Kevin Wang

2019 Semester 2 Progress

Aim

1. For the group, this is a framework for running a scalable workshop.

2. For me, this is a way to learn the general framework of Docker + GC.

Memory usage

- Why did I go through the troubles with all these technologies when laptops, RStudio Cloud, Travis and Binder exists?
- Memory (computation) usage of this workshop:
  - scMerge: 4 GB
  - monocle: 27 GB

Additional info

- Today, we are running two 16 cores, 60GB RAM virtual machines with 50 GB of hard disk space from the US.

- Machine 1:

- Machine 2:

1. Setting up Google Cloud

- Type `source("/home/setup.R")` into your console.

- This copies a **frozen** version of the `SydneyBioX/SingleCellPlus` repo into your working directory.

2. Testing out if we blow up the machine

- Type `source("./scMerge.R")` into your console.

- This will run all the necessary computations in this workshop. Single run should take 2 minutes.

- You could also try `source("/home/user_test.R")`

While we wait

- Docker is a software and a company.
- Running a Docker image is like running a virtual machine, but more lightweight and more easily distributed.
- A Dockerfile tells Docker how to make an image.

While we wait

- Google Cloud takes in a Dockerfile that I wrote in `kevinwang09/scp_docker` and builds that on the Cloud. (i.e. codes are reproducible)
- The built image is available on the GC and can be accessed and installed by any of GC's virtual machines. (i.e. hardwares are reproducible)
- Once a VM is set up, I simply upload all the usernames and passwords.

3. Call me sudo

- System admin can edit all files on disk, including `sudo rm -rf /home/ellis` if Ellis asked too many questions.

- Users can just type `source("/home/omg.R")` into your console to edit the files on WD.