Grid 5k - theory and essential commands

Fanny Ducel

fanny.ducel@inria.fr / https://fannyducel.github.io/teaching/

Main sources of information

https://www.grid5000.fr/w/Grid5000:Home

https://calcul.math.cnrs.fr/attachments/evt/2020-04-formation-g5k/slides.pdf

Inspiration from Chuyuan Li's slides from 2023-2024
https://members.loria.fr/ChuyuanLi/activities/

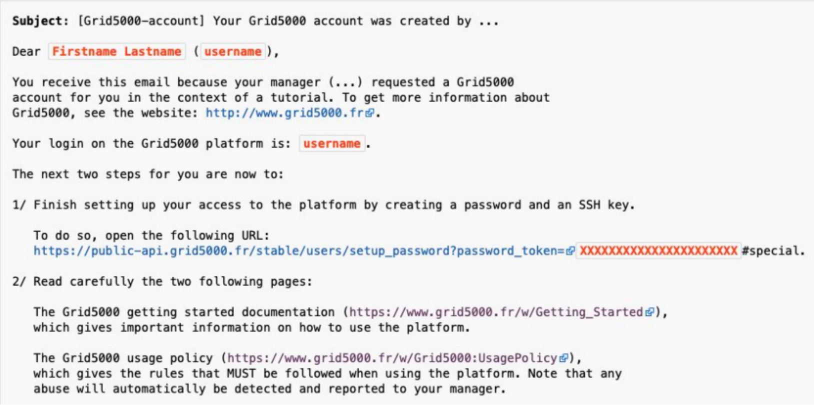
Before we start

What is Grid5k?

What you need to use Grid (essential commands)

Environmental costs

Did you all receive an e-mail about your accounts?



Subject: [Grid5000-account] Your Grid5000 account was created by ...

Dear Firstname Lastname (username),

You receive this email because your manager (...) requested a Grid5000 account for you in the context of a tutorial. To get more information about Grid5000, see the website: http://www.grid5000.fr.

Your login on the Grid5000 platform is: username.

The next two steps for you are now to:

1/ Finish setting up your access to the platform by creating a password and an SSH key.

To do so, open the following URL: https://public-api.grid5000.fr/stable/users/setup_password?password_token=XXXXXXXXXXXXXXXXXXXXXX#special.

2/ Read carefully the two following pages:

The Grid5000 getting started documentation (https://www.grid5000.fr/w/Getting_Started), which gives important information on how to use the platform.

The Grid5000 usage policy (https://www.grid5000.fr/w/Grid5000:UsagePolicy), which gives the rules that MUST be followed when using the platform. Note that any abuse will automatically be detected and reported to your manager.

Warning: no practice for now!

This first morning session is mostly theory.

We will put everything in practice, on your own accounts, **this afternoon**.

Before we start

What is Grid5k?

What you need to use Grid (essential commands)

Environmental costs

Intuitions for understanding what is Grid5k

"A large-scale and flexible testbed for experiment-driven research"
– *Grid5000 HomePage*

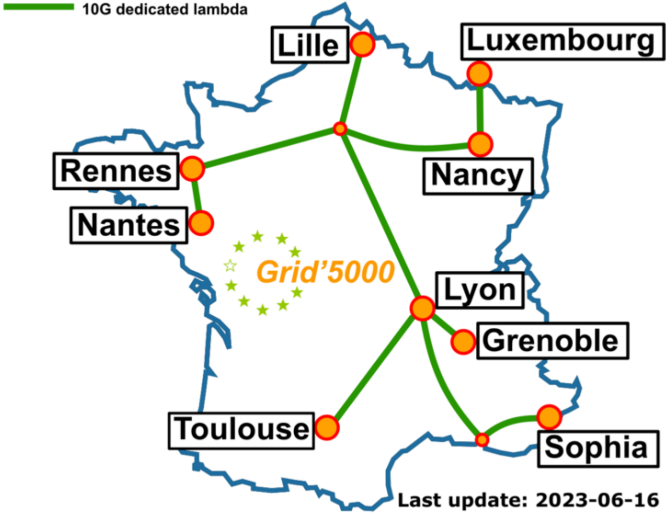
⇒ A distributed cluster of computers

⇒ An infrastructure ("testbed") to compute and run experiments that require large amount of resources

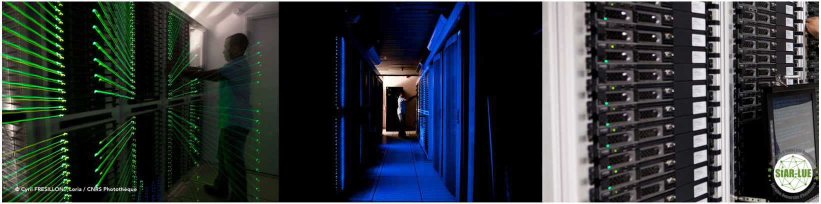
- Example: to use GPUs (e.g., for LLMs), to run time-consuming tasks without occupying 90% of your CPU, for your supervised project (...?)

Key numbers and where is it?

- 8 sites, 31 clusters, 828 nodes, 12328 cores



What does it look like?



Cyril Fresillon, CNRS Photothèque, in Nancy (LORIA)
https://images.cnrs.fr/recherche?direct-query=grid%205000 https://www.grid5000.fr/w/Nancy:Home

Key concepts

- **Site:** Geographically located all over France. For you, this is the "help desk", this is where you go to 'borrow' one or more nodes. Once you have a G5k account, you can get into any site and start working.

Key concepts

- **Site:** Geographically located all over France. For you, this is the "help desk", this is where you go to 'borrow' one or more nodes. Once you have a G5k account, you can get into any site and start working.

- **Cluster:** Group of nodes sitting in one giant server room, spread over the different sites. Different clusters have different kinds of nodes.

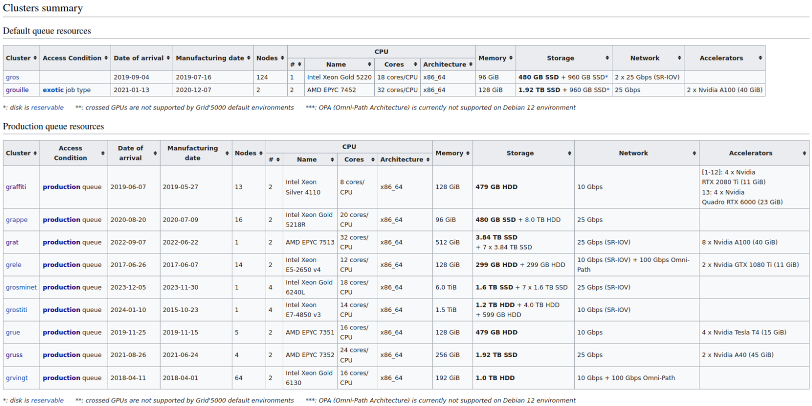
Key concepts

- **Site:** Geographically located all over France. For you, this is the "help desk", this is where you go to 'borrow' one or more nodes. Once you have a G5k account, you can get into any site and start working.

- **Cluster:** Group of nodes sitting in one giant server room, spread over the different sites. Different clusters have different kinds of nodes.

- **Node:** A computer, something that can run code. Typically running Ubuntu/some other Linux system. Having a CPU, RAM, and sometimes a GPU. For your work you would 'borrow' one or more nodes from the "help desk".

Nancy clusters



Clusters summary

Default queue resources

Cluster	Access Condition	Date of arrival	Manufacturing date	Nodes	CPU				Memory	Storage	Network	Accelerators
					#	Name	Cores	Architecture				
gros		2019-09-04	2019-07-16	124	1	Intel Xeon Gold 5220	18 cores/CPU	x86_64	96 GiB	**480 GB SSD** + 960 GB SSD*	2 x 25 Gbps (SR-IOV)	
grouille	**exotic** job type	2021-01-13	2020-12-07	2	2	AMD EPYC 7452	32 cores/CPU	x86_64	128 GiB	**1.92 TB SSD** + 960 GB SSD*	25 Gbps	2 x Nvidia A100 (40 GiB)

: disk is reservable* *: crossed GPUs are not supported by Grid'5000 default environments* ****: OPA (Omni-Path Architecture) is currently not supported on Debian 12 environment*

Production queue resources

Cluster	Access Condition	Date of arrival	Manufacturing date	Nodes	CPU				Memory	Storage	Network	Accelerators
					#	Name	Cores	Architecture				
graffiti	**production** queue	2019-06-07	2019-05-27	13	2	Intel Xeon Silver 4110	8 cores/CPU	x86_64	128 GiB	**479 GB HDD**	10 Gbps	[1-12]: 4 x Nvidia RTX 2080 Ti (11 GiB) 13: 4 x Nvidia Quadro RTX 6000 (23 GiB)
grappe	**production** queue	2020-08-20	2020-07-09	16	2	Intel Xeon Gold 5218R	20 cores/CPU	x86_64	96 GiB	**480 GB SSD** + 8.0 TB HDD	25 Gbps	
grat	**production** queue	2022-09-07	2022-06-22	1	2	AMD EPYC 7513	32 cores/CPU	x86_64	512 GiB	**3.84 TB SSD** + 7 x 3.84 TB SSD	25 Gbps (SR-IOV)	8 x Nvidia A100 (40 GiB)
grele	**production** queue	2017-06-26	2017-06-07	14	2	Intel Xeon E5-2650 v4	12 cores/CPU	x86_64	128 GiB	**299 GB HDD** + 299 GB HDD	10 Gbps (SR-IOV) + 100 Gbps Omni-Path	2 x Nvidia GTX 1080 Ti (11 GiB)
grosminet	**production** queue	2023-12-05	2023-11-30	1	4	Intel Xeon Gold 6240L	18 cores/CPU	x86_64	6.0 TiB	**1.6 TB SSD** + 7 x 1.6 TB SSD	25 Gbps (SR-IOV)	
grostiti	**production** queue	2024-01-10	2015-10-23	1	4	Intel Xeon E7-4850 v3	14 cores/CPU	x86_64	1.5 TiB	**1.2 TB HDD** + 4.0 TB HDD + 599 GB HDD	10 Gbps (SR-IOV)	
grue	**production** queue	2019-11-25	2019-11-15	5	2	AMD EPYC 7351	16 cores/CPU	x86_64	128 GiB	**479 GB HDD**	10 Gbps	4 x Nvidia Tesla T4 (15 GiB)
gruss	**production** queue	2021-08-26	2021-06-24	4	2	AMD EPYC 7352	24 cores/CPU	x86_64	256 GiB	**1.92 TB SSD**	25 Gbps	2 x Nvidia A40 (45 GiB)
grvingt	**production** queue	2018-04-11	2018-04-01	64	2	Intel Xeon Gold 6130	16 cores/CPU	x86_64	192 GiB	**1.0 TB HDD**	10 Gbps + 100 Gbps Omni-Path	

: disk is reservable* *: crossed GPUs are not supported by Grid'5000 default environments* ****: OPA (Omni-Path Architecture) is currently not supported on Debian 12 environment*

https://www.grid5000.fr/w/Nancy:Hardware

Queues and Usage Policy

Default queue:

- Daytime is dedicated to smaller-scale experiments
- Large-scale jobs must be executed during nights or weekends
- Better to use interactive jobs during the day (development phase)
- Read carefully the rules in case of violation of usage

Production queue:

- Smaller set of resources
- Only in the Nancy site
- Suitable for long-running, non-interactive jobs

More information:
`https://www.grid5000.fr/w/Grid5000:UsagePolicy`

Essential commands: entering Grid5k

You will need:

- To create an account
- To access the server
- To send files from your machine (local files) to the server and from the server to your machine
- To reserve resources and run your code

Authentification and access to Grid5k

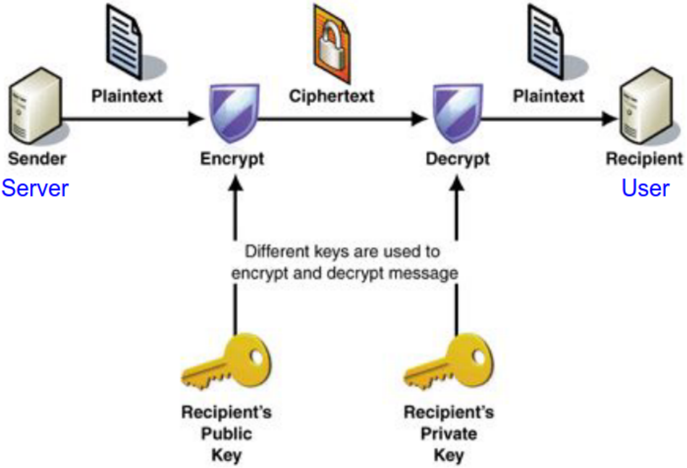
This command is used to connect your machine to the remote server: **ssh login@access.grid5000.fr** [1]

- SSH? = Secure SHell
- To establish a secure communication channel between 2 machines: yours and grid's site
- Public-key authentification (assymmetric cryptography): you have to own a paired private key

[1] or *ssh nancy.g5k* with the alias

Authentification and access to Grid5k

ssh login@access.grid5000.fr



First time on Grid5k

Connect to access machine: **ssh login@access.grid5000.fr**

- Specify a site: **ssh nancy** (or another site: grenoble lille luxembourg lyon nancy nantes rennes sophia toulouse)
- Put in your password
- View machine list of this site

Some warnings

- No BACKUP in g5k, so make sure your important files are stored somewhere outside
- In each site, by default 25 GB storage
- If needed, can ask for more space: *manage account* → *cnrs photohomedir quotas* → *request quota extension*

Transferring files to/from Grid

From your machine to Grid:

```
scp Documents/awesome-file.py  nancy.g5k:~/My-Grid-Folder/
```

From Grid to your machine:

```
scp nancy.g5k:~/My-Grid-Folder/main.py  ~/Documents
```

To transfer folders: add **-r** after **scp**:

```
scp -r Documents/  nancy.g5k:~/My-Grid-Folder/
```

To find different files with a same pattern: *****

```
scp Documents/*.py  nancy.g5k:~/MyGridFolder/
```

Jobs

When you go to a site and ask to borrow one (or more) node(s) to do some work, it is a **job**.

Two types of jobs:

- non interactive (= passive): tell the site to run this program and it will run it in the background.
- interactive: tell the site to let you inside a node where you can do nothing... or run one (or more) programs

Resource reservation with OAR

- OAR: resources and jobs management system (batch manager) in g5k

- Smallest unit of resource: core (cpu core)

→ E.g.: graffiti-1 have 2 CPU with 8 cores/CPU, max used for 16 jobs

- By default a OAR job reserves a host (= node, physical computer with all cpu/cores)

- Ressource reservation syntax: **‘oarsub‘**

Resource reservation with OAR - interactive mode

As soon as a resource is available, you'll be connected to it for 1h.

- Interactive mode: use option **-I** (capital i)
- To reserve GPU (only in production queue in Nancy):
  oarsub -I gpu=1 -l -q production [2]
- Specify number of nodes:
  oarsub -I -l nodes=2
- Specify how many hours:
  oarsub -I -l walltime=3 / oarsub -I -l walltime=00:20

[2] capital I first, then lower L

Resource reservation with OAR - passive mode

Difference: no worries about accidentally terminating your task (terminal closed or network disconnection)

- Passive mode: by default, no option needed

Step 1) Reserve a node and ask it to sleep: **oarsub "sleep 10d"**

Step 2) Allocate a jobID quickly

Step 3) Then use this command to enter the host:
oarsub -C <jobID>

Tip: reserve a specific time in the future with **-r**:
oarsub -l nodes=2, walltime=2 -r '2025-24-09 11:30:00'

Resource reservation with OAR - essential commands

A basic, 2h long job: **oarsub -l walltime=2 "sleep infinity"**[3]

An interactive job with GPU:
oarsub -I -q production -l gpu=1,walltime=2

A job with a specific cluster:
oarsub -q production -p "cluster=’grele’" -l gpu=1,walltime=2 -I

Go back to a job: **oarsub -c <jobID>**

More commands: https://www.grid5000.fr/w/Getting_Started#Reserving_resources_with_OAR:_the_basics

[3] Sleep infinity prevents "unanticipated termination of your jobs in case of errors (terminal closed by mistake, network disconnection)", cf. G5k doc

Other helpful commands - job management

See current jobs state: **oarstat -u**

See current jobs info (nodes, times, ...): **oarstat -f -j <jobID>**

Extend a reservation (if the resource is available):
oarwalltime <jobID> +00:30

Exit a job: **exit** or CTRL+d

Kill a job: exit, then **oardel <jobID>**

TMUX and tips

TMUX: a window manager for the terminal
→ multiple panels/windows
→ session continues even if connection breaks/terminal is closed

Install TMUX: **sudo apt install tmux**

Essential commands:

TMUX and tips

TMUX: a window manager for the terminal
→ multiple panels/windows
→ session continues even if connection breaks/terminal is closed

Install TMUX: **sudo apt install tmux**

Essential commands:

- Check if there is an existing tmux session going on: **tmux ls**
- YES → go to this session: **tmux a -t <session-name>**
- NO → start a new session: **tmux new -s <session-name>**
- Open several panels:
  CTRL+b, then " (vertical split) or % (horizontal split)

Before we start

What is Grid5k?

What you need to use Grid (essential commands)

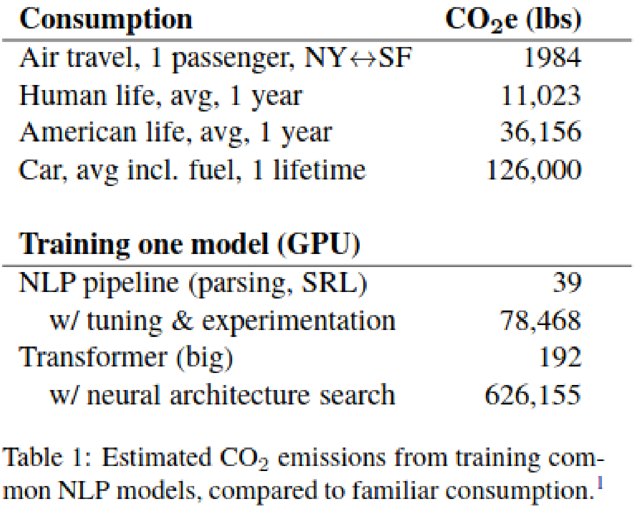
Environmental costs

Running jobs costs money and energy

- Having access to Grid5k as a master's student is a chance and a luxury
- It is professional material
- Running jobs costs money and energy (= environmental resources), it is not magic

⇒ **So you have to be conscious about how you use it!**

Environmental costs - more concretely



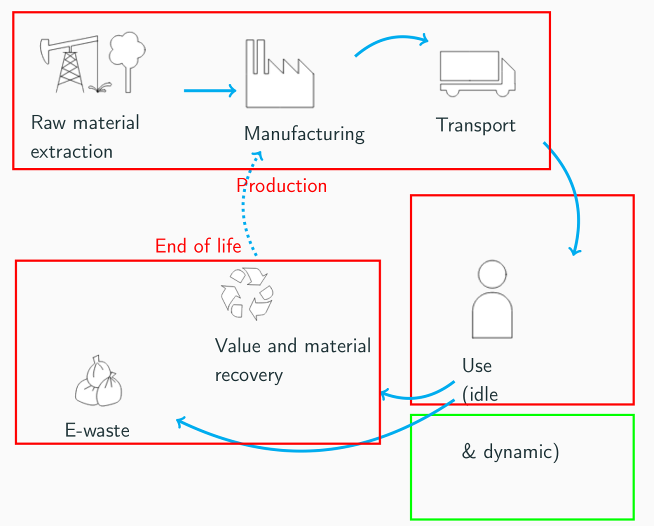
Consumption	CO_2e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000

Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

[Strubell et al., 2019]

Environmental issues are underestimated



A. Névéol

- Carbon footprints measure are underestimations: other sources of emissions, different results [Bannour et al., 2021]

Environmental racism

> *"The negative effects of climate change are reaching and impacting the world's* ***most marginalized communities first*** *[...] Is it fair or just to ask, for example, that the residents of the Maldives (likely to be underwater by 2100) or the 800,000 people in Sudan affected by drastic floods* ***pay the environmental price of training and deploying ever larger English LMs****, when similar large-scale models aren't being produced for Dhivehi or Sudanese Arabic?"*

[Bender et al., 2021]

Estimate the environmental impacts of your projects

- Write down the number of hours and the nodes you used for your project or check your statistics at: https://api.grid5000.fr/stable/users/
- Try and get an estimation from http://calculator.green-algorithms.org/
- Add it to your project and try to see what it represents IRL

Other resources to look up and consider using in your projects:
https://github.com/blubrom/MLCA
https://github.com/saintslab/carbontracker

CPU vs. GPU

- GPU are more powerful and faster
- BUT, GPUs consume more energy than CPU
- Be sure that you **need** a GPU before using one

Your turn: Do some research about the differences between GPUs and CPUs in general, and in terms of environmental impact

Bannour, N., Ghannay, S., Névéol, A., and Ligozat, A.-L. (2021).
Evaluating the carbon footprint of NLP methods: a survey and analysis of existing tools.
In Proceedings of the Second Workshop on Simple and Efficient Natural Language Processing, pages 11–21, Virtual. Association for Computational Linguistics.

Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021).
On the dangers of stochastic parrots: Can language models be too big?
In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, page 610–623, New York, NY, USA. Association for Computing Machinery.

Strubell, E., Ganesh, A., and McCallum, A. (2019).
Energy and policy considerations for deep learning in NLP.

In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 3645–3650, Florence, Italy. Association for Computational Linguistics.