romeoLAB, le portail web HPC : cas d'utilisation pour la pédagogie et les logiciels à la demande

Dr Arnaud RENARD - HPC CTO & Research Engineer
Dr Jean-Matthieu ETANCELIN - HPC Research Engineer
Pr Michael KRAJECKI - HPC Center Director



romeolab@univ-reims.fr

http://romeo.univ-reims.fr
@HPCromeo
#romeoHPC




UNIVERSITÉ
DE REIMS
CHAMPAGNE-ARDENNE




University of Reims

Université de Reims Champagne-Ardenne (URCA)

Multidisciplinary university

- about 27 000 students
- 5 campus : Reims, Troyes, Charleville-Mézières, Chaumont et Châlons-en-Champagne
- a wide initial **undergraduate** studies program
- **graduate** studies and **PhD** program linked with **research labs**

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

romeo LAB


HPC Center
ROMEO
Centre de Calcul Régional

ROMEO
HPC Center for Grand-Est region

Its **mission** is to deliver, for both **industrial and academic** researchers :

CLASSICAL MISSIONS

- **high performance** computing **resources**,
- **secured storage** spaces,
- specific & **scientific softwares**,
- **advanced user support** in exploiting these ressources,
- in-depth **expertise** in different engineering fields: HPC, applied mathematics, physics, biophysics and chemistry, …

SPECIFIC MISSIONS

- promote and diffuse HPC and **simulation** to companies / **SMB**
- identify, experiment and master **breakthrough technologies**
  - which give new **opportunities** for our user
  - from **technology-watching** to **production**
  - for all **research domains**

GPU, DL, Quantum

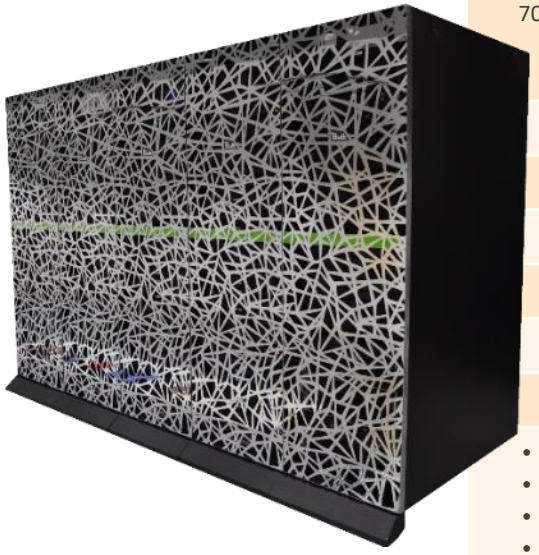
- **Teaching High Performance Computing to Researchers and Students**

Université de Reims Champagne-Ardenne

ROMEO

Principales machines de production

2013 / Bull	*Année / Constructeur*	2018 / Bull-ATOS
2,5 M€	*Investissement*	6,5 M€
2 080 cœurs 700 000 cœurs GPU (260 K20X)	*CPU*	3 220 cœurs (skylake @2,6 GHz) 1 000 000 cœurs GPU (280 P100)
254,0 Tflops	*Puissance*	**1,022 Pflops** x 4
70 KW	*Energie*	120 KW
4 To	*Mémoire*	15 To
200 To (//)	*Stockage*	600 To (//)
130	*Serveurs*	115
40 Gb/s *IB*	*Interconnexion*	100 – 200 Gb/s BXI
• GPU • romeoLAB • TOP500 #151 • GREEN500 #5 • GRAPH500 #105		• DLI & DL • TOP500 #254 • GREEN500 #20 • HPCG #63 • GRAPH500 #??



Motivation: romeoLAB must be *powerfull*

Necessity of computer science courses dedicated to parallel programming on heterogeneous architectures. (Students, Reserchers, European & French projects)

- HPC is more and more complex (different processors, memory, network, …)
- Hardware specific phenomenon (bandwidth bottleneck, cache page issues, …)

➢ We want to execute code on a real HPC facility.
Romeo works perfectly, got software already installed and supported.

REIMS.GPU spring SCHOOL 9 - 13 mai 2016

As an example, we organized a 5 days event dedicated to GPU technologies in 2016

[Arnaud Renard, Jean-Matthieu Etancelin, Michaël Krajecki: romeoLAB: A High Performance Training Platform for HPC, GPU and DeepLearning. CARLA2017: 55-67]

Motivation: romeoLAB must be *powerfull*

But using HPC supercomputer is difficult :

- Ssh, Sftp (+installing clients),
- Password, login, account opening account process,
- Load module environment,
- Deal with Slurm,
- Xserver or VNC client for Graphical interface (profiler, debugger, viewers, …),

Using those tools

- is time-consuming;
- is not pedagogic objectives of courses;
- are obstacles to the pedagogical process efficiency;

Motivation: romeoLAB must be *easy*

A web-based solution is the most easy-to-use solution.

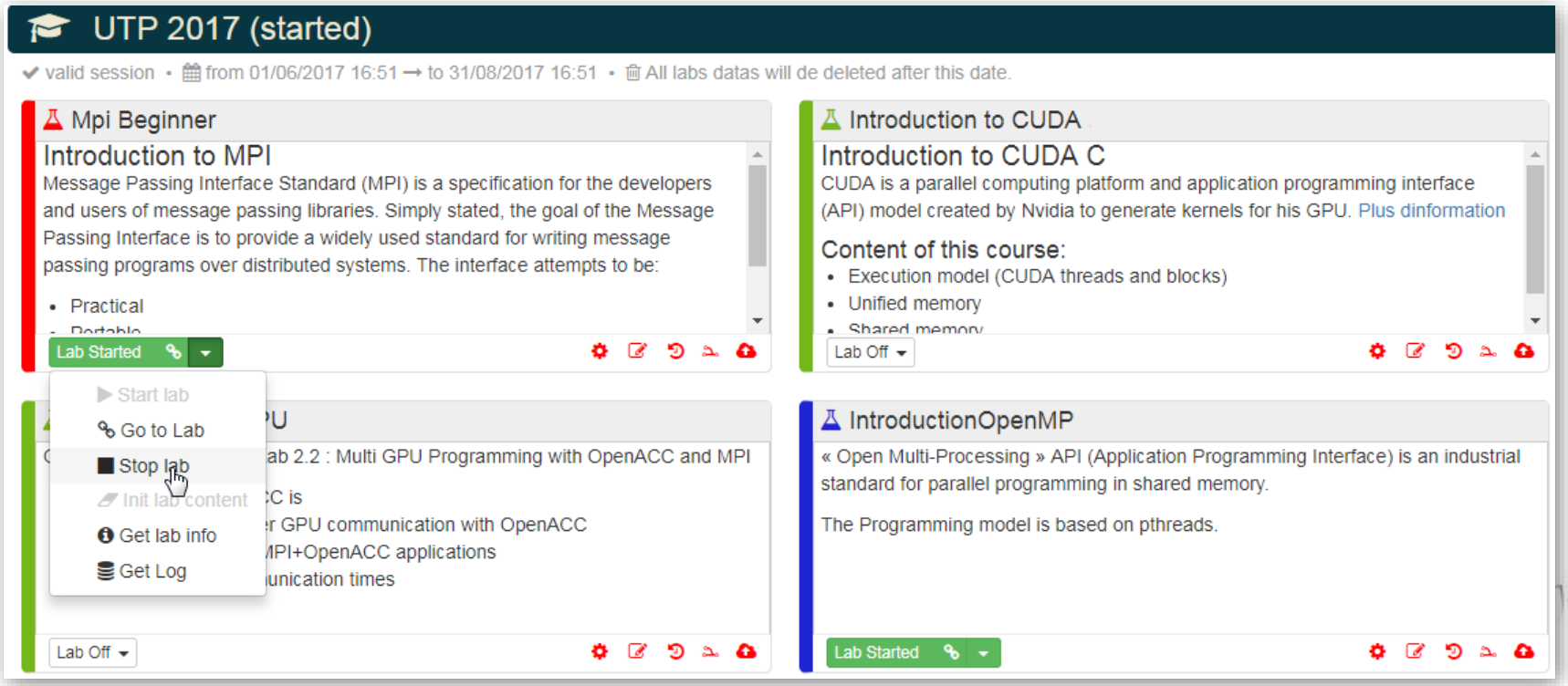
- Users already have a browser
- Lot of references / other experiences
- HTML5 allow almost everything
- Multi-device
- Poor internet connections

User management

- Simple registration process
- Disposable accounts
- Access to session with access code
- Start and Stop labs thru SLURM jobs / reservation

Fonctionnalités :

- romeoLAB
  - Orienté HPC
  - Execution sur notre cluster de prod (2500 coeurs / 260 GPU), slurm (reservation)
  - Cours (session, via token, persistance) = n * exercice (Lab) = n * notebook
  - Contenus dans GIT
  - Interface Enseignant



Motivation: romeoLAB must be *Pedagogic*

In romeoLAB, (like modern MOOC - Massive Open Online Course), development environments is completly integrated into pedagogical content.

- We use Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text.
Uses include: data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more.

Motivation: romeoLAB must be *Pedagogic*

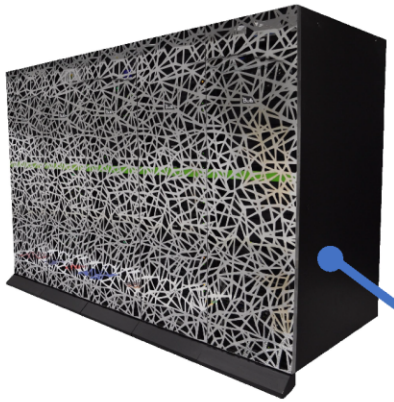
On the same web-page :

- development environments
  - File browser & Editors (native & NodeMirror)
  - Compilation, command lines & modules loading (native)
  - Execution live (native)
  - Execution in batch for large runs (in-house Ipython-batch-execution-magics) : 24 students can each run a 32-nodes MPI runs with 32 nodes in total
- Desktop access to run Graphical software
  - Profilers, Debuggers, Graphical interfaces like Paraview (in-house integration with VNC + x11vnc + VirtualGL)
- Pedagogical content
  - Pdf, Video, content, images, iframed web pages … (native or in-house)

romeoLAB : a High Performance Training Platform for HPC, GPU and DeepLearning

Dr Jean-Matthieu ETANCELIN - jean-matthieu.etancelin@univ-reims.fr
Dr Arnaud RENARD - arnaud.renard@univ-reims.fr
University of Reims Champagne-Ardenne - CReSTIC EA3804 - http://romeo.univ-reims.fr

" Romeo : THE MOST AMAZING GPU CLUSTER IN EUROPE "
Pr Krajecki - GTC2016



Hosted at University of Reims Champagne-Ardenne.
Designed by Atos-Bull.
Powered by 260 GPU Tesla Cards
Ranked 151@Top500 in Nov. 2013

Romeo is dedicated to research, professional services and education,
in the GrandEst Region, France

2,5 M€, funded by

Live code in a web page, running on HPC nodes

Content is versioned and modified by designated teachers

Jupyter NoteBook

GIT

Lab workload managed with native HPC scheduler and mixed with scientific jobs

romeoLAB core

SLURM

Web Server

All your labs available in your browser, https sessions secured with password

Deploy On Demand

Web Proxy

HPC environment with custom software stack, Deep Learning Frameworks via Containers and private cluster

Docker

Virtual Machines

Easy access GUI, editor and Deep Learning frameworks through the secured http proxy

EXAMPLE OF TEACHING SESSIONS :

- Master courses
- Reims GPU Spring School
- GTC2016 and GTC2016-Eu Hands-on Labs
- DeepLearning Schools 2017

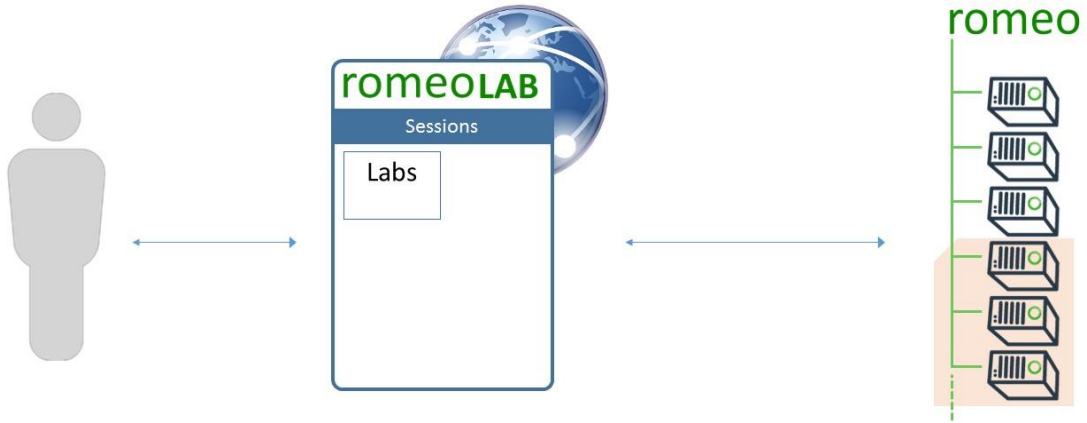
EXAMPLE OF EXISTNG LABS :

- GPU Accelerated Libraraies (cuBLAS, cuFFT, cuRand, ...)
- Multi-GPU Programming with OpenACC
- Advanced Multi-GPU: rCUDA and GPU Direct RDMA
- OpenFoam Tutorial
- OpenMP for HPC, Python programming
- Caffe for DeepLearning

romeoLAB Big Picture 1/3

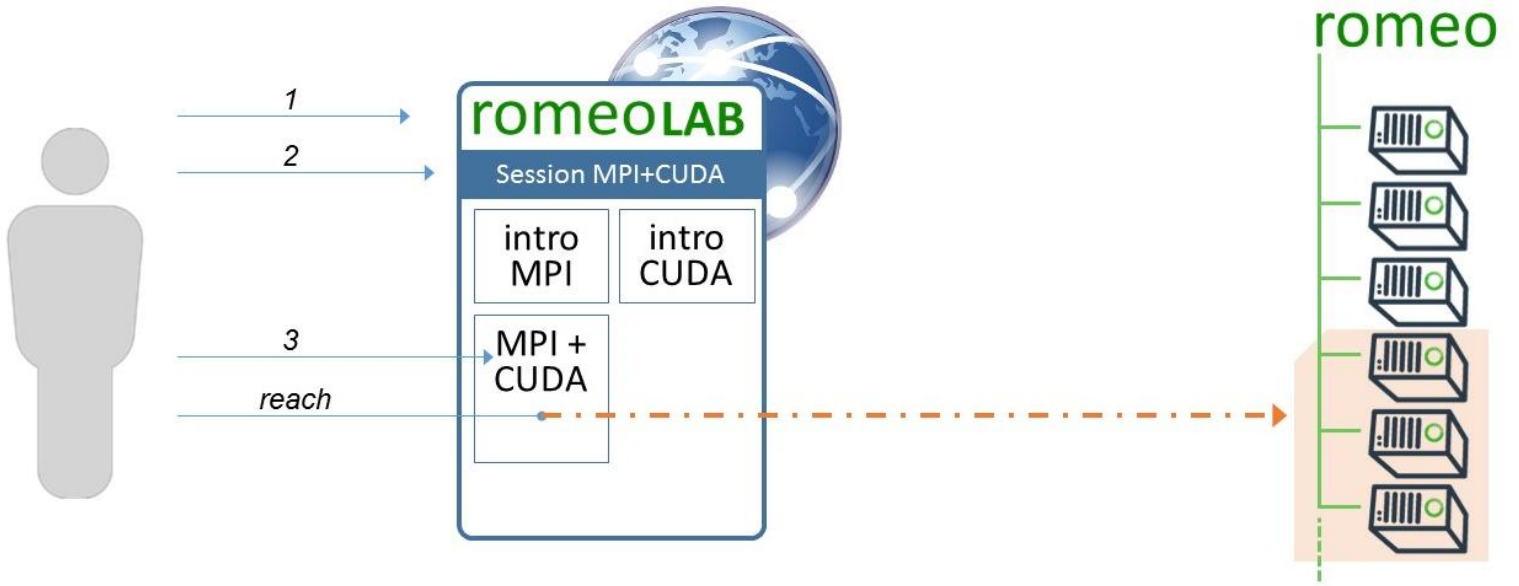
The Jupyter NoteBook is running on the compute node (protected mode)
All the management of *romeoLAB* is made on the web server (FrontEnd)

We need to develop our own server because existing one (like jupyterHub) do not correspond to our needs of customization, integration, evolutivity, dedicated to teaching … Written in PHP + MVC + RedBean + MariaDb + …



Users can reach and leave sessions, start and stop labs, …
Teachers can create sessions or labs, update labs contents, and manage students.

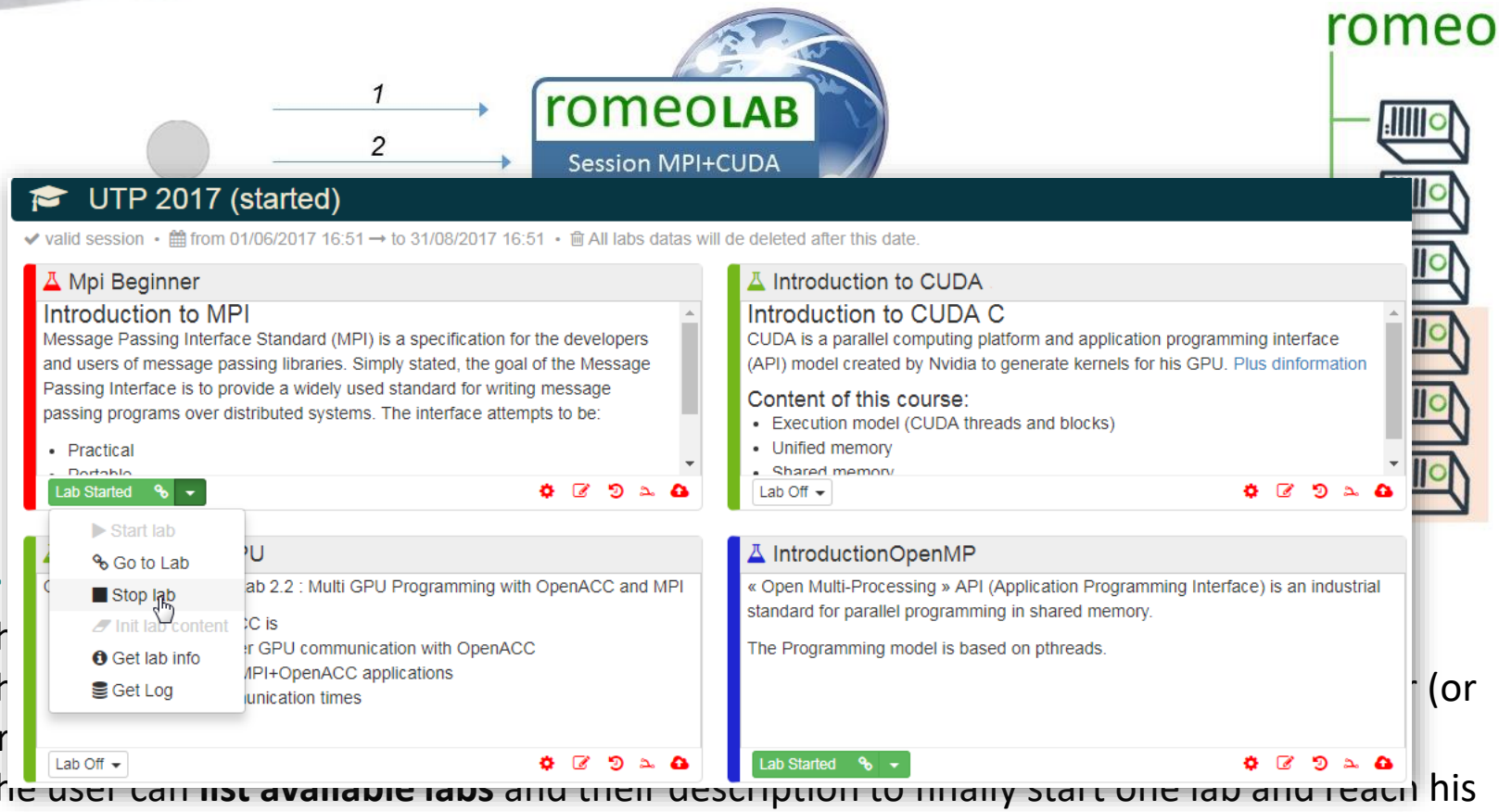
romeoLAB Big Picture 2/3



User view of accessing an interactive content in *romeoLAB*:

1. The user **creates an account**, and log in to the platform
2. The user **reach an active Session (=classroom)** with the access code given by teacher (or provided by an activation link, provided by email, on a webpage or Paypal)
3. The user can **list available labs** and their description to finally start one lab and reach his IPython Notebook running on one compute node. He can watch videos and documents, fill table with performance results, edit, compile and also profile code via a remote desktop.

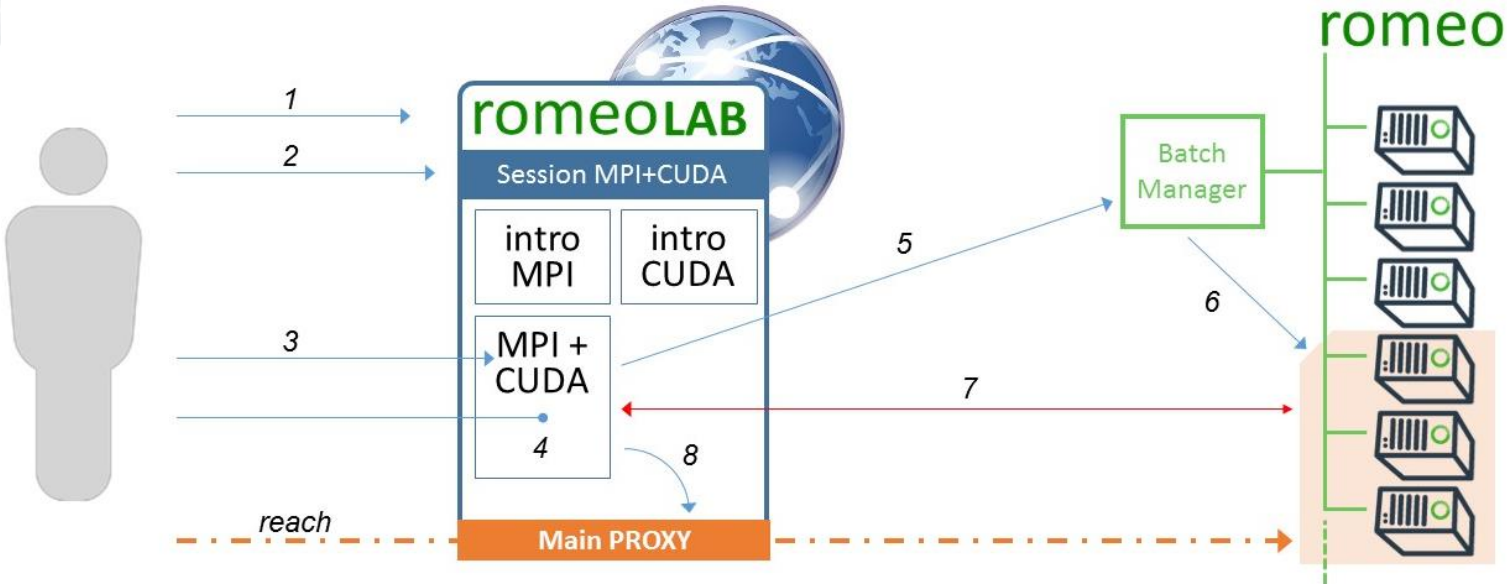
romeoLAB Big Picture 2/3



User

1. Th
2. Th (or pr
3. The user can **list available labs** and their description to finally start one lab and reach his IPython Notebook running on one compute node. He can watch videos and documents, fill table with performance results, edit, compile and also profile code via a remote desktop.

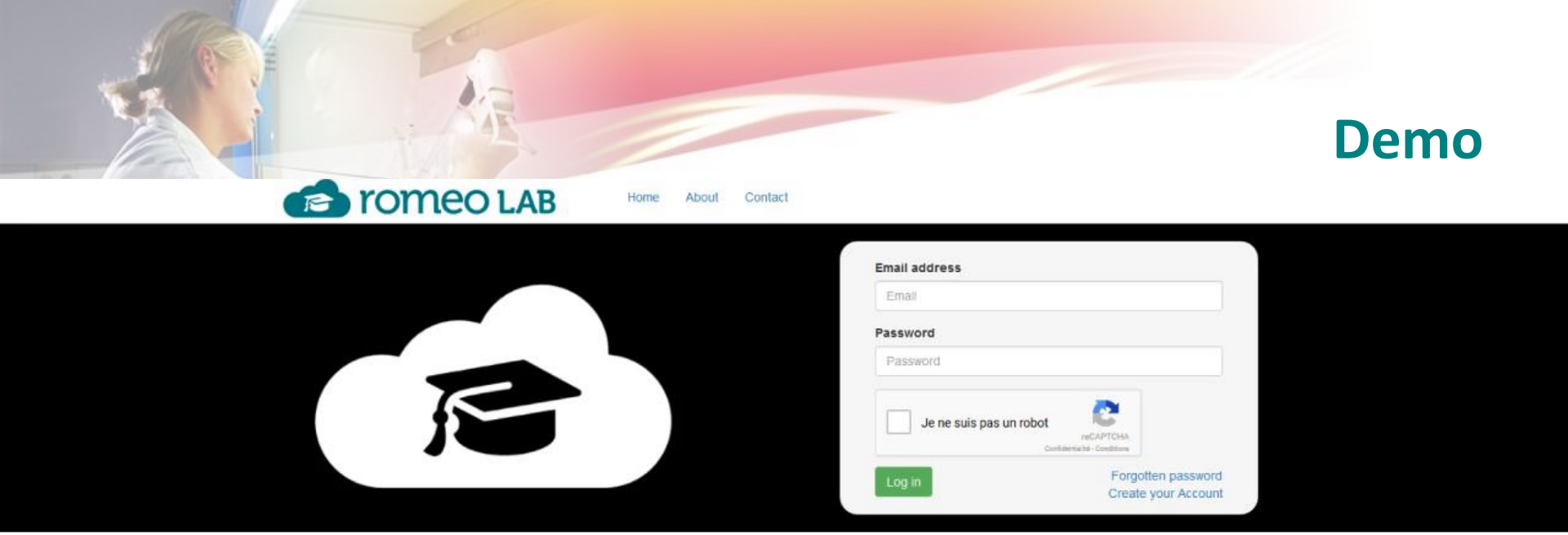
romeoLAB Big Picture 3/3



Internal behavior when a user starts a lab:

4. Server assign a **temporary cluster-user** to the user and dynamically load initial lab content (from the lab repository with GIT protocol).
5. It launches a job through the workload scheduler (and possibly via reserved dedicated resources).
6. This job setting up all resources parameters (available ports), starts all services (notebook, editors, VNC server, …)
7. romeoLAB is waiting / probing the start of those services
8. **Proxy routes** are setup (Main proxy) to provide direct access to these services (we use "Configurable HTTP Proxy", which is a NodeJs tool)

Demo


romeo LAB
Home
About
Contact
Email address
Email
Password
Password
Je ne suis pas un robot
reCAPTCHA
Log in
Forgotten password
Create your Account
This site is the lab platform for the University of Reims Champagne-Ardenne computer science courses. The site is used to facilitate the hands-on lab and is developed by ROMEO HPC Center.
Labs are powered by ROMEO HPC Cluster which is the largest GPU Cluster in France. Courses are provided by ROMEO HPC Center, University of Reims Champagne-Ardenne, and partners worldwide. You are welcome to contribute.
ROMEO HPC Center is dedicated to teaching and research in French Grand-Est region and offers computational ressources, storage, expertise and services for companies. Please contact us for any request : romeo@univ-reims.fr.
Grand Est
ALSACE CHAMPAGNE-ARDENNE LORRAINE
GRAND REIMS
COMMUNAUTÉ URBAINE
Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

USAGE:
Previous events and Sessions

Name	Attendees	Duration	Courses
2017:			
JDEV2017	24	4 days	GPU programming
Groupe Calcul	38	3 days	Advanced Python for HPC
Profiler Days	17	3 days	Profiling tools for parallel codes
2016:			
Master Courses	40	4 months	GPU programming and HPC
OpenFOAM School	20	3 days	OpenFOAM software
10th LoOPS day	60	2 days	C++ (HPX) vs Python (DSLs)
GTCEU2016	55	90 minutes	MPI and OpenACC
GPU Spring School	36	1 week	GPU programming
GTC2016	60	90 minutes	Advanced tools for hybrid cluster

2017/2018 :

linux / Gaussian / Molden / OpenFOAM / industriels / VI-HPS / Abaqus

USAGE: Available content

romeoLAB is addressing a wide range of technologies and audience levels. As we encourage mutualization, this list is growing :

Beginner :

- Introduction to : Python, OpenMP, MPI , CUDA, OpenACC, …
- GPU accelerated applications : CUDA, OpenACC, Python, …
- GPU accelerated libraries : cuBLAS, cuRand, cuFFT, …

Intermediate

- OpenFOAM
- Gaussian / Molden
- OpenCL, CUDA Asynchronism

Advanced

- Profiling : TAU, MAQAO
- Advanced Python : Cython, Numba, Pythran
- CUDA Optimizations
- Multi-GPU with CUDA
- Multi-GPU with OpenACC and MPI
- Multi-GPU with rCUDA & GPU-Direct RDMA

Conclusion

Powerfull : we want to execute code on a real HPC facility, because ours works perfectly and we've got all the software already installed and supported. Where're using a proxy for external users to access to compute nodes.

Easy : romeoLAB is a modern MOOC platform making it possible to run HPC in a simple web browser.
Ssh, ftp and job managers (Slurm) are not part of courses educational objectives nor prerequisites.

Pedagogic : on the same web page, student must find lessons (video, pdf, images, …) and the edition / compilation / execution interfaces. Jupyter Notebook is our solution for strong interactivity.
Teacher can create his courses on the same platform and manage attendees to his courses,

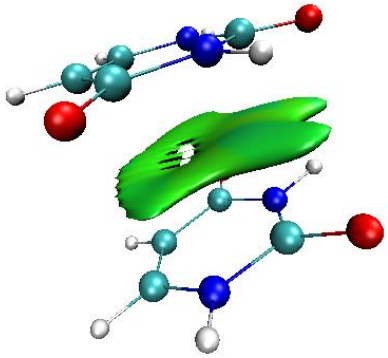
Multi-Application : Compiling and executing code is not enough. We must run profilers, GUIs, and other scientific software.

ICMR
Institut de Chimie Moléculaire de Reims
UMR CNRS 7312

CReSTIC

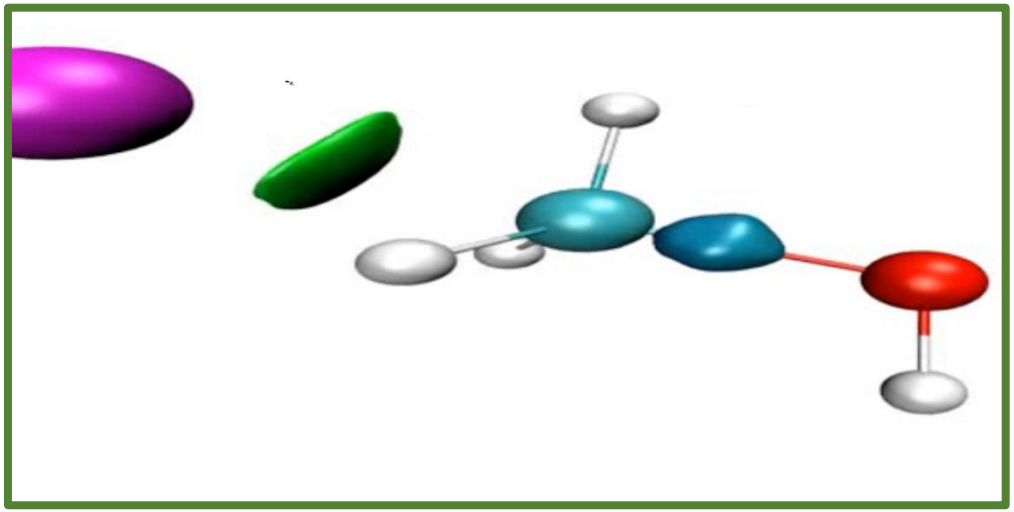
romeoLAB on-demand

Applications IGMPlot

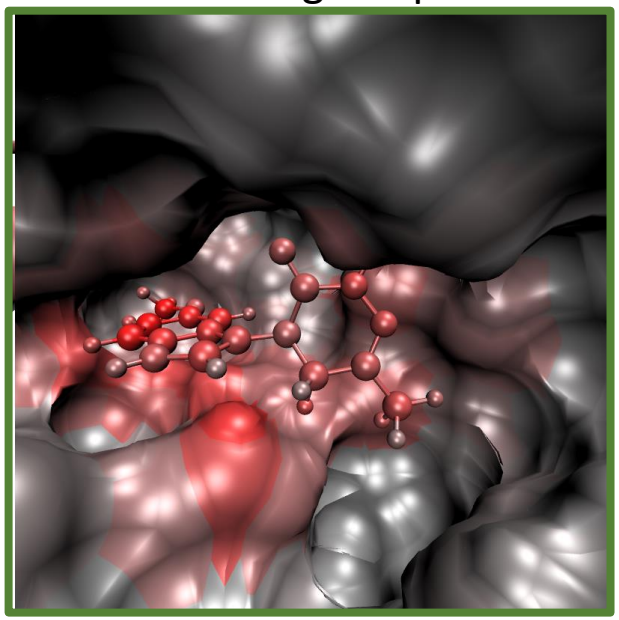


http://igmplot.univ-reims.fr

Réaction chimique SN_2



Interaction ligand-protéine

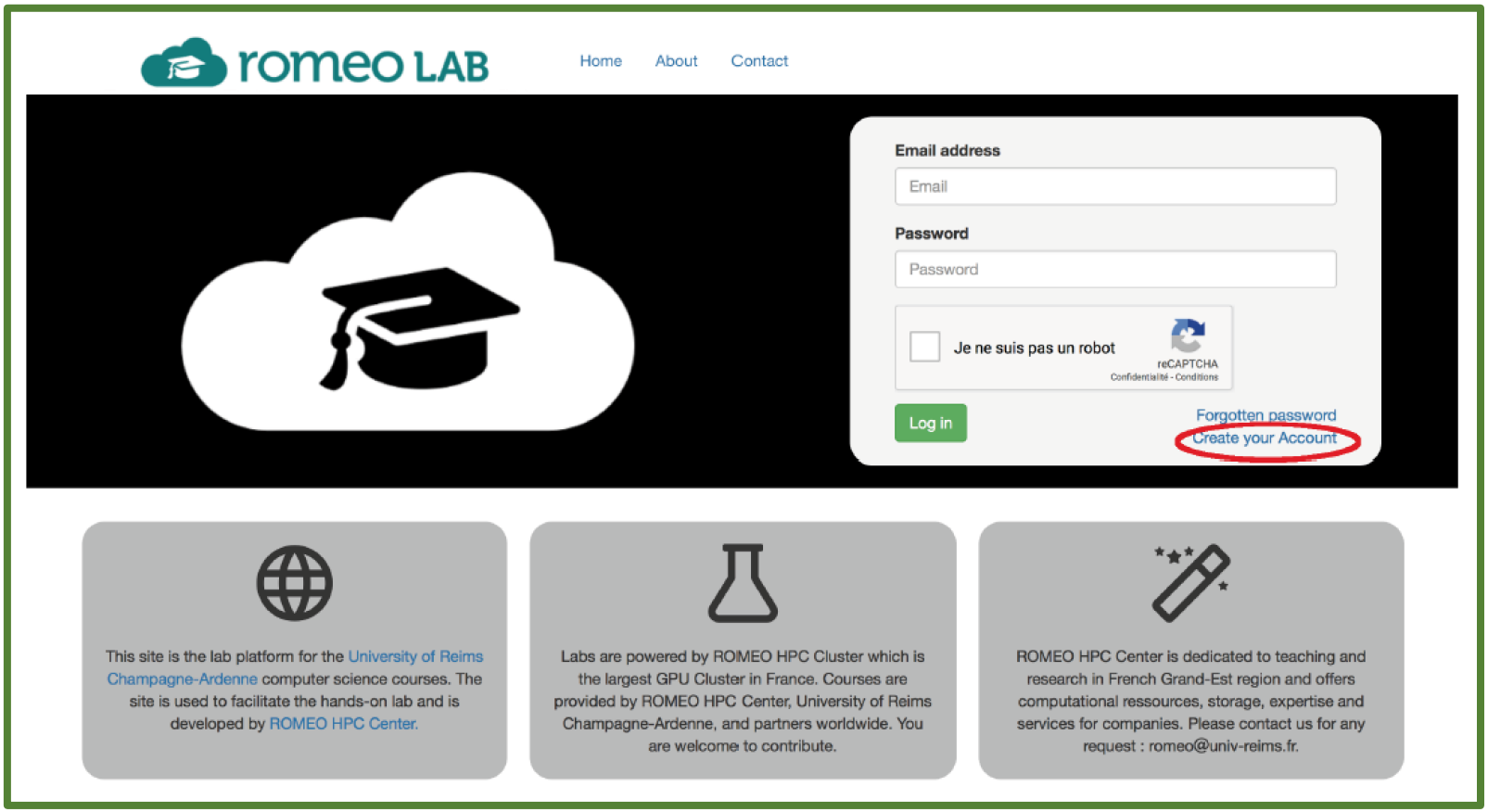


IGMPlot On-demand

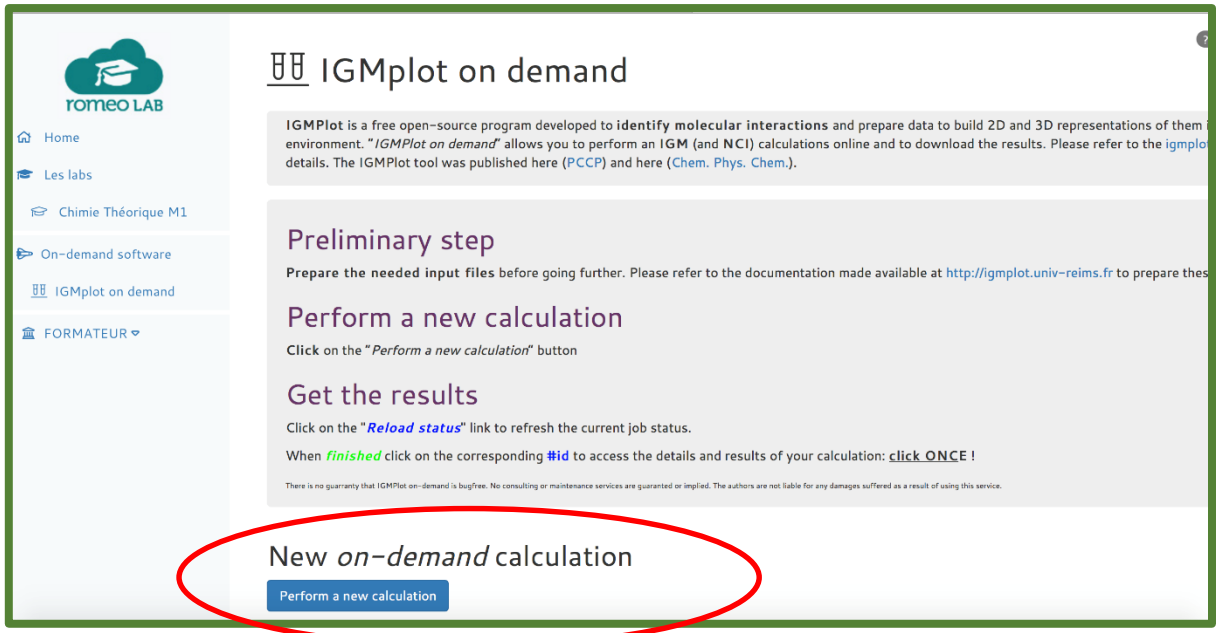
ICMR
Institut de Chimie Moléculaire de Reims
UMR CNRS 7312

CReSTIC

Centre de Calcul ROMEO
MAISON DE LA SIMULATION
de Champagne-Ardenne



IGMPlot On-demand



IGMPlot On-demand

ICMR
Institut de Chimie Moléculaire de Reims
UMR CNRS 7312

CReSTIC

DEMO

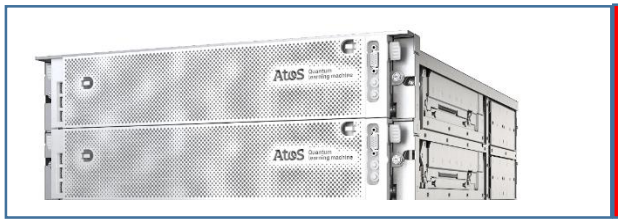
romeoLAB : actualités

PRACE VI-HPS

Université de Reims Champagne-Ardenne — ROMEO

15 – 19 octobre

2 – 4 février



19 – 20 septembre

8 novembre

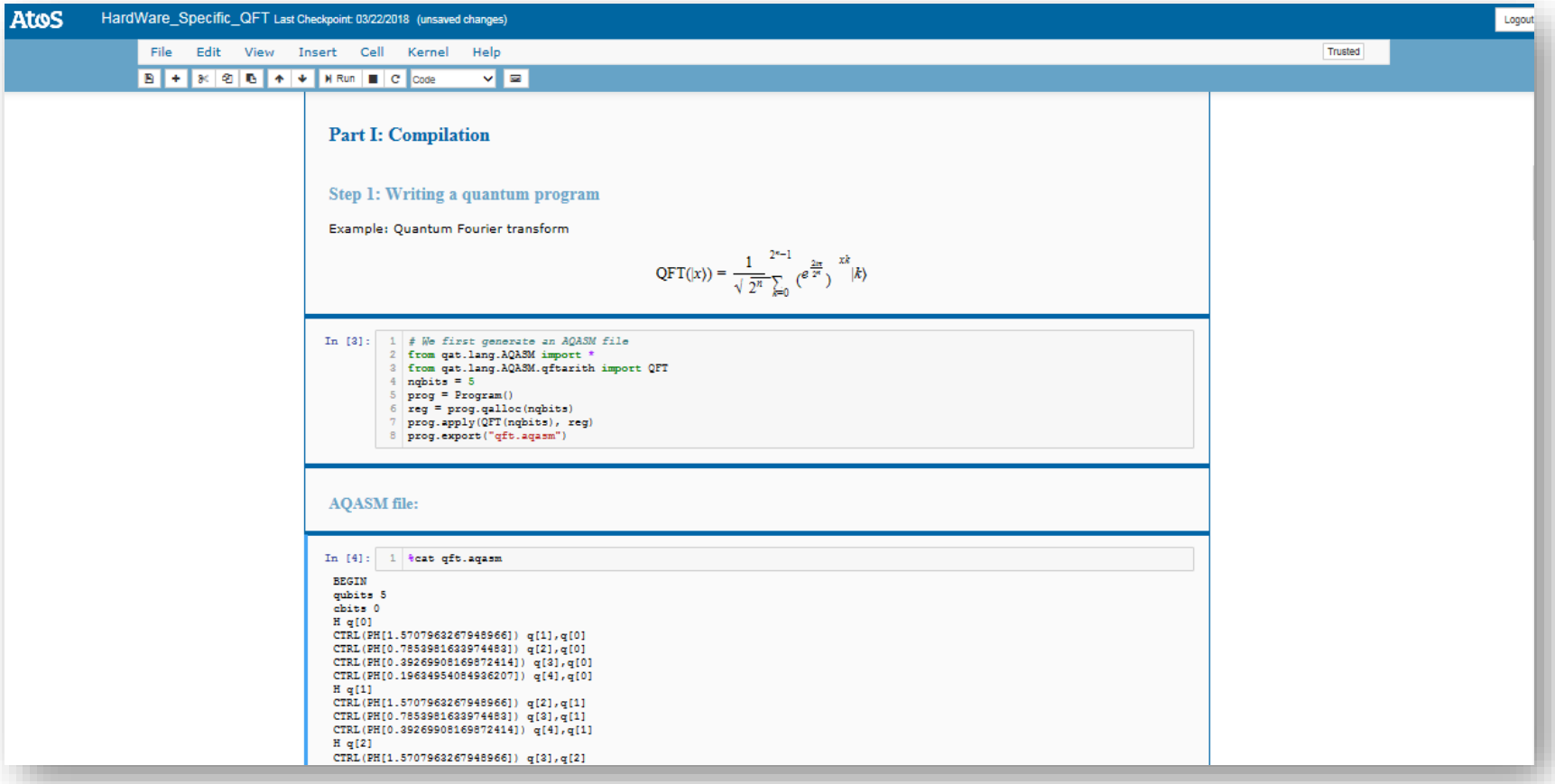
NVIDIA. DEEP LEARNING INSTITUTE

NVIDIA. DEEP LEARNING INSTITUTE UNIVERSITY AMBASSADOR

romeo LAB

Université de Reims Champagne-Ardenne

romeoLAB : access single nodes (DGX - QLM - …)



Perspectives *romeoLAB*

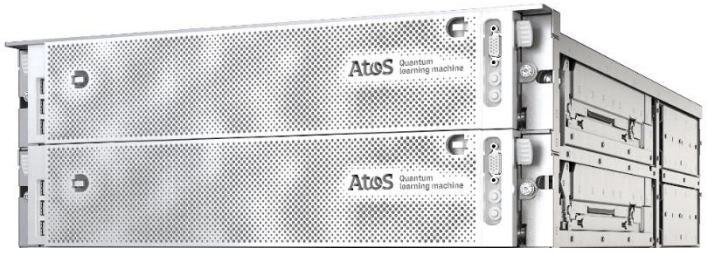
- Maintenir
- Version distribuable
  - nouveau supercalculateur
  - partenaires / opensource
  - rename
- Sécurité
- OpenStack & Single machie

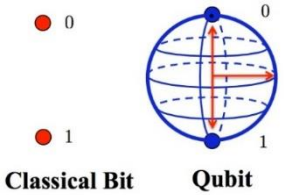
romeo LAB

romeolab@univ-reims.fr

Quantum computing

- Solve difficult problems
  - Classical bit VS Qubits





- **QLM** Group (partners on right)
  - Quantum simulator platform
  - Develop new algorithms (BD, AI, SC, Cyber security)
  - Designing computing architectures
  - Quantum safe cryptography algorithms

cea

OAK RIDGE National Laboratory

Argonne NATIONAL LABORATORY

Université de Reims Champagne-Ardenne

JÜLICH FORSCHUNGSZENTRUM

SURF SARA

fhoö UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA

UNIVERSITY of TASMANIA