The SCIGNE Platform: Services for Open Science

Jérôme Pansanel

cnrs

IPHC
Institut Pluridisciplinaire Hubert CURIEN
STRASBOURG

Université de Strasbourg

The SCIGNE Platform

In a few words

- SCIGNE – a platform offering all-terrain computing and storage services hosted at IPHC
- Managed by a team of engineers providing support to researchers for managing and analysing large amounts of data
- Dedicated support in several scientific fields (physics, chemistry, biology and ecology); open to many communities through scientific collaboration
- Launched in 2007 with the Grid Computing service for ALICE and CMS
- Since 2011, completed with the Cloud Computing service (server and kubernetes as a service), as well as the data management service (iRODS)
- Labeled by IN2P3 in 2017 and by the University of Strasbourg in 2020 (CORTECS)
- Scientific Commitee is shared with the HPC center of the University
- Involved in several national and international scientific projects

https://scigne.fr

The Team

A team with many skills

- 8 highly-skilled engineers
- 4,3 FTEs
- Expertises:
  - Processing and analysis of large amounts of scientific data
  - Computation workflow management
  - Computation reproductility studies
  - Data and software management plans, making the data FAIR
  - Software development, source code opening
  - Building of container apps
  - GPU-enabled software development, artificial intelligence
  - IT security
  - Green computing
  - Network and infrastructure

High-Throughput Computing (HTC)

Resources

- 5000 cores and 4.0 PB storage
- Direct access for local users to the batch scheduler
- CentOS 7-based Linux environment and singularity containers
- Service availability > 99 %
- Interconnected with the European Grid Infrastructure (EGI) at 80 Gb/s

Projects

- Involved in IN2P3 projects (computeOps, DOMA, LCG France)
- Reproductible build of physics software with GUIX (https://guix.gnu.org)
- Co-management of the Biomed virtual organisation
- Partner of BELLE 2, EGI, France Grilles and WLCG
- Leading the technical team of France Grilles, the French grid representative

Server as a Service (Cloud Computing)

Resources

- Large VMs for hosting compute and virtual research environments (up to 128 cores and 1 TB RAM)
- > 1k cores, 6 TB RAM and 1024 TB disk storage
- > 7M CPU-Hours provided in 2022
- Dedicated and isolated network for enhanced security
- GPU and Kubernetes as a Service
- Availability > 99 %

Projects

- Member of the EGI FedCloud technical group
- Participating to the *HEPIX Benchmarking Group*
- Partner of BELLE 2, EGI, France Grilles, IFB (ELIXIR), INRAE and WLCG
- Funded H2020 projects: FAIR-IMPACT, FAIR-EASE, SIESTA, GREENDIGIT and ECHOES

Data Management

Resources

- 200 TB
- Availability > 99 %
- Based on CEPH and iRODS technologies

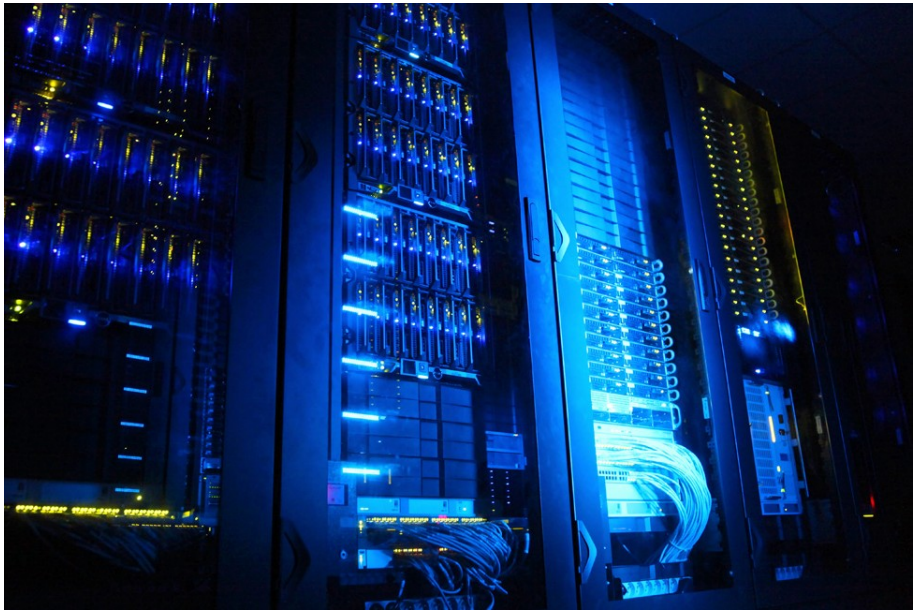
Projects

- Involved in several projects to share our knowledge and skills on iRODS (Data Terra, MesoNET, etc)
- Partner of the France Grilles iRODS distributed infrastructure (FG-iRODS)
- Looking at machine actionable DMP
- Organisation of training sessions
- On-going work for joining identity federation (EduGain)

User services

Services for users

- Training (DIRAC, OpenStack, Docker, iRODS, ...)
- Help with using computing and storage infrastructures (launching production, fixing issues, ...)
- Writing reusable technical documentation (Open License)
- Advice on data management and organisation
- Involved in the local « Atelier de la Donnée », as well as some national WG
- Custom software installation
- Software development and parallelisation
- Analysis and advice on architectural choices
- Server and service hosting



Open Science

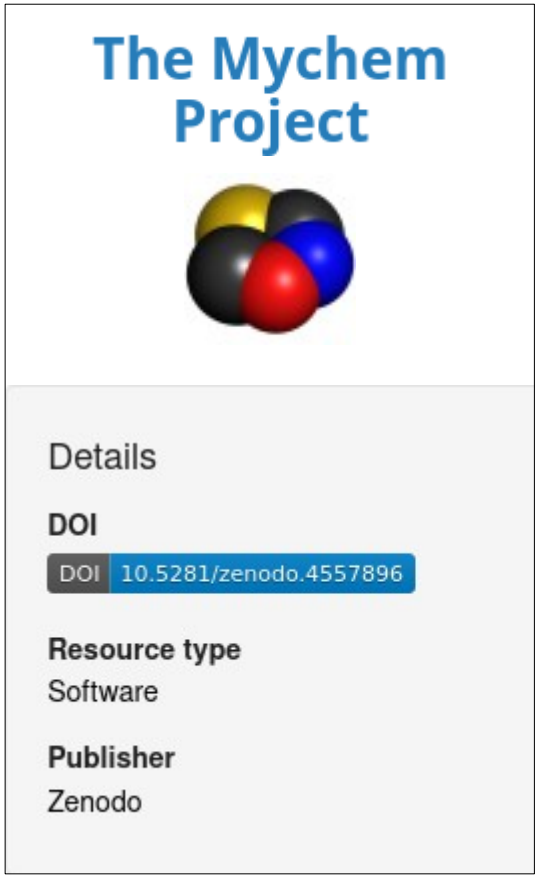
Open Science is part of our DNA

- Maintaining list of Free and Libre Open Source chemistry software since 2001
- Promoting Open and FAIR Data
- All documentation and training material are made freely available with Creative Common licence (https://creativecommons.org/)
- All software development are released under Open Source license (GPL, Apache v2, BSD)
- Involved in the Open Science Steering Committee and the *Atelier de la Donnée* of the University of Strasbourg
- Involved in Open Science technical WG at the French and European levels (France Grilles, EOSC, ...)

Example

Open Source Software

- Mychem – a chemoinformatics extension for MySQL
- C/C++ software that provides a set of functions for handling chemical data within MySQL and MariaDB
- Based on Open Babel, a well-known Open Source software in chemoinformatics
- Released under GPL v2+
- Source code available on GitHub: https://github.com/mychem/mychem-code
- Published on Zenodo: https://zenodo.org/records/4557896
- Software Management Plan openly accessible on DMP OPIDoR: https://dmp.opidor.fr/plans/5940/export.pdf



Want to give a try?

Contact us: scigne@iphc.cnrs.fr