Table of Contents

Getting Started

This guide is designed to get up and running as a new user to the Digital Research Alliance of Canada. It covers the following topics:

-

Below is an overview of the main types of nodes typically found in a DRAC HPC environment. This is intended as a brief over

1. Login Nodes

Purpose

- Entry point to the cluster where you initially land after logging in.
- Lightweight tasks: editing files, preparing and submitting jobs, managing data.

Important Usage Notes

- Not intended for running resource-intensive computations.
- Long-running or CPU-heavy processes can degrade system performance for all users.

References

-
-

2. Compute Nodes

Purpose

- Actual workhorses of the cluster, executing the vast majority of computational tasks.
- Accessed via the job scheduler (e.g., `sbatch`, `srun`, `salloc` for Slurm).

Important Usage Notes

- Must be allocated through the scheduler to ensure fair resource sharing.
- Optimized for heavy computation, memory, or GPU workloads depending on node configuration.

References

-
-

3. Data Transfer Nodes (DTNs) or Data Mover Nodes

Purpose

- Designed to handle file transfers in and out of the cluster.
- Optimized with high-bandwidth connections and specialized configurations to move large datasets efficiently.

Important Usage Notes

- Preferred nodes for large or frequent data transfers (e.g., using `scp`, `rsync`, or Globus (https://www.globus.org/)).
- Helps avoid overloading login nodes with data copy tasks.

Introduction to HPC

A brief overview of what High Performance Computing (HPC) means in the context of the Digital Research Alliance of Canada is presented here. For more details about HPC is available at the Alliance Wiki. For our purposes we will treat this as largely supplementary information.

Nodes

Login Nodes

Compute Nodes

Etc

Systems

The Digital Research Alliance of Canada provides several advanced research computing systems and services to support Canadian researchers:

High-Performance Computing Clusters

The Alliance operates several national HPC clusters (below). Each of these clusters has slightly different behaviours and for specific information the

- **Fir**: Replacing the Cedar cluster
- **Rorqual**: Replacing the Béluga cluster
- **Trillium**: Replacing Niagara and Mist clusters
- **Nibi**: Replacing the Graham cluster

Cloud Computing

The Alliance provides cloud computing services, including:

- **Arbutus**: A virtual infrastructure cloud system
- **Cedar Cloud**: Offering persistent and compute cloud resources[14]

Storage Systems

Large-scale storage is available on the HPC clusters, including:

- **Fir storage**: Offering backup storage capabilities
- Storage on other clusters like Graham2[14]

Additional Services

- **Database servers**

- **Globus file transfer service**
- **Nextcloud cloud storage** [5]

Software and Support

- Hundreds of pre-installed research software packages
- Technical support and training for users
- Access to visualization systems[4]

The Alliance aims to provide a comprehensive digital research infrastructure to advance Canada's research capabilities across all academic disciplines. Researchers can access these systems free of charge, with larger resource allocations available through a competitive application process[15].

Citations: [1] https://www.uvic.ca/systems/researchcomputing/about-us/digital-research-alliance-of-canada/index.php [2] https://alliance.smapply.ca [3] https://ised-isde.canada.ca/site/digital-research-infrastructure/en [4] https://www.ualberta.ca/en/information-services-and-technology/research-computing/digital-research-alliance-canada.html [5] https://docs.alliancecan.ca/wiki/Technical_documentation [6] https://docs.alliancecan.ca/wiki/Infrastructure_renewal [7] https://docs.alliancecan.ca/wiki/Available_software [8] https://ccdb.alliancecan.ca/security/login [9] https://alliancecan.ca/en/latest/news/alliance-receives-85m-advance-canadas-digital-research-infrastructure [10] https://alliancecan.ca/en/services/advanced-research-computing/accessing-resources [11] https://alliancecan.ca/en/services/advanced-research-computing/federation/national-host-sites [12] https://alliancecan.ca/en/services/advanced-research-computing/national-services [13] https://alliancecan.ca/en/initiatives/canadian-digital-research-infrastructure-needs-assessment [14] https://alliancecan.ca/en/services/advanced-research-computing/accessing-resources/resource-allocation-competition/available-resources [15] https://alliancecan.ca/en/services/advanced-research-computing/national-services/clusters [16] https://csc.ok.ubc.ca/catalogue/digital-research-alliance-of-canada/ [17] https://alliancecan.ca/en/our-services [18] https://alliancecan.ca/en/about/alliance [19] https://alliancecan.ca/en [20]

https://alliancecan.ca/en/services/research-software/canadian-research-software-platforms [21] https://engage.alliancecan.ca

Navigating the Environment

Start typing here...

Hello World

To get started with a simple practical example, we assume that you have successfully logged in to one of the nodes listed in Introduction to HPC (Nodes). You should be on a login node ("Login Nodes" in "Nodes") for some general purpose system to start this guide.

To schedule and manage jobs submitted on our clusters, we use the SLURM Workload Manager (https://en.wikipedia.org/wiki/Slurm_Workload_Manager). Documentation can be found here (https://slurm.schedmd.com/documentation.html). For now that will not be needed.

> ⚠ All jobs must be submitted via the scheduler!

The login nodes are not suitable to computation, it will likely be slower and more limiting than running things locally.

Instead, we can dispatch slurm scripts as seen below.

Managing Results

Start typing here...

Multiprocessing

Start typing here...

GPU Jobs

Start typing here...