Compute Canada Manual

Author	Alice Zhang
Department	Centre for Heart Lung Innovation
Principal Investigator	Dr. Zachary Laksman
Email	a.zhang.4@alumni.ubc.ca
Date	August 17, 2020

Contents

1. Summary

This document includes instructions to set up and run several operations remotely using Compute Canada. This includes setting up an SQL database, mounting a database to your local computer, and running programs remotely through the CC Arbutus cloud. All instructions were written under the assumption that a CC account has already been created, and the account has access to a Cloud project. The attached instructions were initially created for the purpose of running a CellProfiler image analysis pipeline remotely. Some sections may not apply to your project type. Because CellProfiler does not work well with the traditional HPC cluster available, we opted to use the CC cloud resources. If your are working with a Linux subsystem and your program runs well using Linux, you can likely run it remotely using the default HPC cluster on CC.

A big thank yout o Lucas Whittington, Mathew Larade, Erming Pei, Jiarui Liu, Venkat Menhadevan, and Ken Bigelow from Compute Canada, who enabled me to compile this document. Janet Liew's assistance was also essential for us in applying for and setting up the CC resources we used. If you need technical support with CC, contact support@computecanada.ca

2. CC MySQL Databases

2.1 Creating and accessing your database

If you need the privileges to create your own database, you will need a MySQL account which will connect to MariaDB: the branch of the Cedar server you'll be accessing. To get an account, please send a request to support@computecanada.ca with the following information:

- Your CC username
- The amount of database space you need

If you are working with a Windows or Mac operating system (OS), you can use an Ubuntu subshell to access and test your database. This step is not strictly necessary, but Ubuntu is a useful resource to access your database, and allows you to use Linux commands without changing your operating system. You can install Ubuntu directly from the Windows store.

Once Ubuntu has been installed, launch it and connect to the database by entering the following commands:

1. ssh -Y alice35@cedar.computecanada.ca
2. Enter your password and a block should appear welcoming you to Cedar.
3. module load mariadb
4. mysql –version
5. mysql
6. show databases;

If you would like to test your account's permissions to write to the database:

1. CREATE TABLE tab (name VARCHAR(20), birth DATE);
2. SHOW TABLES; (the new table should appear)
3. DROP TABLE tab;

2.2 Mounting your database to your local computer

To export files to your database from local programs such as CellProfiler, you will need to mount a remote hard drive to your local computer.

Tools you will need include:

- winFSP
- SSHFS-Win

If you are working with a Windows OS:

1. Open file explorer.
2. Right click on This PC and select 'Map network drive'
3. Choose a drive to mount to.
4. In the folder section type: \\sshfs\yourCCusername@cedar.computecanada.ca

You should see a new drive appear in your Network Locations.

If you are working with a Mac OS:

2.3 Building a VM to run programs remotely

Detailed instructions can be found here for Linux and Windows systems. There are currently no instructions for Mac OS.

Alternative Approach

Alternatively, you could try the command line version of CellProfiler. The benefit is that you can run it in the command line and Linux platform, which is the main platform of CC. So it means you can access to more CPUs, GPUs, RAM. However, it is much more complicated to make it work. For example, you need to install miniconda first. This Linux version is also less stable and different from Windows or Mac version of CellProfiler. So in general, it is an option but for CellProfiler a VM would likely be a better approach.