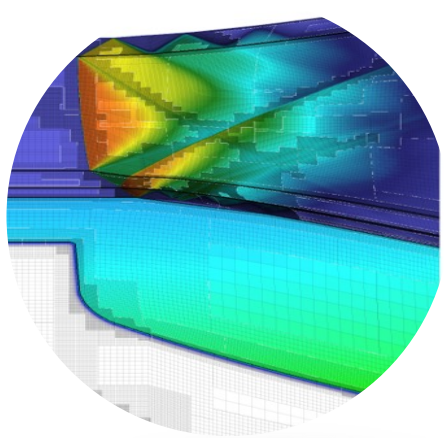
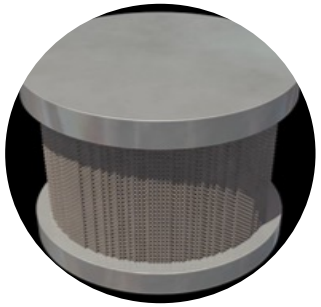
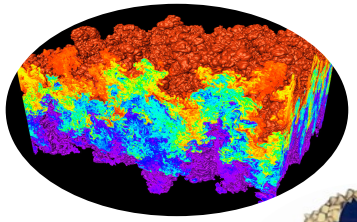
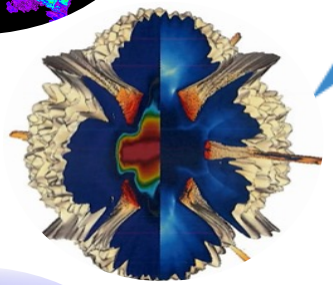
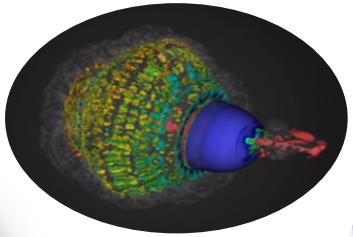
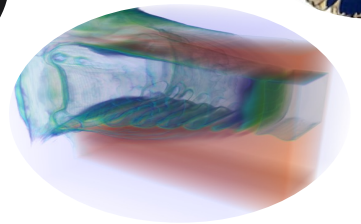









ADVANCED SIMULATION & COMPUTING™

Scientific Workflow Management Challenges and Tools

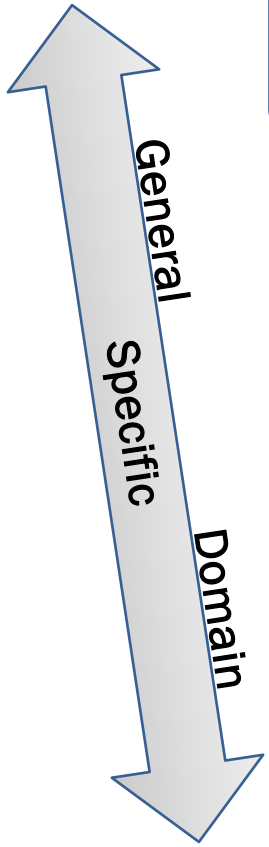
PSAAP-IV Preproposal Meeting
Houston, TX
Aug 8-9, 2023

Rob Neely
ASC Program Director, LLNL

LLNL-PRES-747686

Lawrence Livermore National Laboratory
NNSA National Nuclear Security Administration

“Workflow” - let’s define it…



Workflow

An orchestrated and **repeatable** pattern of business activity enabled by the **systematic** organization of resources into processes

Workflow Management System

An infrastructure for the set-up, performance and monitoring of a defined sequence of tasks, arranged as a **workflow** application.

Scientific Workflow System

A specialized form of a **workflow management system** designed specifically to compose and execute a series of computational or data manipulation steps (workflow) in a scientific application.

Source: Wikipedia The Free Encyclopedia

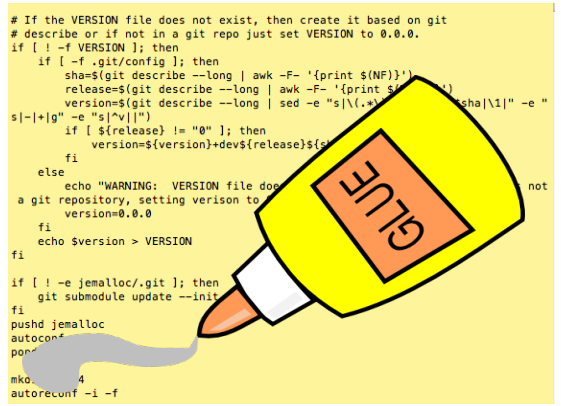
Commonly understood definitions in our community:

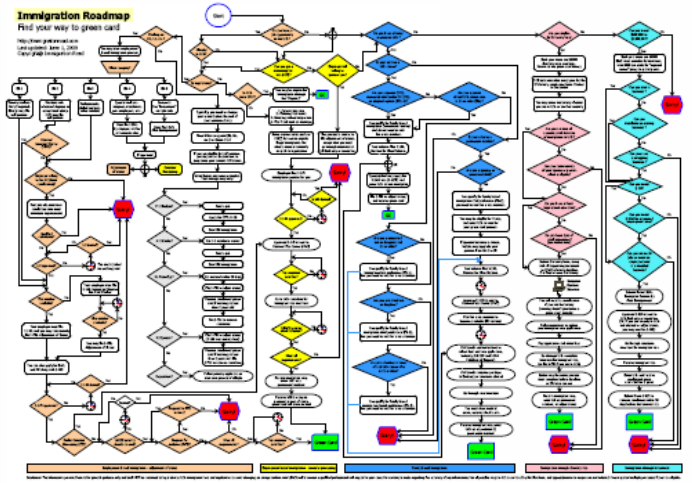
Workflow is the **process** and **data management** activities (both **computational** and **intellectual** in nature) that occur during the course of scientific discovery from **problem concept** to **resolution** and all points in between.

It often involves **multiple systems**, **codes**, **data sources**, and **tools** in a coordinated dance to ensure **accuracy, reliability, reproducibility, efficiency, shareability,** and **credibility.**

Your need for formalized and automated workflow will depend…

- Are you writing a research code designed to solve a single problem? Single platform?
  - Your workflow management is probably manual and easy to describe in a batch script
- Do you have repetitive tasks? Looking for reproducibility? Tying together digital inputs and outputs from multiple sources?
  - Hopefully you're (at a minimum) writing scripts, and managing them with version control
- Is your workflow complex? Data rich? Do you need to document it? Reproduce it for others?
  - Workflow management should be a piece of your project culture
  - Leverage **existing tools** to give your workflow structure and reproducibility





Motivation: Scientific simulation/analysis in NNSA often requires multi-year efforts across large multi-disciplinary teams

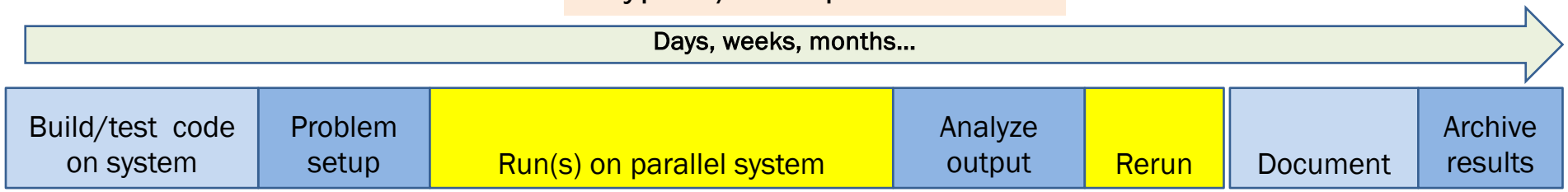
- Problem setup can take weeks/months (in some cases)
- Increasing use of ensembles and advanced statistical methods
- Multiple AI/ML workflows are being developed
- Integration of experimental data is a key component of our work
- Deluge of output – extracting information from data
- Need for reproducible results with *pedigree* and *provenance* of inputs, code versions, analysts
- Exascale systems are required to increase predictive capability but by themselves will not greatly enhance end-to-end analyst productivity

Security and access control requirements can slow adoption of new technologies and infrastructure inside the labs.
PSAAP == Proving Grounds

Courtesy: Dan Laney

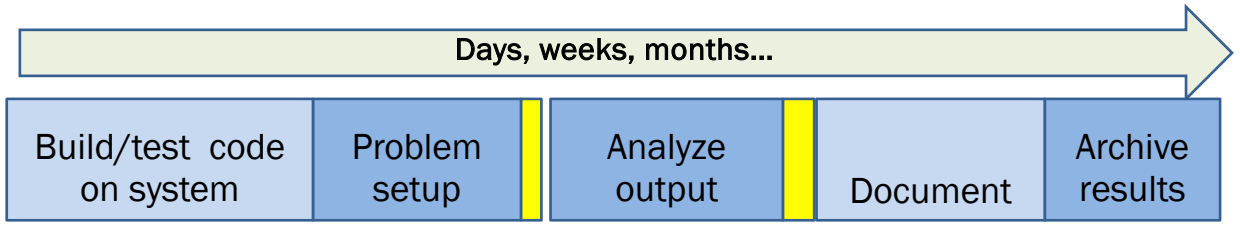
Workflow is like Amdahl's Law
Ironically, we often focus on the least costly parts

A typical/exemplar workflow



We spend a lot of time and money making this piece go fast (can you say "exascale"?)

But even if we're wildly successful, our overall time to solution is still…



I know, I know – this is a contrived example that is overly "waterfall" in nature…

Fast computers are a necessary but not sufficient solution to improved scientific capability and discovery

I/O in our simulation codes is complex and a key challenge for automation

Input Data

Experimental / Test Data
- Captures problem to simulate
- Comparison with simulation results

Output Data

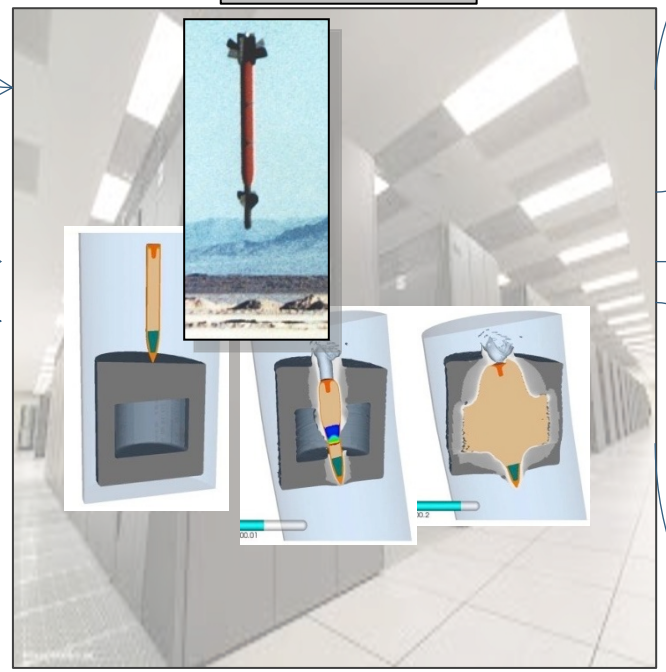
Input Files
- Small
- Version controlled

Material Properties
- Structured
- Small to medium sized

Meshes
- Discretized geometry and regions
- Medium sized

Geometry
- CAD-like descriptions
- Largely unstructured

Simulation



Graphics / Visualization
- Many large files
- Used for detailed analysis

Restart / Checkpoint
- Very large files
- Capture complete state

In-situ Analysis
- Predetermined views for vis
- Increasingly used for machine learning

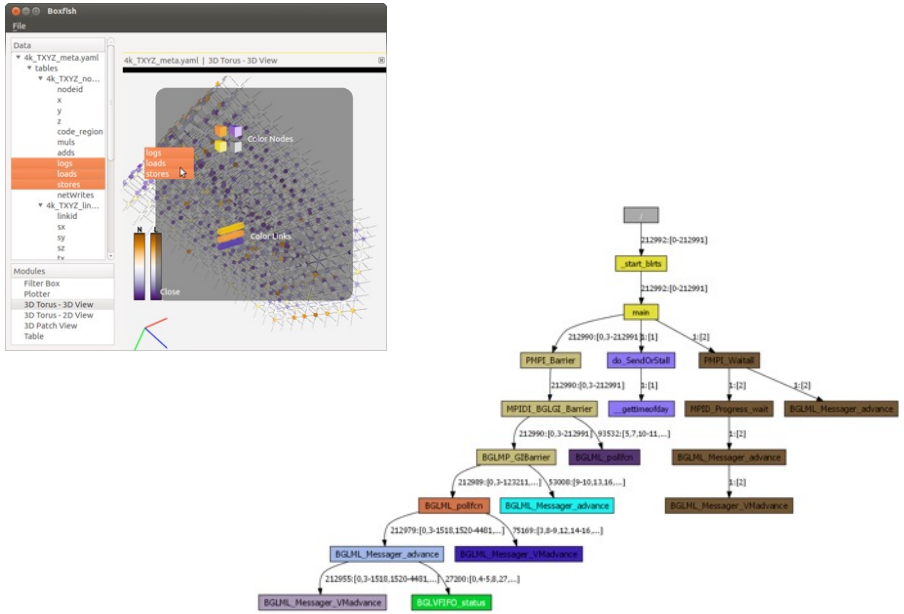
Log Files
- Small unstructured data
- Capture run status

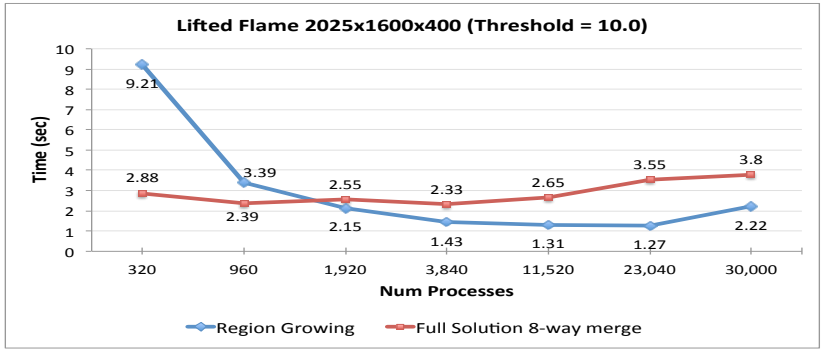
Performance metrics
- Unstructured
- Optional, but important

Most input data is ideally shared, version controlled, and requires complete traceability

Most output data is captured and archived. Sometimes cheaper to reproduce rather than save

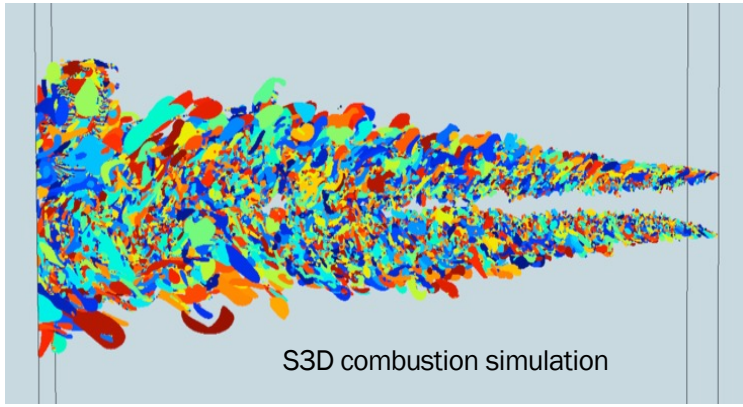
In-situ analysis is an enabling technology for ML and UQ workflows





New in-situ analytics techniques for large scale simulations allow one to bypass the I/O bottleneck (disk) and reduce data sets offering a much higher frequency of analysis results

Research is needed to scale performance analysis and debugging tools to extreme scale systems via continuous data collection, in-situ data reduction methods, and automated analysis


S3D combustion simulation

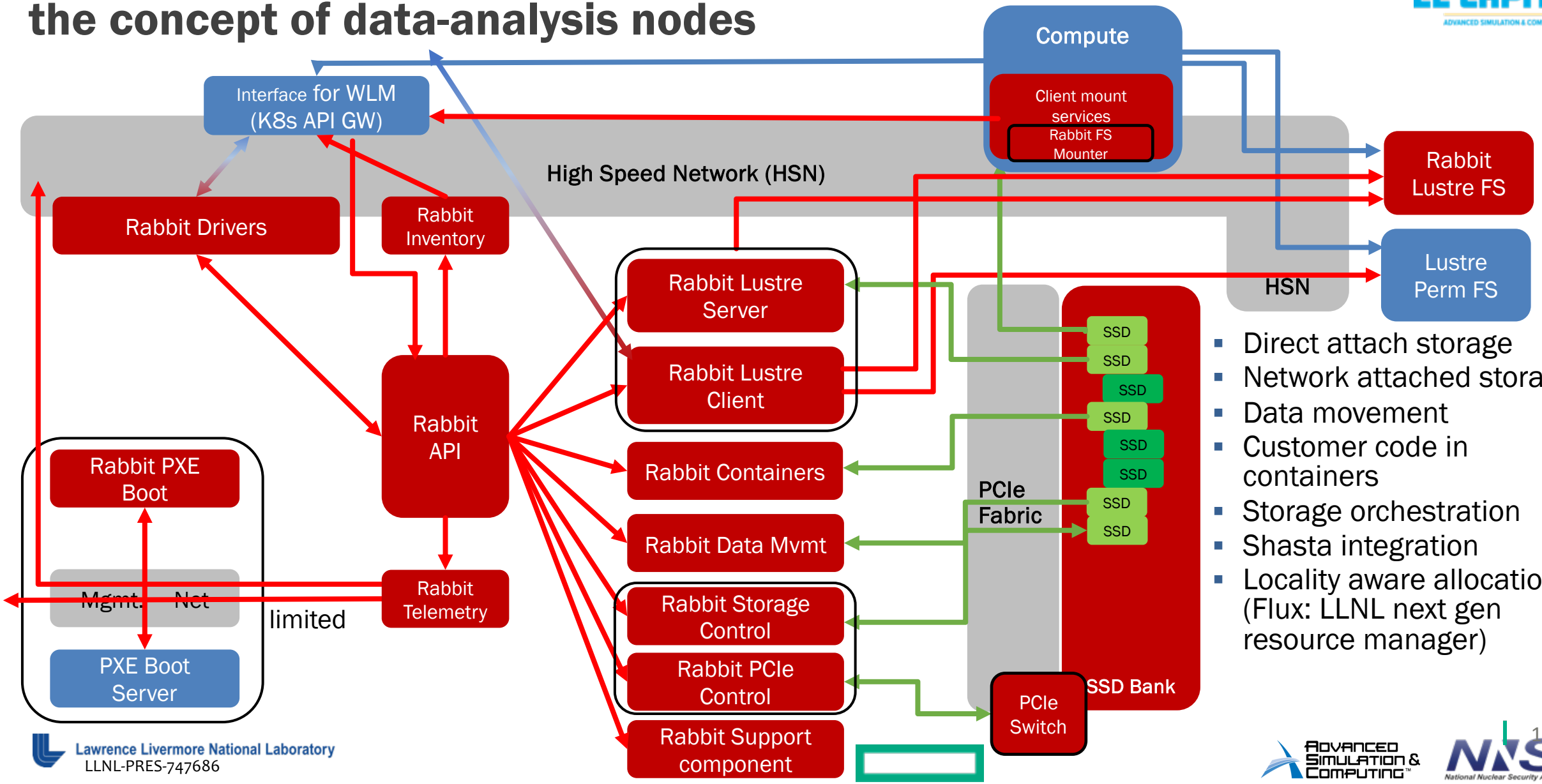
Near-node local storage was a key aspect in El Capitan selection

- *El Capitan* (and it's unclassified PSAAP-available counterpart *Tuolumne*) will deploy one Rabbit module for every compute chassis
- Rabbit modules will:
  - Reduce system interference
  - Enable efficient defensive I/O
  - Likely serve as OS file cache
  - Possibly support more efficient input
    - Particularly for ML training
    - Stage-in of restart files is more complex

- Rabbit modules are one of HPE's essential innovations
  - Many funded under non-recurring engineering (NRE) contract, joint with Oak Ridge National Laboratory
  - Opportunities for other sites to deploy Rabbit modules, extend NRE directions

We will deploy other future heterogeneous system architectures with data analysis nodes

Rabbit software will enable many use cases including the concept of data-analysis nodes



- Direct attach storage
- Network attached storage
- Data movement
- Customer code in containers
- Storage orchestration
- Shasta integration
- Locality aware allocations (Flux: LLNL next gen resource manager)

OK – but why should PSAAP centers care?

- Workflow research and development encompasses much of what the HPC community cares about:
  - Reproducibility
  - Simulation and HPC
  - Artificial Intelligence / Machine learning
  - Large data sources
  - Uncertainty Quantification
  - Software deployment (containers)
  - Productivity
  - Education / knowledge transfer
- Adopt it, adapt it, or invent it if you have new innovations
  - There is continuing *research* needed in scientific workflows
  - Even absent novel research, it's increasingly important to consider as part of your project

Think hard before inventing yet-another-workflow-system from scratch

- There are lots of workflow systems in use across DOE sites
- All may require some effort to get set up and running with, but the cost is lower than writing your own of similar complexity
- Experience indicates that adopting a ‘real’ workflow management system leads to systemizing workflows that makes it easier to port workflows in the future
- Ask your DOE contacts for tools in-use at our labs to try out (a few are noted here)

flux

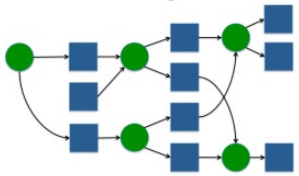
maestro

BEE

Parsl

Merlin

Makeflow



Radical Cybertools

nextflow

snakemake

A non-exhaustive list of potential workflow research topics

- Novel use of storage systems
  - Data lakes, databases, “Rabbits”
- Eliminating or demphasizing I/O
  - In-situ analysis
- Scheduling complex workflows across diverse machines (including cloud resources)
- Using system monitoring data, machine learning, and other data for purpose of optimizing performance or energy
- End-to-end security and encryption
- Flexible data formats for different use-cases (e.g. ML training data generation vs. validation/comparison with experiment)
- Others… ad nauseum

These activities are all being pursued in the NNSA labs, but we don’t have final solutions, and more ideas and research is needed

Some things you can ask me:

Hey (pick your favorite LLM), resubmit that last job on twice as many nodes

Were there any local temperature spikes between 5 and 7 ms of my currently running simulation?

Optimize the function “LUsolve” for the MI300A

Archive all my results tagged “project2029” in my off-site cloud storage account

Launch notebook in /home/neely/prj2029, but change initial density of the steel in region ten to 8 g/cm^3

PSAAP centers can (and should) be incubators for new ideas and technologies, grounded in the broader context of a mission of scientific discovery

Inspire us with your vision
Amaze us with your innovations
Influence us with your solutions
Impact the world with your students