



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

September 27, 2023 11am - Noon



*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

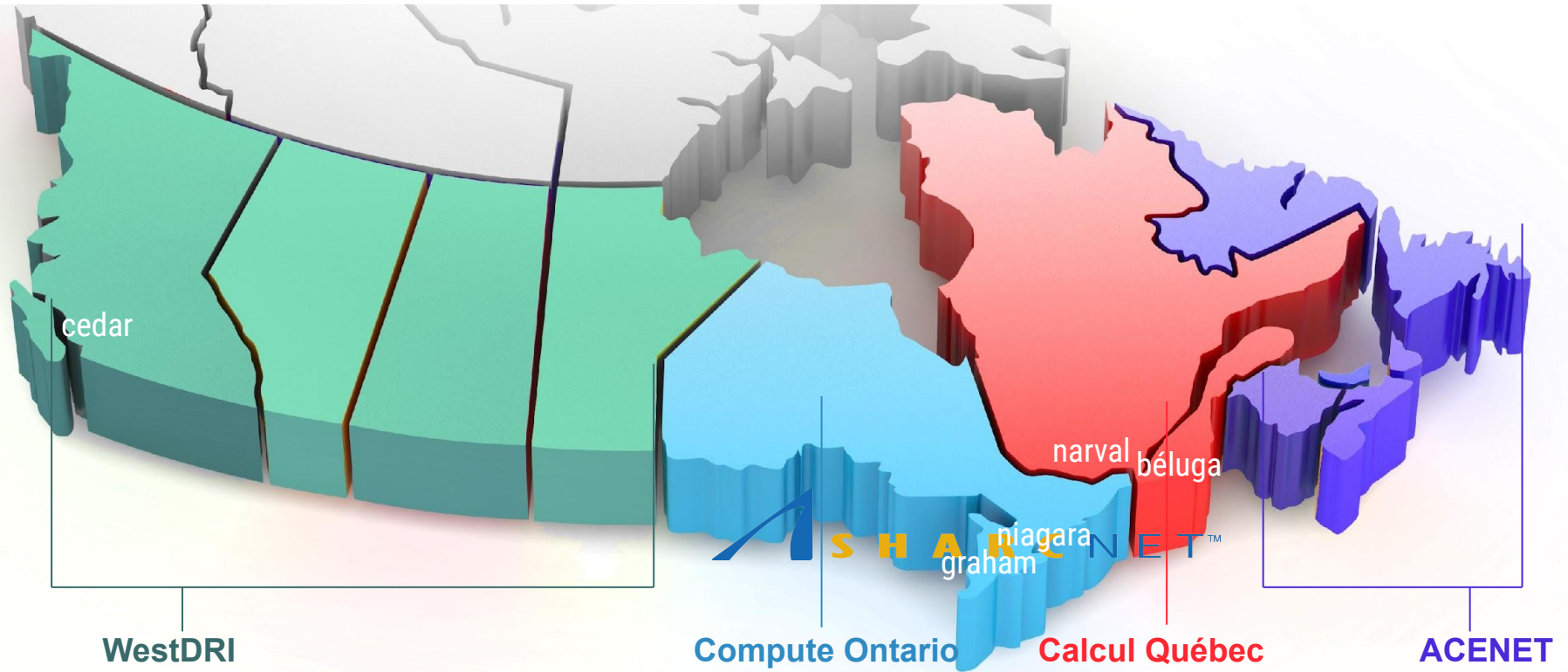
# Supercomputing at Western, SHARCNET and beyond

## Why use supercomputing resources

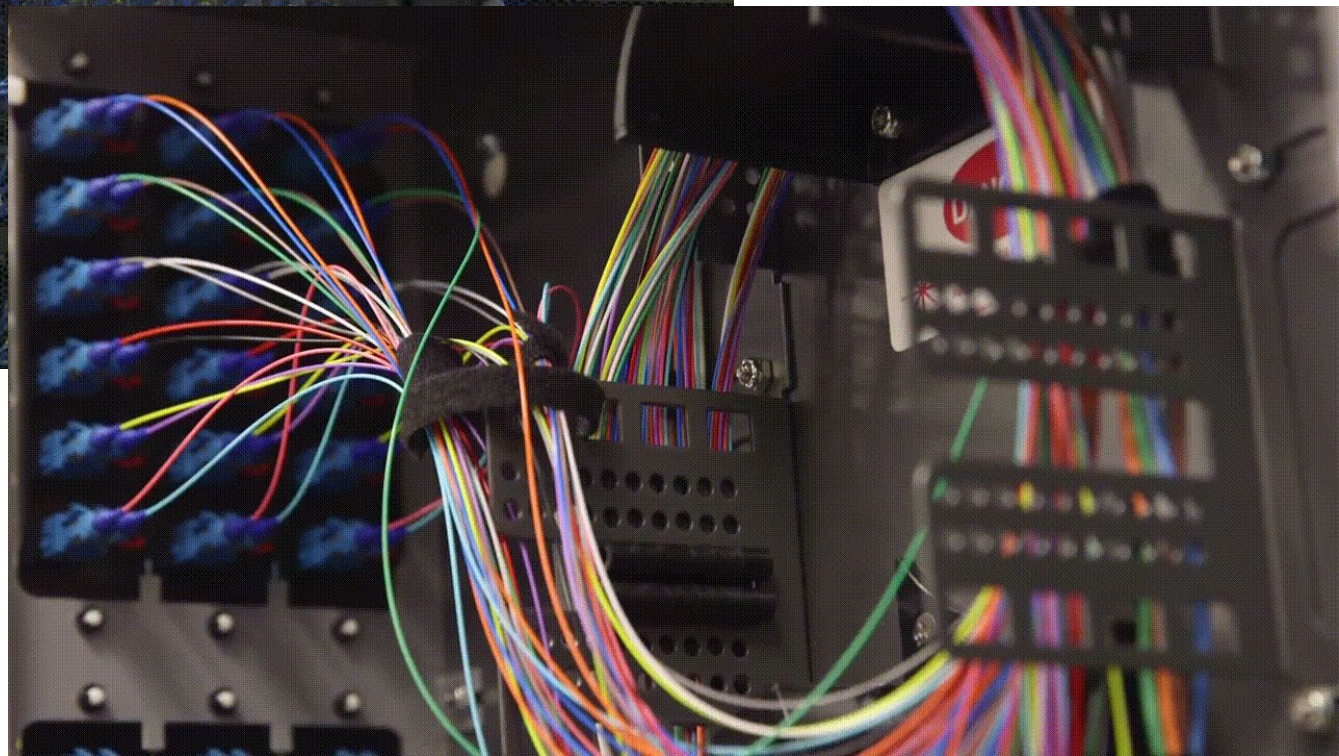
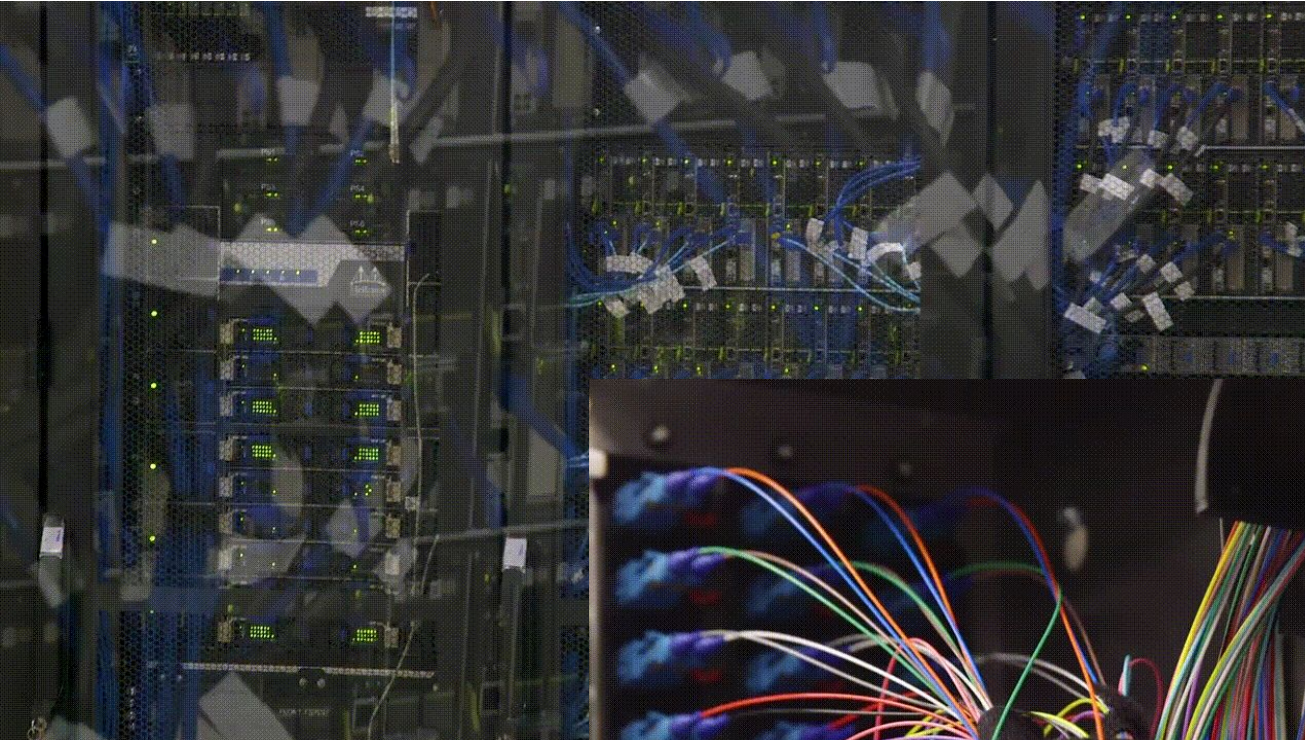
- You do not have many cores or much memory;
- You do not have the GPUs of the type needed;
- You need a huge amount of disk space, e.g. hundreds of TB;
- You need to run large scale of simulations that need hundreds of cores;
- You need to run large amount of simulations concurrently instead of one 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



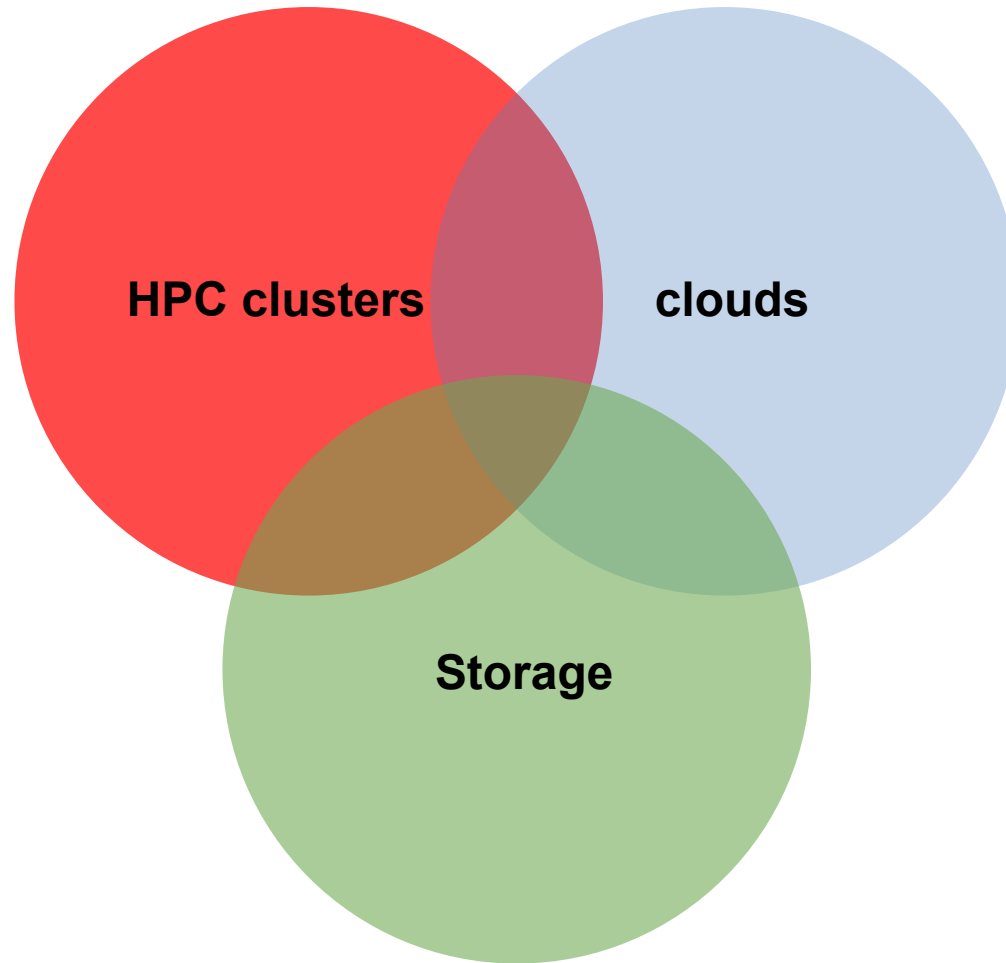
*Single account...*



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



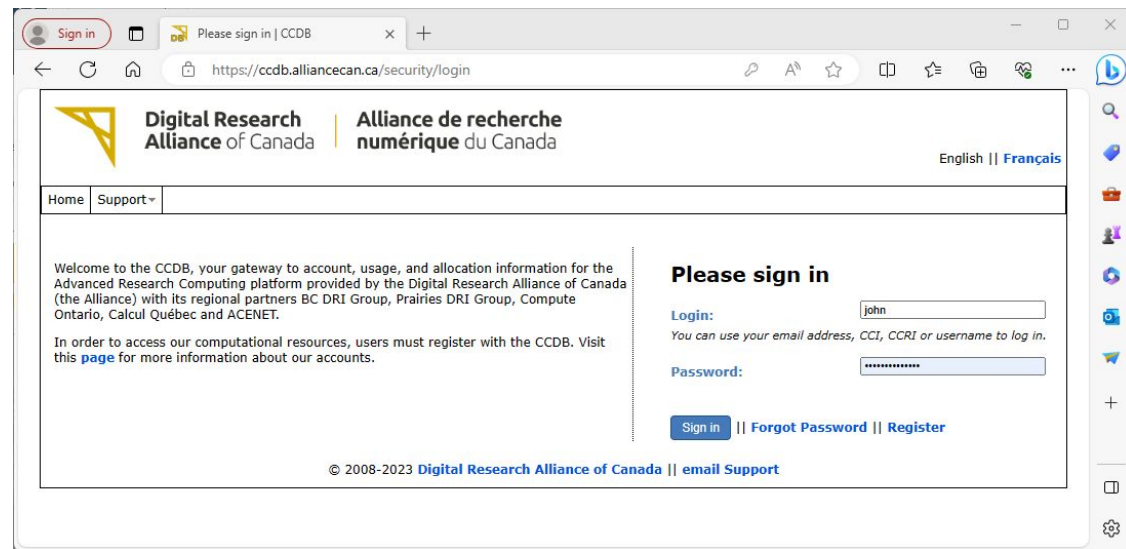
# Supercomputing at Western, SHARCNET and beyond



# 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](http://cedar.alliancecan.ca) (**94,528c**)
- [graham.alliancecan.ca](http://graham.alliancecan.ca) (**41,548c**)
- [niagara.alliancecan.ca](http://niagara.alliancecan.ca) (**80,640c**)
- [beluga.alliancecan.ca](http://beluga.alliancecan.ca) (**39,120c**)
- [narval.alliancecan.ca](http://narval.alliancecan.ca) (**80,912c**)

## Cloud services

- [arbutus.cloud.alliancecan.ca](http://arbutus.cloud.alliancecan.ca)
- [cedar.cloud.alliancecan.ca](http://cedar.cloud.alliancecan.ca)
- [graham.cloud.alliancecan.ca](http://graham.cloud.alliancecan.ca)
- [east.cloud.alliancecan.ca](http://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

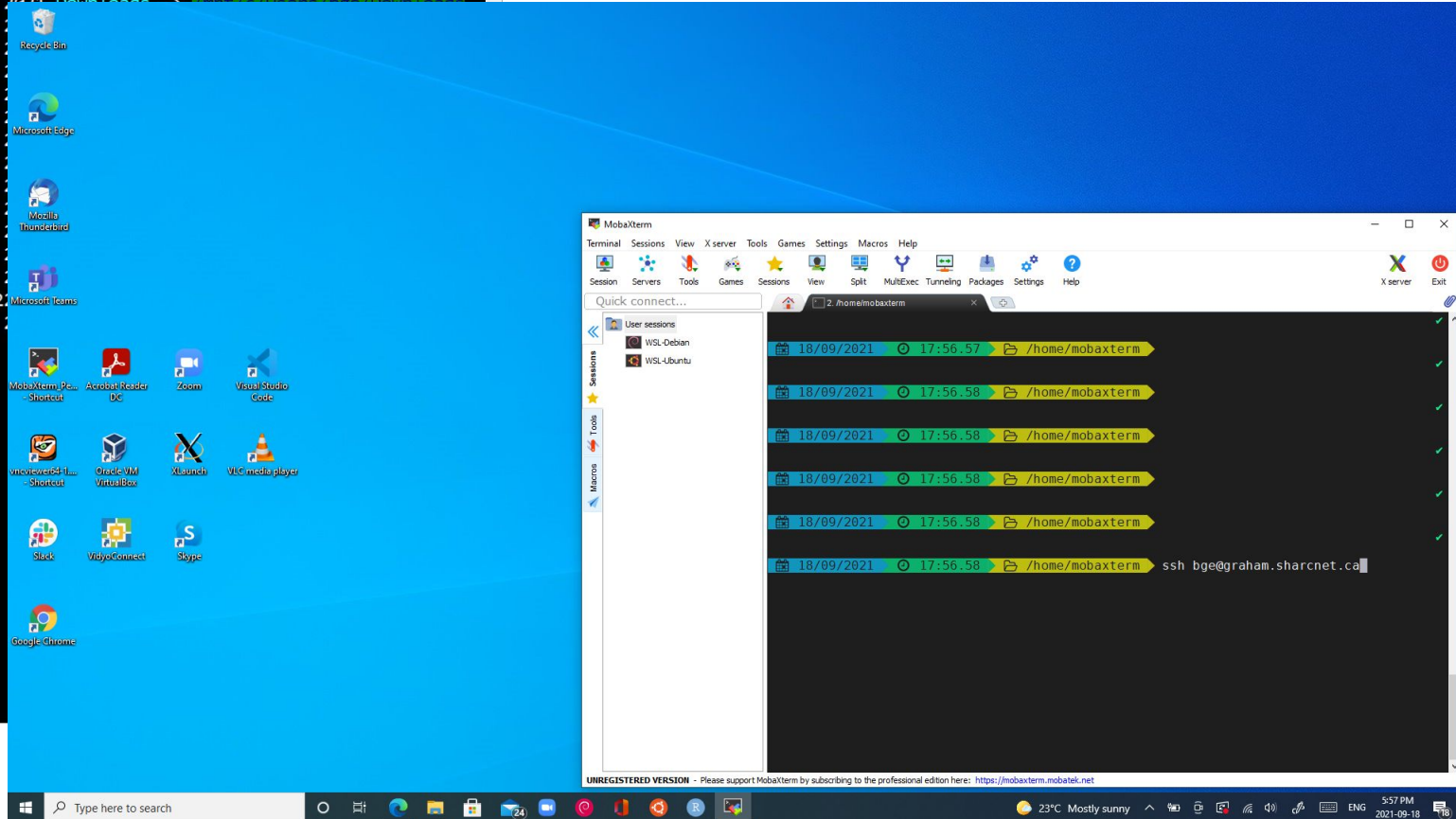
- OS: 64-bit Linux (CentOS)
- Languages supported: Python, R, C/C++, Fortran, Matlab, Java, Julia, CUDA, etc.
- Access to a variety of software packages
- Parallel development support:
  - **MPI**, Chapel: Distributed memory systems (cross compute nodes) and shared memory system (single node)
  - **OpenMP**, 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)

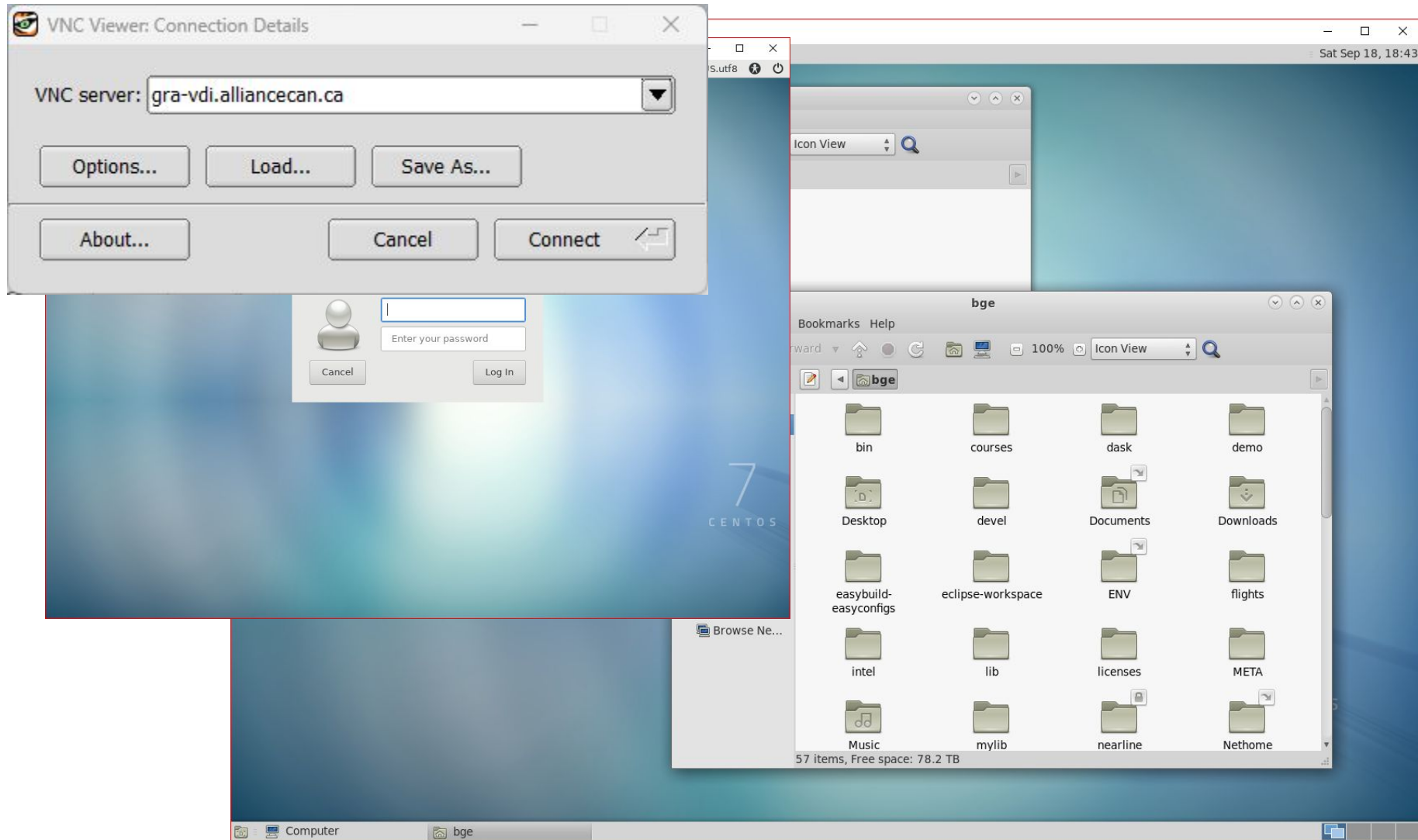
```
bge@crow:~$ ls
bin          Freezer      Music        Pictures     R            tmp
Documents   lib          Nethome     Private     Talks        Videos
Downloads   Library     output      Python      teaching

bge@crow:~$ ls -l
total 28
drwxr-xr-x  2 bge bge 4096 Jun 19 21:17 bin
lrwxrwxrwx  1 bge bge   26 Feb 18 2021 Documents -> /mnt/c/Users/bge/Documents
lrwxrwxrwx  1 bge bge   26 Feb 18 2021 Downloads -> /mnt/c/Users/bge/Downloads
lrwxrwxrwx  1 bge bge   16 May  8 2021 Music -> /mnt/c/Users/bge/Music
drwxr-xr-x  3 bge bge 4096 Jan 26 2021 Pictures
lrwxrwxrwx  1 bge bge   15 Mar  8 2021 Private
lrwxrwxrwx  1 bge bge   23 Oct 24 2021 R -> /mnt/c/Users/bge/R
lrwxrwxrwx  1 bge bge   12 Feb 28 2021 Talks
drwxr-xr-x  3 bge bge 4096 Oct  7 2021 Teaching
lrwxrwxrwx  1 bge bge   25 Feb 22 2021 tmp
lrwxrwxrwx  1 bge bge   16 Feb 10 2021 Videos
drwxr-xr-x  6 bge bge 4096 Jun  7 2021 .
drwxr-xr-x  3 bge bge 4096 Feb 28 2021 ..
lrwxrwxrwx  1 bge bge   18 Jan 15 2021 .ssh
drwxr-xr-x  3 bge bge 4096 Mar 19 2021 .vscode
drwxr-xr-x  4 bge bge 4096 May 31 2021 .wslconfig
lrwxrwxrwx  1 bge bge   23 Feb 22 2021 .wslconfig
bge@crow:~$
```

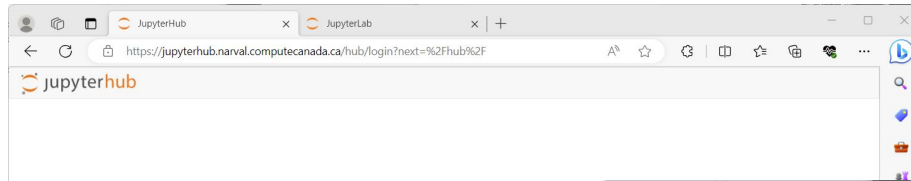


via Linux terminal

# Connecting to GUI desktop



# Connecting to clusters via jupyterhub



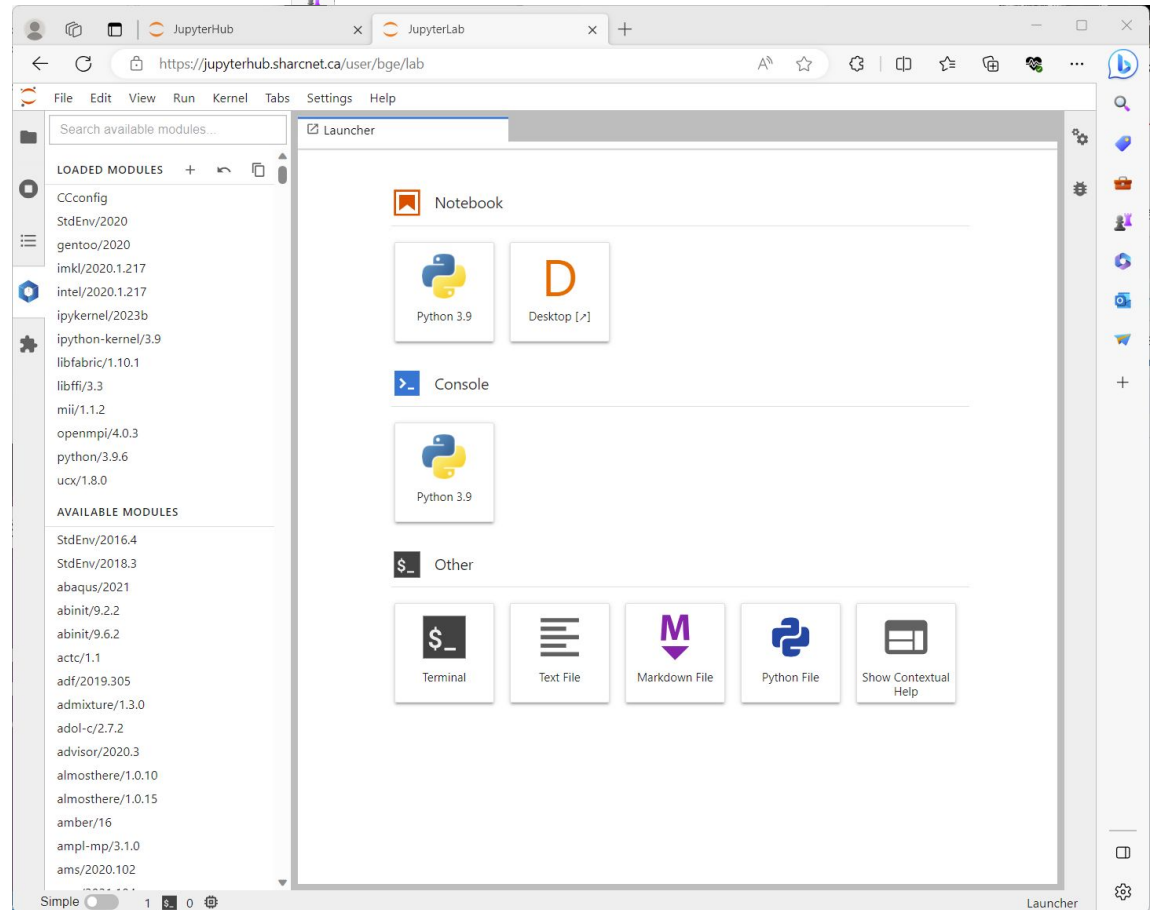
Access to the cluster via jupyter notebook, terminal and desktop in a browser.

Sign in

Username:

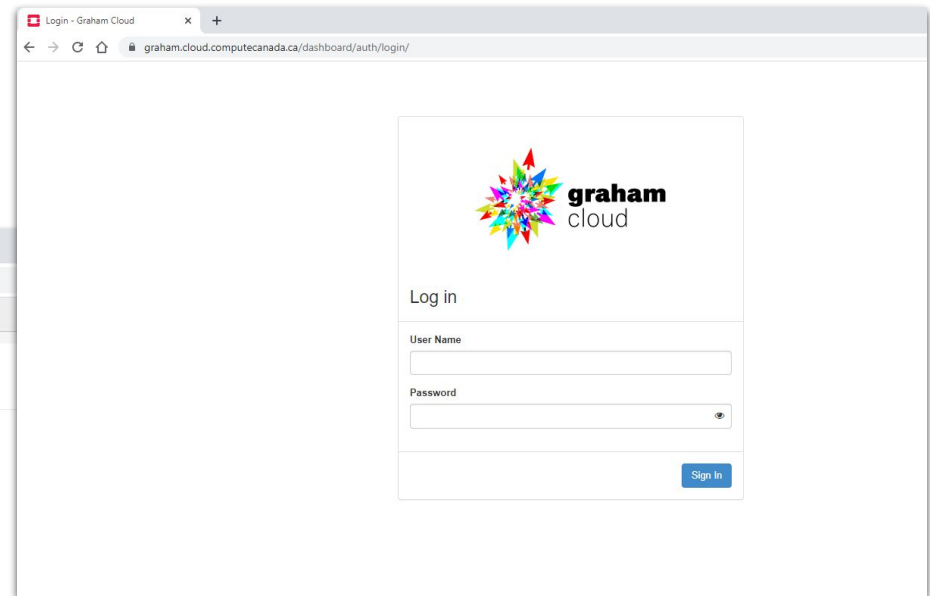
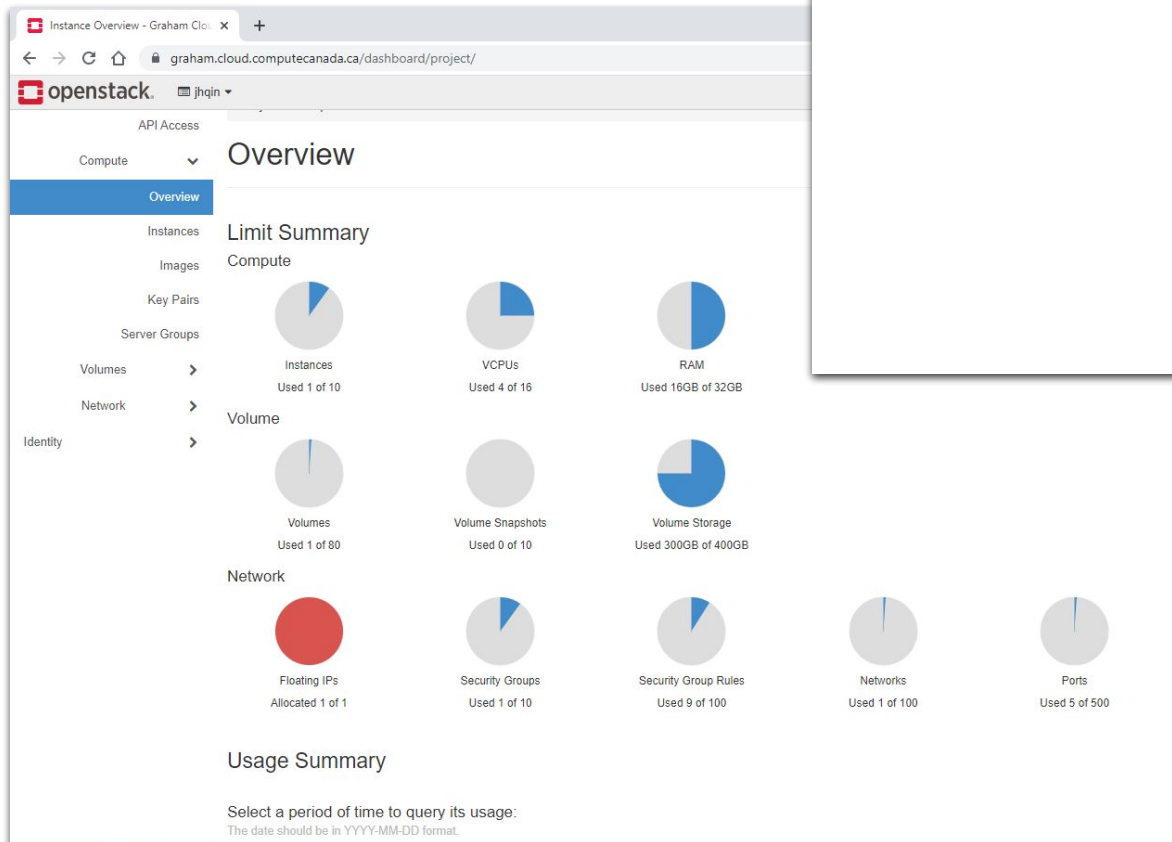
Password:

Sign in



# Connecting to cloud

- [arbutus.cloud.alliancecan.ca](https://arbutus.cloud.alliancecan.ca)
- [cedar.cloud.alliancecan.ca](https://cedar.cloud.alliancecan.ca)
- [graham.cloud.alliancecan.ca](https://graham.cloud.alliancecan.ca)
- [east.cloud.alliancecan.ca](https://east.cloud.alliancecan.ca)



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

# Accessing files on remote systems

The screenshot illustrates a Windows desktop environment with several applications open. The primary focus is on accessing a remote file system. A File Explorer window titled "bge@graham.sharcnet.ca (\\sshfs) (U:)" is open, showing a list of folders and files. The folders include docs, lecture, UWO, 3D Objects, Desktop, Documents, Downloads, Music, Pictures, Videos, Local Disk (C:), SDXC (D:), bge (\\192.168.1...), public (\\192.168...), and bge@graham.sh... The files listed include folders like octave, off, opt, People, perf5, Pictures, project, projects, Public, public, Python, R, rrms, scratch, src, ss2021, ss2020, Templates, and tests, along with their respective dates and sizes.

In the background, another File Explorer window shows the "This PC" view, highlighting network locations. One of the network locations is "bge@graham.sharcnet.ca (\\sshfs) (U:)", which is selected. The desktop also shows various icons for applications like Recycle Bin, Microsoft Edge, Mozilla Thunderbird, Microsoft Teams, MobaXterm, and Google Chrome.

A terminal window in the foreground shows the output of a file listing command. The output is as follows:

```
gr-a-login3 ~]$ ls -l
-rwxr-xr-x 1 bge bge 16 Jul 14 2017 project -> /project/6000114
-rwxr-xr-x 2 bge bge 4 Jul 17 2017 projects
-rwxr-xr-x 4 bge bge 6 Jun 1 17:44 public -> Public
-r-xr-xr-x 4 bge bge 4 Aug 18 14:57 Public
-r-xr-xr-x 4 bge bge 5 Jun 9 09:57 Python
-rwxr-xr-x 3 bge bge 3 Dec 19 2019 R
-rwxr-xr-x 1 bge bge 393 May 25 11:20 README.txt
-rwxr-xr-x 1 bge bge 4599 Feb 26 2018 RpLots.pdf
-rwxr-xr-x 9 bge bge 12 Feb 26 2020 rrms
-rwxr-xr-x 1 bge bge 12 Jul 14 2017 scratch -> /scratch/bge
-rwxr-xr-x 1 bge bge 529 Jan 9 2020 sn_gpu.R
-rwxr-xr-x 4 bge bge 4 Oct 11 2019 src
-r-xr-xr-x 5 bge bge 5 Jun 10 14:08 ss2021
-r-xr-xr-x 3 bge bge 4 Jun 9 2020 sss2020
-rwxr-xr-x 2 bge bge 2 Aug 8 2018 Templates
-rwxr-xr-x 1 bge bge 1435 May 27 15:57 test.f90
-rwxr-xr-x 82 bge bge 121 Oct 8 2018 tests
-rwxr-xr-x 2 bge bge 2 Aug 8 2018 Videos
```

# Transferring large amounts of files using Globus

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

The screenshot displays the Globus File Manager interface. The top navigation bar shows the current collection as 'computecanada#graham-globus' and the destination as 'bge-crow'. The path on the left is '/home/bge/' and on the right is '/C:/Users/bge/Documents/teaching/'. A 'Start' button is visible in the center. Below the navigation bar, there are two panels showing file lists. The left panel shows a list of folders and files, with 'devel' selected. The right panel shows a list of files and folders, including 'intro\_hpc\_2022fall', 'intro\_hpc\_2022fall.odp', 'julia\_2020fall', 'julia\_ccf\_2022-02', 'mpfun', 'mpi\_2021-2022', 'oneapi', 'output', 'pack2.f90', 'pack2a.f90', 'perf\_coss2022', 'R', and 'README.md'. A central menu is open, showing options like 'Share', 'Transfer or Sync to...', 'New Folder', 'Rename', 'Delete Selected', 'Download', 'Open', 'Upload', 'Get Link', 'Show Hidden Items', and 'Manage Activation'.

NAME	LAST MODIFIED	SIZE
a.out	6/1/2021, 01:42 PM	6
a2-1.f90	6/1/2021, 01:46 PM	2
addwa_gpu	1/24/2022, 12:35 PM	4
array_asgmt2d.f90	5/26/2021, 03:27 PM	6
bin	6/1/2021, 06:07 PM	
coss2022	7/13/2022, 01:40 PM	
courseid_114_participants.csv	1/24/2022, 11:29 AM	2
courses	6/5/2019, 12:52 PM	
dask	4/1/2021, 02:24 PM	
demo	7/11/2023, 02:55 PM	
Desktop	8/8/2018, 02:13 PM	
devel	12/4/2019, 09:52 AM	
Documents	7/20/2020, 01:49 PM	

NAME	LAST MODIFIED	SIZE
intro_hpc_2022fall	9/7/2022, 02:47 PM	-
intro_hpc_2022fall.odp	9/7/2022, 02:19 PM	6.70 MB
julia_2020fall	1/22/2021, 09:04 PM	-
julia_ccf_2022-02	2/16/2022, 11:32 AM	-
mpfun	2/17/2021, 04:39 PM	-
mpi_2021-2022	3/27/2022, 11:52 PM	-
oneapi	2/7/2022, 11:17 AM	-
output	10/7/2021, 12:52 AM	-
pack2.f90	5/27/2021, 05:16 PM	632 B
pack2a.f90	5/27/2021, 05:18 PM	632 B
perf_coss2022	7/4/2022, 01:02 AM	-
R	2/27/2021, 02:43 PM	-
README.md	1/22/2021, 09:04 PM	40 B

# Where to look for information and get help

The image displays two browser windows side-by-side. The left window shows the 'Technical documentation' page on the Digital Research Alliance of Canada website. It features a search bar, a language selector (English), and a navigation menu on the left. The main content area includes a welcome message and two columns of links: 'Systems and services' and 'How-to guides'. The right window shows the SHARCNET website homepage, which has a yellow header with navigation links (FACILITIES, SUPPORT, ABOUT US) and a large image of a particle detector. Below the image is a section titled 'Neutrinos by the Numbers: Sudbury's SNOLAB' and a 'GitLab Instance at SHARCNET' announcement. A sidebar on the right lists 'EVENTS' and 'NEWS'.



<https://docs.alliancecan.ca/>

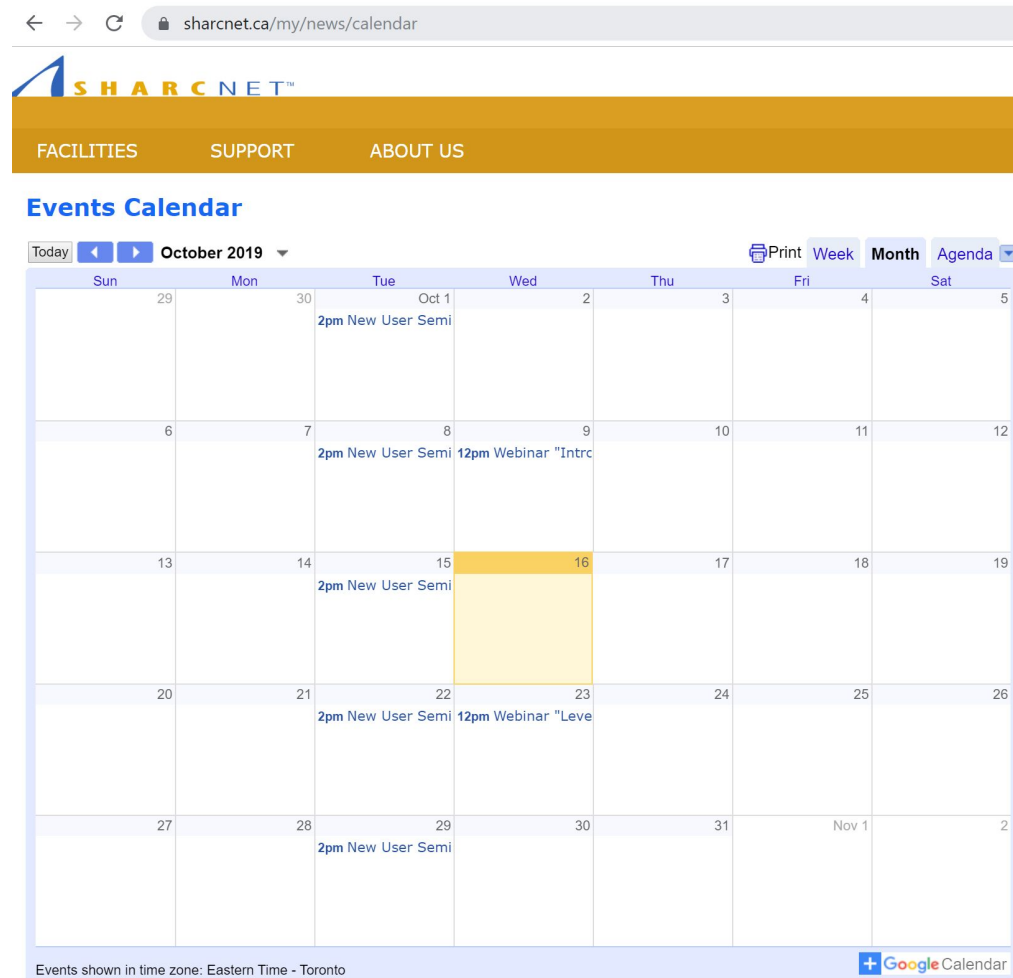
<https://www.sharcnet.ca/>



# 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 <sup>CA</sup> [youtube.sharcnet.ca](https://www.youtube.com/sharcnet.ca)



sharcnet.ca/my/news/calendar


SHARCNET™

FACILITIES SUPPORT ABOUT US

### Events Calendar

Today ← → October 2019 Print Week Month Agenda

Sun	Mon	Tue	Wed	Thu	Fri	Sat
29	30	Oct 1 2pm New User Semi	2	3	4	5
6	7	8 2pm New User Semi	9 12pm Webinar "Intrc	10	11	12
13	14	15 2pm New User Semi	16	17	18	19
20	21	22 2pm New User Semi	23 12pm Webinar "Leve	24	25	26
27	28	29 2pm New User Semi	30	31	Nov 1	2

Events shown in time zone: Eastern Time - Toronto 

# Where to look for information and get help

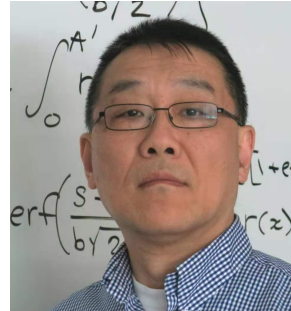
## Local staff



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](mailto: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  $\geq$  200 core years (Cy) per cluster
  - GPU  $\geq$  25 Reference GPU units (RGUs) per cluster
  - Project storage  $\geq$  41 TB, or,
  - Nearline storage  $\geq$  101 TB, or,
  - Persistent cloud storage  $\geq$  1q TB, or,
  - Compute cloud  $\geq$  81 vCPUs, or,
  - Persistent cloud  $\geq$  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](mailto: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

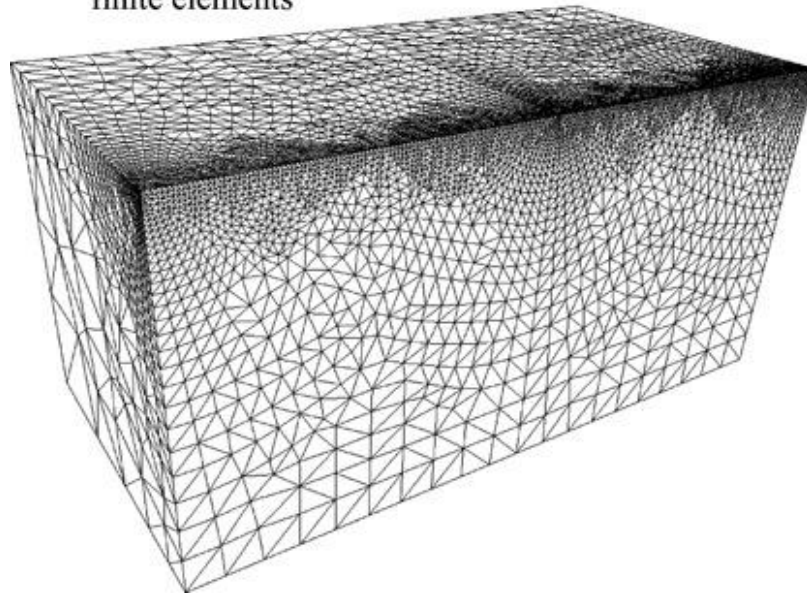


# What every graduate student should know

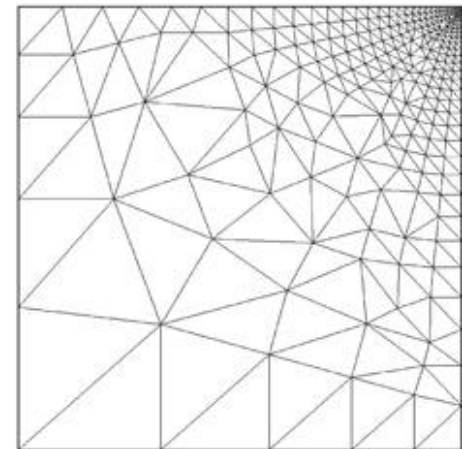
What people do on supercomputers

- 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



# 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.