# New User Seminar

Research Computing
UNIVERSITY OF COLORADO BOULDER



Be Boulder.

**Website:** www.rc.colorado.edu

**Documentation:** https://curc.readthedocs.io

**Helpdesk:** rc-help@colorado.edu

# Meet the User Support Team

Layla
Freeborn

Brandon
Reyes

Andy
Monaghan

Michael
Schneider

John
Reiland

Dylan
Gottlieb

Mohal
Khandelwal

Ragan
Lee

# RC Resources

## Tech Support

- High Performance Computing
- Data Management
- Cloud Computing
- Secure Research

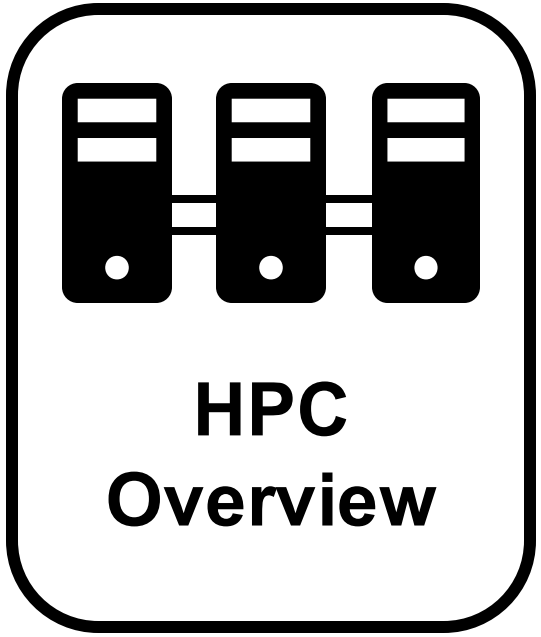
## Human Support

- Training Materials & Workshops
- Consultations & Office Hours
- Help Tickets

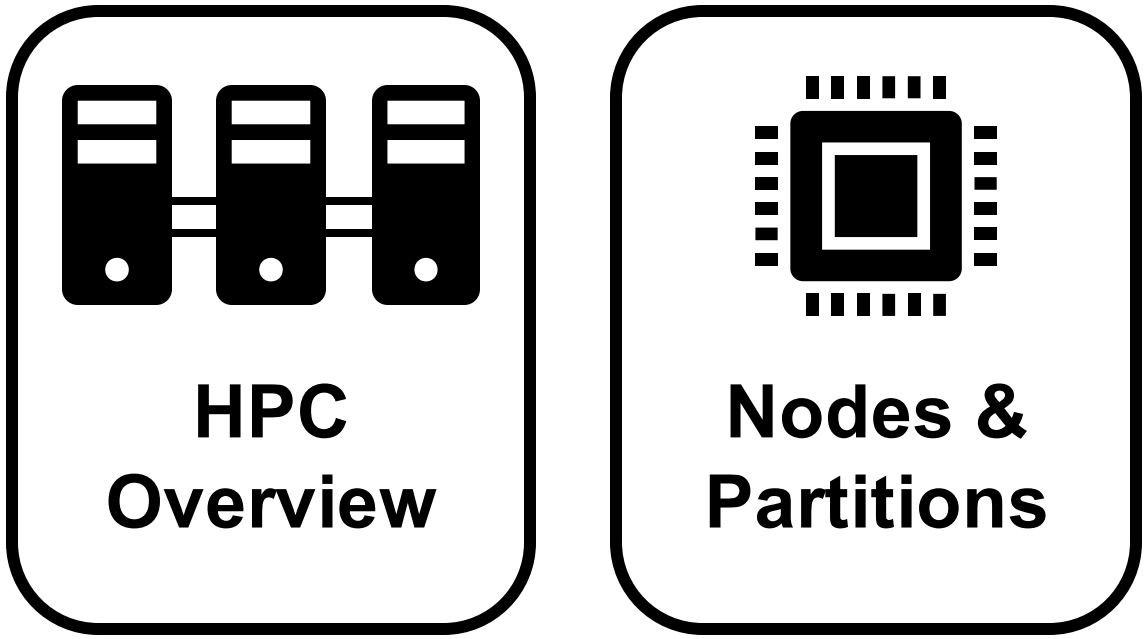
**Ask Questions**

**Discuss Ideas**

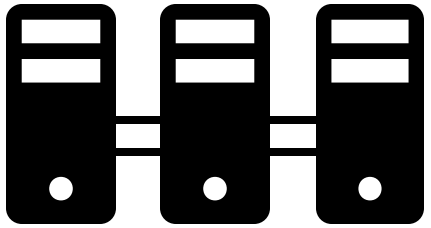
# Learning Goals



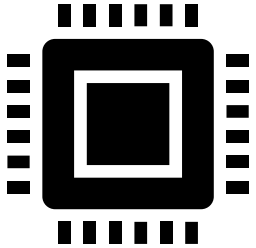
# Learning Goals



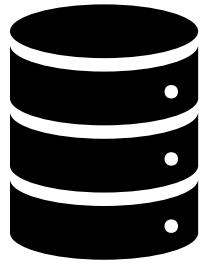
# Learning Goals



**HPC Overview**

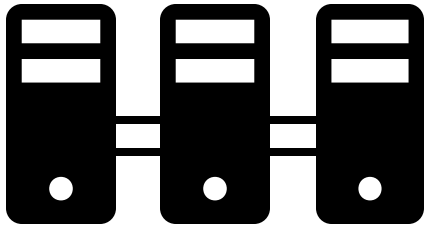


**Nodes & Partitions**

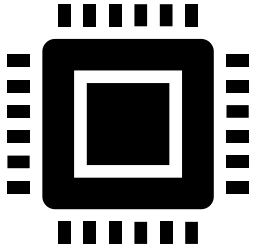


**Data Storage**

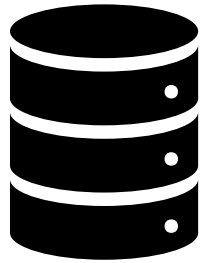
# Learning Goals



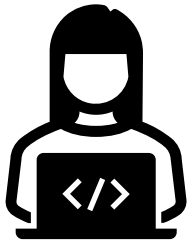
**HPC Overview**



**Nodes & Partitions**



**Data Storage**



**System Access**

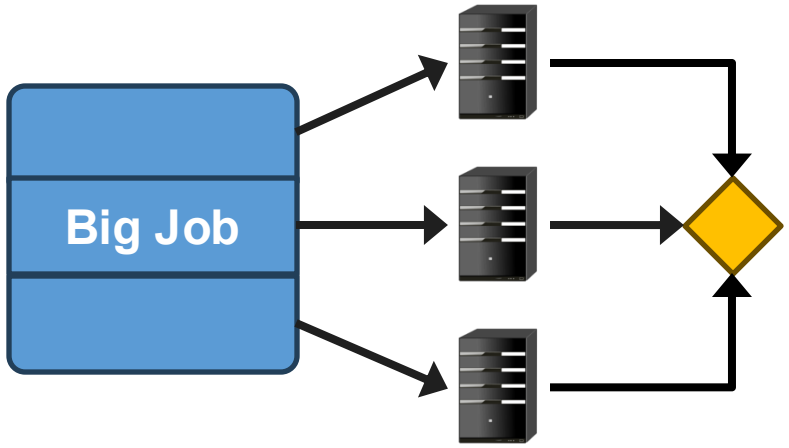
# High Performance Computing

# High Performance Computing

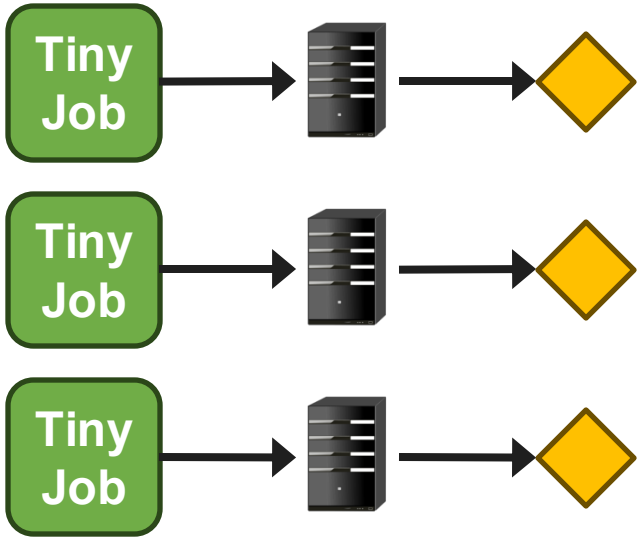
# High Performance Computing

**Scale** **vs** **Speed**

# What can / use HPC for?

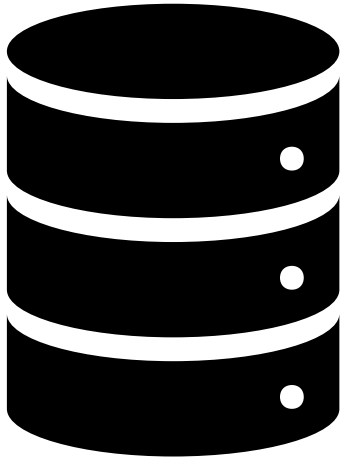



**Parallel Jobs**

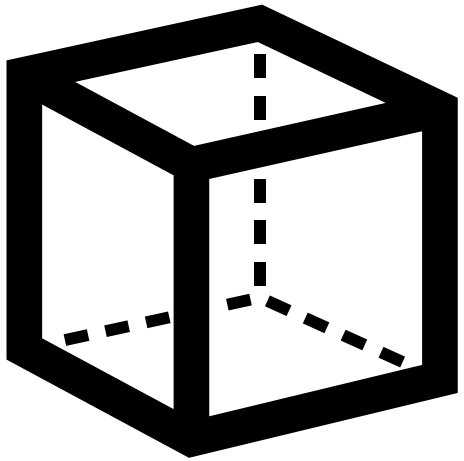



**Serial Jobs**

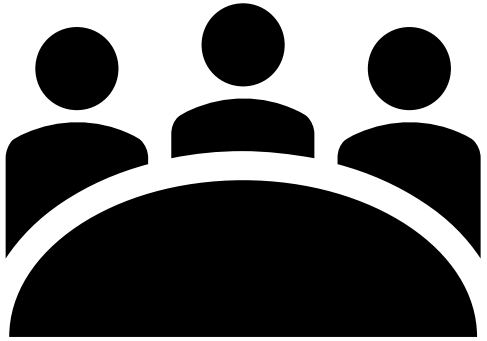
# What can / use HPC for?



**Big Data**



**Viz/Rendering**

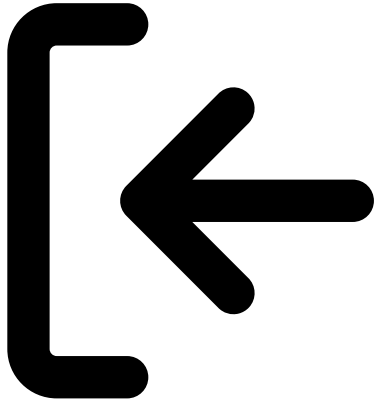


**Shared Work Environment**

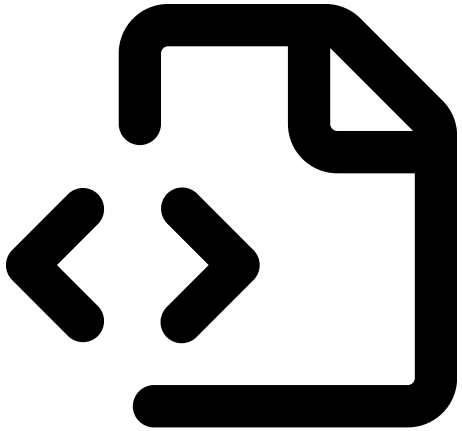
# HPC Cluster: Alpine

- Heterogeneous cluster
- Hardware provided by CU Boulder, CSU, and AMC
- Access available to CU Boulder, CSU, AMC and RMACC users

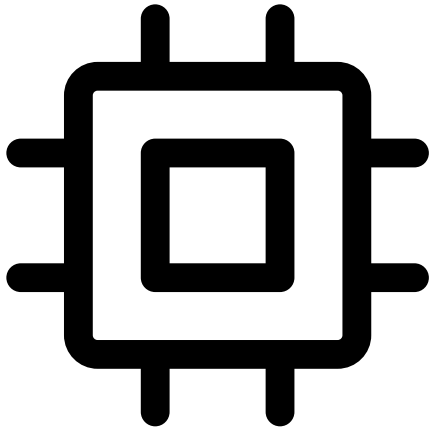
# Node Types



**LOGIN**

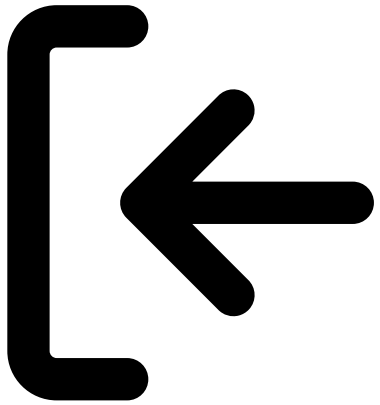


**COMPILE**

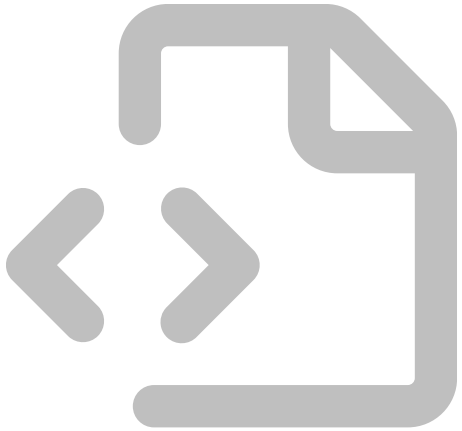


**COMPUTE**

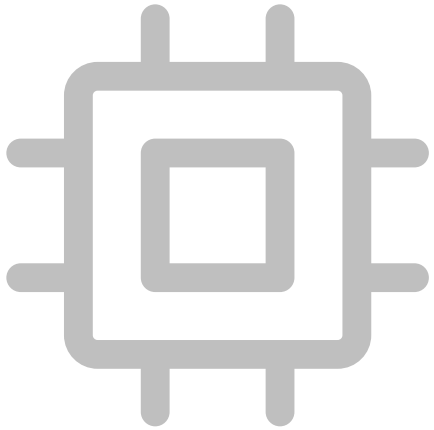
# Node Types



**LOGIN**



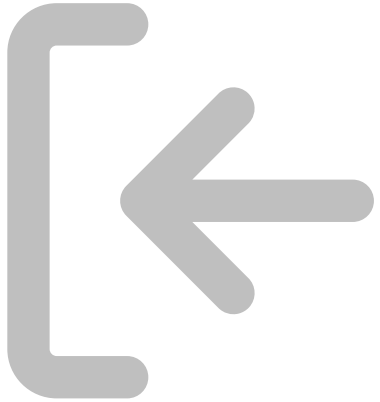
COMPILE



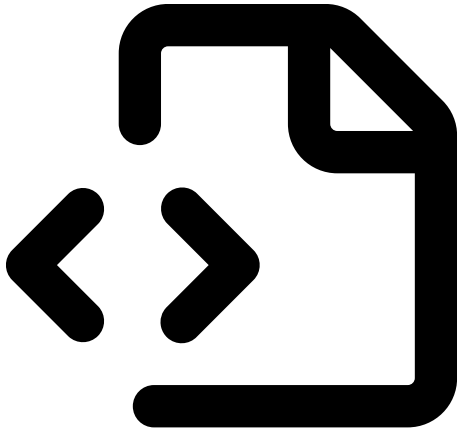
COMPUTE

- Entry to system
- View or edit files
- Submit jobs

# Node Types

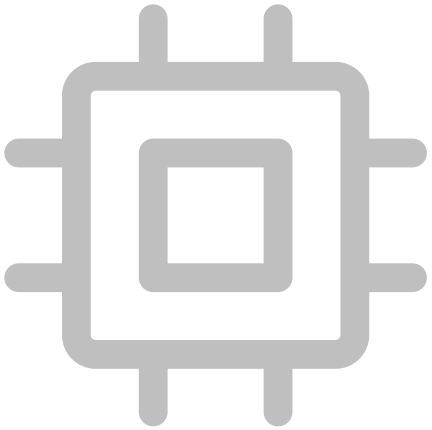


LOGIN



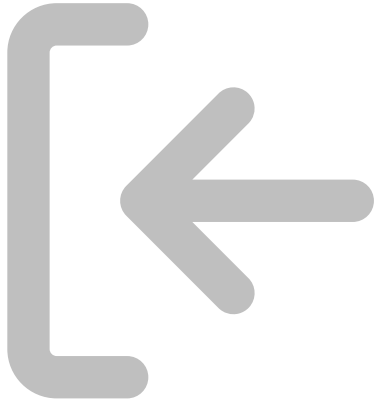
**COMPILE**

- View or edit files
- Submit jobs
- **Compile code**
- **Install software**

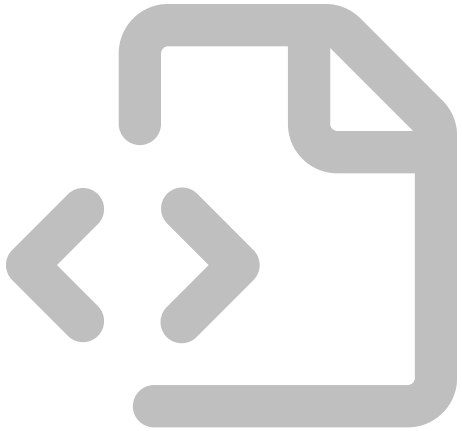


COMPUTE

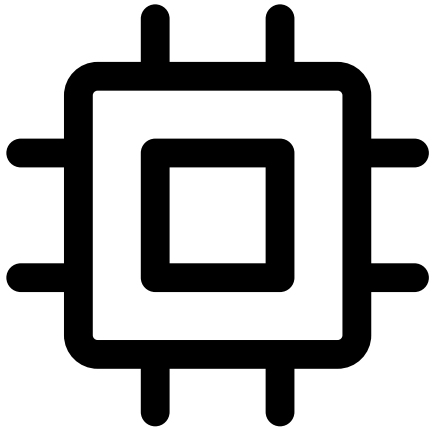
# Node Types



LOGIN



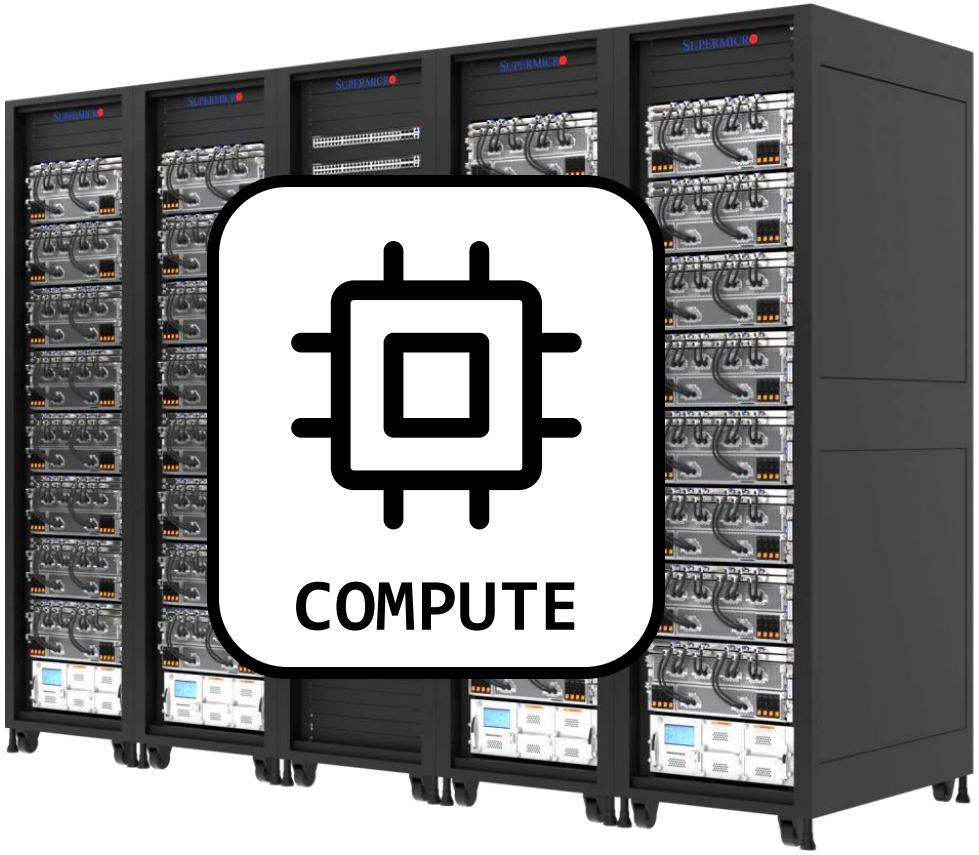
COMPILE



**COMPUTE**

- Run scheduled jobs
- Handle calculations

# Alpine Partitions



# Alpine Partitions

**amilan**

General Usage

# Alpine Partitions

**amilan**
General Usage

**amem**
High Memory

# Alpine Partitions

**amilan**
General Usage

**amem**
High Memory

**aa100**
Nvidia GPU's

# Alpine Partitions

**amilan**
General Usage

**amem**
High Memory

**aa100**
Nvidia GPU's

**ami100**
AMD GPU's

# Data Storage

## Core

- Personal Storage
- Includes 3 Directories
  - /home (2 GB)
  - /projects (250 GB)
  - /scratch (10 TB)

# Data Storage

## Core

- Personal Storage
- Includes 3 Directories
  - /home (2 GB)
  - /projects (250 GB)
  - /scratch (10 TB)

## PL

- PetaLibrary
- Tiered Storage
  - Active, Archive
- Requires Funding
- Starts at 1 TB

# Core

**Home**

- Personal data
- Config files
- OnDemand data
- Do not share!

**2 GB**

# Core

**Home**

- Personal data
- Config files
- OnDemand data
- Do not share!

**2 GB**

**Projects**

- Code, Job scripts
- Installed software
- Shared data

**250 GB**

# Core

**Home**

- Personal data
- Config files
- OnDemand data
- Do not share!

**2 GB**

**Projects**

- Code, Job scripts
- Installed software
- Shared data

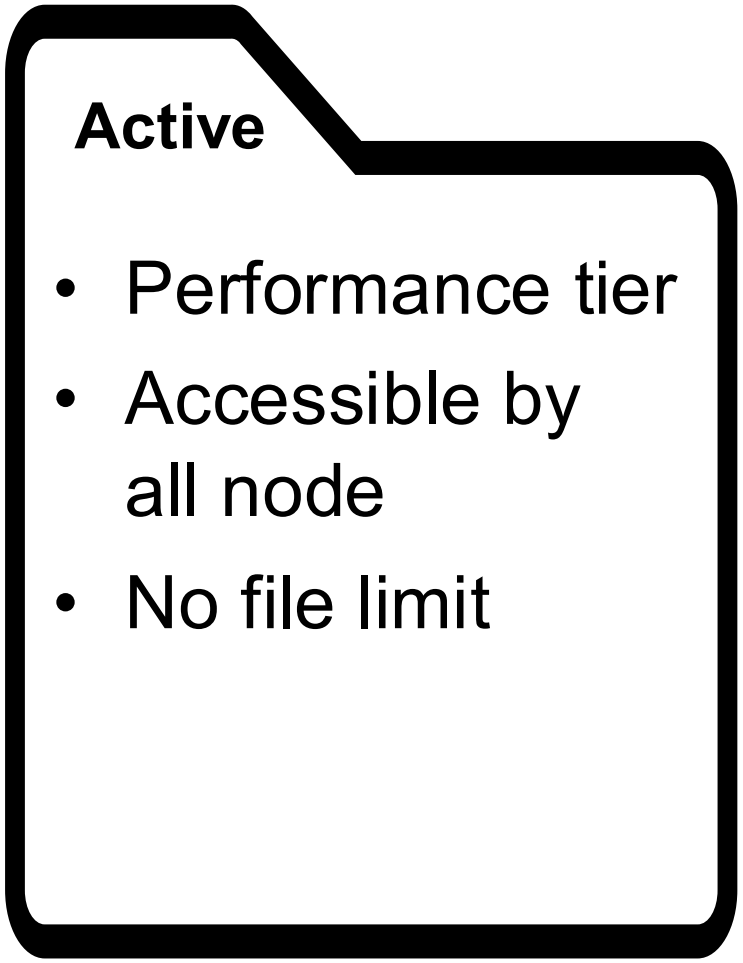
**250 GB**

**Scratch**

- Job data (in/out)
- Shared data
- Deleted after 90 days

**10 TB**

# PetaLibrary



# PetaLibrary

**Active**

- Performance tier
- Accessible by all node
- No file limit

**Archive**

- Integrity tier
- Accessible by login node only
- 10,000 File Limit

# PetaLibrary

**Active**

- Performance tier
- Accessible by all node
- No file limit

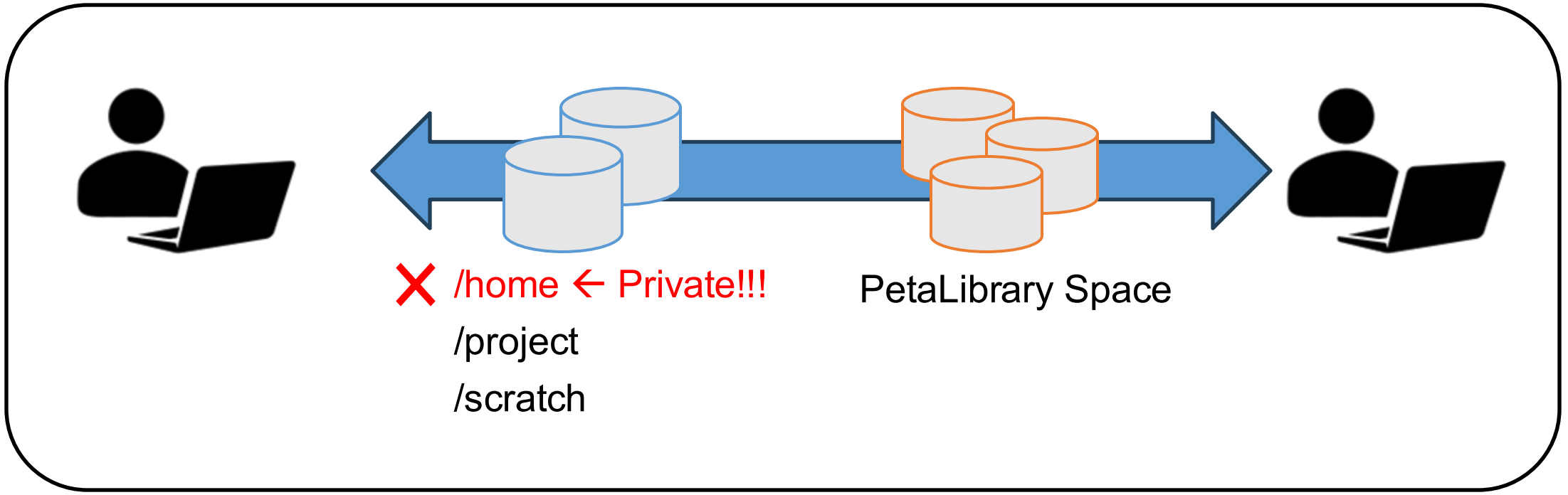
**Archive**

- Integrity tier
- Accessible by login node only
- 10,000 File Limit

**Active + Archive**

- Active synced to archive
- Updated every 15 minutes
- No file limit

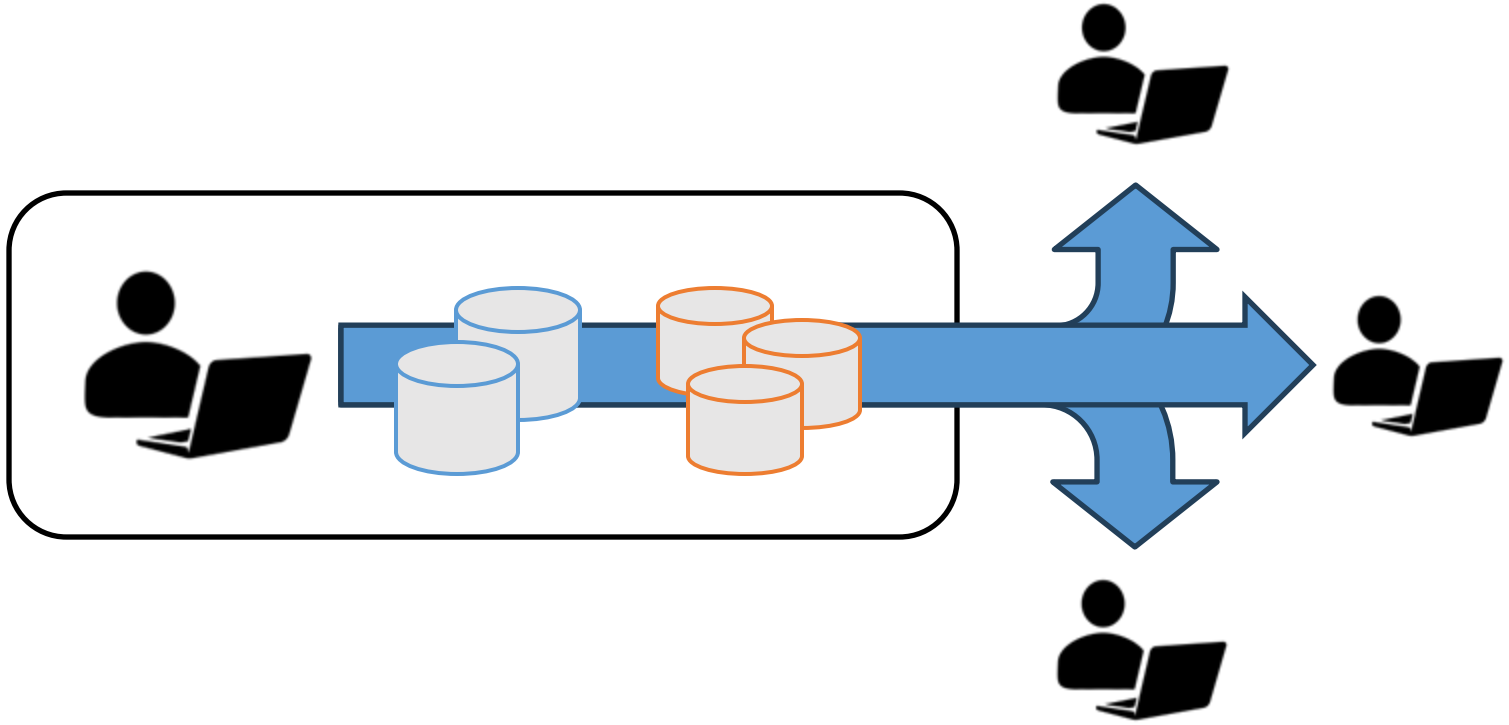
# Data Sharing: Within RC



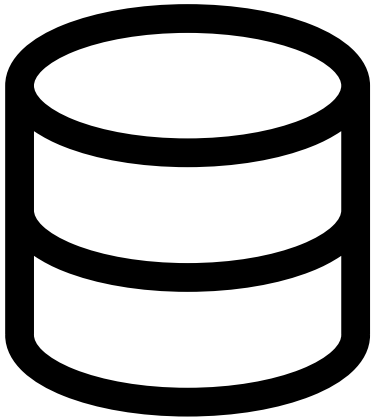
# Data Sharing: Outside RC

**Large Data Transfers:**

- Globus (Recommended)
- Data Transfer Nodes (DTN)
- Terminal/Command Line:
  - rsync
  - rclone
  - sftp
  - scp



# Data Transfer Nodes



**DATA (DTN)**

- Support data transfers
- Can be selected when using scp, sftp, or ssh transfers

# Acceptable data storage and use

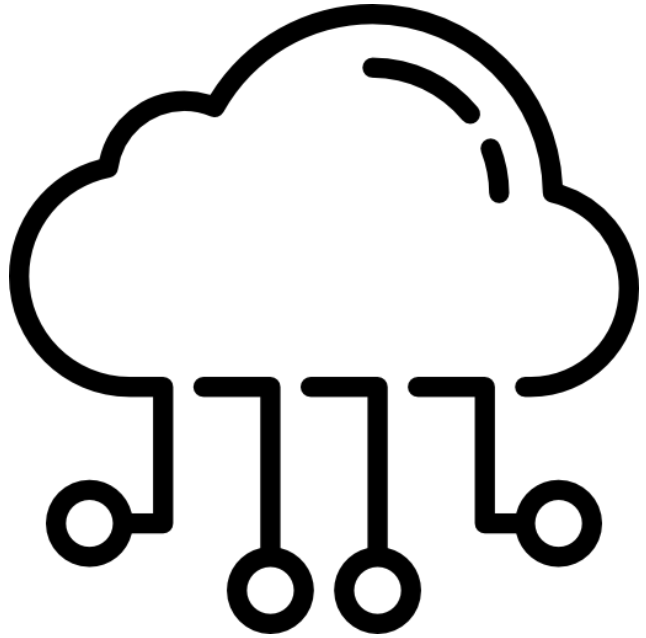
CURC systems and services **should not be used to store** any data that is US government Classified, nor any Controlled Unclassified Information.

For users requiring storage for sensitive data types, please see the secure research computing resources:

https://www.colorado.edu/rc/secure-research-computing-resources

# Cloud Computing

- CURC supports AWS, Azure, and GCP
- Alternative to HPC
- Enhance HPC



# How to Access RC Resources?

1. Get an RC account
2. Set up two-factor authentication with Duo
3. Log in
4. Create greatness! (responsibly)

# Getting an Account

- CU Boulder, CSU users and affiliates:
  - Request an account through the RC Account request portal: https://rcamp.rc.colorado.edu/accounts/account-request/create/organization
- AMC, RMACC users and affiliates:
  - Request an account through the ACCESS-CI User Registration Portal: https://identity.access-ci.org/new-user.html

# Your RC Account

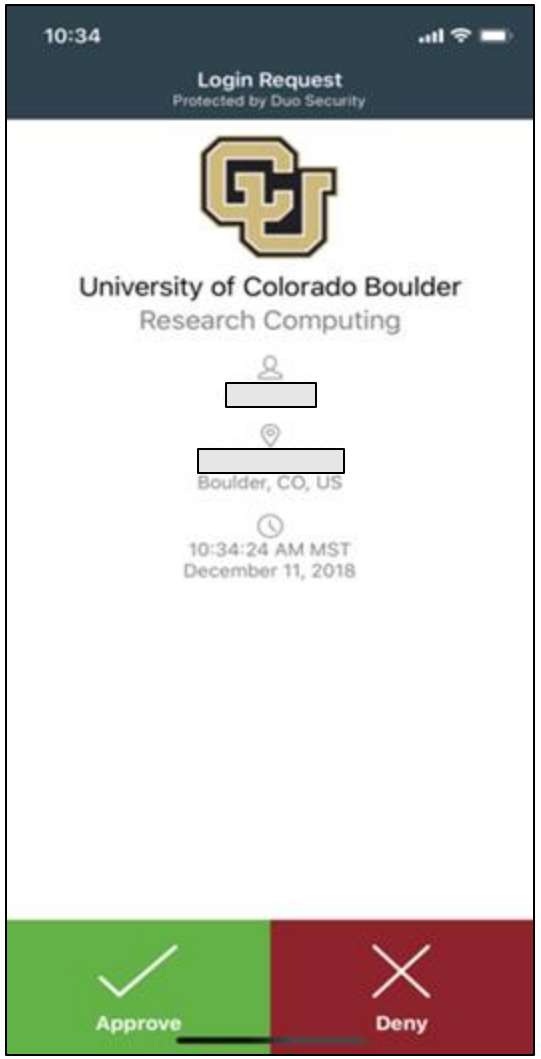
**Access to:**

1. Alpine Cluster
2. Core Storage
3. PetaLibrary Storage*
4. Open OnDemand
5. Approximately 2,000 Service Units (SUs) per month

*If purchased

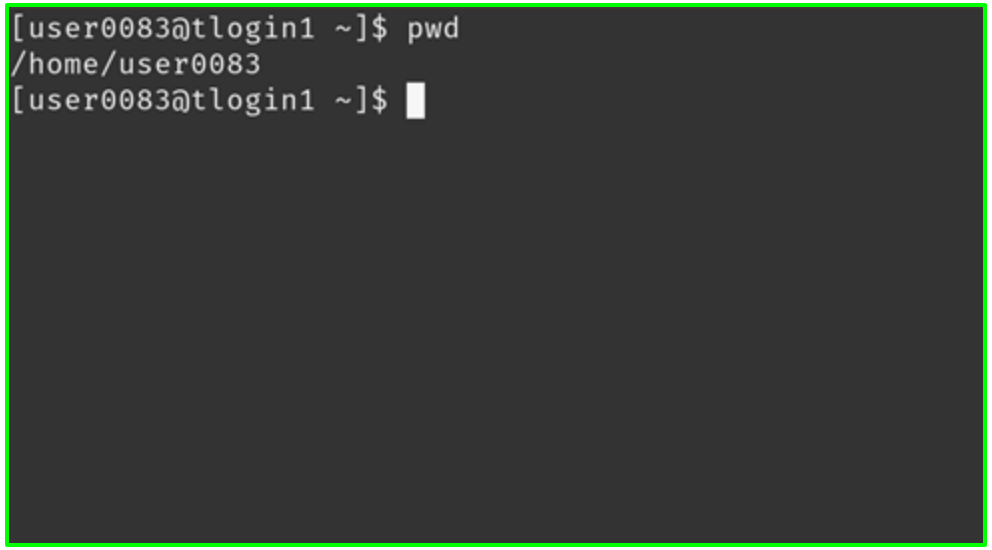
# Duo Authentication

1. **Duo smartphone app (recommended)**
2. Phone Call/Text is an alternatives



# Terminal Access

- Mac or Linux
  - Terminal application
- Windows
  - PuTTY
  - Powershell




- Open OnDemand (*alternative for CU affiliates*)
  - For those less familiar with Linux (ondemand.rc.colorado.edu/)

# Demo: Logging in via Terminal

To login to an RC login node:

```
$ ssh <username>@login.rc.colorado.edu
```

Supply your IdentiKey* password and your Duo app will alert you to confirm the login

* Exclusive to CU and CSU accounts

# Demo: logging in with OnDemand

CURC Open OnDemand is a browser based, integrated, single access point for all of your HPC resources at CU Research Computing.

- CU Boulder: Visit https://ondemand.rc.colorado.edu.
- Other RMACC Institutions: Visit https://ondemand-rmacc.rc.colorado.edu/

# Help! I'm stuck, where do I go?

- **Documentation**: curc.readthedocs.io/
- **Trainings with Center for Research Data and Digital Scholarship (CRDDS)**: https://www.colorado.edu/crdds/
- **Helpdesk**: rc-help@colorado.edu
- **Consult Hours (Tuesday 12:00-1:00 in-person or virtually, Thursday 1:00-2:00 virtually)**

# Helpdesk Tickets

To: rc-help@colorado.edu

Dear Research Computing,

Help! My code won't run!
Help!

Help please,
John

To: rc-help@colorado.edu

Dear Research Computing,

I am running into issues running my Python script. I am using a conda environment called my_python_env with the pytorch software, and I am receiving the following error. I am not sure how to troubleshoot. My job ID is 620350. Let me know what I can try!

sbatch: error: Batch job submission failed: Invalid partition name specified.

Thanks,
John

# Survey and feedback

http://tinyurl.com/curc-survey18