

ARC/HPC @ OntarioTech Fall Update

New Resources, Support and
Training Courses

October 10, 2024, 12:00-1:00 pm

Armin Sobhani

SHARCNET | The Alliance

HPC Technical Consultant






✉ asobhani@sharcnet.ca

✉ armin.sobhani@ontariotechu.ca

🏠 <https://staff.sharcnet.ca/asobhani>



In Today's Update...

-  Supercomputing @ **OntarioTech**
-  What PIs Should Know
-  What Graduate Students Should Know
-  SHARCNET's Online Training
-  Q&A

Supercomputing @ OntarioTech



Academic / Research Computing @ OntarioTech

Campus IT Groups

Academic Computing

<https://itsc.ontariotechu.ca/>

OntarioTechU.Net

SHARCNET – HPC

Advanced Research Computing
(ARC)

Research data management
(RDM, e.g. for librarians)

Training on ARC, data science,
machine learning

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

Academic / Research Computing @ OntarioTech

Campus IT Groups

Academic Computing

<https://itsc.ontariotechu.ca/>

OntarioTechU.Net

SHARCNET – HPC

Advanced Research Computing
(ARC)

Research data management
(RDM, e.g. for librarians)

Training on ARC, data science,
machine learning

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

Academic / Research Computing @ OntarioTech

Campus IT Groups

Academic Computing

<https://itsc.ontariotechu.ca/>

OntarioTechU.Net

SHARCNET – HPC

Advanced Research Computing
(ARC)

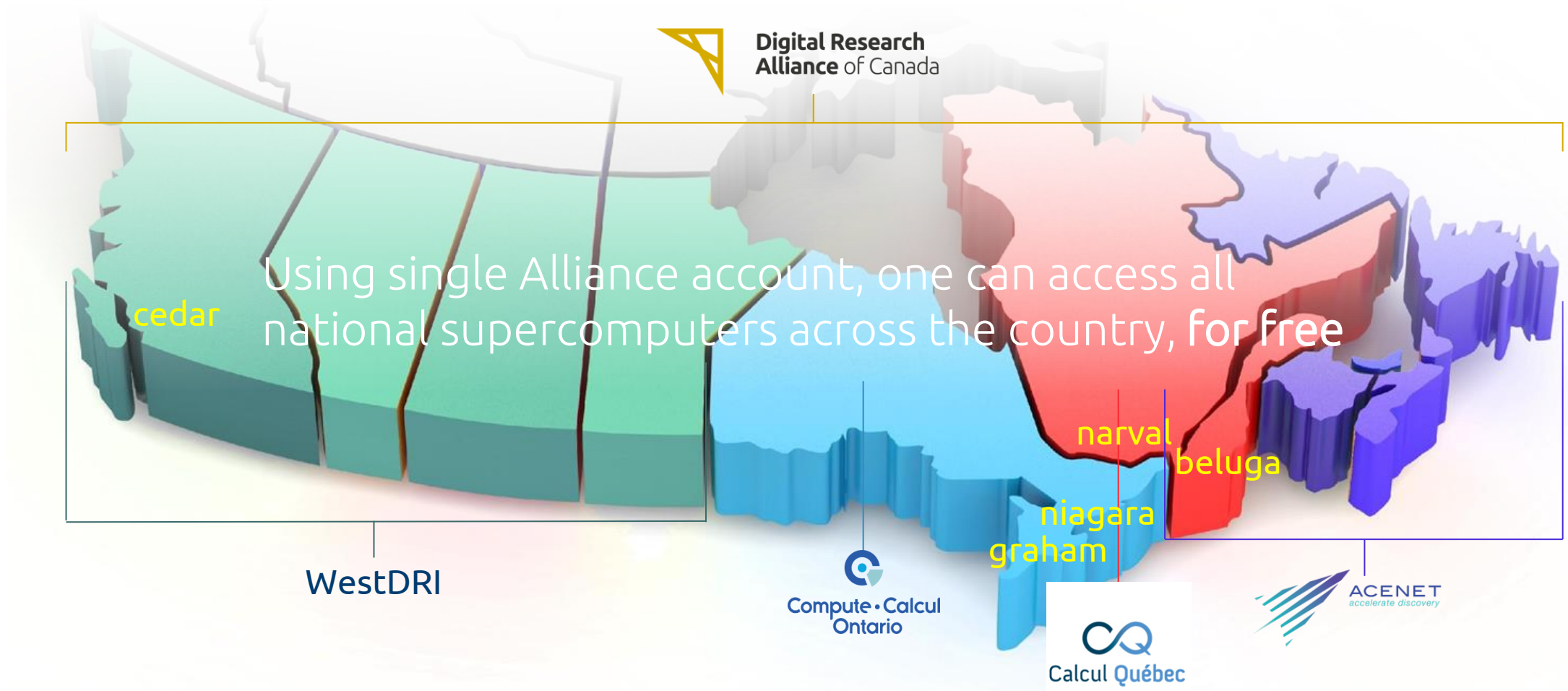
Research data management
(RDM, e.g. for librarians)

Training on ARC, data science,
machine learning

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

Supercomputers – What They Are...

Supercomputers – Where They Are...



Supercomputers – Why You Need Them...

You do not have enough CPU 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 simulations that need many CPU cores

You need to run large number of simulations concurrently instead of one after another

You need to run web services

You need to run SQL databases

You need to run programs on a cloud

Supercomputing – Getting an Account

Sign up for an Alliance account for FREE at:

<http://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

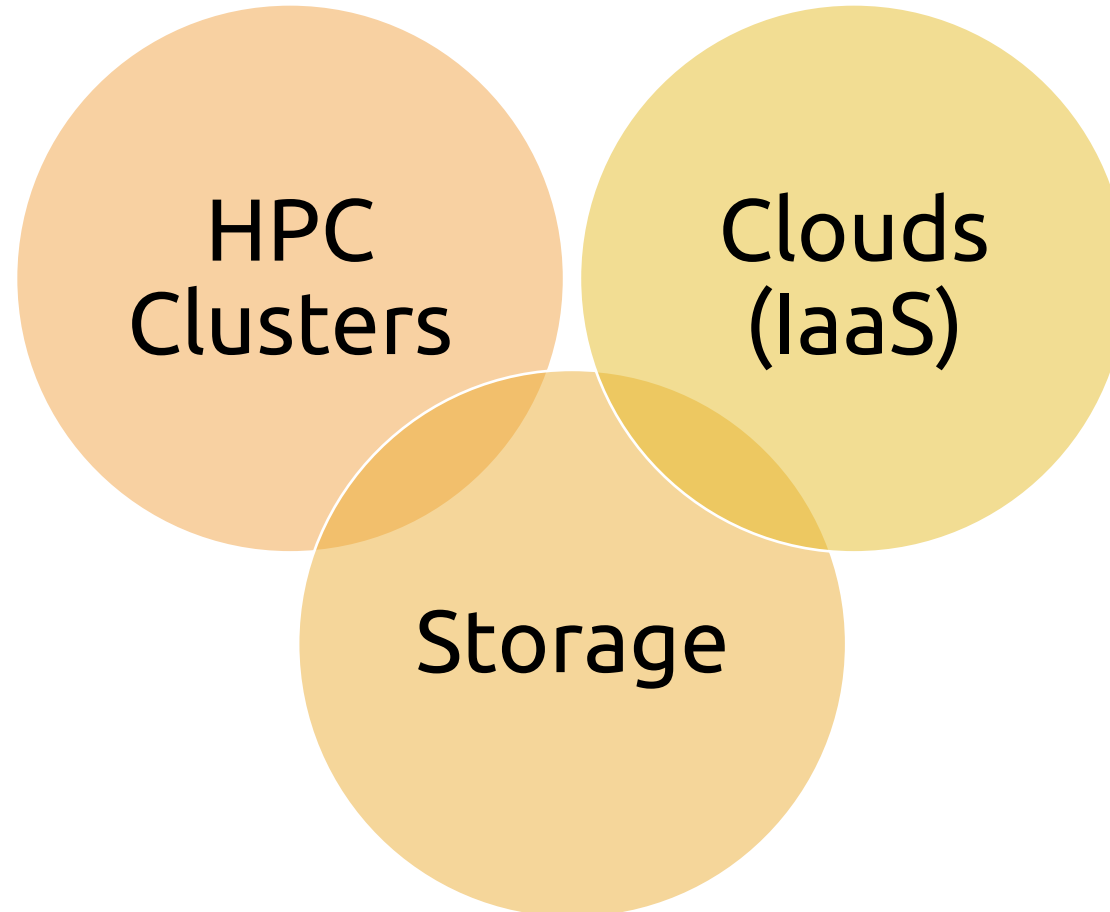


Multi-Factor Authentication (MFA) is now required



Supercomputing – Resources

To run large and exascale simulations that need dedicated access to hundreds of thousands CPU and GPU cores and low latency, fast interconnect fabric; or to run hundreds of thousands of simulations simultaneously



To run web services, databases or use virtual machines (VMs) with full control. The CPU and GPU resources are typically over committed, that is, a CPU core or GPU device might be shared by the running VMs

To store hundreds of thousands TB of research data

Supercomputing – Resources

Clusters across the Country

cedar.alliancecan.ca

graham.alliancecan.ca

niagara.alliancecan.ca

beluga.alliancecan.ca

narval.alliancecan.ca

Cloud Services (IaaS)

arbutus.cloud.alliancecan.ca

cedar.cloud.alliancecan.ca

graham.cloud.alliancecan.ca

beluga.cloud.alliancecan.ca

Storage Space

home
50G, backed up

project
1T per group, backed up

scratch
20T per user, purged 60 days

nearline
(tapes) - archive

Supercomputing – Resources

Clusters across the Country

cedar.alliancecan.ca

graham.alliancecan.ca

niagara.alliancecan.ca

beluga.alliancecan.ca

narval.alliancecan.ca

Cloud Services (IaaS)

arbutus.cloud.alliancecan.ca

cedar.cloud.alliancecan.ca

graham.cloud.alliancecan.ca

beluga.cloud.alliancecan.ca

Storage Space

home
50G, backed up

project
1T per group, backed up

scratch
20T per user, purged 60 days

nearline
(tapes) - archive

Supercomputing – Resources

Clusters across the Country

cedar.alliancecan.ca

graham.alliancecan.ca

niagara.alliancecan.ca

beluga.alliancecan.ca

narval.alliancecan.ca

Cloud Services (IaaS)

arbutus.cloud.alliancecan.ca

cedar.cloud.alliancecan.ca

graham.cloud.alliancecan.ca

beluga.cloud.alliancecan.ca

Storage Space

home
50G, backed up

project
1T per group, backed up

scratch
20T per user, purged 60 days

nearline
(tapes) - archive

Supercomputing – Resources

Clusters across the Country

cedar.alliancecan.ca

graham.alliancecan.ca

niagara.alliancecan.ca

beluga.alliancecan.ca

narval.alliancecan.ca

Cloud Services (IaaS)

arbutus.cloud.alliancecan.ca

cedar.cloud.alliancecan.ca

graham.cloud.alliancecan.ca

beluga.cloud.alliancecan.ca

Storage Space

home
50G, backed up

project
1T per group, backed up

scratch
20T per user, purged 60 days

nearline
(tapes) - archive

Supercomputing – Resources

Current Resources across the Country ¹

Resource	Cedar	Graham	Niagara	Béluga	Narval	Total
CPU cores	100,400	41,548	80,640	39,120	83,216	344,924
GPUs	1,352	520	N/A	688	632	3,192
Storage ²	23PB	16PB	3.5PB	25PB	19PB	-

Source: https://docs.alliancecan.ca/wiki/National_systems

1. The counts may not reflect the actual numbers in service.
2. Listed here are project spaces only. Each cluster has home and scratch, which are relatively smaller. Some systems also have nearline, a tape-based storage system for archive.

Supercomputing – Resources

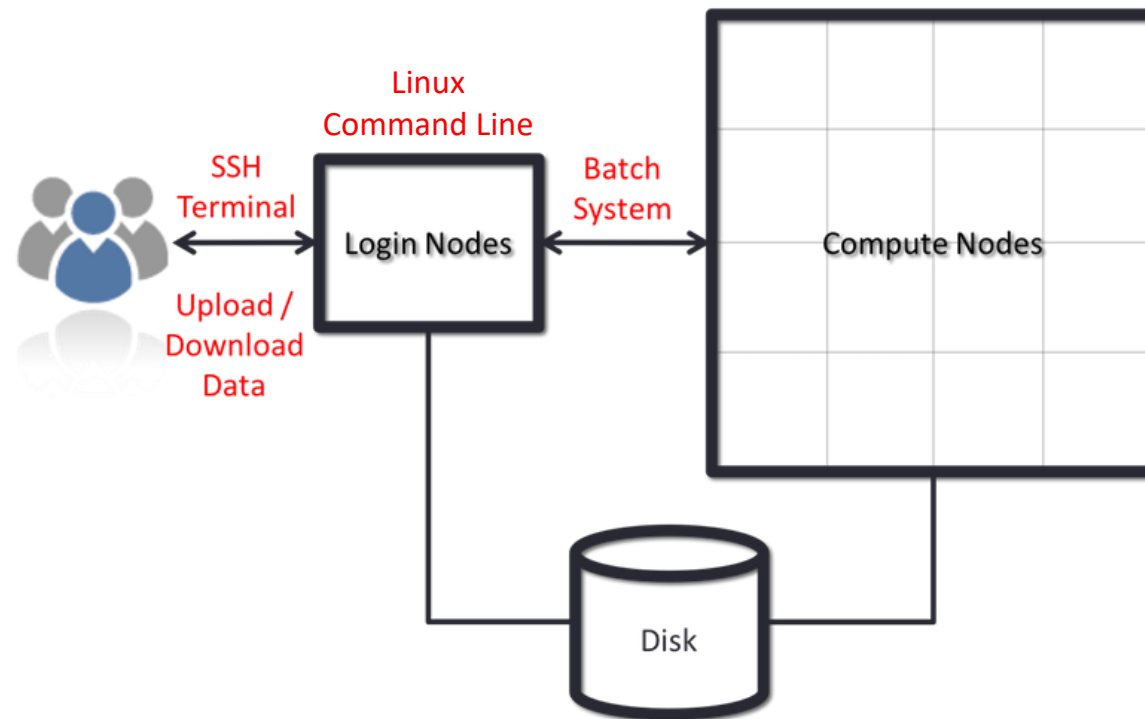
System Refresh by 2025 (Tentative) ¹

Resource	Fir	Graham2	Niagara2	Rorqual	Narval	Total
CPU cores	165,120	134,400	235,008	131,712	83,216	749,456
GPUs (H100)	640	280	240	324	632	2,116
Storage ²	49PB	25PB	29PB	TBD	19PB	-

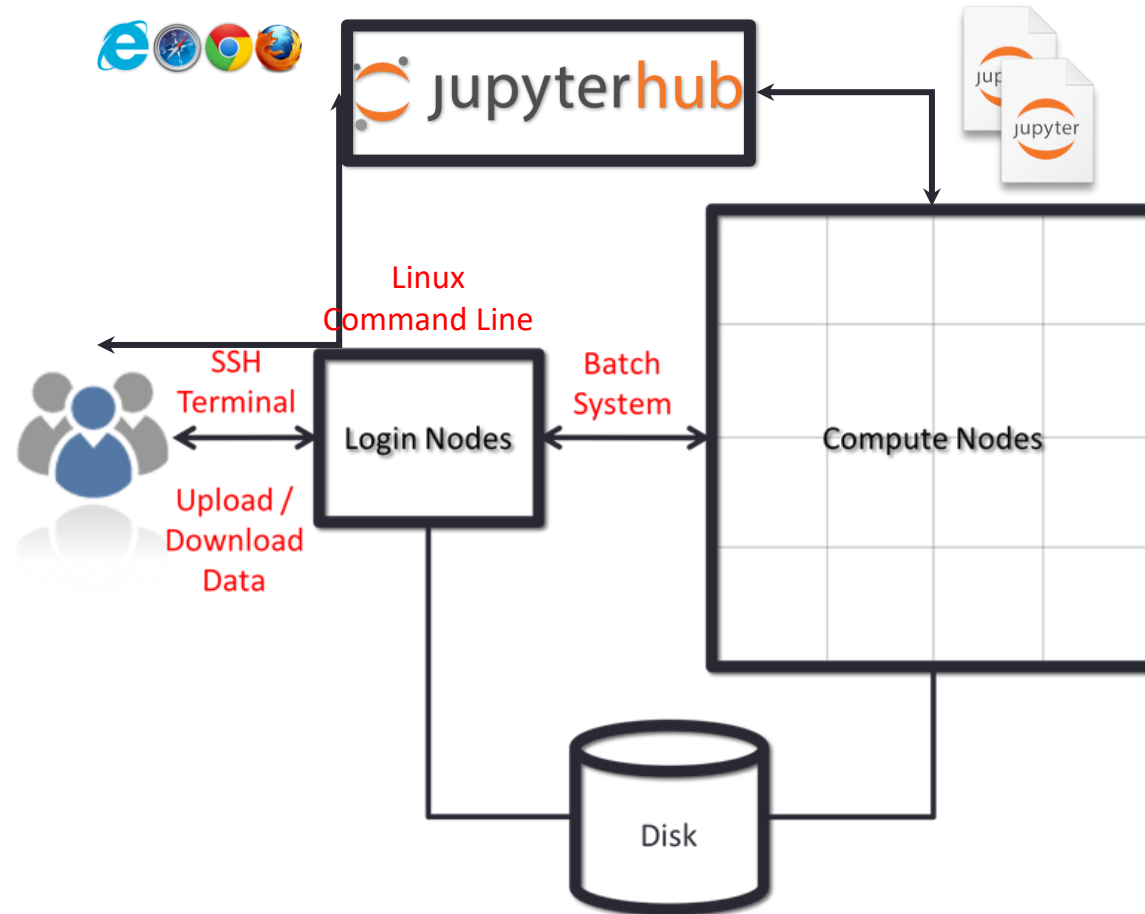
Source: https://docs.alliancecan.ca/wiki/National_systems

1. By July 2025, these new systems will be in operation.
2. These numbers represent the major storage.

Supercomputing – Connecting to Clusters



Supercomputing – Connecting to Clusters



Supercomputing – Connecting via SSH

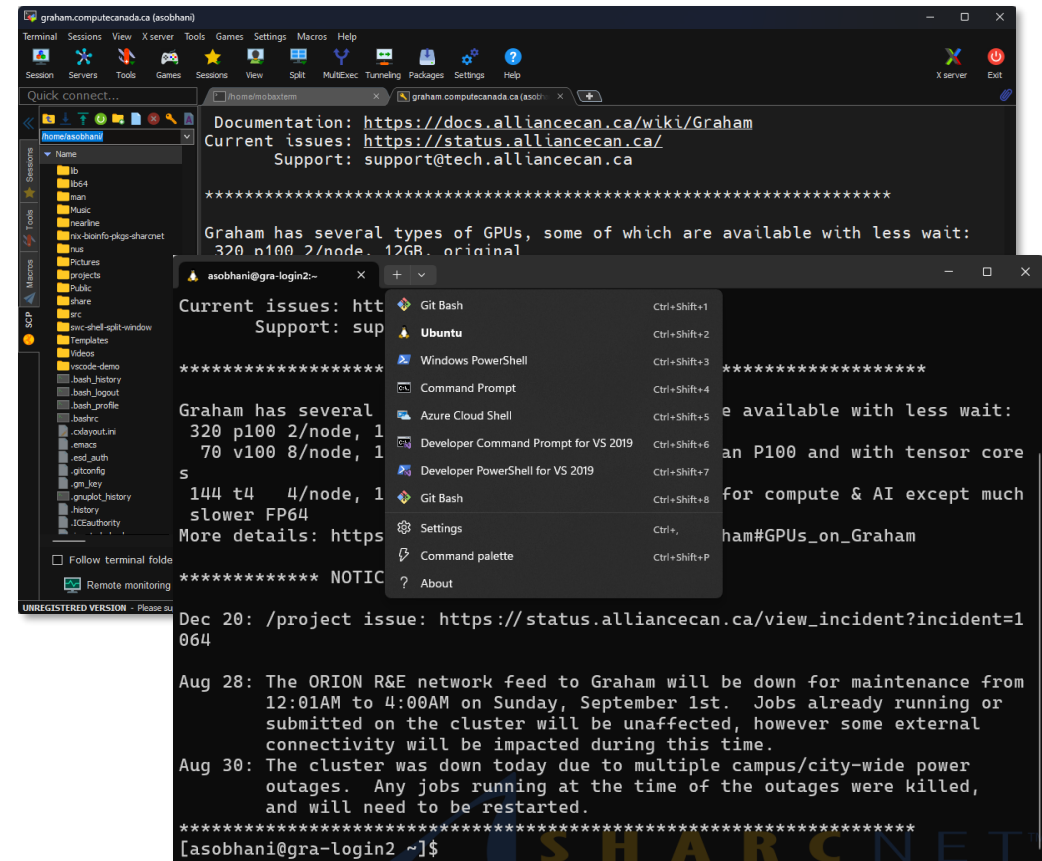
SSH Client

Windows 10 and 11

- Windows Terminal
 - OpenSSH in PowerShell or CMD
 - WSL
 - Git Bash
- MobaXterm

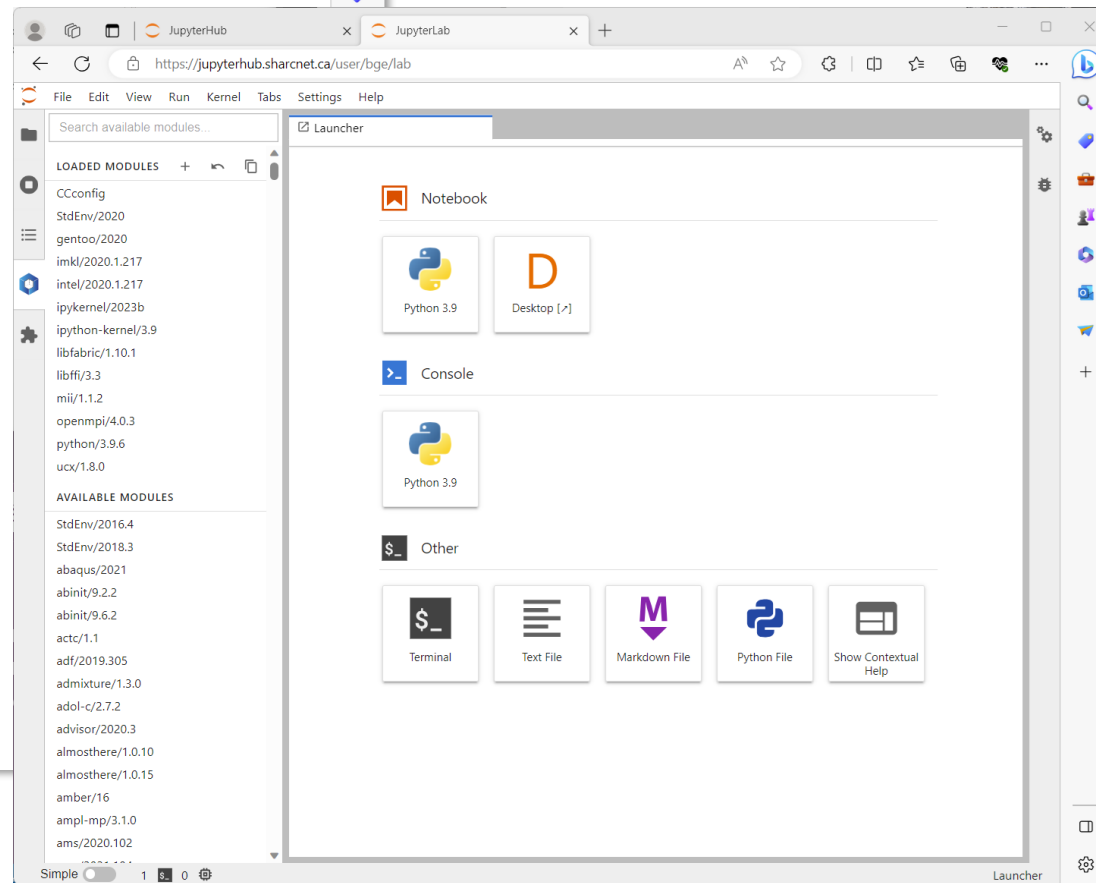
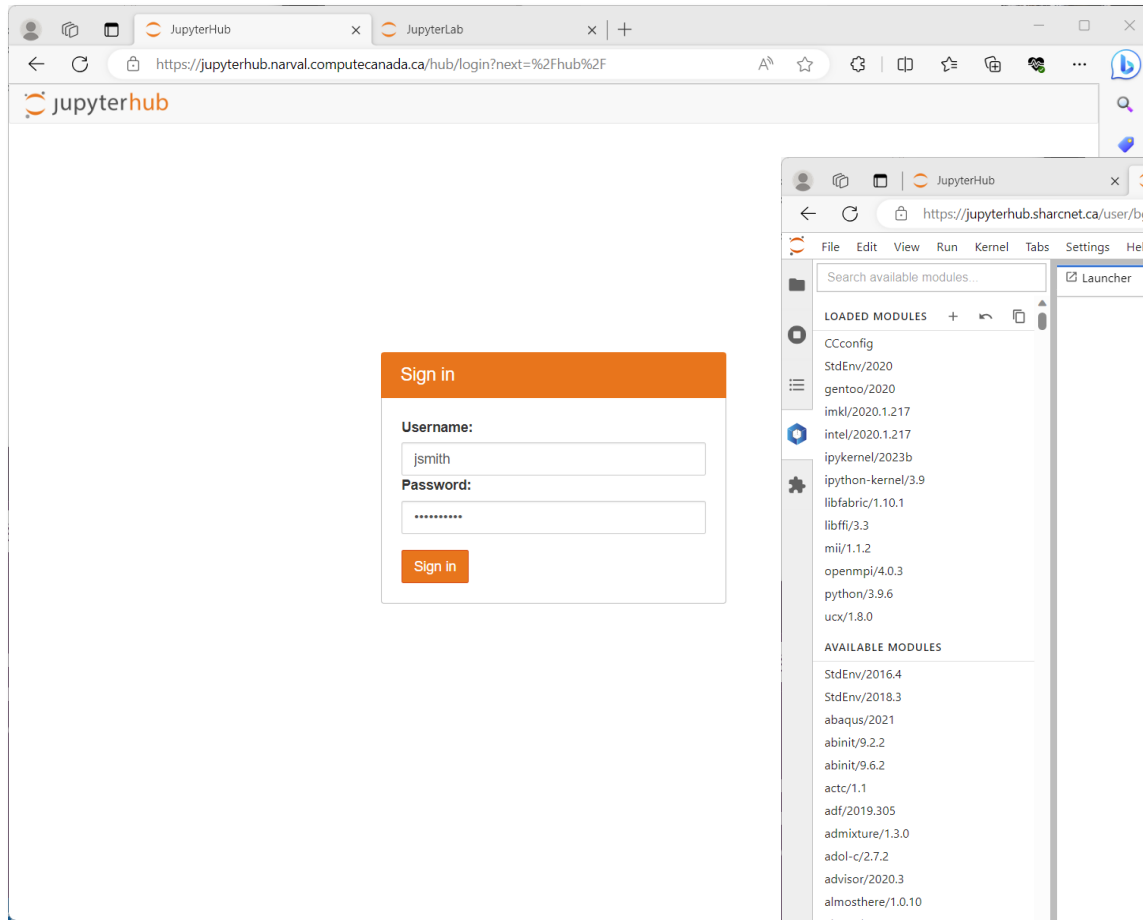
Linux and Mac

- OpenSSH
- ssh username@graham.alliancecan.ca

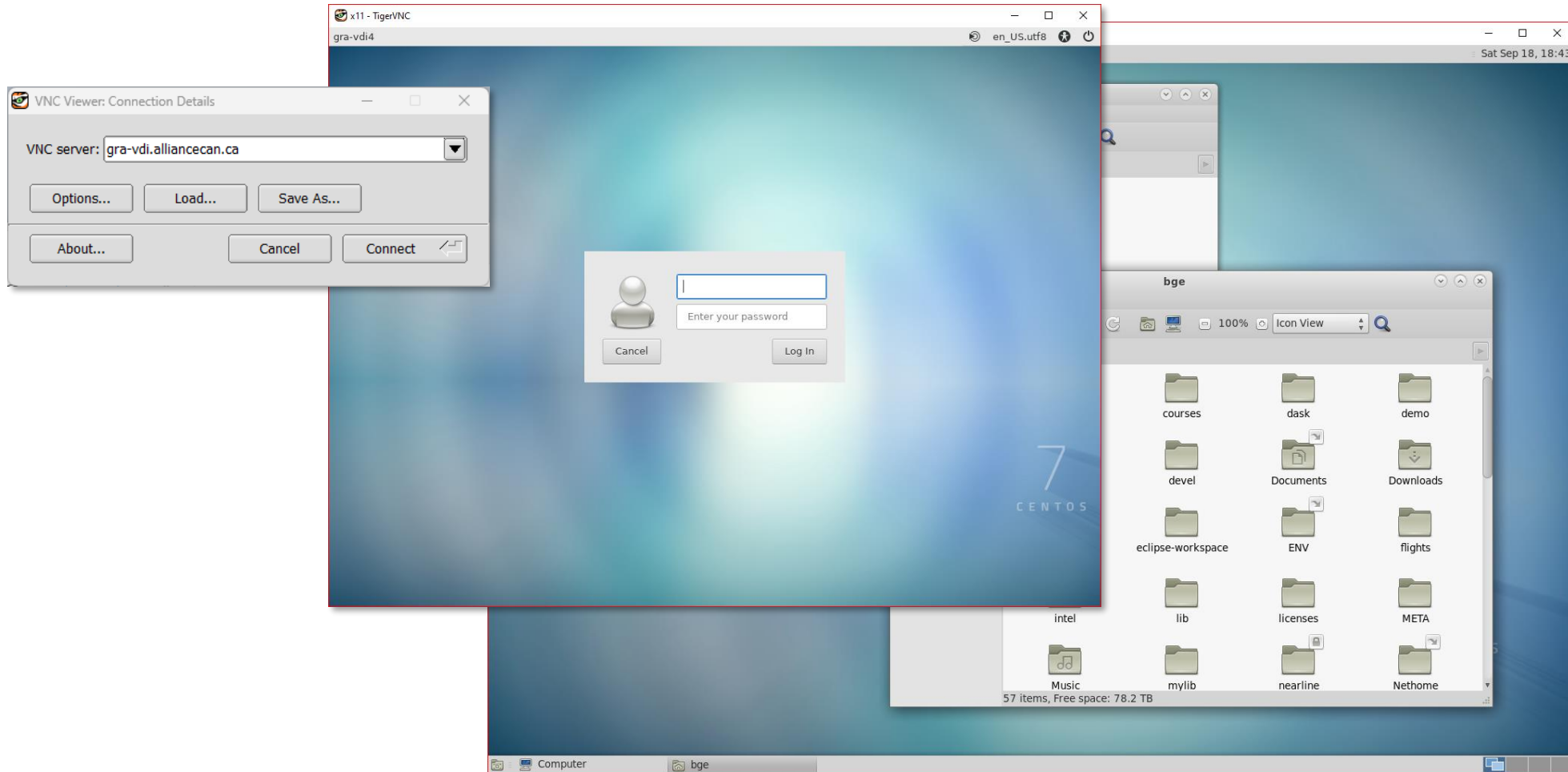


Supercomputing – Connecting via JupyterHub

Access to the cluster via Jupyter Notebook, terminal and desktop in a browser

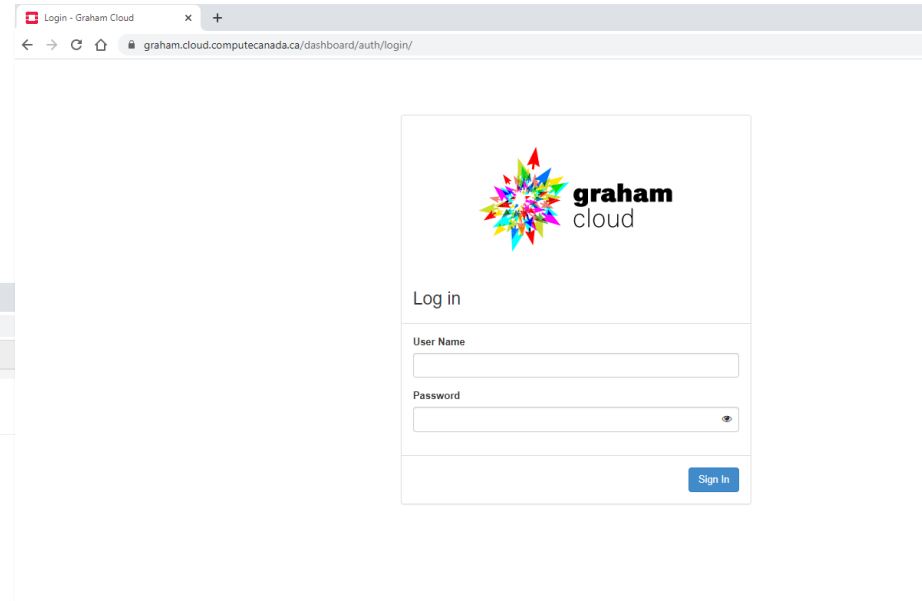
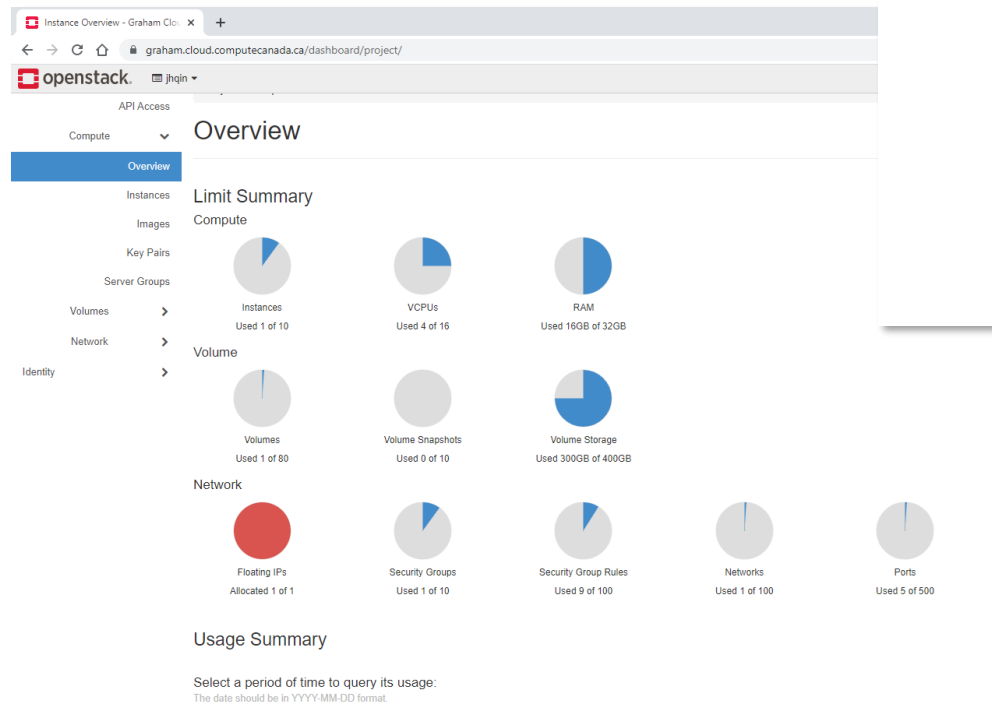


Supercomputing – Connecting via VDI



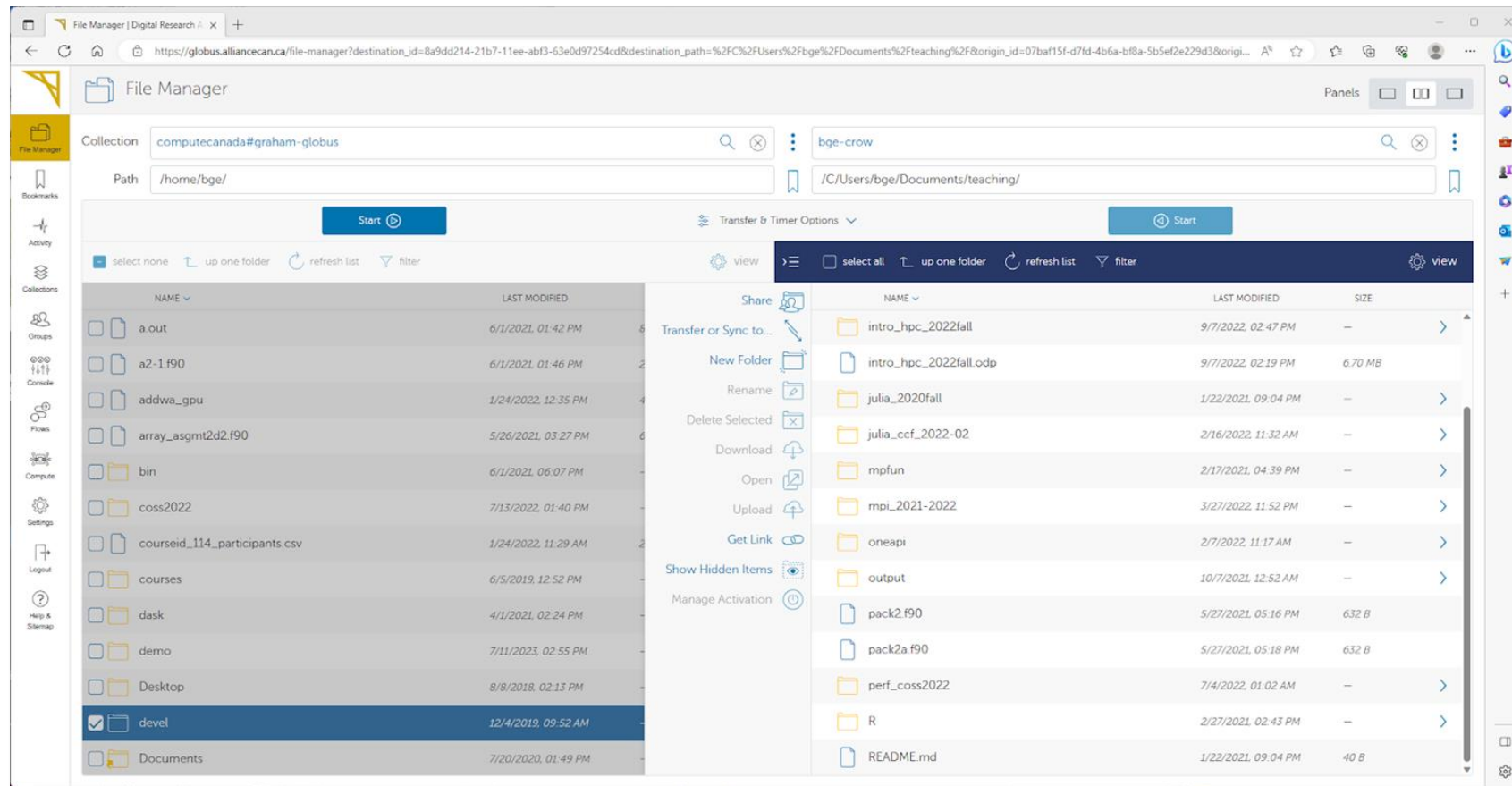
Supercomputing – Connecting to Cloud

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



- A cloud project account is required
- No Windows images, all VMs are **Linux**
- Up to **80** virtual CPU cores
- May have virtual GPUs
- Up to **10 TB** storage
- Up to **25** persistent CPU cores

Supercomputing – Transferring Files via Globus



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

Supercomputing – Getting Help

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

The screenshot shows the 'Technical documentation' page on the SHARCNET wiki. The page title is 'Technical documentation' and it includes a search bar and a language selector set to 'English'. A blue banner states: 'This site replaces the former Compute Canada documentation site, and is now being managed by the Digital Research Alliance of Canada. Ce site remplace l'ancien site de documentation de Calcul Canada et est maintenant géré par l'Alliance de recherche numérique du Canada.' Below this, there are sections for 'Other languages' (English, français) and a welcome message: 'Welcome to the technical documentation wiki of the Digital Research Alliance of Canada (the Alliance). This is the primary source for users with questions on equipment and services of the Alliance. The focus here is on national services and systems. For documentation on services and systems managed by our regional partners, please use the links provided below. This wiki is a constant work-in-progress and some links might take you to pages which do not yet have content; such pages display like this. Our staff works constantly to improve and expand the available documentation; note however that any of our users is free to add new material and edit existing content.' The page is divided into two main columns: 'Systems and services' and 'How-to guides'. The 'Systems and services' column lists 'General-purpose clusters' (Béluga, Cedar, Graham, Narval, Known issues) and 'Niagara, a cluster designed for large parallel jobs'. The 'How-to guides' column lists 'Getting started' (Getting started with the new national systems, Niagara Quick Start Guide, Cloud Quick Start Guide, SSH - How to connect to our compute clusters, Linux introduction), 'Storage and file management' (Transferring data, Scratch purging policy), 'Best practices for data migration', 'Using modules and Standard software environments to access software', 'Running jobs' (Installing software yourself, Programming guide, Visualization, How to get technical support), and 'FAQ: Frequently Asked Questions'.

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

The screenshot shows the SHARCNET website homepage. The header includes 'SHARCNET' and navigation links for 'FACILITIES', 'SUPPORT', and 'ABOUT US'. A large banner features a list of events: 'New user seminar every Tuesday at 2pm EST', 'Weekly colloquium at noon on Wednesdays', '@SHARCNET', and 'YouTube youtube.sharcnet.ca'. Below the banner is a section titled 'Neutrinos by the Numbers: Sudbury's SNOLAB' with a sub-headline: 'Some of the greatest ongoing problems in particle physics may be solved by taking a closer look at theta_13'. The main content area is divided into 'EVENTS' and 'NEWS'. The 'EVENTS' section lists several 'New User / Refresher Webinar' sessions on Tuesdays from September 26 to October 24. The 'NEWS' section includes 'RAC 2024 launches on September 26, 2023', 'SHARCNET Dedicated Programming Support Round XXII Call f...', '2023 Compute Ontario Summer School', 'SHARCNET awards Round XXI Dedicated Programming Support', and 'New Weekly Colloquia Series'. A 'GitLab Instance at SHARCNET' section is also visible, with a sub-headline: 'Version control your work and collaborate using GitLab at SHARCNET. All the public and private repositories that you will need!'. The footer features logos for 'Digital Research Alliance of Canada' and 'Compute - Calcul Ontario', along with social media icons for Twitter and YouTube.

Supercomputing – Getting Help

Local Staff @ OntarioTech



Armin Sobhani

HPC Technical Consultant



Faculty of Science, UA 3020



905-721-8668 x3607



asobhani@sharcnet.ca



armin.sobhani@ontariotechu.ca



<https://staff.sharcnet.ca/asobhani>

Arrange an Office Visit for:

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

...

Supercomputing – Getting Help

Help from beyond **OntarioTech**

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/my/contact/directory>

What PIs Should Know



Resource Allocation Competition (RAC)

RAS

Rapid Access Service

AKA Default Allocation

CPU/GPU: Opportunistic

Storage: 40 TB across all clusters

Cloud compute instance: 1 month

Cloud persistent instance: 1 year

RAC

Resource Allocation Competition

Annual (in fall)

Peer-reviewed
(scientific and technical)

Roughly 80% of resources

Resource Allocation Competition (RAC)

RAS

Rapid Access Service

AKA Default Allocation

CPU/GPU: Opportunistic

Storage: 40 TB across all clusters

Cloud compute instance: 1 month

Cloud persistent instance: 1 year

RAC

Resource Allocation Competition

Annual (in fall)

Peer-reviewed
(scientific and technical)

Roughly 80% of resources

Resource Allocation Competition (RAC)

RAS

Rapid Access Service

AKA Default Allocation

CPU/GPU: Opportunistic

Storage: 40 TB across all clusters

Cloud compute instance: 1 month

Cloud persistent instance: 1 year

RAC

Resource Allocation Competition

Annual (in fall)

Peer-reviewed
(scientific and technical)

Roughly 80% of resources

RAC – Types

RRG

Resources for Research Groups

Resource: **Compute and/or Storage**

CPU \geq 200 core-years, or

GPU \geq 25 reference GPU units (RGU) years, or

or project storage \geq 41 TB, or

or nearline storage \geq 101 TB

RPP

Research Platforms and Portals

Resource: **Cloud**

Persistent cloud storage \geq 1 TB, or

Compute cloud \geq 81 vCPUs, or

Persistent cloud \geq 26 vCPUs

RAC – Types

RRG

Resources for Research Groups

Resource: **Compute and/or Storage**

CPU \geq 200 core-years, or

GPU \geq 25 reference GPU units (RGU) years, or

or project storage \geq 41 TB, or

or nearline storage \geq 101 TB

RPP

Research Platforms and Portals

Resource: **Cloud**

Persistent cloud storage \geq 1 TB, or

Compute cloud \geq 81 vCPUs, or

Persistent cloud \geq 26 vCPUs

RAC – Types

RRG

Resources for Research Groups

Resource: **Compute and/or Storage**

CPU \geq 200 core-years, or

GPU \geq 25 reference GPU units (RGU) years, or

or project storage \geq 41 TB, or

or nearline storage \geq 101 TB

RPP

Research Platforms and Portals

Resource: **Cloud**

Persistent cloud storage \geq 1 TB, or

Compute cloud \geq 81 vCPUs, or

Persistent cloud \geq 26 vCPUs

RAC – 2025 Important Info

Important Dates

Applications accepted:
September 24 to October 30, 2024

Announcement of Results:
late March 2025

Implementation of allocations:
early July 2025

Info Sessions Recordings

RAC



GPU



Cloud



Dedicated Programming by SHARCNET

SHARCNET's staff spending 50% of time working with the PI on specifically defined programming tasks

The DP program runs for about 4 months and can be renewed

There are two to three calls a year for PIs to apply

The applications are reviewed based on the scientific merits and the feasibility of the proposed programming project

Check: <https://www.sharcnet.ca/my/research/programming>



What Grad Students Should Know



What People Do on Supercomputers



Astrophysics Simulations

CFD, Environmental Simulations, etc.

Coupled Simulations

Material Science

Applications of AI Machine/Deep Learning

Economics Finance etc.

Some Facts

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

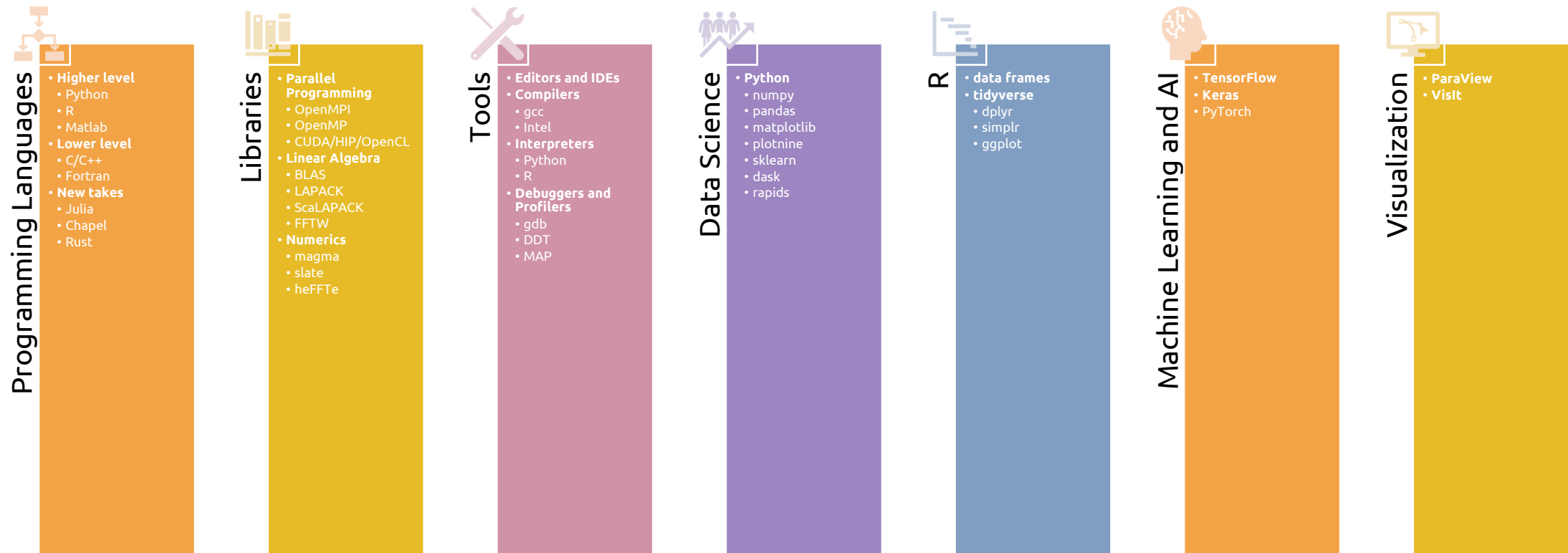
Research supercomputers run Linux and not Windows

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

High performance programming is hard

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

Popular and Lesser-Known Items



SHARCNET's Online Training

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



Introduction to ARC Series

Intro to ARC

Starts in early October

Two 1-hour classes each week

Include lectures and labs

Topics

HPC Python

Julia

C++

PyTorch

Jupyter

Apptainer

CMake

MPI

MATLAB

SHARCNET's Self-Paced Courses (NEW)

Self-Paced Courses

Just google for SHARCNET self-paced

Online self-directed courses, with quizzes and/or home assignments

If successfully completed, the Certificate of Completion will be issued

Available Courses

Introduction to GPU Programming

Introduction to Machine Learning

Introduction to the Shell

Coming Soon

Introduction to SQL

Introduction to Supercomputing

New Alliance User course

Parallel Code Debugging and Profiling

Other SHARCNET's Online Training

Weekly Compute Ontario Colloquia Series

Every Wednesday at 12pm-1pm, online

Topics of general interest to our users

Delivered by SHARCNET, SciNet and CAC staff, plus guest speakers

No registration is required

Alliance account is not required

Compute Ontario Summer Schools

Every June

Three weeks, 2 parallel streams, ~40 courses

Alliance account is not required

SHARCNET's YouTube Channel

<http://youtube.sharcnet.ca>



Hundreds of webinar recordings and short videos

Q&A

