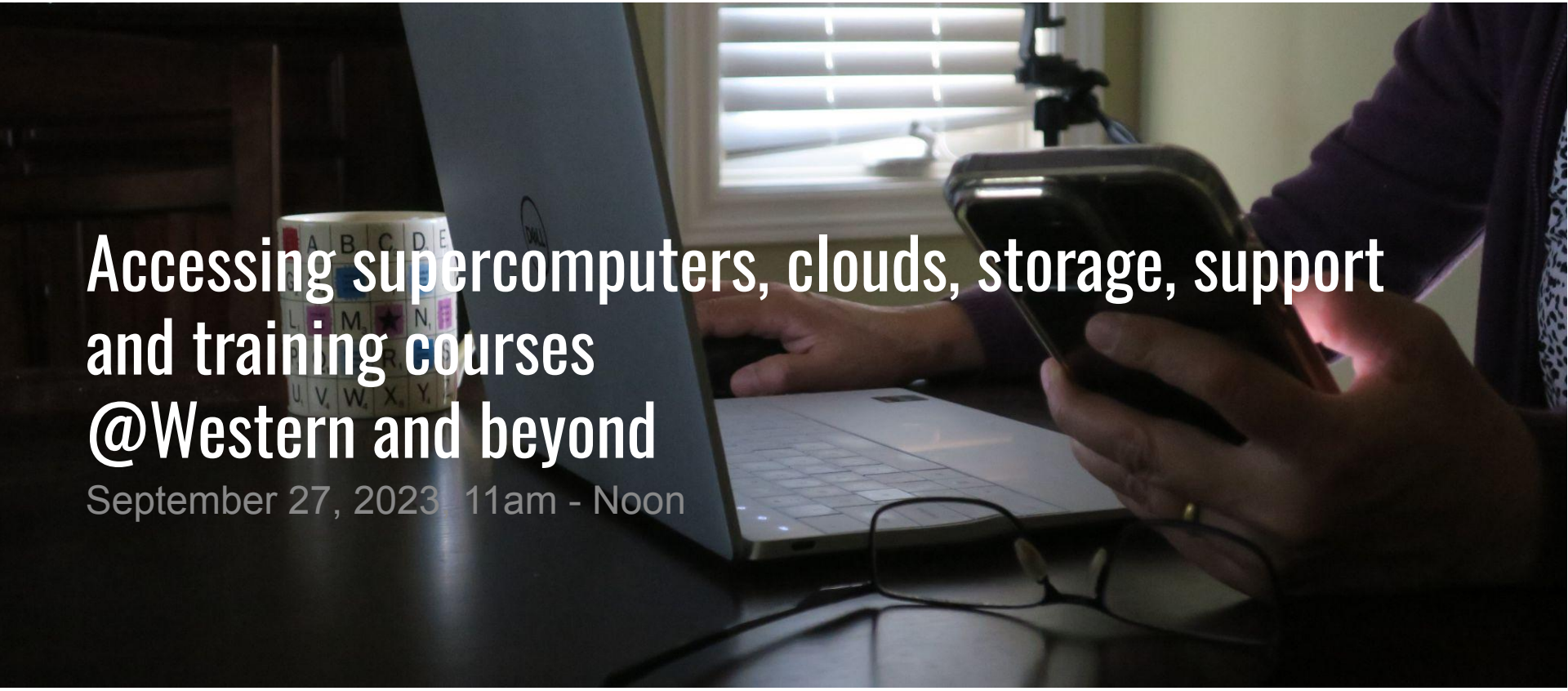
Western UNIVERSITY · CANADA

Digital Research Alliance of Canada



Accessing supercomputers, clouds, storage, support and training courses @Western and beyond

September 27, 2023 11am - Noon

SHARCNET™

Shared Hierarchical Academic Research Computing Network

Supercomputing at Western, SHARCNET and beyond

- Supercomputing at Western, SHARCNET and beyond
  - Why do you need supercomputers
  - What are available
  - How to access supercomputers (clusters) and clouds
  - Accessing and managing files
  - Running programmes - *submitting jobs*
  - Where to get help
- PIs applying for compute, storage and cloud resources
- What every graduate student should know
- Introduction to advanced research computing courses
- Q & A

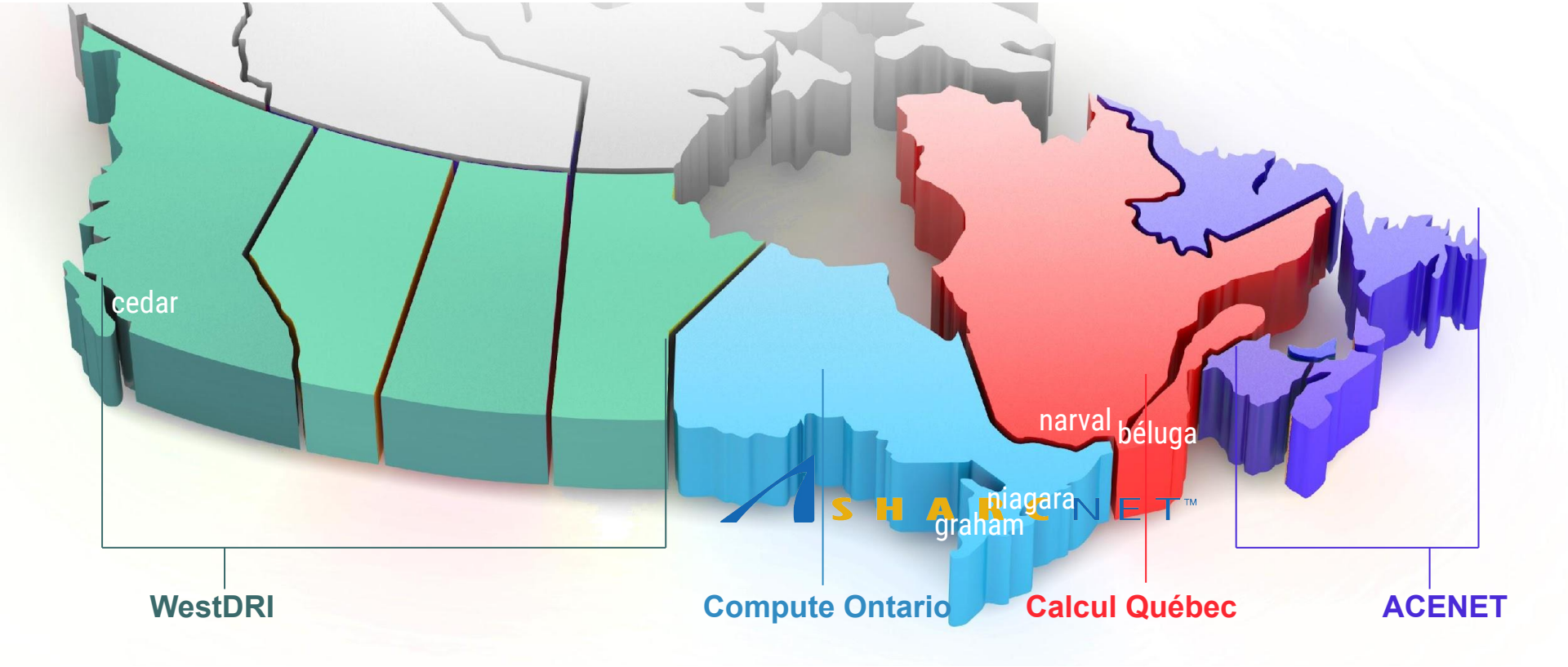
Supercomputing at Western, SHARCNET and beyond

Why use supercomputing resources

- You do not have many cores or much memory;
- You do not have the GPUs of the type needed;
- You need a huge amount of disk space, e.g. hundreds of TB;
- You need to run large scale of simulations that need hundreds of cores;
- You need to run large amount of simulations concurrently instead of one after another;
- You need to run a web services;
- You need to run a SQL database;
- You need to run programs on a cloud;
- All these services are free

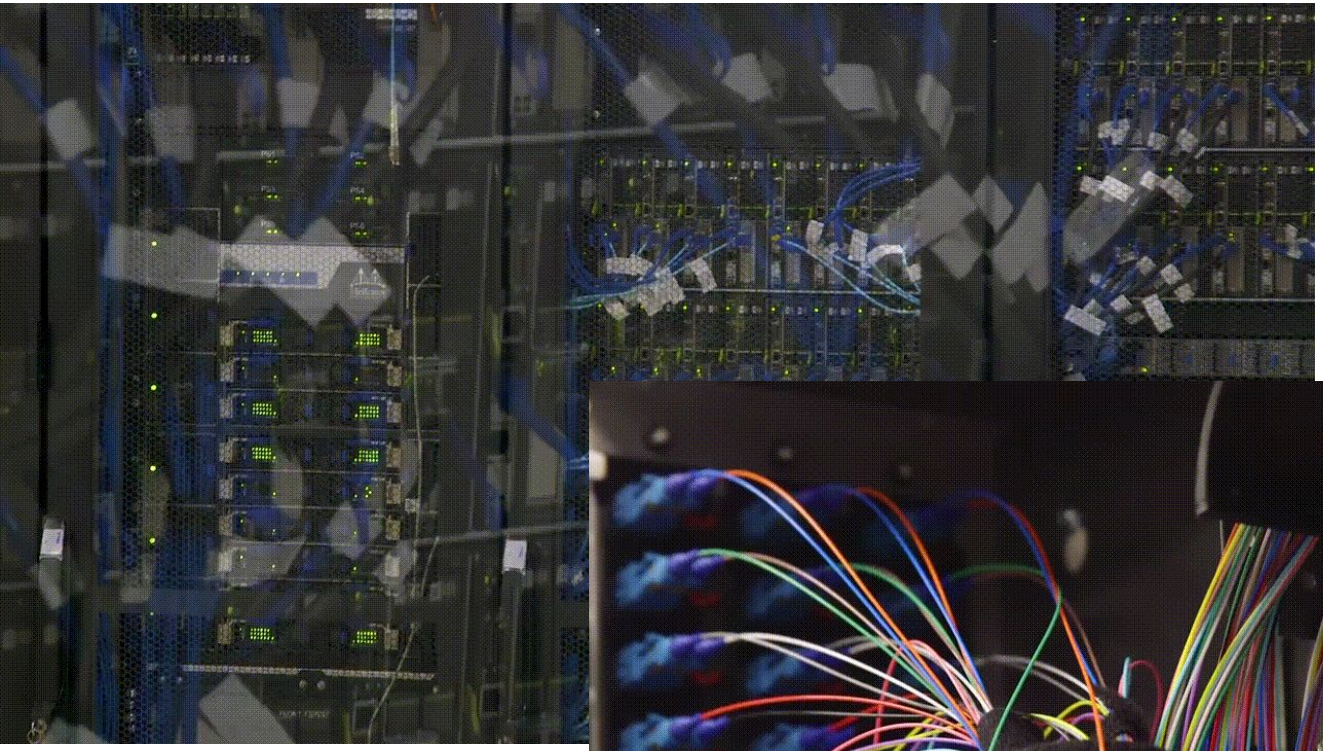
Digital Research Alliance of Canada

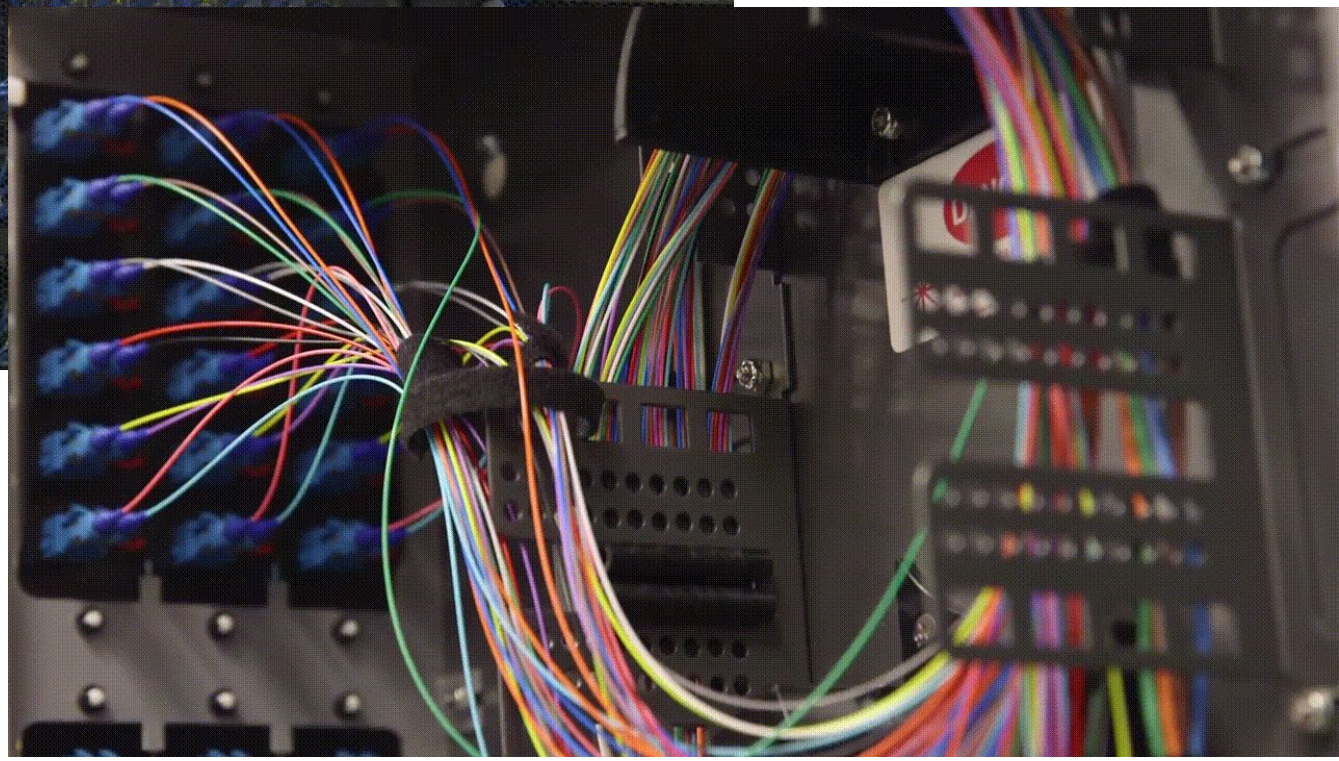
Single account...



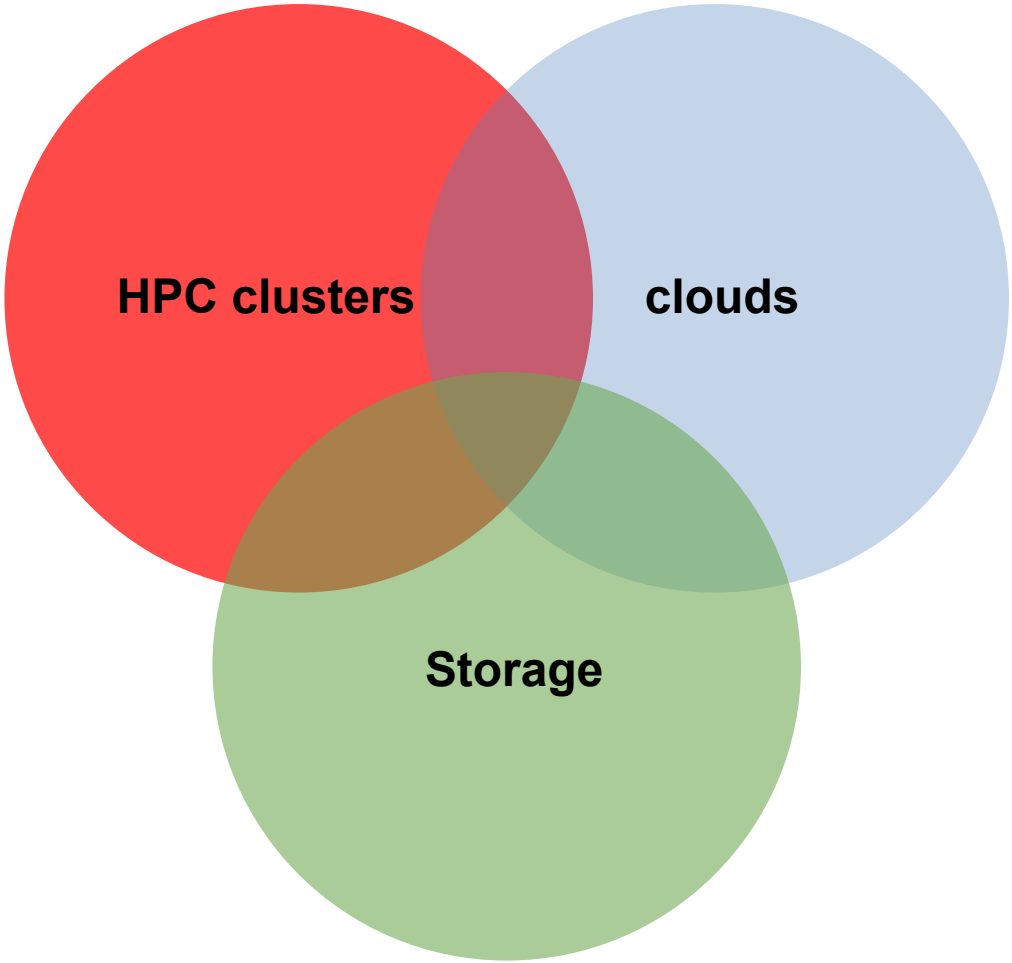
One can access all national supercomputers across the country, for free.

Digital Research Alliance of Canada





Supercomputing at Western, SHARCNET and beyond

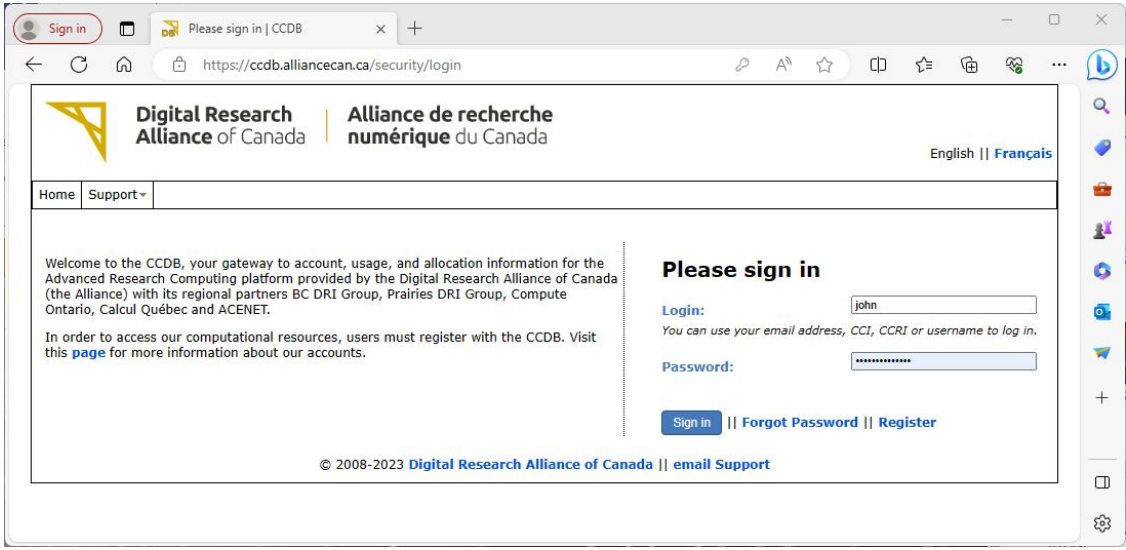


Supercomputing at Western, SHARCNET and beyond

Sign up for an account for FREE at

https://ccdb.alliancecan.ca/

- Your supervisor should have an account.
- Students, postdoc, visiting scholars and other research staff can sign up for an account with supervisor's role ID (CCRI)
- This account allows you to access all the supercomputers and clouds across the country.
- It's FREE.



Multi-factor authentication (MFA) is being introduced, currently to certain groups on a voluntary basis.

Supercomputing at Western, SHARCNET and beyond

Clusters across the country

- cedar.alliancecan.ca (**94,528c**)
- graham.alliancecan.ca (**41,548c**)
- niagara.alliancecan.ca (**80,640c**)
- beluga.alliancecan.ca (**39,120c**)
- narval.alliancecan.ca (**80,912c**)

Cloud services

- arbutus.cloud.alliancecan.ca
- cedar.cloud.alliancecan.ca
- graham.cloud.alliancecan.ca
- east.cloud.alliancecan.ca

GPUs

- P100, V100, A100, T4

Storage space

- /home 50G, backed up.
- /project 1T per group, up to 40T by request; backed up.
- /scratch 20T per user, up to 200T by request; old files are removed in 60 days.
- /nearline (tapes)

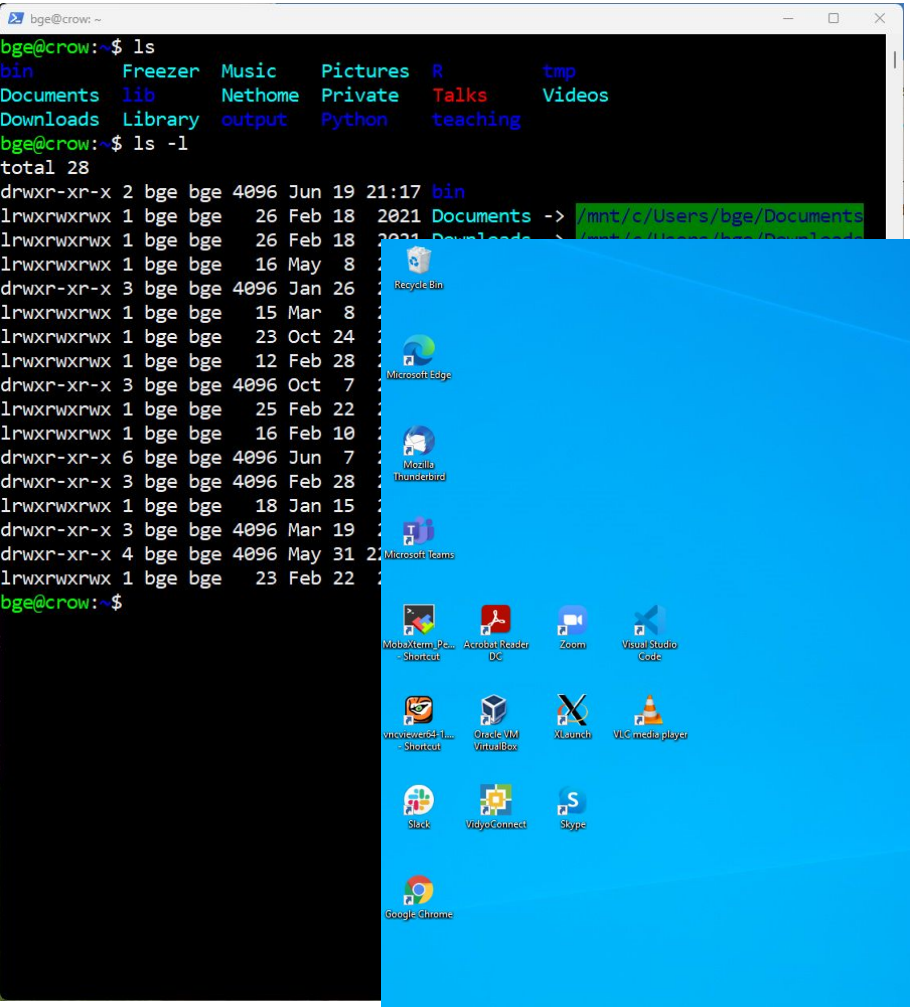
Supercomputing at Western, SHARCNET and beyond

Cluster computing environment

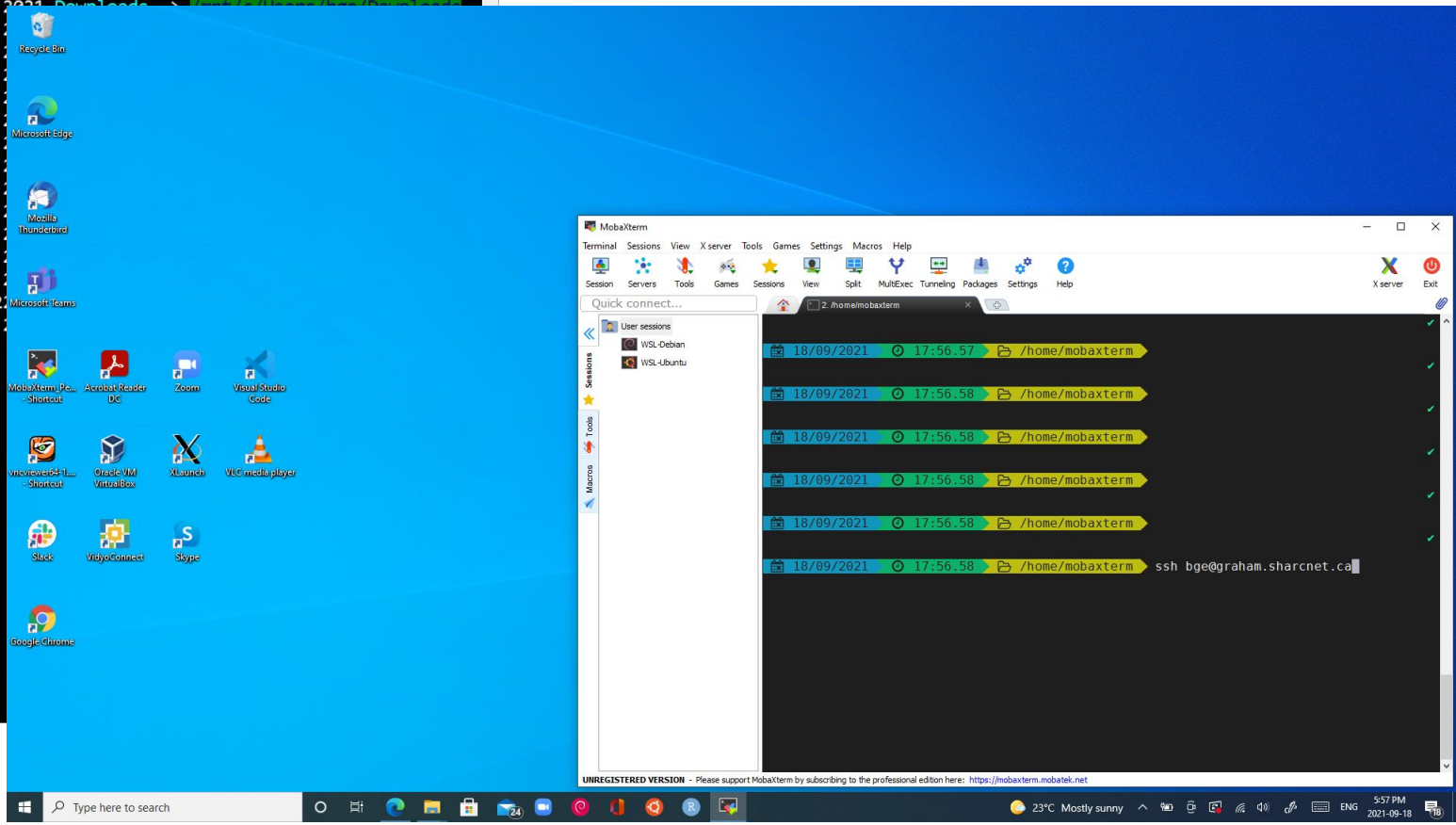
- OS: 64-bit Linux (CentOS)
- Languages supported: Python, R, C/C++, Fortran, Matlab, Java, Julia, CUDA, etc.
- Access to a variety of software packages
- Parallel development support:
  - MPI, Chapel: Distributed memory systems (cross compute nodes) and shared memory system (single node)
  - OpenMP, Pthreads: Multithreading, within a single node
  - CUDA, OpenACC, OpenCL: GPUs and other accelerators on chip
  - C++: Language support for multithreading (since C++-11 standard)
  - Fortran: Language support for parallel programming (since 2003 standard)
  - Julia: Parallel processing constructs, shared and distributed objects
- Data science support:
  - R, Python, Julia, Spark, DASK, etc.
- You must learn how to run programmes in batch via job scheduler slurm

Connecting to clusters via SSH

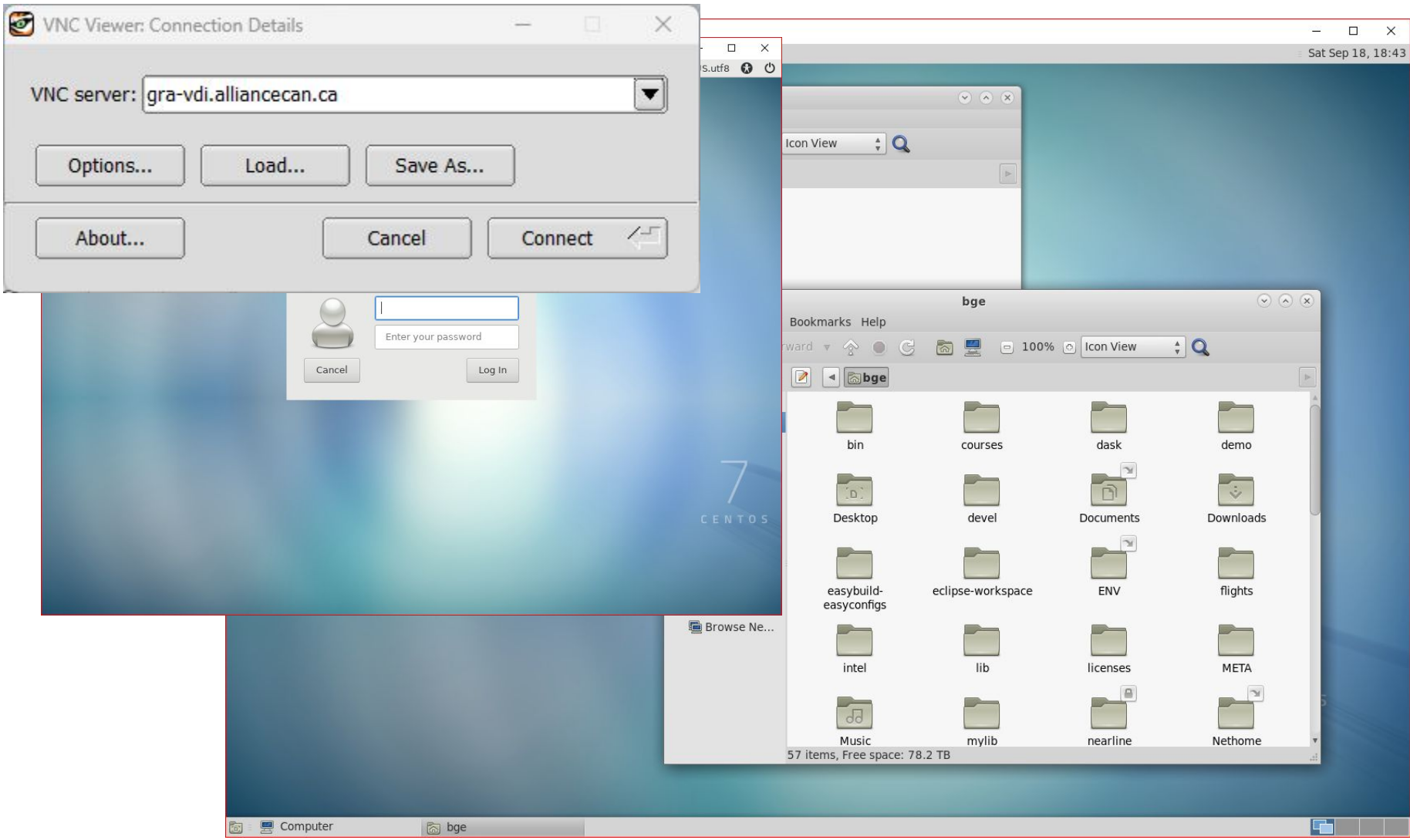
via MobaXterm or Windows Subsystem for Linux (WSL)



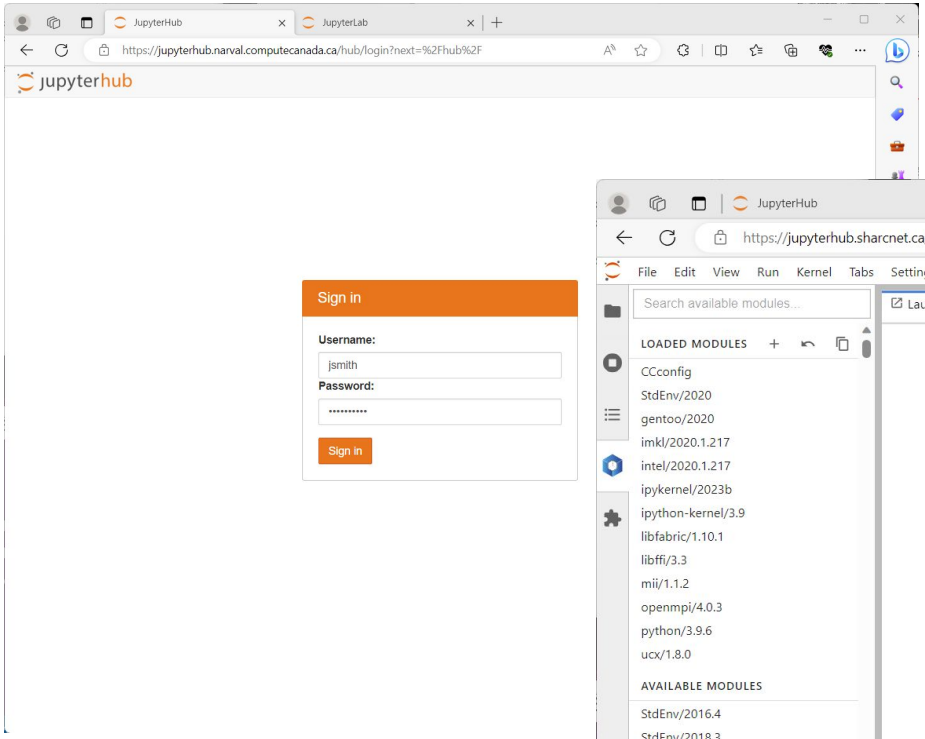
via Linux terminal



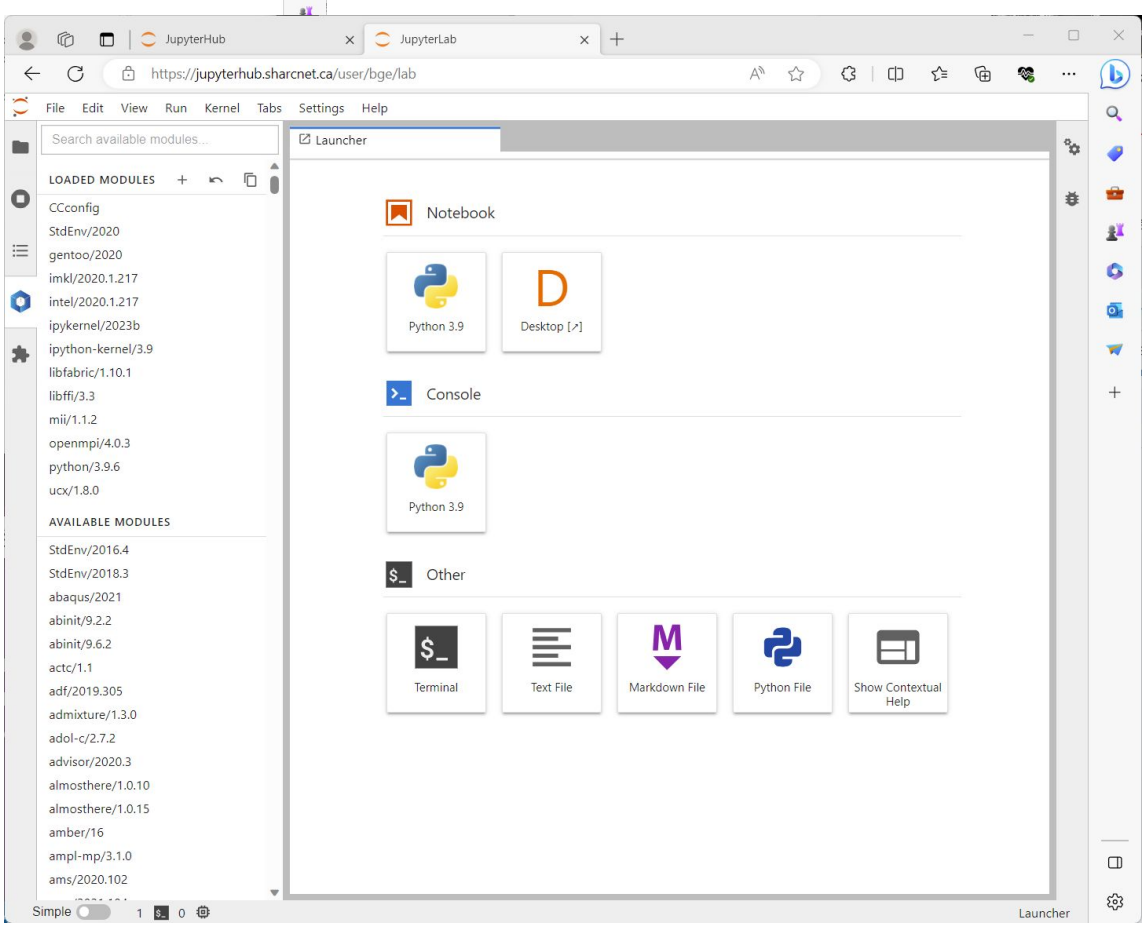
Connecting to GUI desktop



Connecting to clusters via jupyterhub

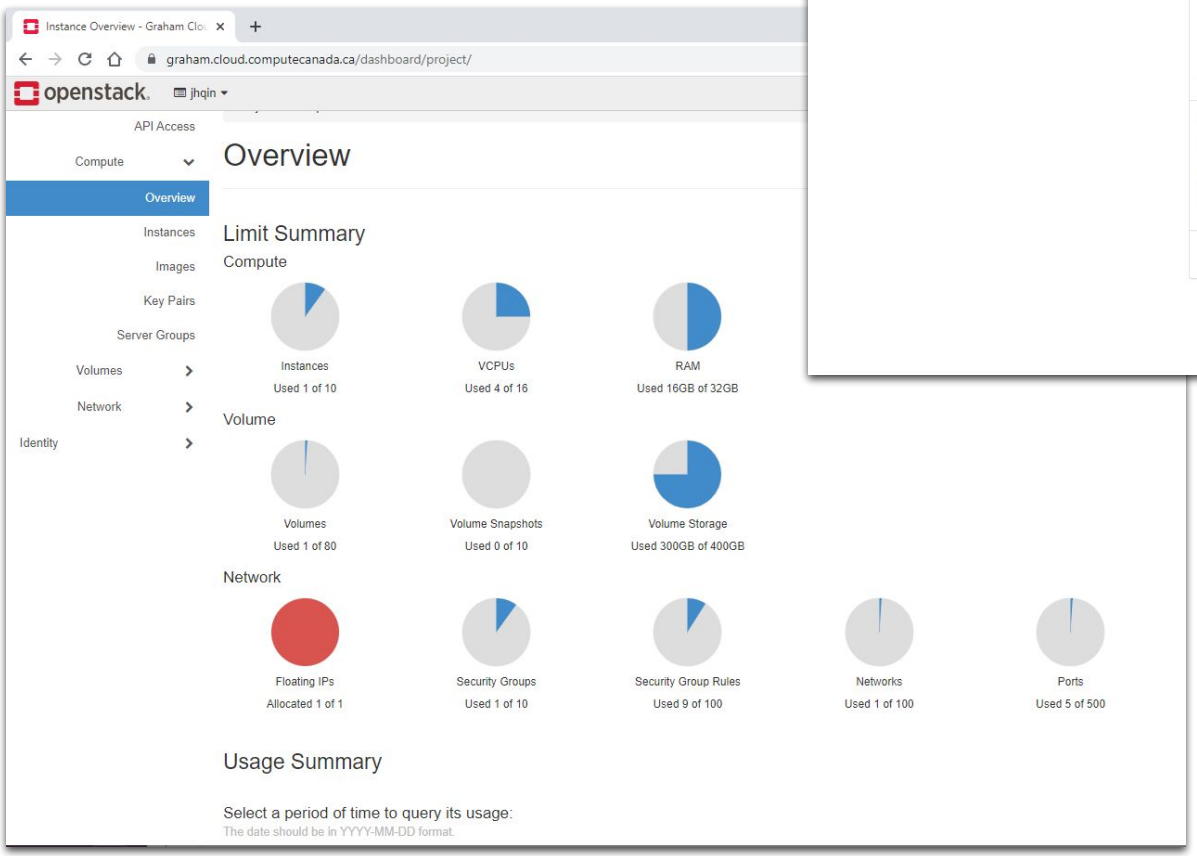


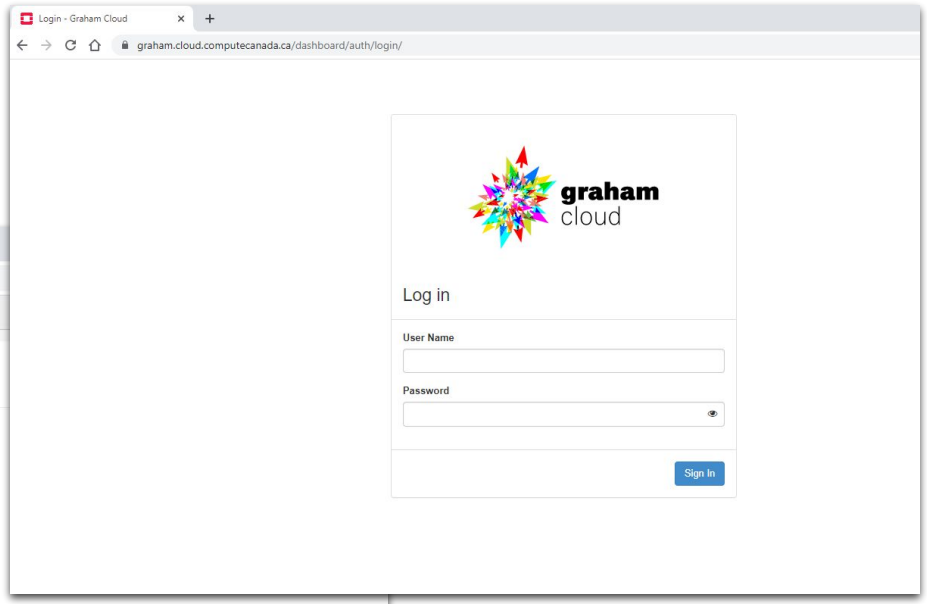
Access to the cluster via jupyter notebook, terminal and desktop in a browser.



Connecting to cloud

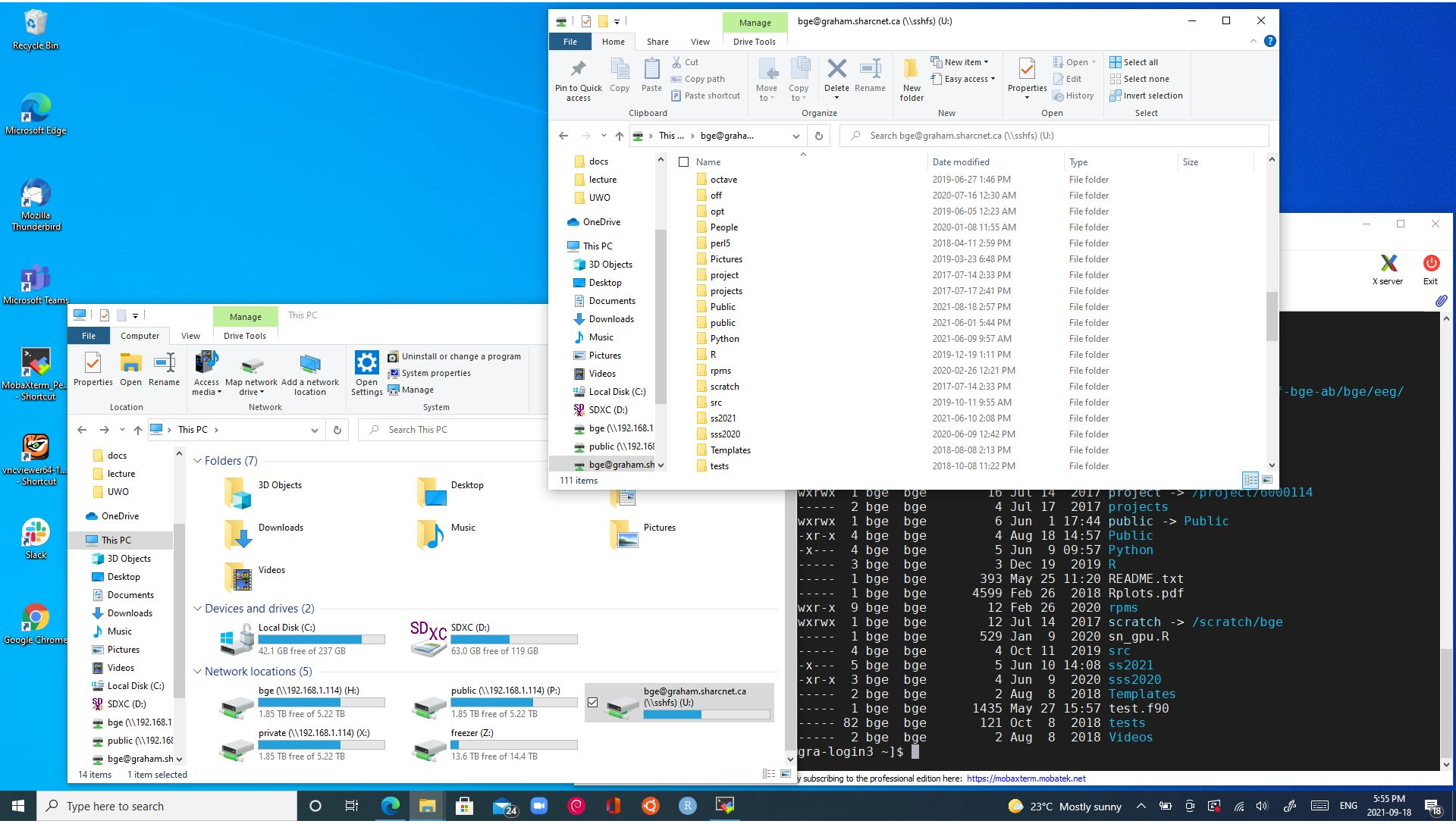
- arbutus.cloud.alliancecan.ca
- cedar.cloud.alliancecan.ca
- graham.cloud.alliancecan.ca
- east.cloud.alliancecan.ca





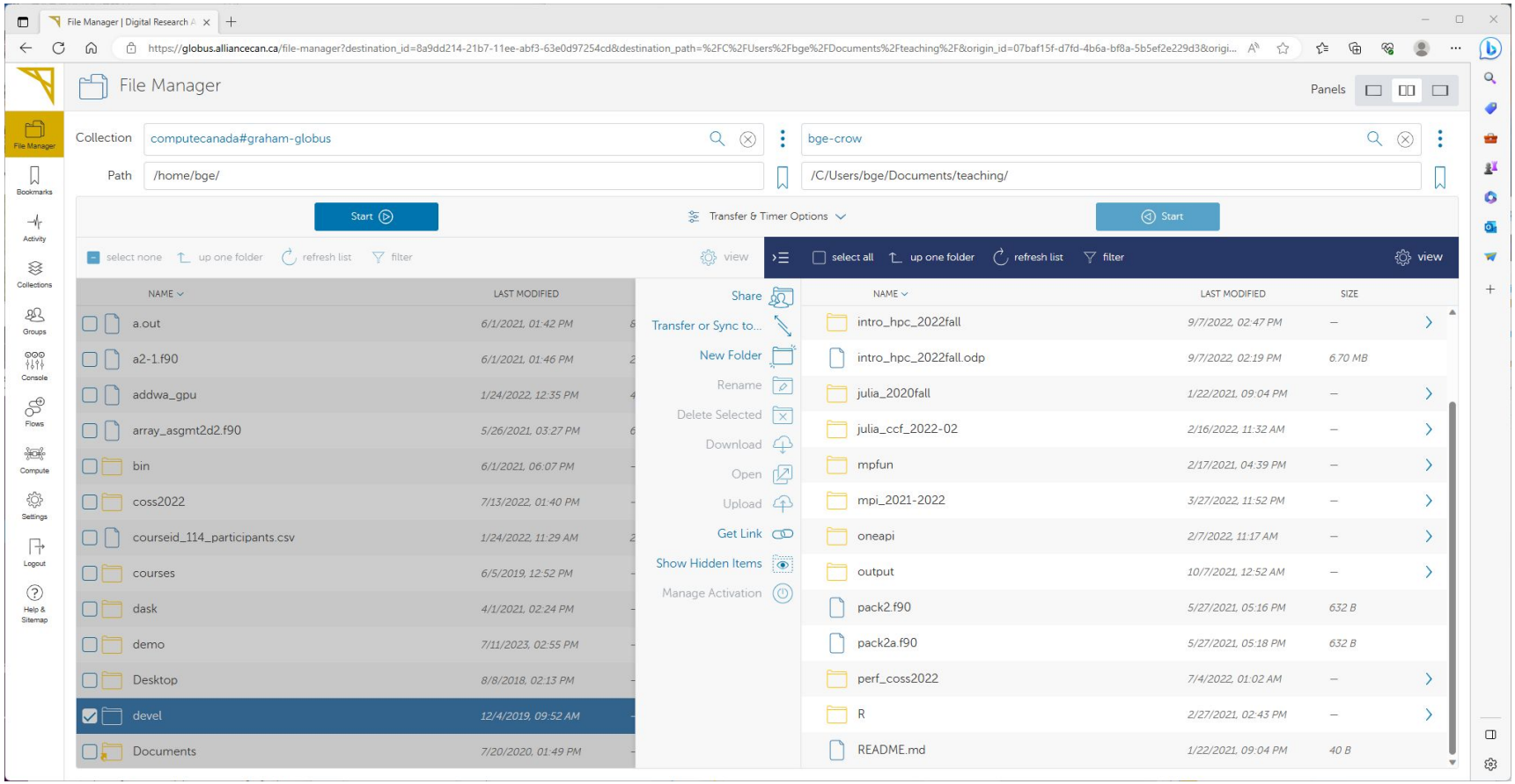
- A cloud project account is required
- Multiple cloud sites are available

Accessing files on remote systems



Transferring large amounts of files using Globus

Check https://docs.alliancecan.ca/wiki/Globus. Go to https://globus.alliancecan.ca/ and follow the instructions

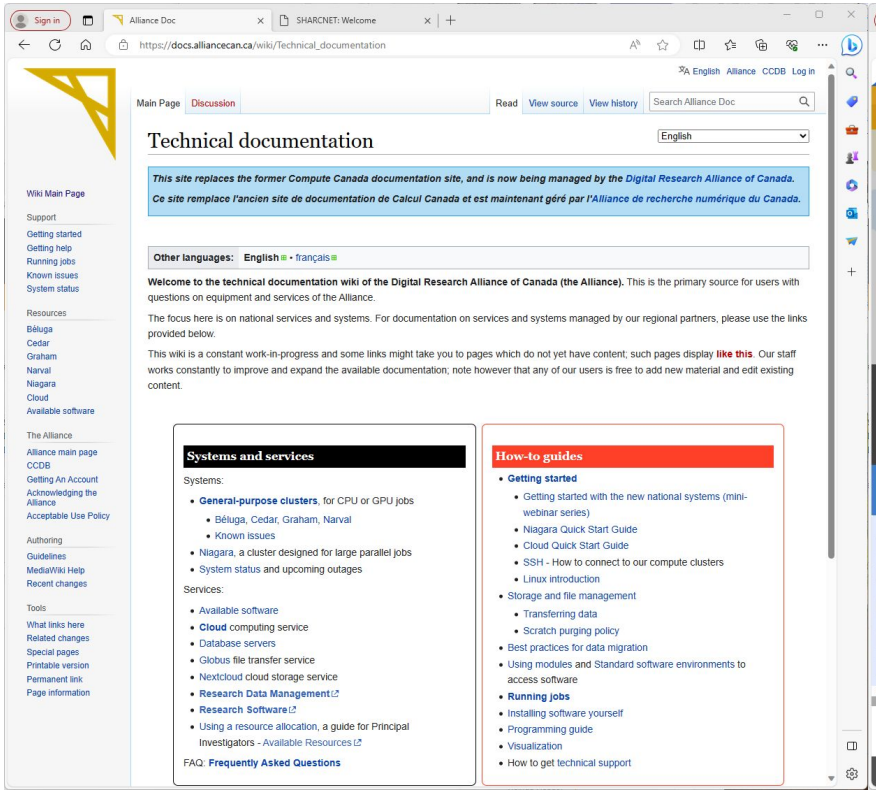


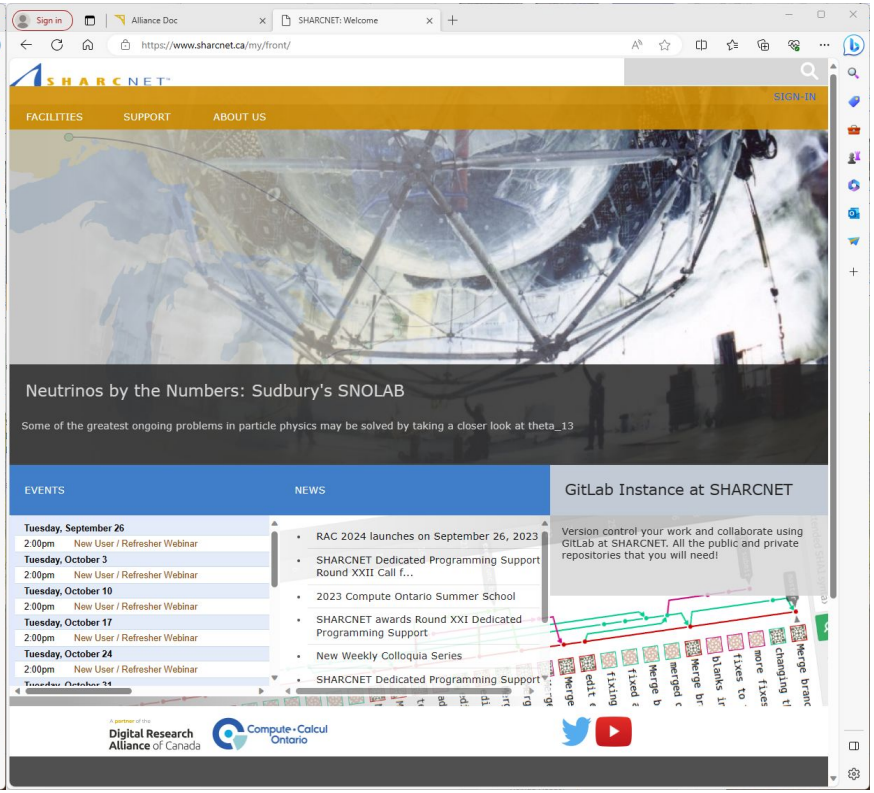
Western University · Canada

SHARCNET™

Digital Research Alliance of Canada

Where to look for information and get help





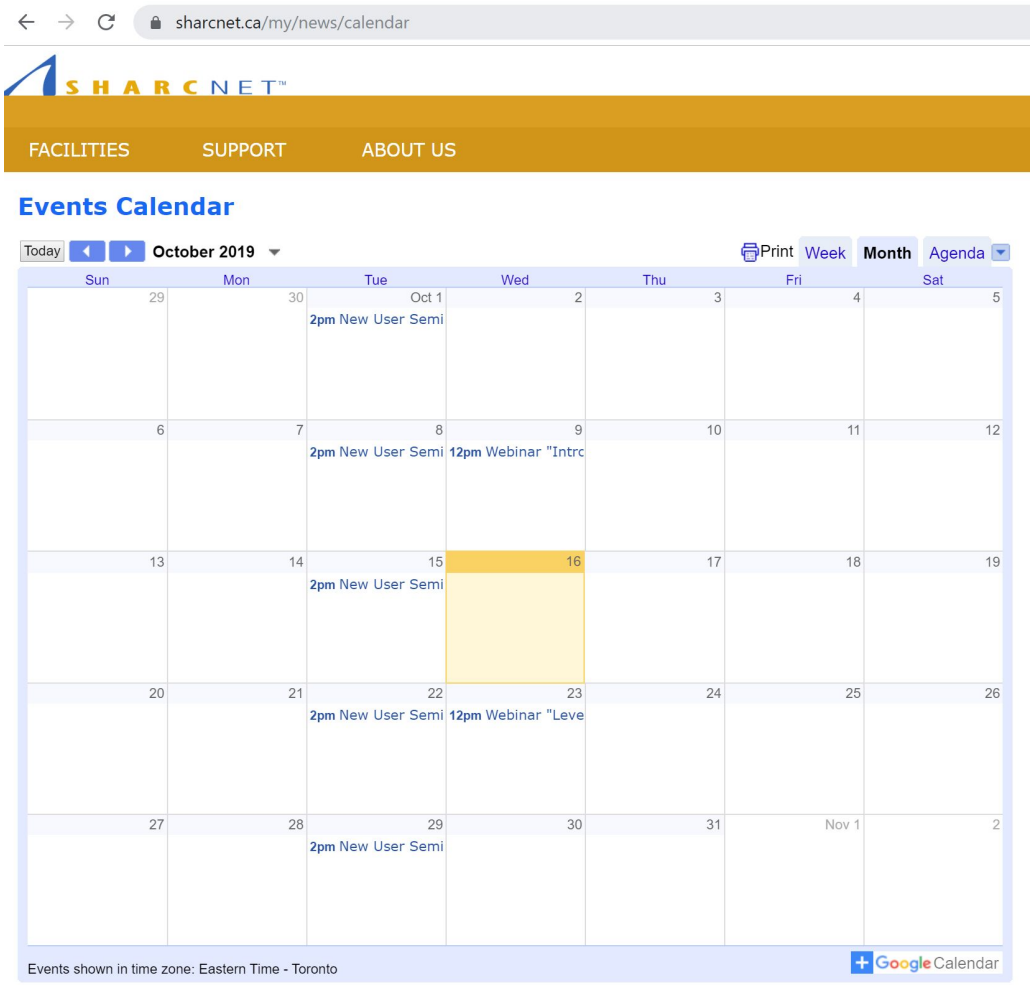
https://docs.alliancecan.ca/

https://www.sharcnet.ca/

Where to look for information and get help

Online

- New user seminar every Tuesday at 2pm Eastern time.
- Bi-weekly general interest seminars at noon on Wednesday.
- @SHARCNET
- youtube.sharcnet.ca

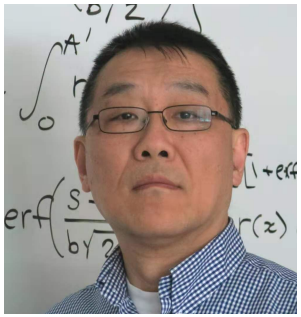


Where to look for information and get help

Local staff

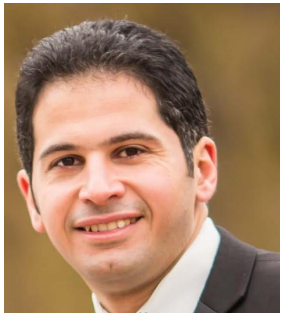
Tyson Whitehead, *HPC*, Math, Stats, EE

Jinhui Qin, *HPC*, Big Data, CS

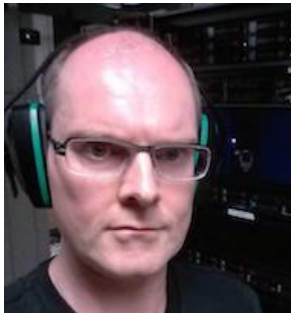


Ge Baolai, *HPC*, Applied Math

Doug Roberts (WLU), *HPC*, CFD, Commercial Software



Mohamed Elsakhawy, *Sysadmin*, Cloud, CS



Fraser McCrossan, *Sysadmin*, CS

Use of systems
Installation of software
Access to commercial software and site licence
Debugging and optimizing code
Programming
Consultation on various research problems
Grant application for compute hardware
… …

Where to look for information and get help

Interactive help

- Weekly new user seminar: https://www.sharcnet.ca/my/news/calendar
- Ticketing system (most recommended): support@tech.alliancecan.ca
- Staff contact info to email or phone: https://www.sharcnet.ca/
- Arrange an office visit

Use of systems
Installation of software
Access to commercial software and site licence
Programming
Debugging and optimizing code
Consultation on various research problems
Grant application for compute hardware
...

Where to look for information and get help

Dedicated programming support

- Staff spending 50% of time working with the PI on specifically defined programming tasks.
- The DP programme runs for about 4 months.
- There are two to three calls a year for PIs for apply.
- The applications are reviewed based on the scientific merits and the feasibility of the proposed programming project.

Where to look for information and get help

Local training events, workshops

- Local workshops
- Annual summer school - week long, multi-streams, many courses, mostly hands-on.
- Online, in-person/self-paced learning training course

Supercomputing at Western, SHARCNET and beyond

- Supercomputing at Western, SHARCNET and beyond
- PIs applying for compute, storage and cloud resources
- What every graduate student should know
- Introduction to advanced research computing courses
- Q & A

Resource allocation competition background

- In principle, researchers are expected to have well defined projects to access resources funded by CFI. CFI requires the allocation of resources be competitive.
- The majority of the resources are allocated through resource allocation competition (RAC) process, the remaining portion is for opportunistic access aka *default allocation* or RAS.
- With RAC allocations, users' queued jobs may start sooner.
- The RAC applications are peer reviewed (scientific and technical reviews).

RAC (cont'd)

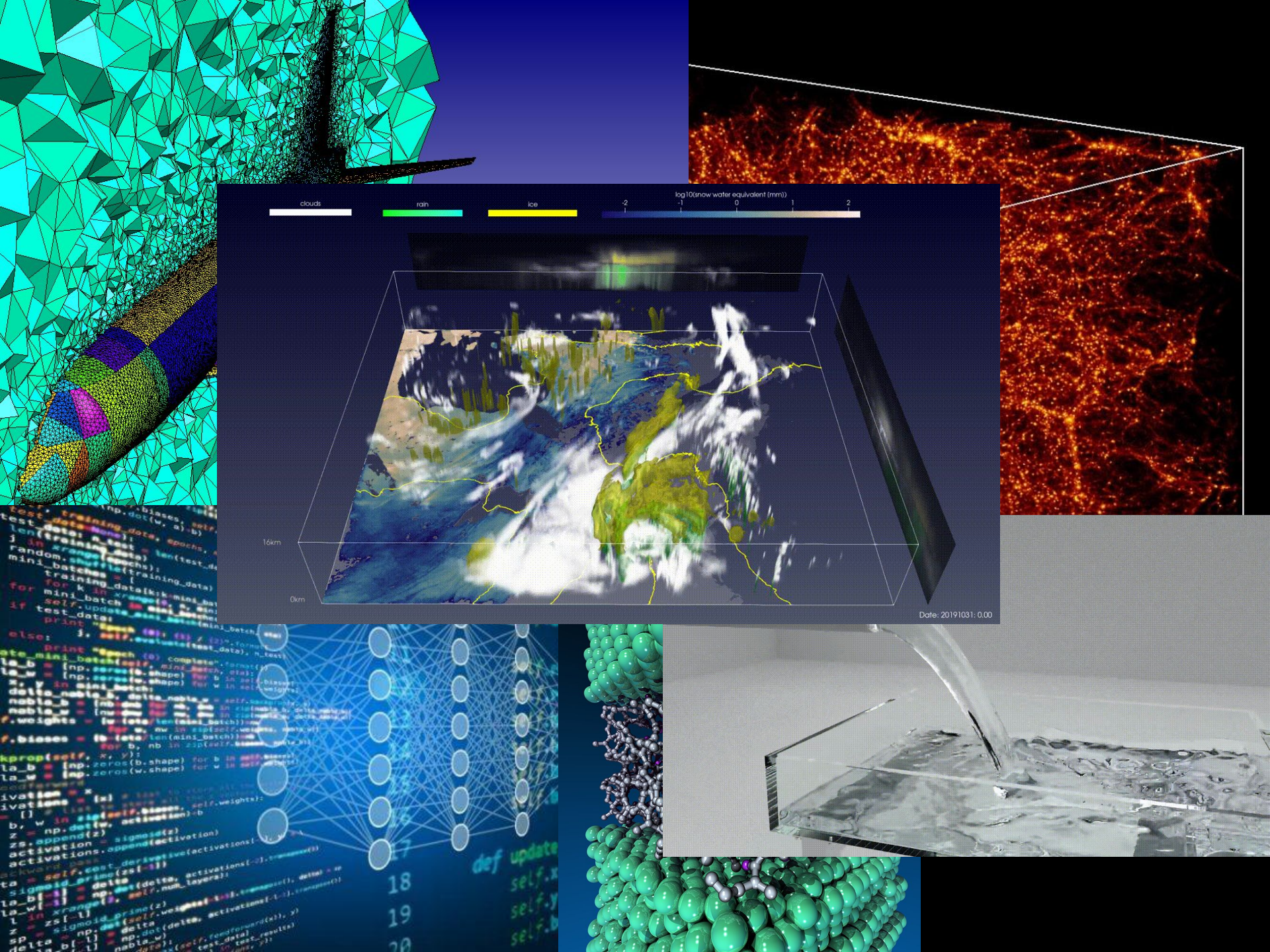
- Resource allocations include compute - CPU/GPU - and storage
- Currently two competitions:
  - Resources for Research Groups (RRG)
  - Research Platforms and Portals (RPP)
- Minimum requirements:
  - CPU ≥ 200 core years (Cy) per cluster
  - GPU ≥ 25 Reference GPU units (RGUs) per cluster
  - Project storage ≥ 41 TB, or,
  - Nearline storage ≥ 101 TB, or,
  - Persistent cloud storage ≥ 1q TB, or,
  - Compute cloud ≥ 81 vCPUs, or,
  - Persistent cloud ≥ 26 vCPUs.
- Requested resources are subject to scale back based on the science score and the amount asked.
- Check the Alliance's documentation for the official numbers and facts.

RAC (cont'd)

- First time applicants must consult with Digital Research Alliance of Canada technical staff for assessment. Send an e-mail to help@sharcnet.ca
- Info session dates:
  - Sept. 28 (English), Sept. 29 (French), 2024 info sessions on RAC applications.
  - Oct. 3 (English), Oct. 4 (French), 2023 info sessions on GPUs.

Supercomputing at Western, SHARCNET and beyond

- Supercomputing at Western, SHARCNET and beyond
- PIs applying for compute, storage and cloud resources
- What every graduate student should know
- Introduction to advanced research computing courses
- Q & A

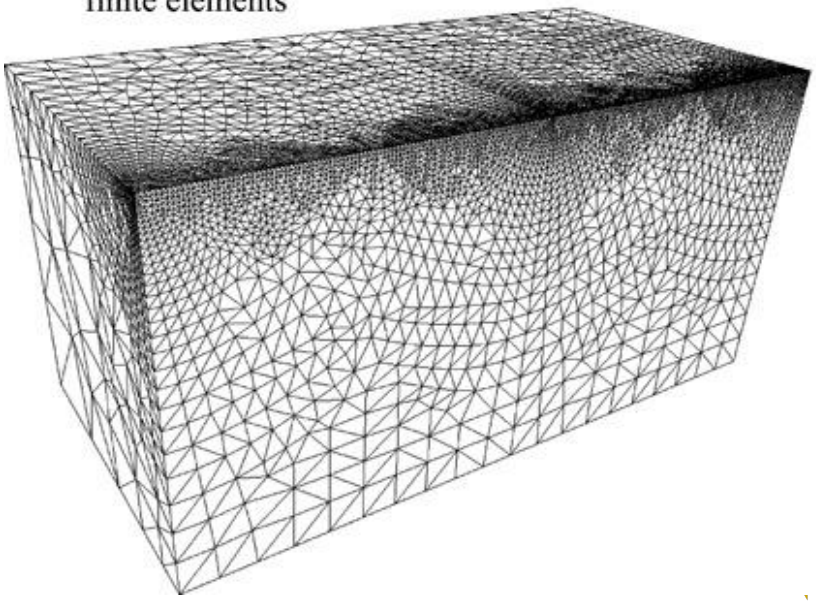

clouds
rain
ice
log10(snow water equivalent (mm))
-2
-1
0
1
2
16km
0km
Date: 20191031: 0.00

What every graduate student should know

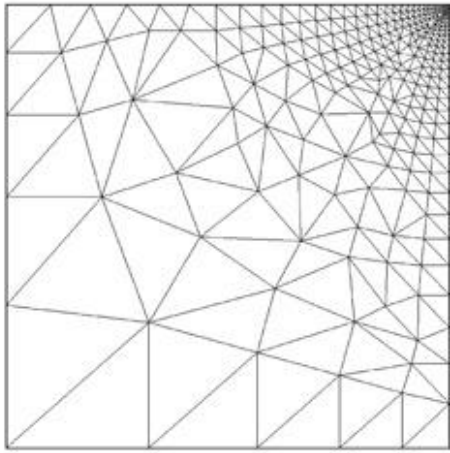
What people do on supercomputers

- Astrophysics simulations
- CFD, Environmental simulations, etc
- Coupled simulations
- Material science
- AI and machine/deep learning
- Economics, finance studies, etc.

3D mesh of linear tetrahedral finite elements



2D mesh of linear triangular finite elements



Western UNIVERSITY · CANADA

What every graduate student should know

A supercomputer is a lot of computers, not a super fast computer

- Thousands of computers, CPU cores, GPUs, and disks
- Requires concurrent processing to get work done faster

Research supercomputers run Linux and not Windows

- software has to support Linux to run on the supercomputer
- Linux is a publicly developed Operating System freely available

Background theory

- Computer architecture
- Algorithms and numerical methods

High performance programming is hard

- Efficient programs, algorithms, and libraries take decades – use them

Linux

Using Linux

- Moving around.
- Using shell, automating tasks.
- Remember, the core utilities of Linux are very fast.
- Connecting to other computers via Secure Shell (SSH).
- Access to file systems.
- Running programs.

Using Linux on Windows (if you are using Windows)

- Install Windows Subsystem for Linux (WSL).
- You are running a true Linux in Windows.
- You can practice, develop and run programs in it seriously.

Popular and lesser known items

Programming languages

- Lower level: C/C++, Fortran
- Higher level: Python, R
- New takes: Julia, Chapel

Libraries

- parallel programming: OpenMPI, OpenMP, OpenCL, CUDA/HIP
- numerics classic: blas, lapack, scalapack, FFTW
- numerics exascale/accelerators: magma, slate, heFFTe

Tools

- Editors and integrated development environments
- Compilers (gcc, Intel), interpreters (python, R), and just in time compilers
- Debuggers and profilers (gdb, DDT, MAP, etc.)

Popular and lesser known items

Data science/data wrangling:

- python: numpy, pandas, matplotlib, plotnine, sklearn , dask, rapids
- R: data frames, tidyverse (dplyr, simplr, ggplot, etc.)
- SQL

Machine learning and AI

- TensorFlow
- Keras and PyTorch

Visualization

- ParaView and VisIt

… and many more …

What every graduate student should know

See our training courses for a variety of topics that might interest you

https://training.sharcnet.ca/

Supercomputing at Western, SHARCNET and beyond

- What every graduate student should know
- Supercomputing at Western, SHARCNET and beyond
- PIs applying for compute, storage and cloud resources
- Introduction to advanced research computing courses
- Q & A

Training courses

SHARCNET offers a training course this fall and winter semester: ***Introduction to advanced research computing (Intro-ARC)***, including a series of modules

- Introduction to supercomputing
- Introduction to Shell
- Introduction to Jupyterlab
- Python for high performance computing
- Parallel programming with Fortran
- Parallel programming with C++
- Parallel programming with GPUs
- Programming distributed system with message passing interface (MPI)
- Introduction to scalable and accelerated data science
- Visualization of scientific data
- More… (https://training.sharcnet.ca/)

Training courses

How to sign up

- One needs to have a supervisor.
- The supervisor has an Alliance account.
- One needs to register for an Alliance account.
- Sign up for the course with your Alliance account.

Format and schedule

- Live classes weekly.
- Self-driven, grading offline courses with course materials, recorded videos, quizzes and assignments.
- Forums on different topics are open for attendees.

Training courses

How one will learn

- Participate in in-person live class via Zoom.
- Or study online materials, lecture recordings via self-paced learning.
- Complete quizzes and homework assignments.