

# Research Computing Support at the University of Utah: An Overview of CHPC

Anita Orendt

Assistant Director

Research Consulting & Faculty Engagement

[anita.orendt@utah.edu](mailto:anita.orendt@utah.edu)

## CHPC's Mission

In addition to deploying and operating high performance computational resources and providing advanced user support and training, CHPC serves as an expert team to broadly **support** the increasingly **diverse research computing needs** on campus.

These needs include support for big data, big data movement, data analytics, security, virtual machines, Windows science application servers, protected environments for data mining and analysis of protected health information, and advanced networking.

- CHPC can help if:
  - You need parallel processing
  - You need access to a single high-powered computer
  - You need the ability to run many individual jobs simultaneously
  - You have a large amount of data to store and/or process
  - You need an application you don't have on your computer
  - Your data is IRB-governed PHI
  - You have other computing needs your local resources cannot meet

# Sources of Useful Information

- Getting Started Guide
  - <https://www.chpc.utah.edu/documentation/gettingstarted.php>
- CHPC policies
  - <https://www.chpc.utah.edu/documentation/policies/index.php>
- Cluster Usage Guides
  - <https://www.chpc.utah.edu/documentation/guides/index.php>
- Application Documentation
  - <https://www.chpc.utah.edu/documentation/software/index.php>
- Programming Guide
  - <https://www.chpc.utah.edu/documentation/ProgrammingGuide.php>
- How to Videos
  - <https://www.chpc.utah.edu/documentation/videos/index.php>

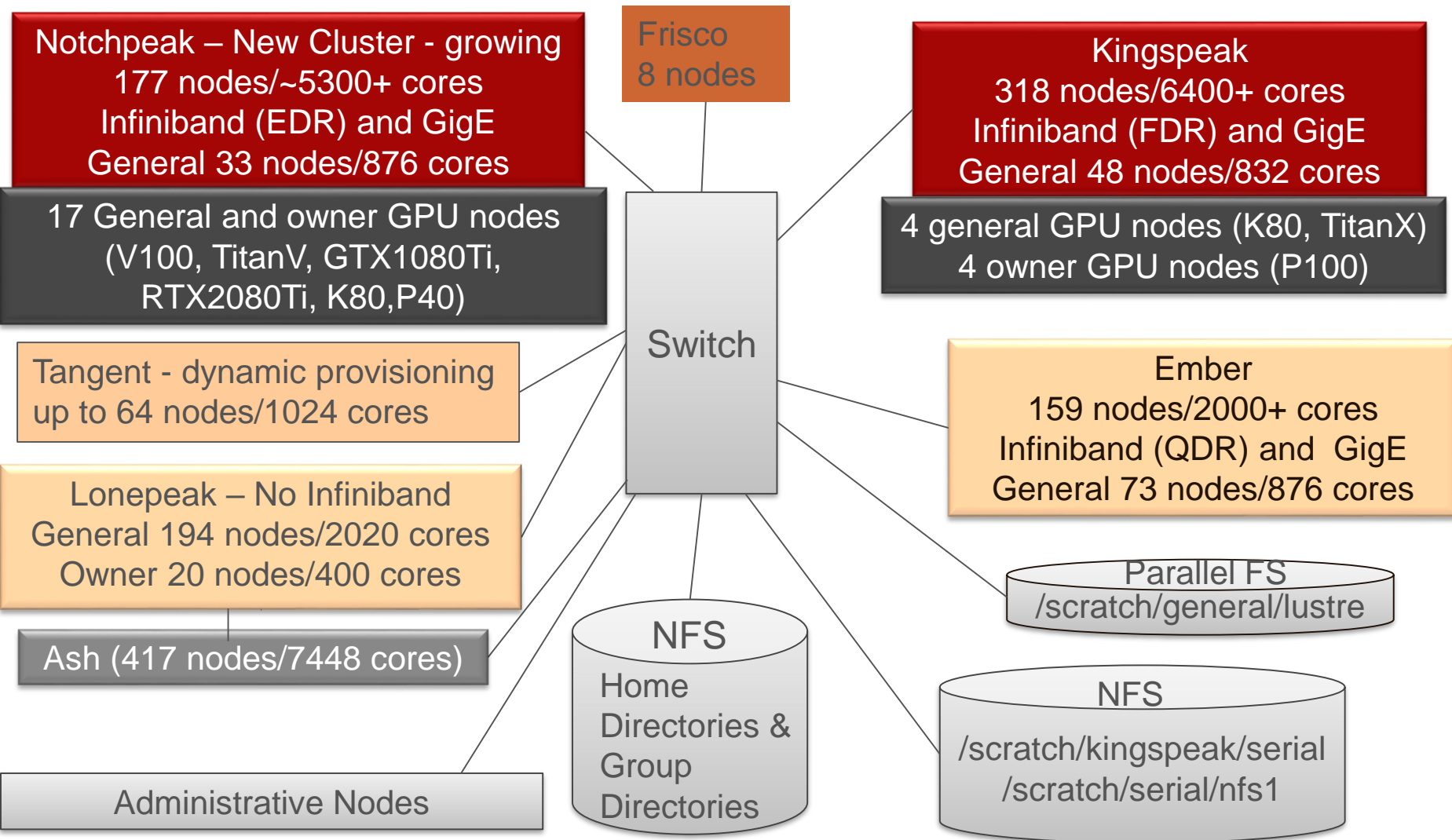
# Downtown Data Center

- Came online Spring 2012
- CHPC completed move to DDC Spring 2013
- Shared with enterprise (academic/hospital) groups
- 92 racks and 1.2MW of power with upgrade path to add capacity for research computing
- Metro optical ring connecting campus, data center, & internet2
- 24/7/365 facility



# CHPC Resources & Services

- **Computational Clusters** – Notchpeak, Kingspeak, Lonepeak, Ember, Ash, Tangent
- **Storage** – home, group, and scratch storage along with tape backup and archive storage options
- **Windows Servers** – mainly statistics usage and windows only applications
- **Virtual Machines** – for needs not met with cluster and windows server
- **Protected Environment** –computational cluster Redwood, storage, VMs, and Windows Server
- **Networking Support** – support compute environment; work with researchers on data movement
- **User Support** – assistance with use of resources; installation of applications; training sessions

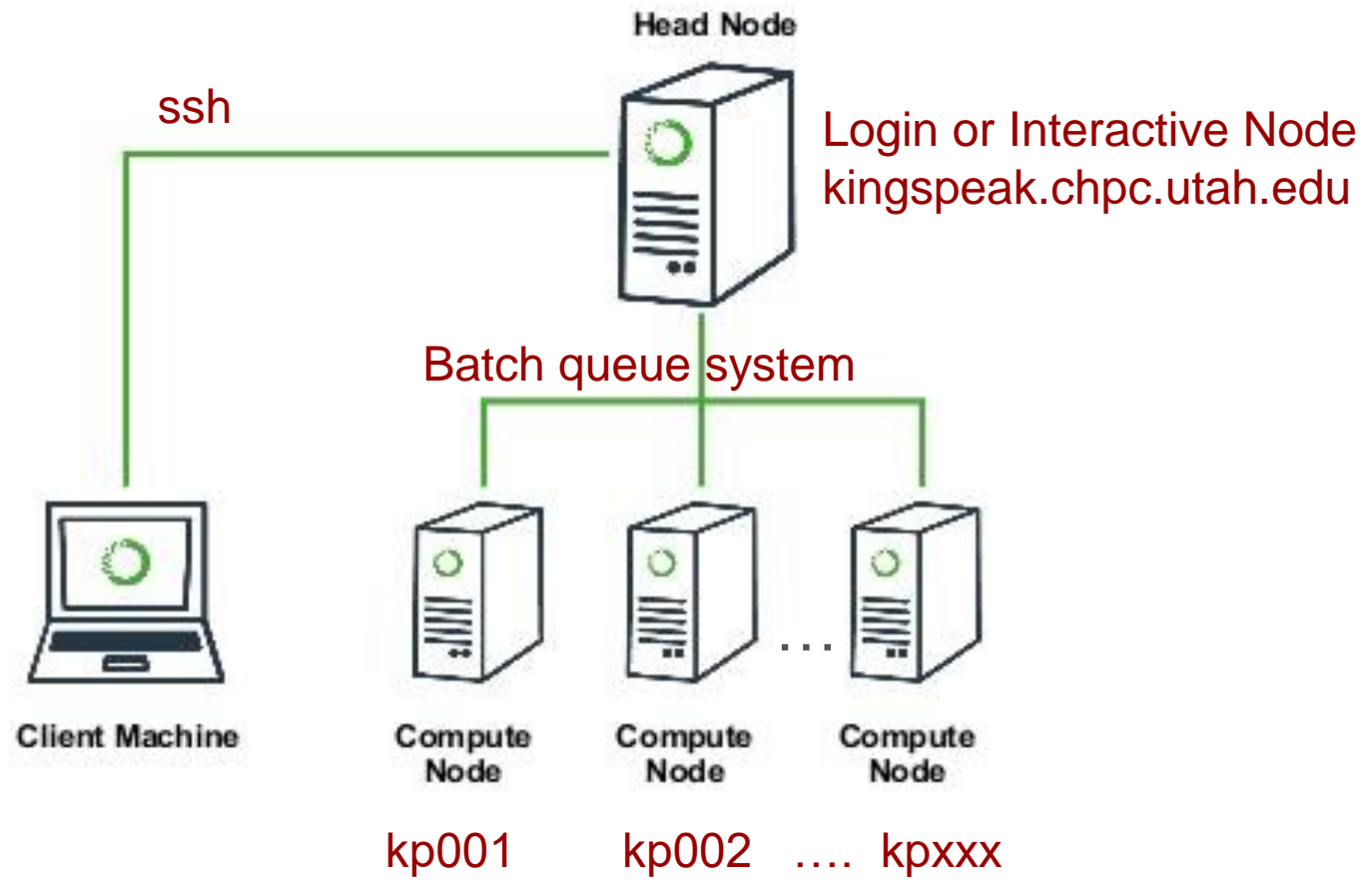


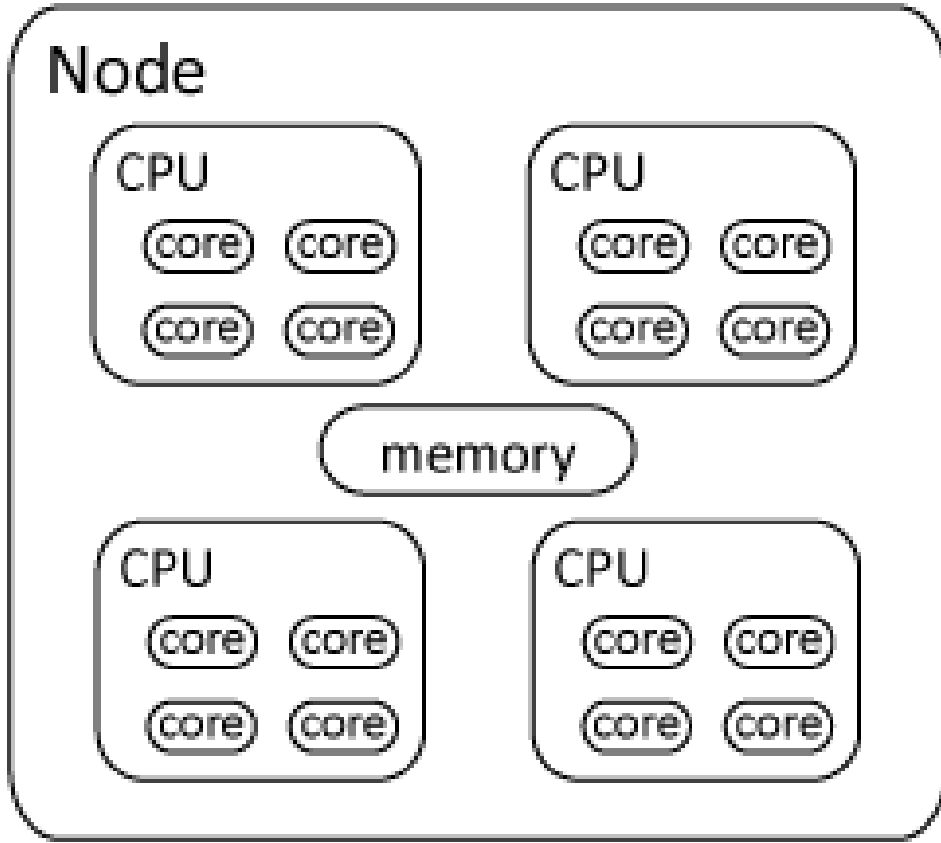






# Cluster Architecture Diagram





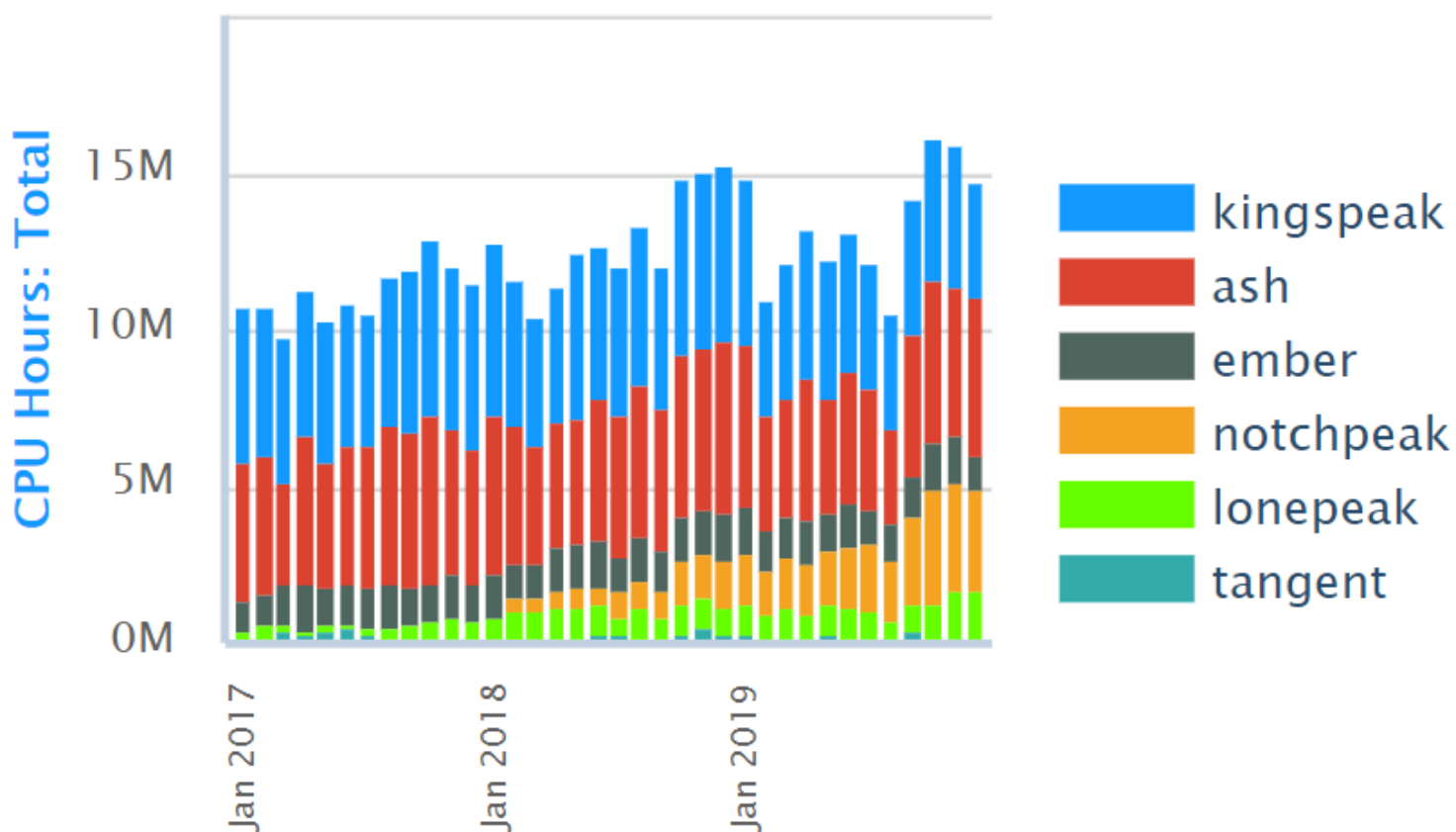
# CHPC Clusters - Condominium Model

- Notchpeak, Kingspeak
  - General resources – Allocation process for cycles
  - Out of allocation – freecycle allowed (with preemption)
- Owner nodes – new purchased added to notchpeak
  - Users from all groups allowed on as owner-guest when not in use (preemption)
  - Found on notchpeak, kingspeak, ember, lonepeak
  - 40 core Intel XeonSP(Cascade) @ 2.1GHz cpu speed, 192GB RAM, 2TB local hard drive, 5 year warranty, EDR IB connectivity @ ~\$6050/node; with 384GB RAM ~\$7700/node
  - 64 core AMD Rome (2.0 GHz base cpu speed), 256 GB memory ~\$6200; with 512 GB, ~\$8600
  - no allocation and no preemption
- Tangent
  - dynamically provisioned cluster resource; no allocation and no preemption
- Ash – Owner cluster
  - All users have guest access as smithp-guest (preemption)

# Core Hour Usage

Over 134M core hours provided in 2017; 160M in 2018; 177M in 2019

Over this time period – over 1300 users from over 375 groups ran more than 11 M jobs



# CHPC Provides Core Infrastructure

- Physical needs (racks, power, cooling)
- Core ethernet and IB fabric; IB to ethernet bridging
- Login/head/management nodes
- Scheduling, batch and allocation systems
- HPC scratch space
- Some application and licensing costs
- High speed data transfer nodes (DTNs)
- 'Science DMZ' network
- CHPC Staff



# Storage Options

- *Home Directories* -- /uufs/chpc.utah.edu/common/home/<uNID>
  - Home directories 50 GB, not backed up
  - Groups can larger home directory per group at \$750/TB for remaining ~3 year warranty; this comes with backup (nightly incremental, weekly full, 2 week retention)
  - Compellent solution - two disk based copies mirrored for HA
- *Group Level File Systems*
  - Group space @\$150/TB
- *Scratch File Systems*
  - For use by all users; scrubbed of files older than 60 days
  - 700 TB Lustre Parallel file system (/scratch/general/lustre)
  - 175 TB NFS mounted file system (/scratch/kingspeak/serial)
  - 600 TB NFS mounted file system (/scratch/serial/nfs1)
- *Disk Based Archive Storage*
  - Archive at\$150/TB

# File Systems

- Access speed based on connectivity
- Local disk fastest – local to each node; varies in size
  - /scratch/local
- Network mounted scratch file systems
  - /scratch/general/lustre
  - /scratch/kingspeak/serial
  - /scratch/serial/nfs1
  - home directories and group spaces (*don't use for large i/o!*)

***Remember NFS mounted spaces – including file systems for group spaces – are a shared resource!***

# Protected Environment

- Refreshed in 2017 with award of NIH Shared instrumentation grant
  - HPC cluster – redwood
  - VM farm – prismatic
  - storage – mammoth, both home and project space; elm for archive object storage
  - windows compute – narwhal
- See <https://www.chpc.utah.edu/resources/ProtectedEnvironment.php>
- Dedicated protected resources for handling of data/projects with protected information
- Currently HIPAA
- Also for projects with other types of sensitive data/restrictions
- Significant area of growth for CHPC
- Preferred location for human genomic – meets NIH dbGaP requirements

# Upcoming Changes

- Adding 32 new nodes, 64 cores each to notchpeak
- Retiring ember, kingspeak going off allocation
- Reworking general nodes on lonepeak
  - Removing the 8 core nodes
  - Adding “new to CHPC” 12 core nodes with either 48 or 96 GB memory
  - Will end up with 192 nodes
- Changes with /scratch/local directory structure
  - Both encrypting content (done)
  - Creation at start of job followed by automatic cleanup at end of job too alleviate issue with having previous jobs impact /scratch/local that is available for use
  - New structure will isolate /scratch/local for individual jobs

# Getting a CHPC Account

- CHPC uses campus uNID and password
- PIs must have account and will need to approve accounts for any members of their research group (can delegate)
- Account Application Procedure – Online process
  - Complete CHPC account form at [https://www.chpc.utah.edu/role/user/account\\_request.php](https://www.chpc.utah.edu/role/user/account_request.php)
  - For collaborators outside of University of Utah must complete affiliate form with HR to get uNID <https://www.hr.utah.edu/forms/affiliate.php> and then use account application



# Security Policies

- No clear text passwords, use ssh and scp
- You may not share your account under any circumstances
- Don't leave your terminal unattended while logged into your account
- Do not introduce classified or sensitive work onto CHPC systems unless on Protected Environment
- Do not try to break passwords, tamper with files etc.
- Do not distribute or copy privileged data or software
- Report suspicions to CHPC ([security@chpc.utah.edu](mailto:security@chpc.utah.edu))
- See <http://www.chpc.utah.edu/docs/policies/security.html> for more details

# Accessing Clusters

- Login or interactive nodes with each cluster
  - **cluster**1.chpc.utah.edu or **cluster**2.chpc.utah.edu where **cluster** is notchpeak, kingspeak, ember, tangent, ash-guest, lonepeak (redwood in PE)
- Interactive nodes only used for short compiles, editing and very short test runs
- ***No more than 15 minutes and no jobs of any length that make heavy use of cpu or memory!***
- Have script which watches running processes and notifies users when in violation of the acceptable usage policy

# Accessing Login nodes

- Use FastX from Mac, Windows, or Linux desktops -- preferred
  - <https://www.chpc.utah.edu/documentation/software/fastx2.php>
- Alternatively:
  - From windows need ssh client
    - PuTTY <http://www.chiark.greenend.org.uk/~sgtatham/putty/>
    - Xshell [http://www.netsarang.com/products/xsh\\_overview.html](http://www.netsarang.com/products/xsh_overview.html)
  - For X forwarding applications also need
    - Xming <http://www.straightrunning.com/XmingNotes/>
      - Look for “mesa” version
  - From mac/linux – use terminal ssh (with -Y for X forwarding)
- Access to protected environment needs Duo two factor authentication (and VPN if off campus)

# FastX – Tool for Remote X

- <https://www.starnet.com/fastx>
- Used to interact with remote linux systems graphically in much more efficient and effective way than simple X forwarding
- Graphical sessions can be detached from without being closed, allowing users to reattach to the session from the same or other systems
- Server on all interactive nodes as well as the frisco nodes; some servers have graphics cards and support OpenGL
- Clients for windows, mac and linux; can be installed on both university and personal desktops.

# FastX

- For FastX – see “To Use” section of documentation at <https://www.chpc.utah.edu/documentation/software/fastx2.php>
- Download client following directions on page
- Do install
- Start program
- Set host to kingspeak1.chpc.utah.edu OR kingspeak2.chpc.utah.edu OR other interactive node OR one of the frisco nodes (frisco1-frisco8.chpc.utah.edu)



# Login scripts

- CHPC provides login scripts (“dot” files) when creating account for both tcsh and bash shells
- These files set the environment so that applications are found, batch commands work – ***Do not remove***
- Choose shell at account creation – can change at [www.chpc.utah.edu](http://www.chpc.utah.edu) (sign in, select edit profile)
- Four files: .bashrc, .tcshrc, .custom.sh, .custom.csh
  - **The first two should not be edited**
  - **The second two is where to add custom module loads**
- Will automatically execute an .aliases file if it exists

# CHPC Uses Modules for Setting Environment

- CHPC provides login scripts (“dot” files) when creating account for both tcsh and bash shells
- These files set the environment so that applications are found, batch commands work – ***Do not remove or edit!***
- <https://www.chpc.utah.edu/documentation/software/modules.php> for information
- Presentation on Modules – Tuesday, February 11

# Batch System Information

- Used to access compute nodes which must be used for any extensive use
- Use SLURM – Simple Linux Utility for Resource Management
- <https://www.chpc.utah.edu/documentation/software/slurm.php>
- Presentation on Slurm – Thursday, February 13

# Software on Clusters

- Have a variety of compilers, mpi packages, math libraries and applications installed
- Some licensing restrictions may apply
- If you need a package we do not currently have installed – ask us!
- Currently we place installations at:
  - /uufs/chpc.utah.edu/sys/installdir
- Have a searchable application database
  - <https://www.chpc.utah.edu/software/chpc/>

# Allocation

- General Allocation Process Information
  - <https://www.chpc.utah.edu/documentation/policies/1.4AllocationPolicies.php>
- Regular allocation form
  - [https://www.chpc.utah.edu/apps/profile/allocation\\_form.php](https://www.chpc.utah.edu/apps/profile/allocation_form.php)
  - Requests due Sept 1, Dec 1, Mar 1, and Jun 1
  - Allocation in core hours
- Quick allocation
  - [https://www.chpc.utah.edu/apps/profile/allocation\\_quick\\_form.php](https://www.chpc.utah.edu/apps/profile/allocation_quick_form.php)
- Check usage -- <https://www.chpc.utah.edu/usage/cluster/current-project-general.php>
- Simplified quick allocation requests & general allocation requests for up to 20,000 core-hours per quarter

# Windows Statistics Server

- Beehive – refreshed 2019
  - 48 physical cores, 512TB memory
- Presently has the following software installed
  - SAS 9.4 with text miner
  - SPSS
  - R
  - STATA
  - Mathematica
  - Matlab
- *If you need other software, please contact us to discuss*

# Virtual Machine Farm

- For needs and applications that do not fit in compute cluster or Windows server
- Multiple VM servers with failover – hardware refreshed 2019
- VM storage
- Have community mysql/mssql VMs, git repositories, web servers, etc
- New – user VMs (not use of community ones) will have a cost, both for the VM and for any customization needed.

Blocks	RAM (GB)	Cores	Storage (GB)	Price
1	4	2	50	\$475
2	8	2	100	\$705
4	16	4	200	\$1175
8	32	8	400	\$2115
16	64	8	800	\$3995

*Additional VM storage available, in 100GB increments, at a cost of \$850/TB.*



# Spring 2020 Presentation Series

In INSCC Auditorium  
can join remotely via zoom

DATE	TIME	PRESENTATION TITLE	PRESENTER
Thursday, January 23	1-2pm	<a href="#"><u>Overview of CHPC</u></a>	Anita Orendt
Tuesday, January 28	1-3pm	<a href="#"><u>Hands on Introduction to Linux, part 1</u></a>	Brett Milash & Wim Cardoen
Thursday, January 30	1-3pm	<a href="#"><u>Hands on Introduction to Linux, part 2</u></a>	Brett Milash & Wim Cardoen
Tuesday, February 4	1-3pm	<a href="#"><u>Hands on Introduction to Linux, part 3</u></a>	Wim Cardoen & Brett Milash
Thursday, February 6	1-3pm	<a href="#"><u>Hands on Introduction to Linux, part 4</u></a>	Wim Cardoen & Brett Milash
Tuesday, February 11	1-2pm	<a href="#"><u>Module Basics</u></a>	Anita Orendt
Thursday, February 13	1-2pm	<a href="#"><u>Slurm and Slurm Batch Scripts</u></a>	Anita Orendt
Tuesday, February 18	1-3pm	<a href="#"><u>Hands-on Introduction to Python, Part 1</u></a>	Brett Milash & Wim Cardoen
Thursday, February 20	1-3pm	<a href="#"><u>Hands-on Introduction to Python, Part 2</u></a>	Brett Milash & Wim Cardoen
Tuesday, February 25	1-3pm	<a href="#"><u>Hands-on Introduction to Python, Part 3</u></a>	Brett Milash & Wim Cardoen

DATE	TIME	PRESENTATION TITLE	PRESENTER
Thursday, February 27	1-3pm	<u><a href="#">Numpy, pt 1 (Hands-on Intro to Python, Pt 4)</a></u>	Wim Cardoen & Brett Milash
Tuesday, March 3	1-3pm	<u><a href="#">Numpy, pt 2 (Hands-on Intro to Python, Pt 5)</a></u>	Wim Cardoen & Brett Milash
Thursday, March 5	1-3pm	<u><a href="#">Hands-on Introduction to Open OnDemand</a></u>	Martin Cuma
Tuesday, March 17	1-3pm	<u><a href="#">Introduction to Containers</a></u>	Martin Cuma
Thursday, March 19	1-3pm	<u><a href="#">Introduction to R</a></u>	Wim Cardoen & Brett Milash
Tuesday, March 24	1-3pm	<u><a href="#">Workflows using Snakemake</a></u>	Brett Milash
Thursday, March 26	1-2pm	<u><a href="#">Open Science Grid</a></u>	Wim Cardoen
Tuesday, March 31	1-3pm	<u><a href="#">Introduction to Parallel Computing</a></u>	Martin Cuma
Thursday, April 2	1-3pm	<u><a href="#">Introduction to Profiling</a></u>	Martin Cuma
Tues&Wed, April 7-8	9am-3pm	<u><a href="#">XSEDE HPC Monthly Workshop: Big Data</a></u>	Martin Cuma (local host)
Thursday, April 9	1:30-4:30pm	Optimizing and Accelerating Your MATLAB Code	Mathworks
Tuesday, April 14	1-2pm	<u><a href="#">Introduction to Debugging</a></u>	Martin Cuma

**<https://www.chpc.utah.edu/presentations/Spring2020CHPCPresentationSchedule.php>**

If you would like training for yourself or your group, CHPC staff would be happy to accommodate your request. Please contact [anita.orendt@utah.edu](mailto:anita.orendt@utah.edu) or [helpdesk@chpc.utah.edu](mailto:helpdesk@chpc.utah.edu)

# Getting Help

- CHPC website
  - [www.chpc.utah.edu](http://www.chpc.utah.edu)
    - Getting started guide, cluster usage guides, software manual pages, CHPC policies
- Ticketing System
  - Email: [helpdesk@chpc.utah.edu](mailto:helpdesk@chpc.utah.edu)
- Help Desk: 405 INSCC, 581-6440 (9-6 M-F)
- Mailing Lists:  
[chpc-hpc-users@lists.utah.edu](mailto:chpc-hpc-users@lists.utah.edu) used to send messages to users