How to connect to today's training account?

Informations for today's training:

Today's on **JeanZay specific training accounts**:
- Connect on "`ipcours22.……`" session.
- Your login: use the `cforXXX` login/password provided to you *(complete the form with your name and signature)*

If you need to **switch between qwerty** and **azerty** you can use the command `alt+shift`.

Introduction to IPSL modeling tools and environment (modipsl and libIGCM)

Institut
Pierre-**S**imon
Laplace

January 2023, IDRIS
IPSL « Platforme » group

Schedule for these 2 days:
- 9am-9:30am welcome ; 9:30am-5:30pm training.
- Lunch around 12:00
- After this presentation, all training will be practical work. Take breaks at your convenience.

Don't hesitate to ask questions during the training!

Objectives of this course:
- Have an overview of the tools used to launch Earth climate models, and to know how to use them.
- Know and understand the environments at your disposal (supercomputers).
- Get an idea of the context at IPSL in terms of work teams and models.

Not seen:
- **We will not explain how each model works** (parameters, specific features), how to launch a zoom or a non-standard resolution.
- You will not see details about parallelisation.

=> look for dedicated trainings!

Institut Simon Laplace (IPSL)

Institut Pierre-Simon Laplace

IPSL gathers 8 laboratories for which research topics concern the global environment.

CEREA / GEOPS / LATMOS / LISA / LMD / LOCEAN / LSCE / METIS

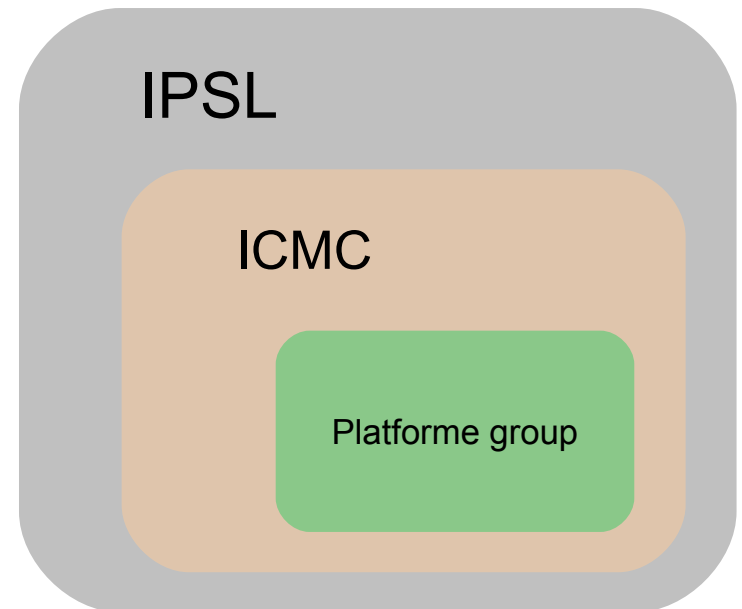**IPSL Climate Modeling Centre** (ICMC https://cmc.ipsl.fr/)

Activities articulated around :

- The development of an integrated model of the Earth system
- Running and analysing climate simulations
- Working groups to share skills
- A scientific expertise

*To be involved in ICMC activities, subscribe to the mailing list **ipsl_cmc@listes.ipsl.fr***

IPSL Platforme group : in charge of the development of modipsl, libIGCM, XIOS usage, metric tools deployment
*To contact the list: **platform-users@listes.ipsl.fr***

IPSL

ICMC

Platforme group

Institut
Pierre-Simon
Laplace

Arnaud Caubel

Anne Cozic

Romain Pennel

Christian Ethé

Jérôme Servonnat

Renaud Person

Laurent Fairhead

Olivier Marti

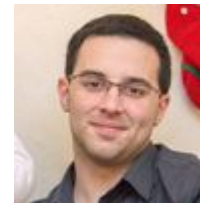**Josefine Ghattas**

Nicolas Lebas

Lola Falletti

Olivier Boucher

Thibaut Lurton

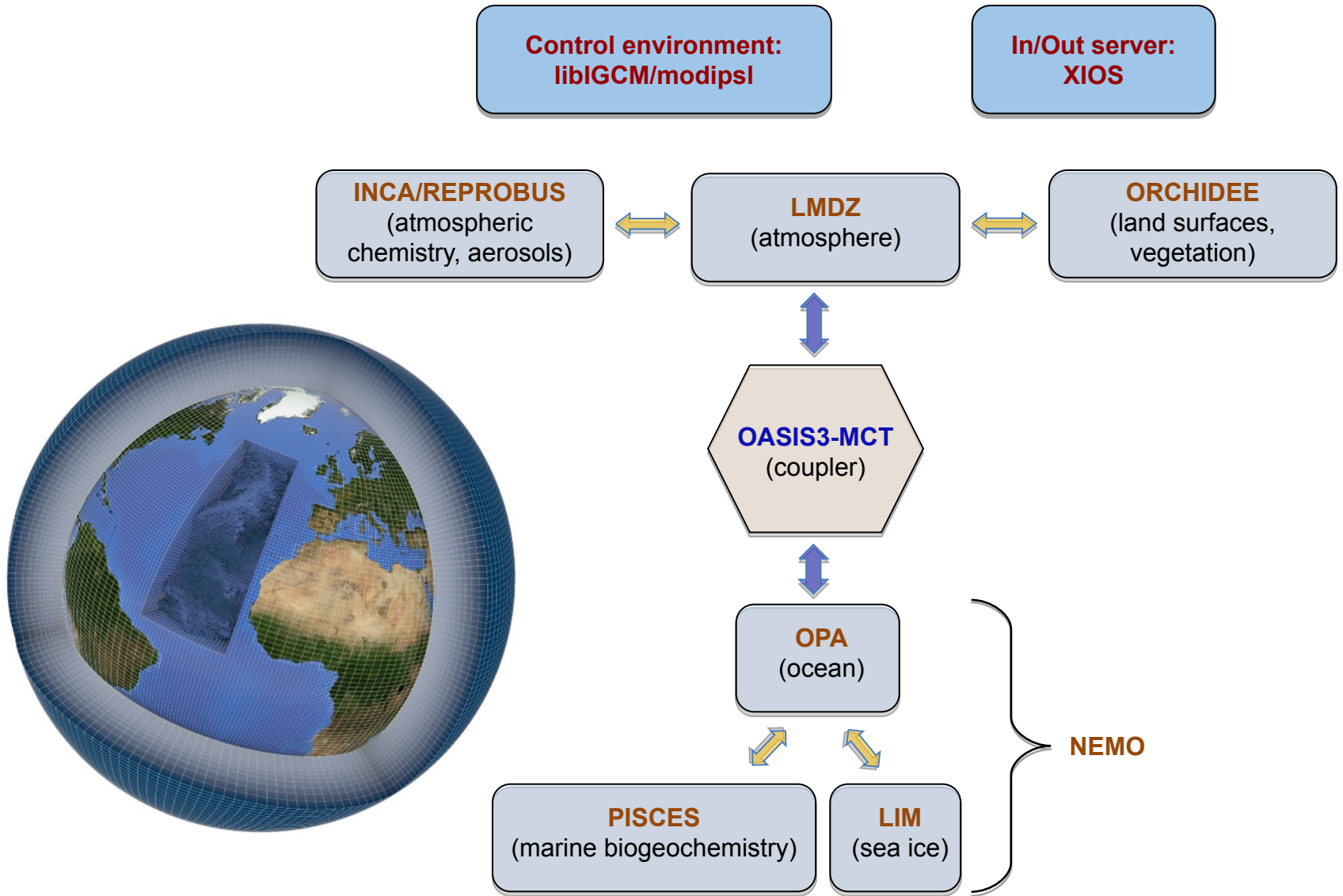**Sébastien Nguyen**

Guillaume Levavasseur

Patricia Cadule

Platform-group members

Arnaud Caubel	Coupled model / CMIP6 wokflow	**Josefine Ghattas**	Orchidee model
Anne Cozic	INCA model	**Nicolas Lebas**	Ensembles, StratAER (LMDZ)
Romain Pennel	Regional model	**Patricia Cadule**	Orchidee model
Christian Ethé	NEMO & PISCES model	**Lola Falletti**	REPROBUS model
Jérôme Servonnat	CliMAF / C-ESM-EP	**Olivier Marti**	weight on coupling
Sébastien Nguyen	paleoclimate model	**Olivier Boucher**	ICMC head
Laurent Fairhead	LMDZ model	**Thibaut Lurton**	CMIP6 coupled simulations
Guillaume Levavasseur	Thredds / Esgf	**Renaud Person**	NEMO & PISCES model

+   For all : libIGCM, modipsl, supercomputers, and lot of things

1. Introduction
2. **IPSL models**
3. High Performance Computing context
4. Which supercomputer(s) for us ?
5. Tools, configurations and performances
6. To go further
7. Now for today

Institut
Pierre-Simon
Laplace

Control environment:
libIGCM/modipsl

In/Out server:
XIOS

INCA/REPROBUS
(atmospheric chemistry, aerosols)

LMDZ
(atmosphere)

ORCHIDEE
(land surfaces, vegetation)

OASIS3-MCT
(coupler)

OPA
(ocean)

NEMO

PISCES
(marine biogeochemistry)

LIM
(sea ice)

Institut
Pierre-**S**imon
Laplace

https://www.nemo-ocean.eu

NEMO (**N**ucleus for **E**uropean **M**odelling of the **O**cean) is a state-of-the-art modelling framework for research activities and forecasting services in ocean and climate sciences, developed in a sustainable way by a European consortium.

3 main components:

- **OPA**: models the ocean {thermo}dynamics and solves the primitive equations
- **LIM / SI3** : models sea-ice {thermo}dynamics, brine inclusions and subgrid-scale thickness variations
- **PISCES**: models the {on,off}line oceanic tracers transport and biogeochemical processes

NEMO_v6.1	*NEMO 3.6*
NEMO_v6.5	*NEMO 4*

http://lmdz.lmd.jussieu.fr

LMDZ (**L**aboratoire de **M**étéorologie **D**ynamique **Z**oom model) is a general circulation model (or global climate model) developed since the 70s at the LMD, which includes various variants for the Earth and other planets (Mars, Titan, Venus, Exoplanets). It is first and foremost a research tool.

2 dynamic cores:

- **Actual**: based on regular Lat x Lon grid. Easy to use but limited in terms of parallelization on actual machines.
- **DYNAMICO**: icosaedric grid that allows very high scalability on HPC machines (still in development).

LMDZOR_v6.2.2	*ATM 144x144x79*
ICOLMDZOR_v7.2	
IPSLCM6.2.2	*ATM 144x144x79 / OCE ORCA1*
	ATM 256x256x79 / OCE ORCA1

https://orchidee.ipsl.fr

ORCHIDEE (**O**rganising **C**arbon and **H**ydrology **I**n **D**ynamic **E**cosystems) represents the state of the art in global land surface modelling. It solves the water-energy-carbon budget, represents the ecosystem in terms of a range of Plant Functional Types and vegetation with a big leaf approach. It uses precipitation, air temperature, wind, solar radiation, humidity and atmospheric CO_2 as forcing data and computes its own phenology.

2 major components:

- **Sechiba**: water and energy budgets
- **Stomate**: biogeochemical and anthropogenic processes

LMDZOR_v6.2.2	*ATM 144x144x79 / ORCHIDEE_2_2*
LMDZOR_v6.4_work	*ATM 144x144x79 / ORCHIDEE trunk*
ORCHIDEE_4_1	*ORCHIDEE offline latest tag*
ORCHIDEE_trunk	*ORCHIDEE offline with the trunk*

http://inca.lsce.ipsl.fr

INCA (**IN**teraction with **C**hemistry and **A**erosols) is a chemistry and aerosol model coupled to General Circulation Model, LMDz. LMDzINCA accounts for emissions, transport (resolved and sub-grid scale), photochemical transformations, and scavenging (dry deposition and washout) of chemical species and aerosols interactively in the GCM. INCA is often coupled to the ORCHIDEE biosphere model in order to determine interactively the exchange of chemical species (emissions, deposition) between the atmosphere and the surface.

LMDZORINCA_v6.2.2 *ATM 96x96x39 (AP) or 144x144x79 (NP)*

REPROBUS model (**RE**active **P**rocesses **R**uling the **O**zone **BU**dget in the **S**tratosphere) coupled with the general circulation atmosphere model LMDz is a 3-D model designs to solve the dynamic and chemistry in the stratosphere in order to study ozone layer and its interactions with climate.
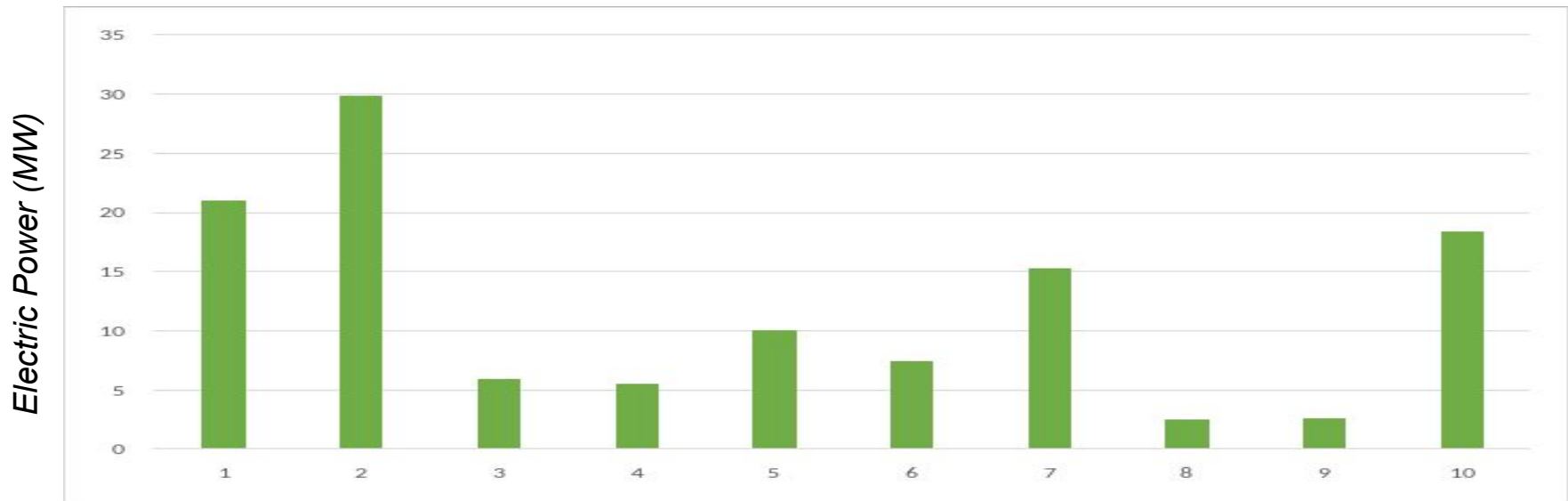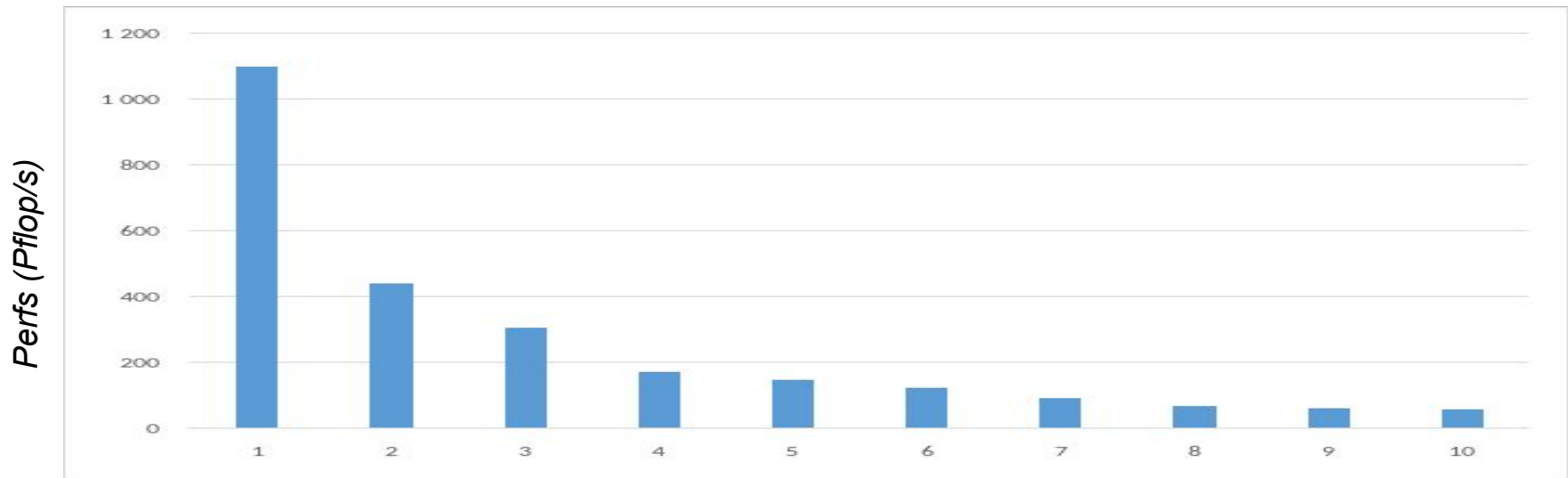
LMDZREPR_v6.2 *ATM 144x144x79*

1. Introduction
2. IPSL models
3. **High Performance Computing context**
4. Which supercomputer(s) for us ?
5. Tools, configurations and performances
6. To go further
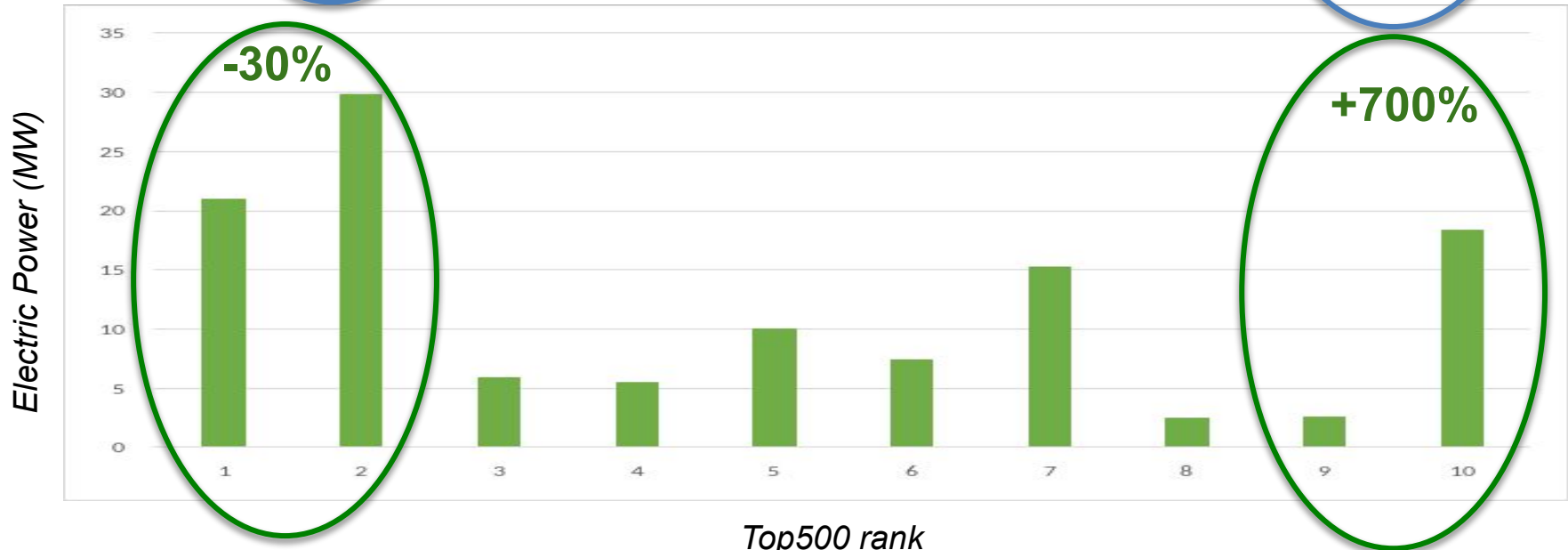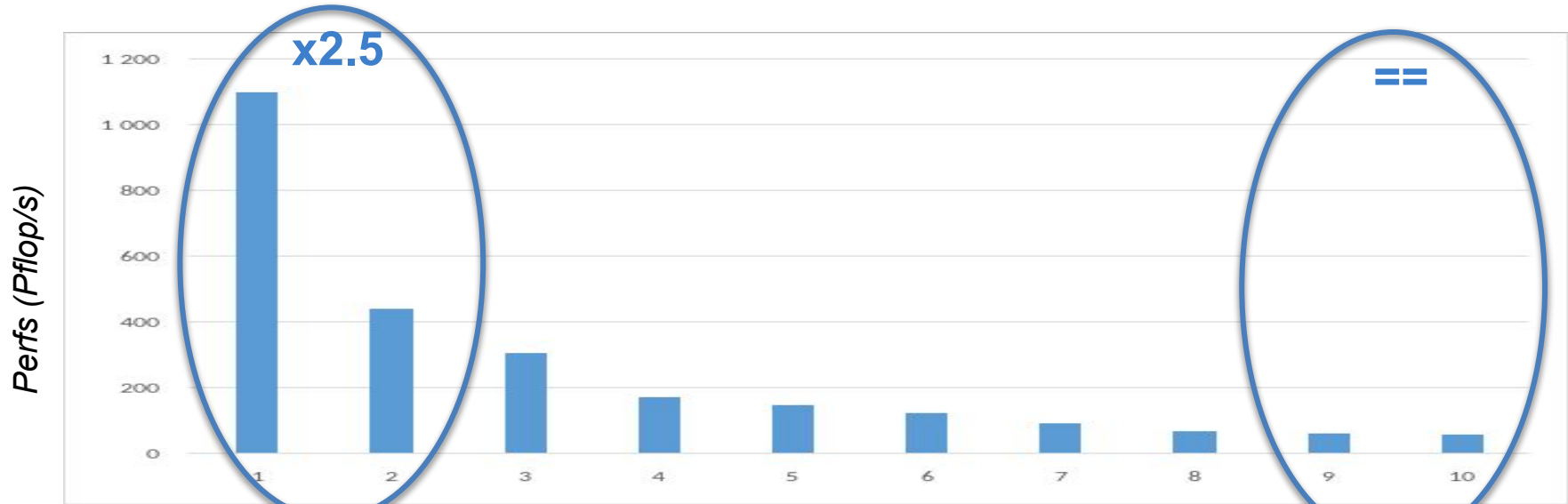7. Now for today

Projected Performance Development
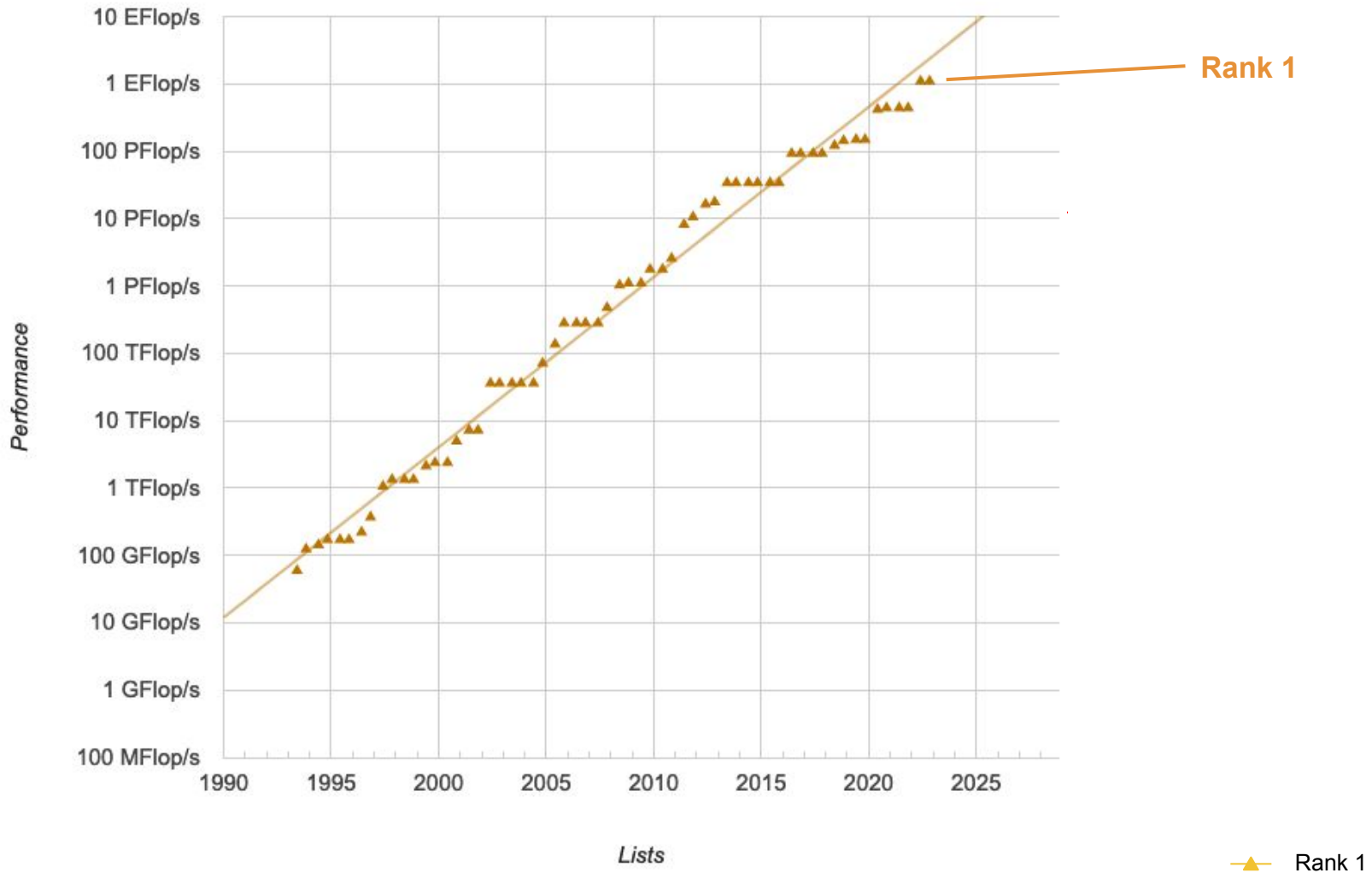
Supercomputer - Power efficiency

Top500 rank

Projected Performance Development
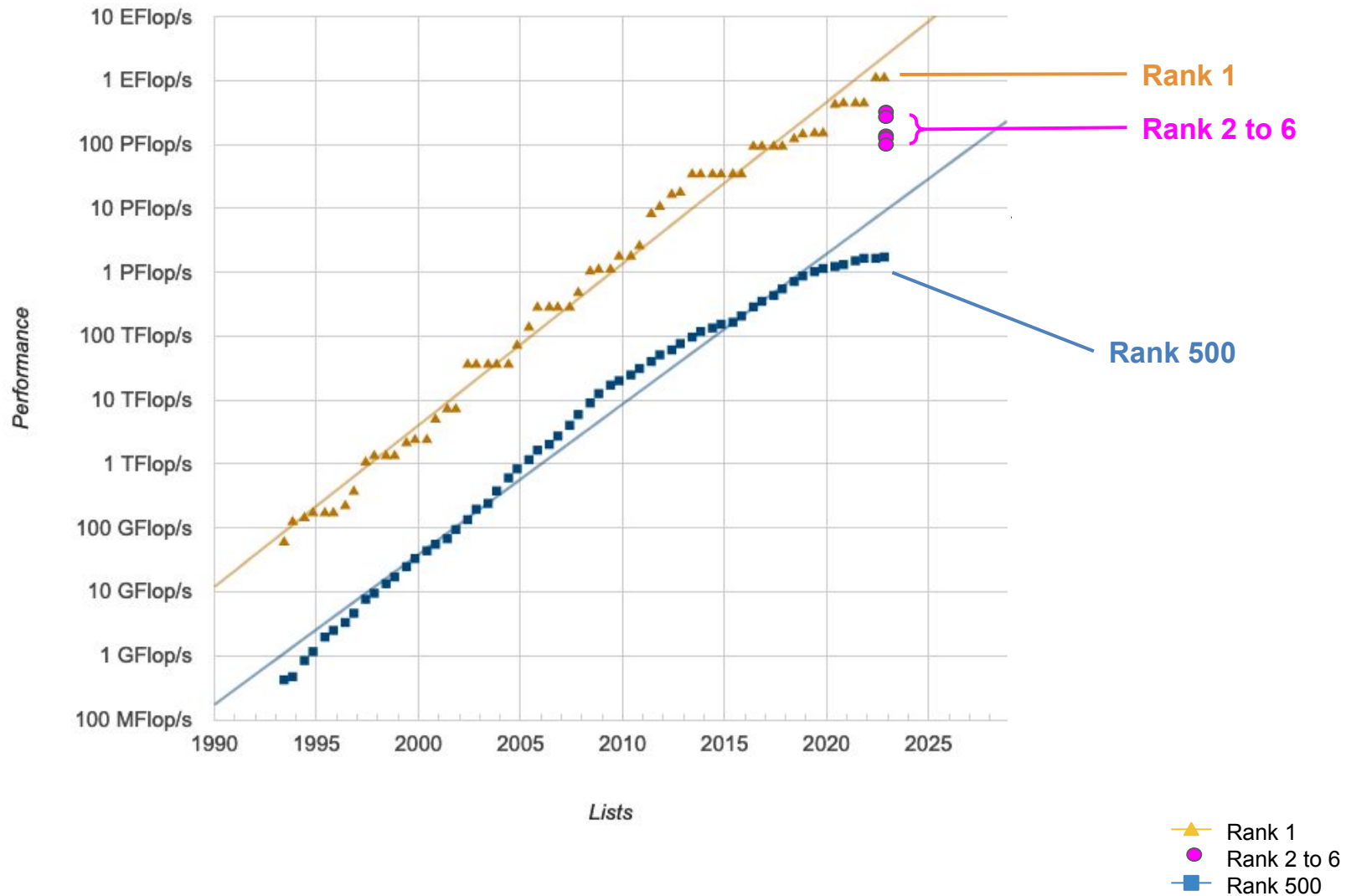
Supercomputer - *green500* ranking

Rank	#Top 500	System	Cores	Rmax (PFlop/s)	Power (MW)	Power Efficiency (GFlops/W)
1	**405**	**Henri** Lenovo - USA	**5 920**	**2.04**	**0.031**	**65.10**
2	32	**Frontier TDS** Cray - USA	120 832	19.20	0.31	62.68
3	11	**Adastra** Cray - France	319 072	46.10	0.92	58.02
4	15	**Setonix GPU** Cray - Australia	181 248	27.16	0.48	56.98
5	68	**Dardel GPU** Cray - Sweden	52 864	8.26	0.15	56.50
6	1	**Frontier** Cray - USA	8 730 112	1 102.00	21.10	52.23
7	3	**LUMI** Cray - Finland	2 220 288	309.10	6.02	51.38

Projected Performance Development
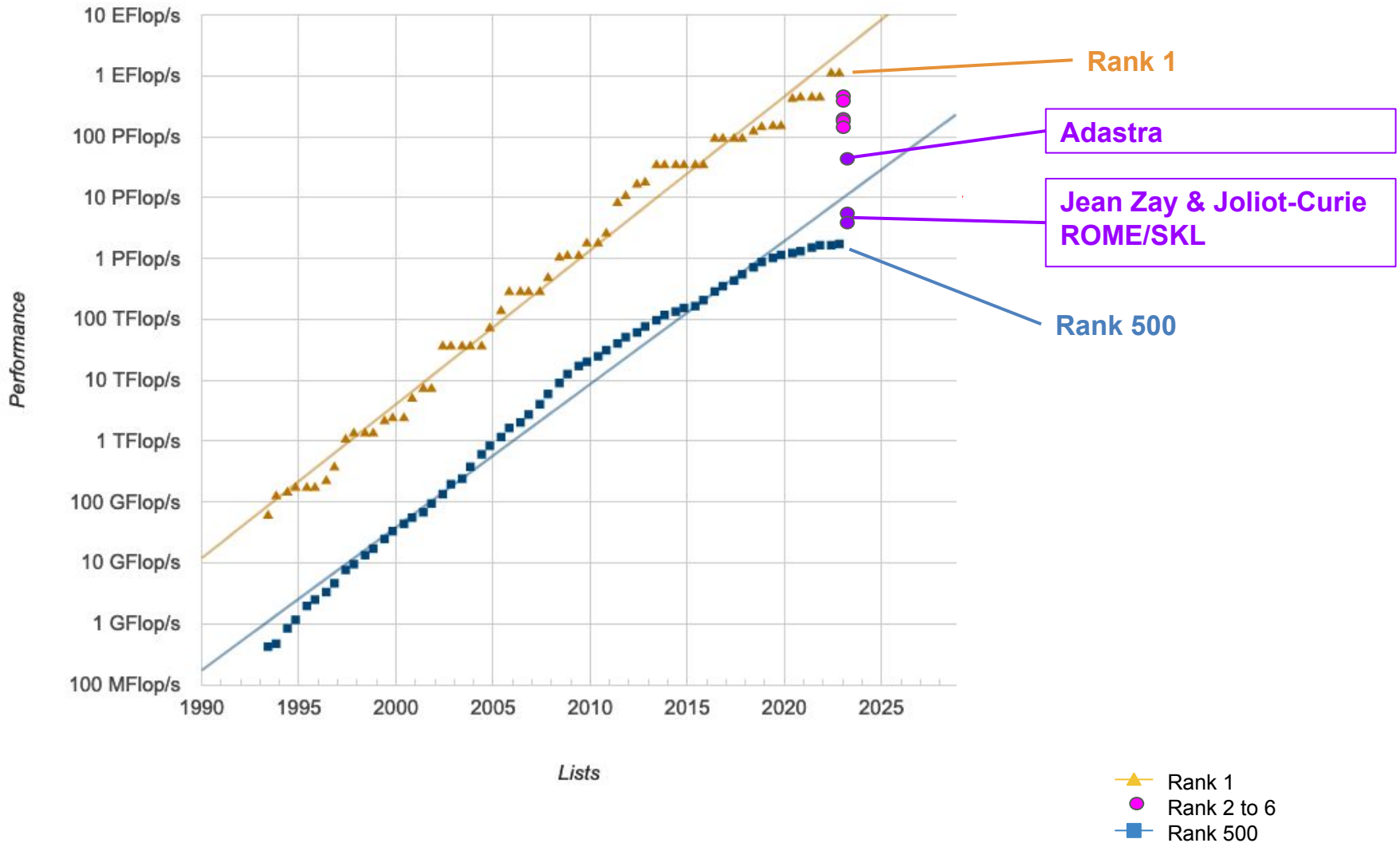
Why do we need supercomputer ?
⇒ *parallelization*!

All models are parallelized with MPI or MPI+OpenMP.

Parallelization allows to run the same executable on several sub-domains to reduce the real time of the execution.

MPI (**M**essage **P**assing **I**nterface) is used to divide **lon/lat grid** splitting latitudes and **OpenMP** (**Open M**ulti-**P**rocessing) to parallelize the **vertical axis** through shared memory threads.

→ The global domain is divided into sub-domains, each core treats one sub-domain
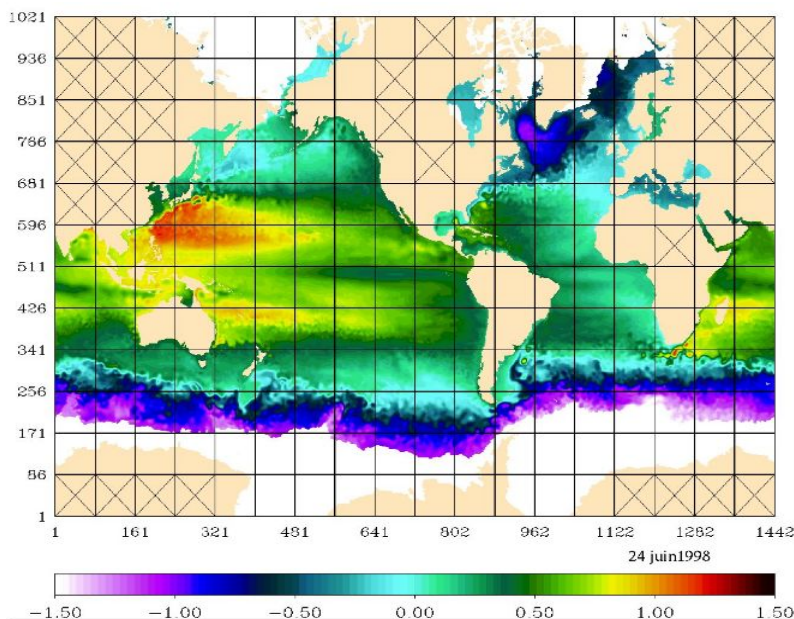
NEMO model parallelism (MPI only)
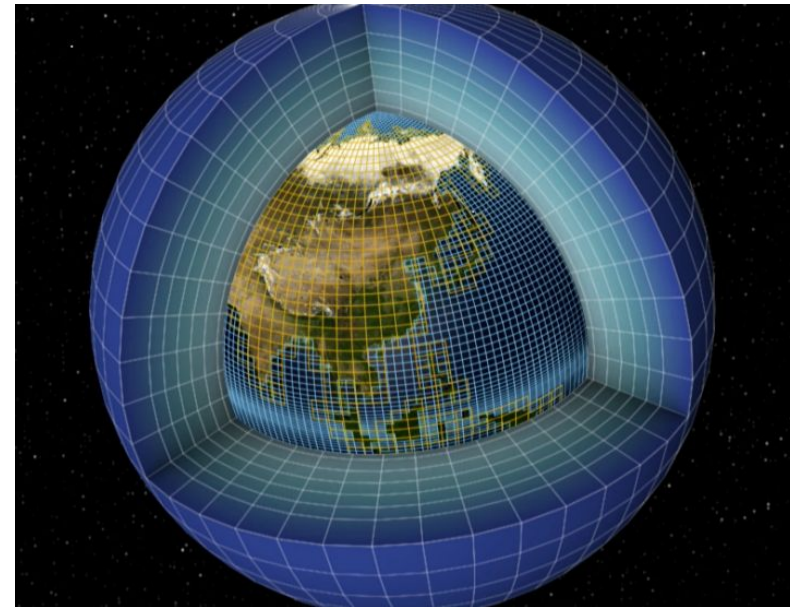
LMDZ model uses hybrid MPI/Open MP parallelization

Institut Pierre-Simon Laplace

Environment:

- A supercomputer is not a personal computer! Each supercomputer is <u>unique</u> and requires a dedicated staff to maintain its hardware and software.

- Its usage requires good skills to understand how to work with it properly.

- This is a very complex system that could implies an increase of instabilities on filesystem, computation nodes, high speed network...

- All resources are **SHARED** between all users (CPU hours, storage, bandwidth). *You need to adopt good practices to avoid to perturb other people. For example, you can be forced to clean your space very quickly if you didn't realize you used too much storage.*

- Computation centers have a **high level security policy**, so you cannot connect to them from everywhere and you need to respect rules.

CPU hours on Tier-1 (national) and Tier-0 (European) centers:

- Computing hours are attributed through bi-annual GENCI or PRACE calls (technical and scientific goals, roadmap, code efficiency and evaluation)

- Dynamical access (all year long) for project with <500 kh CPU or 50 kh GPU

Institut
Pierre-**S**imon
Laplace

What does HPC usage imply for you? (2/2)

General rules and advises:

- <u>Quota</u>: be careful with it! => Computing hours, storage, inodes (=number of file-system object such as file or directory)

- <u>Jobs</u>: priority algorithm between jobs (depending on the resources you request), max number of jobs running at the same time

- Use your computing hours regularly (to avoid peak usage of the machine)

- <u>Security</u>: never share your password!

If rules are not respected, computation centers could:

- remove amount of hours of the project
- block project jobs
- suspend account
- block filesystems (inodes or storage quota)

Where can you use IPSL tools?
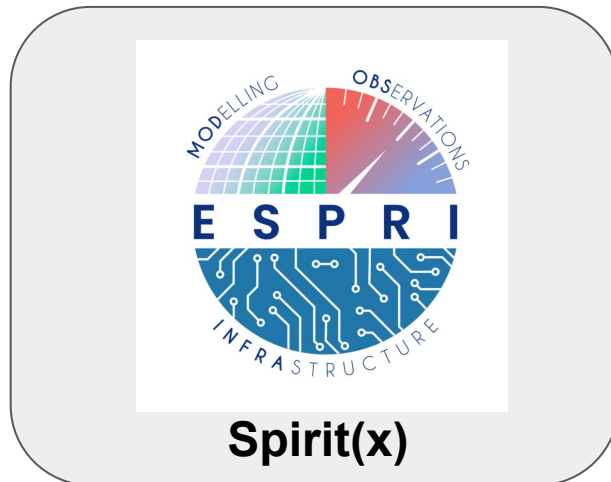
National centers

Jean Zay

**Irene Joliot-Curie
(ROME and SKL)**

Meso-center

Spirit(x)

Local cluster

Obelix

- <u>Filesystems</u>:

  – *HOME : small space, back up*

  – *WORK : working space, no back up, no purge*

  – *STORE : for archive, no back up*

  – *SCRATCH : big working space, is purged after 30 days, not saved*

We advise you to copy the **IPSL platform environment** in the `HOME` of your account and **install** models into your project `WORK`.

- <u>Partitions</u>:

  – Jean Zay CPU: scalar computing nodes for launching simulations

  – Jean Zay pre-post-processing: large memory nodes for pre and post processing tasks (such as compilation)

- <u>Documentation</u>:

  – https://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/IDRIS

  – https://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/IDRIS/Jean Zay

  – http://www.idris.fr (useful cheat sheet : http://www.idris.fr/jean-zay/cheat-sheet.html)

- <u>Assistance</u>: 01 69 35 85 55, assist@idris.fr

- <u>Connexion</u>:

  – ssh -X login@jean-zay.idris.fr   *(JeanZay)*

  – ssh -X login@jean-zay-pp.idris.fr   *(JeanZayPP)*

- The password is the same on *jeanzay* and *jeanzaypp*. Use *passwd* on one of the machines to change it.

- Quota for the whole group. Use *idrquota -s* and *idrquota -w* to check for $STORE and for $WORK.

Quota are attributed for each project for all the group and not individually, so be careful of your own practices to avoid blocking all the group

- <u>Filesystems:</u>

  – *HOME : small space, back up*

  – *WORKDIR : working space and archiving of small files – quota 1Tb, no back up, no purge*

  – *STOREDIR : only for archive of big files – min 10Gb – quota 100 000 inodes, on tape*

  – *SCRATCHDIR : big working space, may be purged after 40 days*

We advise you to copy the **IPSL platform environment** in the `HOME` of your account and **install** models into your project `WORKDIR`.

- <u>Partitions</u>:

  – Irene Joliot-Curie SKL: AMD Rome (Epyc) bi-processors nodes

  – Irene Joliot-Curie Rome: Intel Skylake 8168 bi-processors nodes

  – xlarge nodes: large memory nodes for pre and post processing tasks

- <u>Documentation</u>:

  – https://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/TGCC

  – http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/TGCC/Irene

  – http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/TGCC/IreneAmd

  – Command on irene : irene.*info*

  – https://www-tgcc.ccc.cea.fr (private access for user only) or
  http://www-hpc.cea.fr/docs/userdoc-tgcc-public.pdf (public access)

- <u>Assistance</u>: 01 77 57 42 42, hotline.tgcc@cea.fr

- <u>Connexion</u>:

  – ssh -X login@irene-fr.ccc.cea.fr (SKL)

  or ssh -X login@irene-amd-fr.ccc.cea.fr (ROME)

  – for group quota, use *ccc_quota -g genXXXX*

  – for personal quota, only use *ccc_quota* to check

Quota are attributed for each project for all the group and not individually, so be careful of your own practices to avoid blocking all the group

- Modipsl and libIGCM are also adapted to be used at
  - *Obelix* – LSCE cluster (http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/LSCE)
  - *spirit* and *spiritx* – IPSL clusters (ESPRI mesocenter) (http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/ComputingCenters/ESPRImesocenter)
- Following functionalities are adapted
  - Compilation
  - Computing job
  - Rebuild
  - TS-SE
- Not adapted : *pack, monitoring and full*

1. Introduction
2. IPSL models
3. High Performance Computing context
4. Which supercomputer(s) for us ?
5. **Tools, configurations and performances**
6. To go further
7. Now for today

Software infrastructure based mainly on **modipsl**, **libIGCM** and **XIOS** tools which allow to :

modipsl

- **predefine** and **extract** standard configurations and **compilation** scripts

libIGCM

- **adapt** and **launch** predefined experiments
- **monitor** simulations
- **produce** and **store** results from models
- **produce**, **store** and **distribute** some analysis

XIOS

- **read** input files
- **write** and **interpolate** results from models in parallel

Tools available for usage at TGCC, IDRIS, LSCE and IPSL cluster.

Web documentations & mailing list

- **Modipsl / libIGCM**: http://forge.ipsl.jussieu.fr/igcmg_doc

- **Platform-users**: https://listes.ipsl.fr/sympa/info/platform-users

Community list for communication between all IPSL tools users. Anyone can ask questions and answer his/her colleagues questions.

→ All users need to subscribe

A configuration is a combination of one or several models (components) coupled together

- For example the configuration LMDZOR contains the two models LMDZ and ORCHIDEE.

A configuration can be used for different experiments, using different setups, input parameters, etc.

- For example with the configuration LMDZOR you can run experiments with different parameterizations for the physics in the atmosphere.

- For example with the configuration LMDZOR you can run an experiment with only LMDZ

Institut Pierre-Simon Laplace

Main configuration
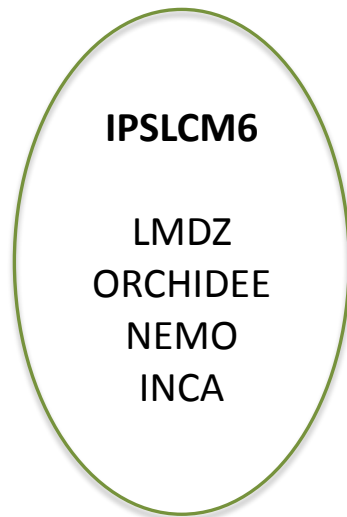
IPSLCM6 sub-configurations

Experiments

IPSLCM6

LMDZ
ORCHIDEE
NEMO
INCA

IPSLCM

IPSLESM

CO_2

AER

LMDZOR

LMDZ

LMDZORINCA

Resolutions

VLR

LR

MR1

MR025

pdControl

piControl

amip

clim

AER

NMHC
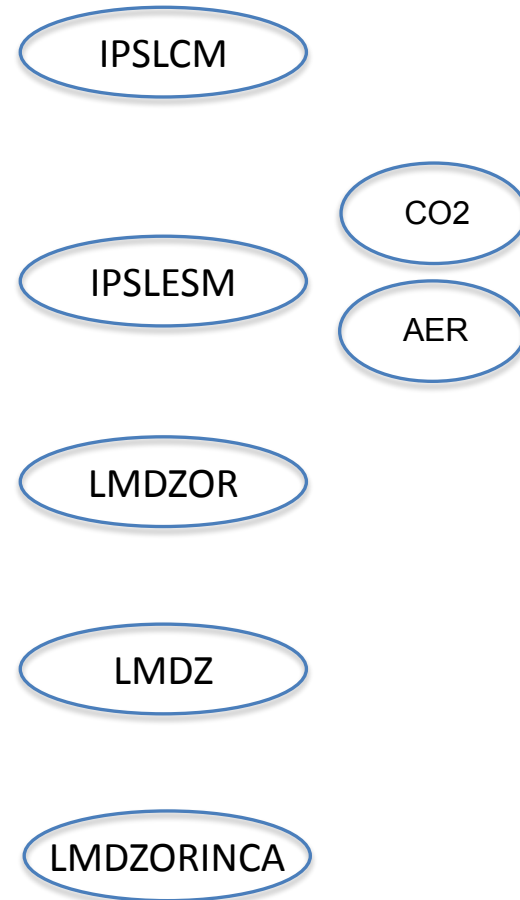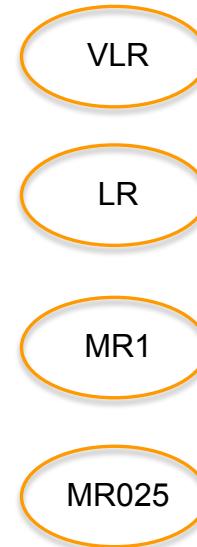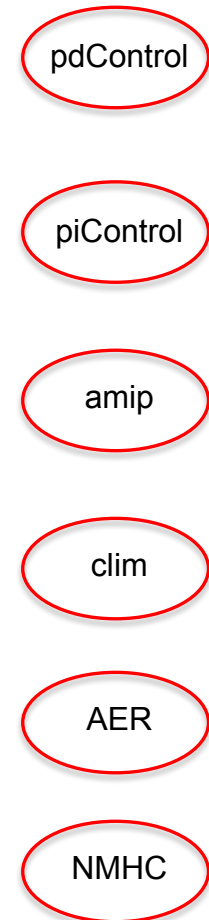
1 main configuration = 1 executable per resolution = several experiments (from sub-configurations)

Actual configs : Recommended version of standard configurations. Parameters set up is the same for a component in all configurations of the "v6 family".

There are 3 types of v6 configurations:

1. v6.1 / v6.1.11   ⇒ versions used for CMIP6
2. v6.2 / v6.2_work   ⇒ versions under development to improve CMIP6 version
3. v6.3 - 4 - 5 ..   ⇒ versions to prepare next generation of configurations
4. *v7*   ⇒ next configurations (DYNAMICO and NEMO4)

Main configurations proposed in this family are declined in v6.1 and several other types.

IPSLCM6

Version of the coupled model used for CMIP6

(currently IPSLCM6.1.11-LR)

Person in charge: A. Caubel

NEMO_v6

Forced ocean model OPA-LIM3-PISCES
Person in charge: C. Ethé

LMDZOR_v6

LMDZ coupled with ORCHIDEE
Person in charge: J. Ghattas

LMDZORINCA_v6

LMDZOR_v6 coupled with INCA
Person in charge: A. Cozic

LMDZREPR_v6

LMDZ_v6 coupled with REPROBUS
Person in charge: L. Falletti

ICOLMDZOR (and ICOLMDZORINCA) DYNAMICO as atmospheric dynamics coupled with LMDZ physics and ORCHIDEE (and INCA)
Persons in charge: A. Caubel, A. Cozic, J.Ghattas

Institut
Pierre-Simon
Laplace

IPSLCM5A2   Previous version of the coupled model (*IPSLCM5*) used on a very low resolution (VLR) grid.
Person in charge : S. Nguyen

IPSLCM5A2-CHT   coupled model in low resolution with an interactive atmospheric chemistry . Person in charge : A. Cozic

ORCHIDEE_trunk/   Forced continental surfaces model ORCHIDEE,
ORCHIDEE_4_1   with latest version on the trunk of ORCHIDEE or tag 4_1. Person in charge: J. Ghattas.

RegIPSL   Regional coupled climate model of IPSL.
Person in charge: R. Pennel.

General recommendation :

• *inform person in charge* before launching new studies based on one of these configurations, especially for coupled models.

• Read model and configuration documentation before using it!!!

Institut
Pierre-**S**imon
Laplace

Configuration	Number of Cores	Simulated Year Per Day
IPSL-CM6.2-MR1 *ATM: 256x256x79 / OCE: eORCA1*	**1200**	**8.8**
IPSL-CM6.1.11-LR *ATM: 144x144x79 / OCE: eORCA1*	**976**	**16**
IPSL-CM5A2-VLR *ATM: 96x96x39 / OCE: ORCA2*	**437**	**95**
NEMO *eORCA1-LIM3-PISCES*	**433**	**20**
LMDZOR_v6.1.10-LR *LMDZ144x144x79*	**576**	**20**
ICOLMDZOR *NBP 40 (~200 km)*	***680***	***40***

Benchmark in January 2021

Institut
Pierre-Simon
Laplace

Configuration	Number of Cores	Simulated Year Per Day
IPSL-CM6.2-MR1 *ATM: 256x256x79 / OCE: eORCA1*	1196	8
IPSL-CM6.1.11-LR *ATM: 144x144x79 / OCE: eORCA1*	1952	24
IPSL-CM5A2-VLR *ATM: 96x96x39 / OCE: ORCA2*	604	97
NEMO *eORCA1-LIM3-PISCES*	640	23.5
LMDZOR_v6.1.10-LR *LMDZ144x144x79*	1136	25

Benchmark in January 2021

IPSL-CM perfs: Jean Zay

Configuration	Number of Cores	Simulated Year Per Day
IPSL-CM6.2-MR1 *ATM: 256x256x79 / OCE: eORCA1*	-	-
IPSL-CM6.1.11-LR *ATM: 144x144x79 / OCE: eORCA1*	1071	24
IPSL-CM5A2-VLR *ATM: 96x96x39 / OCE: ORCA2*	399	93
NEMO *eORCA1-LIM3-PISCES*	428	40
LMDZOR_v6.1.11-LR *ATM: 144x144x79*	711	23

Benchmark in January 2021

1. Introduction
2. IPSL models
3. High Performance Computing context
4. Which supercomputer(s) for us ?
5. Tools, configurations and performances
6. **To go further**
7. Now for today

- **New atmospheric dynamical core DYNAMICO**

  - better computing performances/scalability

  - new architectures (GPUs, MPPs, MICs)

Lon-lat grid → *Icosahedral grid*
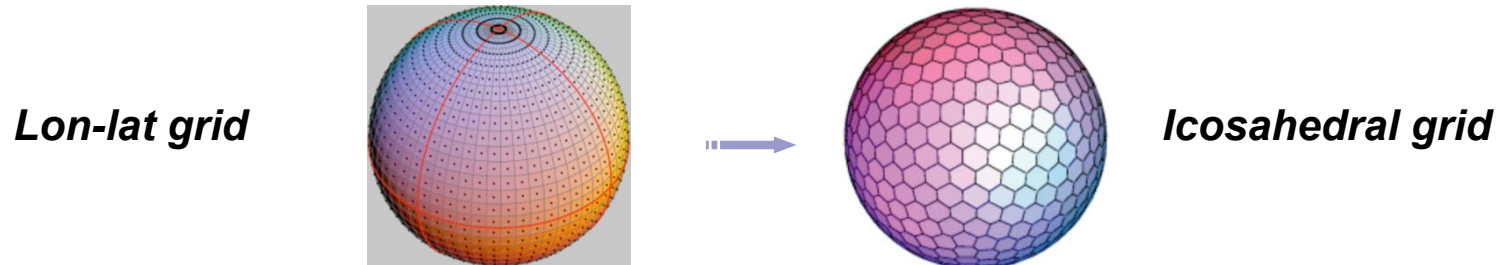
- **NEMO v4** and **SI3** (new sea-ice model), final phase of development

- **CMIP6 workflow**: to integrate the CMIP6 specific workflow (outputs) for "usual runs", available and already used for specific cases

- **StratAer:** config with stratospherical microphysics

- **Ongoing developments**

  - **XIOS 3.0** (XIOS multithreaded (OpenMP) to target « many cores » architectures, coupling functionalities)

  - Ensemble runs (specific I/O design)

  - Zoomed configuration

Training courses at IPSL:
- *IPSL climate modeling: training in simulation tools* (contact platform-training-team@listes.ipsl.fr), current session
- *LMDZ training course* (contact Laurent.Fairhead@lmd.jussieu.fr), last session in January 2023
- *ORCHIDEE Introduction* 2-days course (contact orchidee-help@ipsl.jussieu.fr), next session in February 2023
- XIOS training course (contact xios-team@forge.ipsl.jussieu.fr), past sessions in 2021
- PISCES training course (contact formation_pisces@locean.ipsl.fr), past session in October 2022

Other suggested training courses:
- *MPI, OpenMP and Hybrid MPI/OpenMP* at IDRIS twice a year www.idris.fr
- Training course for using the computer centres (not available actually, usually at TGCC)
- UNIX course
- https://calcul.math.cnrs.fr/category/formation.html→ give an inventory of training course (numeric – calcul – hpc) in France

1. Introduction
2. IPSL models
3. High Performance Computing context
4. Which supercomputer(s) for us ?
5. Tools, configurations and performances
6. To go further
7. **Now for today**

Institut
Pierre-**S**imon
Laplace

Informations for today's training (1/2):

Today's on **JeanZay specific training accounts**:
- Connect on "`ipcours22.……`" session.
- Your login: use the `cforXXX` login/password provided to you *(complete the form with your name and signature)*
- Note that everything in these accounts will be erased after the training.

If you need to **switch between qwerty and azerty** you can use the command **`alt+shift`**.

Download **training_exercises_2023.pdf**,
training_exercises_answers_2023.pdf and
main_commands_summary_2023.pdf at:

http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/Training

Advices and informations for this training(2/2):

- **Not all exercises are meant to be done**: select topics based on your knowledge of modipsl/libIGCM and your needs.

BEGINNER	INTERMEDIATE	SPECIALIZED
0. Introduction 1. Check your quota 2. Installing and compiling 3. Basic simulations	4. Debug 5. Create time series 6. Monitoring and Inter-monitoring 7. How to REDO part of a simulation 8. Modify output using XIOS 9. Output files manipulations	10. Install and run NEMO-PISCES 11. Ensembles 12. Coupled model 13. ICOLMDZOR configuration

Beginner: to be done by everyone, in this order.
Intermediate and **specialized**: depending of your level and your needs, not meant to be done in this order.

- **Take your time to read everything in the doc!** All is explained.

Do not hesitate to ask questions! ☺