions. As part of our DEI activities (See Appendix 4.3), we will include early career HPC tools developers to lead exploration of this aspect of the management challenge, calling upon their own experiences and interests to address the workforce development objective.

Session leads will be assigned in order to facilitate productive conversations that lead to actionable results. Highlights of the discussion topics and proposed actionable results will be recorded. The STEP team will provide a rollup of the outcomes from the *tool survey* and *tool sustainability* sessions. Of immediate use will be actionable tasks targeting near-term implementation. Longer term directions for broader and sustained impact will also be included. Each Town Hall will focus on a subset of the HPC tools for the survey session and a subset of the sustainability challenges, with all topics appearing in at least two of the three town halls to ensure coverage and diversity of view points.

Many potential collaboration efforts do not gain traction because of the difficulty in engaging a core constituency. This proposal, in contrast, is authored by a substantial number of the HPC tool and national labs communities who support the need to tackle these challenges limiting sustainability. Further, many collaborative efforts lose steam when internal priority directions outweigh collaborative gain. Our approach increases the collaborative gain by bringing together communities to address challenges caused by their dependencies. As part of the Objectives addressing the coordination challenge, we include planning for enabling continuing communications among these communities even after the Town Halls.

3.2 STEP Center for Community-Driven Progress

While the Town Hall report will have actionable paths forward, larger and longer engagement is needed to take those ideas and build them into coordinated community-driven development directions. We wish to

Town Hall Topics	STEP Objectives
The Exploding HW Challenge • Exascale architectures will feature more diverse architectural resources and tighter integration among computing components. (e.g., AMD APUs in El Capitan with GPUs, CPUs, caches, and HBM in the same packaging). This explosion of architectures and features increases tool targets. • Hardware support for measuring GPU performance is evolving. GPU vendors provide many event counters that enable tools to gain insight into how code interacts with hardware structures, but fine-grain measurements are needed to associate problems with code (e.g., no existing GPU provides mechanisms correlate particular acceesses with memory hierarchy locations). • Tools development is often *reactive* to rapidly evolving computer architectures. For example, AMD and NVIDIA recently released new tools and sampling APIs, and Intel changed their GPU binary format. More *proactive* design processes, where tool developers and hardware vendors collaborate early and closely, are essential for developing the tooling support that is critical for scientific communities.	• Technology Tracking: Develop processes to rapidly adapt to new technologies. Monitor, assess, and address: 1) the importance of emerging architectures and technologies, 2) the needs projects using particular tools, 3) HW and SW capabilities, implementation, and interfaces that tools depend upon, 4) strategies for adapting tools to planned and emerging architectures; and 5) a coordinated process for deployment and production use. • Early Access to HPC Hardware: Develop plans to support early access and tool development on platforms where classified and restricted access are barriers. • Procurement Input Processes: Develop plans to build early input from the tools community into procurement and NRE contract processes (e.g., need for additional telemetry and requirements that telemetry that is being utilized stays in place.)
The Exploding Use Case Challenge • New opportunities for tools, such as feedback-driven driven dynamic resource management, put new design requirements on tools. • As new programming paradigms and workflows, such as Python-based AI/ML and Julia-enabled parallelism, become more relevant to HPC, the tools community will need to adapt their products to support new use cases. • Increased capabilities for processing closer to architectural resources (e.g., with smart NICs, in-memory compute, etc.) provide new opportunities for tools to not only provide insights into application efficiencies, but also to dynamically remap and reconfigure applications and resources. Latency and interoperability challenges must be addressed to realize this potential, however.	• New Cases and Tools: Identify a) new paradigms and their associated requirements for tooling and b) new use cases and opportunities for tools. • Sustainability Built into New Tools: Establish collaborative community and co-design, development, and management processes to ensure that future tools are developed sustainably from the ground up.
The Coordination Challenge • Tools are uniquely and closely tied to the intricacies of the low-level hardware and software. However, lack of contracts between software and hardware layers affects the composition of tools (e.g., which software and hardware resources participate in a data allocation or usage? In which cases is it safe to monitor resources or hold locks?) • More broadly, there are often undocumented dependencies between tools and low-level computing infrastructure, which can lead to situations where it is difficult or impossible for tools to adapt if and when hardware or system software changes. • Furthermore, most tools are currently developed independently of one another, resulting in a lack of interoperability or functional reuse. The community also lacks standards, which leads to more incompatibilities and uncertain developmental directions.	• Support of Community Co-Dependency: Build recognition of potential dependencies. Obtain commitments for supporting dependencies in development and documentation processes. Establish lists of required issues and changes for notification and documentation. Develop working groups and/or communication channels to ensure continued coordination. • Community Standards: Develop use-case driven plans and processes to standardize data access APIs and data and storage formats. Such standards will enable tools interoperability and facilitate sustainability. Resuable components reduce the per-developer expertise as architectural options expand.
The Management Challenge • A sustainable HPC tools ecosystem must have plans for long-term coordination of development roadmaps, for establishing and ensuring standards, building a skilled workforce, operating budgets, and for ensuring and Diversity, Equity, and Inclusion in the field. • Lack of established education and training in distributed and production-quality code development results in variability in code quality and style, making sustainability more difficult	• Software Lifecycle: Develop tools-aware SW engineering resources to help ensure best practices are being applied, including processes for SW sustainment. • Development roadmaps: Develop plan for determining needs for tools, including External Advisory Boards, and awareness of the needs of ASCR. • Operating budgets: Establish processes involving funding level planning, tracking, and reporting. • Workforce development: Develop training and education plan to make sustainability considerations an integral part of the tools development process. • DEI: develop plan for promoting inclusive and equitable participation (PIER Plan in *Appendix 7*)

Table 1: Sustainability Session Topics (with Motivating Examples) and Potential Outcome Objectives.

ensure quality and breadth over a duration that will have a substantial impact on HPC. To accomplish this the STEP team will take the report outcomes and further develop an initial plan for creating and operating a STEP Center that will ensure such long-term coordination and engagement. We envision this STEP Center as a vehicle for scaling the STEP project activities beyond year one. This plan will include details for coordinating tool development directions, for establishing standards to facilitate development and maintenance, and for developing and hosting training activities in the use of tools. Sustainable funding is at the heart of commitment to any long term endeavour and so will be a major consideration in plans for a STEP Center. The community coordination enabled by the center will help position tools for sustainable funding throughout their lifecycles as opposed to the ad-hoc practices currently limiting tool utility and sustainability.

3.3 Team and Credentials

STEP consists of an accomplished team of HPC tools professionals and representatives from vendors, facilities, and application teams (see Figure 4). The team members include a number of representatives of widely used tools at HPC sites, including application profiling, sub-system, and system monitoring tools; foundational data interfaces; performance analysis tools, and more. They have substantial expertise and multiple years of experience in all aspects of HPC tool development, including developing for high-fidelity data collection; developing for low-latency information gathering, analysis, and response; supporting multiple concurrent and evolving types of processors, accelerators, filesystems, high speed networks; developing interactivity with a variety of software stacks and resource managers; developing interoperability among tools; and developing performance and tool output visualizations. Brief biographies are available in *Appendix 1*.

3.4 Project Organization and Budget

The STEP team is led by Terry Jones, a Senior Research Staff member at ORNL. Terry has experience as an ECP tools project lead, as an organizer for Quantitative co-design workshop series at Supercomputing, as an invited participant in multiple Dagstuhl workshops in the topic area, and as overall lead for 3 ASCR projects as well as an ASCR pathforward.

The team has vast aggregate experience organizing conferences, workshops, seminars in HPC tools and related fields, and are well qualified to conceive of, host, and draw participants for a successful Town Hall series. We will delegate responsibility for sub-tasks to Town Hall site planning leads, break-out leads, session leads, report chapter leads, and leads for other activities and duties as needed. For this project, the STEP team will organize into small groups for the planning and execution of the Town Halls. All members will participate in report writing.

Figure 4 provides a list detailing the breadth of credentials to be used in our planning activities. Our concept envisions building a collaborative structure of tools experts, with HPC vendors, facilities, and application teams as depicted in Figure 5.

The STEP budget will be allocated as follows: $115K set aside for three Town Hall Meetings, and $10K set aside for Diversity/Equity/Inclusion (possible Town Hall travel stipends for under-represented groups).

3.5 Schedule and Tasks

The following are the tasks to be carried out during calendar year 2023 by Co-PIs/institutions. Included with each task are the associated entrance dependencies and the identification of the persons who will be responsible for the item. As described earlier (see *Section 3*), the tasks are logically grouped into ten 'tasks' or thrust areas. A complete Task Table is found in Figure 6.

We will target completion of our planning at 9 months. We believe the 9 month schedule is reasonable for our activities and it provides a degree of flexibility should we encounter unanticipated delays. Moreover, an earlier completion will be useful for responding to subsequent software sustainability FOAs in a timely manner.

Project Members

Role	Name	Title	Credentials	Institution
Lead-PI	Terry Jones	Sr Research Staff	Three-time ASCR PI	Oak Ridge National Lab
Task Lead	James Brandt	Distinguished Technical Staff	LDMS architect/designer lead	Sandia
Task Lead	Philip Carns	Computer Scientist	IO tools developer/designer	Argonne National Lab
Vendors	Kshitij Doshi	Sr Principal Engineer	Intel research manager	Intel Corporation
Vendors	Jonathan Gallmeier	Dir, Development Environment	Director, Development Environment	Hewlett Packard Enterprise
Task Lead	Ann Gentile	Mgr, HPC Development Env.	Monitoring Software Architect	Sandia
Facilities	Kevin Harms	Sr Software Developer	ALCF facilities engagement	Argonne National Lab
Task Lead	Lucho Ionkov	Computer Scientist	Tools/Pub-sub designer/architect	Los Alamos National Lab
Task Lead	Heike Jagode	Research Asst. Professor	PAPI designer/architect lead	University of Tennessee
Task Lead	Mike Jantz	Associate Professor	Memory tools designer/architect	University of Tennessee
Task Lead	Matthew Legendre	GL, Tools Development Grp	Adiak+Caliper designer/architect lead	Lawrence Livermore Natl Lab
Vendors	Keith Lowery	Sr Fellow, Research	Lead for performance analysis tools	Advanced Micro Devices
Task Lead	John Mellor-Crummey	Professor	HPCToolkit designer/architect lead	Rice University
Task Lead	Barton Miller	Vilas Distinguished Professor	Paradyn and tools designer/architect	University of Wisconsin
Vendors	José Moreira	Distinguished Research Staff	HW architect for Power processors	IBM
Applications	Erdal Mutlu	Computer Scientist	Tensor Algebra Many-body Methods	Pacific Northwest Natl Lab
Facilities	Phil Roth	GL, Algorithms & Perf Group	OLCF tools and engagement	Oak Ridge National Lab
Task Lead	Sameer Shende	Research Professor & Director	TAU and tools designer/architect	University Oregon
Task Lead	Shane Snyder	Software Engineer	Darshan designer/architect lead	Argonne National Lab
Task Lead	Devesh Tiwari	Professor	AI driven tools for data center	Northeastern University
Applications	Theresa Windus	Senior Scientist	Lead for NWChemEx	Ames National Lab

Figure 4: Project members and their credentials.

3.6 Measure of Success and Midterm/Final 'Exams' to check success

Our success will be measured against **(a) Midterm:** whether or not we are able to realize the three town hall meetings in the Summer of 2023; **(b) Final:** whether or not we are able to produce the subsequent STEP reports detailing the information, surveys, plans and processes resulting from the town halls within 90 days of the last town hall meeting.

4 Impact

4.1 Impact on HPC tools: a Darshan example

As an example of the potential impact of a sustainable tools ecosystem, consider the Darshan HPC I/O characterization tool [2]. Darshan has been sustained using a variety of methods over its lifetime; it therefore provides not just an example of the state of the practice, but also an example of how the STEP proposal could benefit such a tool.

Darshan was created in 2008 under the auspices of ASCR computer science base funding, with significant contributions from the ALCF, for the express purpose of understanding I/O behavior on the ALCF's flagship computing platform. Today, Darshan is used on a variety of HPC platforms around the world to optimize applications, improve facility utilization, and inform the research community. However, connecting these two points was not straightforward. Darshan was created and maintained using resources assembled

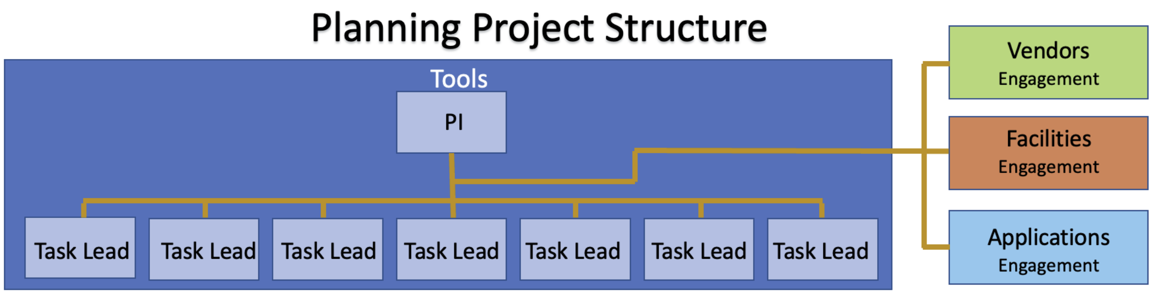

Figure 5: Organizational chart of project structure.

Schedule and Tasks

Due Date (assumes Apr-2023 start)	Description	Entrance dependency	Responsible person
2023 Q2	Task 1: Project kickoff and initiate bi-weekly videocons	none	T. Jones
	Task 2: plan and attendee list for Town Hall Meeting #1	Task 1	T. Jones
2023 Q3	Task 3: Town Hall Meeting #1 (Eastern Timezone)	Task 2	T. Jones
	Task 4: plan and attendee list for Town Hall Meeting #2	Task 2	T. Windus
	Task 5: Town Hall Meeting #2 (Central Timezone)	Task 4	T. Windus
	Task 6: plan and attendee list for Town Hall Meeting #3	Task 2	J. Brandt
2023 Q4	Task 7: Town Hall Meeting #3 (Western Timezone)	Task 6	J. Brandt
	Task 8: STEP Final report (includes Town Hall findings)	Task 3,5,7	T. Jones
	Task 9: STEP Co-design Center Blueprint	Task 3,5,7	T. Jones
	Task 10: Project sunset and documentation	Task 8,9	T. Jones

Figure 6: Tasks table with entrance dependencies and responsible PI/Co-PI.

over many years from limited-duration ASCR research grants, NSF collaborations, and the ECP program. As a consequence of this model, Darshan development was often reactive and had to be painstakingly aligned to fit within the scope of evolving research deliverables. Unfortunately, Darshan's story is far from unique. Rather, it serves as just one example of the past and future life-cycle challenges faced by many of the DOE's most critical performance tools.

The advent of ECP addressed many sustainability issues by bringing productization, platform support, and integration planning to the forefront as first-class objectives. However, as ECP comes to its conclusion, the future path for balancing research and sustainability is uncertain, and key venues for stakeholders to share guidance and responsibility are being eliminated. This STEP proposal seeks not only to fill this void but also to unlock new opportunities as well. For example, STEP's agile, community-based responsiveness would help foster new innovations; even tools like Darshan are difficult to bootstrap without agile, community-driven management support. Broadening focus beyond exascale platforms will enable wider benefits for the community. Indeed, the impact of this project will extend to additional DOE facilities, facilitate workforce development beyond just the communities with access to leadership facilities, and facilitate contributions from the broader open source world. Finally, a STEP town hall approach that caters to the HPC tool community would aid in focusing communication for key stakeholders, especially vendors and facilities that influence or are influenced by numerous HPC tools.

4.2 Impact on scientific computing at large

Improving the efficiency of the software on HPC platforms not only provides faster time-to-solution for the scientist, but also effectively decreases the cost per computation. This increased efficiency and science throughput has a direct impact on the DOE mission and the new innovations that science initiates.

Another substantial benefit of this project is that it will engender closer and more frequent communication between HPC tool developers and scientific applications, which will reduce redundant effort in both domains. Application developers will be able to focus on their application's requirements, rather than trying to develop ad-hoc tools to analyze resource usage or performance. At the same time, tool developers will receive better and more direct feedback by deploying their tools in real-world applications, rather than trying to approximate these conditions in simulation or through proxy applications.

Finally, a sustainable tools ecosystem will significantly broaden the use of tools in scientific applications. For example, standard and interoperable interfaces among tools will make it easier for developers to adopt and combine tools with different purposes. Moreover, developers will have better assurance that the capabilities their applications rely upon will not disappear if one of the tools in the ecosystem is discontinued. Thus, a sustainable tools ecosystem will also lead to more sustainable scientific software and increase the productivity of scientific software development.

4.3 Contributions and Artifacts

The principal contribution of this project is the establishment of a strong foundation for coordinating multiple facets of software sustainability across the HPC tools ecosystem.

The primary artifact of the Town Halls (Section 3.1) is a report of the outcomes, which include:

- Results from the tool surveys, cataloging characteristics of and directions for tools and determining opportunities for community interactions to increase sustainability
- Highlights of the discussions, further clarifying the Challenges (listed and exemplified in **Table 1 (left)**) and determining approaches to overcome them.
- Actionable paths-forward and plans for implementing those approaches (exemplified by the Objectives in **Table 1 (right)**)
- Plans for enabling continuing community interaction by establishing effective communication channels such as rapid dissemination of anticipated vendor interface changes to tool teams and mechanisms for facility operators and application developers to communicate needs.

Bringing the requisite communities together in Town Hall discussion will provide greater awareness of the tools, their dependencies in the ecosystem, a basis for cross-community collaborations, and a significant increase in aggregate knowledge.

An additional artifact of this proposal is an initial high-level plan for a STEP Center, scaling the STEP activities beyond year one, with the goal of ensuring the long-term coordinated community-driven development directions necessary to sustain the HPC tools ecosystem. This plan will be articulated in a report.

STEP will actively engage the scientific community through a number of outreach and dissemination activities:

- *Report*: We will actively communicate STEP findings in all appropriate venues.
- *Presentations*: We will give posters and presentations at windows of opportunity (e.g., conferences and workshops relevant to the DOE science community). We also anticipate giving invited presentations at other national labs and academic institutions.
- *Direct engagement with other software ecosystem sustainment planning efforts*: We will actively work with any of the software sustainment teams interested in aspects of our technology andor approach.
- *Web page*: We will host a publicly available STEP web page that will provide a project overview and all of our documents.