

# Université Paris-Saclay

## Mésocentre et Hébergement

8 mars 2021



1. L'Université Paris-Saclay
2. Construction du mésocentre de l'Université Paris-Saclay
3. Hébergement –Projets structurants

*Remarque liminaire : l'utilisation du masculin est appliquée quand il n'est pas possible d'utiliser un nom épicène, ceci sans aucune discrimination et dans le seul but d'alléger le texte*

POINT 1

# L'Université Paris-Saclay

e

# L'Université Paris-Saclay



## Organisation générale



## Quelques chiffres

### FORMATION ET RECHERCHE

**48 000**  
étudiants  


**24 000**  
en premier  
cycle



**12 000**  
en master

**4 600**  
en doctorat

**18** graduate schools  
et institut  


**8 100**  
chercheurs et  
enseignants-chercheurs

  
**13 000**  
publications par an

**8 500**  
administratifs et  
personnels techniques

  
**275**  
LABORATOIRES

  
**13%** de la  
recherche  
française

**500**   
plateformes  
expérimentales

### INNOVATION

  
**100**  
startup créées  
PAR AN

**1** réseau de  
fablab

**6** incubateurs

**1** société d'accélération  
de transfert de  
technologie  


**7 000**   
étudiants sensibilisés  
à l'entrepreneuriat  
par an

**1** Centre de  
Design



## Graduate Schools

Chimie	Computer Science		Métiers de la Recherche et de l'Enseignement Supérieur	Éducation, Formation, Enseignement
Géosciences, Climat, Environnement, Planètes	Mathématiques	Droit	Institute for the Sciences of Light	Biosphera
Physique	Engineering and Systems Sciences	Humanités – Sciences du Patrimoine	Economics & Management	Life Sciences and Health

# L'Université Paris-Saclay

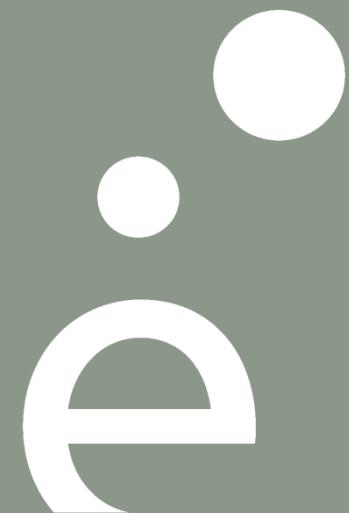


A la pointe de la recherche mondiale...



POINT 2

# Construction du mésocentre de l'Université Paris-Saclay



# Construction du mésocentre de l'Université Paris-Saclay



## Moyens actuels

- Simulation HPC - Moulon@PS, LabIA+FactoryIA, VirtualData
- Cloud : ~10000 cœurs (forte composante astro-cosmologie)
- GPU : plus de 50 nœuds pour plus de 120 cartes
- Stockage : ~ 2 250 To
- CPU MPI : ~ 12 000 cœurs
- CPU OpenMP-Sequentiel : ~12 000 cœurs – mémoire max 1,5 T
- Réseau : Infiniband 10 et 56 Gbit/s, Omni-Path 100 Gbit/s

## Objectifs du mésocentre

- Plateforme dédiée aux utilisateurs des communautés scientifiques
- Moyens mutualisés pour répondre à tous les besoins
- Coordination et animation scientifique par les utilisateurs
- Administration technique et système s'appuyant sur les acteurs
- Pilotage par les composantes et les établissements composantes

# Construction du mésocentre de l'Université Paris-Saclay



## Schéma global

### Communautés scientifiques Unités de recherche Formation

Coordination  
Scientifique

Administration  
Système

Comité Technique  
(utilisateurs)

### Moyens mutualisés

GPU / Stockage / CPU MPI / CPU OpenMP-Sequentiel

➡ Réseau dédié ➡

Hébergement sécurisé sur 2 sites proches

### PILOTAGE

Comité des tutelles

Composantes/Établissement Composantes

# Construction du mésocentre de l'Université Paris-Saclay



## Comment ?

- Mise en commun de moyens humains, matériels et logiciels sur le calcul scientifique intensif et le stockage massif de données
- Rationalisation des moyens et des outils d'analyse de données
- Mise en commun de bonnes pratiques, expertise de la Maison de la Simulation
- Rationalisation de l'accompagnement des chercheurs
- Proximité (indispensable) entre les chercheurs et le personnel du mésocentre

## Avantages

- Capacité de formation des chercheurs, bénéfice aussi pour les étudiants
- Contribue à une organisation optimale du supercalcul en France (MésoNet)
- Élément fort de la stratégie scientifique de l'Université : renforce la coopération entre ses composantes et établissements composantes
- Meilleure gestion des données et des codes (utilisation systématisée de Data Management Plans, etc...) dans un contexte de Science Ouverte

# Construction du mésocentre de l'Université Paris-Saclay



## Financement

- Près de 6 M€, dont 3,7 M€ CPER, Sésame IdF, CNRS, Université Paris-Saclay, DIM ACAV, CEA, IGN, IMT, ISCPPIF...
- Plusieurs centaines de k€/an d'apport : ETP (coordination, animation, administration), et hébergements (VirtualData et IDRIS)

## Utilisateurs

- Potentiellement toutes les unités de recherche de l'Université Paris-Saclay
  - Lien fort avec la Faculté des Sciences d'Orsay, CentraleSupélec, ENS Paris-Saclay, CEA)
  - Mais aussi Pharmacie, Médecine, Droit et Gestion, STAPS...
- Plusieurs centaines d'utilisateurs actifs

# Construction du mésocentre de l'Université Paris-Saclay



## Organisation en cours de construction

- Site web commun
- Charte unifiée
- Formulaire de demande d'accès unique
- ...

**université  
PARIS-SACLAY**

**Charte d'utilisation du mésocentre  
Paris-Saclay**

**Présentation et objectif**  
Le mésocentre Paris-Saclay est mis en place pour faciliter la collaboration entre les établissements de recherche et les utilisateurs en relevant par exemple au profit d'un code de calcul interne et commun normes et méthodes utilisées dans le cadre d'un projet.

**Utilisation et conditions de travail**  
utilisation de ressources renouvelables et régulières. Le comité d'application et les délégués doivent faire preuve de volonté et responsabilité dans leur utilisation et leur dépense.

**Financement**  
Des ressources sont prévues par l'Université Paris-Saclay pour l'usage des plateformes et des services de soutien à la recherche et à l'innovation. Les utilisateurs peuvent également bénéficier d'aides financières via les partenaires et les fonds de recherche.

**Utilisateurs éligibles**  
Les utilisateurs éligibles sont les membres de l'Université Paris-Saclay et d'un responsable scientifique et d'un ou plusieurs administrateurs. Les utilisateurs peuvent bénéficier de certains avantages si bien qu'ils ont des ressources informatiques favorables.

**Qualité de service**  
En dépit d'une politique de best effort de la part de toutes les organisations et de disponibilité des différents services en vise de qualité.

**Assistance**  
Les utilisateurs peuvent solliciter de l'aide pour résoudre leurs demandes via l'interface de support technique ou par e-mail à l'adresse support@mesocentre-paris-saclay.fr.

**Finalité**  
Les utilisateurs de l'interface s'engagent à faire une utilisation des ressources numériques fournies dans le cadre des recherches scientifiques (encadrées et supervisées) et pédagogiques de l'Université Paris-Saclay.

**Données à caractère personnel/sensible**  
Les plateformes du mésocentre Paris-Saclay ne sont actuellement pas labellisées pour héberger des données sensibles (type de données de santé ou des données identifiantes comme de faux noms). Des demandes d'autorisation pour l'usage d'application sont invitées à contacter le comité de gestion du mésocentre pour être prise en compte dans leur fonctionnement.

**Confidentialité**  
Les utilisateurs du mésocentre s'engagent à respecter les politiques de contrôle d'accès mises en place par l'Université Paris-Saclay et les plateformes. Il est permis à accéder une annexion des données appartenant à des personnes auxquelles il ne sont pas destinées, ou autrement à des pages auxquelles il ne participent pas.

MesoUPSSaclay Working Groups Services Support News

**université PARIS-SACLAY** **école supérieure paris-saclay**

**CentraleSupélec** **CNRS**

**Mesocentre Paris Saclay**

**Plateformes**  
Le mésocentre regroupe **3 plateformes** techniques majeures de calcul scientifique / traitement de données issues des établissements partenaires dans le projet Paris-Saclay.

**Accès**  
Les ressources de calcul sont accessibles à tout.e chercheur.se rattaché à un laboratoire appartenant au périmètre de l'**Université Paris-Saclay** conformément à la **charte du mésocentre**. Les demandeur.es sont invité.e.s à renseigner le **formulaire de demande d'accès** et à l'adresser à mesocentre@universite-paris-saclay.fr.

**Projets**  
Les études réalisées sur les différentes plateformes du mésocentre vont de la physique de l'infiniment petit aux sciences humaines et sociales.



- **Spécificités**
  - Ressources de calcul performantes et homogènes
  - Environnement logiciel pour le développement (outils de compilation, de débogage et de profiling)
  - Visualisation à distance
  - Support de proximité
- **Volontés**
  - amener des nouveaux acteurs au calcul scientifique
  - stimuler le développement de méthodes numériques et d'algorithmes
  - proposer un lieu de mise au point des codes pour le passage à l'échelle
- **Ressources humaines**
  - coordination scientifique : **2 EC à temps partiel (0.3 ETP)**
  - administration de la machine, gestion de l'environnement logiciel, support aux utilisateurs, accompagnement des projets : **3 ingénieurs calcul et 2 ingénieurs système (2.5 ETP)**
  - correspondants au sein des laboratoires : **1 par laboratoire impliqué dans le mésocentre** (contact utilisateurs, validation des projets)

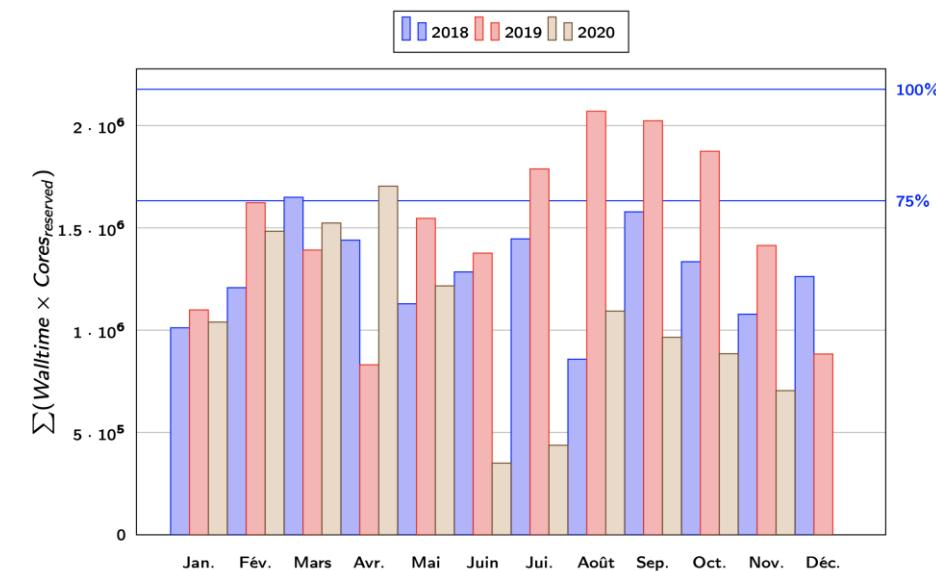


# Mésocentre : Simulation HPC - Moulon@PS

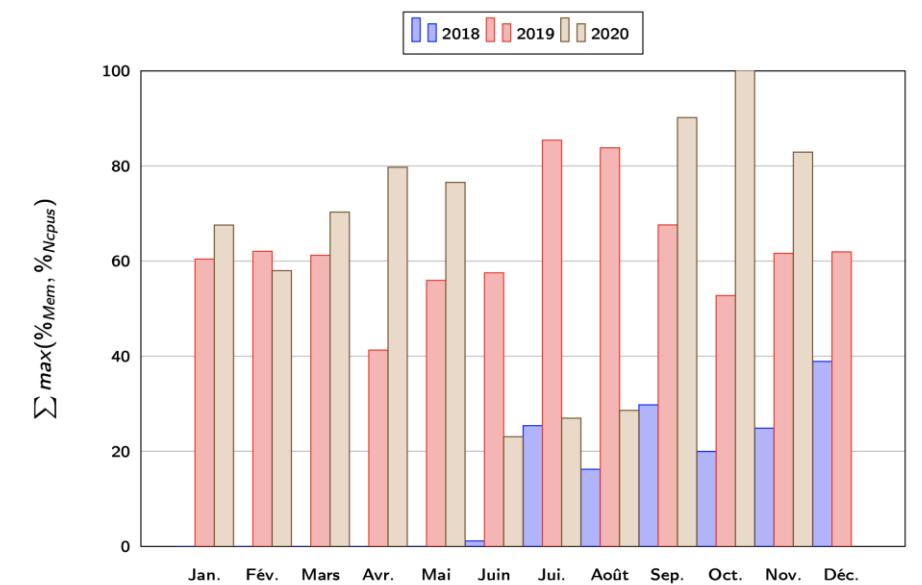


- **Financement CPER** : ~3,7 M€
- **Apport des établissements** (hébergement (IDRIS), maintenance du matériel, licences logiciels, RH) : plusieurs centaines k€/an
- **Utilisation** : 100 à 150 utilisateurs actifs/mois (plus de 25 unités de recherche), > 600 utilisateurs continus

Exploitation **fusion** : occupation des nœuds ICEX



Exploitation **fusion** : nœuds large mémoire



# Mésocentre : LabIA + FactoryIA



- Deux grappes de calcul dédiées à la recherche en Intelligence artificielle
  - **LabIA** - CNRS : plateforme de recherche ~ 35 cartes GPU, 240 To de stockage (*BeeGFS*), hébergée à l'IDRIS/CNRS (Orsay)
  - **FactoryIA** - CEA: plateforme à vocation mixte recherche + projets industriels, ~150 cartes GPU, 80 To de stockage (SSD), hébergée au TGCC (Bruyère le Châtel)
- Investissement total 1.3ME : région IDF, CEA, CNRS, Univ. Paris-Sud, IMT
- Moyens humains: 2 éq. TP ingénieur (CEA, CNRS) + correspondants au sein des laboratoires impliqués (gestion comptes utilisateurs)
- Accès largement ouvert aux laboratoires de l'Univ. Paris-Saclay
- Vocation: mutualiser l'expertise et les ressources dédiées à la recherche en IA (traitement d'images, traitement des langues, traitement du signal, etc) ; préparer les utilisateurs (M2, doctorants, permanents) au passage à l'échelle sur l'infrastructure nationale

# Mésocentre : VirtualData...



- Des ressources/services mutualisées hébergés au sein d'un datacenter moderne, construit autour de l'efficacité énergétique
  - 51 racks disponibles (28 occupés) sur un potentiel de 90 racks
  - Conçu pour l'hébergement de racks denses (jusqu'à 30 kW/rack)
  - Double induction électrique pour le matériel critique (jusqu'à 300 kW)
- Mutualisation des services organisés autour d'une plateforme cloud
  - ~10000 coeurs
  - Développement récent financé par le DIM ACAV+
  - Quelques machines à grande quantité de mémoire (1,5 TB)
  - ~20 serveurs à 256 cœurs
  - Des utilisateurs commencent à financer du matériel en remplacement de leurs ressources de laboratoire
- Cloud VirtualData partie prenante de l'infrastructure nationale France Grille et Cloud
  - Réflexion en cours sur la façon de contribuer à l'EOSC

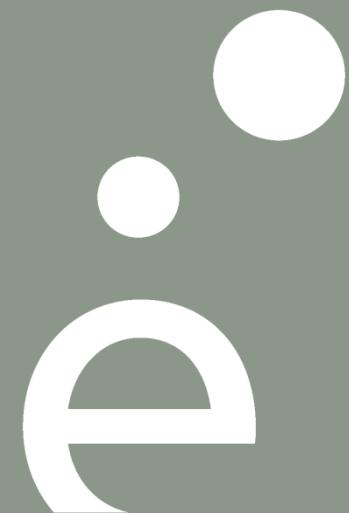
# ... Mésocentre : VirtualData



- 1 infrastructure de stockage distribué basée sur Ceph
  - Financé par le Labex P2IO pour les besoins de ses laboratoires mais potentiellement extensible à d'autres
  - Répartie sur 3 sites : Orsay, Palaiseau (X), Saclay (CEA/Irfu) pour permettre des stratégies de résilience quand c'est nécessaire
  - Usages multiples : object storage, block devices (~iSCSI), file system distribué
  - 1,5 PO utile en cours de mise en production : retard à cause des problèmes du réseau Paris Saclay avec le 100G
- Services avancées pour l'analyse de données : Spark et JupyterHub
  - JupyterHub : forte utilisation pour l'enseignement (~10 enseignants, O(1000) étudiants), un service en forte croissance
  - Cluster Spark à la demande : au coeur du projet Fink d'event broker pour LSST (élection attendue avant l'été), ouvert à d'autres usages pour le traitement de grande masse de données

POINT 3

# Hébergement – Projets structurants



## *2 lieux d'hébergement sur l'Université Paris Saclay*

- VirtualData (Univ. Paris-Saclay, Orsay Vallée) et IDRIS (CNRS, Plateau)
- Proximité géographique
- Connaissances mutuelles des équipes



## Pourquoi ?

- Accès direct des chercheurs aux experts applicatifs **et** aux experts des plateformes (accompagnement du mésocentre)
  - Impose que les ressources humaines du mésocentre soient localisées sur Orsay-Gif.
- Prendre en compte les besoins spécifiques concernant la qualité du service et l'arbitrage des ressources en cas de saturation
- Favoriser l'utilisation de plateformes et services mutualisés plutôt que de déployer des solutions communautaires spécifiques



## Besoins - Ressources existantes

- Certaines unités, gros consommateurs de calcul, affectent déjà un de leurs ingénieurs à l'accompagnement de leurs chercheurs pour utiliser le mésocentre,
- 1 ingénieur dédié au support de l'informatique scientifique à la DSI de l'université Paris Saclay (ex UPSud)
- Possibilité que d'autres laboratoires, moins gros consommateurs, se regroupent pour permettre à court terme le recrutement d'autres ingénieurs sur cette fonction d'accompagnement,
- Pérennisation du CDD Mésonet par l'Université Paris-Saclay renforcera la dynamique en cours. Possibilité de soutien RH d'autres tutelles ?
- Hébergement actuel : qualité de connexion – critique pour les utilisateurs (*expérience des prestataires durant le confinement pour la visio-conférence pas toujours « optimale »*) :
- Acteurs de l'hébergement = maîtrise de la redondance, sécurité des données...

## Importance des projets structurants

- ***Demande de CPER***
- Continuité de mutualisation – structuration et renforcement des moyens HPC/IA/DATA
- 1 infrastructure de stockage globale à l'ensemble des plateformes
- Hébergement redondant local
- Stratégie de l'Université Paris-Saclay pour accompagnement de la recherche et de ses besoins calcul et données
- ***Equipex+/PIA MesoNet (coordination GENCI)***
- Répondre aux besoins des chercheurs universitaires et industriels avec le développement d'équipements numériques structurants
- Infrastructure distribuée dédiée à la coordination du calcul scientifique
- Mésocentre Université Paris-Saclay : référent IdF, inscrit dans la cohérence nationale
- Financement : uniquement RH (avec pérennisation)

# CONCLUSION

e

# CONCLUSION



- **Mésocentre Université Paris-Saclay**
  - Fait partie de la stratégie de l' Université Paris-Saclay pour accompagner ses chercheurs
  - Ambition d'un service adapté aux besoins des chercheurs
- **Bénéficie d'investissements importants de développements significatifs depuis plusieurs années**
  - Bâtiments IDRIS et VirtualData)
  - Simulation HPC - Moulon@PS (Supercalcul)
  - Maison de la Simulation (Développements logiciels pour le supercalcul)
  - Cloud VirtualData

# CONCLUSION



- **Objectif : être opérationnel pour la rentrée/septembre 2021**, pour une ouverture **progressive** aux communautés scientifiques Université Paris-Saclay (calcul, hébergement,...)
- **Nouvelles demandes issue d'autres communautés pour utiliser le mésocentre** (e.g. Données Sciences du Patrimoine - pas développé ici)
- Renfort d'un CDD de 3 ans à venir financé par Equipex+/PIA « MesoNet », qui sera **pérennisé**

# CONCLUSION



- **Dynamique de mutualisation** (initiée séparément il y a 5/6 ans) : **nouvelle dimension dans le contexte de l'Université Paris Saclay** (complémentarité des plateformes antérieures)
  - **démarche exemplaire** de ce que la DGRI et les régions mettent en avant : des services mutualisés s'appuyant sur une infrastructure commune
  - **contexte particulier** : densité du tissu recherche
  - **appui sur des infrastructures maîtrisées** (optimisation des coûts d'exploitation)
- **Enjeu structurant de la demande CPER** pour soutenir cette dynamique de mutualisation.
- **Vague 3 transformation numérique du MESRI** : une opportunité pour intégrer des initiatives telles que la nôtre ?

# Backup Slides



# VirtualData (cloud) : ressources délivrées en 2020



## Cloud: utilisation 2020

