Getting Started With GLiCID: Beginner Session

Mir Junaid

January 11, 2024







OUTLINE

- Introduction to High Performance Computing (HPC)
 - What's HPC?
 - HPC Use Cases
- Introducing NAUTILUS
 - Architecture of Nautilus
- Working With A Supercomputer
 - Basic Linux Commands
 - SSH Connections and access to Nautilus
- SLURM Workload Manager
 - Basic Slurm Commands
 - Batch Scripting
- Modules
- Guix Package Manager
- Data Management

What's HPC?

Data, Data, Everywhere

Key Statistics 2023

- 3.5 quintillion bytes of data is created every single day (Source: Earthweb)
- 333.2 billion emails are sent per day
- 100 billion messages are sent through WhatsApp in a day
- 5 billion Snapchat videos and photos are shared per day
- 456,000 tweets are made on Twitter each minute of the day
- o 500 million daily story users on Instagram every day
- People spend \$1 million per minute online

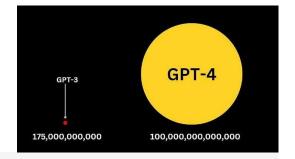


Data, data everywhere, but not a byte to use.



There are only 10 types of people in the world:
Those who understand binary and those who don't.

Data, Data, Everywhere



Tencent ML Images



Tencent AI has now released the largest open-source, multilabel image dataset – **Tencent ML Images**. It contains nearly 18 million images, multi-labeled with up to 11,166 categories.

Neurohive.io

https://neurohive.io/en/datasets/tencent-dataset/

Tencent Released The Largest Multi-Labelled Image Dataset - neur...

We propose **EAGLE**, a large-scale dataset of ~1.1 million 2D meshes resulting from simulations of unsteady fluid dynamics caused by a moving flow source interacting with nonlinear scene structure, comprised of 600 different scenes of three different types.

https://eagle-dataset.github.io/

EAGLE Dataset



Data, data everywhere, but not a byte to use.



There are only 10 types of people in the world:
Those who understand binary and those who don't.

What to do with this data?

- It is through data that
 - groundbreaking scientific discoveries are made,
 - o game-changing innovations are fueled, and
 - o quality of life is improved for billions of people around the globe.
- But we need huge computing power/resources to analyze this humongous data
- **HPC** gives us the power to deal with this data



What is High Performance Computing (HPC)?

- HPC is the ability to process data and perform complex calculations at high speeds
- Laptop/desktop (3 GHz processor) can perform around 3 billion calculations/sec
- HPC solutions can perform quadrillions of calculations/sec (million times faster)
- HPC is the foundation for scientific, industrial, and societal advancements

What is High Performance Computing (HPC)?

- Best-know types of HPC solutions is the Supercomputer
- It is made up of thousands of computers that work together
- Fastest Supercomputer is the US-based Frontier, with a processing speed of
 1.102 exaflops, or quintillion floating point operations per second (flops)
- HPC solutions can be deployed on-premise, at the edge, or even in the cloud



Top 500

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, Atos EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096

Source: https://www.top500.org/ 9

How does HPC work?

- A standard computing system solves problems primarily using serial computing
- It divides the workload into a sequence of tasks, and then executes the tasks one after the other on the same processor
- In contrast, HPC leverages
 - Massively parallel computing
 - Compute clusters (also called HPC clusters)
 - High-performance components

How does HPC work?

Massively parallel computing

• Parallel computing using tens of thousands to millions of cores

Compute clusters/HPC clusters

- Consists of multiple high-speed computer servers networked together
- The computers, called nodes, use either high-performance multi-core CPUs or, more likely today,
 GPUs (graphical processing units)
- Well suited for rigorous computations and graphics-intensive tasks

High-performance components

 Other computing resources in an HPC cluster - networking, memory, storage and file systems - are high-speed, high-throughput and low-latency components that can keep pace with the nodes and optimize the computing power and performance of the cluster

HPC: Use cases

Al and ML

 HPC supports training deep neural networks, processing large datasets, and accelerating machine learning algorithms

Weather and Climate Modelling

 HPC is used to run complex atmospheric models, simulate weather patterns, and predict climate change phenomena

Engineering and Design Optimization

 HPC is employed to optimize engineering designs, analyze structural integrity, simulate fluid dynamics, and enhance product performance

Astrophysics and Cosmology

 HPC facilitates large-scale simulations of the universe, including galaxy formation, stellar evolution, and gravitational wave analysis

HPC: Use cases

Drug Discovery and Molecular Dynamics

• HPC enables the simulation of drug interactions, protein folding, and molecular dynamics, aiding in the development of new pharmaceuticals.

Financial Modeling and Risk Analysis

- HPC helps in analyzing complex financial models, running Monte Carlo simulations, and assessing investment risks.
- And many more...



Introducing Nautilus



Nautilus Cluster

- Nautilus is the fictional submarine belonging to Captain Nemo featured in Jules Verne's novels Twenty Thousand Leagues Under the Seas (1870) and The Mysterious Island (1874).
- Verne took the name "Nautilus" from one of the earliest successful submarines, built in 1800 by Robert Fulton, who also invented the first commercially successful steamboat.





Nautilus Architecture

- Nautilus has 3 main components:
 - Set of nodes communicating with each other
 - Fast interconnect using Infiniband 100 Gb/s technology with high bandwidth and low latency
 - Shared Storage (scratch) 427 TB (IBM/Spectrum Scale- GPFS)

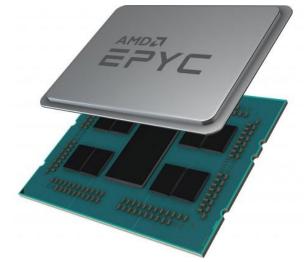


Nautilus Architecture

Each node consists

- Red Hat Operating System (RHEL 8.7)
- 2 AMD EPYC 9474F processors @3.6GHz (4.1GHz Max) with 48 CPU cores
- TDP (Thermal Design Power)/Power Consumption: 360W
- 384 GB RAM







Nautilus Architecture

#Computing nodes	Processor and Speed	RAM	#Cores
40 cnode[301-340]	BullSequana X440 (2 AMD EPYC 9474@3.6GHz 48c)	384 GB	3840
8 cnode[701-708]	BullSequana X440 (2 AMD EPYC 9474@3.6GHz 48c)	768 GB	768
4 visu[1-4]	BullSequana X450 (2 AMD EPYC 9474@3.6GHz 48c) with Nvidia A40 (48G) 2 GPUs per node	768 GB	384
4 gnode[1-4]	4 BullSequana X410 (2 AMD EPYC 9474@3.6GHz 48c) with Nvidia A100 (80G) 4 GPUs per node	768 GB	384

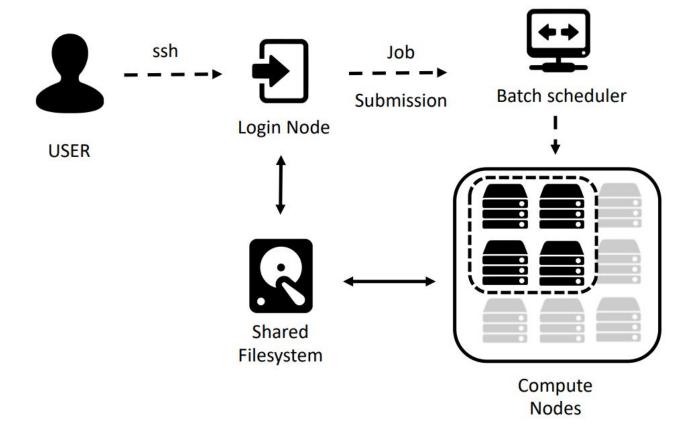
Is NOT like this...







· Status page https://ckc.glicid.fr |imir@ec-nantes.fr@nautilus-devel-001 ~l\$ | jmir@ec-nantes.fr@nautilus-devel-001:~





Login node(s)

- Editing and transferring files
- Compile programs
- Prepare simulations



Compute nodes

- Multicore nodes
- Large memories
- High-speed interconnections



Batch scheduler

- Resource allocation
- Job queueing
- Accounting and



File system

- Parallel FS
- Efficient I/O
- Node local disks

Getting Started with GLiCID

Prerequisites

- macOS
 - Terminal (pre-installed)
- Windows
 - MobaXterm
 - PowerShell
- Linux
 - You are already well equipped :)

LINUX COMMAND LINE



Linux Command Line - Brief History



- One of the earliest operating systems was called **Unix**
- Designed to run as a multi-user system on mainframe computers
- Users connecting to it remotely via individual terminals
- Terminals were pretty basic: just a keyboard and screen
- Send keystrokes to the server and display any data they received on the screen
- No mouse, no fancy graphics, not even any choice of colour
- Everything was sent as text, and received as text
- Programs that ran on the mainframe had to produce text as an output and accept text as an input

Linux Command Line - Brief History





IBM Mainframe, Late 1960's/Early 1970's

Linux Command Line - Brief History



- Linux is a sort-of-descendant of Unix
- The core part of Linux is designed to behave similarly to a Unix system
- Most of the old shells and other text-based programs run on it quite happily
- Most of the <u>Top 500</u> supercomputers use Linux

What's A Command Line?



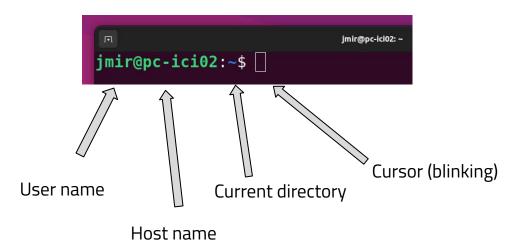
- The Linux command line is a text interface to your computer
- Often referred to as **shell, terminal, console, prompt** or various other names
- It can give the appearance of being complex and confusing to use
- But it is not so scary as it looks
- You just need to memorize a few basic commands



Launch the Terminal

```
jmir@pc-ici02: ~
jmir@pc-ici02:~$
```

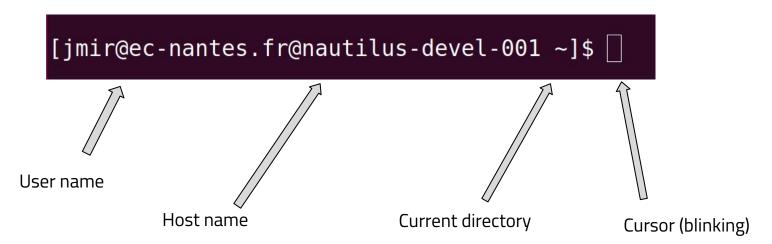
Structure of a linux commands



The system is ready to accept commands



Structure of a linux commands (in Nautilus)



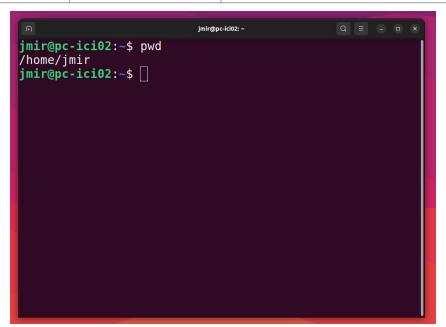
The system is ready to accept commands

Command	Syntax	Description
Print Working Directory	pwd	Print present working directory
List	ls	List files and directories at path
Change directory	cd	Change current directory
Make directory	mkdir	Create new directory
Create empty file	touch	Create new file or update timestamp
Move	mv	Move or rename files and directories
Сору	ср	Copy files or directories from source to destination
Remove	rm	Remove files
Text editor	vim	Vim is a highly configurable text editor



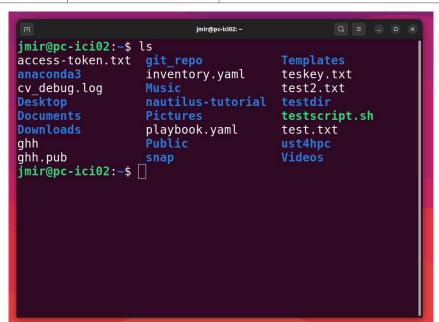


Command	Syntax	Description
Print Working Directory	pwd	Print present working directory



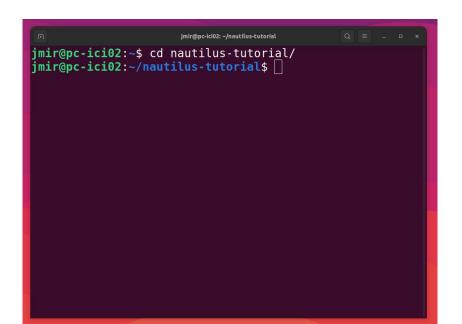


Command	Syntax	Description
List	ls	List files and directories at path





Command	Syntax	Description
Change directory	cd	Change current directory





Command	Syntax	Description
Make directory	mkdir	Create new directory

```
jmir@pc-ici02: ~/nautilus-tutorial
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet
jmir@pc-ici02:~/nautilus-tutorial$ mkdir test-dir
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet test-dir
jmir@pc-ici02:~/nautilus-tutorial$ |
```



Command	Syntax	Description
Create empty file	touch	Create new file or update timestamp

```
jmir@pc-ici02: ~/nautilus-tutorial
jmir@pc-ici02:~/nautilus-tutorial$ touch test-file
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet test-dir test-file
jmir@pc-ici02:~/nautilus-tutorial$ [
```



Command	Syntax	Description
Move	mv	Move or rename files and directories

```
jmir@pc-ici02: ~/nautilus-tutorial/cheat-sheet
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet test-dir
jmir@pc-ici02:~/nautilus-tutorial$ mv test-dir/ cheat-sheet/
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet
jmir@pc-ici02:~/nautilus-tutorial$ cd cheat-sheet/
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$ ls
test-dir test-file
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$
```



Command	Syntax	Description
сору	ср	Copy files or directories from source to destination

```
jmir@pc-ici02: ~/nautilus-tutorial/cheat-sheet
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet test-file
jmir@pc-ici02:~/nautilus-tutorial$ cp test-file cheat-sheet/
jmir@pc-ici02:~/nautilus-tutorial$ cd cheat-sheet/
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$ ls
test-dir test-file
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$
```



Command	Syntax	Description
Remove	rm	Remove files

```
jmir@pc-ici02: ~/nautilus-tutorial/cheat-sheet
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$ ls
test-dir test-file
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$ rm test-file
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$ ls
test-dir
jmir@pc-ici02:~/nautilus-tutorial/cheat-sheet$
```

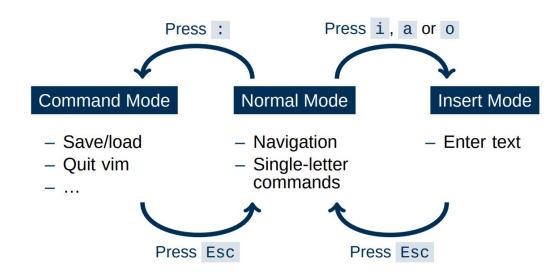


Command	Syntax	Description
Text editor	vim	Vim is a highly configurable text editor

```
jmir@pc-ici02: ~/nautilus-tutorial
#!/bin/bash
mkdir new-dir && cd new-dir
echo "Ciao"
```

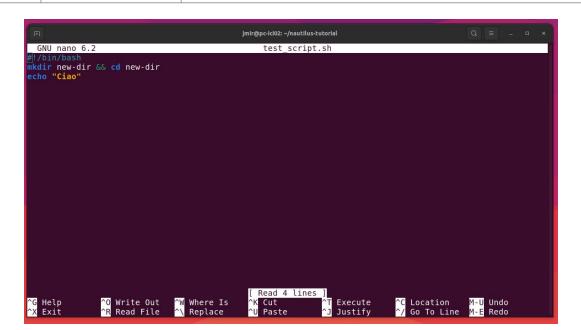


Command	Syntax	Description
Linux editor	vim	Vim is a highly configurable text editor



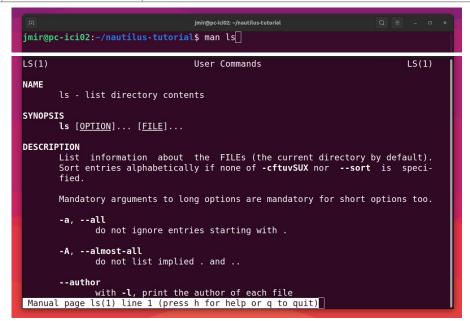


Command	Syntax	Description
Text editor	nano	Comparatively easier (Ctrl+Option)





Command	Syntax	Description
User Manual	man	Displays whole manual of the command



Basic Linux Commands - File Management



Path to folders and files

Relative Path

Absolute path

```
jmir@pc-ici02: ~/nautilus-tutorial/test-dir/
jmir@pc-ici02: ~/nautilus-tutorial/test-dir/
jmir@pc-ici02: ~/nautilus-tutorial/test-dir$
```

Basic Linux Commands - File Management



Local to Remote

scp <file_name> nautilus:/scratch/users/<username>

```
jmir@pc-ici02:-/nautilus-tutorial

jmir@pc-ici02:-/nautilus-tutorial

jmir@pc-ici02:-/nautilus-tutorial

jmir@pc-ici02:-/nautilus-tutorial

jmir@pc-ici02:-/nautilus-tutorial
```

• Remote to Local

scp nautilus:/scratch/users/<username>/<file_name> /<local_path>

```
jmir@pc-ici02:~/nautilus-tutorial
jmir@pc-ici02:~/nautilus-tutorial$ scp nautilus:/scratch/users/jmir@ec-nantes.fr/filename /home/jmir
```

Basic Linux Commands - Large Files



Compress

```
$ tar -czvf <folder_name.tar.gz> <foldername>
```

• Decompress

```
$ tar -xzvf <folder_name.tar.gz>
```

Basic Linux Commands - File Management



Local to Remote

```
$ scp -r folder_name nautilus:/scratch/users/username
```

Remote to Local

```
$ scp -r nautilus:/scratch/users/username/folder_name /local_location
```

Note: Run both commands from the local machine.

Basic Linux Commands - Bash Scripting



- What if we want to run many bash commands?
- ... maybe in a workflow???
- Important part of process automation in Linux
- Plain text file that contains a series of commands
- Any command you run on the command line can be put in a script and vice-versa
- Executed like a program

Basic Linux Commands - Bash Scripting



Simple bash script

```
#!/bin/bash
mkdir test-dir && cd test-dir
echo "Ciao"
```

- Save as test_script.sh
- To execute ./test_script.sh

File Permissions

chmod u+x <filename>

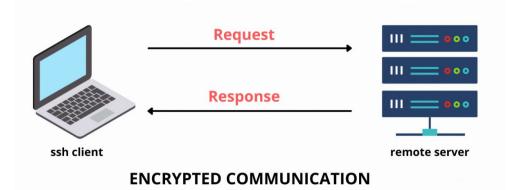
```
jmir@pc-ici02: ~/nautilus-tutorial
jmir@pc-ici02:~/nautilus-tutorial$ touch test script.sh
jmir@pc-ici02:~/nautilus-tutorial$ vi test script.sh
jmir@pc-ici02:~/nautilus-tutorial$ ./test script.sh
bash: ./test script.sh: Permission denied
jmir@pc-ici02:~/nautilus-tutorial$ chmod u+x test script.sh
jmir@pc-ici02:~/nautilus-tutorial$ ./test script.sh
Ciao
jmir@pc-ici02:~/nautilus-tutorial$ ls
cheat-sheet new-dir python pytorch geometric test-dir test script.sh
jmir@pc-ici02:~/nautilus-tutorial$ [
```

SSH CONNECTIONS



What is SSH Key?

- SSH is a secure shell (terminal) connection to another computer
- You connect from your computer to the LOGIN NODE
- Security is given by public/private keys
- A connection to the supercomputer needs a
 - Key,
 - Configuration,
 - Key/IP address known to the supercomputer



How to access GLiCID cluster?

- Create an account on https://clam.glicid.fr
- Account validation by an administrator
- User uploads SSH key to CLAM portal (in profile's SSH Access tab)
- SSH connection configuration on local PC

How to configure SSH connection?



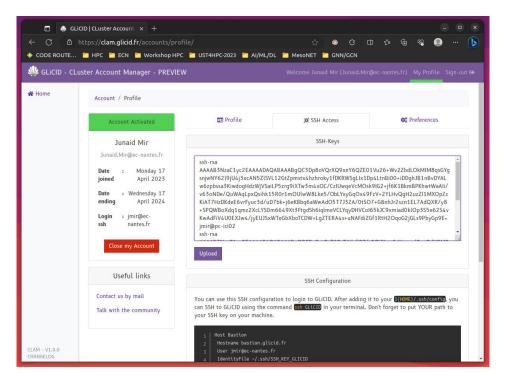
Generate SSH key and copy the public key (id_ed25519.pub)

```
jmir@pc-ici02:~$ ssh-keygen -t ed25519
Generating public/private ed25519 key pair.
Enter file in which to save the key (/home/jmir/.ssh/id_ed25519): y

jmir@pc-ici02:~/.ssh$ cat id_ed25519.pub
ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAID7Tm0MUiYv62VbE/uyk1Gcan9Wfu1IEsg7sBX8
R6Fjw junaid.mir@ec-nanes.fr
jmir@pc-ici02:~/.ssh$
```

How to configure SSH connection?

Upload this SSH key to the CLAM





How to configure SSH connection?

- Copy Contents to the config file and save it
- Replace <my_username> with your username

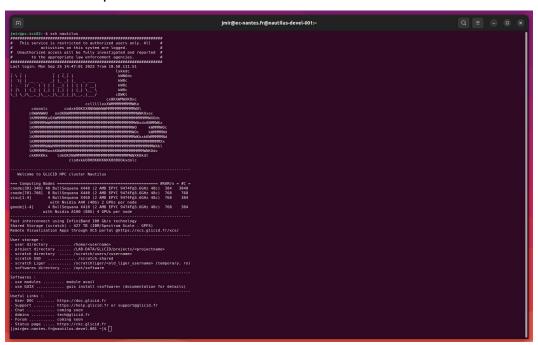
```
jmir@pc-ici02:~/.ssh$ cat config
Host Bastion
        Hostname bastion.glicid.fr
        User jmir@ec-nantes.fr
        IdentityFile ~/.ssh/id ed25519
        ForwardAgent yes
Host glicid
        Hostname login-001.glicid.fr
        User jmir@ec-nantes.fr
        ProxyJump Bastion
        IdentityFile ~/.ssh/id ed25519
Host nautilus
        Hostname nautilus-devel-001.nautilus.intra.glicid.fr
        User jmir@ec-nantes.fr
        ProxyJump glicid
        IdentityFile ~/.ssh/id ed25519
```



- Login using SSH by typing this command in the terminal
 - ssh glicid and then press <Enter>

```
jmir@pc-ici02:~/.ssh$ ssh glicid
The authenticity of host 'bastion.glicid.fr (194.167.60.10)' can't be established.
ED25519 key fingerprint is SHA256:0Szy+0r30Rkizt8TXgKeLqD4qRn8Xq+0YmEE4EsfqrU.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'bastion.glicid.fr' (ED25519) to the list of known hosts.
The authenticity of host 'login-001.qlicid.fr (<no hostip for proxy command>)' can't be established.
ED25519 key fingerprint is SHA256:OSzv+Or3ORkizt8TXqKeLqD4qRn8Xq+OYmEE4EsfqrU.
This host key is known by the following other names/addresses:
    ~/.ssh/known hosts:1: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'login-001.glicid.fr' (ED25519) to the list of known hosts.
Last login: Wed Nov 29 14:07:00 2023 from 194.167.60.12
jmir@ec-nantes.fr@guix-devel-001 ~$ ls
env.yml KEYS ml-container test-containers TP ContainerWorkshop wget-log wget-log.1
jmir@ec-nantes.fr@quix-devel-001 ~$
```

- Login directly to nautilus
 - ssh nautilus and then press <Enter>





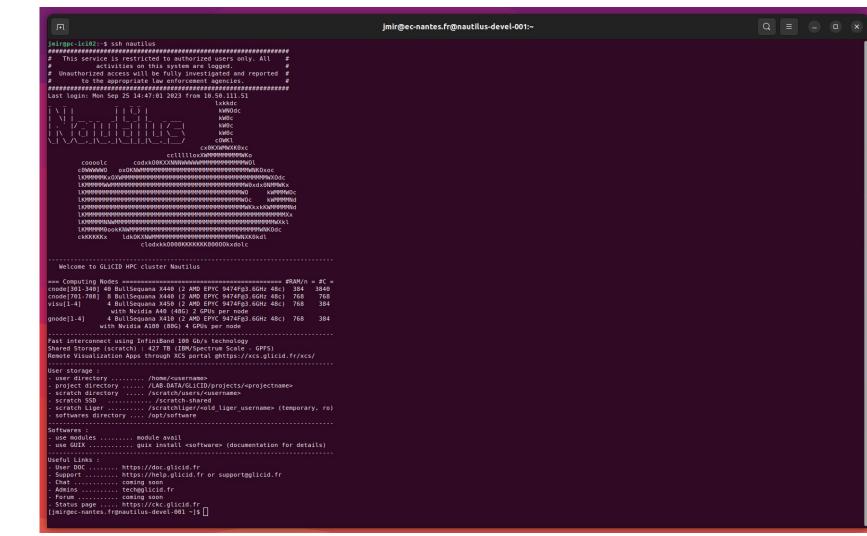


- Switch from glicid to nautilus
 - ssh nautilus-devel-001 and then press <Enter>
 - It will not work but ask for a password



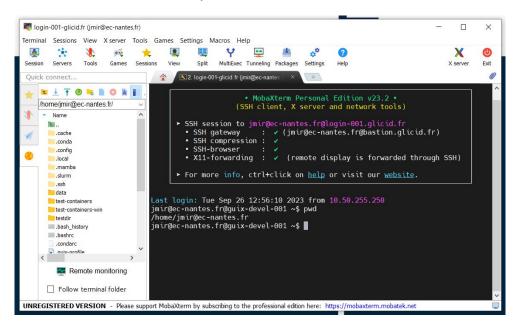
- Switch from glicid to nautilus
 - To switch to nautilus, set-up authorized_keys
 - cd ~/.ssh after logging in to glicid
 - Either generate a new key or copy the private key id_ed2259 (which is not a good idea)

```
jmir@ec-nantes.fr@guix-devel-001 ~/.ssh$ ssh-keygen -t ed25519
Generating public/private ed25519 key pair.
Enter file in which to save the key (/home/jmir@ec-nantes.fr/.ssh/id ed25519):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/jmir@ec-nantes.fr/.ssh/id ed25519
Your public key has been saved in /home/jmir@ec-nantes.fr/.ssh/id ed25519.pub
The key fingerprint is:
SHA256:7bx0Ee+80Lh9ykhp3/5DGxRPzPXa5DkJqQodBSU6/L0 jmir@ec-nantes.fr@guix-devel-001.waves.intra.glicid.f
The key's randomart image is:
+--[ED25519 256]--+
        .Sooo. 0+0
         .o.B. o.
          .XE+ . o
          + 0 0.0
          .= *+0.0
imir@ec-nantes.fr@guix-devel-001 ~/.ssh$ ls
id ed25519 id ed25519.pub known hosts known hosts.old
jmir@ec-nantes.fr@quix-devel-001 ~/.ssh$ cat id ed25519.pub > authorized keys
jmir@ec-nantes.fr@quix-devel-001 ~/.ssh$ ls
authorized keys id ed25519 id ed25519.pub known hosts known hosts.old
jmir@ec-nantes.fr@guix-devel-001 ~/.ssh$ ssh nautilus-devel-001
```





- On Windows
 - Use MobaXTerm
 - Install MobaXterm (Free version) https://mobaxterm.mobatek.net/download.html



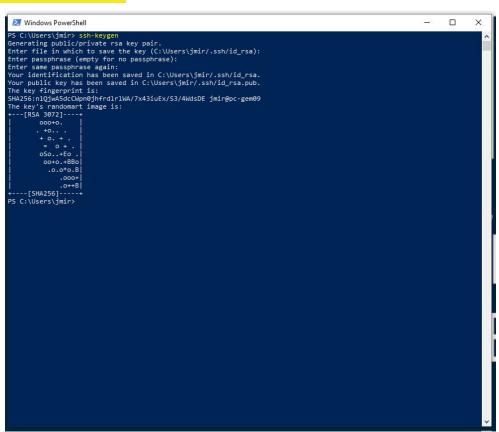


- Open Windows PowerShell
 - Run ssh-keygen (Preferably use ed_25519)
 - Save it in C:\Users\username\.ssh\id_rsa (normally it will be automatic)
 - Don't enter any passwords just press enter (easier)
 - You'll find two files
 - id_rsa.pub and id_rsa.pkk
 - Create an account on https://glicid.clam.fr and upload the public key(id_rsa.pub) (remove spaces if any)(be careful, don't delete anything by mistake)

- Open MobaXterm
- On SSH, configure
 - Remote-host: login-001.glicid.fr
 - Username: <u>imir@ec-nantes.fr</u> (DON'T USE MINE)
 - Port: 22 (automatic)
 - Click Advanced SSH settings
 - Use private key (upload your private key)(id_rsa.pkk)
 - Go to Network Settings -> SSH gateway (jump host)
 - Gateway host: bastion.glicid.fr
 - Username: <u>imir@ec-nantes.fr</u> (DON'T USE MINE)
 - Port:22
 - Use SSH key -> upload private key (id_rsa.pkk)
 - Click OKAY
 - Click OKAY in Session Settings
 - o Double click OR right click and execute a session
 - You'll be logged in.
 - To exit -> Either write exit in the terminal and press enter or click on the exit option.
- Be careful: Everything is case-sensitive and don't use MobaXterm for key generation

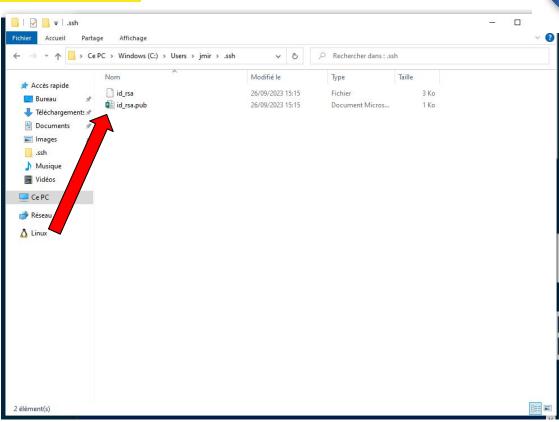
Open PowerShell

ssh-keygen

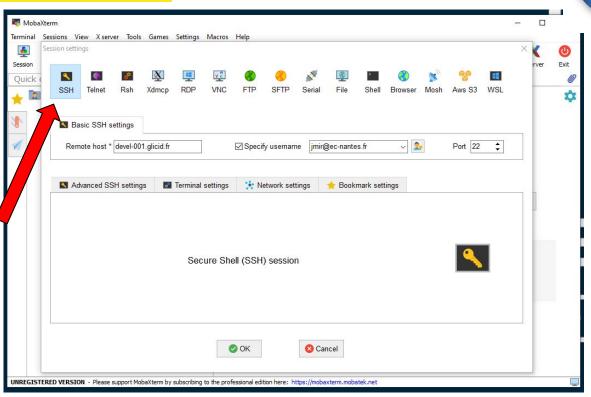




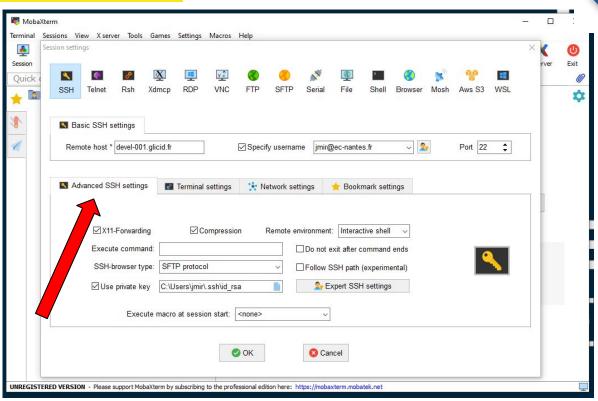
- Keys generated
 - Private key
 - Public key



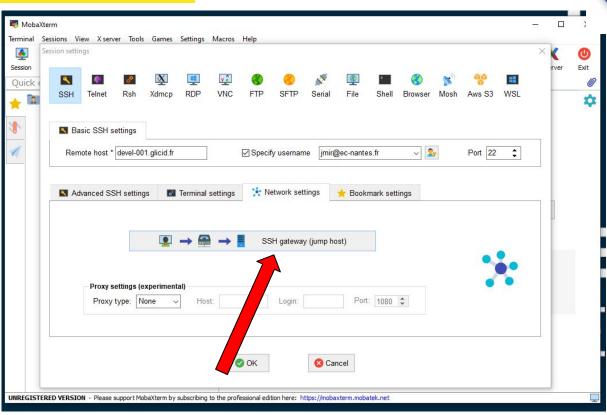
- Open a session
- Configure SSH



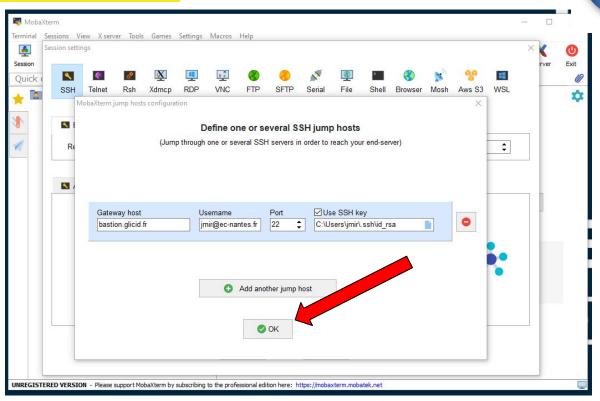
- Advanced SSH
 - Upload id_rsa



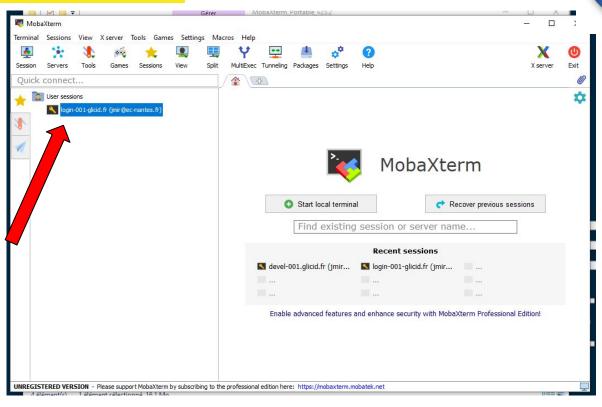
SSH Gateway



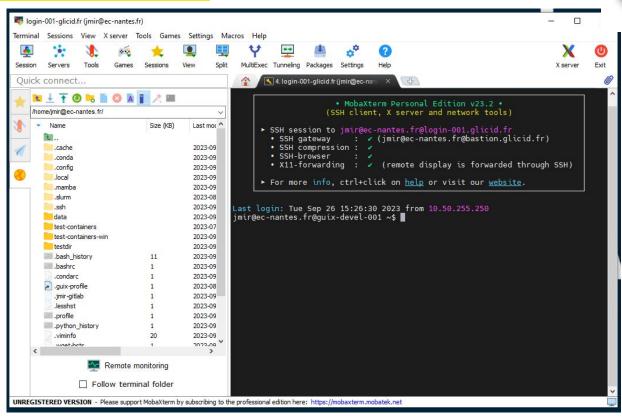
- SSH Gateway
 - Upload id_rsa



Execute Session



Here you go...



WINDOWS: File Management



File Management

- Just drag and drop
- Or using upload option
- Inside the terminal, it's Unix
- o So if you didn't pay interest in the command line session :(

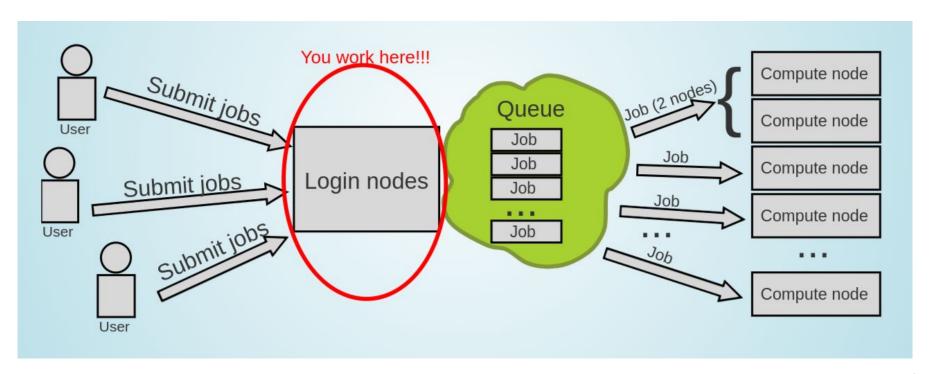
Get Your Hands Dirty



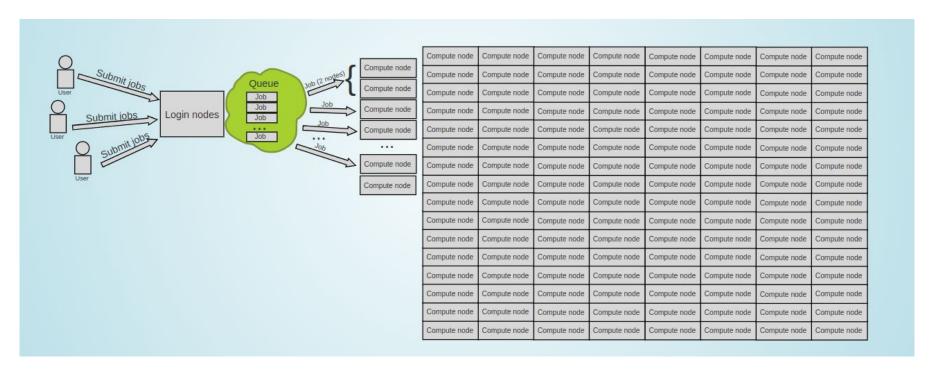
SLURM Workload Manager Slurm



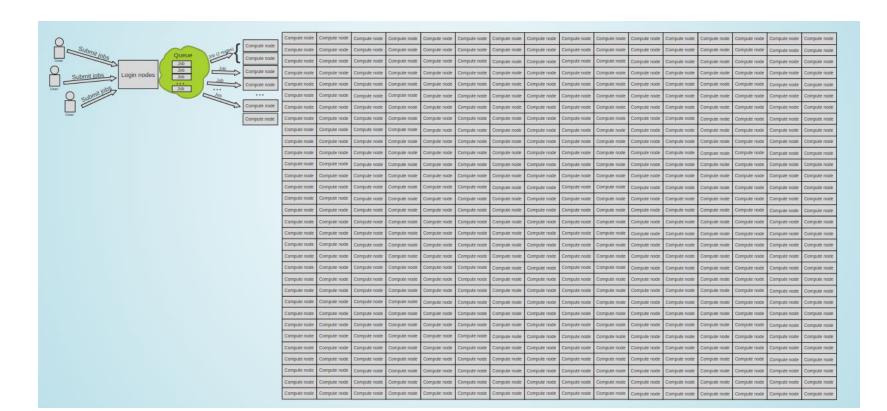
You submit jobs



But you don't use the whole Supercomputer



There are many more users

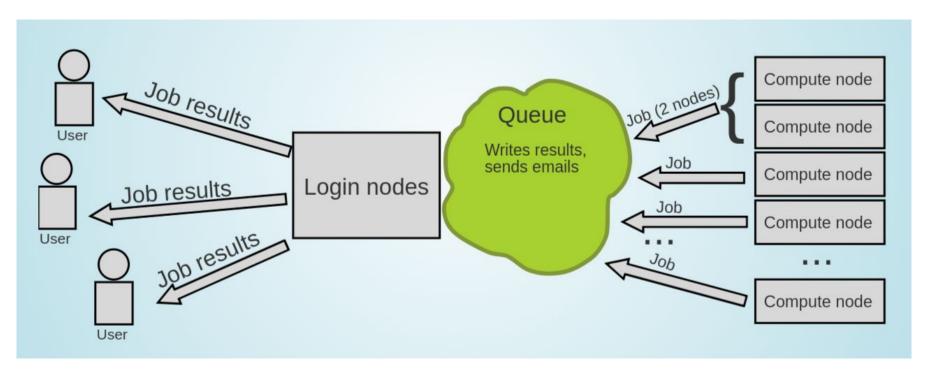


Enter the queue, and wait

- Your job(s) enter the queue,
 and wait for its turn
- When there are enough resources for that job, it runs



Results



This it how it works

- User submits jobs
- Job enters the queue
- When it can, it runs
- Sends results back to user

- CAUTION
- Login nodes are for submitting jobs, move files, compile, etc.
- NOT FOR TRAINING NEURAL NETS

Who will manage this workload?

- Need software that will distribute the jobs appropriately and manage the resources
- Keeps track of what nodes are busy/available, and what jobs are queued or running
- Tells the resource manager when to run which job on the available resources



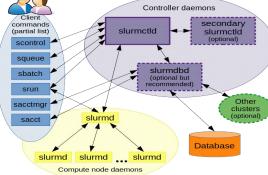
SLURM - Workload Manager



- Simple Linux Utility for Resource Management (SLURM)
- Open source, fault-tolerant, and highly scalable cluster management and job scheduling system for large and small Linux clusters
- It has centralized manager, slurmctld, to monitor resources and work

• Each compute node has a **slurmd daemon**, which can be compared to a remote shell: it waits for work, executes that work, returns status, and waits for more

work.





Command	Syntax	Description
sbatch	sbatch <job_id></job_id>	To submit job script for later execution (batch mode)
sinfo	sinfo	Get information about available nodes
squeue	squeue -u	Show information about jobs
scancel	scancel <job-id></job-id>	To terminate queued or running jobs
srun	srun <resource-parameters></resource-parameters>	To run jobs interactively
sacct	sacct	Show information about current and previous jobs

To submit a job

sbatch job.slurm

```
jmir@ec-nantes.fr@nautilus-devel-001:/scratch/users/jmir@ec-nantes.fr/nautilus-tutorial
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ ls
cuda-script.sh job.slurm numpy-arrays.py pytjob.sh pytorch-tensors.py script.py submit.sh [jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ sbatch job.slurm
Submitted batch job 125254 on cluster waves
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$
```



Get information about available nodes.

sinfo

```
imir@ec-nantes.fr@nautilus-devel-001:~
[jmir@ec-nantes.fr@nautilus-devel-001 ~]$ sinfo
CLUSTER: nautilus
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
                                 1 inval cnode339
standard
                  infinite
                                    down* cnode[323,329]
standard
                  infinite
                                10 drain cnode[302-308,337-338,340]
                  infinite
standard
                                      mix cnode310
standard
                  infinite
                                    alloc cnode[301,309]
                                    idle cnode[312-322,325-328,330-336]
standard
                  infinite
standard
                  infinite
                                     down cnode[311,324]
                  infinite
                                    drain cnode[703-704]
biamem
                                    alloc cnode[701-702,706]
bigmem
                  infinite
bigmem
                                     idle cnode[705,707-708]
                  infinite
                                     mix gnode[1-4]
                  infinite
                  infinite
                                   alloc visu1
visu
                  infinite
                                     idle visu[2-4]
all*
                  infinite
                                    inval cnode339
all*
all*
all*
                  infinite
                                   down* cnode[323,329]
                                12 drain cnode[302-308,337-338,340,703-704]
                  infinite
                  infinite
                                      mix cnode310, gnode[1-4]
                  infinite
                                6 alloc cnode[301,309,701-702,706],visu1
all*
                  infinite
                                     idle cnode[312-322,325-328,330-336,705,707-708],visu[2-4]
                  infinite
                                     down cnode[311,324]
CLUSTER: waves
PARTITION AVAIL
                 TIMELIMIT NODES STATE NODELIST
all*
                   9:00:00
                                     unk* budbud018
all*
                   9:00:00
                                      mix budbud020
all*
                   9:00:00
                                     idle budbud[014-017,019,021-022]
             up 4-04:00:00
                                      mix budbud020
                                     idle budbud[021-022]
[imir@ec-nantes.fr@nautilus-devel-001 ~]$ [
```





To check Priority and MaxWall Time

sacctmgr show qos format="name%20,priority,MaxJobsPerUser,MaxWall"

```
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ sacctmgr show qos format="name%20,priority,MaxJobsPerUser,MaxWall"
              short
             medium
          unlimited
                            10
           priority
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ 🗌
```

Submit your slurm script

sbatch -M nautilus -p standard -qos=short <script-name>.slurm

```
jmir@ec-nantes.fr@nautilus-devel-001:/scratch/users/jmir@ec-nantes.fr/nautilus-tutorial
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ ls
myscript.slurm python test
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ sbatch -M nautilus -gos=short myscript.slurm
Submitted batch job 437079 on cluster nautilus
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$ ls
myscript.slurm python test slurm-437079.out test dir
[jmir@ec-nantes.fr@nautilus-devel-001 nautilus-tutorial]$
```



Slurm - Batch Script

SIUCM workload manager

Sample script to run python code using conda environment

```
#!/bin/bash
#SBATCH --job-name=myjob
                                 # create a short name for your job
#SBATCH --nodes=1
                                 # node count
#SBATCH --ntasks=1
                                 # total number of tasks across all nodes
#SBATCH --cpus-per-task=1
                                 # cpu-cores per task
#SBATCH --mem-per-cpu=2G
                                 # memory per cpu-core
#SBATCH --gres=gpu:4
                                 # number of gpus per node
#SBATCH --time=00:05:00
                                 # total run time limit (HH:MM:SS)
cd /scratch/user/<username>
                                 # go to your working directory / optional
hostname
python myscript.py
```

Software Modules

Software Modules

Modules

- Lot of useful software packages
- Different versions
- Maintained by experts
- Optimized for the architecture
- Users cannot install a module
- Have to request the administrator

How to use Modules?

Useful commands

Command	Description
module avail	List modules
<pre>module avail <package_name></package_name></pre>	List all installed versions of python
module load <package_name></package_name>	Load the default python version
<pre>module load <pakage_name 3.11.5=""></pakage_name></pre>	Load a specific version of python
module unload <package_name></package_name>	Unload python
module list	List currently loaded modules

How to use Modules?

\$module avail

```
jmir@ec-nantes.fr@nautilus-devel-001:~
[imir@ec-nantes.fr@nautilus-devel-001 ~1$ module avail
 ------/usr/share/Modules/modulefiles/applications
castem/2021 castem/2023 gaussian/g16-revA01 hyperworks/2022.2 lammps/15Jun2023 turbomole/7.41
                         ------/usr/share/Modules/modulefiles/libraries
aocl-blis/4.0
                                 intel/ccl/2021.9.0
                                                           intel/dpl/latest
                                                                                          intel/mkl32/2023.1.0
boost/1.82.0 gnu
                                 intel/ccl/latest
                                                           intel/intel ipp ia32/2021.8.0
                                                                                          intel/mkl32/latest
                                 intel/dnnl-cpu-gomp/2023.1.0 intel/intel ipp ia32/latest
cuda/12.2.0 535.54.03
                                                                                          intel/tbb/2021.9.0
                                                           intel/intel ipp intel64/2021.8.0
fftw/3.3.10 intel serial
                                 intel/dnnl-cpu-gomp/latest
                                                                                          intel/tbb/latest
fftw/3.3.10 intel serial sp
                                 intel/dnnl-cpu-iomp/2023.1.0 intel/intel ipp intel64/latest
                                                                                          intel/tbb32/2021.9.0
fftw/intelmpi/3.3.10 intel intelmpi
                                                           intel/intel ippcp ia32/2021.7.0
                                                                                          intel/tbb32/latest
                                 intel/dnnl-cpu-iomp/latest
fftw/omp/3.3.10 intel omp
                                 intel/dnnl-cpu-tbb/2023.1.0
                                                           intel/intel ippcp ia32/latest
                                                                                          libtool/2.4.6 gnu
gmsh/4.11.1 gnu
                                                           intel/intel ippcp intel64/2021.7.0 netcdf/c-4.9.2 gnu
                                 intel/dnnl-cpu-tbb/latest
hdf5/1.14.1-2 gnu
                                 intel/dnnl/2023.1.0
                                                           intel/intel ippcp intel64/latest
                                                                                          netcdf/f-4.6.1 qnu
hdf5/1.14.1-2 intel
                                 intel/dnnl/latest
                                                           intel/mkl/2023.1.0
                                                                                          rdma/46.0 gnu
hdf5/intelmpi/1.14.1-2 intel intelmpi intel/dpl/2022.1.0
                                                           intel/mkl/latest
                                                                                          ucx/1.14.1 anu
                  ------/usr/share/Modules/modulefiles/compilers
            intel/compiler-rt/2023.1.0
amd/4.0.0
                                      intel/compiler-rt32/latest intel/compiler32/2023.1.0 intel/icc/latest
cmake/3.26.4 intel/compiler-rt/latest
                                      intel/compiler/2023.1.0
                                                              intel/compiler32/latest
                                                                                     intel/icc32/2023.1.0
qcc/13.1.0
           intel/compiler-rt32/2023.1.0 intel/compiler/latest
                                                              intel/icc/2023.1.0
                                                                                     intel/icc32/latest
                                      quix/latest
                     intel/clck/latest
                                          intel/debugger/latest
                                                                    intel/dpct/latest
                                                                                            intel/inspector/latest intel/oclfpga/latest
intel/advisor/2023.1.0 intel/dal/2023.1.0
                                          intel/dev-utilities/2021.9.0 intel/init opencl/2023.1.0 intel/itac/2021.9.0
                                                                                                                 intel/vtune/2023.1.0
intel/advisor/latest
                    intel/dal/latest
                                          intel/dev-utilities/latest
                                                                    intel/init opencl/latest
                                                                                            intel/itac/latest
                                                                                                                 intel/vtune/latest
intel/clck/2021.7.3
                    intel/debugger/2023.1.0 intel/dpct/2023.1.0
                                                                    intel/inspector/2023.1.0
                                                                                            intel/oclfpga/2023.1.0
                                    intel/mpi/2021.9.0 intel/mpi/latest openmpi/ucx/4.1.5 gcc 8.5.0 ucx 1.14.1 rdma 46.0
[imir@ec-nantes.fr@nautilus-devel-001 ~]$ □
```

Guix Package Manager

What is Guix?

Guix

- Package building system/Package manager
- Works on GNU/Linux
- Allows each user to manage his/her own packages
 - without root privilege
 - without interfering with other users
- Easy creation of isolated environments with designated packages
 - useful for per-project dependency management

Guix Package Manager



Useful commands

Command	Description
guix pull	You need to run this at least once(maybe weekly :p)
<pre>guix search <package_name></package_name></pre>	Look for a package to install
<pre>guix install <package_name></package_name></pre>	To install a package
<pre>guix remove <package_name></package_name></pre>	To remove a package
guix package -l	List of installed packages

How to use Guix?

\$guix package -1

```
jmir@ec-nantes.fr@nautilus-devel-001:~
[jmir@ec-nantes.fr@nautilus-devel-001 ~]$ quix package -l
Generation 1
               août 01 2023 15:53:10
 graphviz
                       out
                                /qnu/store/8ljq5ipy0qs6w69rjiqz1lvf01zqzaxh-qraphviz-7.0.1
               7.0.1
Generation 2
                sept. 07 2023 16:33:51
+ python
               3.10.7 out
                                /qnu/store/3hnmbi6yyn06w4xz7wpsw5nzb12l6xjv-python-3.10.7
               sept. 20 2023 09:35:38
Generation 3
+ conda
               22.9.0 out
                                /gnu/store/gxrlqm3qqnz3jwz79kwl1125i3lhicb9-conda-22.9.0
Generation 4
                sept. 20 2023 09:37:35
                                /gnu/store/kz02cd8dcgmryb2fk3ylniky2z333yi2-python-numpy-1.23.2
+ python-numpy 1.23.2 out
Generation 5
               sept. 20 2023 10:10:55
Generation 6
                sept. 20 2023 11:52:44
                                        /qnu/store/pq68s6204m0sf7q085qa937wwc6nf65m-python-pandas-1.4.4
+ python-pandas
                       1.4.4 out
Generation 7
               sept. 20 2023 12:58:54
+ python-pytorch
                       1.13.1 out
                                        /qnu/store/icwxw62i5xpfdczv4sih0f58qvw9ah89-python-pytorch-1.13.1
Generation 8
               sept. 20 2023 15:25:47
+ python-numba 0.56.4 out
                                /qnu/store/llns1jkzdwsdvpclgkq37kqr7qp0vq4i-python-numba-0.56.4
Generation 9
               sept. 20 2023 16:20:15
+ python-torchvision 0.15.2 out
                                        /gnu/store/gjmf484aq6kw6gfd02474nl1fwi2wvix-python-torchvision-0.15.2
Generation 10 sept. 20 2023 16:32:05
+ python-pytorch-lightning
                                2.0.2
                                        out
                                                /gnu/store/v2hvz552j2ywrhx6vs8cjyh51ldhx2hd-python-pytorch-lightning-2.0.2
Generation 11
               sept. 22 2023 09:48:01 (current)
+ conda
               22.9.0 out
                                /gnu/store/vf3j4n995jk9zh8yzbvm0mmwajdq71fx-conda-22.9.0
               22.9.0 out
                                /qnu/store/qxrlqm3qqnz3jwz79kw11125i3lhicb9-conda-22.9.0

    conda

[jmir@ec-nantes.fr@nautilus-devel-001 ~]$
```

Data Management

Data management

- HOME (Personal Space/But don't train your neural network here)
- SCRATCH (Train it here)
 - o HDD
 - o SDD
- LAB-DATA
 - Users
 - Projects

```
imir@ec-nantes.fr@nautilus-devel-001:/LAB-DATA/GLICID/projects
[jmir@ec-nantes.fr@nautilus-devel-001 ~]$ pwd
/home/jmir@ec-nantes.fr
[jmir@ec-nantes.fr@nautilus-devel-001 ~]$ cd /scratch/users/jmir@ec-nantes.fr/
[jmir@ec-nantes.fr@nautilus-devel-001 jmir@ec-nantes.fr]$ pwd
/scratch/users/jmir@ec-nantes.fr
[jmir@ec-nantes.fr@nautilus-devel-001 jmir@ec-nantes.fr]$
[jmir@ec-nantes.fr@nautilus-devel-001 jmir@ec-nantes.fr]$ cd
[jmir@ec-nantes.fr@nautilus-devel-001 ~]$ cd /LAB-DATA/
[jmir@ec-nantes.fr@nautilus-devel-001 LAB-DATA]$ ls
[jmir@ec-nantes.fr@nautilus-devel-001 LAB-DATA]$ cd GLiCID/
[jmir@ec-nantes.fr@nautilus-devel-001 GLiCID]$ ls
[jmir@ec-nantes.fr@nautilus-devel-001 GLiCID]$ cd projects/
[jmir@ec-nantes.fr@nautilus-devel-001 projects]$ ls
[jmir@ec-nantes.fr@nautilus-devel-001 projects]$
```

Get Your Hands Dirty Again



Thank you. Any questions?



Please answer the survey if you haven't yet https://forms.gle/B4dto4axGm4EVPwaA

Useful links:

User Doc: https://doc.glicid.fr

Support: https://help.glicid.fr or support@glicid.fr

Chat: On CLAM website

Admins: tech@glicid.fr

Forum: Coming soon

Status page: https://ckc.glicid.fr