

# Workflows At NERSC



New User Training  
September 28, 2022

Bill Arndt  
Data Science Engagement Group  
warndt@lbl.gov

# Agenda

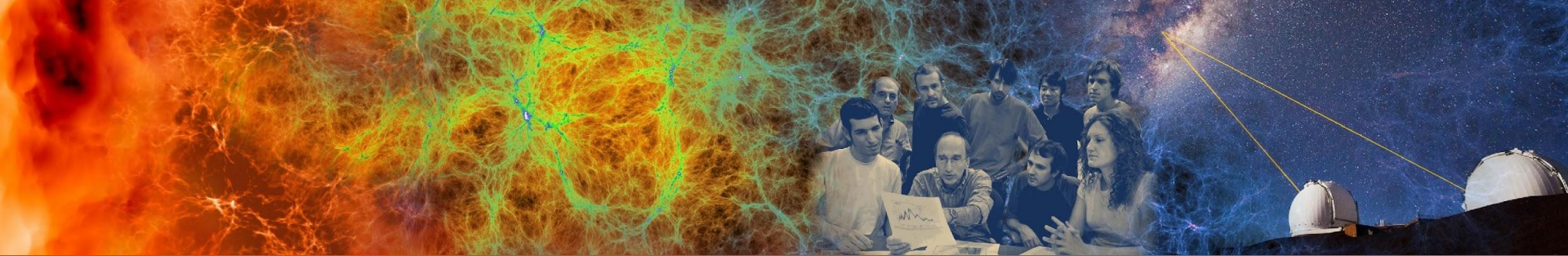
- What are Workflows?
- Workflow Resources at NERSC
- Best Practices

# What Are Workflows?

- A workflow is a problem better solved by inserting automation tools between the user and computational resources.
  - Workflow Management Tools (WMT) are the software systems that perform that automation.
- What can get better?
  - Perform repetitive, urgent, or recovery tasks automatically.
  - Streamline the use of Slurm or transferring data.
  - Organize large amounts of data.
  - Collect and report information about workflow progress.

# Generic Workflow Examples

- “I need to run my application thousands of times.”
- “My data needs several stages of processing with different applications running in an ordered sequence.”
- “My applications each have a 2% chance of crashing and needing to be rerun.”
- “I want my application to run once a month.”



# Workflow Resources at NERSC



BERKELEY LAB



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

# What NERSC Is Doing To Support Workflows

- Specialized infrastructure, software, and support
- Workflows Working Group
  - Formed September 2019 - Laurie Stephey (DAS), Bjoern Enders (DSEG), Bill Arndt (DSEG)
  - Thorough evaluation of many WMT
  - Documentation and guidance refresh
  - Outreach to users, facilities, tool developers, and infrastructure providers

# WMT Documentation And Guidance

- <https://docs.nersc.gov/jobs/workflow-tools/>
  - Detailed information, examples, pitfalls, and suggestions regarding specific tools and use cases
  - Expanding and refining as our tool evaluation continues
- **We *want* to get tickets about workflow management tools**
  - We gain experience and knowledge about what users need
  - We can then use that experience to help other users



# Slurm Crontab (`scrontab`)

- `crontab` is the standard linux solution for running something on a schedule
- On Perlmutter, `crontab` usage provided by `scrontab`
- Interface is same as `crontab`, but it isn't tied to a single machine; `scrontab` anywhere sees the same list and losing a node does not prevent them from running
- It is provided by slurm; there are `#SCRON` directives to communicate requested wall time, account to link to usage, etc.
- Current request limits are 2 cores and 24 hours walltime



# scrontab and workflows

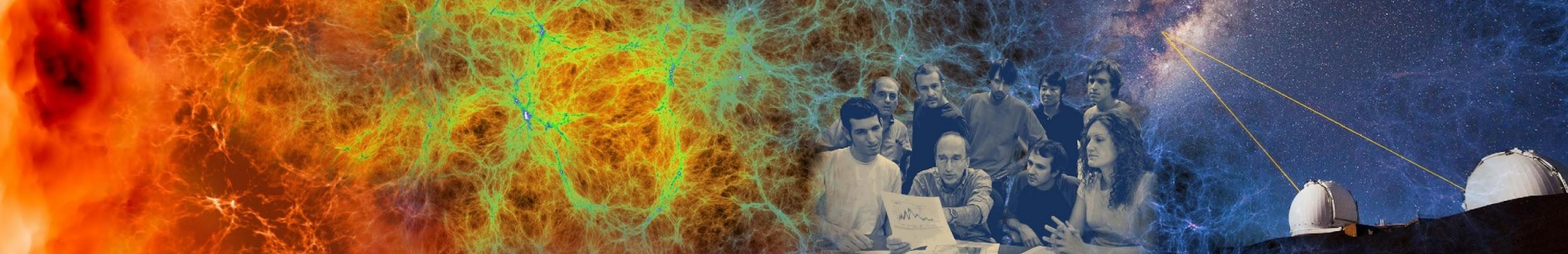
- A workflow QOS is being added for scrontab jobs
- Much longer maximum wall time
  - At least a month but policy may allow longer
- Up to 32 cores available
- Access requires filling out a service now form at [help.nersc.gov](https://help.nersc.gov)

# SPIN

- SPIN is a platform for hosting services related to use of NERSC resources and data hosted by NERSC
- Use SPIN to deploy your own science gateways, workflow managers, databases, API endpoints, and other network services.
- Training sessions and specialized support are available if you decide to set up a service in SPIN.
  - The next SPIN training session will be on October 5th

# Superfacility API

- NERSC provides an application programming interface (API) that allows resources outside of Perlmutter to make requests and issue commands there
- A service running in SPIN (or completely outside of NERSC) can submit HPC jobs by using this interface



# Best Practices



BERKELEY LAB



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

# GNU Parallel Is Great for Many Small Tasks

```
elvis@cori07:~> seq 1 5 | parallel -j 2 'echo \  
> "Hello world {}!"; sleep 10; date'  
Hello world 1!  
Thu Jun 11 00:21:00 PDT 2020  
Hello world 2!  
Thu Jun 11 00:21:00 PDT 2020  
Hello world 3!  
Thu Jun 11 00:21:10 PDT 2020  
Hello world 4!  
Thu Jun 11 00:21:10 PDT 2020  
Hello world 5!  
Thu Jun 11 00:21:20 PDT 2020  
elvis@cori07:~>
```

- Packed jobs have massively reduced total queue wait
  - Can also pack single-node tasks into multiple node jobs
- No risk of Slurm overload
- Run combinations of tasks in parallel and sequence
- Easy input substitution
  - If you need it, *much* more power is available
- Superior to task arrays, too
- See documentation

# Data Centric Workflow Management Tools

- “I have many different applications and data types chained together in a network of dependencies.”
- *Plenty* of options. Snakemake and Parsl are two examples
  - See documentation
- Pitfalls:
  - Many tools assume cloud availability and can't cope well with queue waiting
  - Often lack job packing
  - Naive Slurm integration can use too many requests
  - Risks with networked filesystems

Thank You and  
Welcome to  
NERSC!

