

C3G Analysis Workshop: RNA-Seq

Concepts from *personal* to *super* computer

January 22-23, 2019



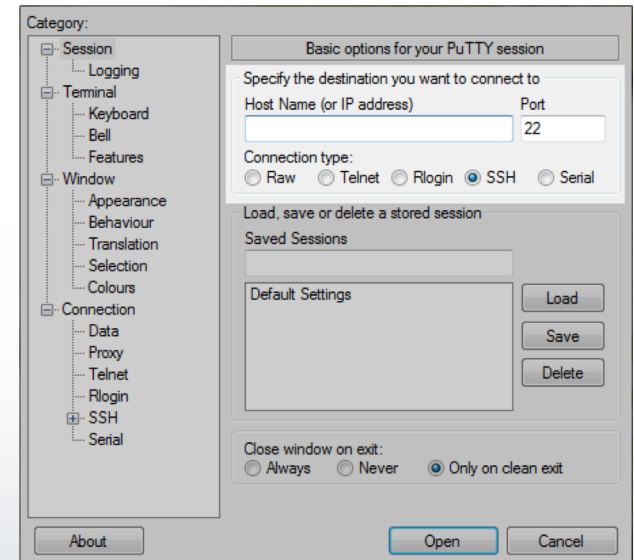
Canadian Centre for
Computational
Genomics

Test your account

Open a bash terminal and type:

```
ssh my_cc_account@mp2.ccs.usherbrooke.ca
```

On older windows system, use putty:



Overview



Canadian Centre for
Computational
Genomics

Calcul Québec, Calcul/Compute Canada

A simple computer

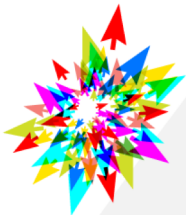
- User Interface
- A system that counts
- A system that stores

Supercomputers

- Pipelines
- Genpipes

Pause

Exercises



compute
canada | calcul
canada

Each researcher in
Canada has a 10 TB
space and 50 Core-
years on all sites

CCF national data
centres support
researchers across
Canada

Arbutus/UVic

Cedar/SFU

Beluga/ETS-McGill

Niagara/UofT

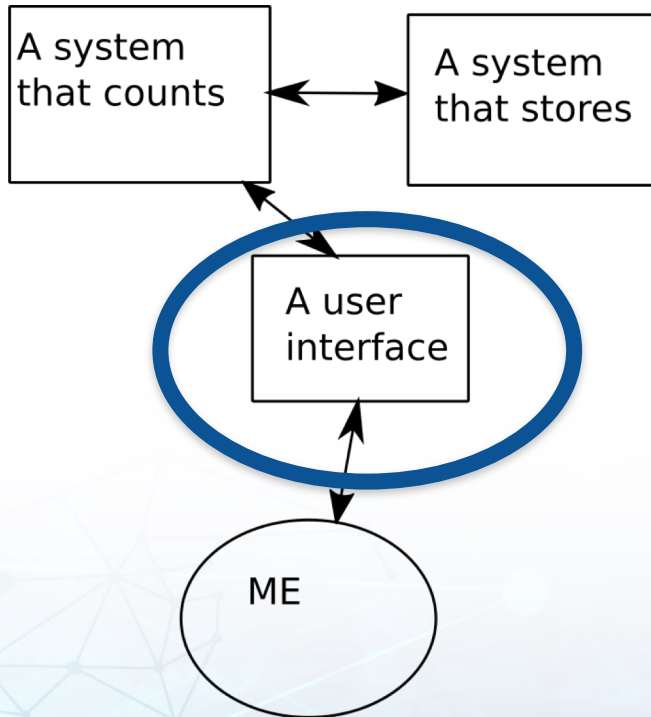
Compute • Calcul
Ontario

Calcul Québec

Graham/Waterloo



What is a computer?



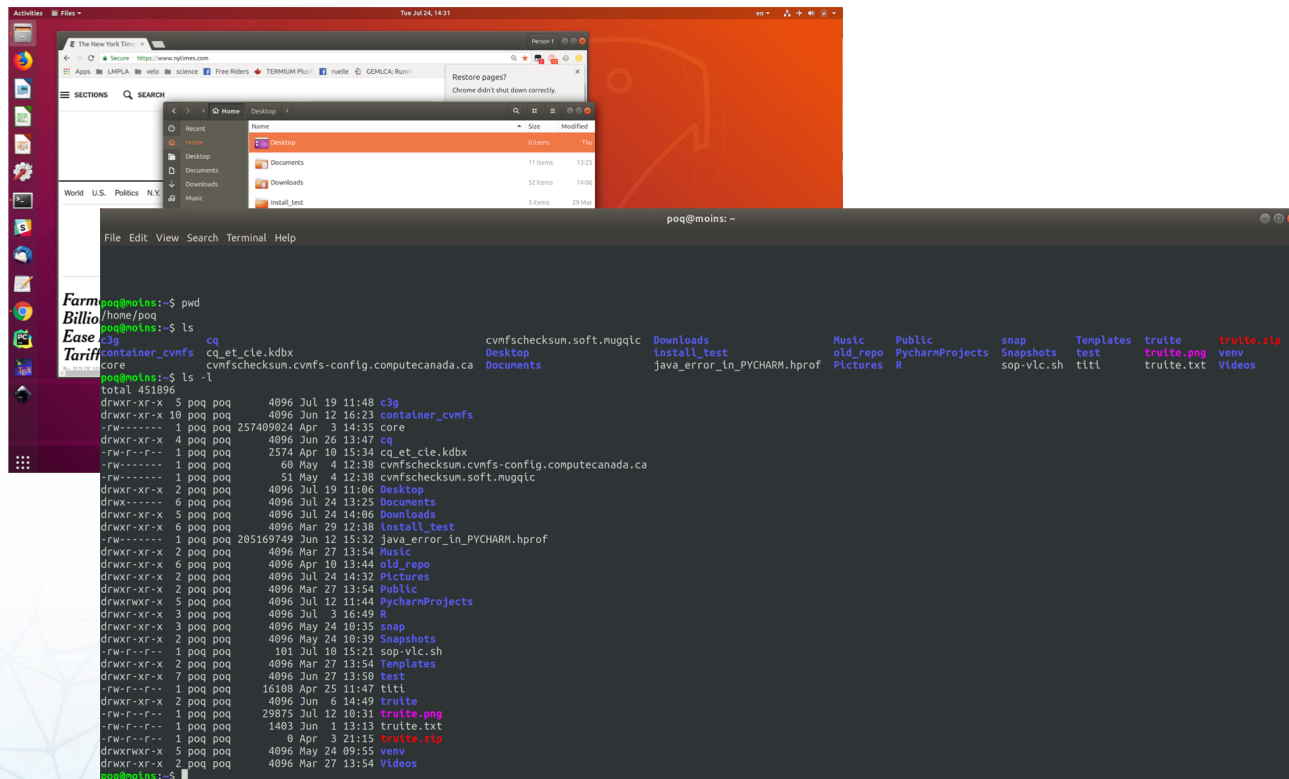
Here is a rather
simplified version of
one

A User Interface (UI)



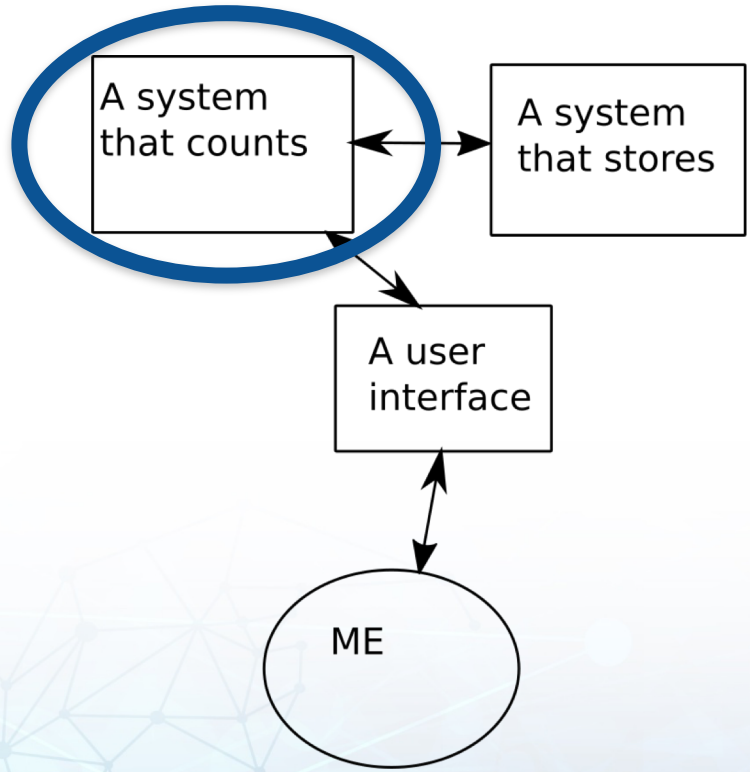
Canadian Centre for
Computational
Genomics

GUI



UI

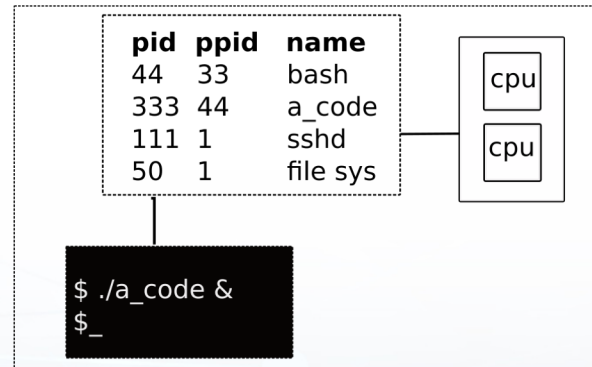
Here it is again



A system that counts

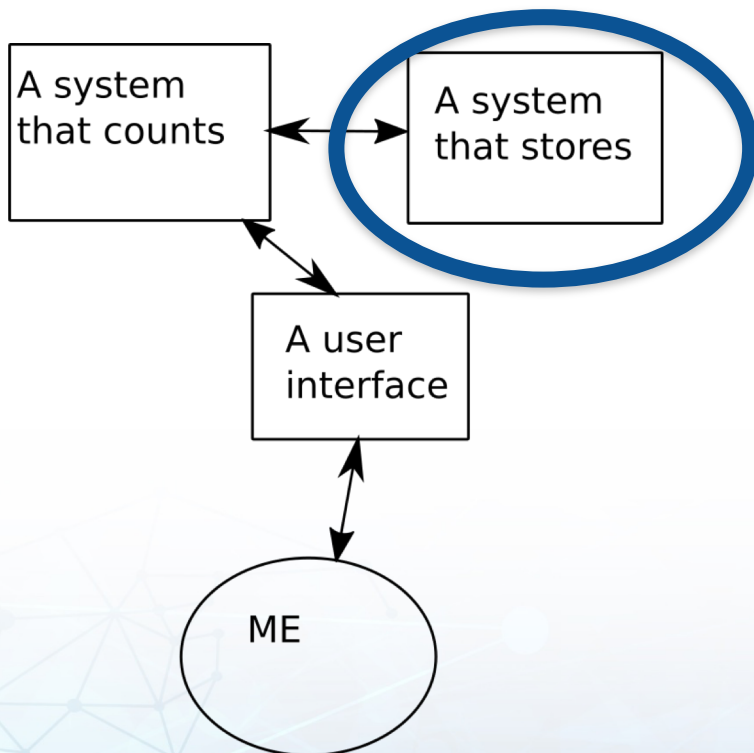


- When you type commands, you will type them in the terminal, or for that matter every time you click on a new icon on your computer, a script/binary/software is *executed*
- Once executed, it *transforms* into a process



- This process has ways to talk to the external world: streams. There are three "standard" streams called stdin, stdout and stderr.

One more time

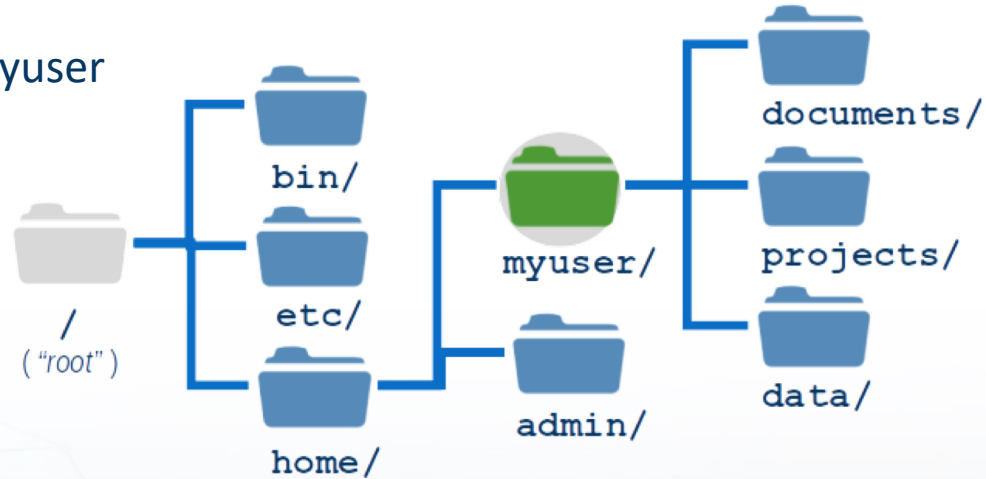


A system that store

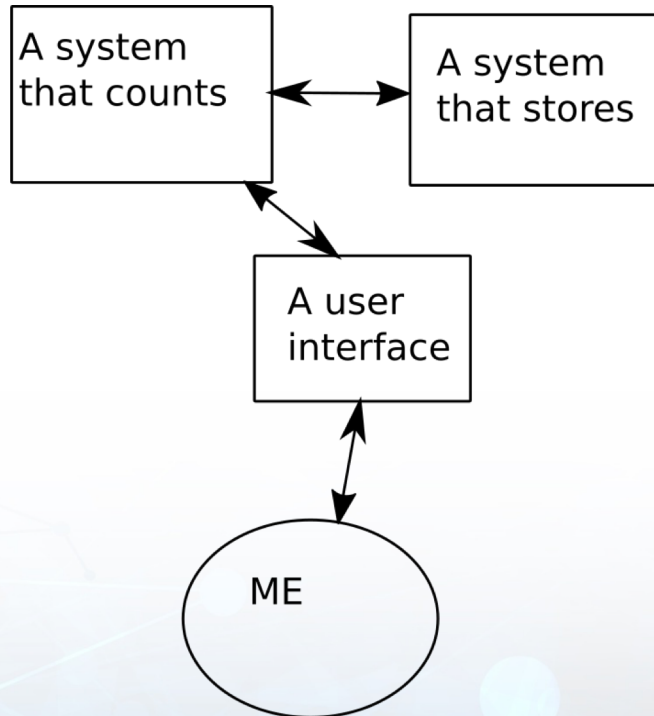


Some definitions

- Root -> /
- \$HOME -> /home/myuser
- Full path
- Working directory
- Relative path



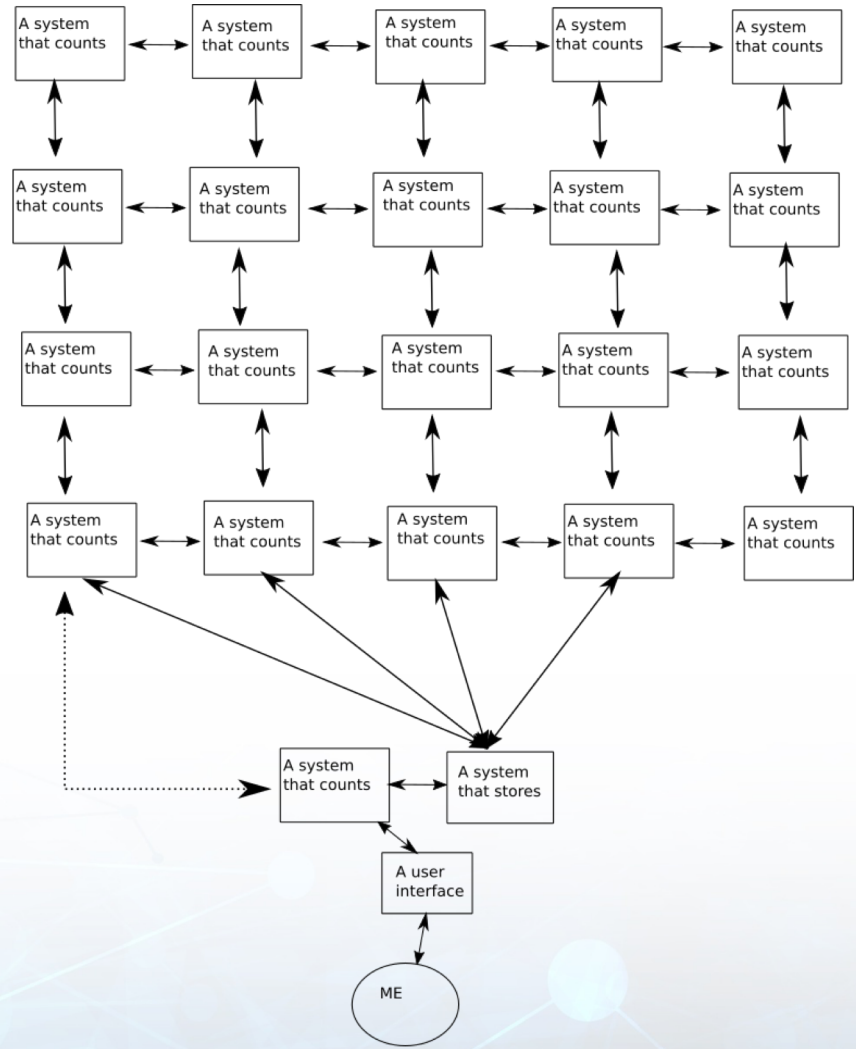
What is a supercomputer?

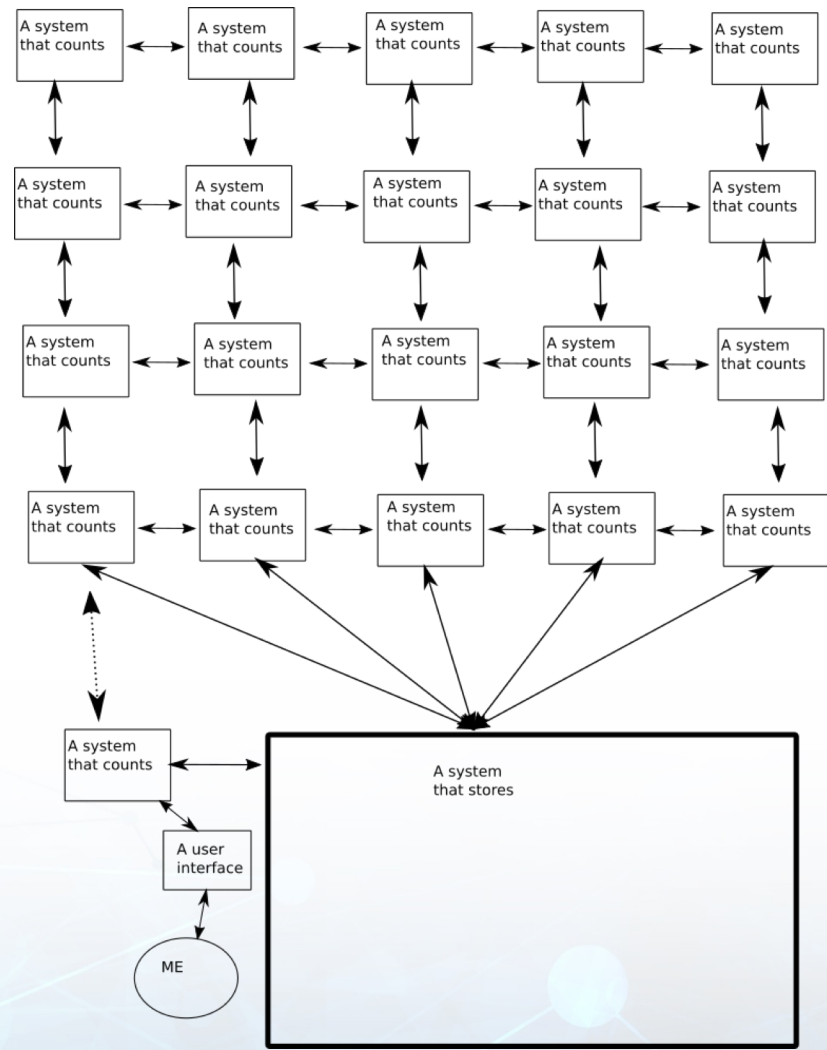


Simple computer



Supercomputer for Data science





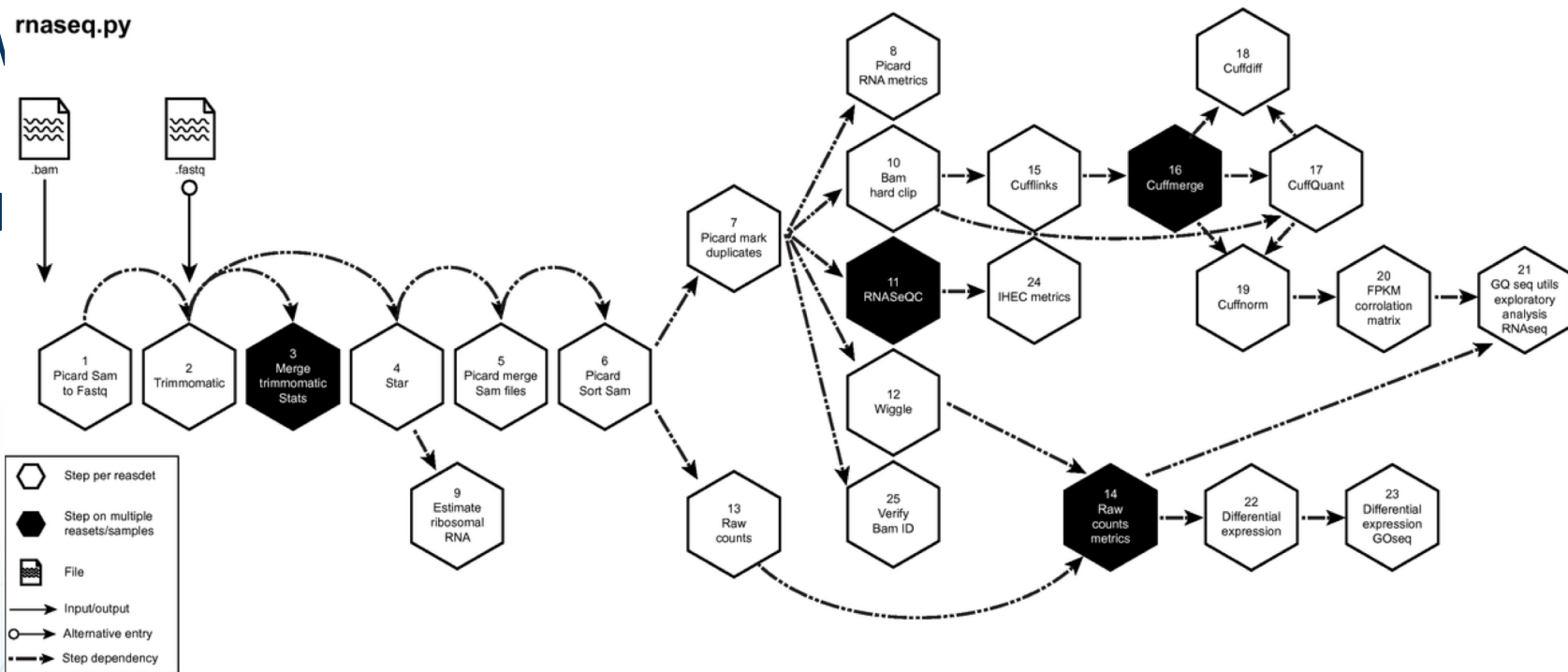
Supercomputer for
Data science



What can we do with an HPC



rnaseq.py



How to make the HPC work?

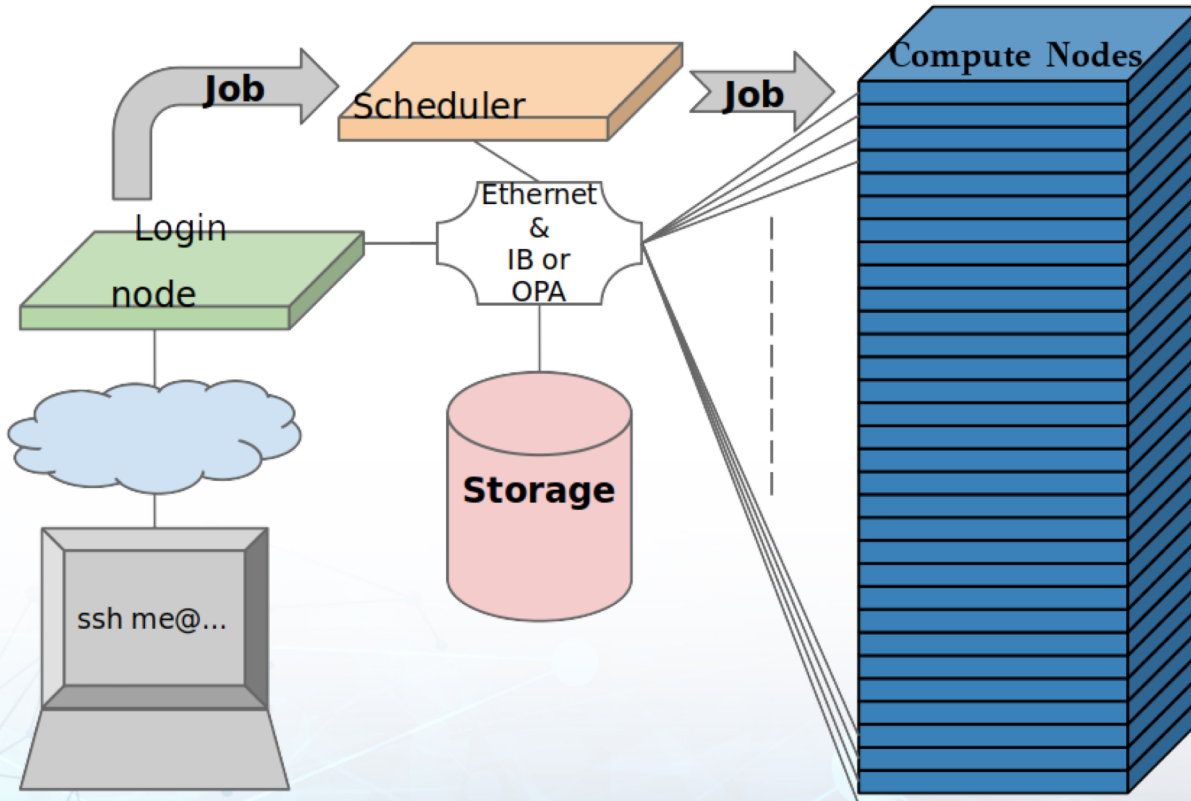


Canadian Centre for
Computational
Genomics

For many tasks, you **don't** have to **think that a supercomputer** is super... In fact, **GenPipes** is taking care of that part for you!

- **You run Genpipes on a single (simple) computer and it creates a script that will control the supercomputer for you**
- You just ask the supercomputer to execute it and then look at your data being generated.

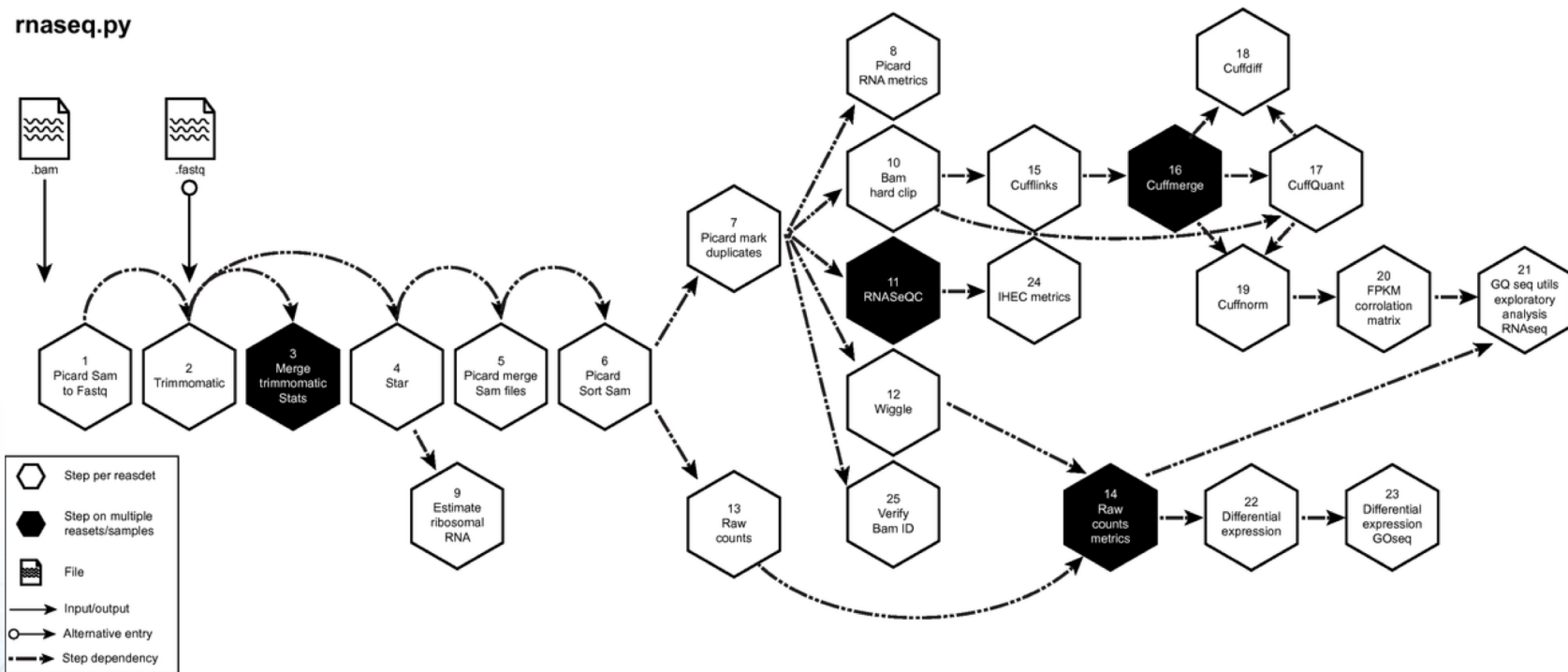
More HPC details



What is this workshop about?



rnaseq.py





Genome Québec



Genome Canada



Ontario Genomics



McGill



Canadian Centre for
Computational
Genomics

SickKids®

MiC:M McGill initiative in
Computational Medicine

Thank you!

 HPC4Health

 **IHEC**
International Human Epigenome Consortium

compute | calcul
canada | canada


bioinformatics.ca

 Terry Fox
PROFYLE


GenAP




Calcul Québec


canarie