THE UNIVERSITY OF UTAH

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

Brett Milash

Scientific Consultant

brett.milash@utah.edu

CHPC's Mission

- Innovate, design, engineer, deploy, and operate **cost-effective**, **functional**, **high performance**, and **compliant research computing and data resources** for the university communities we support.

- Continue to innovate, optimize, secure, and embrace the **heterogeneous** and **rapidly changing IT landscape**.

- Seek out opportunities for **workforce development** and **continuous learning** for **student employees** and **staff**.

- Provide a **robust set of service offerings** to **support researchers** including **user support**, **facilitation**, **training**, **documentation**, and **collaboration**.

Who We Are

- Staff of 39 full-time professionals and 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 if:

- 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

THE UNIVERSITY OF UTAH

CHPC Resources & Services

- *Compute Clusters* – Granite, Notchpeak, Kingspeak, Lonepeak

- *Data Storage* – home directory, group spaces, scratch space, archive storage

- *Windows Servers* –windows-only applications (e.g. statistics programs)

- *Virtual Machines* – for needs not met with cluster and windows server

- *Protected Environment* – for sensitive data, includes compute cluster, storage, virtual machines, and Windows Server

- *Networking Support* – supports compute environment; high-speed data transfers, work with researchers on data movement

- *User Support* – assistance with use of resources; installation of applications; training; consultations

CHPC Linux Clusters

Environment	Cluster	Compute Nodes	CPU Cores
General	notchpeak	496	27,828
General	kingspeak	220	5,200
General	lonepeak	135	2,748
General	granite	10	864
Protected	redwood	223	8,244

- "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 system for batch or interactive computing

- Access clusters with ssh, fastx, or OnDemand

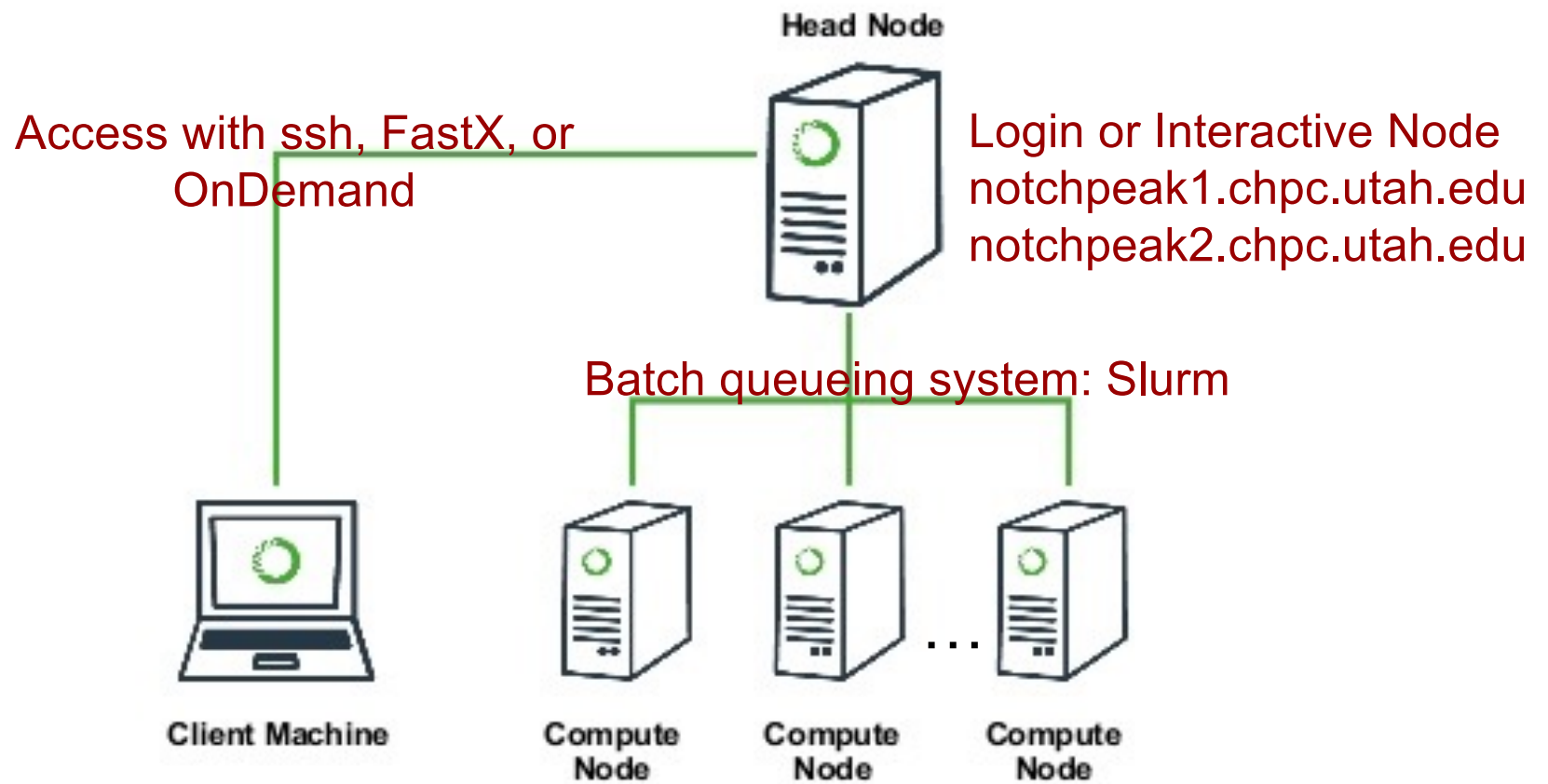* on granite, notchpeak and redwood clusters only
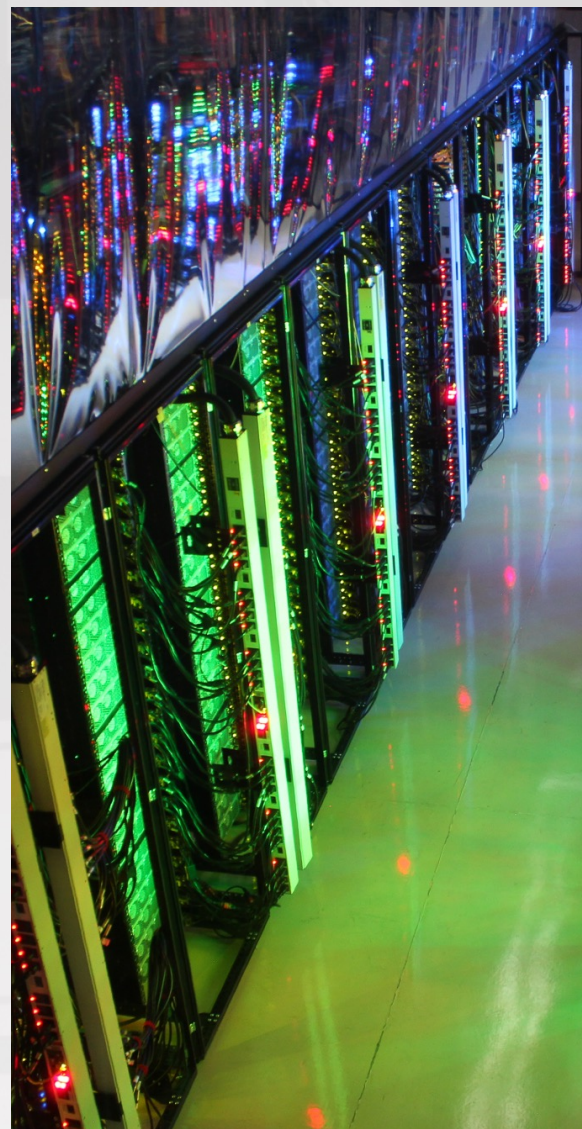
Downtown Data Center

- On-line in Spring 2012, CHPC completed move in Spring 2013
- 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

Cluster Architecture Diagram

Access with ssh, FastX, or OnDemand

Head Node

Login or Interactive Node
notchpeak1.chpc.utah.edu
notchpeak2.chpc.utah.edu

Batch queueing system: Slurm

Client Machine

Compute Node

Compute Node

Compute Node

notch001   notch002 … notchxxx

Open OnDemand

- Open OnDemand (OOD) web portal provides access to CHPC resources
- View, edit, upload and download files
- Create, edit, submit and monitor 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
- Our Presentation Series includes a class on OnDemand

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

Slurm

- 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

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

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 $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

Windows Servers

- Beehive (general environment) – refreshed 2019
  - 48 CPU cores, 512GB memory

- Narwhal (protected environment)
  - 24 CPU cores, 512GB memory

- Both have the following software installed
  - SAS 9.4 with text miner
  - R
  - STATA
  - Mathematica
  - Matlab

- *If you need other software, please contact us to discuss*

Virtual Machines

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

- Multiple VM servers with failover – hardware refreshed 2019; expanded 2021, includes data storage

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

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

Blocks	RAM (GB)	Cores	Storage (GB)	Price/5 yrs
1	4	2	50	$425
2	8	2	100	$615
4	16	4	200	$990
8	32	8	400	$1745
16	64	8	800	$3250

Additional VM storage available, in 100GB increments, at a cost of $1100/TB.

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

THE UNIVERSITY OF UTAH

Networking

- High Speed

- Reliable

- Low Latency

- Enables:
  - Multi-node jobs using Message Passing Interface (MPI) libraries
  - Access to all your files from any node
  - High-speed data transfer nodes
    - Globus, rsync, rclone, sra-toolkit, s3cmd, aspera
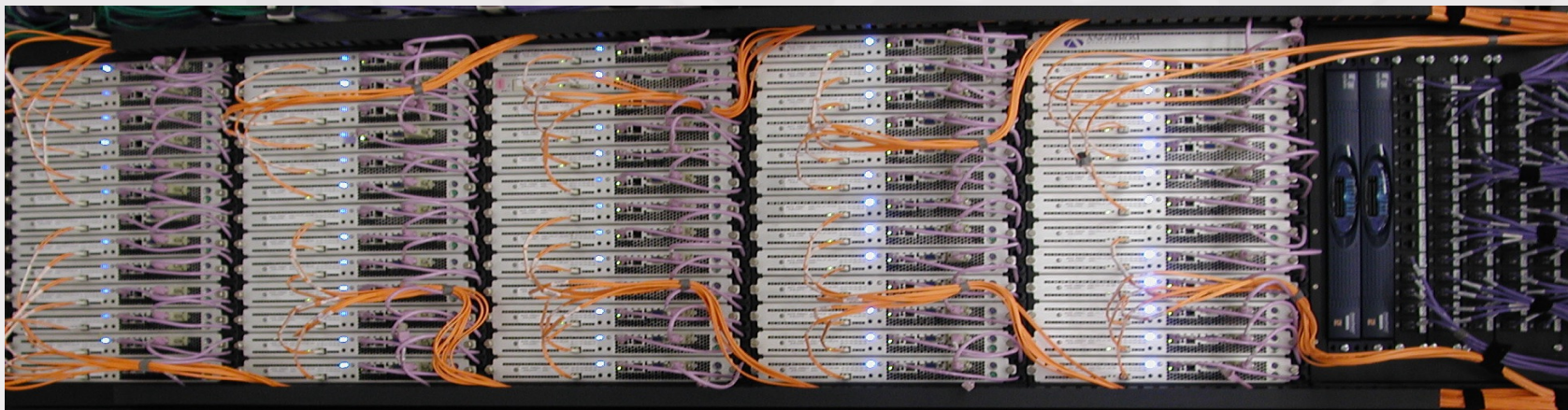
Secure Computing

Protected Environment (PE)

- Dedicated secure resources for handling data & projects with protected information
- Currently PHI and projects with other types of sensitive data/restrictions
- Preferred location for human genomic data – meets NIH dbGaP requirements
- Refreshed in 2017 with award of NIH Shared instrumentation grant
- HPC cluster (redwood), VM farm, data storage (home, group space, archive, scratch)
- Windows compute server (narwhal)
- See www.chpc.utah.edu/resources/ProtectedEnvironment.php

Cybersecurity Maturity Model Certification (CMMC) Enclave

- For Controlled Unclassified Information (CUI)

Software: installed by CHPC

- Over 600 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

- Git, gitlab for version control

- CUDA, CuPy for GPU programming

- If you need a package installed - ask us!

Software: installed by you

- anaconda, mamba
- pip (python)
- R
- Compilers for numerous languages
- Create your own modules

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 $425 / 5 years
Owner CPU Compute Nodes	Varies, from ~$12,000 (as of 9/25/24) 64 cores, 384 GB RAM, 6.2 TB SSD local scratch, email us for a quote

Getting an Account

- www.chpc.utah.edu -> Documentation -> Getting Started
  - Requires a U of Utah uNID, uses your campus password
  - All user accounts linked to a Principal Investigator
  - Affiliate accounts (uNIDs) available for other universities, collaborators
  - Automated process, requires PI approval, email confirmation
- Account provides:
  - 50 GB free home directory space
  - Login scripts: .bashrc & .custom.sh or .tcshrc & .custom.csh
  - Access to clusters, 50 TB scratch space, windows server, software
  - Subscription to mailing list chpc-hpc-users@lists.utah.edu

Getting an Allocation

- www.chpc.utah.edu -> User Services -> Allocations
  - Provides priority access to notchpeak, redwood clusters
  - Large and Small allocations - applications reviewed each quarter
  - Quick allocations (very small) - 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: www.chpc.utah.edu/usage
- View allocation and cluster access: *myallocation* command

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