



# **OLCF Best Practices & New User Overview**

Bill Renaud



ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

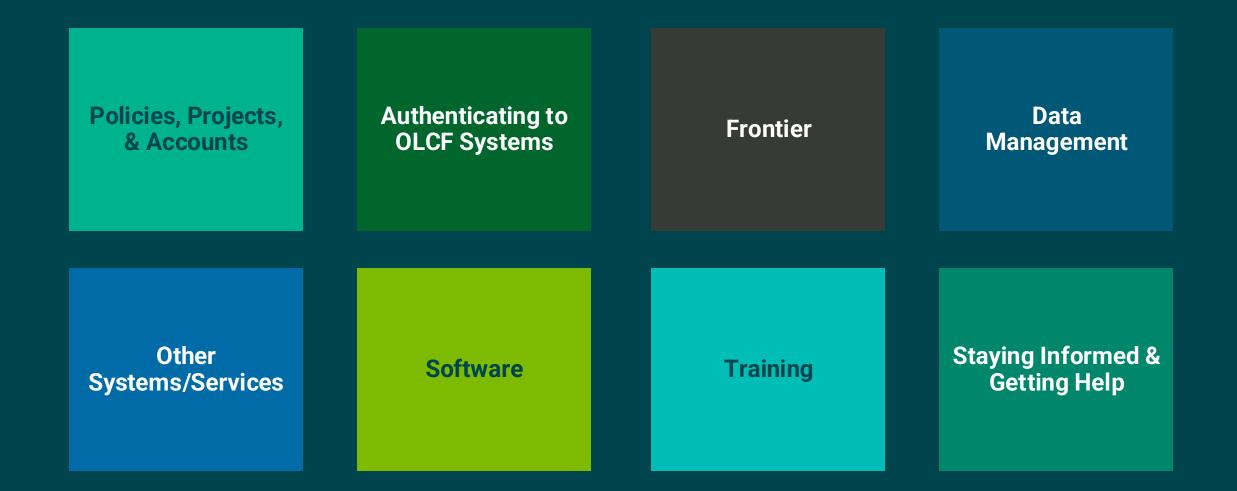


# **General Information**

- This presentation gives a brief introduction to using OLCF resources, covering systems, policies, and more
- This is not all-inclusive, so feel free to ask questions
- Primary focus is on our moderate systems (Frontier, Andes, Orion, etc.); we'll mention the open enclave
- A lot of this may be a reminder, and it's largely based on previous presentations https://www.olcf.ornl.gov/calendar/userconcall-[jun2019|feb2020|feb2021|feb2022]/ https://www.olcf.ornl.gov/calendar/2020-olcf-user-meeting/ https://www.olcf.ornl.gov/calendar/summit-new-user-training/ https://www.olcf.ornl.gov/calendar/frontier-training-workshop-august-2023/
- Aggregation of info from docs.olcf.ornl.gov...bookmark that site!



#### **Overview**





## **Policies, Projects, & Accounts**



### Policies, Projects, & Accounts Best Practices

- 1) Know where to find basic policies & the OLCF Acknowledgement
- 2) Understand how system/group access works at OLCF
- 3) Understand allocation policies
- 4) Be prepared for the end of the project



### **OLCF Policies**

- Users must agree to various policies in order to have an account and/or project

Computing Policy	Data Management Policy
Security Policy	Project Policies (Reporting, Allocation Utilization, User Agreement, etc.)

- You can review these policies at <u>https://docs.olcf.ornl.gov/accounts/olcf\_policy\_guide.html</u>
- Our Acknowledgement Statement is on that page, too

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-000R22725.



## **System and Group Access**

#### Projects

- Are granted access to systems
- Are associated with Unix groups

#### Users

- Are assigned to projects
- Inherit the project's system accesses & groups

So to join a group or obtain access to a system, you join a project that has access to that group/system. This also means you lose access to groups/systems when you leave that project (or it ends) unless you have them via another project.



## **Allocation Policies & Project Closeout**

#### Allocations

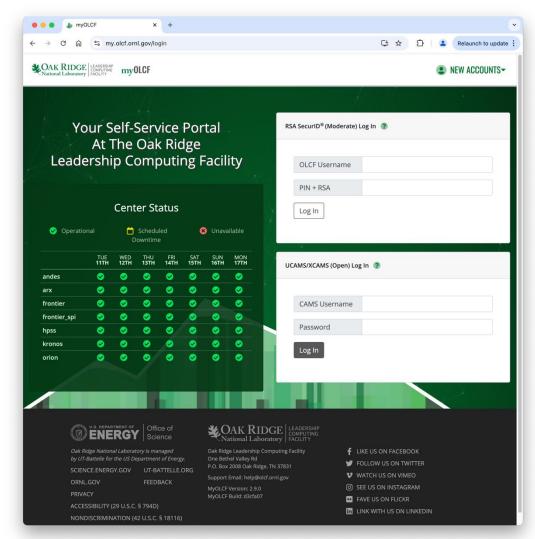
- Projects *are not* disabled for exceeding their allocation
- Priority reduction and/or job limits are used to prioritize projects with hours remaining while keeping the system busy if there's work to do
- We don't issue "refunds" but can delay priority reduction if appropriate (which has the same effect)

#### **Project Closeout**

- When projects end you'll lose access to systems/groups granted by that project (unless you have them from another project)
- Projects have a 1-month "data-only" mode to retrieve important data
  - No HPC access...only DTNs/Globus
  - Don't wait to transfer data! (More later...)



# Managing Accounts & Projects with myOLCF



- Apply for/view/renew projects
- Apply for/renew user accounts
- Get usage information
- Update contact information
- Sign up for RSA verifications
- Sign up for Office Hours
- NEW! Interact with tickets

#### https://my.olcf.ornl.gov

Documentation: https://docs.olcf.ornl.gov/services\_and\_applications/myolcf/index.html



## Authenticating to OLCF Systems



## Authentication Best Practices

- 1) Be familiar with different authentication systems
- 2) Know how to handle lost/forgotten credentials
- 3) Know common SSH failures



# **Authenticating to OLCF Systems**

- Interactive access is via Secure Shell (SSH)
- For the moderate enclave
  - Systems use 2-factor authentication via user-selected PINs and RSA SecurID tokens
  - Password, publickey, etc. are not permitted
  - RSA provides one-time credentials...you can't use the same PASSCODE twice in succession
- For the open enclave
  - Systems use password authentication (via XCAMS password)



### Authenticating to OLCF Systems: Lost or Forgotten credentials

- For Moderate systems (i.e. RSA token issues), contact <u>help@olcf.ornl.gov</u>
- For Open systems, a self-serve password reset is available <u>https://xcams.ornl.gov</u>



## Common SSH issues: Wrong Username

- Normally, SSH neither prompts for a username nor tells you what it's using
  - Some clients such as PuTTY do
  - It'll also be displayed if you request verbose output
- Several ways to fix
  - Specify on command line (ssh user@host.domain)
  - Specify in ~/.ssh/config
  - Specify via graphical client (each client will be a little different)



### **Common SSH issues: Token Out of Sync**

Enter PASSCODE:

Wait for the tokencode to change, then enter the new tokencode :

- You may get prompted for the next tokencode
  - A tokencode is the 6-digit number from your SecurID token
  - A PASSCODE is your PIN followed by the current tokencode
- If so, enter *only* the next tokencode (not a full PASSCODE) here
- Tip: Web clients (myOLCF) do not support next tokencode mode, so if you've had a few failures there try connecting somewhere via ssh to resync, then try myOLCF again
- Tip: Even if you're not in an "active processing" phase of your project, log in every couple of weeks to keep everything in sync



### Common SSH Issues: Password Prompt

Enter PASSCODE: Enter PASSOCDE: Enter PASSCODE: userl@frontier.olcf.ornl.gov's password:

- Sometimes SSH will prompt for a password after (three) PASCODE failures
- This is fallback behavior of SSH...you do not have a password so there's nothing to enter here
- If this happens, type Ctrl+C and try again
- Tip: If you've failed twice, let the tokencode change and try w/the new one



## Frontier



## **Frontier Best Practices**

- 1) Know basic system design
- 2) Understand the build environment
- 3) Know Slurm features/settings

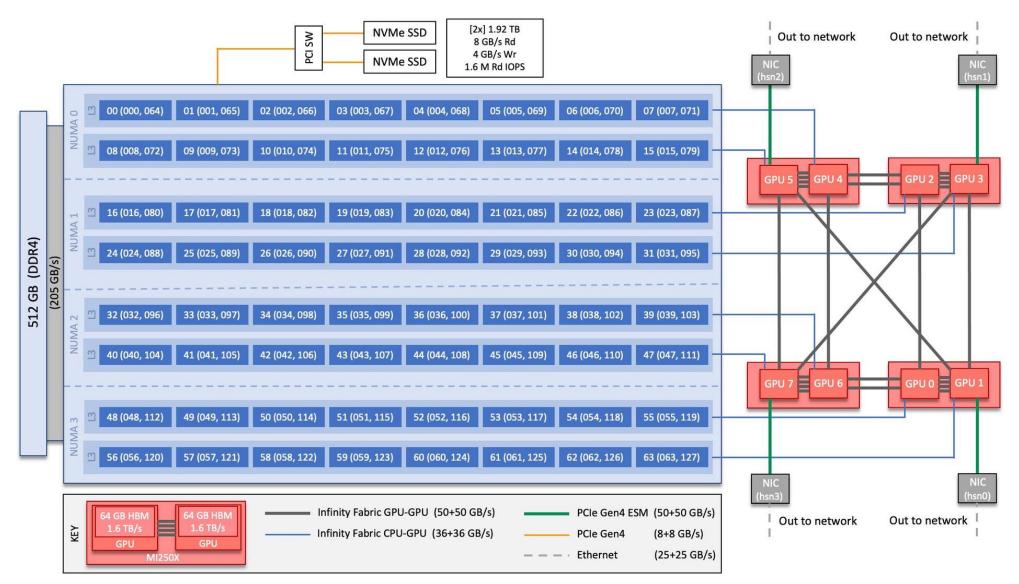


### **Frontier**

- Hybrid system using AMD EPYC processors and MI250X GPUs
- Slurm scheduler/job launcher
- Similar to a commodity cluster (especially as compared to older Cray systems) but with some important differences
- NFS areas are read-only on compute nodes



## **Frontier Node Layout**





# **Building on Frontier**

- Like our older Cray systems, Frontier uses compiler wrappers
  - C:cc, C++:CC, Fortran: ftn
  - Automatically link in MPI (and many other libraries, typically from cray-\* modules)
  - Can cause headaches for some build environments (cmake, autoconf)
- Build environment is controlled by a number of related modules
  - Be aware of compatibility issues between versions (Check the website, release notes, etc.)
- Various compile options (MPI, OpenMP Offloading, HIP, GPU-Aware MPI, etc.) are discussed in the Frontier User Guide



# **Building on Frontier**

#### - Controlled by a combination of modulefiles

cpe	Loads specific release of the Cray Programming Environment*	
PrgEnv-?	Provides a specific compiler toolchain*	
craype	Provides compiler drivers	
amd/gcc/cce	Provides the back-end compiler	
cray-mpich	Provides MPI libraries	
rocm	Provides the ROCm toolchain	
craype-accel-amd-gfx90a	Tells compiler drivers to target MI250x	

\*Both of these are "meta-modules" that load/swap a number of other modules



# **Running on Frontier**

- Traditional batch scripts (sbatch) and interactive jobs (salloc) are supported
- Whole nodes are allocated, so Slurm requests in the initial allocation can be simpler
  - Ask for nodes, time, etc.
  - Use srun to handle layout
- There are some important options for certain features, though
  - Mustuse -- threads-per-core=2 if you'll use that in srun
  - Must use -C nvme to enable NVMe use
- Consider using sbcast for conda environments or large numbers of .so files
  - Per-node loads from NVMe can be much faster than many-node loads from Lustre
- Note that by default 8 cores (1 per L3 domain) are reserved for OS
  - Overriding this is possible, but not recommended



# Data Management



### Data Management Best Practices

- 1) Know the different storage areas
- 2) Understand Lustre basics
- 3) Develop a data policy
- 4) Know data transfer options



#### **Storage Areas**

#### By Who Can (Potentially) Access

By Intended Use

- User
- All members of the project
- All users on the system

- Semi-permanent/home directory
- Short-term/scratch
- Long-term/archive

#### By Storage Technology

- NFS
- Lustre
- GPFS
- Nearline (GPFS+Tape)

#### Everything you ever wanted to know about storage @ OLCF: https://docs.olcf.ornl.gov/data/index.html



### **Data Storage Areas**

Storage Area	Technology	Purpose of Area
User home	NFS	Frequently accessed user data
Project home	NFS	Frequently accessed project data
User work	Lustre (Moderate)/GPFS (Open)	User's scratch files <sup>1</sup>
Project work	Lustre (Moderate)/GPFS (Open)	Project-shared scratch files
Global work	Lustre (Moderate)/GPFS (Open)	"World"-shared scratch files
User archive	Nearline (GPFS+tape)	User's archive files <sup>1</sup>
Project archive	Nearline (GPFS+tape)	Project-shared archive files
Global archive	Nearline (GPFS+tape)	"World"-shared archive files

<sup>1</sup> Even though these are called "user" areas, they're still considered project data



## **Moderate Data Storage Locations**

Storage Area	Path <sup>1</sup>	Default Permissions	Change Permissions?	Backed Up?
User home	/ccs/home/\$USER	0700	Yes	Yes
Project home	/ccs/proj/PROJID	0770	No	Yes
User work <sup>2</sup>	/lustre/orion/PROJID/scratch/\$USER	0700		No
Project work <sup>2</sup>	/lustre/orion/PROJID/proj-shared	2770	No	No
Global work <sup>2</sup>	/lustre/orion/PROJID/world-shared	2775	No	No
User archive	/nl/kronos/olcf/PROJID/users/\$USER	0700		No
Project archive	/nl/kronos/olcf/PROJID/proj-shared	2770	No	No
Global archive	/nl/kronos/olcf/PROJID/world-shared	2775	No	No

<sup>1</sup> These are recommended paths, not necessarily absolute paths <sup>2</sup> Alternatively, \$MEMBERWORK/PROJ, \$PROJWORK/PROJ, \$WORLDWORK/PROJ



## **Open Data Storage Locations**

Storage Area	Path <sup>1</sup>	Default Permissions	Change Permissions?	Backed Up?
User home	/ccsopen/home/\$USER	0700	Yes	Yes
Project home	/ccsopen/proj/PROJID	0770	No	Yes
User work	/gpfs/wolf2/olcf/PROJID/scratch/\$USER	0700		No
Project work	/gpfs/wolf2/olcf/PROJID/proj-shared	2770	No	No
Global work	/gpfs/wolf2/olcf/PROJID/world-shared	2775	No	No

<sup>1</sup> These are recommended paths, not necessarily absolute paths



## **Using Lustre**

- Lustre is designed for handling large, parallel files
- Files are striped across multiple *Object Storage Targets* (OSTs)
- May be more complicated than previous Lustre systems you've used
- Default striping tries to optimize striping/tier based on file size
  - For files < 512GB, let the system handle striping
  - For files larger than 512GB, explicit striping may help (see docs site for suggestions)
  - Don't stripe across > 450 OSTs
  - Special striping considerations
    - Using lfs setstripe -c 0 will use OLCF-default striping
    - Using lfs setstripe -c -1 will use max allowed stripe (450)



# Farewell, HPSS. Hello, Kronos!

- After several decades of service, it is time for HPSS to enjoy its retirement
- It has been replaced by our new nearline resource, Kronos
  - Multi-level (disk + tape) storage system
  - Unlike HPSS, Kronos is directly mounted (on DTNs only)
  - Standard commands can be used to move data back and forth
  - Available via the OLCF Kronos Globus Collection/Endpoint





#### NFS areas are backed up to a limited extent via the .shapshot subdirectory

\$ ls -al /ccs/proj/abc123/important\_data |grep .snapshot

\$ ls /ccs/proj/abc123/i	mportant_data/.snapshot		
daily.2022-02-10_0010	hourly.2022-02-15_1605	hourly.2022-02-16_0005	hourly.2022-02-16_0805
daily.2022-02-11_0010	hourly.2022-02-15_1705	hourly.2022-02-16_0105	hourly.2022-02-16_0905
daily.2022-02-12_0010	hourly.2022-02-15_1805	hourly.2022-02-16_0205	hourly.2022-02-16_1005
daily.2022-02-13_0010	hourly.2022-02-15_1905	hourly.2022-02-16_0305	hourly.2022-02-16_1105
daily.2022-02-14_0010	hourly.2022-02-15_2005	hourly.2022-02-16_0405	hourly.2022-02-16_1205
daily.2022-02-15_0010	hourly.2022-02-15_2105	hourly.2022-02-16_0505	hourly.2022-02-16_1305
daily.2022-02-16_0010	hourly.2022-02-15_2205	hourly.2022-02-16_0605	hourly.2022-02-16_1405
hourly.2022-02-15_1505	hourly.2022-02-15_2305	hourly.2022-02-16_0705	weekly.2022-02-13_0015

- Lustre and GPFS are not backed up (AND are subject to purge)
- Kronos can be a backup for NFS/Lustre, but it is not backed up itself
- ALWAYS have an off-site backup for important data, and archive data as soon as possible after creation



# **Developing a Data Plan**

- It's easy to generate lots of data very quickly, so it is crucial to develop a project data plan early (It's easier to fix 100 files than 100 million)
- Considerations
  - What files are stored where (user areas, project areas, directory naming/structure, etc.)
  - File/directory attributes (permissions, umask settings, group, etc.) (How can you make things seamless when users leave the project?)
  - What happens when the project ends?
  - Where should data be copied, and who is doing that? (It will take longer than you think...start early!)
- Once you develop your plan, do spot-checks during the year
- Note that proj-shared & world-shared areas *facilitate* sharing but don't *automatically* share files stored there
  - You still need to ensure appropriate permissions/groups/etc.
  - Think of directories as the front door to the building...you still need to make sure people have access to the right rooms/offices.

## Who owns the data?

- Data in the user home directory is generally considered the user's data
  - Normally the user will need to approve someone else getting access
  - No project data should go here
- Data in other directories (/ccs/proj, /lustre and /nl) belongs to the project
  - PI approval is sufficient to change permissions, move data, etc.
  - Yes, even user scratch and user archive areas belong to the project!



# **Transferring Data**

- <u>https://docs.olcf.ornl.gov/data/index.html#transferring-data</u>
- Globus Collections/Endpoints will likely give you the fastest, most convenient experience
  - OLCF DTN (Globus 5) provides access to NFS and Lustre
  - OLCF Kronos provides access to Kronos/nearline
  - **NCCS Open DTN** provides access to Open storage areas
- Data transfer nodes (dtn.ccs.ornl.gov / opendtn.ccs.ornl.gov) provide command-line access
  - Internal transfers (e.g. to/from nearline) via rsync/cp
  - External transfers via sftp/scp (If Globus is an option, it's likely to be *much* faster)
- Transfer important data early and often



# **Monitoring Storage Usage**

- For user home directories, use the quota command
- For project home areas, cd into the directory and use the df . command (Each /ccs/proj area behaves like it is a distinct filesystem in this regard)
- We do not currently have good tools for Lustre and Archive areas
  - And you likely can't see everything due to permissions, so du isn't ideal
  - If you need to know, reach out to <u>help@olcf.ornl.gov</u>



# **Other Systems/Services**



#### Support System Best Practices

1) Know what additional systems and services are available to your project



# **Support Systems**

- Data Transfer Nodes (DTNs)
  - Moderate: dtn.ccs.ornl.gov
  - **Open**: opendtn.ccs.ornl.gov
- Other systems
  - Moderate data analysis system: andes.olcf.ornl.gov



# **Other Services**

#### - Slate

- Kubernetes/Openshift cluster
- See <a href="https://docs.olcf.ornl.gov/services\_and\_applications/slate/getting\_started.html">https://docs.olcf.ornl.gov/services\_and\_applications/slate/getting\_started.html</a> for instructions on requesting and using
- Both Moderate (<u>https://marble.ccs.ornl.gov</u>) and Open (<u>https://onyx.ccs.ornl.gov</u>) instances available
- Jupyter
  - All projects should have access
  - Moderate (<u>https://jupyter.olcf.ornl.gov</u>) and Open (<u>https://jupyter-open.olcf.ornl.gov</u>) instances available



# Software



# Software Best Practices

- 1) Know where to find software
- 2) Understand considerations for building software



# **Finding and Using Software**

- Basic commands are part of the default environment
- Many other tools are provided via Lmod (<u>https://lmod.readthedocs.io</u>)
- Compilers, debugging/optimization tools, math/scientific libraries, some scientific apps provided
  - If you think we should add something, request it via <u>help@olcf.ornl.gov</u>
- Special case: Python
  - We provide python+miniforge module that provides basic packages
  - Recommend that + conda/venv to build specific modules you need
  - <u>https://docs.olcf.ornl.gov/software/python/index.html</u>



# **Building Software**

- Projects list the software they'll use as part of the project application
  - A review may be needed prior to adding software that wasn't listed, and you must comply with relevant requirements (licensing, export control, etc.)
  - Adding software may change your project's classification
- Consider building in either NFS or ramdisk and installing in NFS
  - Avoids the scratch purge
  - These areas typically provide better performance for compiling than lustre does
  - Be sure to set appropriate permissions
- Be careful when deploying AI tools...they may require careful review



# Training



# Training Best Practices

- 1) Know what training options are available
- 2) Find upcoming and historical training

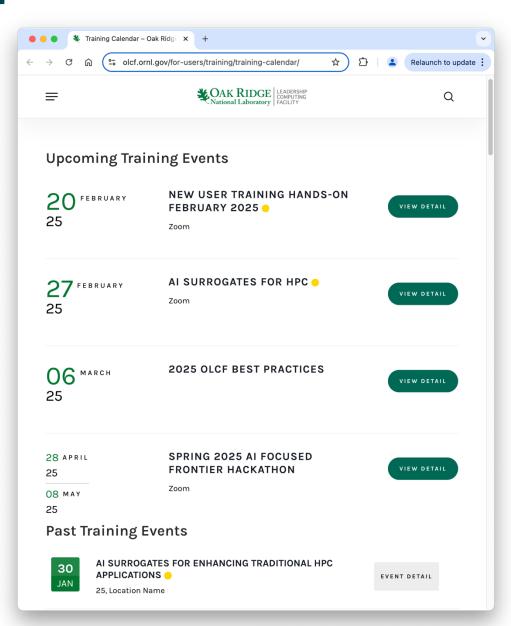


# **Training Opportunities**

- OLCF supports multiple training opportunities throughout the year
- Mix of virtual, hybrid, and in-person
- Hackathons: <u>https://www.olcf.ornl.gov/frontier-hackathons</u>
- Have an idea for training? Reach out: <u>help@olcf.ornl.gov</u>
- How to find upcoming events
  - Weekly Update
  - OLCF homepage
  - Training calendar



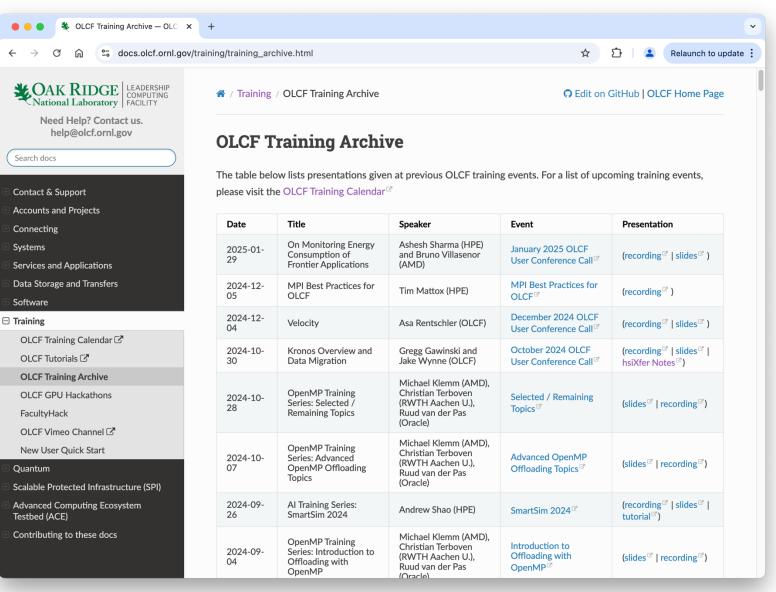
# **Training Calendar**





https://www.olcf.ornl.gov/for-users/training/training-calendar/

# **Training Archive**





#### https://docs.olcf.ornl.gov/training/training\_archive.html

# **Staying Informed & Getting Help**



### **Staying Informed & Getting Help Best Practices**

- 1) Know how we communicate with you
- 2) Working efficiently with tickets
- 3) Use Office Hours when necessary
- 4) Find documentation



# **Staying Informed: Email Lists**

- Announce Lists
  - Major announcements, Weekly Update, etc.
  - All users are required to be members
  - Emails come from <u>olcf@olcf-communications.org</u> Yes, that address looks strange, but it is indeed us!
- Notice Lists
  - Typically, 'recent' users
  - Permanent opt-in/out-out possible
  - Less critical updates, status information, etc.



# Staying Informed: Weekly Update

- Typically sent each Wednesday
- Announcements, upcoming downtimes, training opportunities, etc.
- You should be receiving this...if not, check your spam/junk folders and/or reach out to us

#### OLCF Weekly Update February 12, 2025

#### In This Message

- <u>Upcoming Downtimes</u>
  - Moderate Center-wide Outage (February 18th)
  - Open Center-wide Outage (February 18th)
- <u>Center Announcements</u>
  - Frontier Software Upgrades, Rebuild Recommended (February 18th)
  - HPSS Decommissioned on January 31st, 2025
- Meetings & Workshops
  - Frontier Al Hackathon Call For Participation
  - OLCF Hands-On New User Training
  - Al Surrogates for HPC
  - OLCF Best Practices
- OLCF Highlights
  - ORNL Staffers Named 'Vanguards of HPC-AI'
- <u>Community Events</u>
  - HPSF Conference Registration Open

#### **Upcoming Downtimes**

Frontier, Andes, Home, Moderate-DTNs, Orion, HPSS, Kronos, Frontier-SPI, SPI-DTNs will be unavailable from 6:00 AM Tuesday, February 18th, until 8:00am on Wednesday, February 19th.



### **Staying Informed: Website/Social Media**

- Check the OLCF Website for announcements about meetings as well as articles about the science done on OLCF Systems
- We maintain a presence on multiple social media platforms, too (announcements, news stories, etc.)
  - <u>https://www.facebook.com/oakridgeleadershipcomputingfacility</u>
  - <u>https://x.com/OLCFGOV</u>
  - <u>https://www.linkedin.com/showcase/computing-at-ornl</u>
  - https://vimeo.com/olcf
  - <u>https://www.instagram.com/olcfgov</u>
  - <u>https://www.flickr.com/photos/olcf</u>



# **How To Get Help**

- Tickets
- Office Hours
- Website/Online documentation



# **Submitting Tickets**

- Send as many errors as possible (better yet, capture them in a file & direct us to it)
- It's not necessary to send reproducers (tell us where they are, set up test directory/.tar file, etc.)
- Report new issues in new tickets rather than replying to old ticket
- Remember to tell us what system you're asking about (we support several)



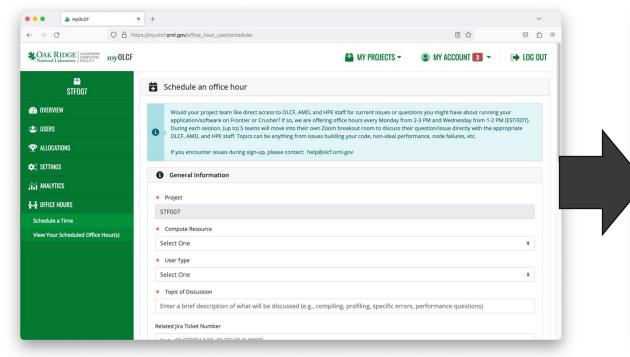
# **Requesting Policy Exceptions**

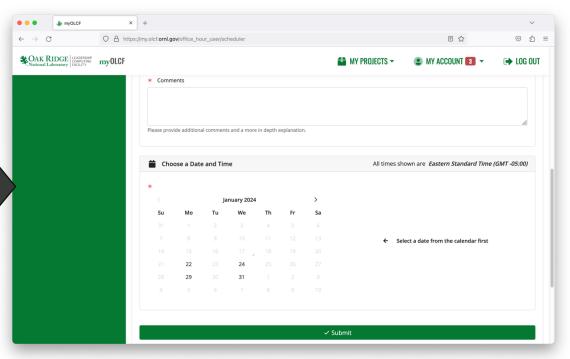
- The Resource Utilization Council (RUC) accepts requests for temporary exemptions from some policies (job limits, quotas, purge exemptions, priority, etc.)
- Requests should be sent to <u>help@olcf.ornl.gov</u>
- Submit requests well in advance to allow for time to review
- Be sure to explain why existing limits are an issue
- If requesting a priority boost, be sure the job is already submitted! (The queue will sometimes surprise you)



### **Office Hours**

- Offered Monday 2-3PM and Wednesday 1-2PM (Eastern)
- Zoom meeting w/users, center staff, and vendors
- Sign up via myOLCF







#### **Documentation**

- OLCF Documentation Site: <u>https://docs.olcf.ornl.gov</u>
  - User Guides
  - Slate/Jupyter documentation
  - Training archive
- HPE/Cray Docs: <u>https://cpe.ext.hpe.com/docs/latest/index.html</u>
- ROCm Docs: <u>https://rocm.docs.amd.com</u>









We're here to help you!

Questions/comments/etc. can be sent to OLCF User Assistance Center

- Staffed Mon-Fri (exclusive of ORNL holidays) from 9AM-5PM US Eastern Time
- Ticket web form (please don't send formatted text)
- <u>help@olcf.ornl.gov</u> (often the best option)

Computer operations staff available 24x7 for login/RSA issues via phone: (865) 241-6536







ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY