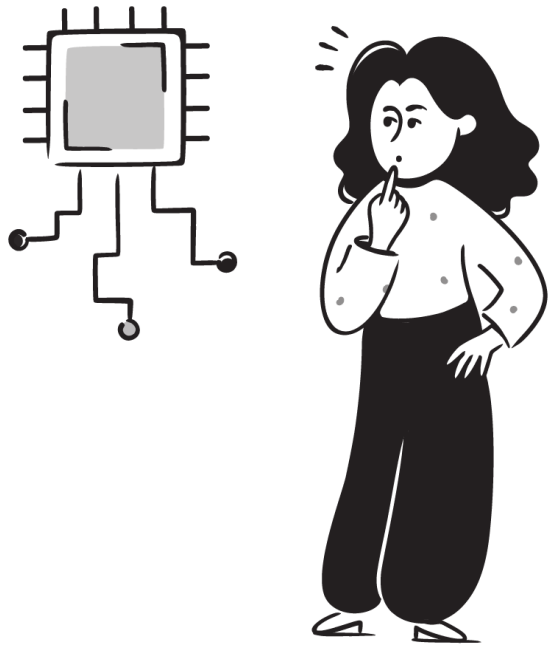
LIÈGE université
GIGA

Working with High Performance Computing Clusters & Parallel Processing

Computational Biosciences Days
13th February 2025
Martin Grignard, PhD, Eng.

GIGA bioinformatics team
bioinfo.giga@uliege.be

What's (in) a computer?



Components of a computer

CPU

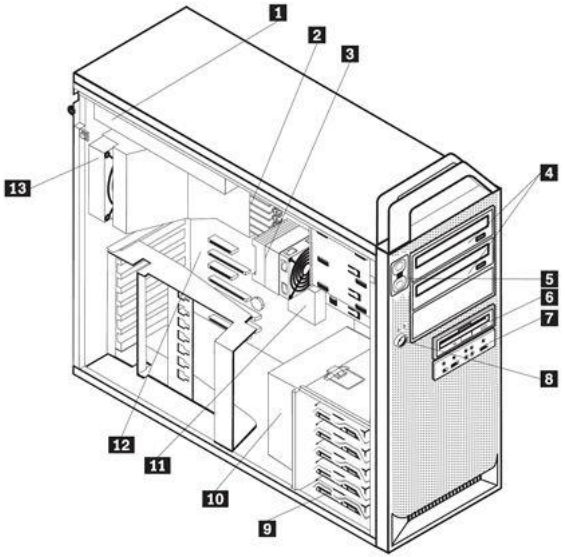
Primary component acting as the control center (*i.e.* the active brain of the computer).

RAM

Short term volatile memory where data is stored during computing.

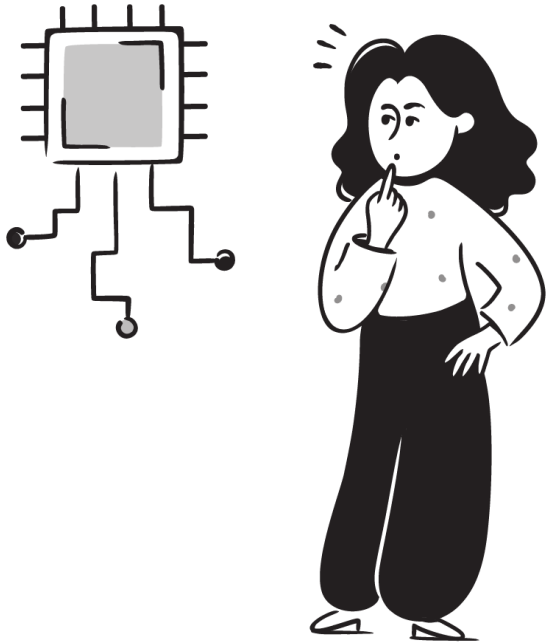
GPU

Specialized component optimized to accelerate graphics and parallel tasks.



And lots more! (Motherboard, storage, power supply...)

What's (in) a high performance computer?



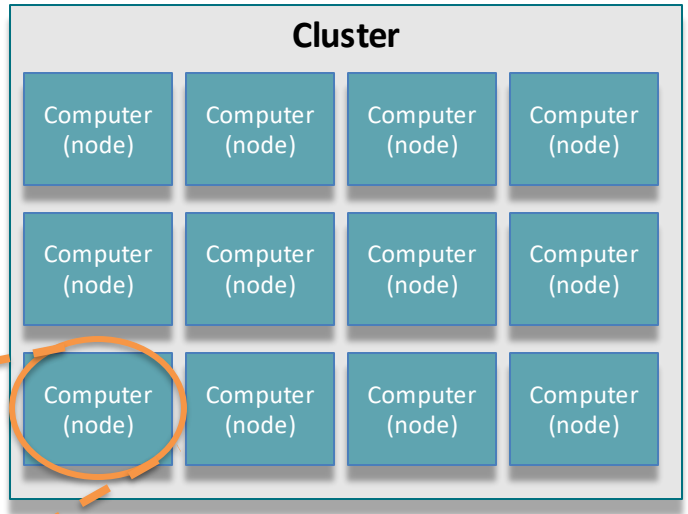
High Performance Computing (HPC)

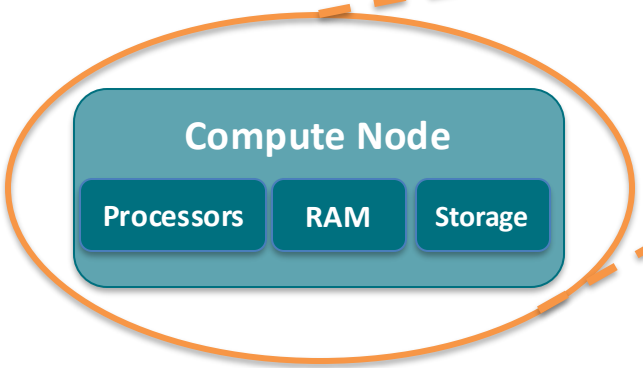
Cluster

Group of linked computers, working together closely so that in many respects they form a single computer

Node

Part of a cluster (equivalent to a high-end workstation)





High Performance Computing (HPC)

Definition

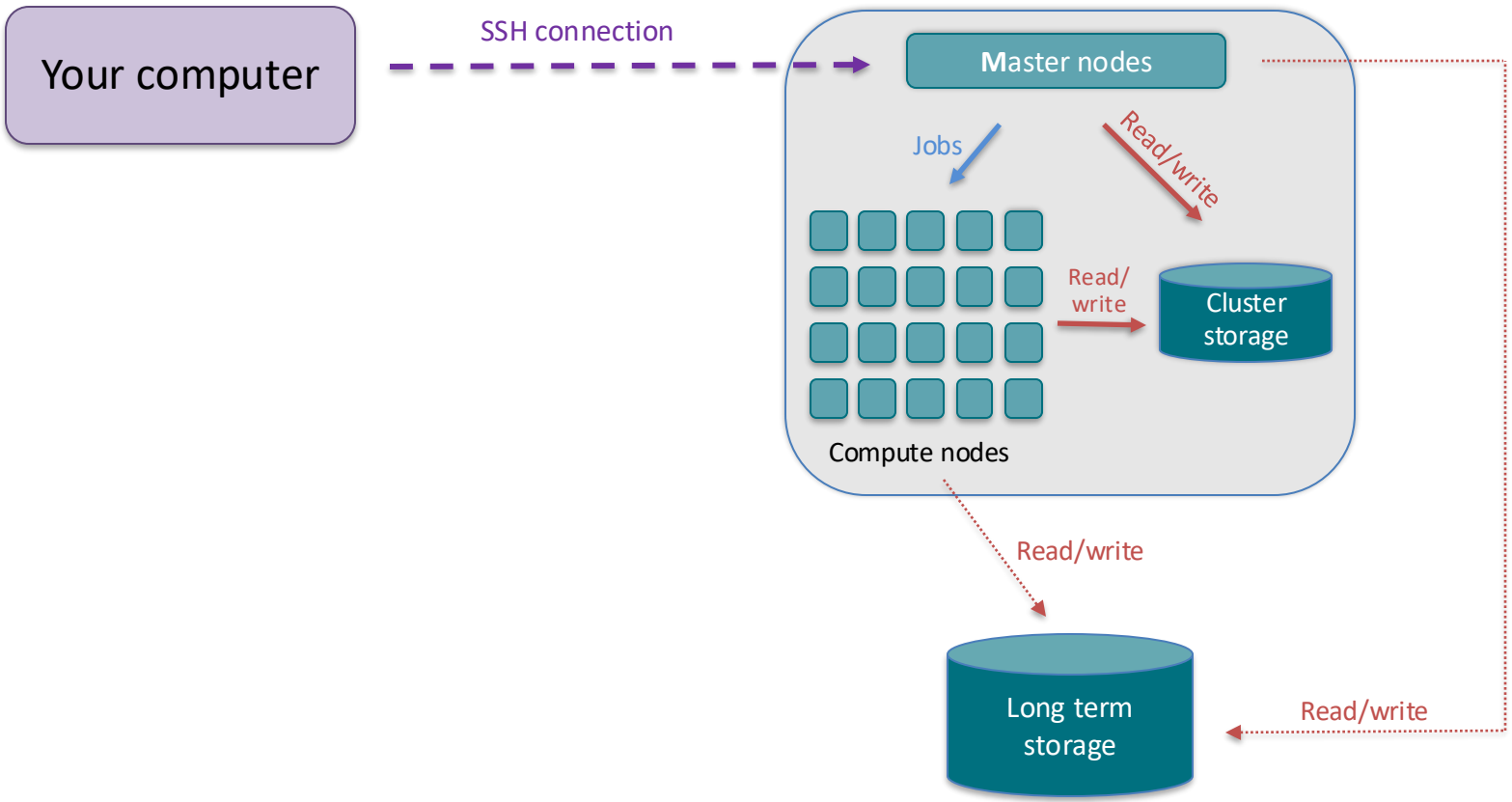
Computing system with extremely high computational power that is able to solve hugely complex problems.

- Analysis of huge volume of data (WGS, high resolution images, etc)
- Compute-intensive processes (simulations, determination of relationship between observations, etc)

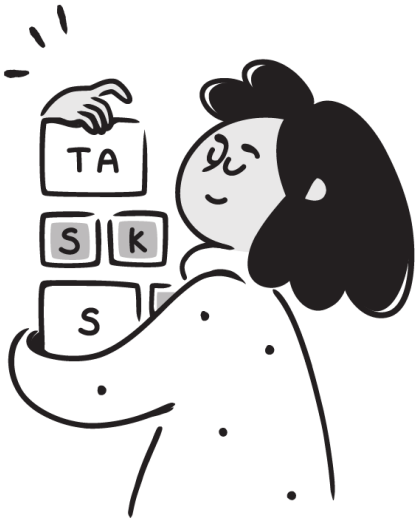
How to achieve high computational power ?

- Provide powerful machine
- Group several machines together
- Share them and optimize usage

High Performance Microwaving


Your computer
SSH connection
Master nodes
Jobs
Read/write
Read/
write
Cluster
storage
Compute nodes
Read/write
Long term
storage
Read/write

Why would you need a cluster?



When do I need to use a cluster ?

- I need to run the same analysis again and again (on hundreds of samples or testing hundreds values of a given parameter)
- My data don’t fit my disk or my computer’s memory
- The program I use require resources my computer doesn’t have

Use cases : population study

I have to apply the same process to many samples.

Illustration with numbers

- Data from 360 subjects
- Require 1 day/subject/core

Workstation (4 cores) : 90 days

Cluster (720 cores) : 1/2 day

Use cases : whole genome sequencing

My data won’t fit in my computer memory.

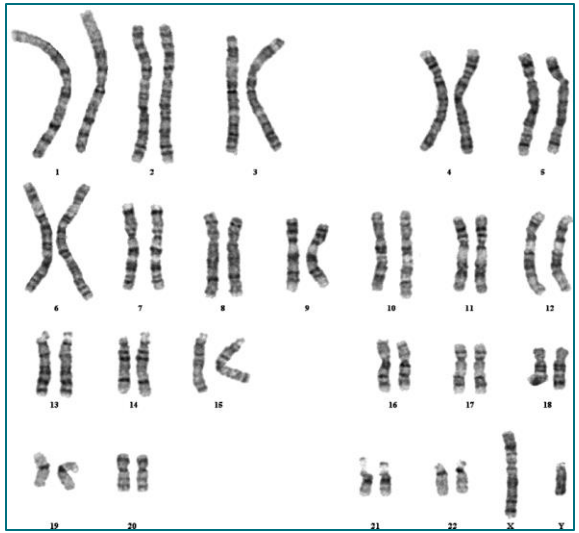
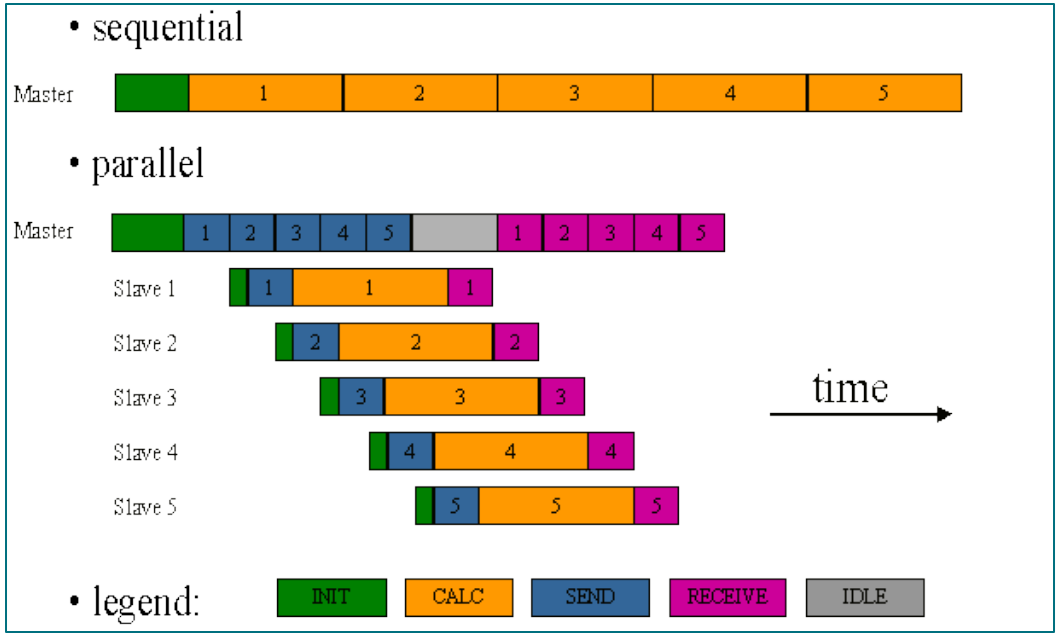


Illustration with numbers

- Human genome = 6 billion bases
- A single person's whole genome > 300GB and processing it will require > 300GB RAM
- Of note: in some cases, analysis could be split by chromosome and parallelized

How much can I parallelize in practice ?



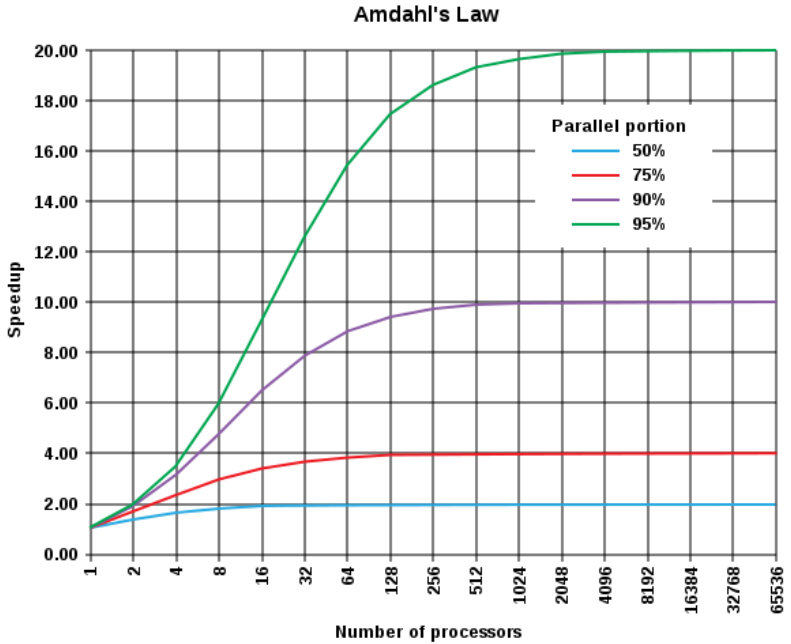
- Worth if time(subtask) >> time(overheads)
- Time saved α fraction parallelizable (Amdahl's law)

Amdahl's law

$$\text{Speedup} = \frac{1}{(1 - F) + \frac{F}{N}}$$

F = parallelisable fraction
N = number of nodes
Assume no overhead for

- Scheduling
- Networking
- Synchronisation

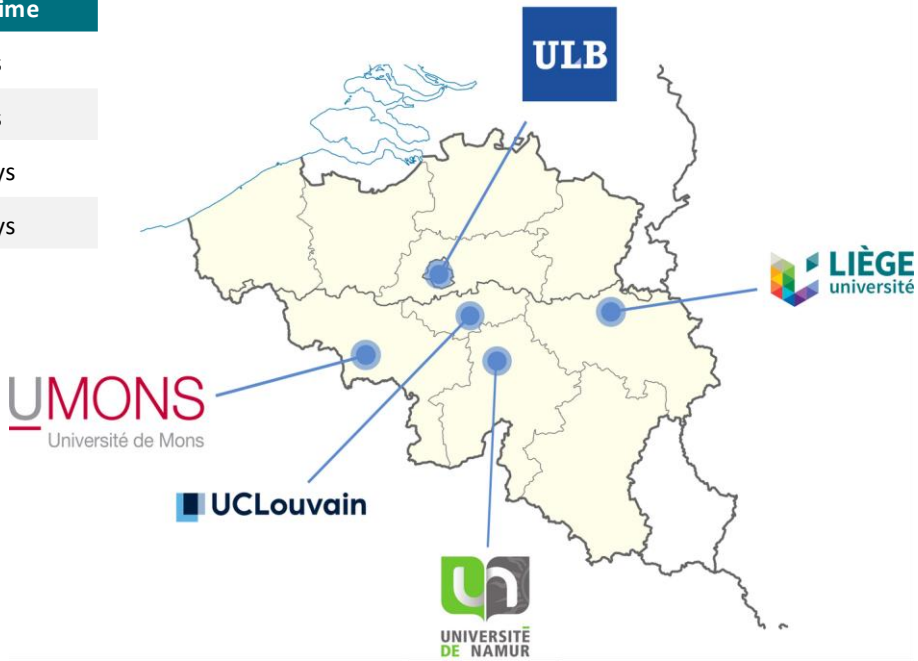


Which clusters do I have access to?

CECI

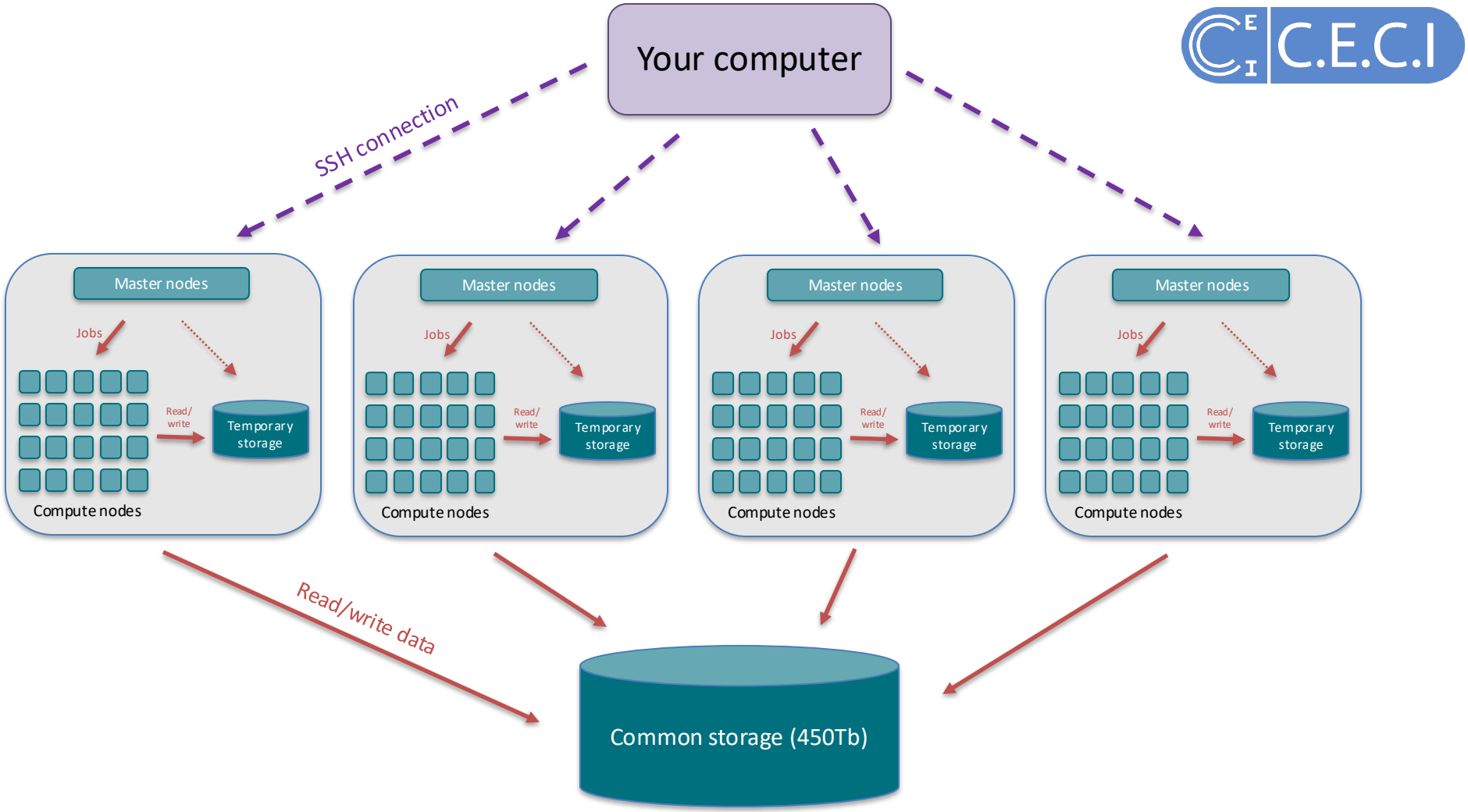
(Consortium des Équipements de Calcul Intensif)

Cluster	Host	CPUs	RAM	GPU	Max time
NIC5	ULiège	4672	256 GB … 1 TB	No	2 days
Lemaitre4	UCLouvain	5120	766 GB	No	2 days
Hercules2	UNamur	1024	256 GB … 2 TB	Yes	15 days
Dragon2	UMons	592	192 … 384 GB	Yes	21 days



Website

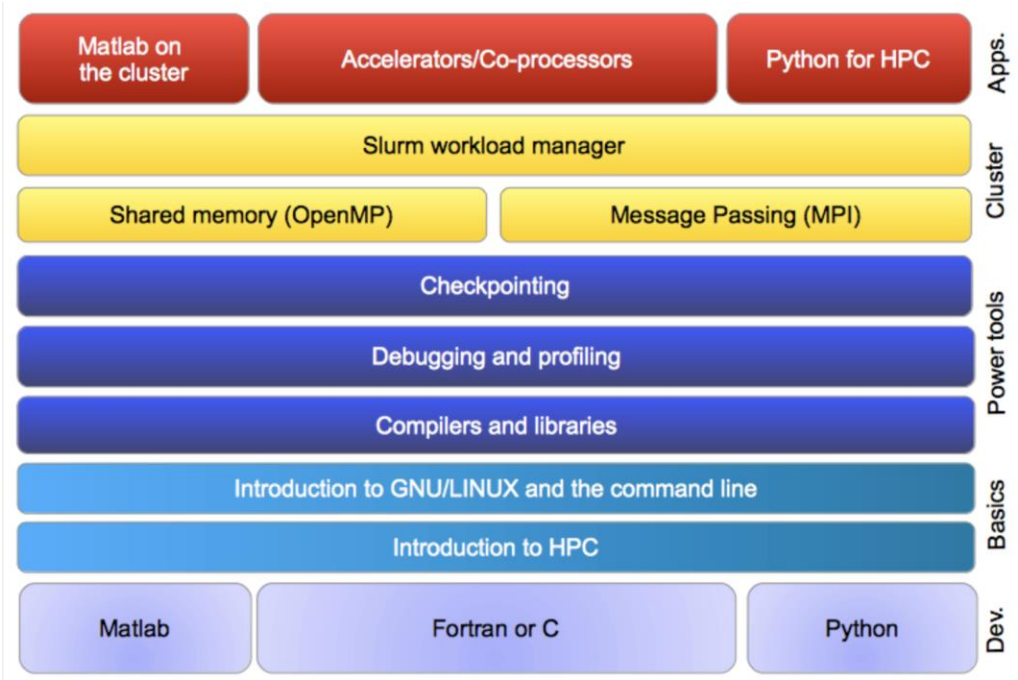
http://www.ceci-hpc.be/


Your computer
SSH connection
Master nodes
Jobs
Read/
write
Temporary
storage
Compute nodes
Master nodes
Jobs
Read/
write
Temporary
storage
Compute nodes
Master nodes
Jobs
Read/
write
Temporary
storage
Compute nodes
Master nodes
Jobs
Read/
write
Temporary
storage
Compute nodes
Read/write data
Common storage (450Tb)

CECI website and training

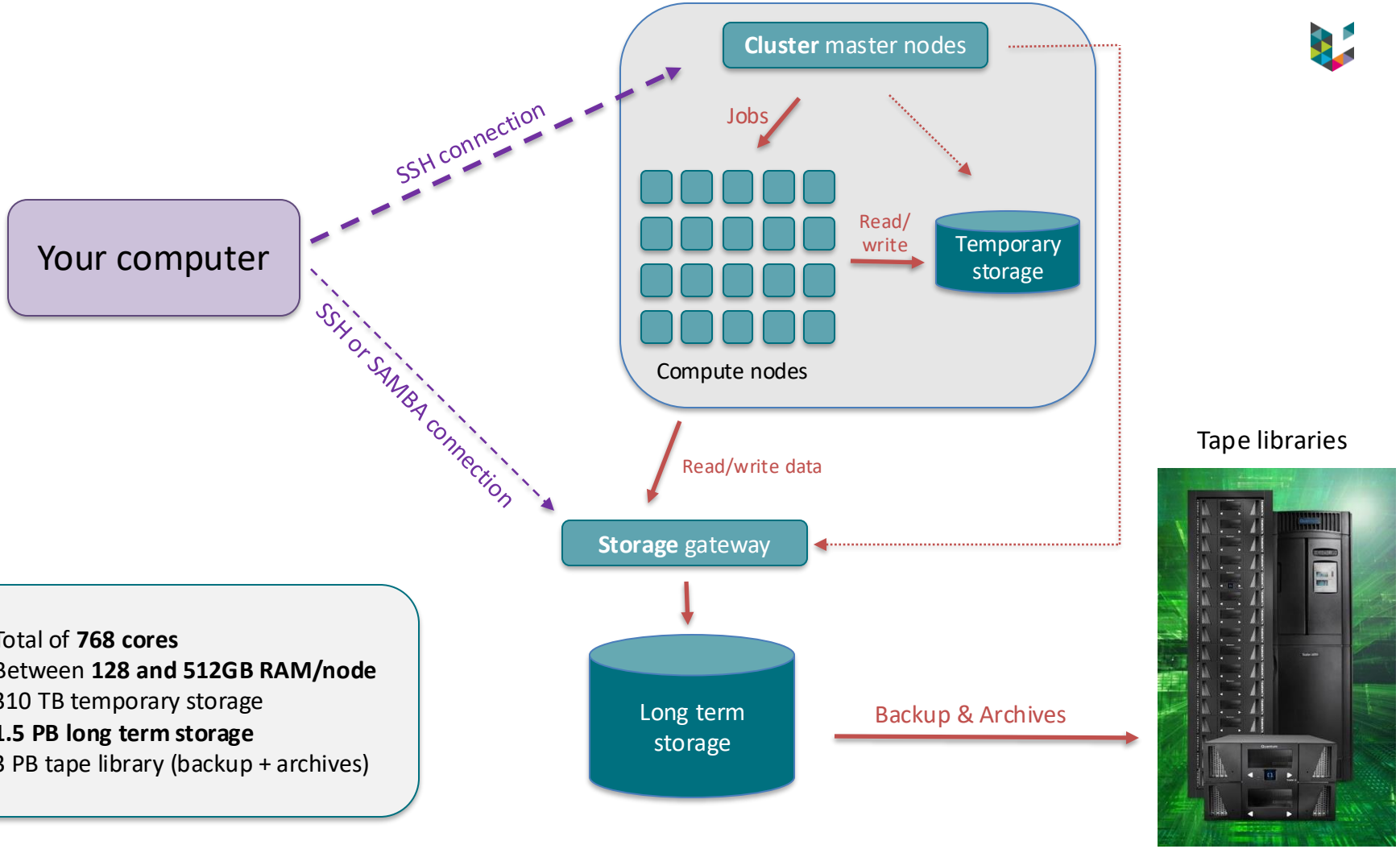
http://www.ceci-hpc.be/

- Documentation and FAQ
- Trainings
  - courses given every year (Oct-Nov)
  - videos from previous year on youtube channel
- Support



GIGA

- Accessible to all GIGA members and CHU bioinformatic team
- Directly linked to the GIGA mass storage (1.5 petabyte)
- Documentation



- Total of **768 cores**
- Between **128 and 512GB RAM/node**
- 310 TB temporary storage
- **1.5 PB long term storage**
- 3 PB tape library (backup + archives)

Which clusters should I use?

Which cluster to use ?

	NIC5 (CECI)	GIGA
Nodes	73 - brand new	24
Total nb of cores	4,672	768
RAM	256 GB - 1TB	128-512 GB
Server disks	All SSD	slower disks
Number of users	high => waiting time can be long	lot less (~60) but waiting time still sometimes long
Linked to mass storage ?	No	Yes

Which cluster to use ?

Example with sequencing run demultiplexing (several hundreds samples)

	NIC5 (CECI)	GIGA
Data transfer (1.6 TB)	2h	0h
Waiting time (job in the queue)	5min	0min
Demultiplexing	9h	16h
Output transfer (1.6TB)	2h	0h
TOTAL	**13h**	**16h**

How to access a cluster?

The interface between the user and the cluster: the command line terminal

All CECI and GIGA clusters use CentOS (Linux)

Linux™ CentOS

From MAC and Linux

Terminal

Connection to cluster from a Windows computer

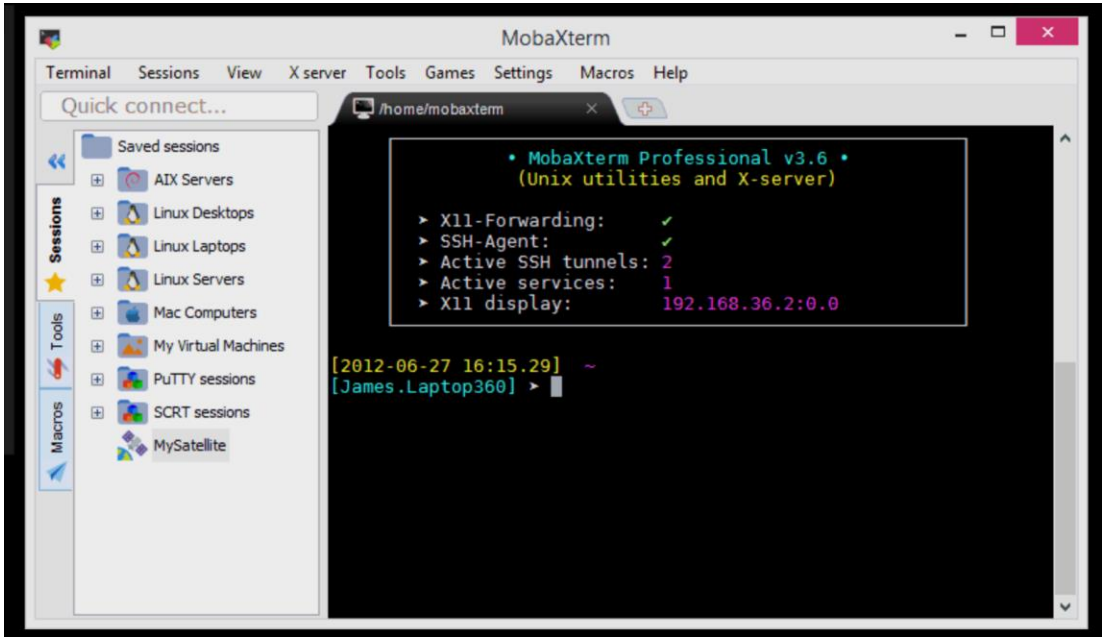
Windows SSH clients

PowerShell
- on Windows 10 and higher,
- looks for it in start menu

MobaXterm
- download from https://mobaxterm.mobatek.net/
- easy to use
- command line interface + interface for file transfer + allow use of graphical applications remotely.

MobaXterm

MobaXterm (recommended by CECI)



How do I connect to the CECI or GIGA cluster ?

CECI cluster

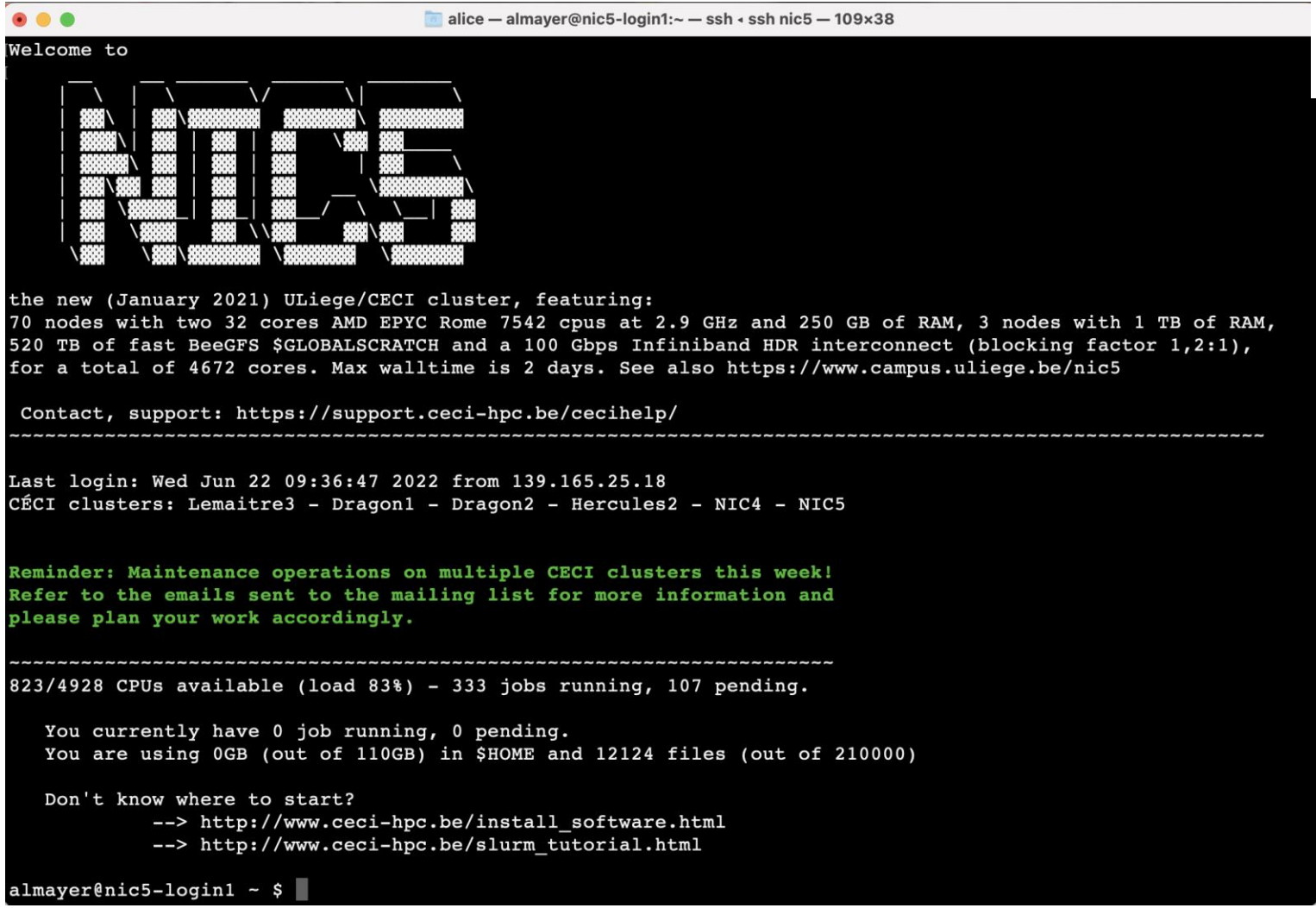
1. Get a CECI account: https://support.ceci-hpc.be/doc/_contents/QuickStart/CreatingAnAccount/index.html
2. Connection instructions: https://www.ceci-hpc.be/assets/pdf/renew-create-account.pdf

GIGA cluster

Connection instructions (GIGA members):
https://giga-bioinfo.gitlabpages.uliege.be/docs/mass-storage-and-cluster/getting_started/connect.html
(The very first time, it's mandatory to connect to mass storage using SAMBA protocol)

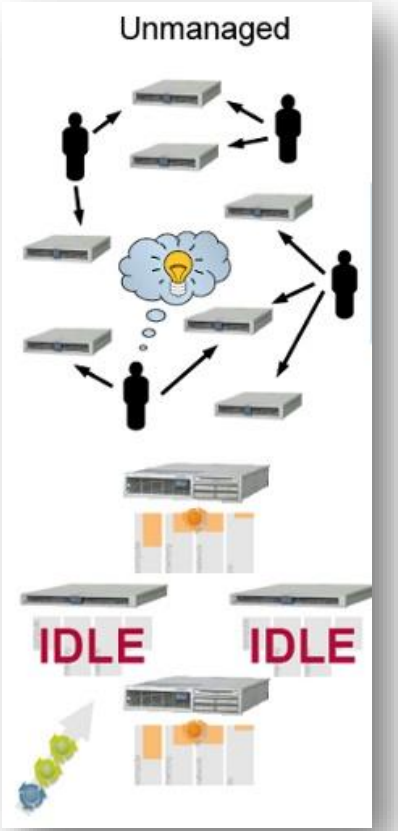
In both cases, if you are outside of university network:

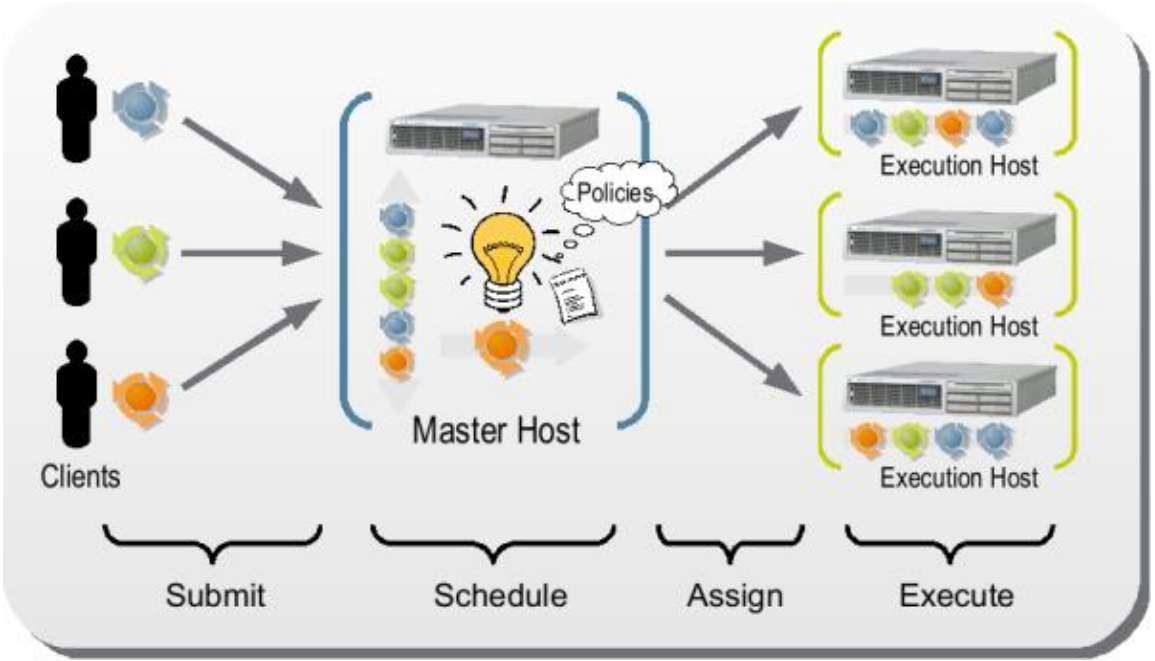
https://my.segi.uliege.be/cms/c_11650735/en/mysegi-vpn-f5-big-ip


alice — almayer@nic5-login1:~ — ssh ‹ ssh nic5 — 109×38
Welcome to
the new (January 2021) ULiege/CECI cluster, featuring:
70 nodes with two 32 cores AMD EPYC Rome 7542 cpus at 2.9 GHz and 250 GB of RAM, 3 nodes with 1 TB of RAM,
520 TB of fast BeeGFS $GLOBALSCRATCH and a 100 Gbps Infiniband HDR interconnect (blocking factor 1,2:1),
for a total of 4672 cores. Max walltime is 2 days. See also https://www.campus.uliege.be/nic5
Contact, support: https://support.ceci-hpc.be/cecihelp/
Last login: Wed Jun 22 09:36:47 2022 from 139.165.25.18
CÉCI clusters: Lemaitre3 - Dragon1 - Dragon2 - Hercules2 - NIC4 - NIC5
Reminder: Maintenance operations on multiple CECI clusters this week!
Refer to the emails sent to the mailing list for more information and
please plan your work accordingly.
823/4928 CPUs available (load 83%) - 333 jobs running, 107 pending.
You currently have 0 job running, 0 pending.
You are using 0GB (out of 110GB) in $HOME and 12124 files (out of 210000)
Don't know where to start?
--> http://www.ceci-hpc.be/install_software.html
--> http://www.ceci-hpc.be/slurm_tutorial.html
almayer@nic5-login1 ~ $

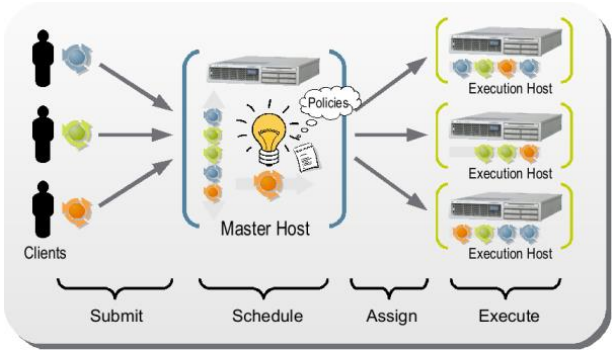
How to use a cluster?

Now that I’m connected, how do I run an analysis?





Slurm basics



Info you need to give to slurm

- Partition (or queue) you want to send your job to
- number of CPUs and RAM
- maximum time your job will take
- optional info such as job name, log file path, when to send email and to which address, etc.

Slurm priorities

- if you ask for shorter time, your job will have higher priority
- users who launch very high number of job have lower priority

Two types of slurm sessions

Interactive sessions

- Several short tasks
- Tasks that require user input
- Typically: when developing/optimizing pipeline

Batch sessions

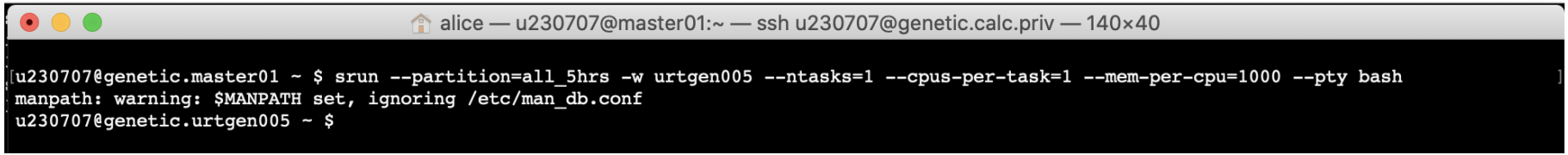
- Longer running processes
- Parallel processes

Slurm interactive sessions

How to start one ?

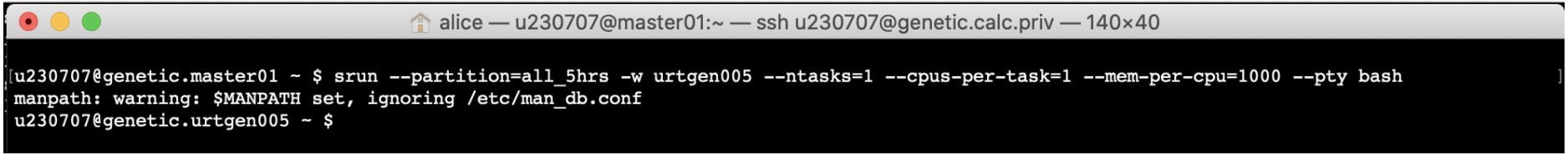
$ srun --partition=all_5hrs --cpus-per-task=1 --mem-per-cpu=1000 --pty bash

- Asking for 1 core and 1 Gb of RAM on a node of the all_5hrs partition
- Asking to have bash session on the allocated node



Notice the change of prompt, from **u230707@genetic.master01** to **u230707@genetic.urtgen005** !!!!

Slurm interactive sessions



Slurm interactive session (srun)

You are now on a node

You can perform analysis there

If you use more resources than requested, slurm will kill your session on the node

If you lose your internet connection, your session will be aborted, and your program will crash

Batch jobs

myscript.sh

Resources requested
(http://www.ceci-hpc.be/scriptgen)

```
#!/bin/bash
#
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH –mem-per-cpu=1000
#SBATCH –partition=all_5hrs
#SBATCH --time=1:00:00
#SBATCH --mail-user=my@email.com
#SBATCH --mail-type=FAIL
```

Instructions (Shell script, Python...)

```
# Do some stuff
# preceded by srun if you want to have statistics for that step
srun echo “Hello”
```

submit

```
$ sbatch myscript.sh
```

Batch jobs

Monitor jobs

while job is still running, give info on resources, nodes, etc
$ scontrol show job <JOB_ID>

After job finished, info on resources used
$ sacct --format="JobId,JobName,NodeList,State,Elapsed,CPUTime,MaxRSS,AveRSS,ReqMem, ReqCPUS, Submit,Start" -j <JOB_ID>

Cancel jobs

$ scancel <jobID>

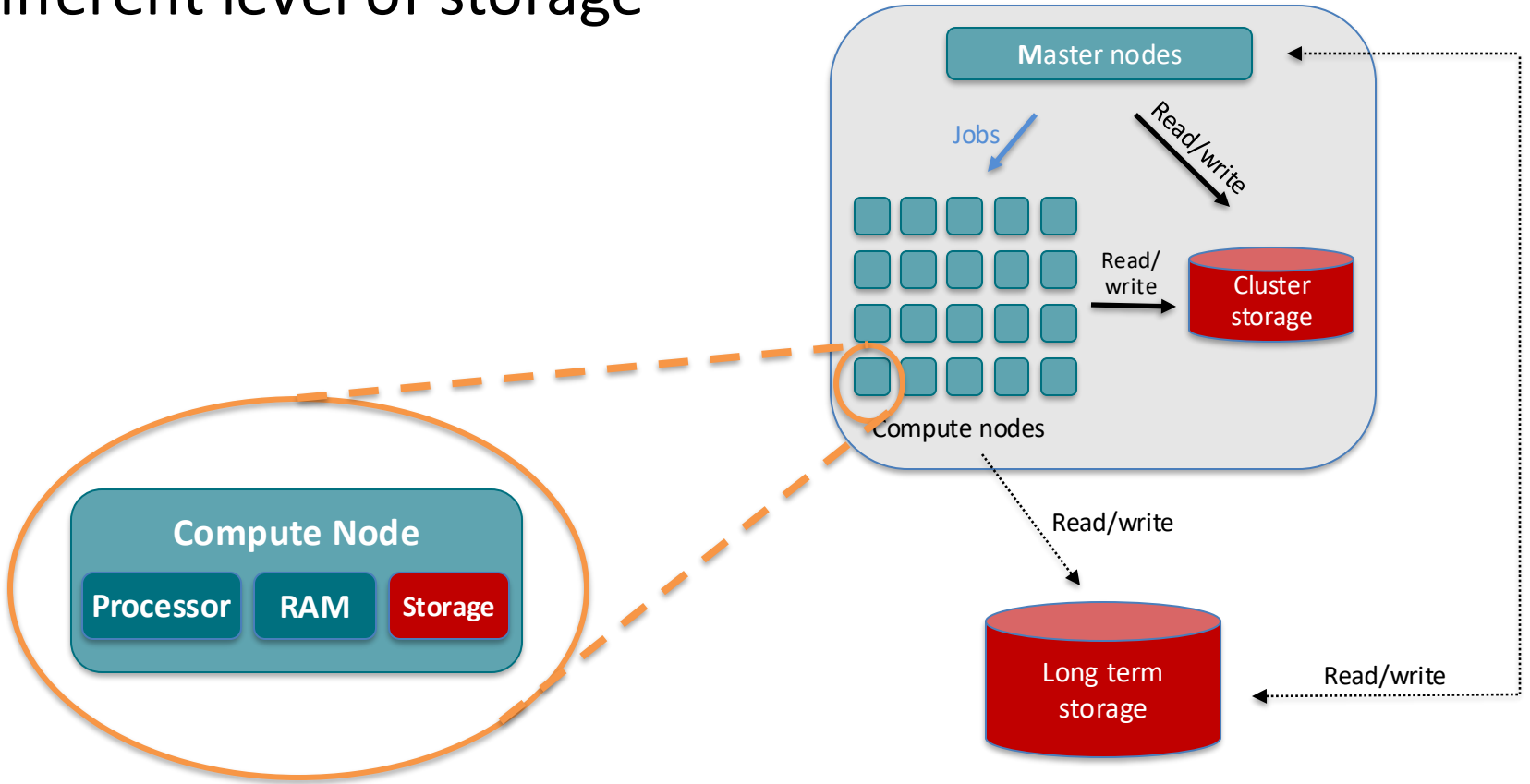
Using slurm array to parallelize

This script will launch 4 jobs (by 2). Each one will write its number in the log and wait 2 minutes.

```
#!/bin/bash
#
#SBATCH --job-name=Test
#SBATCH --cpus-per-task=1
#SBATCH --mem-per-cpu=1000    # in Mb (could also write 1G)
#SBATCH --time=1:00:00        # "hours:minutes:seconds"
#SBATCH --partition=all_5hrs
#SBATCH --output=test_%j.log  # path + name of log file, %j = job ID
#SBATCH --mail-user=alice.mayer@uliege.be
#SBATCH --mail-type=FAIL
#SBATCH --array=1-4%2

date
# This will be printed in the log file of each job
echo ""
echo "Hello, I'm the job number ${SLURM_ARRAY_TASK_ID}"
srun sleep 120     # do nothing during 2 minutes
date
```

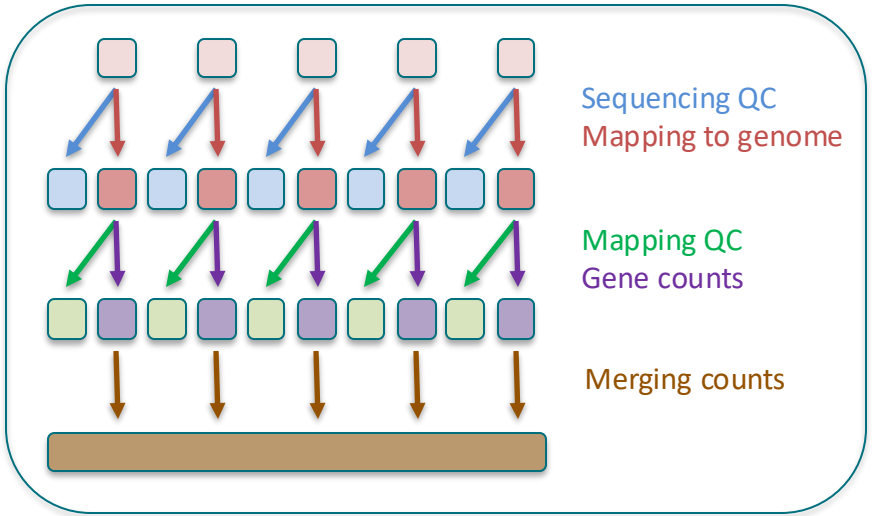
Different level of storage



Workflow manager (ex: nextflow)

Tools developed to process several samples through several analysis steps,
while optimizing resources usage

nf-core/rnaseq



https://nf-co.re/rnaseq/output
https://www.nextflow.io/docs/latest/tracing.html

Let's conclude !

Take-home messages

- Cluster = group of powerful compute nodes linked together
- Clusters are very useful when an analysis is not possible or too slow on our desktop computer
- When using a cluster,
  - don’t calculate on master node but use slurm to send jobs to compute nodes instead
  - write temporary and intermediate files on node or cluster temporary storage and not directly on mass storage
- When your analysis is finished
  - Transfer final output to long term storage
  - Delete temporary and intermediate files from node and cluster storage

Useful links

CECI clusters

- CECI website: http://www.ceci-hpc.be/
- documentation: https://support.ceci-hpc.be/doc/ (including slurm tutorial and FAQ)
- Training: http://www.ceci-hpc.be/training.html (courses in Oct-Nov every year)
- Youtube channel: https://www.youtube.com/channel/UCaFutwpSbKmWOxDwEqUpSqA/featured

GIGA clusters

- Bioinformatics team's website: https://www.gigabioinformatics.uliege.be/cms/c_8464757/en/gigabioinformatics
- wiki: https://giga-bioinfo.gitlabpages.uliege.be/docs/mass-storage-and-cluster/index.html
- slurm page: https://giga-bioinfo.gitlabpages.uliege.be/docs/mass-storage-and-cluster/tools/job/slurm.html
- slurm manual: https://slurm.schedmd.com/quickstart.html
- Contact: https://sam.med.uliege.be/ (choose UDI-MED or BIOINFO-GIGA as category)


LIÈGE université
GIGA