CentraleSupélec

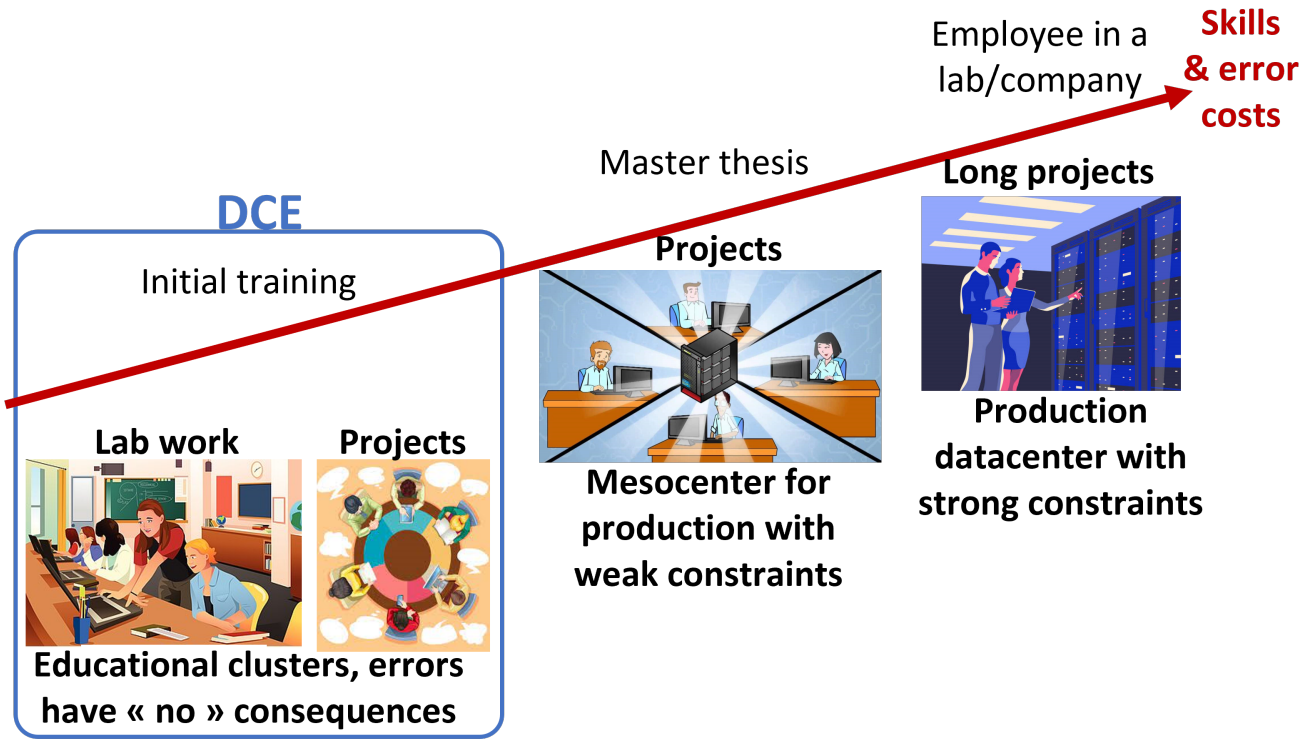
Feedback from a data center for education at CentraleSupélec engineering school

J. Fix, S. Vialle, R. Hellequin, C. Mercier, P. Mercier, J.-B. Tavernier

jeremy.fix@centralesupelec.fr

IPDPS - Edupar Workshop 2022

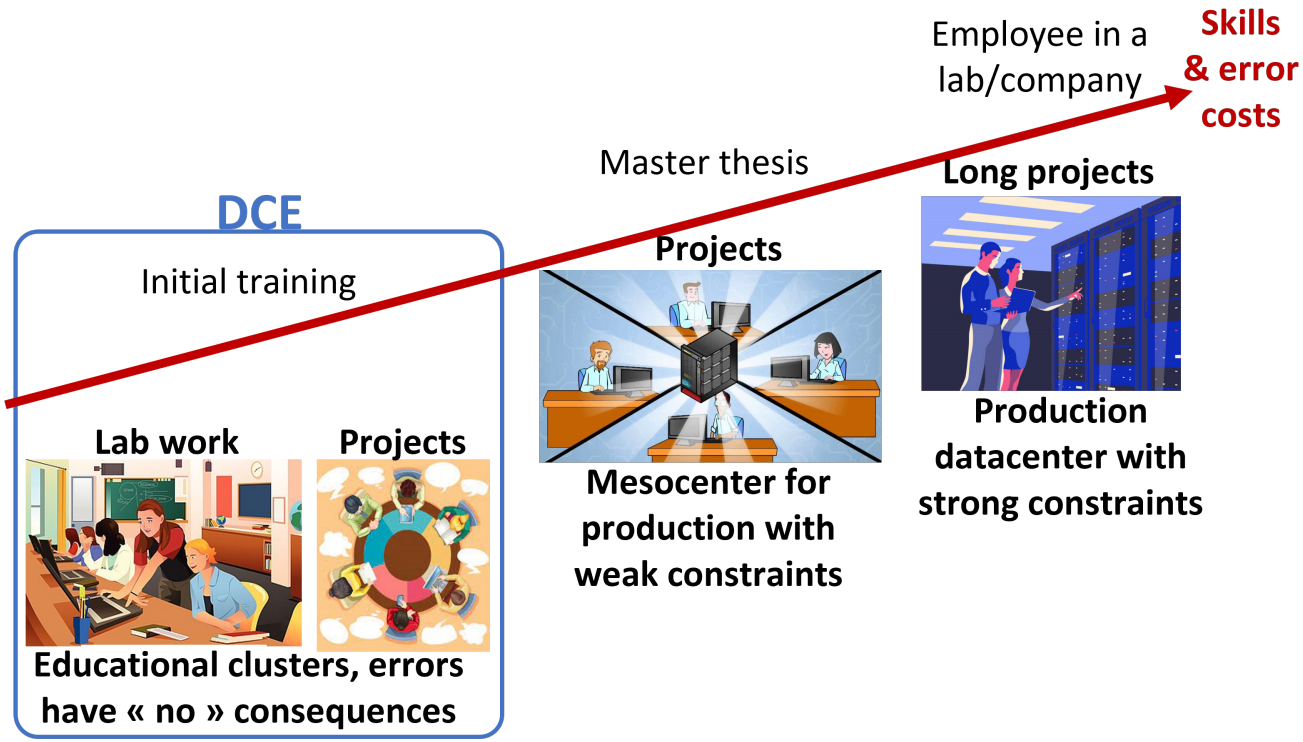
Motivations



For teaching, for beginners

- powerfull shared computing resources

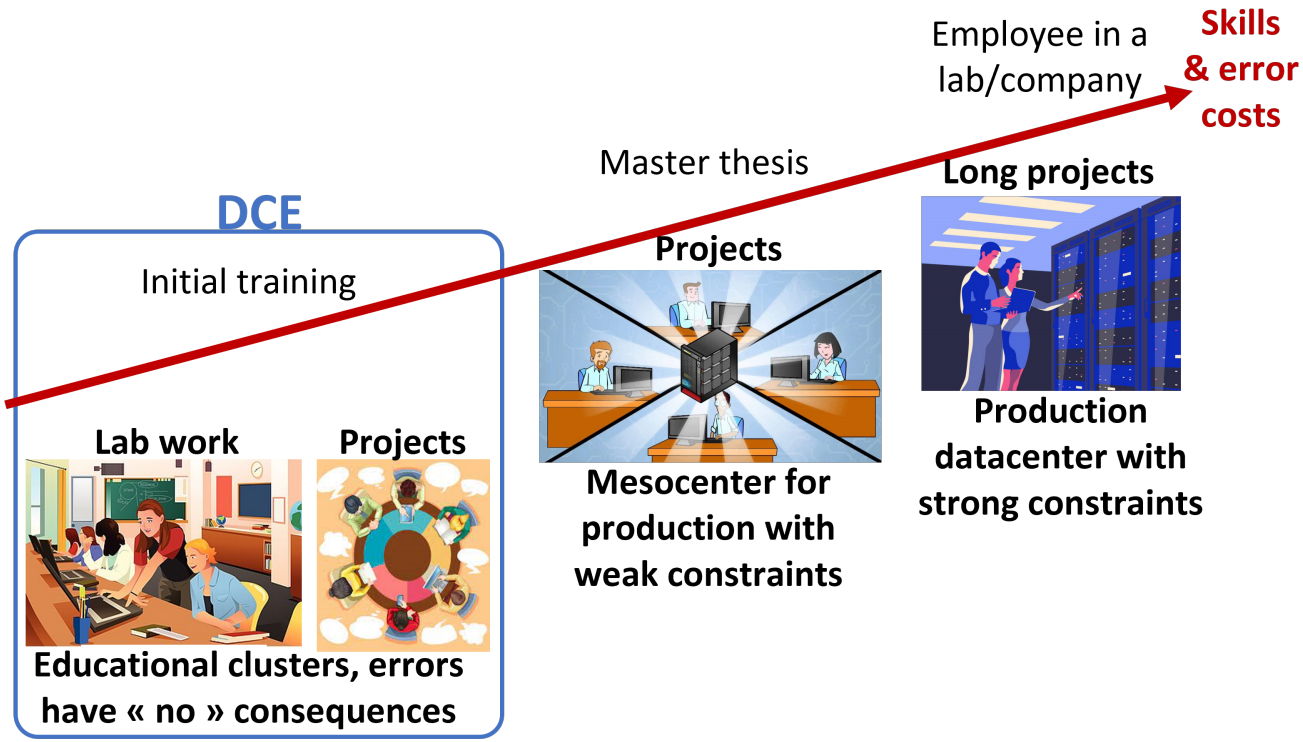
Motivations



For teaching, for beginners

- powerfull shared computing resources
- no bill to the end users

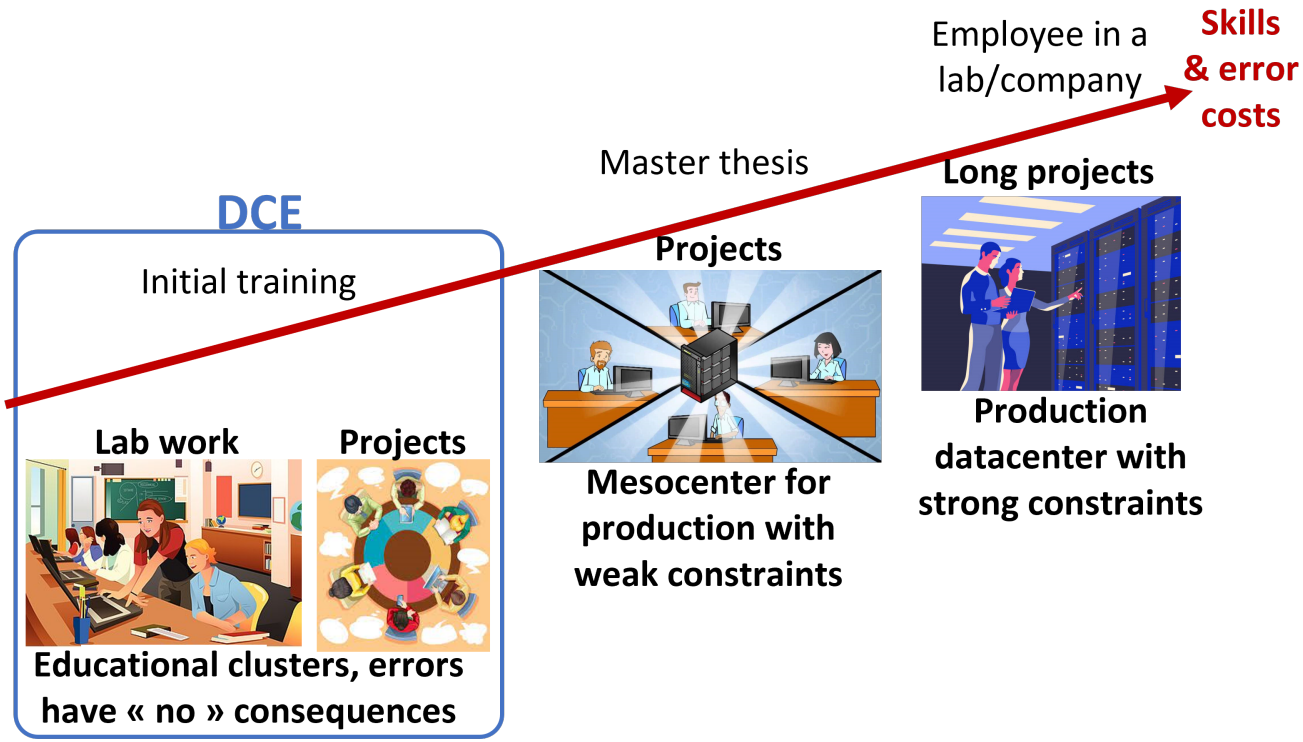
Motivations



For teaching, for beginners

- powerfull shared computing resources
- no bill to the end users
- errors without consequences

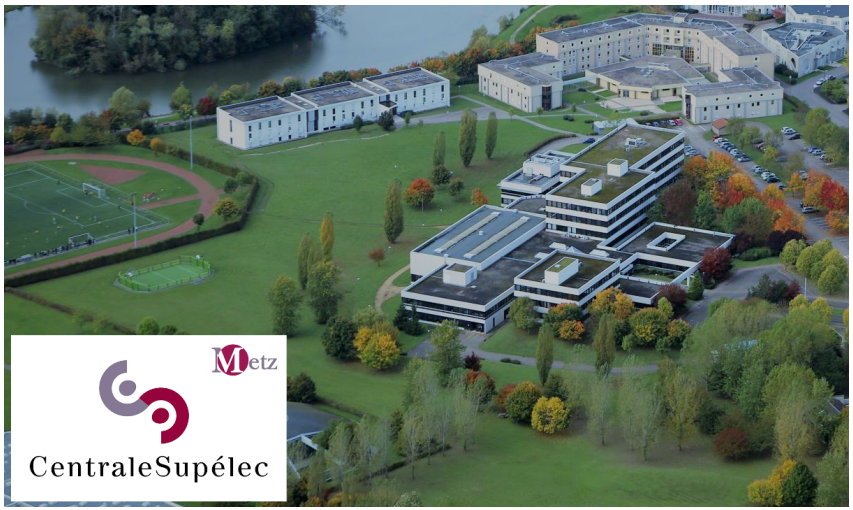
Motivations



For teaching, for beginners

- powerfull shared computing resources
- no bill to the end users
- errors without consequences
- access as simple and ergonomic as possible

Where is the Data Center for Education (DCE)



What is the Data Center for Education (DCE)

CPU clusters

- Kyle : 68 × 16 physical cores, 64 GB
- Sarah : 32 × 8 physical cores, 32 GB

Eth 10Gb/s

AVX, AVX2, AVX-512

What is the Data Center for Education (DCE)

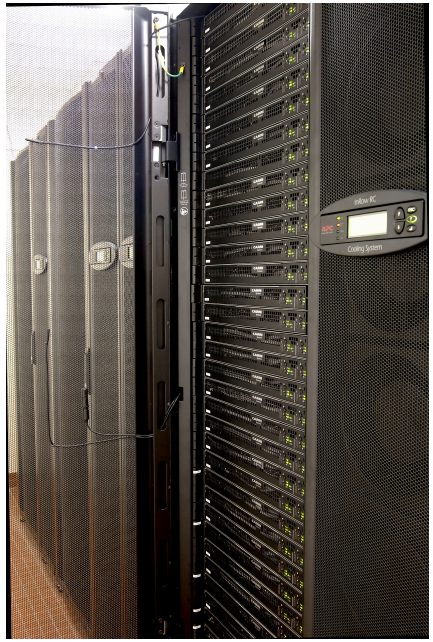
CPU clusters

- Kyle : 68 × 16 physical cores, 64 GB
- Sarah : 32 × 8 physical cores, 32 GB

Eth 10Gb/s
AVX, AVX2, AVX-512

GPU servers

- Cameron : 17 GPU servers (1080)
- Tx: 17 GPU servers (2080Ti)
- Sh: 17 GPU servers (3090)

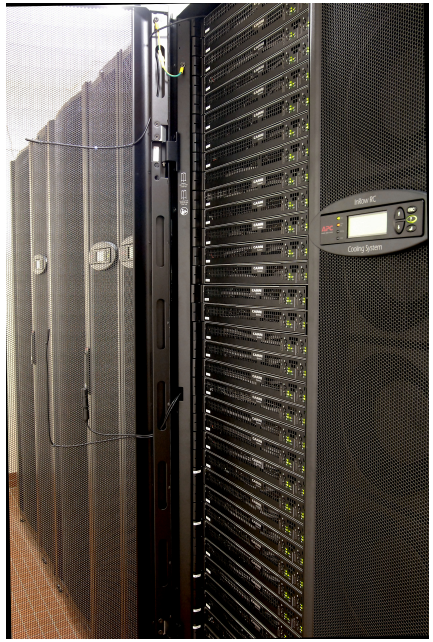


What is the Data Center for Education (DCE)

CPU clusters

- Kyle : 68 × 16 physical cores, 64 GB
- Sarah : 32 × 8 physical cores, 32 GB

Eth 10Gb/s
AVX, AVX2, AVX-512



GPU servers

- Cameron : 17 GPU servers (1080)
- Tx: 17 GPU servers (2080Ti)
- Sh: 17 GPU servers (3090)

Admin machines

1(+1)×Gateway, 1(+1) frontal SLURM, NFS server + 1 storage bay

Software environment

Administration

- ressources are scheduled with **slurm**
- the installations are synchronized with **ansible**
- the nodes are monitored with **Zabbix**
- labs are planed on website `www.dce-cs.fr`
- users access the nodes with home made software dcejs

Software environment

Administration

- ressources are scheduled with **slurm**
- the installations are synchronized with **ansible**
- the nodes are monitored with **Zabbix**
- labs are planed on website `www.dce-cs.fr`
- users access the nodes with home made software dcejs

Teaching

- Hadoop / Spark pre-installed distributed servers (on Sarah)
- Intel OneAPI, gcc, python, OpenMP, OpenMPI, OpenBLAS, CUDA, cublas, cudnn, ...
- Tensorflow, PyTorch, pandas, numpy, scikit-learn
- Robot Operating System (ROS) ecosystem

Custom envs. with conda (soon with module/singularity)

Costs and man months

Hardware costs

- Kyle (68 nodes): 160k€
- Tx (17 nodes + 2080Ti) : 50k€

Costs and man months

Hardware costs

- Kyle (68 nodes): $160k€$
- Tx (17 nodes + 2080Ti) : $50k€$

Energy costs

- CPUs clusters : $\approx 103MWh/y$
- GPUs servers : $\approx$~~80~~ $120MWh/y$ (Cam,Tx + Sh)

Total estimated : ~~182~~ $222MWh/y$ ($\approx 25K€$)
Not accounted : the free cooling with large fans

Costs and man months

Hardware costs

- Kyle (68 nodes): $160k€$
- Tx (17 nodes + 2080Ti) : $50k€$

Energy costs

- CPUs clusters : $\approx 103MWh/y$
- GPUs servers : $\approx$~~80~~ $120MWh/y$ (Cam,Tx + Sh)

Total estimated : ~~182~~ $222MWh/y$ ($\approx 25K€$)
Not accounted : the free cooling with large fans

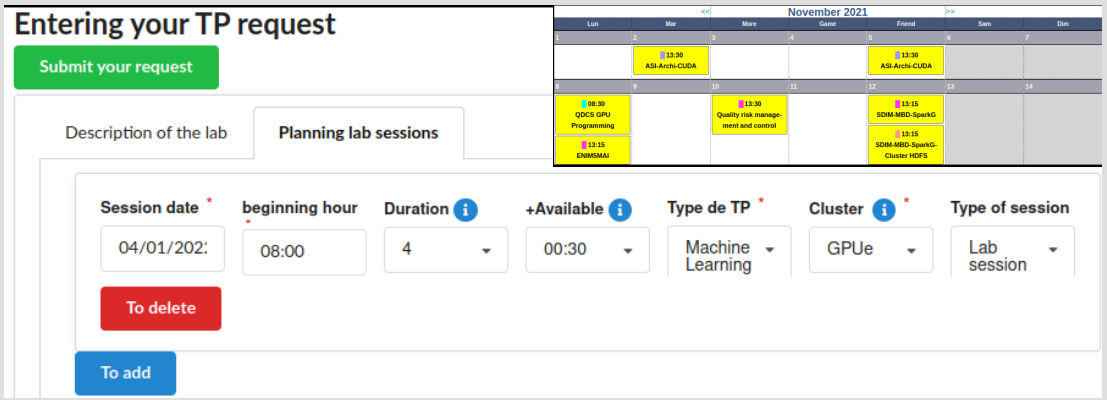
Humans

3 engineers (40%, 5%, 5%) and 2 professors (10%, 5%)
$\approx$ 7.8 (french-)man months/year

Using the DCE

Planning and accessing the DCE

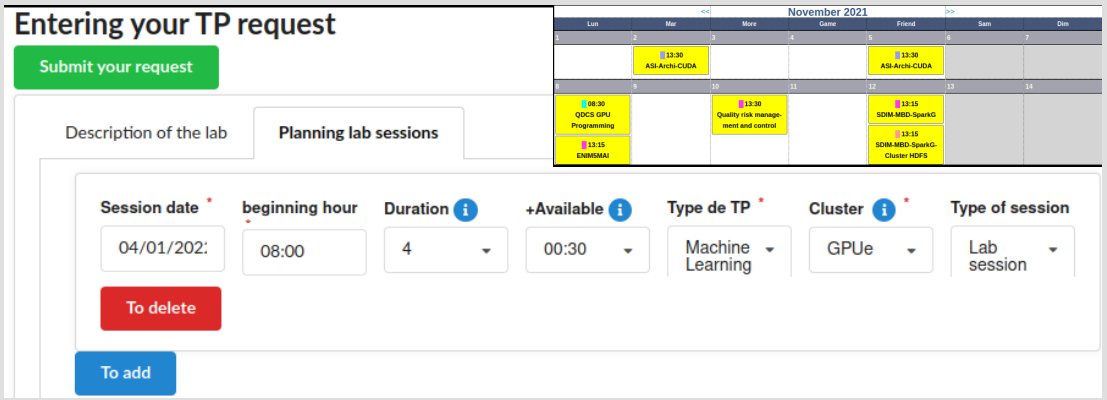
Registering a lab



Teacher requests are validated by the admin team.

Planning and accessing the DCE

Registering a lab



Teacher requests are validated by the admin team.

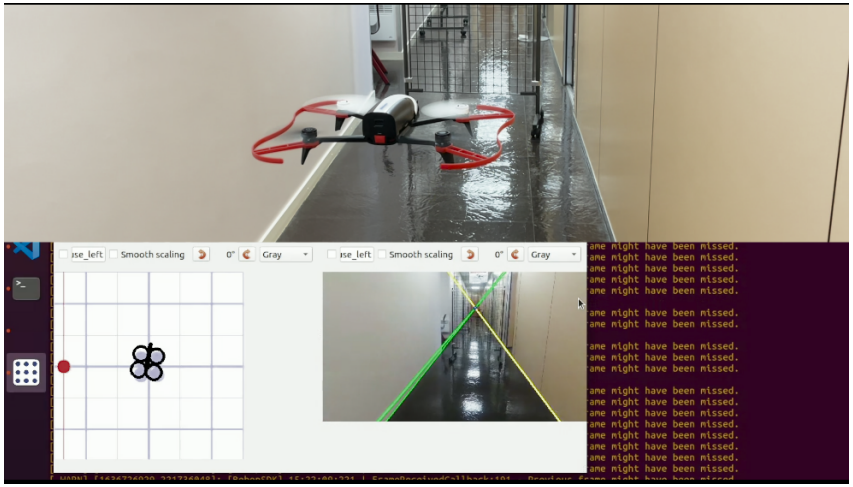
Co-hosting labs and projects

Partition	Max time	# nodes	Purpose
gpu_tp_resa	4h30	34	Scheduled lab works
gpu_tp	2h	34	Lab works live coding
gpu_prod_night	12h	16	Batch jobs, night& weekends
gpu_inter	2h	42	Project live coding
gpu_prod_long	48h	17	Long batch jobs

Description of the use case

The students experiment algorithms on robotic platforms:

- written in ROS with multiple interacting processes,
- computationnally intensive processes (MonteCarto / Particle filters for SLAM, planning and navigation, Gazebo simulator),
- requires an interactive session

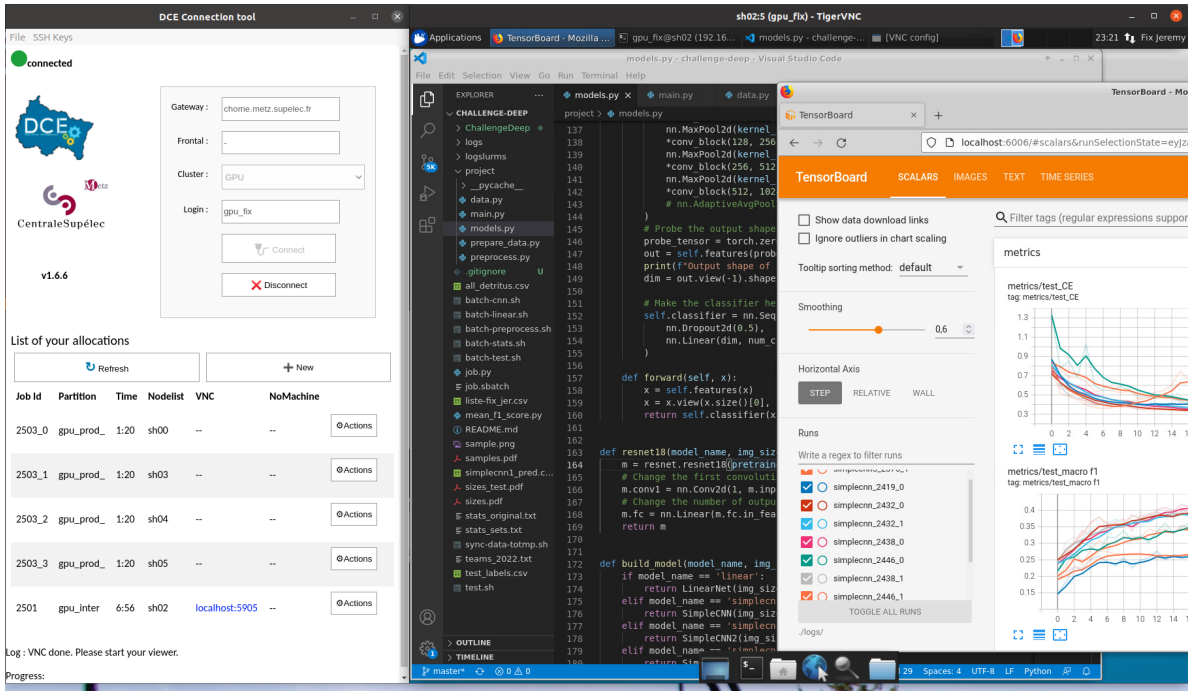


Also labs with on-site platforms during Covid-19.

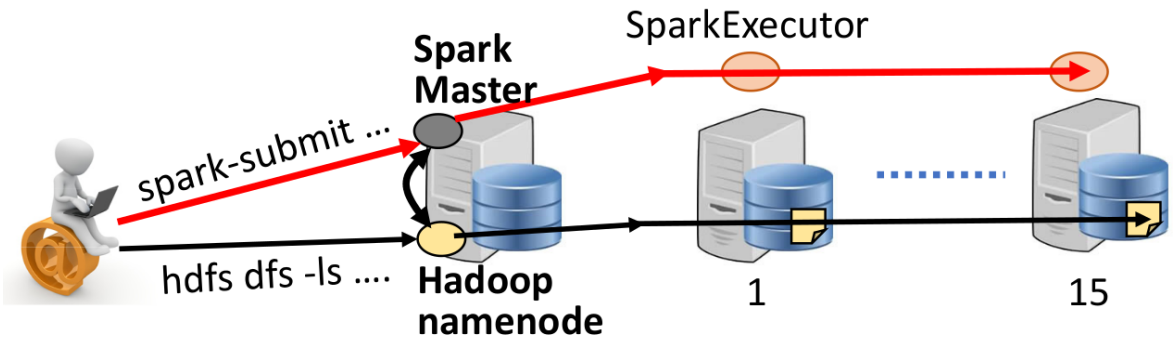
Description of the use case

The student train deep neural networks (pytorch/tensorflow):

- during ML labs, or projects/challenges (see Kaggle 3MD4040)
- both interactive and batch mode (long running jobs)

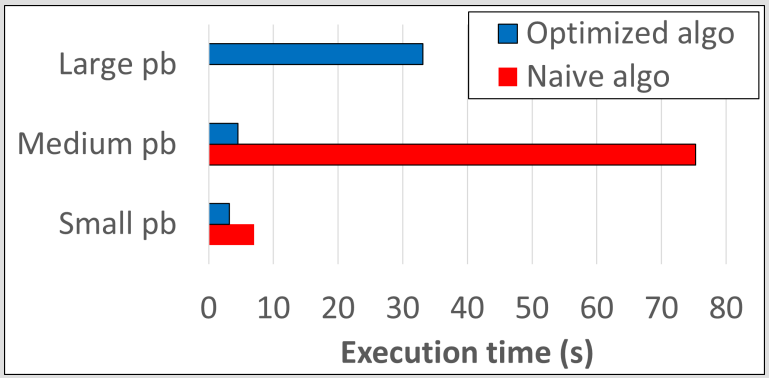


III. Distributed Hadoop/Spark on CPU cluster[1/2]



Description of the use case

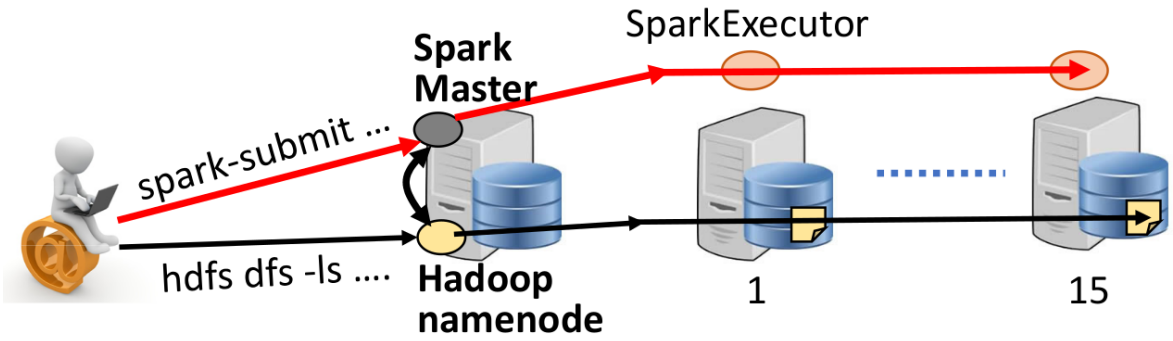
Design of a Map-Reduce with limited network traffic



Bad algorithms on real distributed systems ⇒ huge network cost

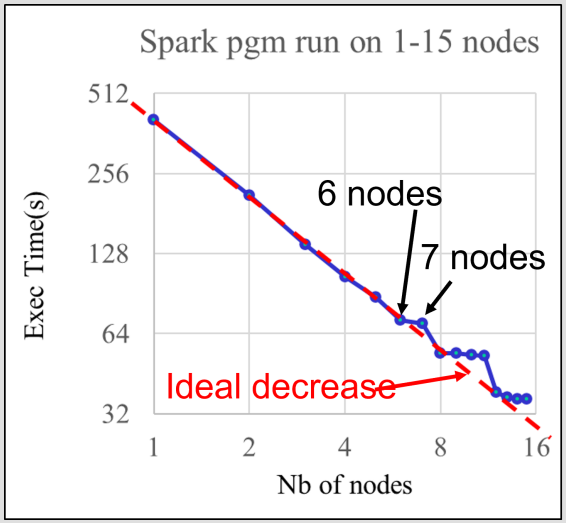
Teaching use cases

III. Distributed Hadoop/Spark on CPU cluster[2/2]

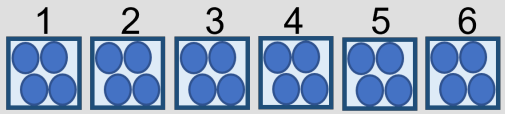


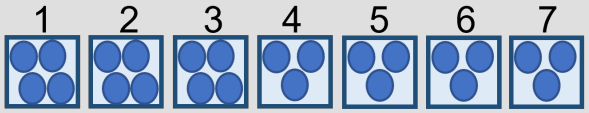
Description of the use case

Observe load balancing impact on speedup of Spark programs.



Ex: Load balancing of a **24-task** Spark pgm:





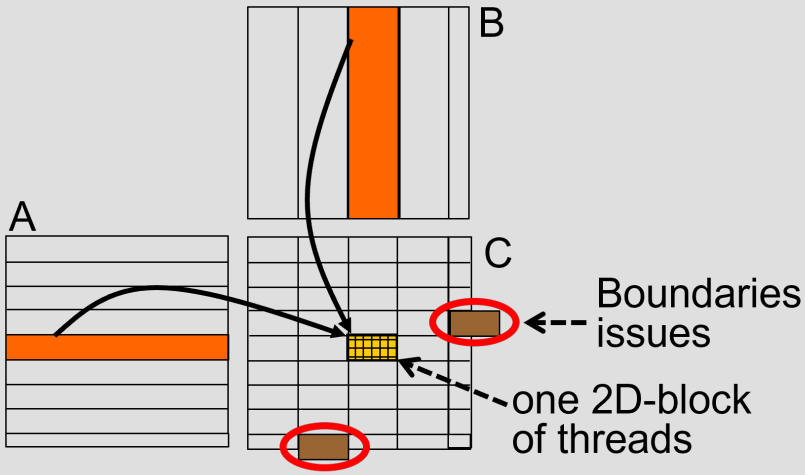
T(7 nodes) = T(6 nodes) = T(4 tasks)

Description of the use case

First steps with CUDA on GPU: dense matrix product

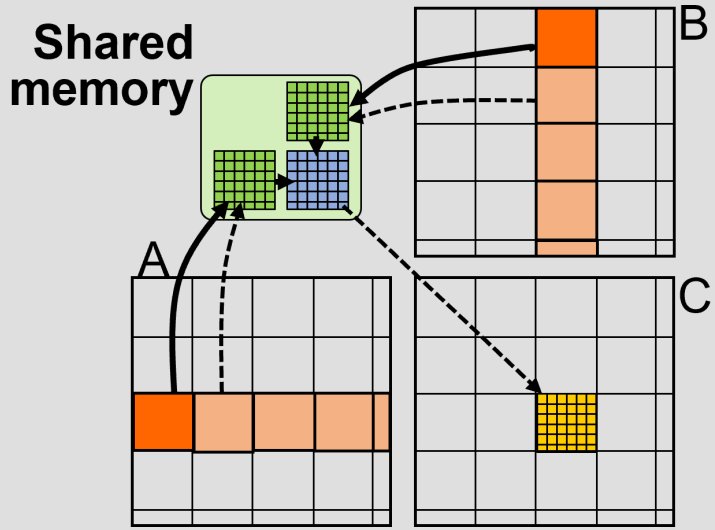
Issues:

- Grid of blocks of threads
- First CUDA kernels
- Coalescent mem. accesses



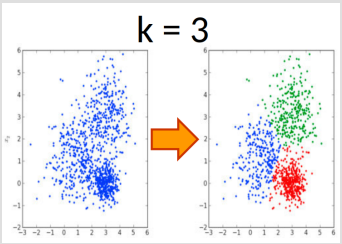
Issues:

- Use the shared memory
- Increase the coalescence
- Decrease the number of global memory accesses



Description of the HPDA use case

K-means clustering on GPU with CUDA

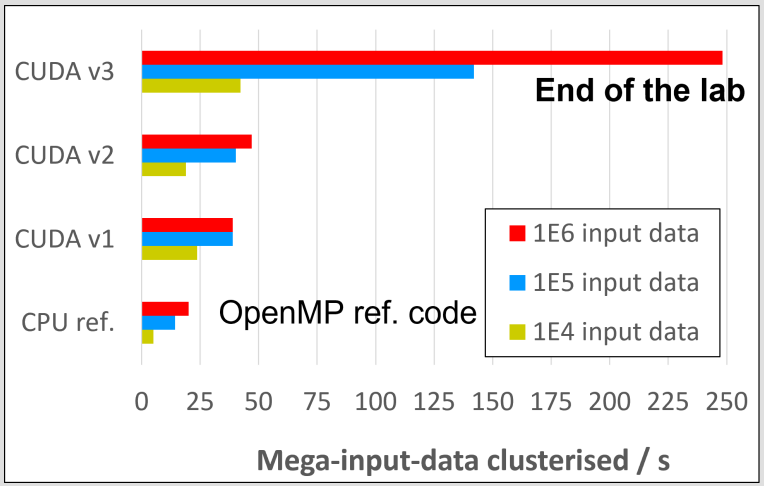


Issues adressed:

- Coalescence
- Shared memory
- Reduction
- Atomic functions

Lab: Design the right data structures and associated computations to obtain optimized kernels ... many combinations exist !

After 2 lab sessions :



Many CUDA optimizations succeed.

Questions

Thank you for your attention !

We are gratefull for the support of

More details on
`https://dce.pages.centralesupelec.fr/`.

Feedback from a data center for education at CentraleSupélec engineering school, J. Fix, S. Vialle, R. Hellequin, C. Mercier, P. Mercier, J.-B. Tavernier
`jeremy.fix@centralesupelec.fr`