European Commission

Trillion

INFN
Istituto Nazionale di Fisica Nucleare

Introduction to CINECA supercomputers

Alexei Sytov

Catania, 17/09/21

Briefly about me

- **2018: PhD** in Experimental Physics, University of Ferrara and in Theoretical Physics, Belarusian State University
- **2019-2021**: **Post-doctoral Fellow** in Experimental Physics at the INFN Division of Ferrara.
- Since **2020** involved in **MC_INFN**
- Since **02/09/2021**: Marie Sklodowska-Curie Action Global Individual Fellowships, project 

- **Main field**: interaction of charged particles interaction with crystals; channeling, channeling radiation, accelerator physics.
- **My computing experience**: HPC Monte Carlo simulations, usage of CINECA resources since 2015, PI of 4 projects.
- **Additionally**: Fortran, C/C++, Mathematica, Python, Geant4, Keras deep learning framework.

Marie Sklodowska-Curie Action Global Individual Fellowships by A. Sytov in 2021-2024, Project TRILLION

Main goal: The **implementation** of both physics of **electromagnetic processes in oriented crystals** and the design of specific applications of crystalline effects into **Geant4** simulation toolkit as Extended Examples to bring them to a large scientific and industrial community and under a free Geant4 license.

Group:

- **A. Sytov** – project coordinator
- **L. Bandiera** – INFN supervisor
- **J. Perl** – SLAC supervisor
- **G. Kube** – DESY supervisor
- **I. Chaikovska** – IJCLab Orsay supervisor

Location:

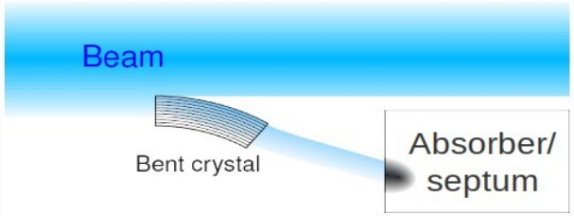
- 2 years at **SLAC** (partner organization).
- 1 year at **INFN Section of Ferrara** (host organization).
- 1 month of secondment at **DESY** (partner organization).
- 1 month of secondment at **IJCLab Orsay** (partner organization).

We thank the **MC_INFN** project for the support!

Marie Sklodowska-Curie Action Global Fellowships by A. Sytov in 2021-2024, Project TRILLION

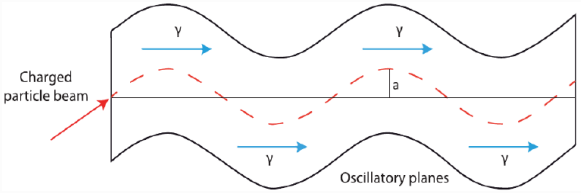
Specific applications to implement into Geant4:

- **Crystalline deflector to extract a charged particle beam from an accelerator** (electron synchrotron, hadron collider) to supply fixed-target experiments by an intense low-emittance beam.



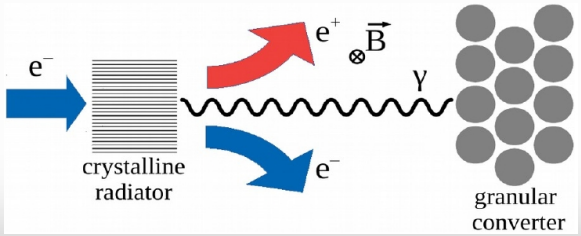
Directly connected with CINECA Project **LEADER**

- **Crystalline source of hard X-ray and gamma radiation, crystalline undulator (CU)**.



Directly connected with milestones 2021 of **MC-INFN** as well as INFN **ELIOT** and **STORM**; European Commission **N-LIGHT** and **PEARL** projects

- **Crystal-based hybrid positron source** for both linear and circular e+e- colliders (**ILC**, **FCC-ee**) as well as for **muon colliders**.



MC-INFN and **LEADER** goals are necessary preparisons for **TRILLION** which will help to accomplish INFN **ELIOT** and **STORM** as well as **N-LIGHT** **=> synergy of projects**

CINECA*

- **Cineca** is a non profit Consortium, made up of **70 Italian universities**, 5 Italian Research Institutions (including INFN) and the Italian Ministry of Education.
- the **largest Italian computing centre**, one of the most important worldwide
- **established** in **1969** in Casalecchio di Reno, Bologna
- **member of PRACE** – The Partnership for Advanced Computing in Europe
- **institutional mission** is to support the Italian scientific community through supercomputing and scientific visualisation tools

*http://www.cineca.it/

What is Marconi-100*?

- The Italian **Tier-0 system** is based on the **IBM Power9 architecture** and **Volta NVIDIA GPUs**
- available in Cineca since **April 2020**
- **14th** position in the Top500 list (**3rd** in **EU**) with a sustained performance of **21.640 Pflops** (peak performance up to **~29.354 Pflops**)
- **10^5-10^6** times faster than a personal computer
- **Location**: Cineca, Casalecchio di Reno, Bologna, Italy
- Acquired by Cineca within PPI4HPC European initiative

*https://www.hpc.cineca.it/hardware/marconi100

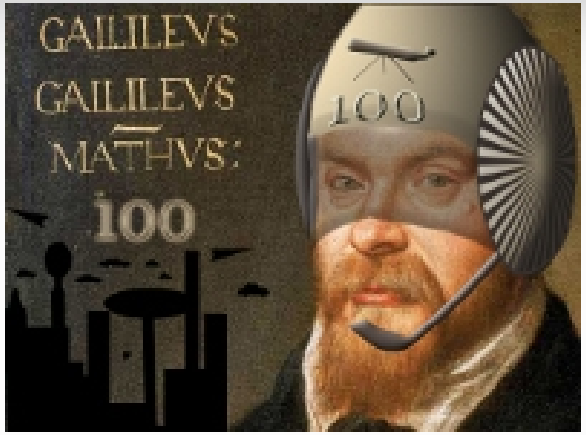
Marconi-100 hardware*

- **Nodes**: 980
- **Processors**: 2 16-cores IBM POWER9 AC922 at 2.6(3.1) GHz
- **Cores**: 32 cores/node, 128 logical cores/nodes, **125440 cores** in total
- **Accelerators**: 4 x NVIDIA Volta V100 GPUs/node
- **RAM**: 256 GB/node
- **Internal Network**: Mellanox IB EDR DragonFly++
- **Disk Space**: 8 PB of local scratch
- **Peak Performance**: 29 PFlop/s

*https://www.hpc.cineca.it/hardware/marconi100

What is Galileo-100*?

- The Italian **Tier-1 system** co-designed by CINECA and engineered by DELL
- **Location**: Cineca, Casalecchio di Reno, Bologna, Italy
- Co-funded by the European ICEI (Interactive Computing e-Infrastructure) project.



- **Nodes**: 528
- **Processors**: 2xCPU x86 Intel Xeon Platinum 8276-8276L
- **Cores**: 48 cores/node, **25344 cores** in total
- **RAM**: 384 GB/node
- **180 nodes** for data processing with 3.0TB Optaine (enhanced RAM)
- **36 GPU nodes** with 2x NVIDIA GPU V100

*https://www.hpc.cineca.it/hardware/galileo100

Other CINECA clusters*

- **MARCONI-A3**: Intel SkyLake 2x Intel Xeon 8160 @2.1GHz 24 cores each, **3216 nodes**, **154368 cores**
- **DGX**: It is the NVIDIA A100 accelerated system suitable for Deep Learning frameworks. AMD 3 nodes; **8 NVIDIA A100 Tensor Core GPUs** per node.
- **CLOUD.HPC**: CINECA cloud service, Intel Broadwell
  2x Intel Xeon E5-2697 v4 @2.3GHz 18 cores each; **80 nodes**

In the future:

- **LEONARDO – pre-exascale cluster (200 Pflops)**

- Also some resources are available for **quantum computing**, emulators on **Marconi 100** and an access to D-Wave systems

*https://www.hpc.cineca.it/content/hardware

How to ask for HPC resources: CINECA ISCRA

- **Class B Projects**: national level project; duration 12 month, application call few times per year, more than 1 M core*h
- **Class C Projects:** small projects, duration 9 month, application call every month, simplified application form, ~100 kcore*h
- **Class D Projects:** Data storage Project, up to 50 TB, maximum duration 36 month.

- **At the end of the project**: a short report + published papers:

  We acknowledge the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support

*https://www.hpc.cineca.it/services/iscra

Dedicated supercomputer time on Marconi-100: project LEADER, Cineca ISCRA Class B National Italian project, no. HP10BHSQLS

```
**********************************************************************************
* Welcome to MARCONI100 Cluster /                                                *
*            IBM Power AC922 (Whiterspoon) -                                     *
*            Red Hat Enterprise Linux Server release 8.1 (Ootpa)                 *
*                                                                                *
* 980 compute nodes with:                                                        *
*       - 2x16 cores IBM POWER9 AC922 at 3.1 GHz                                 *
*       - 4 x NVIDIA Volta V100 GPUs, Nvlink 2.0, 16GB                           *
*       - 256 GB RAM                                                             *
```

1.5 Mh for 1 year

Main goal: simulation of electron beam crystal-based extraction from the DESY-II synchrotron

PI A. Sytov

One scope of the project is simulations the development of radiation loss model: **synergy with MC-INFN**

Project collaborators

INFN Section of Ferrara
- Dr. Laura Bandiera
- Prof. Vincenzo Guidi
- Dr. Andrea Mazzolari
- Mattia Soldani

DESY
- Dr. Gero Kube

INFN, Laboratori Nazionali del Sud
- Prof. Pablo Cirrone

Institute for Nuclear Problems, Minsk, Belarus
- Viktar Haurylavets
- Prof. Victor Tikhomirov

Ask me if you need CPU time!

Our future project MIRACLE, Cineca ISCRA Class B National Italian project

MIRACLE
Medical physIcs and RAdiation in Crystals simuLation with gEant4
Main goal: to supply Italian Geant4 community and their international collaborators by CINECA HPC resources necessary to accomplish MC_INFN and TRILLION projects.

Italian organizations involved

- INFN Sezione di Catania
- INFN Sezione di Ferrara
- INFN Laboratori Nazionali del Sud
- INFN Napoli
- INFN Roma1
- Istituto Superiore di Sanità
- University of Messina
- University of Napoli

Foreign organizations involved

- ELI-Beamlines, Institute of Physics, (FZU), Czech Academy of Sciences
- Institute for Nuclear Problems, Belarusian State University
- University of Surrey

PI A. Sytov

Still under evalutation: decision soon

We can add collaborators later

How to use Marconi-100

First touch to Marconi-100

```
alexei@alexei:~$ ssh asytov00@login.m100.cineca.it
asytov00@login.m100.cineca.it's password:
Last login: Wed May 26 12:50:07 2021 from sytov1.fe.infn.it
*******************************************************************************
*  Welcome to MARCONI100 Cluster /                                            *
*            IBM Power AC922 (Whiterspoon) -                                  *
*            Red Hat Enterprise Linux Server release 8.1 (Ootpa)              *
*                                                                             *
*  980 compute nodes with:                                                    *
*        - 2x16 cores IBM POWER9 AC922 at 3.1 GHz                             *
*        - 4 x NVIDIA Volta V100 GPUs, Nvlink 2.0, 16GB                       *
*        - 256 GB RAM                                                         *
*                                                                             *
*  Disk Space: 8PB GPFS storage                                               *
*                                                                             *
*  Internal Network: Mellanox Infiniband EDR DragonFly+                       *
*  SLURM 20.11                                                                *
*                                                                             *
*  For a guide on Marconi100:                                                 *
*  https://wiki.u-gov.it/confluence/pages/viewpage.action?pageId=336727645    *
*  For support: superc@cineca.it                                              *
*******************************************************************************
*  This system is in its complete configuration and is in full-production     *
===============================================================================
IN EVIDENCE:
 - You can find the installed software on modules organized with different
   profiles. Use the "modmap" command to identify the correct profile for
   your module ("modmap -h" for help)
 - spack/0.14.2 is installed. Spack is a package management tool designed to
   support multiple versions and configurations of software on a wide variety
   of platforms and environments.
 - On March 24th, 2021 the cluster has been updated to OS RedHat 8.1 .
   You may need to recompile your code if you experience any problem that
   wasn't present before the upgrade.
===============================================================================
```

account	start	end	total (local h)	localCluster Consumed(local h)	totConsumed (local h)	totConsumed %	monthTotal (local h)	monthConsumed (local h)
IscrB_LEADER	20201014	20211014	1500000	468084	468084	31.2	123287	74635

```
(base) [asytov00@login01 ~]$
```

- Login:

ssh username@login.m100.cineca.it

- How much resources are you given:

  saldo -b

- Logout:

  logout

Download/upload your files

- **Download** data from Marconi-100:

scp username@login.m100.cineca.it:/m100_work/IscrB_LEADER/
/path_to_file_on_your_personal_mashine

- **Upload** data to Marconi-100:

scp /path_to_your_file_on_your_personal_mashine/file1 username@login
.m100.cineca.it:/m100_work/IscrB_LEADER/

- **Mount** your workdirectory at Marconi-100 to your file manager:

sshfs username@login.m100.cineca.it:/m100_work/IscrB_LEADER/
/path_to_mounted_directory_on_your_personal_mashine

- Do not forget **unmount** the directory or you may have system errors:

fusermount -u /path_to_mounted_directory_on_your_personal_mashine

- Your can also keep your files in $CINECA_SCRATCH and $WORK

  (all files older than ?? days are automatically deleted from CINECA_SCRATCH)

cd $CINECA_SCRATCH
cd $WORK

Module structure

- Load of the proper module (do it every time to avoid errors!)

module load gnu

- Module commands*:

COMMAND	DESCRIPTION
module av	list all the available modules
module load <module_name(s)>	load module <module_name>
module list	list currently loaded modules
module purge	unload all the loaded modules
module show <module_name>	print the env. variables set when loading the module

*From presentation at 12th Advanced School on Parallel Computing, 2016, Bologna

- Install your soft in $WORK directory:
  download your soft: wget ...link…
  upload necessary modules: module load …
  install your soft
  upload necessary modules when launching your codes

Software at Marconi-100*

Two **modes** of work:

- with terminal via **ssh**
- **Remote visualization** via RCM (Remote Connection Manager)

Supported **compilers** (Fortran, C, C++):

- **xl**
- **hpc-sdk**
- **gnu (I recommend this)**

Compilers for **MPI** : spectrum_mpi

Galileo-100 (Intel CPU cluster) main difference: **Intel Compiler**. **My experience**: speed up with a factor **2** w.r.t. **GNU**

Scientific libraries examples

- cuda; python; TensorFlow; Keras; Qiskit, Blas; LAPACK; FFTW; …

Installed software examples

- QuantumEspresso (Solid state physics); ...
- **Visualisation software**: Paraview; MatPlotLib; ...

*https://www.hpc.cineca.it/softwares/All%20software

SLURM job script (only CPU cores) for Macroni-100

```
#!/bin/bash                                  Just write it
#SBATCH --time 23:54:00                      Requested time of calculations
#SBATCH --nodes=1                            Requested number of nodes;
#SBATCH --ntasks-per-node=128                Number of cores per node;
#SBATCH --mem=64GB                           Number of RAM per node;
#SBATCH --partition=m100_usr_prod            Queue name
#SBATCH --job-name=Crystal51                 Job name (to choose any name)
#SBATCH --err=myJob.err                      File for runtime errors
#SBATCH --out=myJob.out                      File for output
#SBATCH –account=IscrB_LEADER                Our account name on Marconi-100
#SBATCH –mail-type=ALL                       Instruction for email
#SBATCH –mail-user=youremail@gmail.com       Your email

cd $SLURM_SUBMIT_DIR                         enter to your code directory
module load gnu/8.4.0                        load compilers
module load cmake
Source $WORK/Geant4new/Geant4/geant4.10.05.p01MT-install/bin/geant4.sh
srun ./mainMCMC run1.mac                     Launch your code
```

Run your program

sbatch runmyprogram.sh ← Start calculations;

runmyprogram.sh is a file with your SLURM job script

squeue -l -u $USER ← see all your currently running jobs or jobs in the queue, elapsed time and ID of your job

squeue -l ← see all the queue

scancel 1234567 (Marconi-100) ← Abort evalution; instead of 1234567 write the proper job ID from squeue -l -u $USER

sinfo -l ← resource information

Important advices:

Do not ask too much nodes, otherwise you may wait in the queue too long.

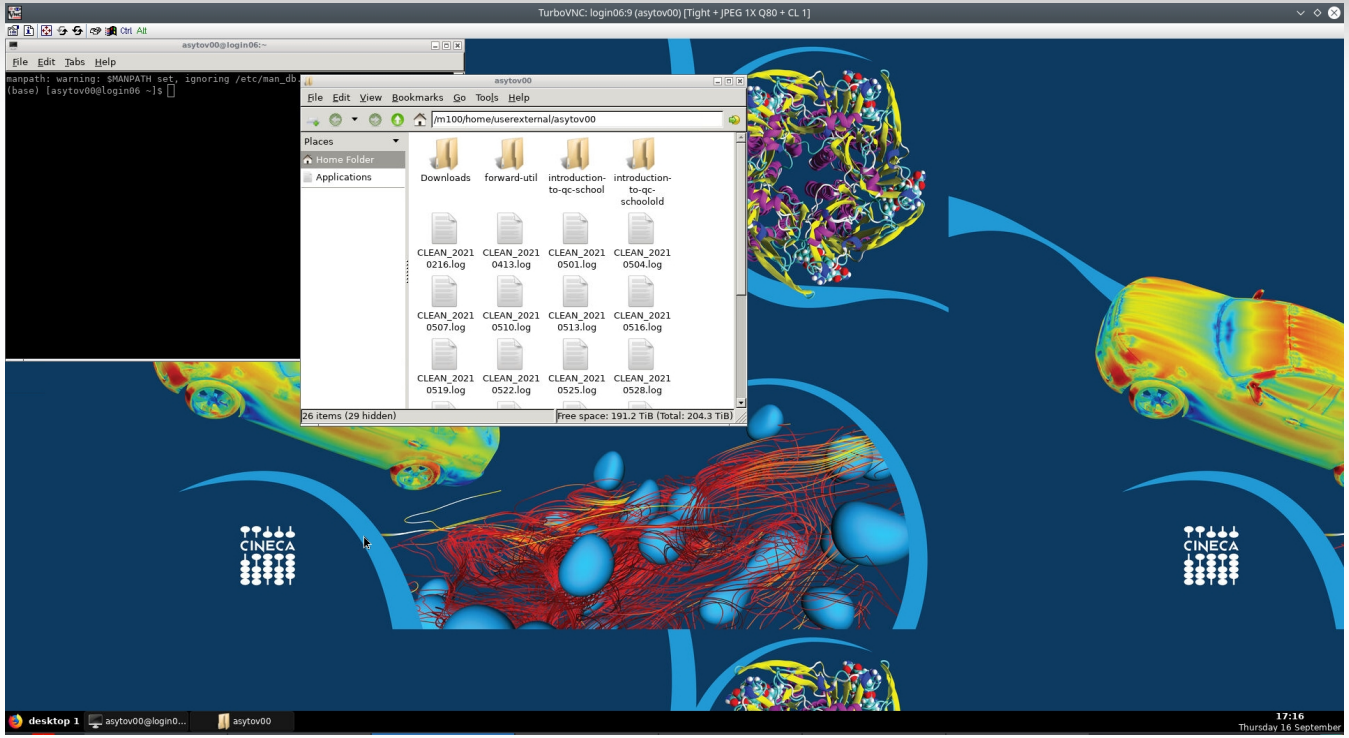
Do not ask too much time, you may wait too long also. Max time per job is 24 hours.

Try to run your application for a small but time scalable problem using 1-2 nodes to estimate the time for your calculations. Ask 20-30% more time for your job than you need.

Be careful with your data amount generated. The limit is 2Tb.

Do not hope to find a free supercomputer on weekends and holidays!

Remote visualization via RCM



Advantages:

- Debug and profile your code in studio
- Analyze your simulated data on supercomputer
- Visualize your graphics (possibly Geant4 UI)
- Use Jupyter Notebook for Python in Firefox browser

Useful links and documentation

MPI and OpenMP examples: http://www.hpc.cineca.it/content/online-training

Supercomputers in Cineca: http://www.hpc.cineca.it/content/hardware

Marconi-100 in general: https://www.hpc.cineca.it/hardware/marconi100

Marconi-100 reference guide:
https://wiki.u-gov.it/confluence/display/SCAIUS/UG3.2%3A+MARCONI100+UserGuide

Batch Scheduler SLURM:
https://wiki.u-gov.it/confluence/display/SCAIUS/UG2.6.1%3A+How+to+submit+the+job+-+Batch+Scheduler+SLURM

Remote visualisation at Marconi-100:
https://wiki.u-gov.it/confluence/pages/viewpage.action?pageId=358200249

Software installed in Cineca: http://www.hpc.cineca.it/content/application-software-science

Training in Cineca 2021: https://www.hpc.cineca.it/content/training

Do not trap in infinite loop at supercomputer!
Thank you for attention!