The Center for High Performance Computing: Support for Research Computing and Data

Martin Čuma

Scientific Consultant, CHPC

m.cuma@utah.edu

CENTER FOR HIGH PERFORMANCE COMPUTING
THE UNIVERSITY OF UTAH

SCIENTIFIC COMPUTING & IMAGING INSTITUTE

SCI
THE UNIVERSITY OF UTAH

Who We Are
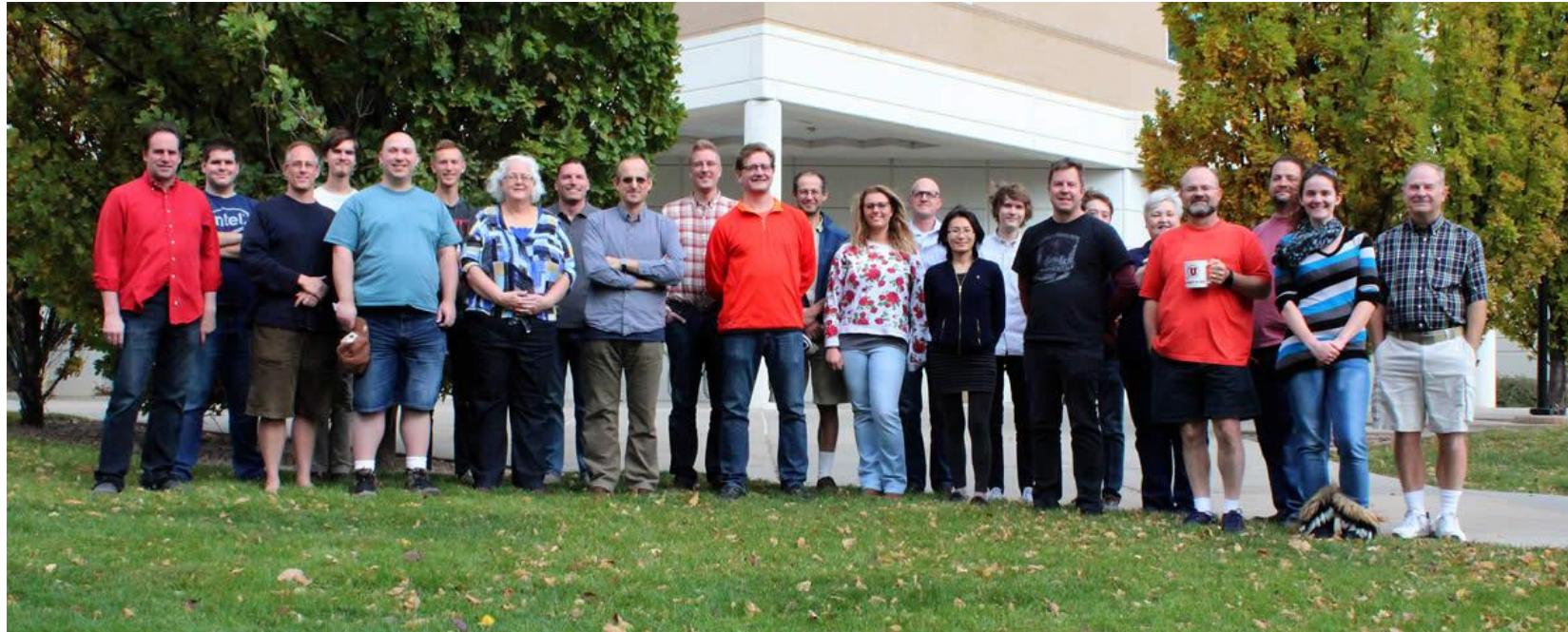
- Staff of 30 full-time professionals and 17 part-time students
- Backgrounds in a variety of sciences, engineering, management
- Expertise in scientific computing, networking, data storage, software development, system administration, ...

CHPC can help advance your research!

- You need parallel processing

- You need access to a single high-powered computer

- You need to run many individual jobs simultaneously

- You have a large amount of data to store and/or process

- You need software you don't have on your computer

- Your data is sensitive/restricted:

  - protected health information, IRB-governed, Controlled Unclassified Information, or otherwise sensitive/restricted

- You have other computing needs your local resources cannot meet.

Center for High Performance Computing (CHPC)
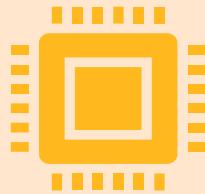
Mission: Provide advanced computing and data systems and services to accelerate discovery and innovation across the UofU and state, contributing to the regional and national research enterprise.

Resources and services

- *Computational resources* – High performance and compliant computing systems

- *Data and networking* – storage, archiving, scratch, home directory, collaboration spaces, data serving, high-speed data transfer

- *Environments* – Windows servers, Linux clusters, Virtual Machines (VMs), and protected environments (enclaves for sensitive data, compute, storage, …)

- *User Support* – onboarding, consultation and facilitation, training, software environment building, allocations, and collaboration.

- *Workforce development* through extensive student employment.

- *Partnering* - in UofU, regional and national initiatives and collaborations.

CHPC resources dashboard

- Last 10 years (where data exist)
- New analytics engine
- Live data
- User data WIP

charts.chpc.utah.edu/grafana/d/chpc-summary/chpc-site-summary

CHPC Linux Clusters
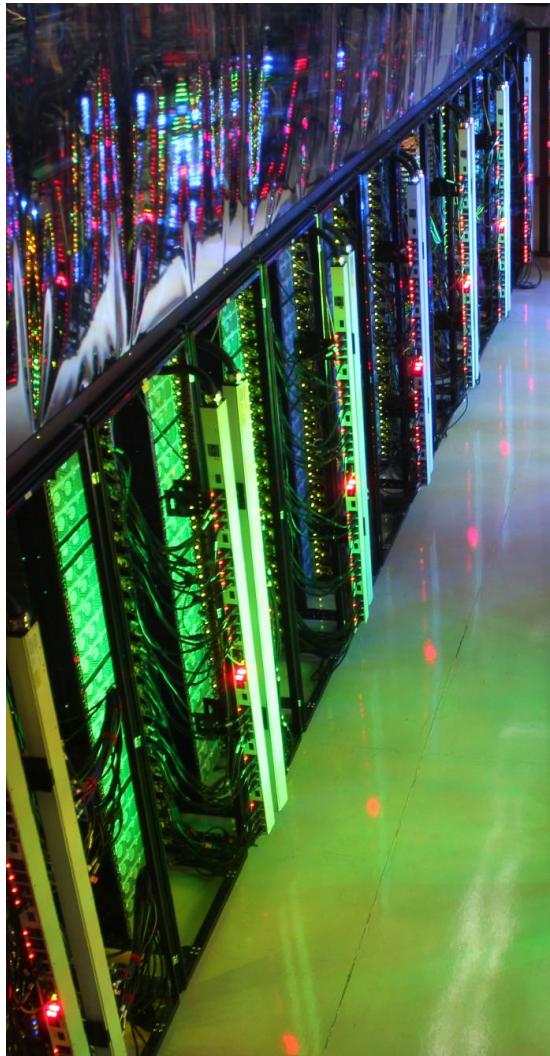
- "Condominium" model
  - CHPC-purchased nodes, available to all, priority access with time allocation[*]
  - Faculty-purchased nodes, priority access to owner, guest access to others
- Interactive, compute, and GPU nodes
- Manage jobs with Slurm scheduler for batch or interactive computing
- Access via web portal, remote desktop or terminal.

www.chpc.utah.edu/resources/HPC_Clusters.php

Environment	Cluster	Compute Nodes	CPUs	GPUs
General	notchpeak	501	27,908	246
General	kingspeak	268	6,352	24
General	lonepeak	124	2,456	124
General	granite	72	6,016	72
Protected	redwood	232	8,804	232

[*] on notchpeak, granite, and redwood clusters only

Cluster Architecture Diagram
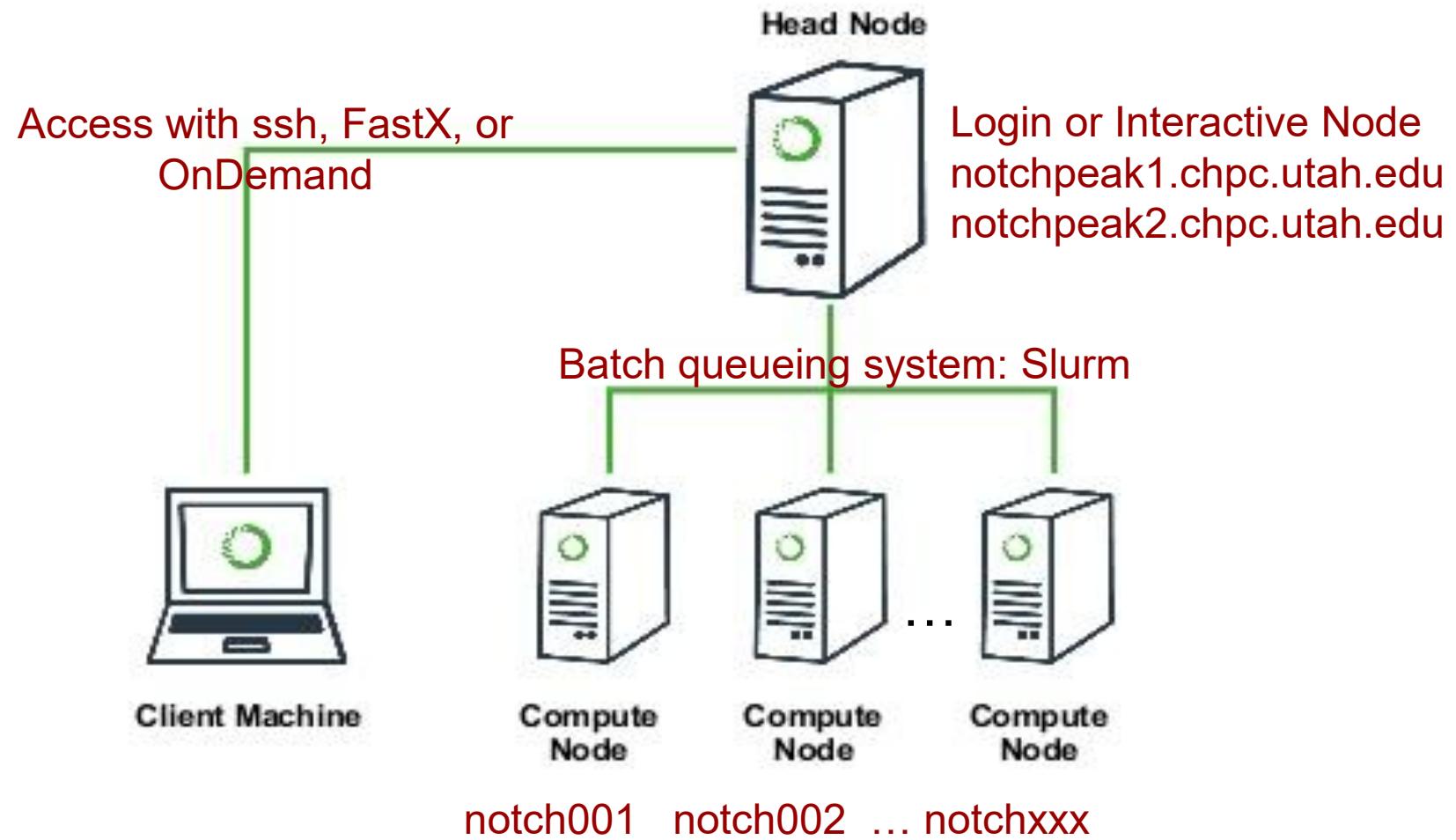
Downtown Data Center

- Houses most CHPC gear since 2012
- Shared with enterprise (academic/hospital) groups
- 92 racks and 1.2MW of power, upgrade path to add capacity for research computing
- Fiber optic network connects campus, data center, & internet2
- 24/7/365 facility
- Power, cooling, network connectivity, security

Data Storage

- *Home Directories -- /uufs/chpc.utah.edu/common/home/<uNID>*
    - Located on a high-performance storage system named VAST
    - Home directories hold up to 50 GB, **not backed up**
    - Groups can buy larger home directory space at $360/TB for 2 years (originally $900/TB for 5 years), backed up (nightly incremental, weekly full, 2-week retention)

- *Group Level File Systems*
    - NFS mounted group space: $450/TB/5 years, backed up, or $150/TB/5 years **not backed up**

- *Scratch File Systems*
    - For use by all users; 50 TB quota; **files older than 60 days removed**
    - 600 TB NFS scratch space (/scratch/general/nfs1), 1.6 PB VAST scratch space (/scratch/general/vast)
    - Local scratch on compute nodes, up to 1TB (TMPDIR=/scratch/local/$USER/$SLURM_JOB_ID)

- *Archive Storage*
    - Archive space costs $150/TB/5 years, similar to cloud storage, but on-site
    - State archive prototype – automatic syncing from SLC to St. George.

www.chpc.utah.edu/resources/storage_services.php

CENTER FOR HIGH
PERFORMANCE COMPUTING
THE UNIVERSITY OF UTAH

SCIENTIFIC COMPUTING
& IMAGING INSTITUTE
THE UNIVERSITY OF UTAH

Example data-intensive projects supported by CHPC

Project	PI and group
Sloan Digital Sky Survey (CHPC is primary data site and replicates to other institutions)	Joel Brownstein, Phy & Astro
VERITAS (Very Energetic Radiation Imaging Telescope Array System) (CHPC is a mirror site)	David Keida, Cosmic Ray
Dark Energy Survey Instrument (DESI) (CHPC is a mirror site)	Kyle Dawson, Phy & Astro
HRRR (High Resolution Rapid Refresh), and Mesowest	John Horel, Atmospheric Science
CTSI (Clinical & Translational Science Institute)	Ram Gouripeddi, Randy Madsen, BMI
BMIC (BioMedical Informatics Core)	Julio Facelli, BMI
UPDB (Utah Population Database)	Nicki Camp, UPDB
SPARC (Surgical Population Analysis Research Core)	Rupam Das, SPARC
RMCDS (Rocky Mountain Cancer Data Systems)	Daniel Denhalter, RMCDS
Icecube	Carsten Rott, Phy & Astro

Secure Computing

- Protected Environment (PE)
  - Dedicated secure resources for handling data & projects with protected information
  - Currently PHI and projects with other types of sensitive data/restrictions
  - Careful with human genomic data – does not meet NIH dbGaP requirements
  - HPC cluster (redwood), VM farm, data storage (home, group space, archive, scratch)
  - Windows compute server (narwhal)
  - www.chpc.utah.edu/resources/ProtectedEnvironment.php

- Citadel
  - Temporary measure for dbGaP (NIST 800-171 rev 2)

- Regulated environment
  - Working on demo, planning together with campus stakeholders

Windows Servers

- Beehive (general environment) – refreshed 2019
  - 48 physical CPU cores, 2.3 Ghz, 512GB memory
- Narwhal (protected environment)
  - 24 physical CPU cores, 3.0 Ghz, 512GB memory
- For use with Windows specific programs
  - ArcGIS
  - Microsoft Office
  - Microsoft SQL server
- ***If you need other software, please contact us to discuss***

www.chpc.utah.edu/documentation/guides/beehive.php

www.chpc.utah.edu/documentation/guides/narwhal.php

Virtual Machines

- For needs and applications that do not fit in compute cluster or Windows server model

- Multiple VM servers with failover – hardware refreshed 2025, includes data storage

- Community VMs for mysql, mssql, git repositories, web servers, etc, free of charge

- Other VMs (not the community VMs) have a cost, both for the VM and for any customization needed.

www.chpc.utah.edu/resources/virtualmachines.php

Blocks	RAM (GB)	Cores	Storage (GB)	Price/ 5 yrs
Small	8	2	50	$550
Medium	16	4	100	$800
Large	32	8	200	$1250
Xlarge	64	12	400	$2200
2XLarge	96	12	800	$4100

Additional VM storage available, in 100GB increments, at a cost of $1300/TB.

CHPC provides dedicated network routes within and beyond the UofU

- 100G and 400G connections to Utah Education and Telehealth Network (UETN)

- UETN connects to Internet2 via 2x400G in SLC and 2x100G in Las Vegas

- CHPC also connects directly to three national testbeds: FABRIC (100G), Cloudlab (100G), and POWDER (100G)

- We provide transfer tools such as Globus, rclone, s3cmd,...

www.chpc.utah.edu/documentation/data_services.php
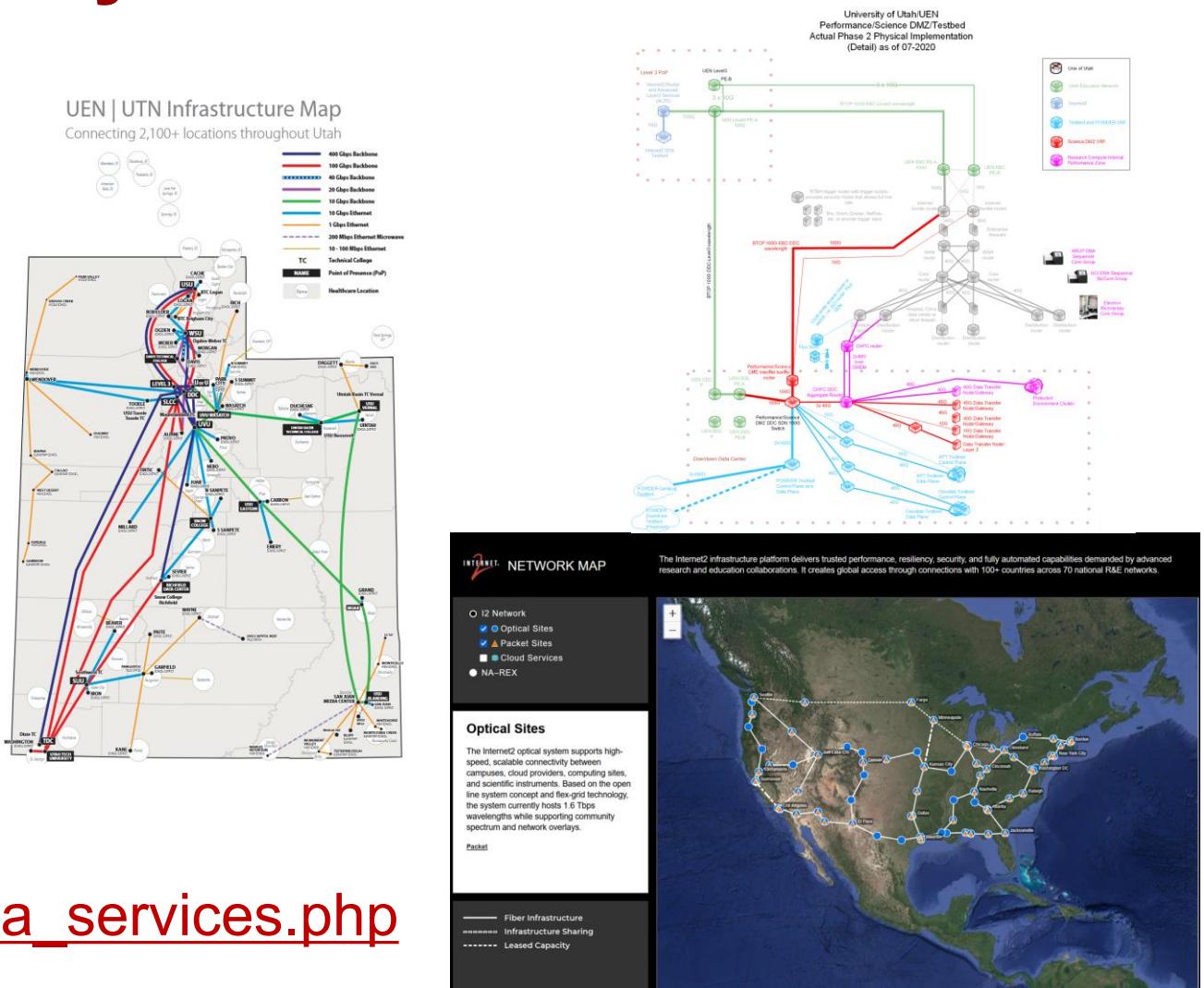
Networking and data transfers

- Fast network among HPC components
  - high bandwidth, low latency
  - compute - Multi-node jobs using Message Passing Interface (MPI) libraries
  - storage - access to all your files from any node
- Data transfer nodes
  - Dedicated network routes inside, outside the campus
  - Transfer tools - Globus, rclone, s3cmd, …

www.chpc.utah.edu/documentation/software/mpilibraries.php
www.chpc.utah.edu/documentation/data_services.php
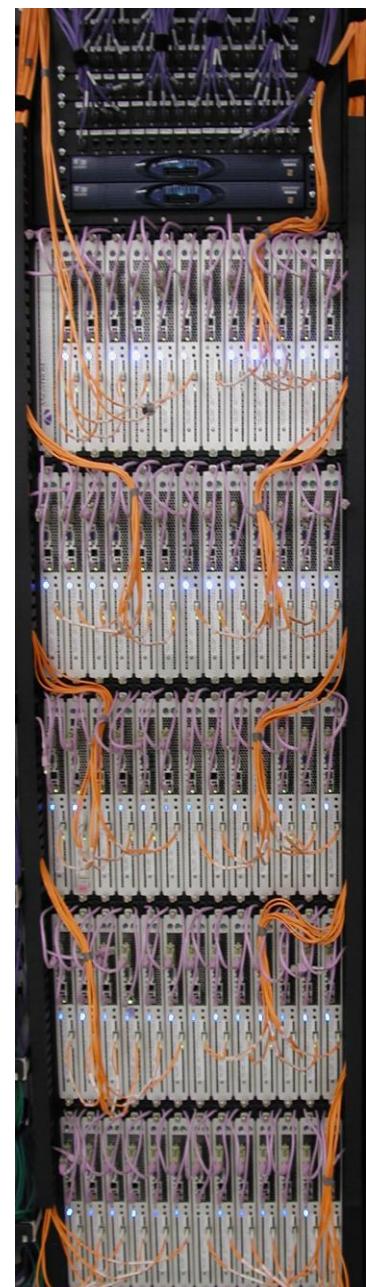
How to get started

1. Request a CHPC account.
   - Go to https://www.chpc.utah.edu/userservices/accounts.php
   - Requires UofU uNID, and every user account is linked to a Principal Investigator.
   - Account provides 50 GB free home directory space, login scripts, access to clusters, 50 TB scratch space, windows server, software.

2. E-mail helpdesk@chpc.utah.edu to request a consultation about your research needs and available services.

3. Try free CHPC resources (kingspeak, lonepeak, owner-guest) and learn from CHPC documentation and training sessions (see www.chpc.utah.edu)

4. Consider:
   - As PI, consider delegating your duties to a group member (account approvals, allocations,...).
   - Request an allocation as needed for your research.
   - Purchase compute resources (owner nodes), storage (group, archive), and pay for dedicated Virtual Machines depending on your project and long-term needs.

What if: I'm a new group member

- Apply for CHPC account, choosing your PI as PI
- If possible, work with your PI or peers on getting established at the CHPC
- Try CHPC free resources (kingspeak, lonepeak, owner-guest)
- Look at www.chpc.utah.edu for documentation
- Ask CHPC AI chatbot (coming soon)
- E-mail helpdesk@chpc.utah.edu if you need help

Getting an Allocation

- www.chpc.utah.edu -> User Services -> Allocations
    - Provides priority access to notchpeak, granite, redwood clusters
        - GPU nodes on granite also require an allocation as of 1/1/25
    - Regular allocations - applications reviewed each quarter
    - Quick allocations (< 30k CPU, <300 GPU SUs) - reviewed immediately
    - Allocations last up to 1 year (4 quarters)
    - Application requires description of research, estimated usage
    - Allocations managed on per-lab basis, not per-individual or per-project
- View allocation usage:
  https://portal.chpc.utah.edu/slurm/allocations/overview/chpc
  https://portal.chpc.utah.edu/slurm/allocations/overview/chpcpe
- View allocation and cluster access: *mychpc batch* command

Software: installed by CHPC

- ~850 different applications, variety of disciplines, multiple versions of each, most (but not all) open source

- Packages managed with software "modules"
  - "module avail" lists available modules
  - "module spider *keyword*" to search for modules

- Variety of compilers, debuggers, MPI & math libraries, containers

- Commonly used Python, R environments

- Git, gitlab for version control

- If you need a package installed - ask us!

www.chpc.utah.edu/documentation/software/index.php

Software: installed by you

- Python+ environments - miniforge, mamba

- Python modules - pip (python)

- R libraries

- Download existing binaries

  - make sure to choose RHEL8, CentOS8, Rocky Linux 8,…

  - or Ubuntu <=20, Ubuntu 22 is too new for Rocky Linux 8

- Pull and run containers using Apptainer

- Build your own programs

- Create your own modules

Open OnDemand

- Open OnDemand (OOD) web portal provides access to CHPC resources

- View, edit, upload and download files

- Create, edit, submit and monitor Slurm jobs

- Run applications including RStudio and Jupyter Lab

- Connect via a web browser, requires minimal knowledge of Linux and Slurm commands.

- Available in both General and Protected Environments
  - General environment: https://ondemand.chpc.utah.edu
  - Protected environment: https://pe-ondemand.chpc.utah.edu

www.chpc.utah.edu/documentation/software/ondemand.php

Slurm scheduler

- Slurm (Simple Linux Utility for Resource Management) controls access and schedules jobs on the cluster

- You request the kinds of resources you need (how much, for how long) and Slurm connects you to them, or puts you in the queue to wait for them

- Accessed with a few simple Linux commands:
  - sbatch or salloc - start a job (either batch or interactive)
  - scancel - stop a job
  - squeue - check on jobs
  - Also sacct, sinfo, scontrol …

- CHPC provides several easy replacements:
  - mychpc - show my access to computing resources
  - mysinfo - show the status of those resources
  - mysqueue - show the status of my jobs

www.chpc.utah.edu/documentation/software/index.php#scheduler

Training, Support and Consultation

- Presentation Series
  - www.chpc.utah.edu/presentations
  - Fall, Spring, Summer semesters
  - Free, open to everyone
  - Mix of lectures and hands-on sessions (linux, python, R, git, OnDemand)
  - If you have suggestions for other topics contact us
  - If you are interested in presentations for classes, research groups, contact us
- Documentation
  - www.chpc.utah.edu/documentation
- Ticketing System: email helpdesk@chpc.utah.edu
- Consultations: email helpdesk@chpc.utah.edu

Costs

Accounts	Free
Cluster access	Free
Priority cluster access	Free (with time allocation)
50 GB home directory	Free
Scratch space access	Free
Training, Support & Consultation	Free
Group & archive disk space	$450/TB/5 years (backed up) or $150/TB/5 years (not backed up)
Virtual Machines	Varies, from $550 / 5 years
Owner CPU Compute Nodes	Varies, from ~$14,000 (as of 8/15/25) 96 cores, 384 GB RAM, 6.4 TB SSD local scratch, email us for quotes
Owner GPU Compute Nodes	Costs vary dramatically depending on number and model of GPU. Email us for quotes

What to watch on the horizon...

- **Research Computing and Data (RCD)** vision and strategy – will be engaging researchers and other U stakeholders.

- **Responsible AI (RAI)**-supported computational capabilities (NVIDIA nodes)

- **U-State-Huntsman-HPE Sovereign Partnership -** A $50 million collocated compute capability and managed services.

- **Regulated Environment (RE)** – Development of compliance program and delivery of full RE deployment in the next year (NIST 800-171 Rev2 → Rev 3)
  - Downtown Data Center (DDC) upgrades - CITADEL expansion and 12 new racks for RE.

- **Innovations**: On-prem cloud concepts for a more agile architecture than VMs

- **Data in Flight:** Core network upgrades to sustain 400G backbone