PRACE-5IP

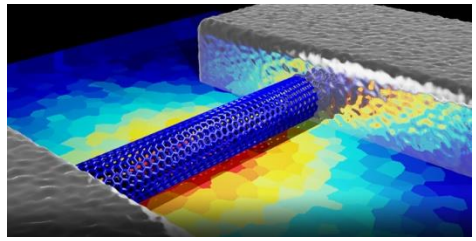
Link with Large Scale Scientific Instruments

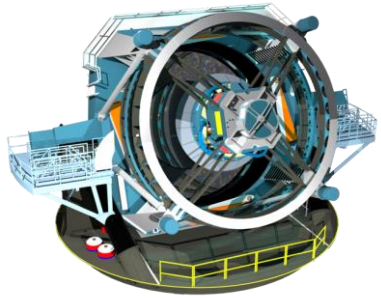
CERN, October 22nd, 2018

(on behalf of F. Suter)

Motivations

- Science discoveries driven by data produced by instruments
- Results require HPC for
  - Post-processing
  - Analysis
  - Visualization
- Associated challenges
  - Data storage and transfer
  - Access to HPC resources as an institute





Outcomes from PRACE-4IP

- Interactions with 5 LSSIs
  - NSC – UiO (DNA sequencers)
  - ESRF – CaSToRC (Synchroton)
  - LHC – NCSA (Particle physics)
  - ELI-ALPS – NIIF (Laser)
  - LSST – CC-IN2P3 (Telescope)
- Focus on **data transfer** (small files, bulk, orchestration, …)
- Open issue: **‘institutional’ access’** to PRACE HPC infrastructure
  - Asked by **three out of five** instruments
  - To **offload** data analysis work

Objectives for PRACE-5IP

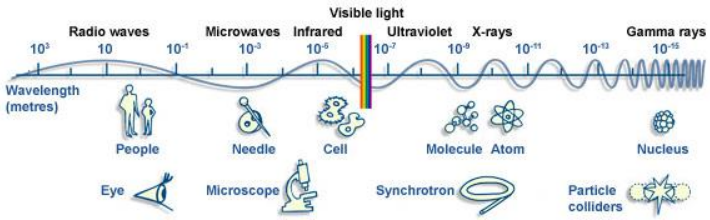
- Address the question of an **institutional access**
  - Usually do not qualify for PRACE Project Access calls
  - Computational work is nonetheless decisive for research
  - Propose **pilot use cases** to identify potential issues
  - Define a **formal PRACE – LSSI collaboration framework**
- Postpone the technical issues
- Select two LSSI
  - ESRF
  - **CERN**

Contributing partners

- CNRS / IN2P3 Computing Centre
  - (CC-IN2P3, **Frédéric Suter**, France)
- National Centre for Supercomputing Applications
  - (NCSA, Nevena Ilieva-Litova, Bulgaria)
- Computation-based Science and Technology Research Centre
  - *(CaSToRC, Andreas Pantelli, Cyprus)*

ESRF: European Synchrotron Radiation Facility

- Shoot brilliant X-ray beams at matter
  - 10^{13} times brighter than hospital X-ray source
  - Study atomic scales: imaging, composition
- Applications: material science and engineering, chemistry, medicine, condensed matter physics, biology
- Users travel to Grenoble, perform experiment, gather data, analyze it partly at home, partly at source



Work done with ESRF

- Preparatory meeting on 26th July 2017
- Joint PRACE – GEANT – ESRF on 11th September 2017
  - Definition of what would be a first pilot
  - Discussion on how to access PRACE resources
  - Define a roadmap for subsequent pilots and further collaboration

Call for Volunteer Tier-1 centres

- Published in November 2017
- Needed resources
  - Up to 256 cores
  - For 2-3 months (from January 2018)
  - ~500M core.hours
- 5 sites answered (Poland (2), France, Cyprus, Germany)
  - Two were selected by ESRF
  - PNSC (HPC Cluster) and MPCDF (GPU cluster)

First outcomes of the PRACE-ESRF collaboration

- ESRF leveraged the PRACE specific resources
  - No need for GPUs or specific hardware on site
  - Scalability tests made possible
- Real benefit of the user support from PRACE centres
  - Porting codes and adapting the environment to the codes
- Work to be done
  - Enable production ready pipelines
  - Automate the offloading to PRACE resources

Work done with CERN

- Preliminary meeting on September 21st, 2017
  - With CERN-IT, WLCG, and experiments
- Expression of needs and expectations (on both sides)
  - Not only resources, but also (on-demand) training
- Identify an opportunity for a concrete collaboration

Next steps with CERN

- Definition of what could be a first pilot
  - Leveraging PRACE specific resources
- Detail the needs for training
  - Help CERN experiments to benefit of HPC resources
- Discussion on how to access PRACE centres
- Identify **actors** and define **roadmap** for further collaboration