LIÈGE université
GIGA

Clusters and parallel computing

GIGA doctoral school 2020

Alice Mayer, PhD
GIGA bioinformatic team
bioinfo.giga@uliege.be

Clusters and parallel computing

LIÈGE université
GIGA

- Basic notions
- When to use a cluster ?
- Which clusters are accessible to ULG/GIGA members ?
- How to use them ?
- Where to find more information ?

Basic notions

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

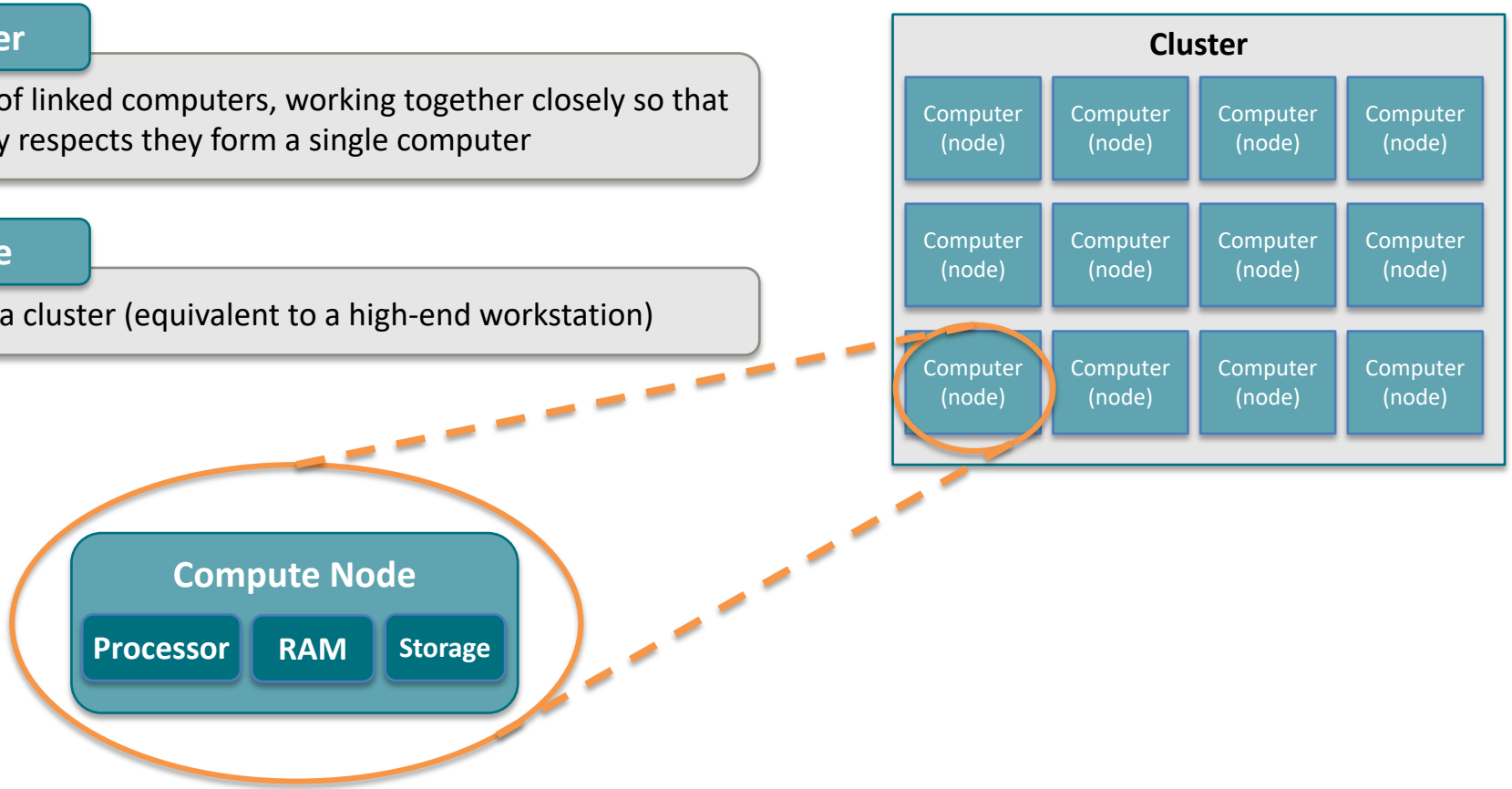
High Performance Computing (HPC)

Cluster

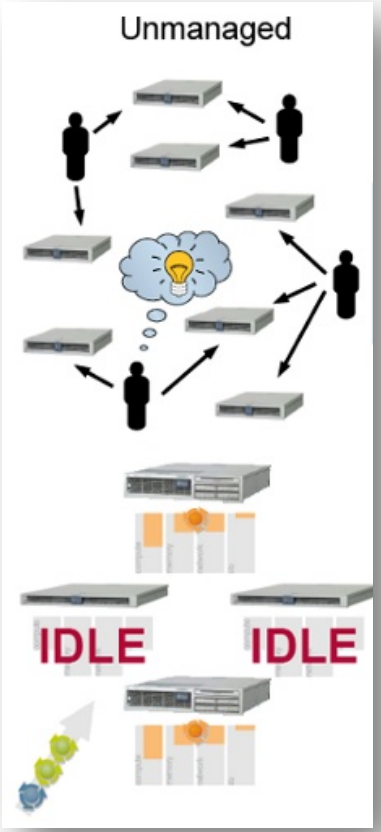
Group of linked computers, working together closely so that in many respects they form a single computer

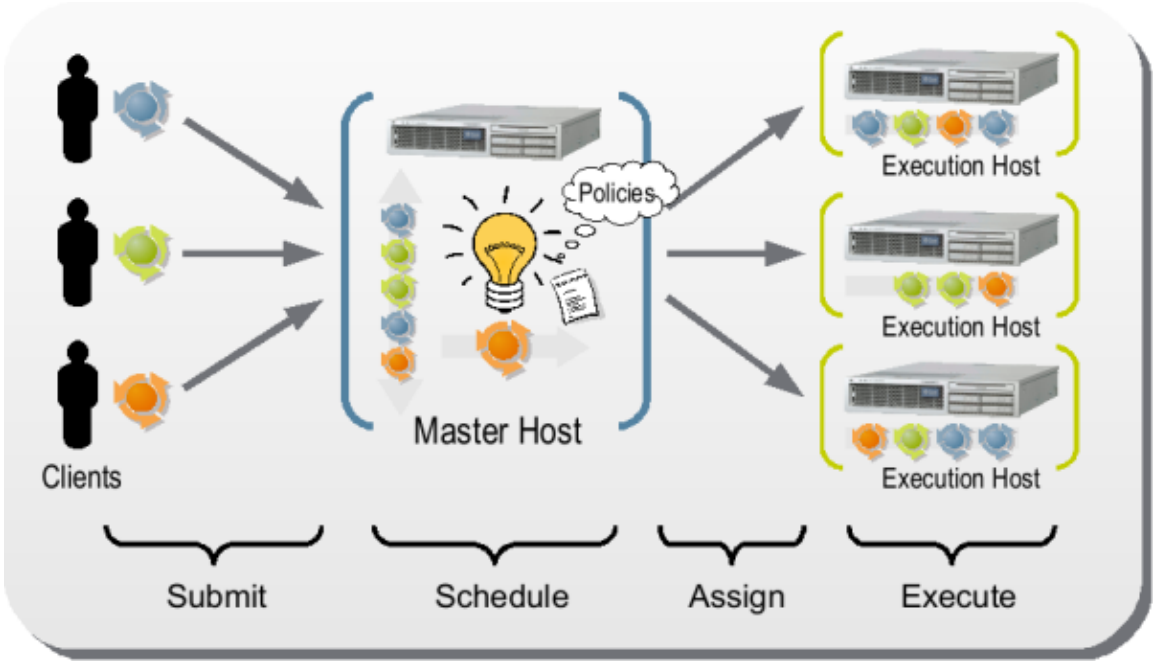
Node

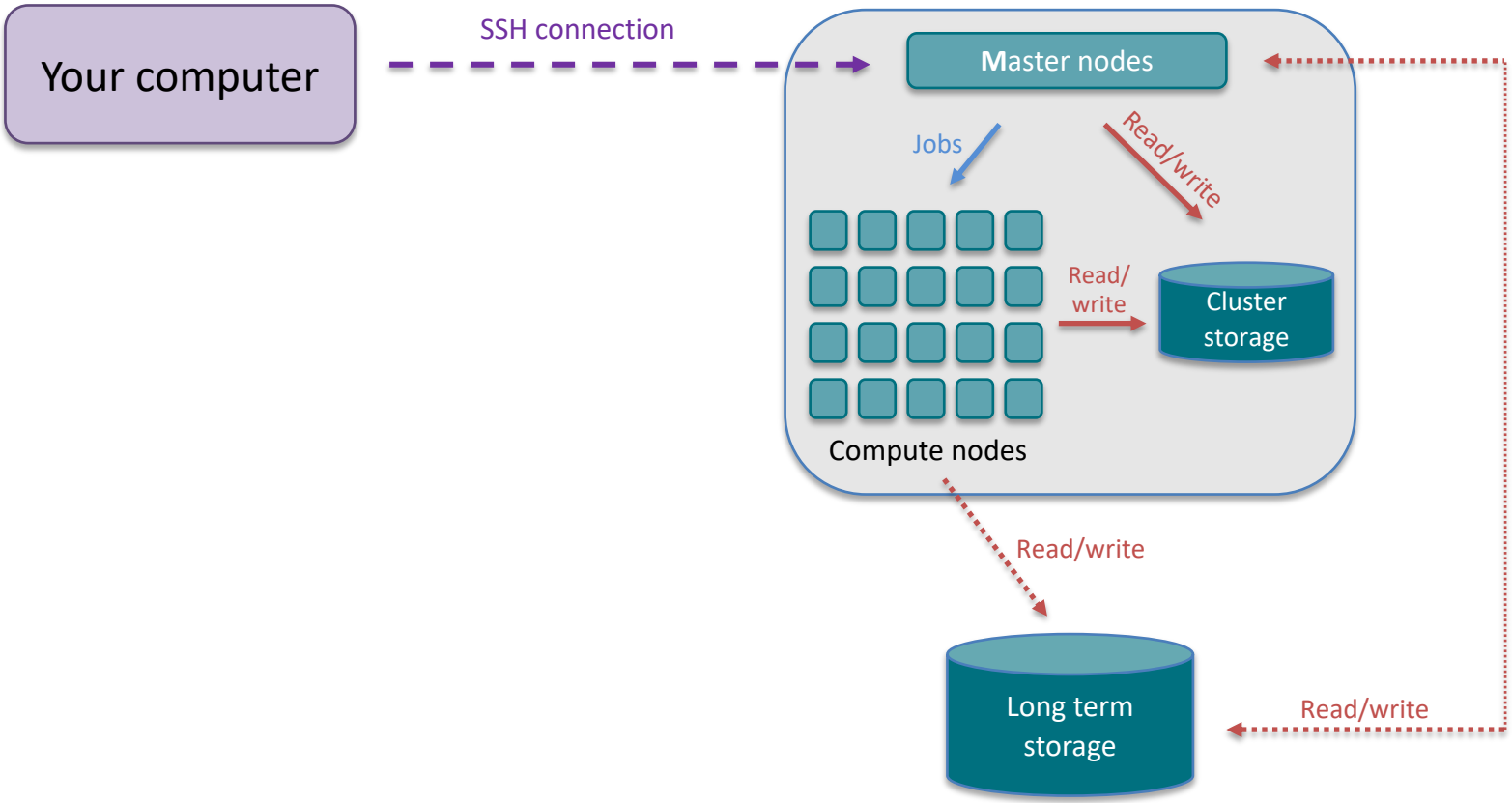
Part of a cluster (equivalent to a high-end workstation)



Cluster organisation






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

Why and when do I need to use 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/CPU

Workstation (12 CPUs) : 30 days

Cluster (360 CPUs) : 1 day

Use cases : whole genome sequencing

My data won't fit in my computer memory.

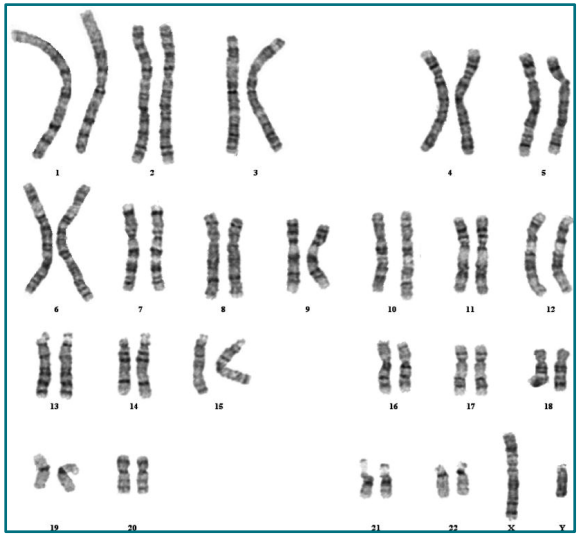


Illustration with numbers

- Human genome = 6 billion bases

  (NB: 6 x 10^9 Seconds = 190 years)
- 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

Which clusters do I have access to ?

- CECI cluster
- GIGA cluster

CECI

(Consortium des Equipements de Calcul Intensif)

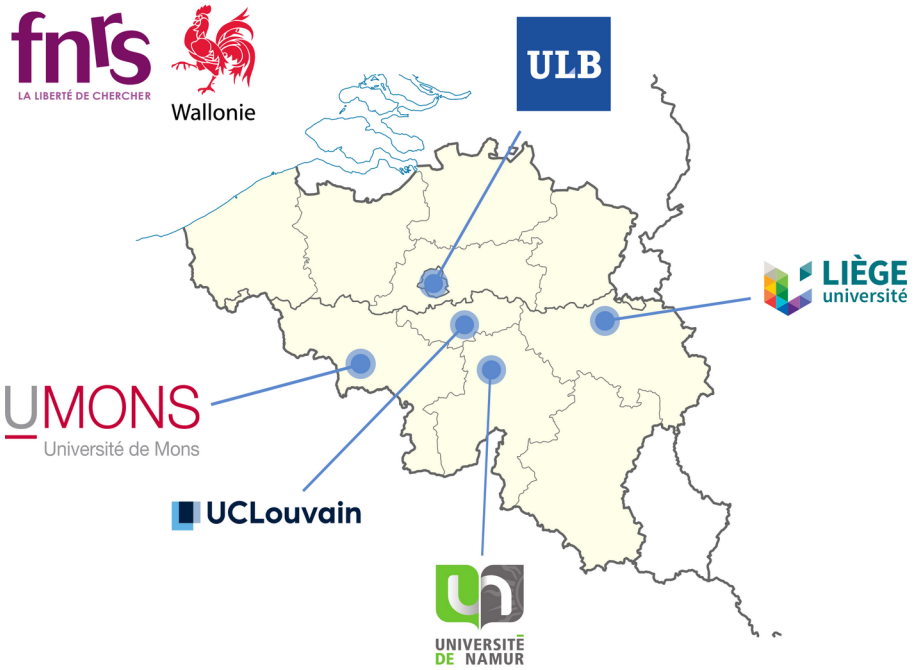
5 universities

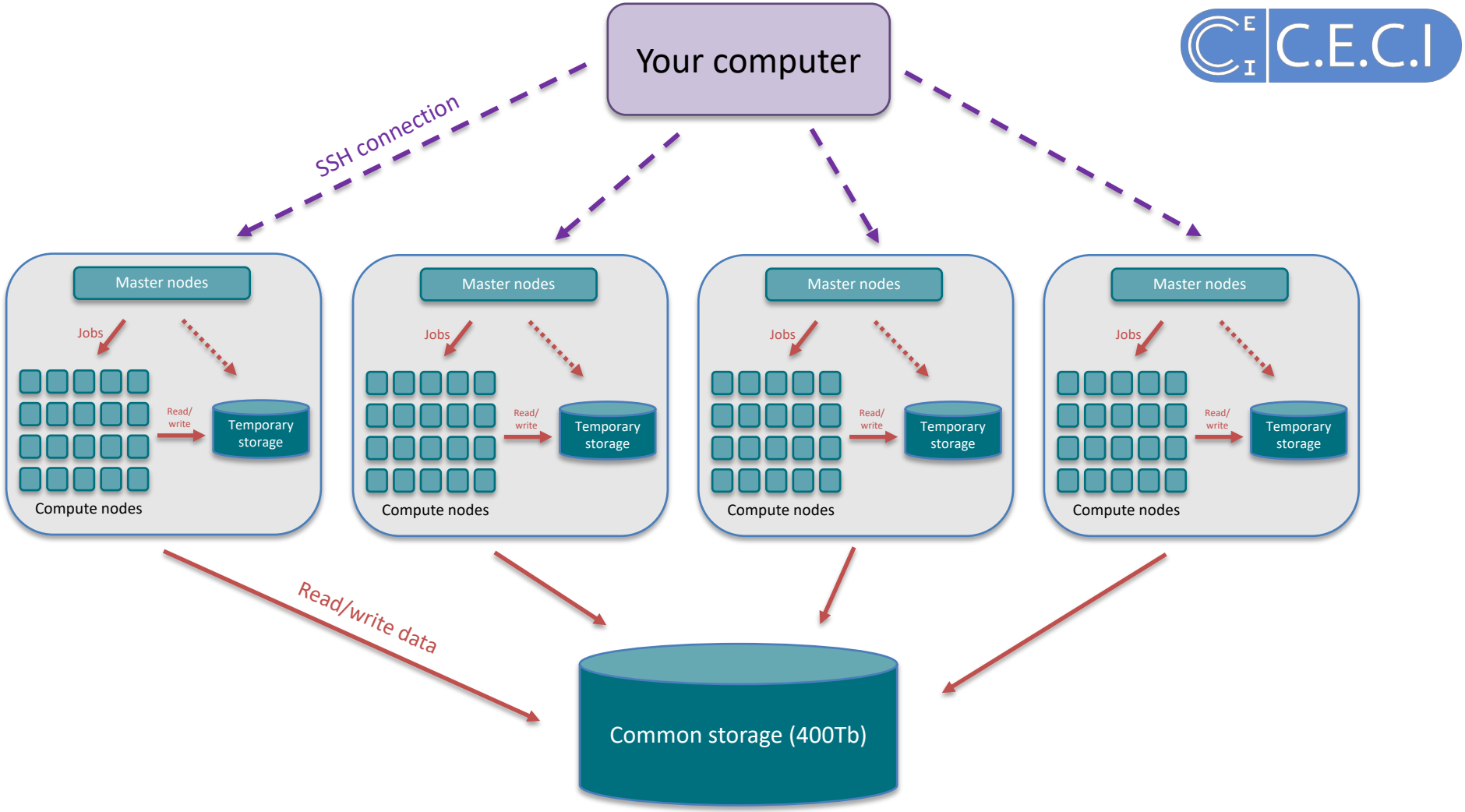
Uliège, UCLouvain, ULB, Umons, UNamur

Website

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

Support, training, documentation




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 (400Tb)

C.E.C.I

UCLouvain Université · LIÈGE université · ULB · UNIVERSITÉ DE NAMUR · UMONS

Lemaitre 3
2008 cores
Skylake
Haswell

95 GB RAM

Omnipath

Q2 2018

NIC4
2064 cores
Ivybridge
Sandybridge

64 GB RAM

QDR IB

Vega
2112 cores
Bulldozer

256 GB RAM

QDR IB

Hercules 2
1536 cores
Sandybridge
Epyc

2 TB RAM

10 GbE

Q3 2019

Dragon 2
592 cores
Skylake
Tesla V100

384 GB RAM

10 GbE

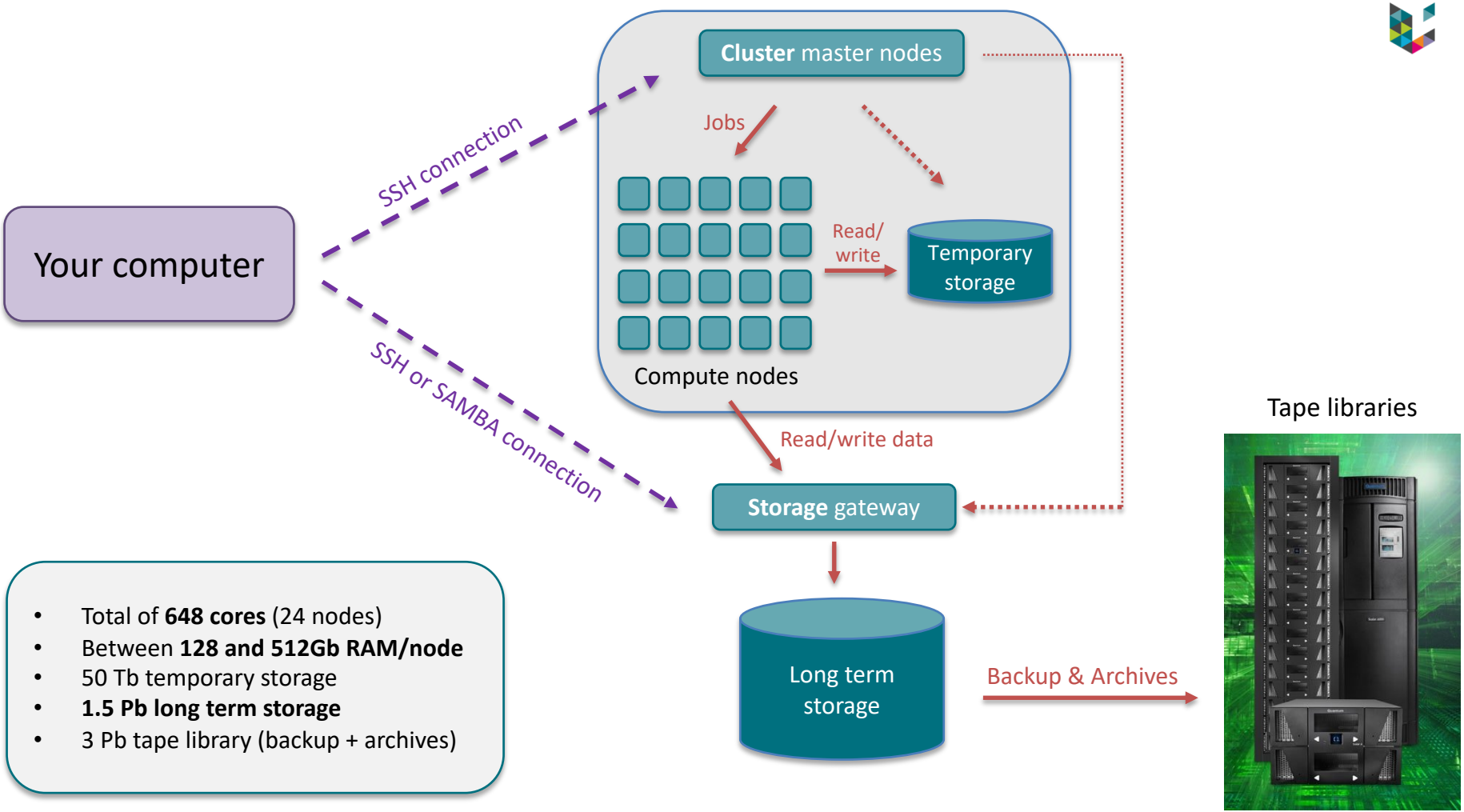
Q1 2019

8312 cores total

LIÈGE université
GIGA cluster

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

  https://gitlab.uliege.be/giga-bioinfo/user-guides-wiki/wikis/home



- Total of **648 cores** (24 nodes)
- Between **128 and 512Gb RAM/node**
- 50 Tb temporary storage
- **1.5 Pb long term storage**
- 3 Pb tape library (backup + archives)

How to use the GIGA or CECI clusters ?

How do I connect to the CECI or GIGA cluster ?

CECI cluster

1. Get a CECI account: https://login.ceci-hpc.be/init/
2. Connection instructions: https://support.ceci-hpc.be/doc/_contents/QuickStart

GIGA cluster

Connection instructions (GIGA members):
https://gitlab.uliege.be/giga-bioinfo/user-guides-wiki/wikis/mass-storage/mass-storage-connection
(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://gitlab.uliege.be/giga-bioinfo/user-guides-wiki/wikis/vpn-connection

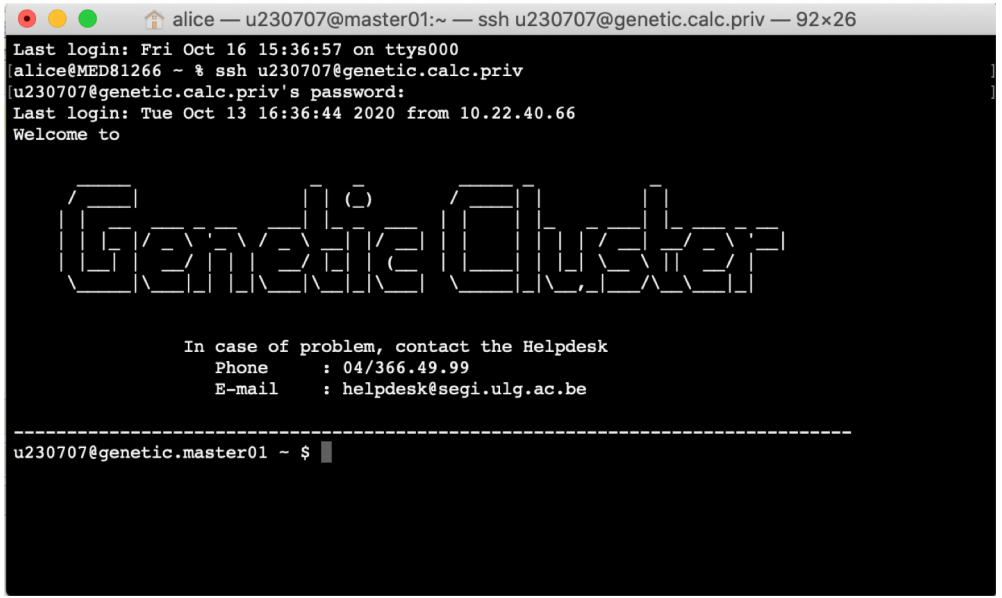
The interface between the user and the cluster: the command line terminal

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

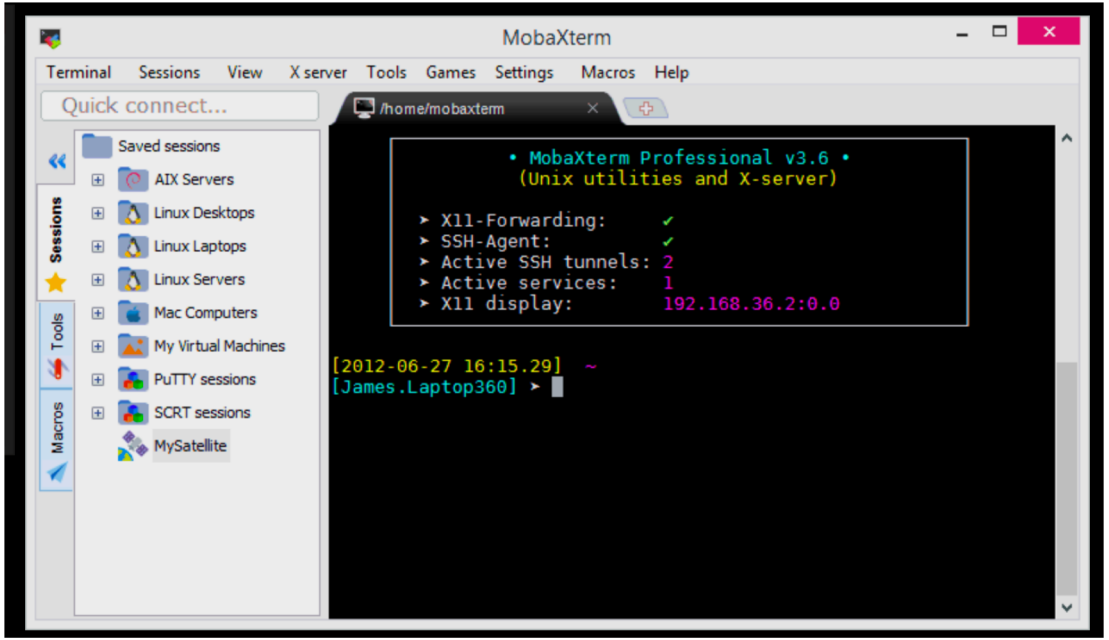
MobaXterm

PuTTY

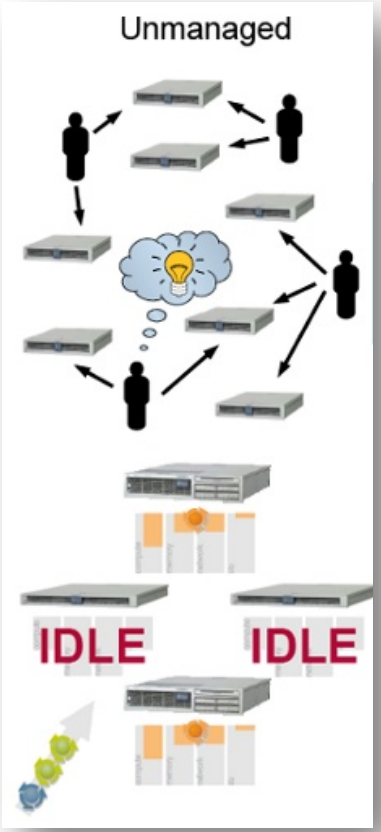
KiTTY

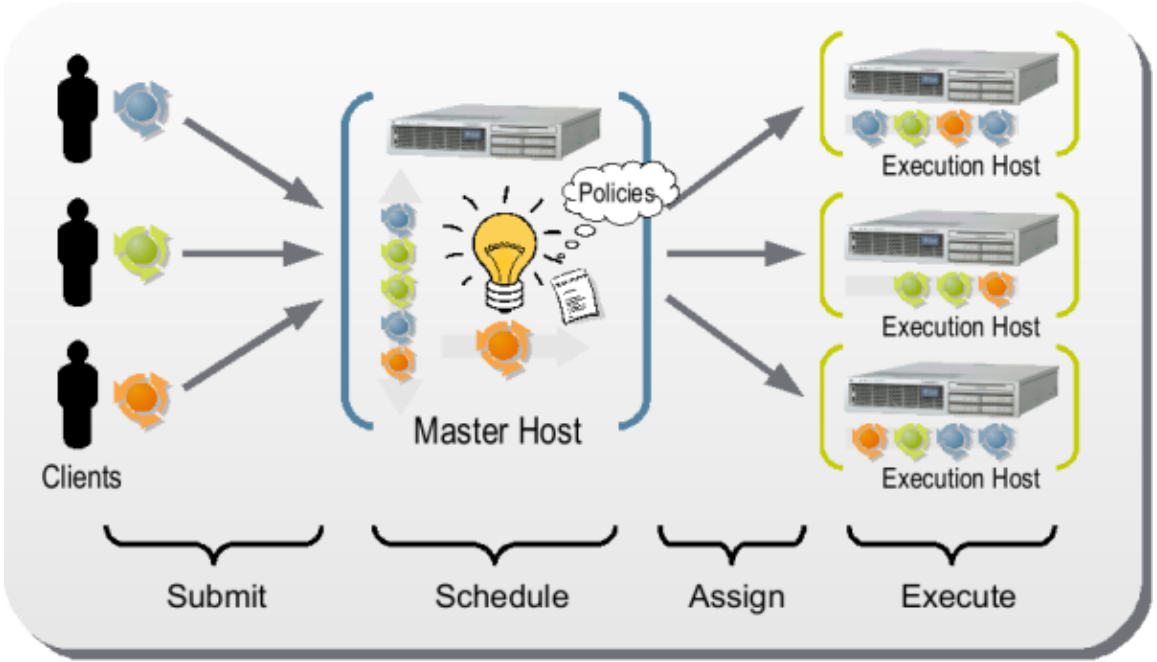
OpenSSH on Windows 10

MobaXterm (recommended by CECI)



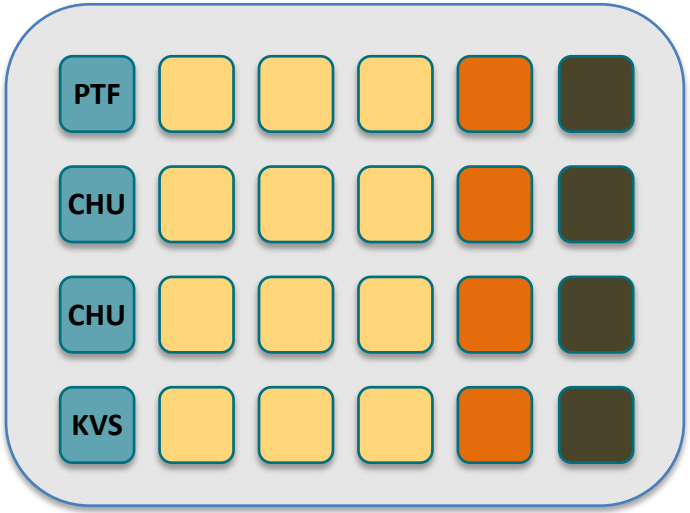
Now that I’m connected, how do I run an analysis?





Nodes partitions

- Available to all with different time limit (all_5hrs, all_24hrs, kosmos)
- Restricted to a group of users (chugen, ptfgen, urtgen)



all_5hrs

all_5hrs + all_24hrs

all_5hrs + kosmos (infinite)

PTF restricted to 1 team

Slurm basics

GIGA cluster partitions

- Available to all with different time limit (all_5hrs, all_24hrs, kosmos)
- Restricted to a group of users (chugen, ptfgen, urtgen)

Before using slurm

```
$ module load slurm
```

Getting info about nodes

```
$ sinfo
$ cat /etc/slurm/slurm.conf | grep ^Node
$ squeue
```

Two types of slurm sessions

Interactive sessions

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

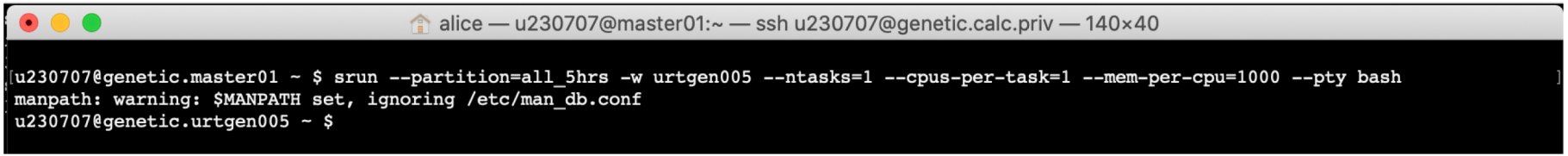
Batch sessions

- Longer running processes
- Parallel processes

Slurm interactive sessions

How to start one ?

$ srun --partition=all_5hrs -w urtgen005 --ntasks=1 --cpus-per-task=1 --mem-per-cpu=1000 --pty bash



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

Don’t forget to close it when you’ve finished !!!!

$ exit

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
echo “Hello”
```

submit

```
$ sbatch myscript.sh
```

```
#!/bin/bash
#
#SBATCH --job-name=Test
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --mem-per-cpu=1000
#SBATCH --time=1:00:00
#SBATCH --partition=all_5hrs   # change to defq on nic4 (CECI cluster)
#SBATCH --output=test_%j.log
#SBATCH --mail-user=alice.mayer@uliege.be
#SBATCH --mail-type=FAIL

date
# Run stuff here
echo "Hello"

#########################################################
### Printing out info about slurm job   #############
#########################################################
echo ""
echo "scontrol show job ${SLURM_JOB_ID} output:"
echo ""
scontrol show job ${SLURM_JOB_ID}
echo ""
date
```

Slurm basics

Monitor jobs

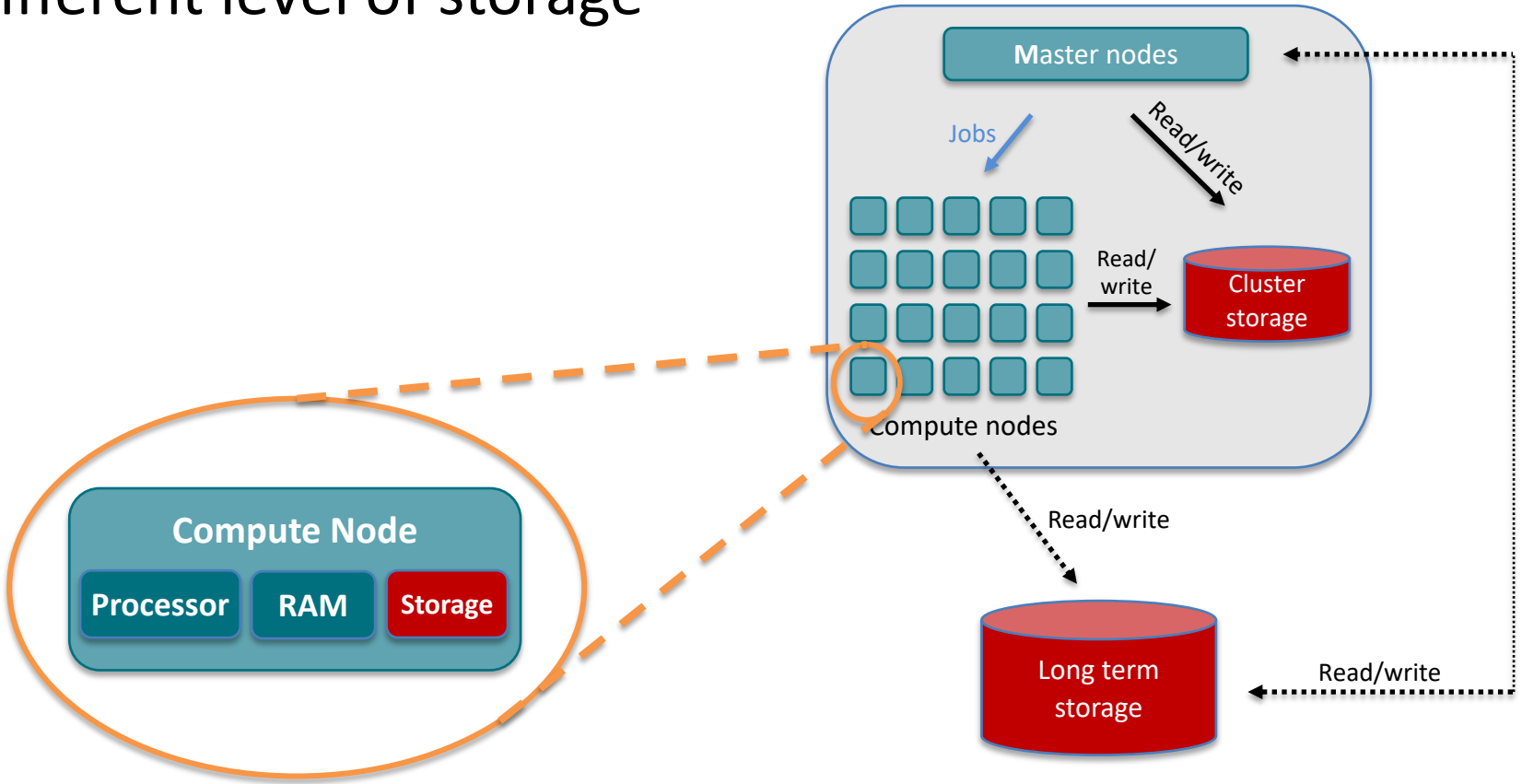
while job is still running, give info on resources, nodes, etc
$ scontrol show job <JOB_ID>

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

Cancel jobs

$ scancel <jobID>

Different level of storage



Different level of storage on GIGA cluster

	Path from node	Speed	Space	Accessible	Backup
Node storage	/local	**+++**	**2 Tb**	Only from 1 node	no
Cluster storage	/gallia/scratch	**+**	**50 Tb**	From all nodes	no
Mass storage	/massstorage	**-**	**1500 Tb**	From all nodes	yes

At the end of your job, don't forget to transfer and delete

- all your files from node storage
- everything you won't need anymore from the cluster storage

Different level of storage on CECI clusters

	Path from node	Speed	Space	Accessible	Backup
Node storage	Local scratch	+++	-	Only from 1 node	no
Cluster storage	Workdir	++	+	From all nodes of 1 cluster	no
	Home	+			
Global storage	GlobalHome + Transfer	-	+++	From all nodes and clusters	no

For more information: https://support.ceci-hpc.be/doc/_contents/ManagingFiles/Storage.html

WARNING: Data in the Workdir can be removed at any time especially during maintenance periods.

Take-Home message and useful links

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 (session starting today, including "Efficient use of Matlab on the cluster" on 25th November)

GIGA clusters

- wiki: https://gitlab.uliege.be/giga-bioinfo/user-guides-wiki/wikis/home
- slurm page: https://gitlab.uliege.be/giga-bioinfo/user-guides-wiki/wikis/cluster/slurm/slurm_home
- slurm manual: https://slurm.schedmd.com/archive/slurm-14.11.11/quickstart.html
- Contact: https://sam.med.uliege.be/ (choose UDI-MED or BIOINFO-GIGA as category)

LIÈGE université
GIGA

Thank you for your attention !
Questions ?

Alice Mayer, PhD
GIGA bioinformatic team
bioinfo.giga@uliege.be


LIÈGE université
GIGA

HPC in Europe

Definition

Computing systems with extremely high computational power that are able to solve hugely complex and demanding problems.

EU priority

HPC is one of the key digital domains where **the EU's investment is due to significantly increase** […]. Moreover, supercomputing will play a key role in Europe's path towards recovery, as it has been identified a **strategic investment priority**.

https://ec.europa.eu/digital-single-market/en/high-performance-computing

Applications

- monitoring and mitigating the effects of climate change
- producing safer and greener vehicles
- **advancing the frontiers of knowledge** in nearly every scientific field
- **drug design**, from testing drug candidate molecules to repositioning existing drugs for new diseases
- **understand the origins and evolution of epidemics and diseases**.

Example

Fighting coronavirus: European supercomputers join pharmaceutical companies in hunt for new drugs

High Performance Computing

Evolving concept

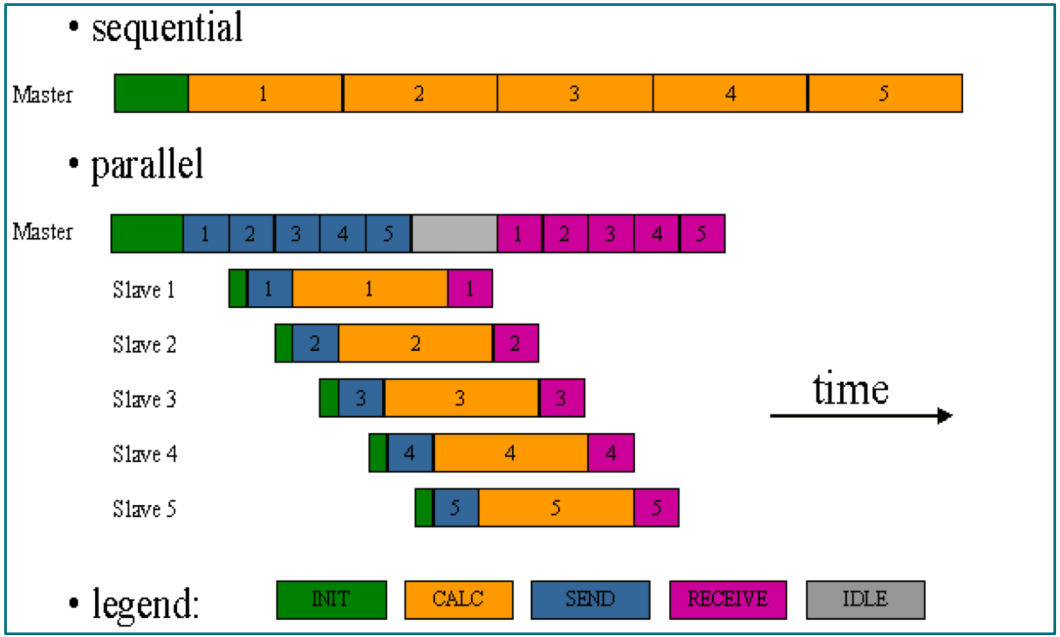
VS

Partnership for Advanced Computing in Europe

- 26 member countries
- 7 supercomputers (5 host countries)
- Ex: Juwels (Germany):
  - 287,136 Gb RAM
  - 123,408 cores

https://prace-ri.eu/hpc-access/hpc-systems/

How much can I parallelize in practice ?



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

Definitions

- Amdahl's law (1967)
  - Theoretical speed up for parallel programs
  - Assumptions:
    - Fraction F of your code perfectly parallelizable
    - Therefore, fraction 1-F of your code is sequential
    - There are no overheads for
      - Scheduling
      - Networking
      - Synchronisation

Definitions

- Amdahl's law
  - Time to run on N cores
    - Example: 4h program
    - Where 3h are parallelisable
    - Then F=75%. Assume N=3
    - Result is 0.5

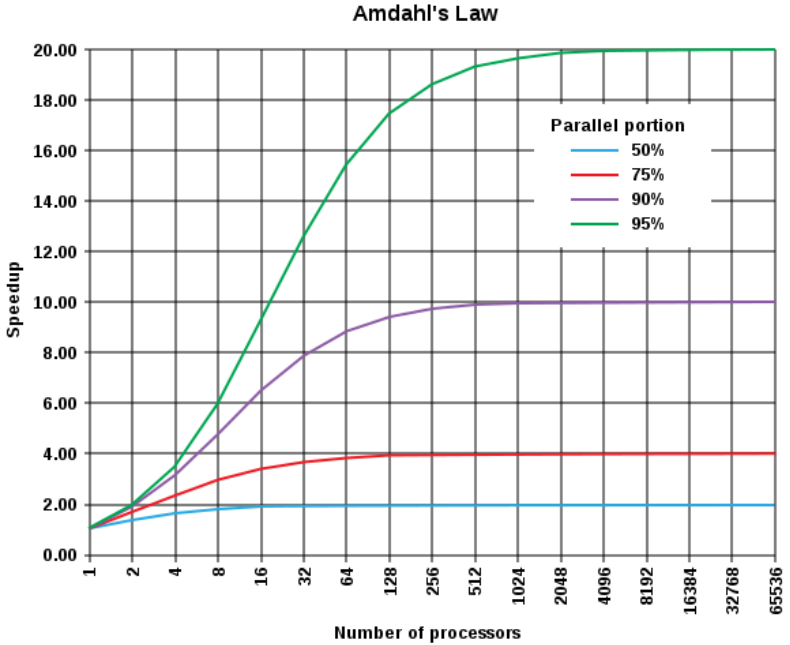
$$(1-F)+\frac{F}{N}$$

  - Speed up
    - Example above
    - Speed up is 2 times faster

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

Definitions

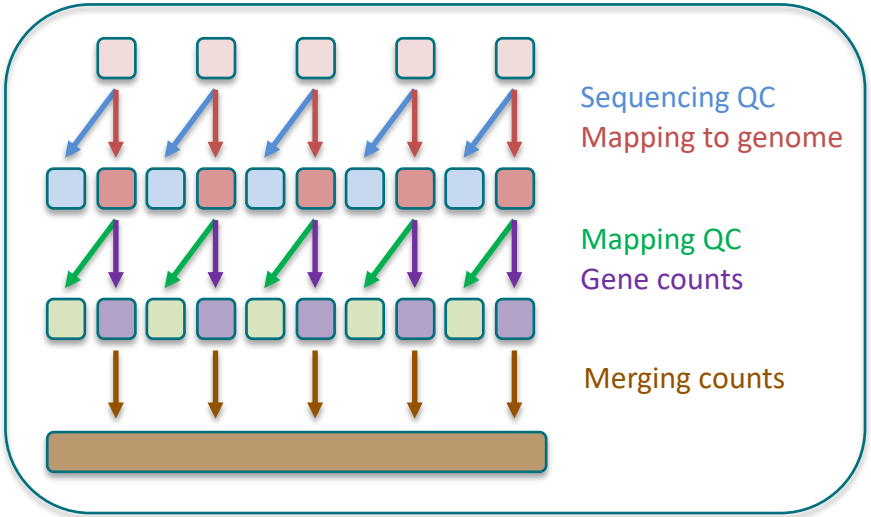
- Amdahl's law



Sometimes, very good tools designed for cluster analysis exist

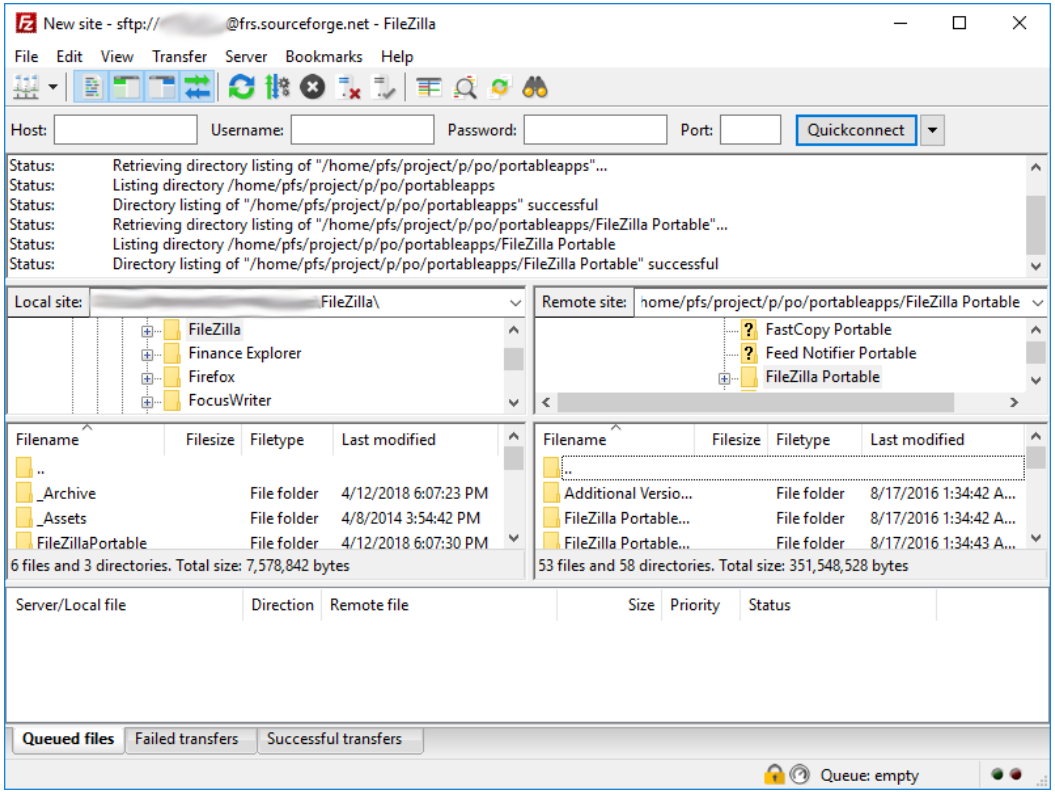
Tools developed to process several samples through several analysis steps,
while optimizing resources usage

nf-core/rnaseq



How do I transfer files from my computer to the CECI cluster or GIGA mass storage ? (or the other way round)

FileZilla



From the GIGA mass storage to the CECI cluster
(or the other way round):
Use **rsync** on command line terminal

How do I transfer files from the CECI cluster to the GIGA mass storage ? (or the other way round)

rsync

1. Open a terminal and log in to the CECI cluster by ssh

2. Define variables for Source and Destination folders:
   - To transfer from CECI cluster to GIGA mass storage
     ```
     $ SourceFolder=Path/To/Source/Folder
     $ DestFolder=u123456@massstorage.giga.priv:Path/To/Folder/from/home
     ```
   - To transfer from GIGA mass storage to CECI cluster
     ```
     $ SourceFolder= u123456@massstorage.giga.priv:Path/To/Folder/from/home
     $ DestFolder=Path/To/Destination/Folder
     ```

3. Synchronise both folders:
   ```
   $ rsync -PHlvrt ${SourceFolder}/<FileName> ${DestFolder}
   ```

NB: Replace u123456 by your university ID