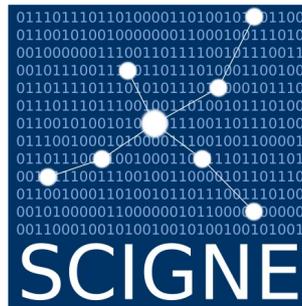


# La plateforme SCIGNE



J. Pansanel

[jerome.pansanel@iphc.cnrs.fr](mailto:jerome.pansanel@iphc.cnrs.fr)

3 juin 2024





# La plateforme SCIGNE

## En quelques mots

- Une plateforme Calcul et Données hébergées à l'IPHC (campus de Cronenbourg, Strasbourg)
- Services et accompagnement pour le traitement et la gestion des données de plusieurs domaines scientifiques (physique, chimie, biologie et environnement)
- Issue de la consolidation des services de la grille de calcul, du Cloud Computing et du service de gestion de données iRODS
- Labellisée par l'IN2P3 en 2017 et par l'Université en 2020 (Cortecs)
- 8 ingénieurs / 4,3 FTE
- Conseil scientifique commun avec le mésocentre de l'Université de Strasbourg
- Impliquée dans plusieurs projets nationaux (PIA, IR) et internationaux (EGI, EOSC, ELIXIR, etc)

→ <https://scigne.fr>

# L'équipe

## Une équipe, des compétences

- 8 personnes du service informatique impliquées
- 4,3 FTEs
- Expertises :
  - Traitement et analyse des données scientifiques massives
  - Reproductibilité des calculs
  - Plan de gestion des données, FAIRisation
  - Bonnes pratiques en développement logiciel, ouverture des codes
  - Gestion des workflows de calcul
  - Conteneurisation des applications
  - Développement GPU, intelligence artificielle
  - Sécurité des systèmes et des applications
  - Étude des impacts environnementaux
  - Réseau et infrastructure

# Calcul à haut-débit (HTC)

## Ressources

- 5000 cœurs et 3,0 Po de stockage
- Accès direct pour les utilisateurs locaux (PBS)
- CentOS 7 avec possibilité d'utiliser des conteneurs singularity
- Migration Alma 9 en cours
- Disponibilité > 99 %
- Accessible à travers différentes VOs
- Interconnecté à la grille de calcul EGI à 20 Gb/s

## Projets et responsabilités

- Projets IN2P3 (DOMA, DIRAC)
- Implication dans la VO Biomed (shifts)
- Partenariat BELLE 2, EGI, France Grilles et WLCG
- Direction technique de France Grilles



# Serveur à la demande (Cloud Computing)

## Ressources

- Hyperviseurs récents (jusqu'à 128 cœurs et 1024 Go de RAM)
- ~ 1600 cœurs, 6 To de RAM et 1024 To de stockage
- > 7 millions d'heures de calcul consommés en 2023 (78 % d'occupation)
- Réseau dédié et isolé
- GPUs (France Grilles, INRAE)
- Disponibilité > 99 %

## Projets et implications

- Développement des sondes de supervision pour France Grilles
- Expertise CEPH, Docker / Kubernetes, sécurité OpenStack
- Hébergement du projet CSAN
- Partenariat BELLE 2, EGI, France Grilles, IFB (ELIXIR), INRAE et WLCG
- Projet H2020 EGI-ACE, FAIR-IMPACT et FAIR-EASE, SIESTA, GreenDigit et ECHOES



# Gestion des données

## Ressources

- 200 To utiles
- Disponibilité > 99 %
- Basé sur les technologies CEPH et iRODS

## Projets et implications

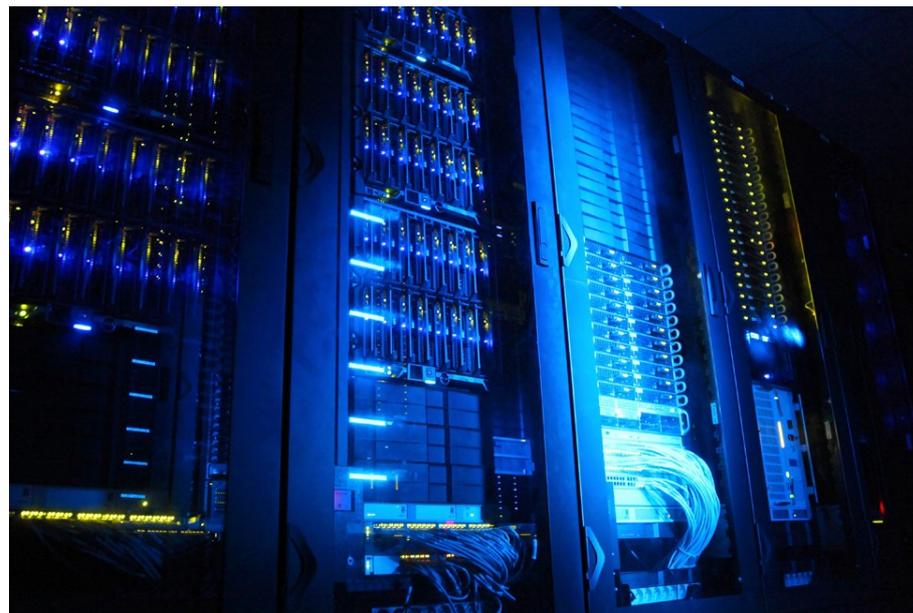
- Développement des sondes de supervision pour France Grilles
- Partenaire de la plateforme distribuée FG-iRODS et de l'IFB
- Expertise iRODS et gestion des *workflows* de données
- Travaux en cours sur l'utilisation d'OpenID
- Implémentation de PGDs
- Organisation de formations iRODS (~ 80 personnes en 2021)
- Partenariat Data Terra, IFB et Mesonet
- Évaluation OneData

011011101101000011010010 01100  
0110010100100000001100010111010  
001000000110011011100101110011  
0010111001 01011010101100100  
0110111011101010110 000101110  
0110110111000100101110100  
0110010100101 11001101110100  
01110010001000000001100001  
01101110 00100011 110110101  
01 11001100100110000101101110  
011001000110100101110110100  
001010000011000000111000 0000  
00110001001010010010100101001

# Services aux utilisateurs

## Services proposés aux utilisateurs

- Formation (DIRAC, OpenStack, Docker, iRODS, ...)
- Aide lors de l'utilisation des infrastructures de calcul et de stockage (lancement de production, résolution des problèmes, ...)
- Rédaction de documentation technique, documentation en ligne (licence libre)
- Conseil sur la gestion et l'organisation des données
- Implication dans l'atelier des données ADELE
- Installation de logiciels
- Développement logiciel et aide à la parallélisation de logiciels
- Analyse et conseil sur les choix d'architecture
- Hébergement de services et de serveurs



## Au coeur de notre ADN

- Maintenance d'une liste de logiciels libres pour la chimie depuis 2001
- Promotion des licences libres et des données FAIR
- Toutes les documentions et supports de cours sont diffusés sous licence Creative Commons (<https://creativecommons.org/>)
- Tous les logiciels développés sous licence libre (Apache v2, BSD ou GPL)
- Implication dans l'atelier des données du site strasbourgeois (ADELE) et dans le comité de pilotage de la Science Ouverte de l'Université de Strasbourg
- Implication dans des groupes de travail en lien avec la Science Ouverte au niveau national et européen

011011101101000011010010 01100  
01100101001000000011000100111010  
001000000110011011100101110011  
00101110011 01011010101100100  
01101110111010101110 000101110  
01101110111000010100101110100  
01100101001011 11001101110100  
0111001000101000100001100001  
01101110 00100011 1101101101  
00 11001100100110000101101110  
0110010001101001011011100110100  
0010100000110000001011000 0000  
00110001001010010010100100101001

# Questions ?