OAK RIDGE National Laboratory

Scientific Workflows: State-of-the-art and Challenges

Rafael Ferreira da Silva, Ph.D.

Senior Research Scientist
National Center for Computational Sciences

https://rafaelsilva.com - silvarf@ornl.gov

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

U.S. DEPARTMENT OF ENERGY

Outline

- Scientific Workflows and Leadership Class Computing
- Workflows: state-of-the-art and Needs
- Workflows: tackling (some of) the challenges
- Workflows Community: sustaining a community roadmap

OAK RIDGE National Laboratory

Scientific Workflows and Leadership-Class Computing

Leadership Computing Facility (LCF)

*Provide the **computational** and **data resources** required to solve the most challenging problems*

2-centers / 2-architectures

address diverse and growing computational needs of the scientific community

user-facility program

collaborative DOE Office of Science program at ORNL and ANL

highly competitive user allocation programs (INCITE, ALCC)

liaisons / catalysts

LCF centers partner with users to enable science & engineering breakthroughs

Projects receive 10x to 100x more resource than at other generally available centers

Oak Ridge Leadership Computing Facility

Top 10 supercomputer in every year
since LCF was founded in 2005

Jaguar, Titan, Summit, and Frontier are the only DOE/SC systems to be **ranked #1** on the **TOP500** list of fastest computers

10^{15}

Jaguar
Cray XT5 2.3 PF
AMD CPU / 7 MW

2009

10^{16}

Titan
Cray XK6 27 PF
NVIDIA GPU, AMD CPU
9 MW

2012

10^{17}

Summit
IBM 200 PF
6 NVIDIA GPUs,
2 Power CPUs
13 MW

2017

10^{18}

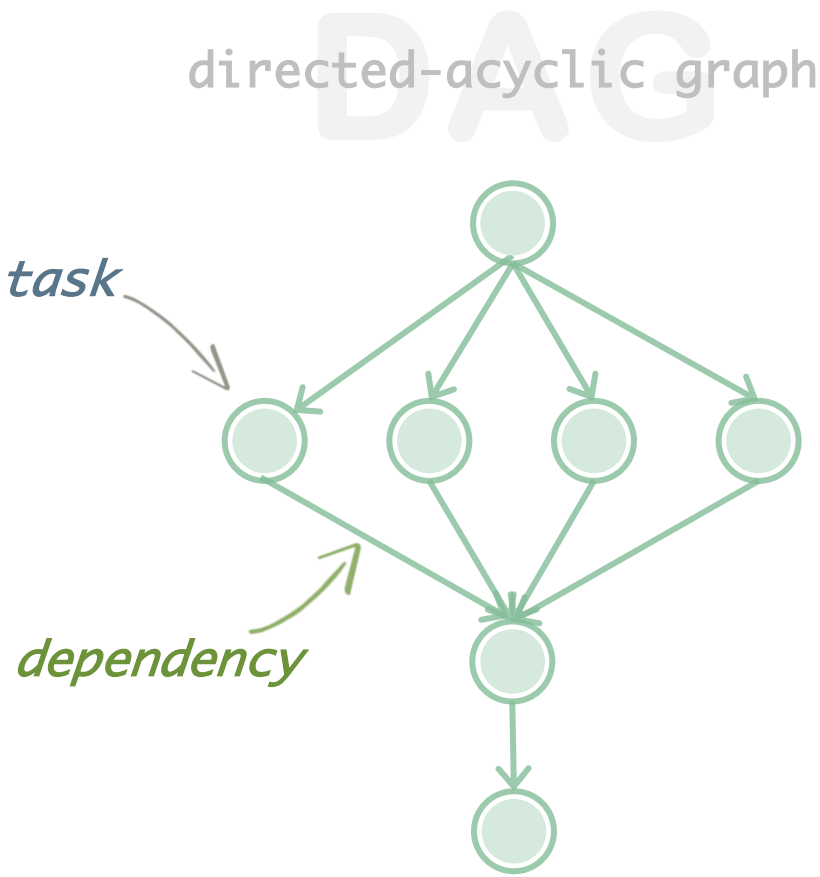
Frontier
Cray Shasta >1.6 EF

+ 4 Purpose-Built AMD Radeon Instinct GPU
1 HPC and AI Optimized AMD EPYC CPU
~10 PB of high bandwidth and DDR memory
Cray four-port Slingshot network (100 GB/s)
2-4x performance and capacity of Summit's I/O subsystem
29 MW

2022

Petascale | **(Exascale)**

Scientific Workflows *(traditional view)*

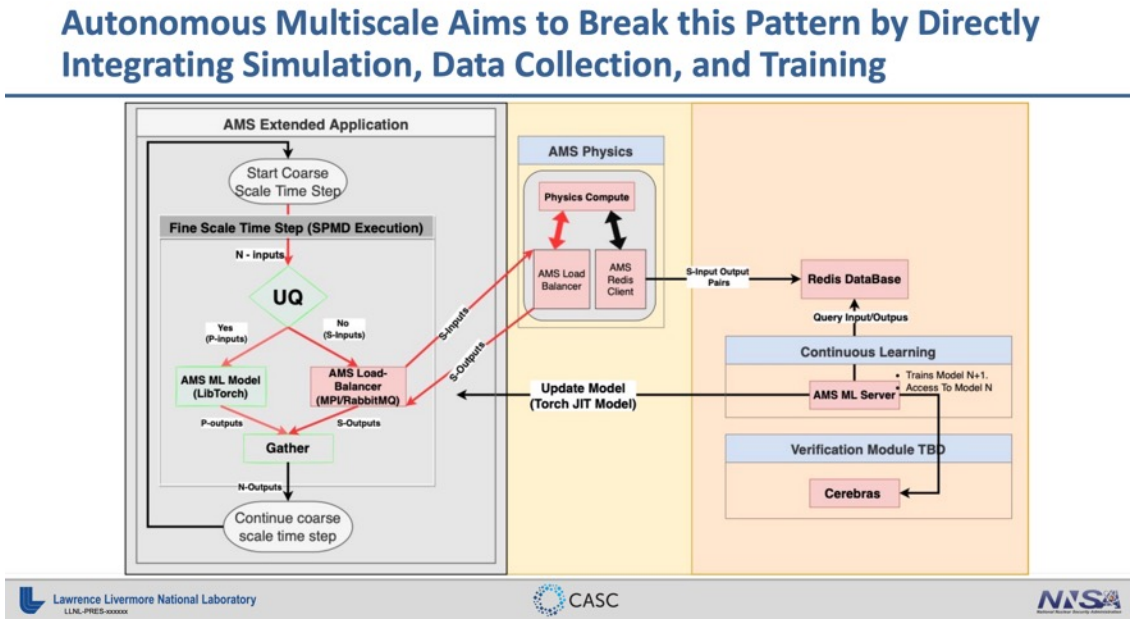


A task often represents a **program** (or script) written in any programming language (**closed box****)

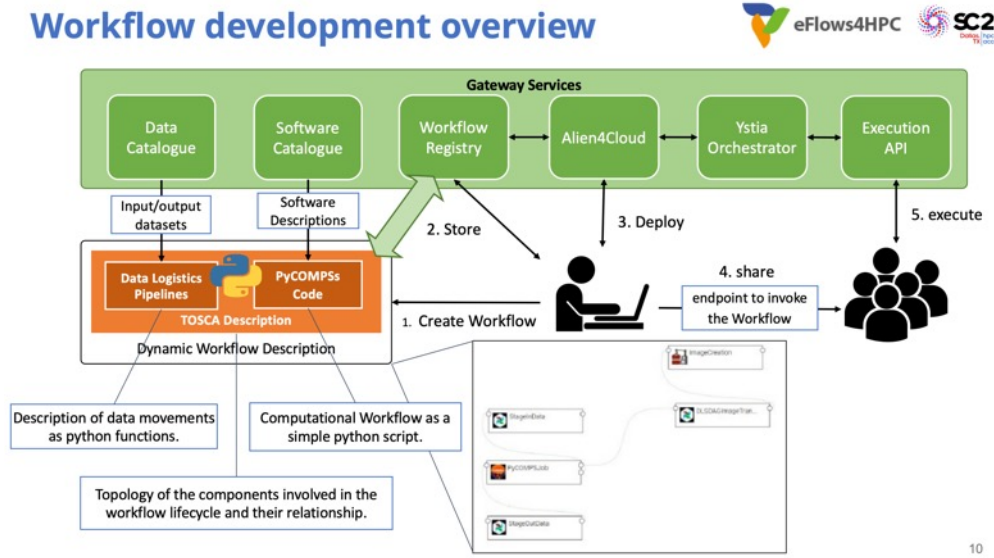
Dependencies are typically based on the **data flow**. It can also be expressed as **conditions**, **exceptions**, **user triggered action**, etc.

**https://www.duncannisbet.co.uk/removing-harmful-language-from-my-lexicon

Modern Scientific Workflows

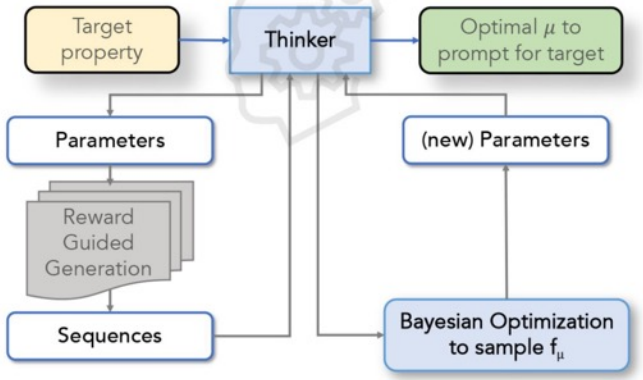


Courtesy of Timo Bremer (LLNL)



Courtesy of Rosa M. Badia (BSC)

A workflow that allows for continuous update of rewards for new sequences generated

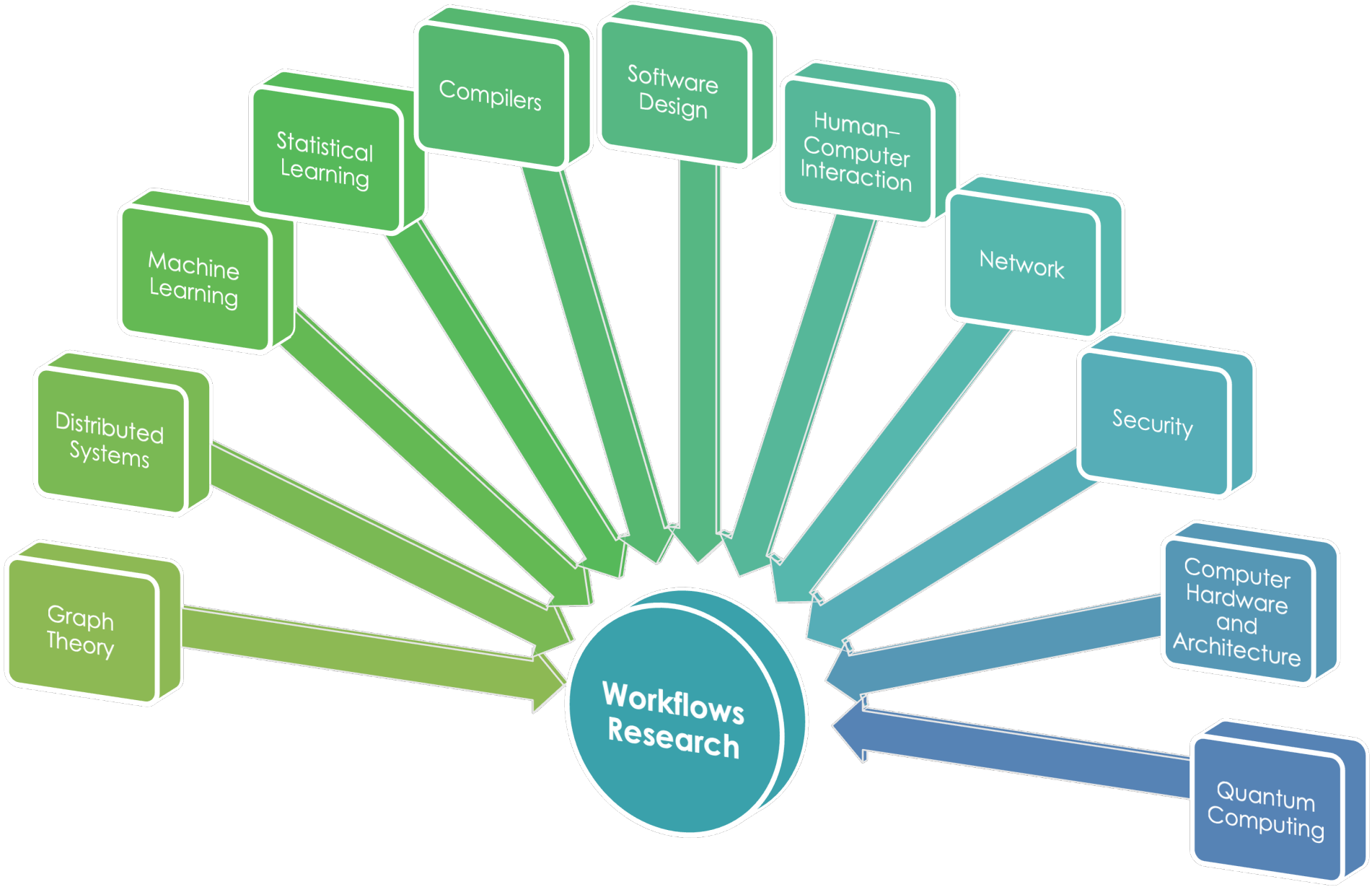


- To learn more visit us at the Gordon Bell Award talk for COVID-19 research
  - C142-145 Wed, Nov 16, 10.30 AM
- The workflow has many aspects that are unique to biological data analyses
- Implemented with Colmena which also executes across multiple sites
- Scaling out across sites and across multiple architectures can be challenging
- Not limited to one learning approach but can potentially exploit multi-modal, multi-view learning from data
- Smart and AI-enabled optimization of workload resources

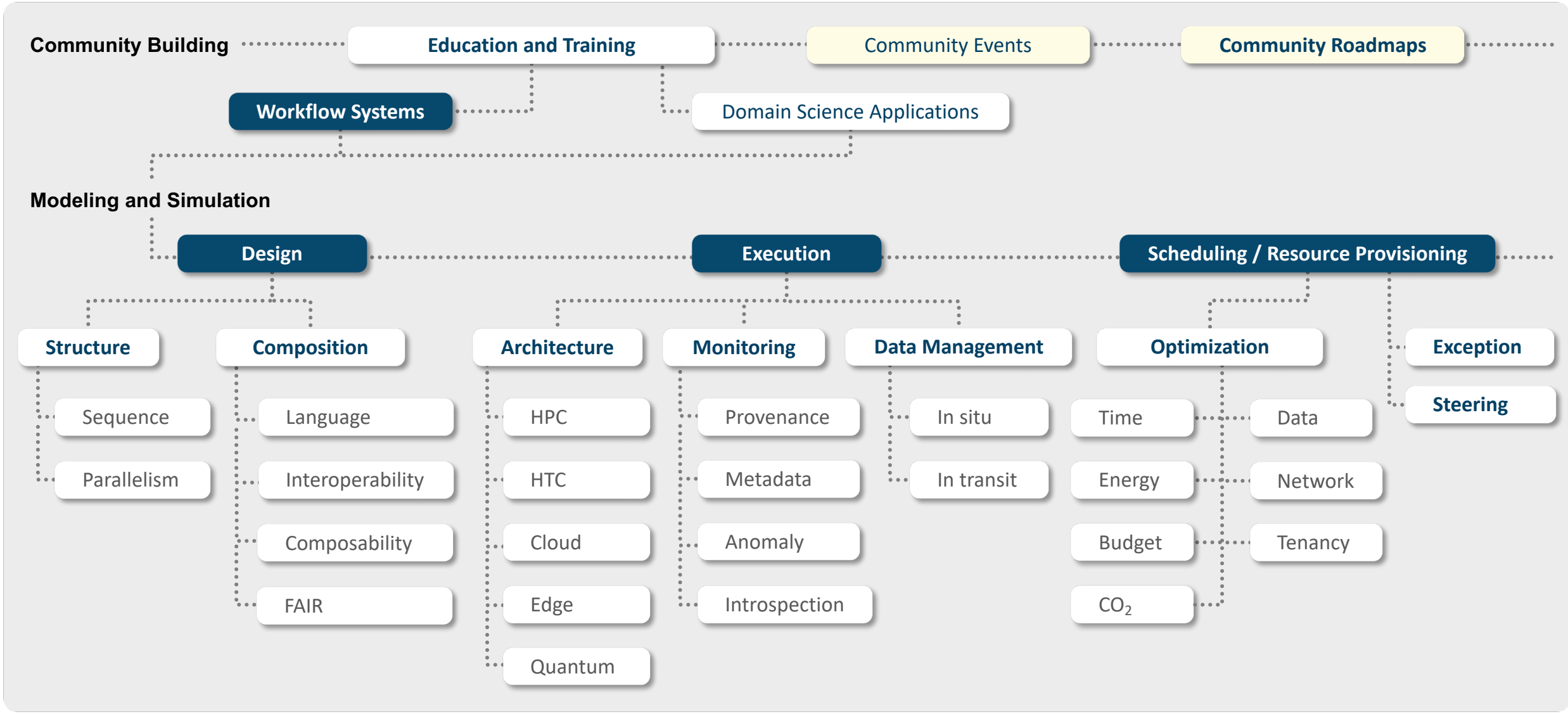
Courtesy of Arvind Ramanathan (ANL)

OAK RIDGE National Laboratory

Interaction with Computer Science Disciplines



Overview of Scientific Workflows Research Challenges

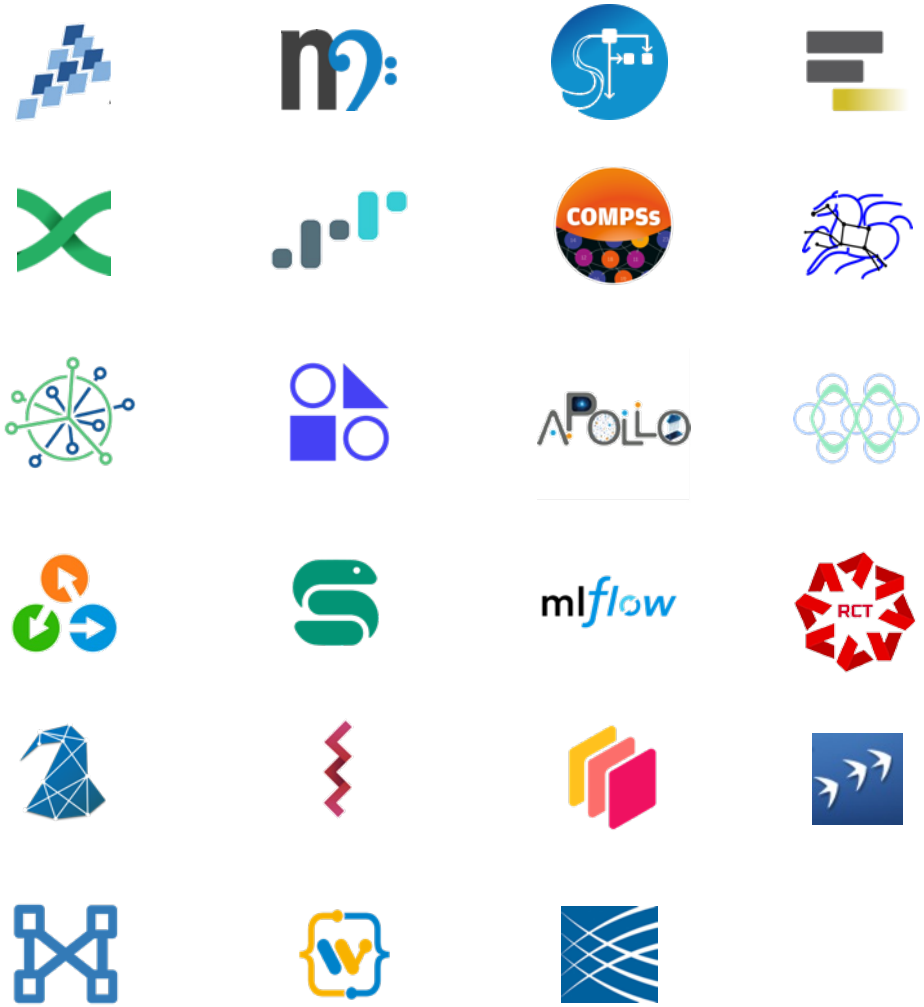


The Universe of Workflow Systems

Forcing applications to fit a particular system

Needs a developer to be "shipped" with the workflow system

The term "workflow" is overloaded



https://s.apache.org/existing-workflow-systems

325 workflow systems as of Mar 9, 2023

The 2021 Community Roadmap

arXiv.org > cs > arXiv:2110.02168

Search... Help | Advance

Computer Science > Distributed, Parallel, and Cluster Computing

[Submitted on 5 Oct 2021 (v1), last revised 8 Oct 2021 (this version, v2)]

A Community Roadmap for Scientific Workflows Research and Development

Rafael Ferreira da Silva, Henri Casanova, Kyle Chard, Ilkay Altintas, Rosa M Badia, Bartosz Balis, Tainã Coleman, Frederik Coppens, Frank Di Natale, Bjoern Enders, Thomas Fahringer, Rosa Filgueira, Grigori Fursin, Daniel Garijo, Carole Goble, Dorran Howell, Shantenu Jha, Daniel S. Katz, Daniel Laney, Ulf Leser, Maciej Malawski, Kshitij Mehta, Loïc Pottier, Jonathan Ozik, J. Luc Peterson, Lavanya Ramakrishnan, Stian Soiland-Reyes, Douglas Thain, Matthew Wolf

The landscape of workflow systems for scientific applications is notoriously convoluted with hundreds of seemingly equivalent workflow systems, many isolated research claims, and a steep learning curve. To address some of these challenges and lay the groundwork for transforming workflows research and development, the WorkflowsRI and ExaWorks projects partnered to bring the international workflows community together. This paper reports on discussions and findings from two virtual "Workflows Community Summits" (January and April, 2021). The overarching goals of these workshops were to develop a view of the state of the art, identify crucial research challenges in the workflows community, articulate a vision for potential community efforts, and discuss technical approaches for realizing this vision. To this end, participants identified six broad themes: FAIR computational workflows; AI workflows; exascale challenges; APIs, interoperability, reuse, and standards; training and education; and building a workflows community. We summarize discussions and recommendations for each of these themes.

Comments: arXiv admin note: substantial text overlap with arXiv:2103.09181
Subjects: **Distributed, Parallel, and Cluster Computing (cs.DC)**
Cite as: arXiv:2110.02168 **[cs.DC]**
(or arXiv:2110.02168v2 **[cs.DC]** for this version)

We summarize the discussions and findings by presenting a consolidated view of the **state of the art**, **challenges**, and potential efforts, which we eventually synthesize into a **community roadmap**

https://arxiv.org/abs/2110.02168

Challenges

Theme	Challenges
FAIR Computational Workflows	• Define FAIR principles for computational workflows that consider the complex lifecycle from specification to execution and data products • Define metrics to measure the FAIRness of a workflow. • Engage the community to define principles, policies, and best practices
AI Workflows	• Lack of support for heterogeneity of compute resources and fine-grained data management features, versioning, and data provenance capabilities • Lack of capabilities for enabling workflow steering and dynamic workflows • Integration of ML frameworks into the current HPC landscape
Exascale Challenges and Beyond	• Resource allocation policies and schedulers are not designed for workflow-aware abstractions, thus users tend to use an ill-fitted job abstraction • Unfavorable design of resource descriptions and mechanisms for workflow users/systems, and lack of fault-tolerance and fault-recovery solutions
APIs, Reuse, Interoperability, and Standards	• Workflow systems differ by design, thus interoperability at some layers is likely to be more impactful than others • Workflow standards are typically developed by a subset of the community • Quantifying the value of common representations of workflows is not trivial
Training and Education	• Many workflow systems have high barrier to entry and lack training material • Homegrown workflow solutions and constraints can prevent users from reproducing their functionality on workflow tools developed by others • Unawareness of the workflow technological and conceptual landscape
Building a Workflows Community	• Define what is meant by a "workflows community" • Remedy the inability to link developers and users to bridge translational gaps • Pathways for participation in a network of researchers, developers, and users

Proposed Community Activities

Theme	Community Activities
FAIR Computational Workflows	• Review prior and current efforts for FAIR data and software with respect to workflows, and outline rules for FAIR workflows • Define recommendations for FAIR workflow developers and systems • Automate FAIRness in workflows by recording necessary provenance data
AI Workflows	• Develop comprehensive use cases for sample problems with representative workflow structures and data types • Define a process for characterizing the challenges for enabling AI workflows • Develop AI workflows as a way to benchmark HPC systems
Exascale Challenges and Beyond	• Develop documentation in the form of workflow templates/recipes/miniapps for execution on high-end HPC systems • Specify benchmark workflows for exascale execution • Include workflow requirements as part of the machine procurement process
APIs, Reuse, Interoperability, and Standards	• Identify differences and commonalities between different systems • Identify and characterize domain-specific efforts, identify workflow patterns, and develop case-studies of business process workflows and serverless workflow systems
Training and Education	• Identify basic sample workflow patterns, develop a community workflow knowledge-base, and look at current research on technology adoption • Include workflow terminology and concepts in university curricula and software carpentry efforts
Building a Workflows Community	• Establish a common knowledge-base for workflow technology • Establish a *Workflow Guild*: an organization focused on interaction and good relationships and self-support between workflow developers and their systems

Roadmap

Thrust	Roadmap Milestones
Definition of common workflow patterns and benchmarks	• Define small sets of workflow pattern and benchmark deliverables, and implement them using a selected set of workflow systems • Investigate automatic generation of patterns and configurable benchmarks (to enable weak and strong scaling experiments) • Establish or leverage a centralized repository to host and curate patterns and benchmarks
Identifying paths toward interoperability of workflow systems	• Define interoperability for different roles, develop a horizontal interoperability (i.e., making interoperable components), and establish a requirements document per abstraction layer • Develop real-world workflow benchmarks, use cases for interoperability, and common APIs that represent workflow library components • Establish a workflow systems developer community
Improving workflow systems' interface with legacy and emerging HPC software and hardware stacks	• Document a machine-readable description of key properties of widely used sites, and remote authentication needs from the workflow perspective • Identify new workflow patterns (e.g. motivated from AI workflows), attain portability across heterogeneous hardware, and develop a registry of execution environment information • Organize a community event involving workflow system developers, end users, authentication technology providers, and facility operators

National Academy of Sciences, Engineering, and Medicine Report

The needs and demands placed on science to address a range of urgent problems are growing. The world is faced with complex, interrelated challenges in which the way forward lies hidden or dispersed across disciplines and organizations. For centuries, scientific research has progressed through iteration of a ***workflow built on experimentation or observation and analysis of the resulting data****. While computers and automation technologies have played a central role in research workflows for decades to acquire, process, and analyze data, these same computing and* ***automation technologies can now also control the acquisition of data****, for example, through the design of new experiments or decision making about new observations.*

https://doi.org/10.17226/26532

OAK RIDGE National Laboratory

Workflows: Tackling (some of) the Challenges

Modeling and Simulation

Cyberinfrastructure Simulation Workbench

WRENCH

https://wrench-project.org

SIMULATORS

SIMULATORS OF WORKLOAD EXECUTIONS WITH RUNTIME SYSTEMS

USER API

SIMULATED RUNTIME SYSTEMS

- General-purpose runtime systems
- Workflow management systems
- Research prototype runtime systems

DEVELOPER API

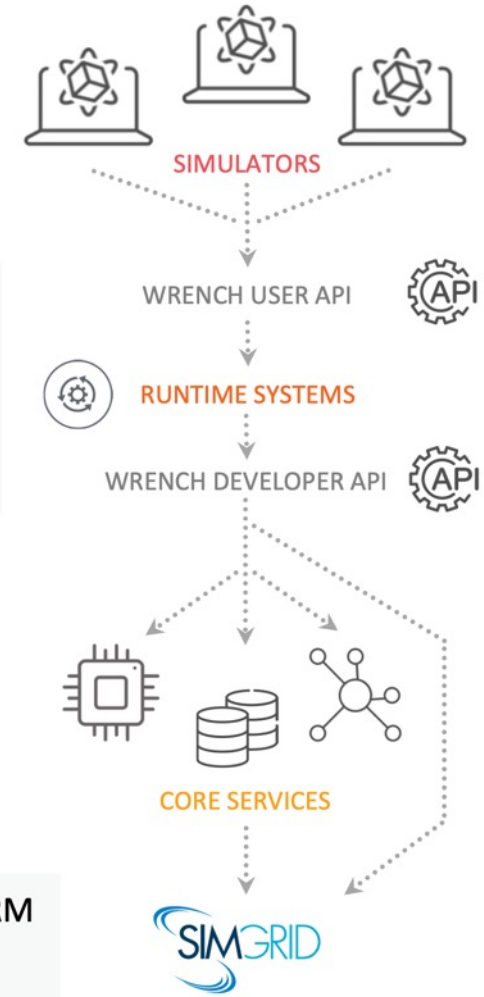
SIMULATED CORE SERVICES

- Computation (Cloud, Batch, Rack)
- Storage (XRootD, Globus, FTP)
- Network Monitoring (Vivaldi)
- Data Location (Replica Catalog)

SIMGRID

SIMULATED PLATFORM

- SimGrid S4U API



SIMGRID

https://simgrid.org

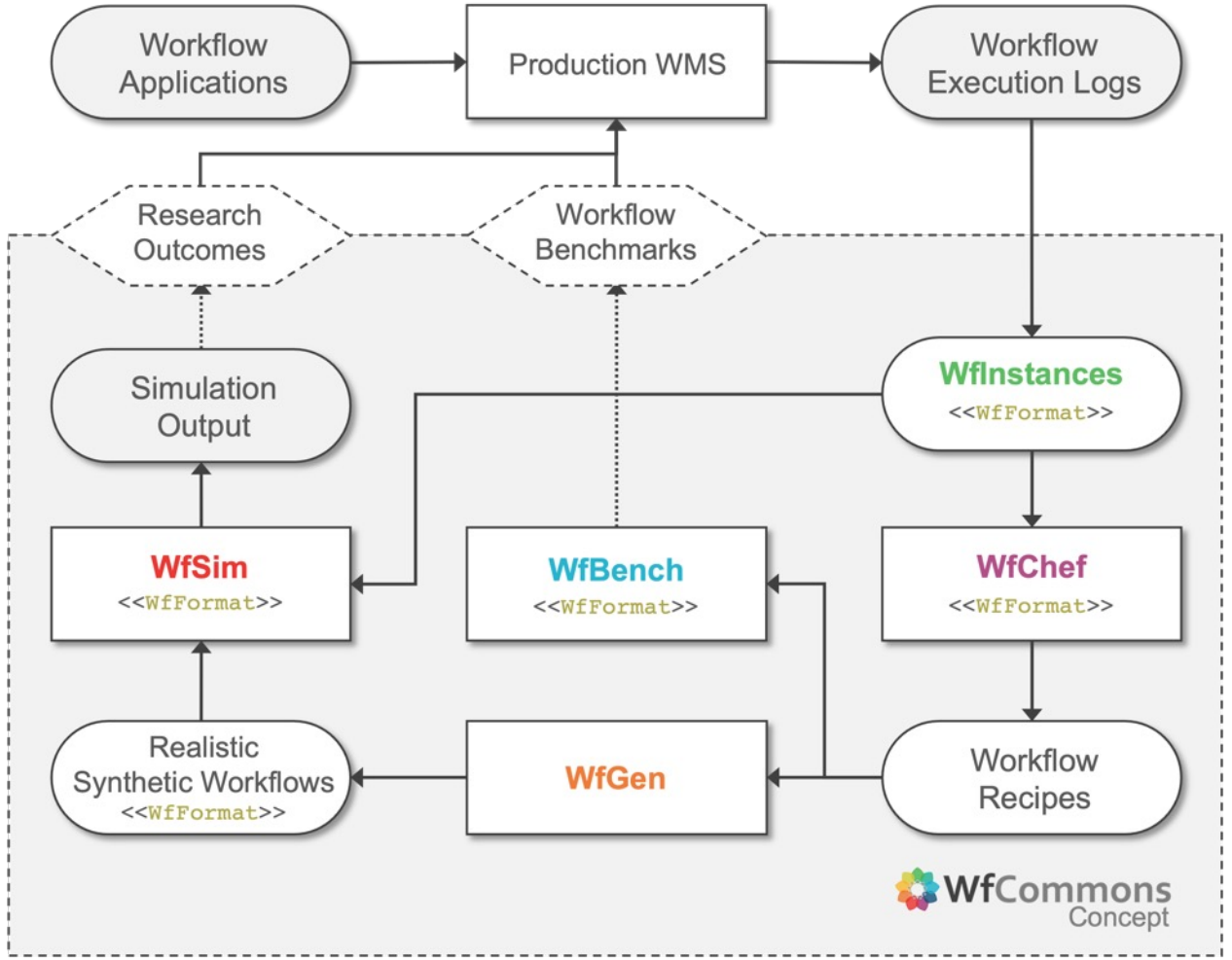
Accurate and **scalable simulation models** of hardware/software stacks

OAK RIDGE National Laboratory

Framework for Enabling Workflow Research and Development

WfCommons
https://wfcommons.org

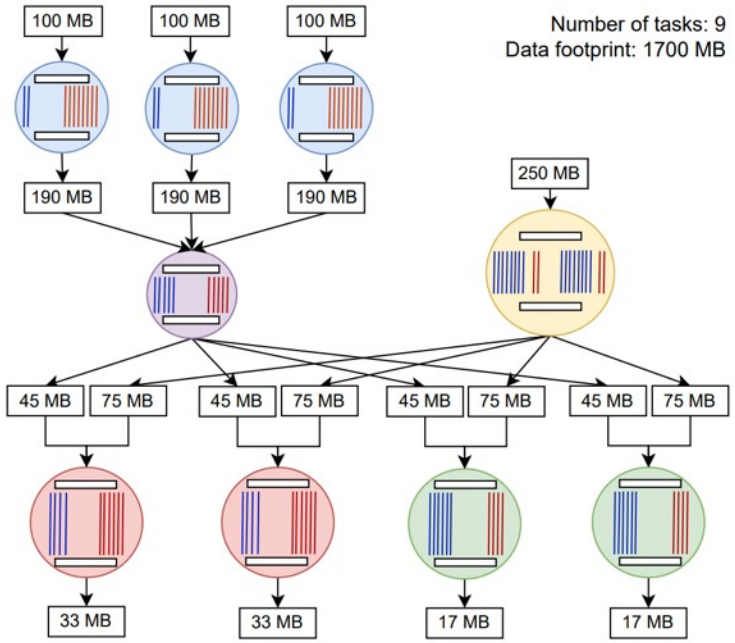
WfCommons is a community framework that provides a collection of tools for analyzing workflow **execution instances**, producing realistic **synthetic workflow instances**, and **simulating workflow** executions



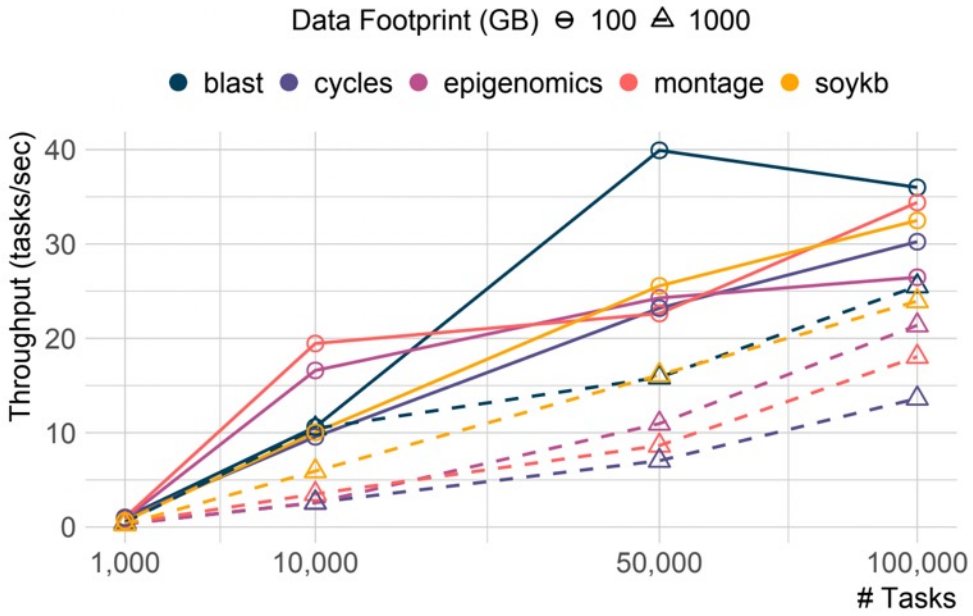
Open source Python package to **analyze instances** and generate representative, realistic **synthetic instances** from workflow recipes

OAK RIDGE National Laboratory

WfBench: Workflow Benchmarks



Representative **tasks** and **workflow** benchmarks



↑ data footprint ↓ throughput
↑ #tasks ↑ throughput

Analysis of workflow system **overhead**

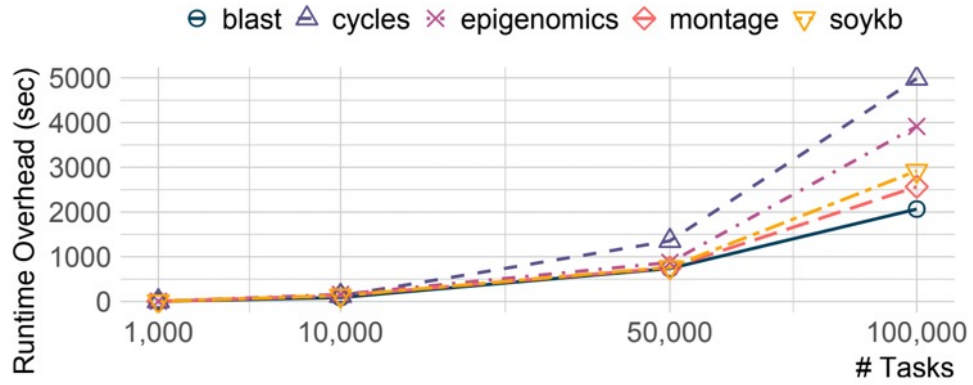
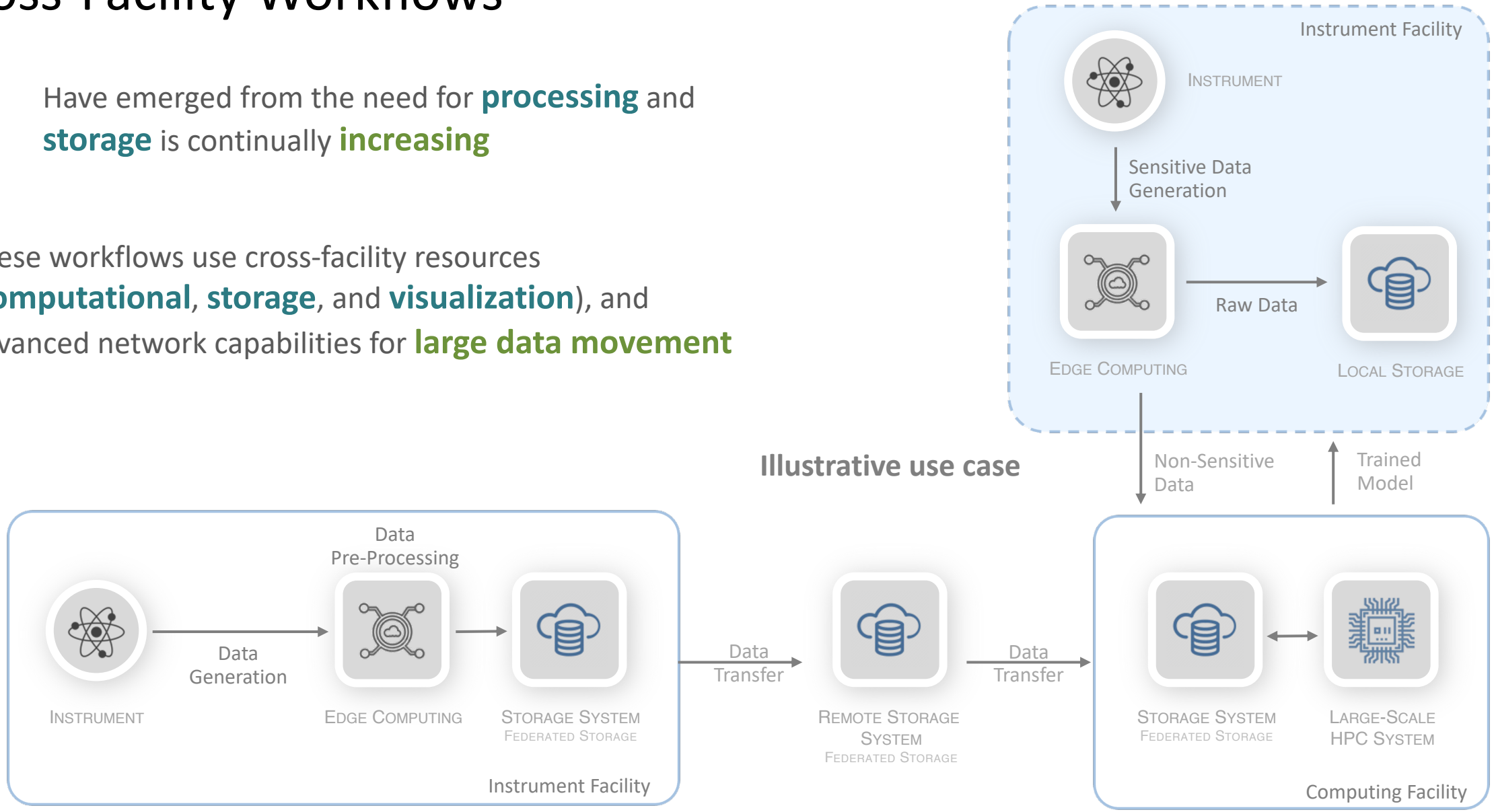


Fig. 10. Workflow execution time (or total overhead) vs. number of tasks.

Cross-Facility Workflows

Have emerged from the need for **processing** and **storage** is continually **increasing**

These workflows use cross-facility resources (**computational**, **storage**, and **visualization**), and advanced network capabilities for **large data movement**



Cross-Facility Workflows: Challenges

Data movement across facilities require **high bandwidth** and dedicated links

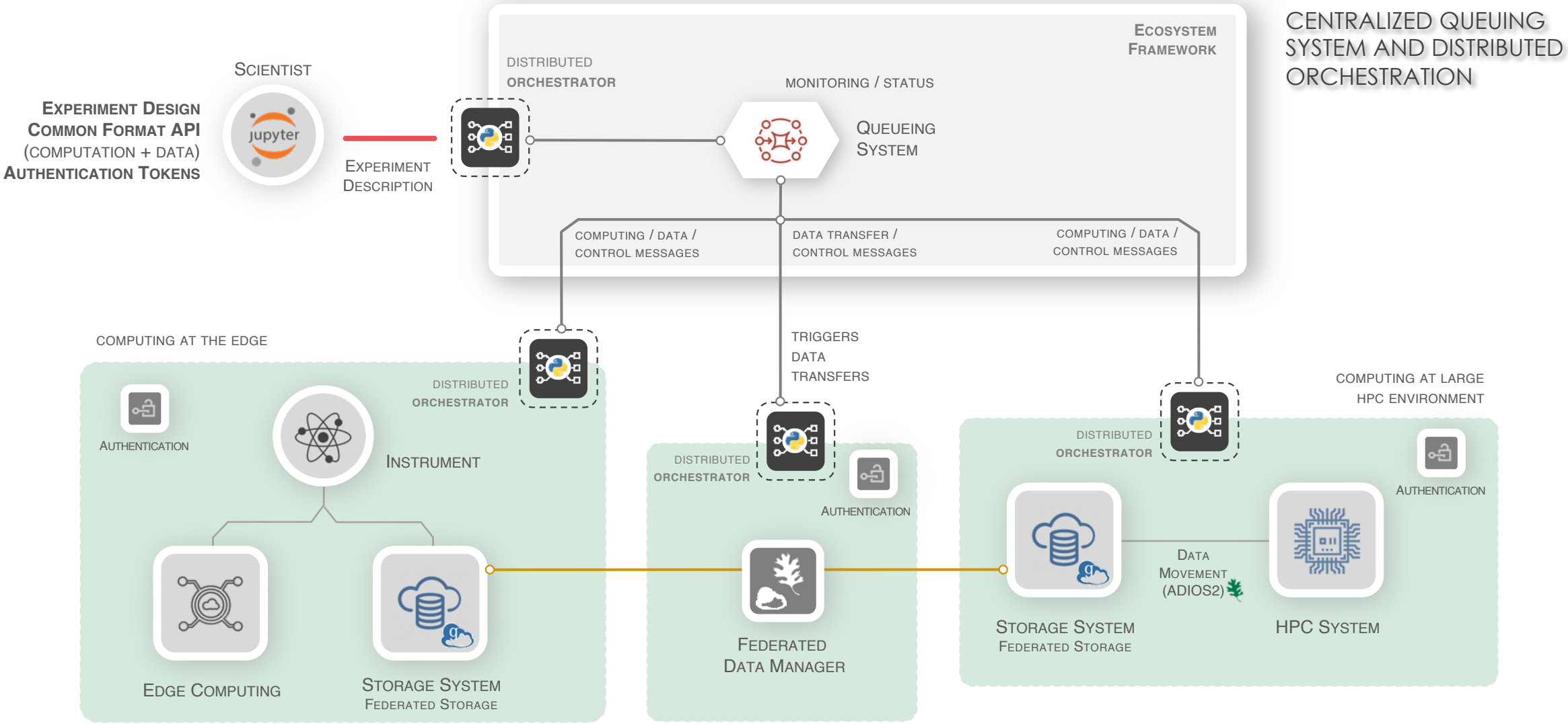
Lack of **common representations** (or interfaces) to describe workflows

Complex **authentication**, **authorization**, and **accounting** (AAA) mechanisms within and across facilities

Policies for **sensitive data** may limit the ability for external workload processing

Facilities have different **software stacks** and more importantly different sets of **policies**

Zambeze Framework



OAK RIDGE National Laboratory

ExaWorks

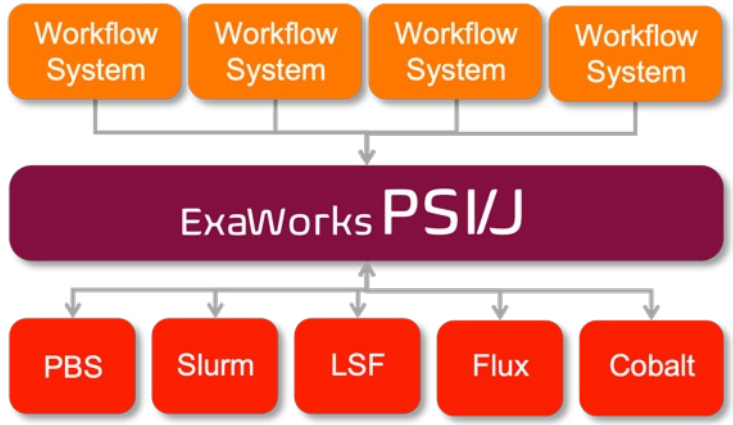
PSI/J, a Portable Submission interface for jobs

Motivated by a sequence of **interviews and community meetings** identifying job launch/management as low hanging fruit

Community generated a **light-weight user-space API Specification** via a public process hosted on GitHub

ECP ExaWorks team created an Initial **Python reference implementation**

ECP ExaWorks creating a public dashboard and CI infrastructure for multiple DOE compute centers that could be extended by community




```python
import jpsi

jex = jpsi.JobExecutor.get_instance('slurm')

def make_job():
    job = jpsi.Job()
    spec = jpsi.JobSpec()
    spec.executable = '/bin/sleep'
    spec.arguments = ['10']
    job.spec = spec
    return job

jobs = []
for i in range(N):
    job = make_job()
    jobs.append(job)
    jex.submit(job)

for i in range(N):
    jobs[i].wait()
```

OAK RIDGE National Laboratory

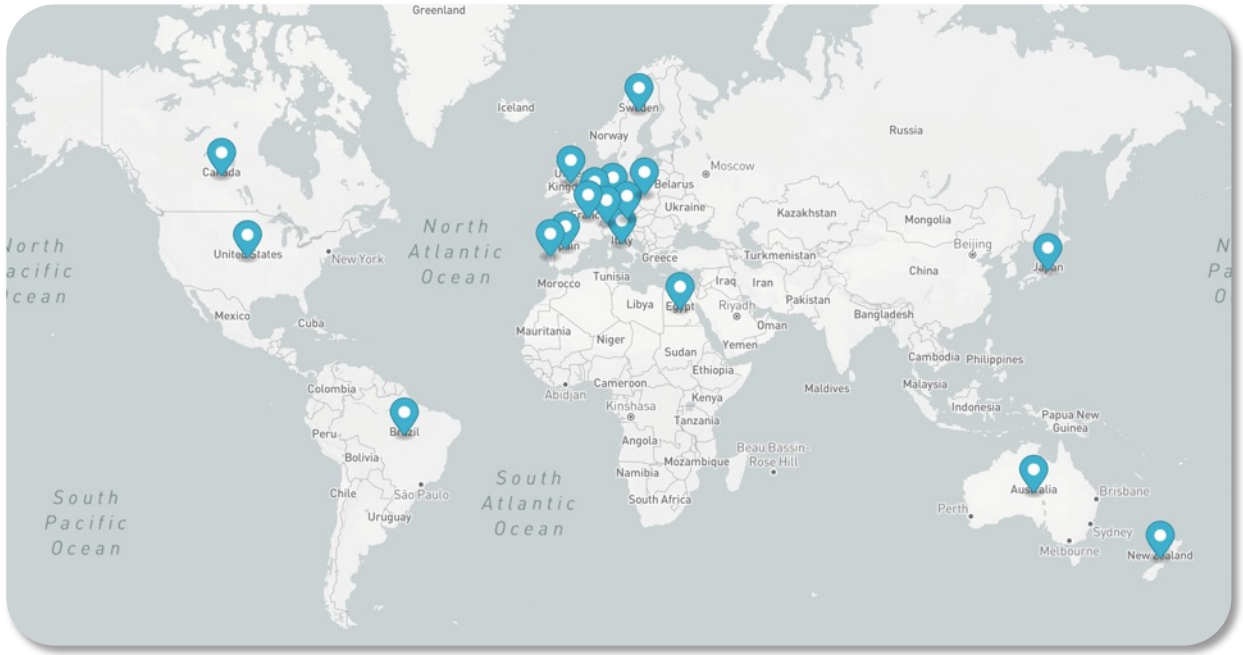
Workflows Community: Sustaining a Community Roadmap

Workflows Community Initiative

https://workflows.community

*The goal of the **Workflows Community Initiative** (WCI) is to bring the workflows community together (**users, developers, researchers, and facilities**) to provide community resources and capabilities to enable scientists and workflow systems developers to discover software products, related efforts, events, technical reports, etc. and engage in community-wide efforts to tackle workflows grand challenges.*

We are always looking for volunteers…



25 workflow systems

155 community members

6 working groups

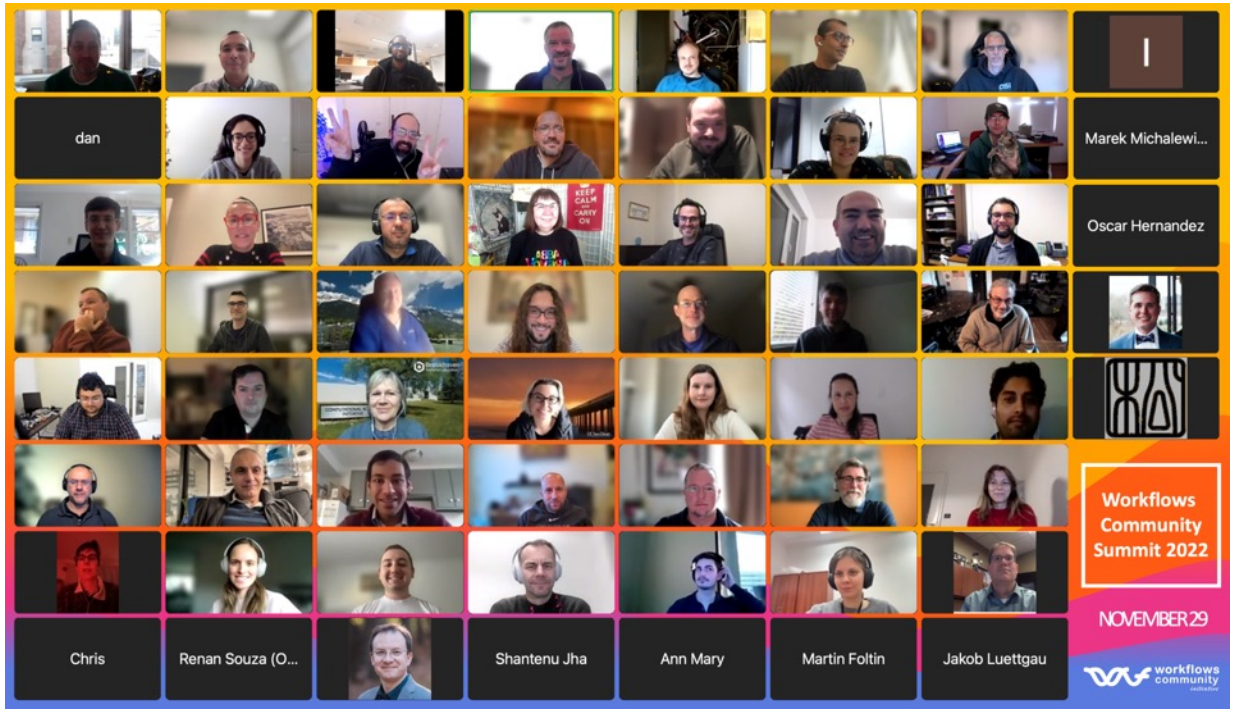
4 research frameworks

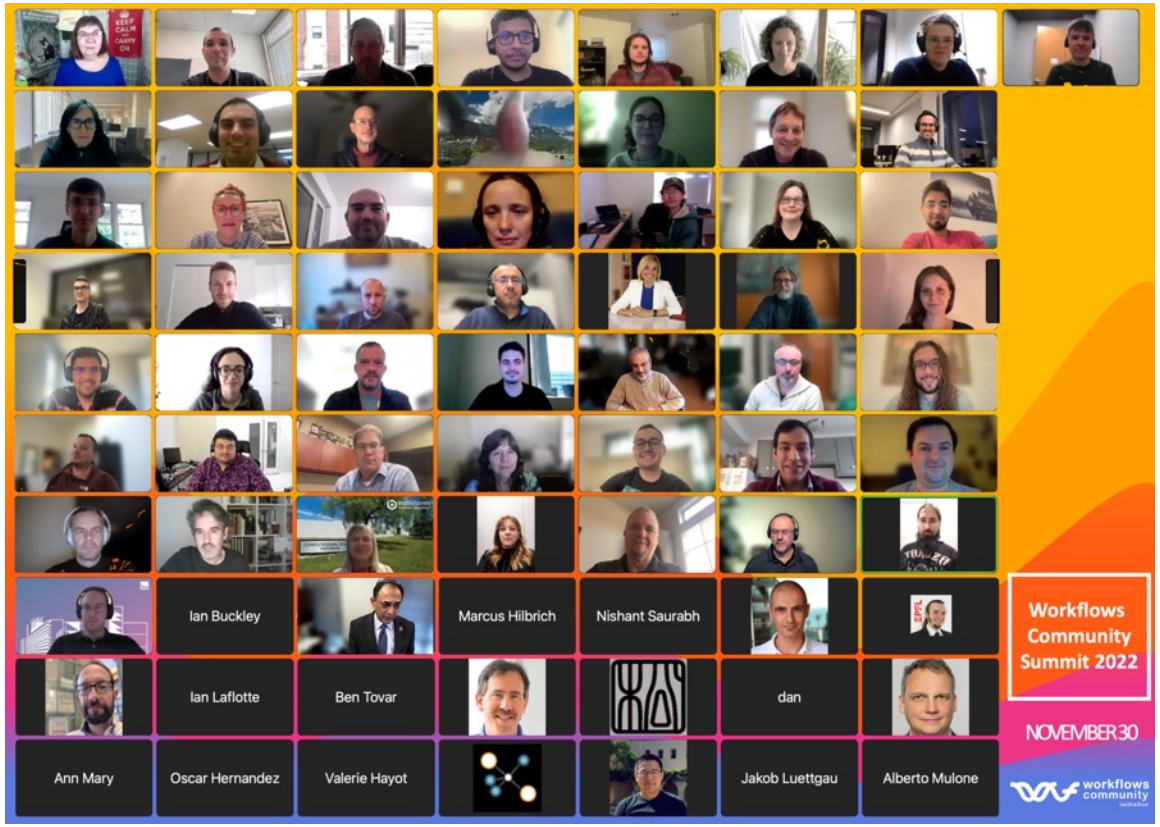
OAK RIDGE National Laboratory

Workflows Community Summit 2022

Nov 29 and 30, 2022

https://workflows.community





100+ international participants

OAK RIDGE National Laboratory

Revised Roadmap: Initial Thoughts

Specifications, standards, and APIs

Standards are often constrained or **hard to implement**
Define **common terms**, building blocks, and concepts
Define some standard for giving **data** from user to workflow systems/operators

AI workflows

Unreliability of models – training processes often have a **human-in-the-loop**
Types of workflows – *workflows to **create AI*** and *workflows that **use AI***
Challenges: **random access** to datasets in training and management of **small files**

High performance data management and in situ workflows

Edge to cloud continuum and data exchange through data objects
Data usage is more **fine-grain** than a typical HPC code
Adaptive **compression** to reduce data necessary to represent the problem domain

Revised Roadmap: Initial Thoughts

HPC and Quantum workflows

Community do not know how to transfer information to the QC system
Limited resources (hard to access), long queues & expensive
Heterogeneity in quantum devices (vendor specific APIs)

FAIR workflows

Standards for expressing the **inputs** of the workflow
Limited availability of **metadata**
FAIR data and FAIR workflows are **intertwingled**

Workflows for continuum and cross-facility computing

Describe **aggregate I/O** needs of a workflow
Coordinate **communication** between sites (different **security** domains)
People who design **experiment facilities** are not necessarily computing experts

OAK RIDGE National Laboratory

Scientific Workflows: State-of-the-art and Challenges

Rafael Ferreira da Silva, Ph.D.

Senior Research Scientist
Data Lifecycle and Scalable Workflows

https://rafaelsilva.com – silvarf@ornl.gov

This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.