
Introduction to local and National HPC at UManitoba

How to use available software and run jobs efficiently.

UofM-Autumn-Workshop 2022
October 2022

Ali Kerrache and Grigory Shamov

- Connect via GoogleMeet (best in Chrome)
- Meeting link: <https://meet.google.com/jrn-sxmp-oub>
- Mute your mic please, unless asking questions!
- Slides will be published at Grex documentation website

Presenters: Grigory Shamov (HPC Site Lead)
Dr. Ali Kerrache (HPC Specialist)

- Updates on Grex HPC machine status, new hardware and services (30 min, Grigory)
 - Digital Research Alliance updates
 - Grex Resources update: available compute and GPU
 - Conditions on using contributed systems
 - Grex storage update, new /project and related changes
- Beginner how-to using HPC machines Part I (40 min, Ali)
 - Getting account and connecting, transferring data
 - Finding software, SLURM scheduler, kinds of jobs
- Break (10-15 min)
- Beginner how-to using HPC machines Part II (40 min, Ali)
 - Using HPC clusters efficiently
 - SLURM Partitions, memory
 - Handling many small jobs

- HPC software overview and best practices (1 h, Ali)
 - How to find software on HPC
 - Environment Modules
 - Installing software from sources
 - Installing libraries for Python and R
 - Software stacks on Grex, CVMFS
- Break (10-15 min)
- Using OOD Web portal on Grex (30 min, Grigory)
 - Connecting to OOD
 - Examples of using interactive Desktop sessions
 - Examples of using Applications (Matlab and Jupyter)
- Announcements of related courses, Q/A (10-30 min, Everyone)

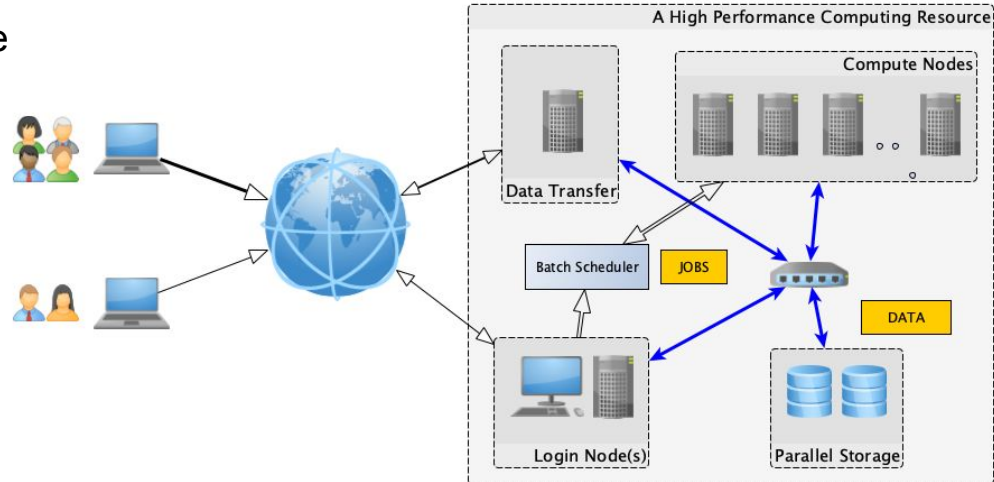


- Know about available High-performance Computing options for UManitoba
 - Local HPC (GreX) , National DRI (The Alliance) resources
 - Updates about their status and capabilities
 - Available user support and documentation
- Being able to use HPC systems (Scheduler, Software, Storage, etc.)
- Being more efficient in getting most of HPC systems
- Covers some GreX-specific topics like using GPU and OOD Web portal



What is High Performance Computing

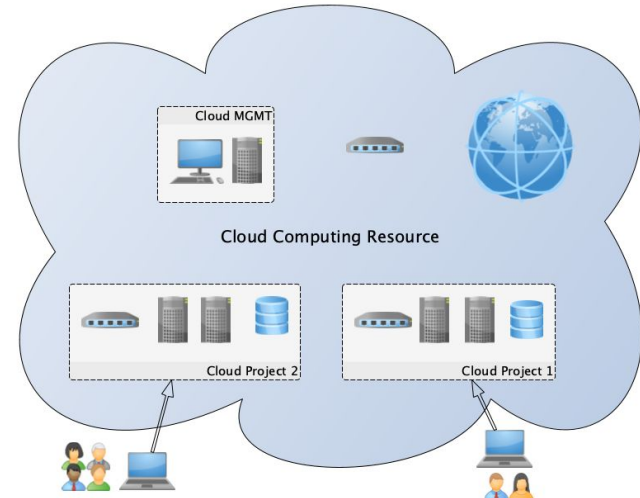
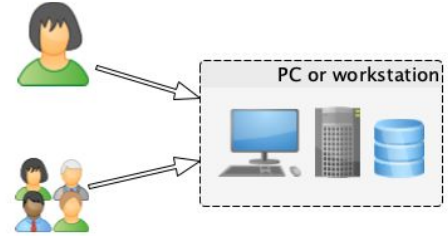
- HPC clusters architecture and how is it different from Cloud and PC computing
 - HPC clusters are *shared* systems with *remote access*
 - They are also “distributed computing systems”, built from many “nodes”, linked with a fast interconnect and under management of a Scheduler software
 - Command line vs Graphical user interface
 - Batch mode vs Interactive mode
 - Software delivery on HPC
 - Dealing with Data (storage, transfer etc.)





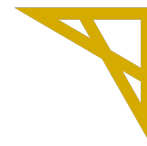
What is not (trad.) High Performance Computing

- Individual PCs or Workstations, or sets of Workstations
 - These can be *shared* systems with *remote access*
 - Scaling up, limited by resources of a single node
 - Interactive workloads, resource management problematic
- Public and Community Cloud computing
 - A way of set up a cluster of nodes.
 - Flexible, elastic, Provides isolation of tenants
 - Can be configured as Workstations of HPC on the cloud
 - Can be configured for burst usage and autoscaling
 - Difficult to set up, DIY or costly
- In this Workshop, we will focus on HPC instead of the above!





**University
of Manitoba**



**Digital Research
Alliance** of Canada

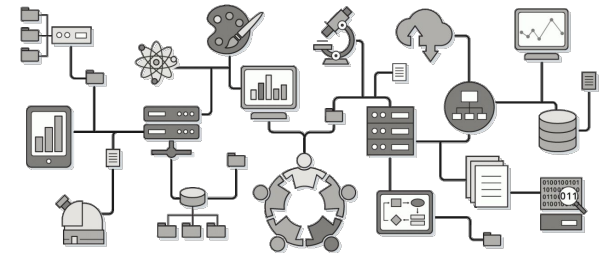
Updates for local and National HPC

Grigory Shamov
October 26, 2022 HPC workshop



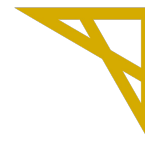
Digital Research Alliance of Canada (DRAC, The Alliance) succeeded ComputeCanada

- Website, contacts migrated
 - <https://alliancecan.ca/en/services/advanced-research-computing>
 - <https://alliancecan.ca/en/services/advanced-research-computing/accessing-resources/resource-allocation-competition> : RAC 2023 call is on: submission window Sept 22 to **Nov 2**.
 - <https://ccdb.computecanada.ca> – CCDB is being used for account management and RAC
 - **~20K** active roles total in CCDB as of now
- Support contact is now support@tech.alliancecan.ca
- Westgrid Corporation is dissolved
 - Used to be our regional consortium as part of CC Federation
 - Split to two regions: BC and Prairies DRI Groups.
 - **2495** CCDB roles from Prairies, **2966** from BC DRI
 - Manitoba is now part of the Prairies DRI.





| System, kind | Commission date | CPU cores | GPU devices | Project storage, PB |
|---------------------|------------------------|------------------|--------------------|----------------------------|
| Arbutus, Cloud | 09. 2016 | 16,008 | 108 | 12 |
| Beluga , HPC | 09. 2019 | 28,960 | 688 | 17 |
| Cedar, HPC | 03. 2017 | 94,528 | 1352 | 19.5 |
| Graham, HPC | 06. 2017 | 34,784 | 498 | 13 |
| Narval, HPC | 09. 2021 | 61,760 | 524 | 14 |
| Niagara / Mist HPC | 03. 2018 | 75,840 | 64 | 2.7 |
| | Total: | 295,872 | 3,126 | 66,230 |



Digital Research Alliance of Canada (DRAC, The Alliance) succeeded ComputeCanada

<https://alliancecan.ca/en/services/advanced-research-computing>

- Operations and Security migrated from ComputeCanada to DRAC.
 - National Hosting Sites are ran by Regions/Universities, supported by “the Federation”
 - Most National Teams (RSNT, IONT etc.) are at Federation, some (RDM) are at DRAC
- Funding proposal and business plans submitted to ISED
 - Crated with major help of the Regions
 - Two years left to go for DRAC in this funding cycle
- The Alliance’s MYFP aims at HPC Critical Infrastructure Renewal in 2023/4, 2024/5.
 - Refresh of the ComputeCanada infrastructure planned to start in the 2023 fiscal year
 - Subject to availability of “matching” funds. **~230M** total if matched.
- Several other projects submitted (Research Data Management , RSE).
 - FRDR (pronounced as “further”) <https://www.frdr-dfdr.ca/repo/>
 - “Innovative funding opportunities”, Quantum computing, etc.



GreX is a High-Performance-Computing machine

Hardware:

- The “legacy” compute part from 2010, 240 nodes left
- Infiniband interconnect (mix of 100, 56 and 40 GB/s)
- UM added 54 compute and GPU nodes in 2019-2021
- Contributed GPU nodes
 - 3x NVIDIA V100*4 NVlink (Physics)
 - 1x NVIDIA V100*16 NVSwitch (CompSci)
 - 2x NVIDIA A30*2 AMD (Agriculture)
- Storage : **15TB** NVMe (/home); **418 TB** Parallel Lustre FS (/global/scratch)
- **New storage: 1.1 PB Parallel Lustre FS (/project)**





GreX is a High-Performance-Computing machine

Software:

- Linux (CentOS7)
- SLURM scheduler
- User management software from CC (CCDB)
- Curated, maintained local software stack
- ComputeCanada CVMFS software stack
- Singularity/Apptainer containers (NVIDIA NGC cloud, etc.)
- **OnDemand Web Portal (OOD)**



Support: ComputeCanada OTRS:

- support@tech.alliancecan.ca
- <https://support.computecanada.ca/otrs/>

Documentation:

<https://um-grex.github.io/grex-docs/>

Status page:

<https://grex-status.netlify.app/>

- Migration to UM Network from BCNet : Complete
 - New domain .hpc.umanitoba.ca replaces .westgrid.ca
 - External IP addresses and network connectivity changed to UMan campus.
- Enabling new 1.1PB Lustre /project fs
 - Will have hierarchical structure similar to CC, with directory quota
- Accepting Contributed systems according to our Guidelines.
- **A 2022/23 Local RAC was called on Sept 20 and concluded on Oct 15**
 - 13 applications for over 3K Core years received, will be scaled to fit



- Enabling new 1.1PB Lustre /project fs
 - Will have hierarchical structure similar to CC, with directory quota
 - */project/dev-PIname/user* or */project/Faculty/dev-PIname/user*
 - Under user's */home/\$USER/Projects* , symbolic links will be created for their projects
- **Storage policy implications of the change**
 - The /project data is associated with PI's accounting group/allocation, and not individual users
 - Members of more than one project will have more than one /project space and quota
 - The data on /project is implicitly shared/owned by the PI of the allocation.
 - Previously , and currently, data on /home/\$USER and /scratch/\$USER were owned by individual users by default
 - Data sharing between members of a PI's group / allocation is easy.
 - Data sharing between different group requires ACLs or Linux groups setup , as before.

- Many University-level HPC systems are expanded with the help of the Researchers community
 - Sherlock at StanfordU is an example <https://www.sherlock.stanford.edu/>
 - CFI had mandated that computing systems over 100K worth to be integrated into National systems (Cedar, Graham, etc.)
 - At UManitoba, we have already 4 contrib. system integrated to Grex, and more coming
- **Contrib systems policy**
 - The PI owns the systems. Their group has preferential access to them
 - Preferential means preemption with SLURM scheduler, at a grace period (1h currently)
 - ComputeCanada/DRAC has a different mechanism, 12h max job time
 - Other users can use the contrib systems when they are idle
 - A good way to run a lot of shorter GPU jobs on Grex (MD, ML, Matlab etc.)



Grex and CC HPC usage over last RAC year

April 30, 2021 to May 1, 2022 (not counting Storage use)

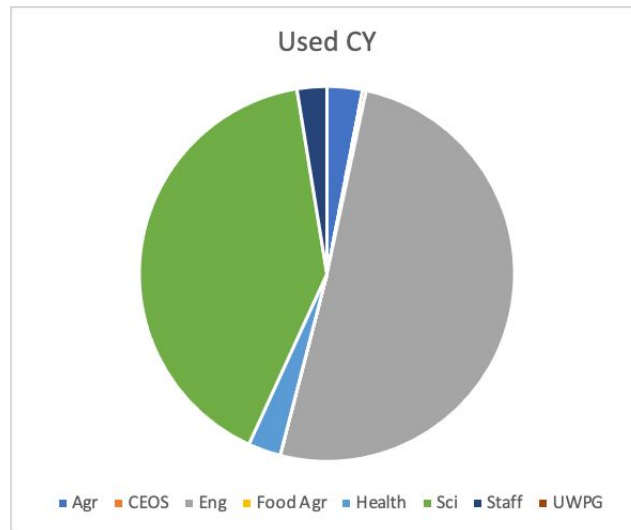
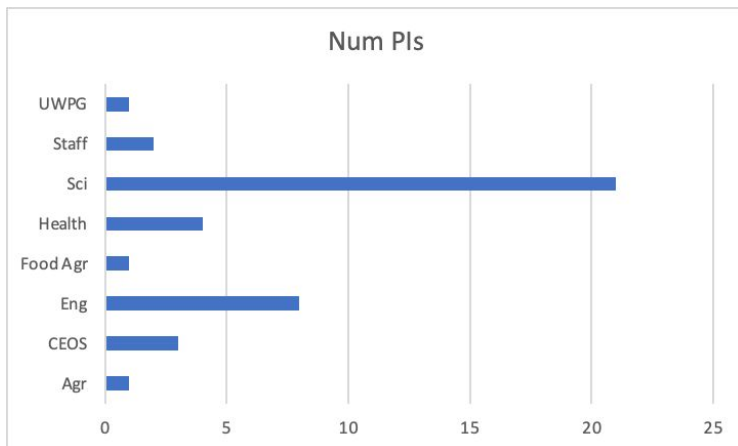
- Grex Total CPU Usage: **3224.31** Core Years from **41** PIs
- ComputeCanada CPU Usage: **2297.09** Core Years from **54** PIs
- ComputeCanada GPU Usage: **2.13** GPU Years from **12** PIs

\$ ~**320K** (Azure D32 v3 : 1.12M)

\$ **279K** (CC RAC estimate)

\$ **5.2K** (CC RAC estimate)

Grex CPU usage per Faculty



Thank you for your attention!