# Intro to HPC

**Introduction**: This workshop will introduce the core concepts behind High Performance Computing and the various services that ARCC provides. After the workshop, participants will understand:

- The various services that ARCC provides.
- The basic infrastructure that makes up a cluster.

**Course Goals:**

- To introduce ARCC's (HPC Center) Mission and Services.
- What are Clusters: ARCC and NWSC.
- What is HPC?
- What does a HPC/Cluster Architecture look like.
- Different types of storage.

**Topics**:

- ARCC (HPC Center) Mission and Services.
- Clusters: ARCC and NWSC.
- What is HPC?
- HPC/Cluster Architecture.
- Different types of storage.

---

## Sections


1. Introducing Available HPC Resources
2. Introducing UW ARCC
3. What is HPC?
4. Next Steps: Getting Started

# Introducing Available HPC Resources

**Goal**: Provide new HPC users with a list of common research computing resources available to the UW research community.

---


- ARCC
- NCAR
- NWSC
- ANL
- With many HPC options, what should you use?
- Next Steps


---

# ARCC

## ADVANCED RESEARCH COMPUTING CENTER

The Advanced Research Computing Center (ARCC) is the primary research computing facility for the University of Wyoming. We provides centralized scientific computing resources, including HPC and research storage, and our center is a gateway to other research institutions within Wyoming and across the nation.

## LINKS TO OUR RESOURCES

| | | |
|---|---|---|
| ABOUT ARCC/CONTACT US | GETTING STARTED | SERVICE REQUEST PORTAL |
| DOCUMENTATION & HELP | NEWS & ANNOUNCEMENTS | TRAINING & WORKSHOPS |
| PARTNERS | LIST OF SERVICES | STUDENT EMPLOYMENT |

# NCAR

**NCAR: National Center for Atmospheric Research:**

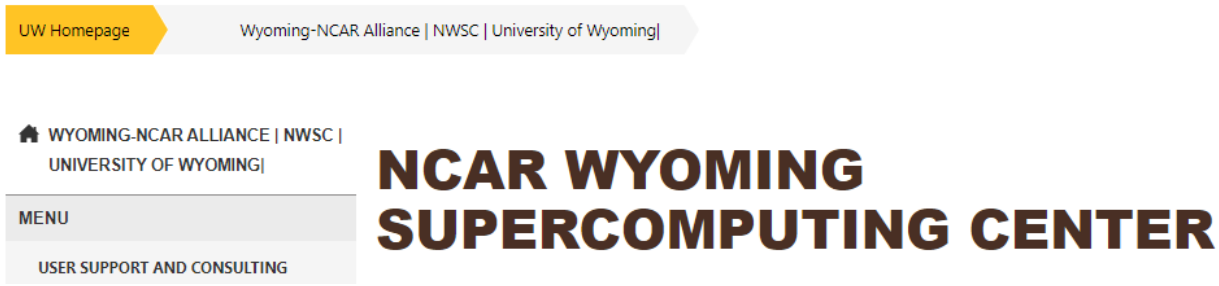
- **A center of research excellence in Earth system science sponsored by the National Science Foundation.**
- **World-Class Research in Earth System Science**


**CISL: Computational & Information System Lab**

- **Help for researchers using NCAR facilities**

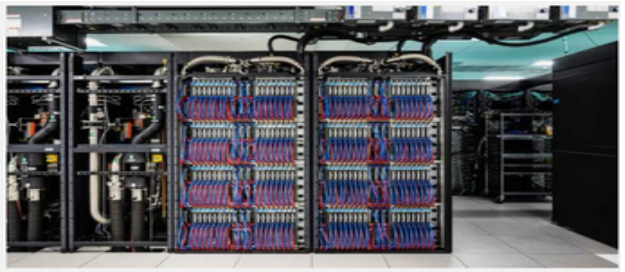
# NWSC

The NCAR Wyoming Supercomputing Center (NWSC) represents a collaboration between NCAR and the University of Wyoming.




---

# ANL

| | |
|---|---|
| Argonne NATIONAL LABORATORY<br>(ANL) |  <br>Argonne Leadership Computing Facility<br>(ALCF) |

| | |
|---|---|
| A multidisciplinary science and engineering research center.<br>Serving as a science and energy laboratory distinguished by R&D capabilities with a powerful suite of experimental and computational facilities. | The Argonne The Argonne Leadership Computing Facility enables breakthroughs in science and engineering by providing supercomputing resources and expertise to the research community. |

---

# With many HPC options, what should you use?

The answer depends on your needs:

- HPC Experience
  - ARCC offers more support for new users
- Specific hardware requirements?
  - Specific models of GPU?
  - Capacity?

Generally, if you're just starting out and wanting to learn HPC, ARCC HPC's offer a good "learning environment" with access to significant computing resources, more user-friendly interfaces, and more extensive user support.
Additional information on NWSC and ANL are available at the links below.

- NWSC
- ANL

---

# Introducing UW ARCC

**Goal**: Introduction to UW ARCC and our services.

- ARCC: Mission
- Core Service 1: High Performance Computing: HPC
- Core Service 2: Research Data Storage
  - Core Service 2: Research Data Storage: Which One?
  - Core Service 2: Research Data Storage Changes:
- Core Service 3: End User Support
- Other Services
- Next Steps


---

# ARCC: Mission

Provide support for research computing endeavors including:

- high performance computing
- large research data storage
- and consulting

to further the University of Wyoming and State of Wyoming's strategic priorities by enabling researchers with the necessary computational needs.

---

# Core Service 1: High Performance Computing: HPC

We maintain a number of **clusters** for the purpose of allowing researchers to perform a variety of use cases such as running:

- Computation-intensive analysis on large datasets.

- Long large-scale simulations.
- 10s/100s/1000s of small short tasks - nothing is too small.
- and lots of other use case…

---

# Core Service 2: Research Data Storage

**Safe and secure storage** and **transfer of data** that researchers can **share and collaborate** on with others within UW, and other institutions across the world.

1. **Alcova:**
    1. High performance data storage geared toward **project-oriented data**.
    2. Storage for **published research data**.
2. **Pathfinder:**

    1. Low-cost storage solution that enables a **Cloud-like presence** for research data hosted by ARCC.
    2. Hosting **onsite backups** and enabling data sharing and collaboration.

---

## Core Service 2: Research Data Storage: Which One?

**Considerations: Cost vs Usability:**

1. **Alcova:**
    1. Consider as more traditional storage that can be accessed via SMB/AD via a *traditional* Windows File Explorer/Globus.
    2. Access follows the idea of a project that users are part of and authenticated via username/AD.
2. **Pathfinder:**

1. A cheaper storage solution that is accessed either via a client and/or programmatically that uses S3 to provide object storage via *buckets*.
2. Access it provides is via access/secret key tokens, that can be time based.
3. Data can be made publicly available.
4. It does not user the notion of projects/usernames.

**Come and discuss what your needs and use cases are…**

---

# Core Service 2: Research Data Storage Changes:

1. **Data Portal:**
    1. Effective June 1, 2024, ARCC introduced the 'ARCC Data Portal' serving the dual purpose of providing high performance back end storage for the MedicineBow HPC system and a data storage solution for researchers needing a centralized data repository for ongoing research projects.
    2. Data Portal storage is FREE up to the default allocation quota.
    3. ARCC's Data Portal is compromised of VAST data storage compromised of high speed all-NVMe storage, housing 3 petabytes of raw storage. VAST storage employs data de-duplication allowing the system to logically store more than the raw 3PB available.
    4. **MedicineBow vs Alcova Spaces:**
        1. Alcova storage on the ARCC Data Portal can be thought of as the "new Alcova" and will replace the prior Alcova storage space listed here. This space is intended for use as collaborative data storage space using SMB protocol for interactive access. This space is backed up by ARCC and can only be used by researchers with a uwyo.edu account.
        2. MedBow space can be thought of as the root level directory of the HPC system, separated into home, project, and gscratch directories, intended for use with HPC workflows where speed and minimal overhead are prioritized over backups.
            1. MedicineBow Data Storage is available upon the go-live of MedicineBow on July 15th.

**– the essence of these services will remain**

**– but the underlying systems are being updated**

# Core Service 3: End User Support

Available to all researchers at UW: User Services

- Website/Wiki - examples and suggested best practices.
- Service Portal:
- Zoom office hours.
- One-on-one consultation.
- Scheduled in-person and (online) trainings.
- YouTube Channel.

---

# Other Services

- OnDemand: Web based access to the clusters - included Jupyter Notebooks.
- Linux Desktop Support.
- Hosting of services - R Shiny application…
- A GitLab service.
- Proposal Development.

- As research needs grow, so will services that we offer.

- We're always open to constructive feedback and suggestions.
- We're here to provide the services for ***you***…

---

# What is HPC?

**Goal**: Provide new users with an understanding of what HPC is, how it works, and why it's useful.

---


- HPC: High Performance Computing
- What is a Cluster
- Compute Nodes
- Core Service 1: HPC: What does this look like?
- UW IT Data Center
- Types of HPC systems
- Cluster and Partitions
- Reservations
- Condominium Model
- Next Steps


---

# HPC: High Performance Computing

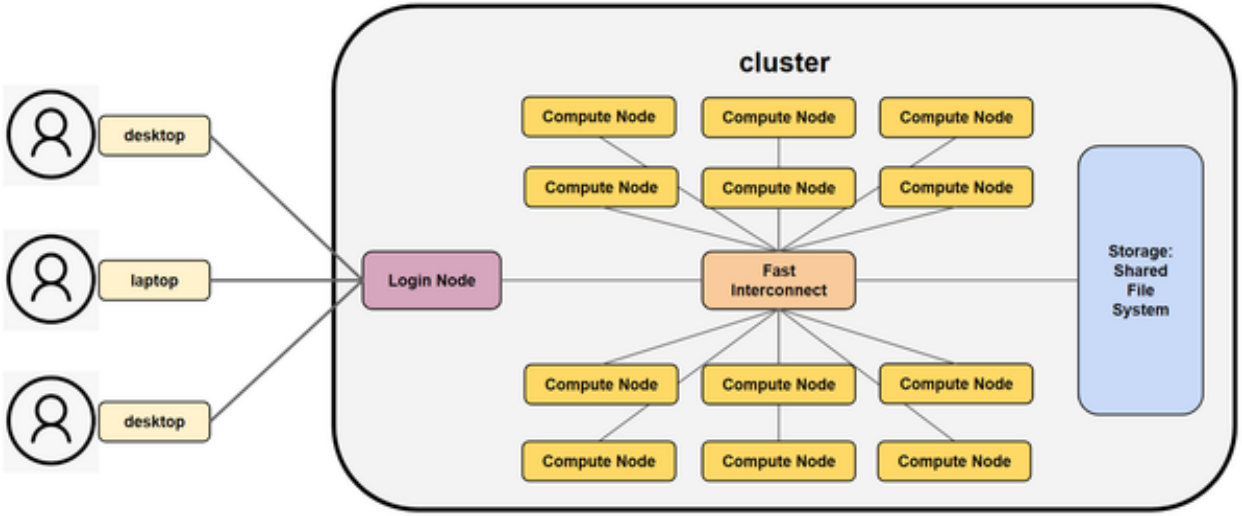
"High Performance Computing most generally refers to the **practice of aggregating computing power** in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business."

**HPC ≠ Desktop**

**HPC >> Desktop**

---

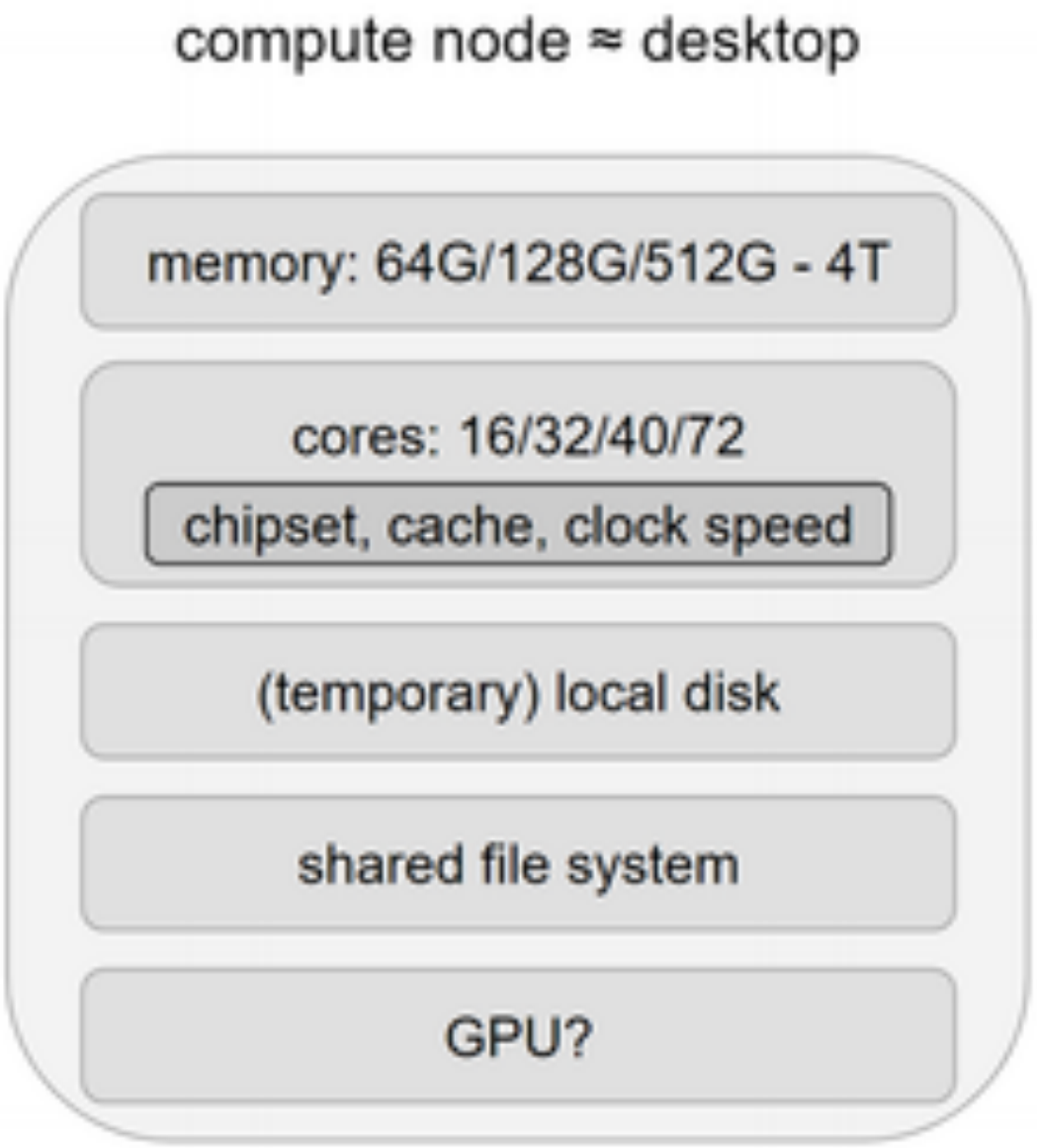
# What is a Cluster




- Users log in from their clients (desktops, laptops, workstations) into a **login node**.
- In an HPC Cluster, each compute node can be thought of as it's own desktop, but the **hardware resources of the cluster are available collectively as a single system**.
- Users may request specific allocations of resources available on the cluster - beyond that of a single node.
- Allocated resources may include CPUs (Cores), Nodes, RAM/Memory, GPUs, etc.

---

# Compute Nodes


compute node ≈ desktop
memory: 64G/128G/512G - 4T
cores: 16/32/40/72
chipset, cache, clock speed
(temporary) local disk
shared file system
GPU?

- We typically have multiple users **independently** running jobs concurrently across compute nodes.
- Resources are shared, but **do not interfere with any one else's resources**.
  - i.e. you have your own cores, your own block of memory.
- If someone else's job fails it does NOT affect yours.
- **Example**: The two GPU compute nodes part of this reservation each have 8 GPU devices. We can have different, individual jobs run on each of these compute nodes, without effecting each other.

---

# Core Service 1: HPC: What does this look like?

We maintain a number of **clusters** for the purpose of allowing researchers to perform a variety of use cases such as running:

- Computation-intensive analysis on large datasets.
  - Megabytes / Gigabytes / Terabytes.
  - On the filesystem in one / many files.
  - In memory.
  - CPU only vs GPU enabled.
- Long large-scale simulations.
  - Hours, days, weeks…
  - Single job across multiple nodes each using multiple cores.
- 10s/100s/1000s of small short tasks - nothing is too small.
  - Seconds, minutes, hours…
  - Single node - one to many cores.
- and lots of other use case…

---

# UW IT Data Center

---

# Types of HPC systems

There are generally two type of HPC systems:

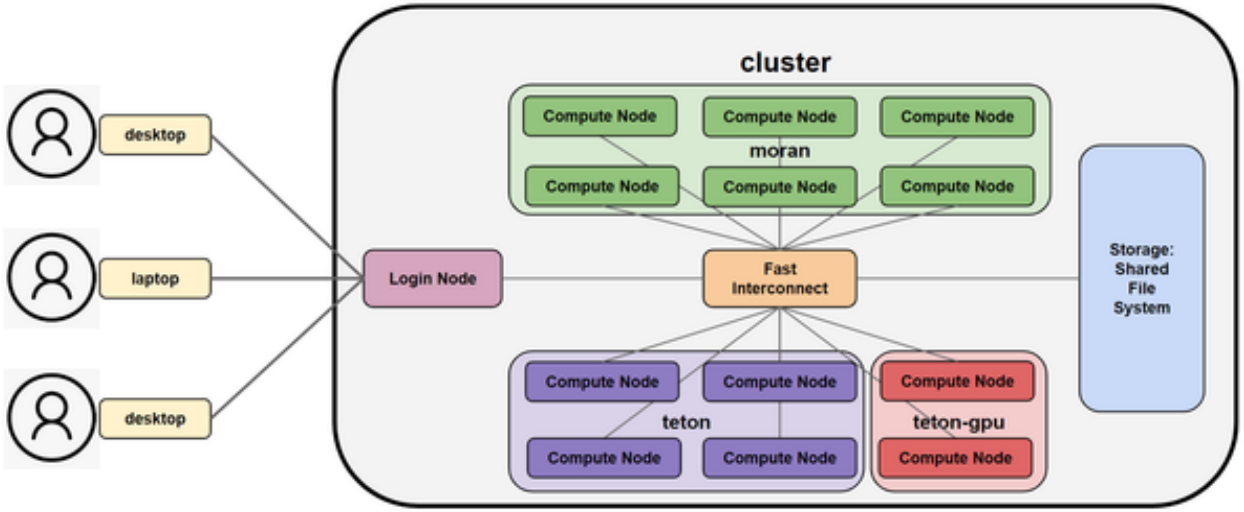
1. **Homogenous**: All compute nodes in the system share the same architecture. CPU, memory, and storage are the same across the system.
    1. Derecho:
    2. Cheyenne:
2. **Heterogenous**: The compute nodes in the system can vary architecturally with respect to CPU, memory, even storage, and whether they have GPUs or not.

        1. Typically, similar compute nodes are grouped via **partitions**.

2. Beartooth Hardware Summary Table

---

# Cluster and Partitions




---

# Reservations

A reservation can be considered a **temporary partition**.

It is a set of compute nodes reserved for a period of time for a set of users/projects, who get priority use.

For example, a reservation would look like the following:

```
ReservationName = biocompworkshop
```

```
StartTime = 06.09-09:00:00
EndTime   = 06.17-17:00:00
Duration  = 8-08:00:00
Nodes     = mdgx01,t[402-421],tdgx01 NodeCnt=22 CoreCnt=720
Users     = Groups=biocompworkshop
```

---

# Condominium Model

The "**condo model**".

- Allow researchers to **invest into the cluster** - purchasing additional compute nodes that they get **priority to use**.
- '**preempt**' jobs outside of the investor's project - allow the investor to start their jobs immediately.
  - "*immediately*" if no other jobs from that investment project are already using the investment.
  - A preempted job is stopped and automatically re-queued. When it starts will be determined by the current cluster utilization.
  - Consider the idea of **check-pointing** which allows a job to continue analysis at the point where it was stopped.
- This is managed by defining '**investor partitions**'.
- ARCC Investment Program

---

# Next Step: Getting Started

Goals: Provide information for how to ask for help, and how to get started with ARCC HPC Resources

---


- How to Get Help
- Next Steps: Request an Account with ARCC
- Summary
- Next Steps

# How to Get Help

- Service Portal and arcchelp@uwyo.edu
- Zoom Office Hours.
- When requesting help:
  - Be as clear and specific as you can.
  - Provide enough detail so we can replicate your issue.
  - Provide job ids, log files, working folder paths…
  - Software and module versions.
  - Links to your software homepage/repo and what you've tried.
  - Have you consulted online communities?
  - When using other people's code: try and understand what it is doing before asking for help.
- We'll always do our ***best effort***, but we're not domain experts (supporting *all* researchers).

# Next Steps: Request an Account with ARCC




- Web: [Advanced Research Computing Center | Overview]
- Wiki: [ARCC Wiki]
- Portal: https://arccwiki.atlassian.net/servicedesk/customer/portals

---

# Summary

Covered:

- ARCC (HPC Center) Mission and Services.
- Clusters: ARCC and NWSC.
- What is HPC?
- HPC/Cluster Architecture.
- Different types of storage.

---

# Next Steps

| Previous | Workshop Home |
| --- | --- |
| What is HPC? | Intro to HPC |

Use the following link to provide feedback on this training: https://forms.gle/oSYCKBWSchotoJDG7 or use the QR code below.