

# Mésocentre UPSaclay

Journée de lancement 06/12/2022

## Le projet

Le projet de mésocentre mutualisé de l'Université Paris-Saclay (UPSaclay) consiste à **construire un outil d'appui à la recherche (et selon les forces, à l'enseignement)** dans les domaines de la **simulation numérique**, de **l'hébergement de données**, du **traitement de données** et de **l'IA** au bénéfice des diverses communautés de l'Université Paris-Saclay.

Il s'agit de **mutualiser des moyens de calcul et de stockage** ainsi que **des moyens humains** autour d'un objet commun au service de la communauté scientifique.

Le projet se construit, dans un premier temps, à partir des plateformes **VirtualData (cloud)** (CNRS/IN2P3, UPSaclay), **Ruche** (ENS, CS) et **LabIA** (CNRS/INS2I, UPSaclay), avec l'accompagnement de la **Maison de la Simulation**.

Hébergement sur 2 sites: **IDRIS** qui héberge **Fusion/Ruche** et **LabIA** d'une part et le **bâtiment 206** qui héberge le cloud **VirtualData** d'autre part

# Objectifs

➔ Apporter une **infrastructure de calcul et de traitement de données**, un **support de proximité** et un **environnement scientifique interdisciplinaire** permettant le développement de **nouveaux outils de simulation**, de **nouvelles approches dans le traitement et l'analyse de données** et **le développement de la connaissance** dans un domaine très compétitif

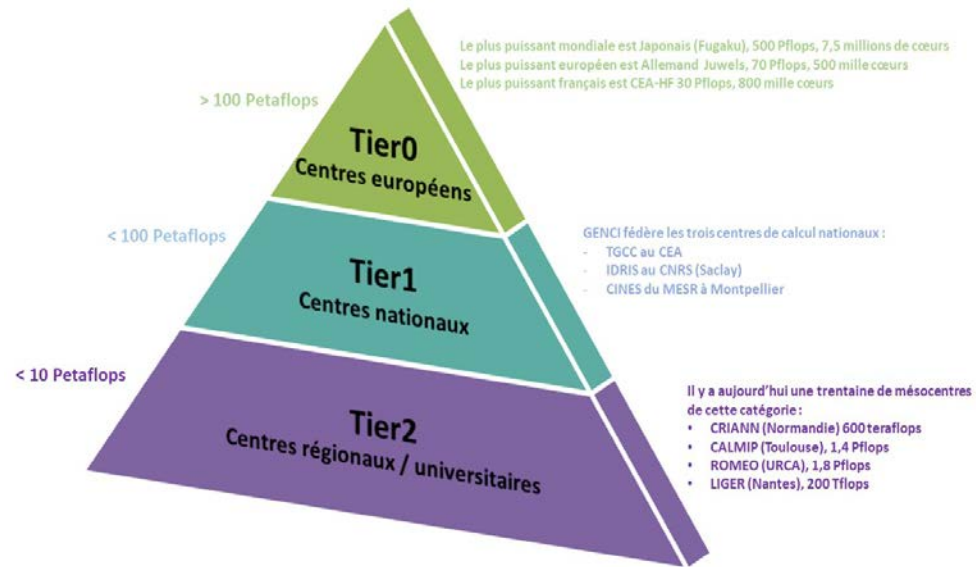
➔ **Mise en commun** et une **palette d'approches** du calcul numérique qui permet de choisir la meilleure combinaison d'outils au service des projets scientifiques, dans une **culture de Science Ouverte** homogène pour l'Université,

➔ Répondre à la nécessité de **mutualiser/rationnaliser les moyens de calcul** d'une part avec une **vision éco-responsable** (*économie d'échelle, d'extensibilité du système, politique d'achats groupés et d'optimisation des coûts de fonctionnement –dont consommation électrique-, ...*), **ainsi que les postes d'ingénieurs de haut niveau** (*mise en réseau de l'ensemble des ingénieurs et techniciens, échange de compétences et de savoir-faire,...*) d'autre part.

# Objectifs

- ➔ Permettre aux scientifiques de **se préparer aux environnements de production des niveaux supérieurs** (capacité de formation/accompagnement des chercheurs, ingénieurs, étudiants,...),
- ➔ Etre un pôle pour **l'animation scientifique, la formation** (projets pédagogiques pour enseignants et étudiants de l'Université, formation continue,...) et **la diffusion des bonnes pratiques** pour le calcul scientifique et la simulation numérique.
- ➔ Assurer une **redondance** entre les deux sites pour le volet de stockage des données,
- ➔ Offrir une **ouverture** vers le calcul scientifique et le traitement de données massives **à des disciplines encore neuves ou débutantes** dans ce domaine du fait de l'environnement scientifique et du support de proximité,
- ➔ Permettre à l'Université de **contribuer en nature à de grands projets collaboratifs nationaux et internationaux** (inscription dans un cadre national/organisation optimale du supercalcul en France: **MésoNet**),

Le positionnement des mésocentres **universitaires** ou **régionaux** se situe aujourd'hui sur une échelle de 5000 à 10 000 cœurs.



Positionnement de quelques mésocentres français

Le mésocentre d'UPSaclay, avec, seulement pour la partie Ruche par exemple, plus de **10.000 coeurs CPU** et plus de **20 noeuds GPU** (environ 650 TFlop/s CPU + 700 TFlop/s GPU), se situe au niveau Tier 2. VirtualData (cloud et pas HPC) double le nombre de cœurs.

# Genèse

A partir de:

- **Virtual Data** (financements Labex P2IO, DIM ACAV, tutelles,...)
- **Fusion/Ruche** (financements par CPER 2015-2020, tutelles, CEA,...)
- **LabIA** (financement Sesame IdF, tutelles,...)

- demande financement **CPER 2021-2027**, déposée par E. Augé en 2020 (avec IDRIS)

## Le groupe de travail (créé début 2021)

- *M. Jouvin* (**VirtualData**), *P.-A. Boucard* et *R. Vicquelin* (**Fusion/Ruche**), *F. Yvon* (**LabIA**), responsables des trois principales grandes installations locales de calcul intensif et de stockage de données dans le périmètre d'UPSaclay,
- *E. Audit* et *M. Baboulin* (**Maison de la Simulation**),
- *M. Leoni*, représentant de la **DSI d'UPSaclay**,
- *G. Bernadat*, **chargé de mission du projet « mésocentre UPSaclay »** par la VP Recherche,
- *M. Guidal* (**Vice-Président Recherche adjoint Sciences et Ingénierie** UPSaclay), *B. Bourguignon/E. Trizac* (**Vice-Doyen Recherche de l'UFR des Sciences**), *E. Augé* (**Vice-Président Recherche adjoint Science Ouverte** UPSaclay), *P. Maitre* (**Vice-président Recherche ENS-PS**) et *O. Gicquel/P.-H. Cournède* (**Directeur de la Recherche CS**),
- et, plusieurs fois, *L. Bertrand* (**PPSM/DIM sciences du patrimoine**), *J.-Y. Berthou/P. Lubrano* (**INRIA**), *A. Refloch* (**ONERA**), *N. Sabouret/S. Haar* (**GS ISN**).

# Formalisation

Après passage au CoDiRE de février 2022, une lettre de mission a été rédigée pour former un comité de pilotage transitoire en charge de définir les bases de fonctionnement du mésocentre.

➡ Document rédigé de 14 pages par le comité de pilotage avec propositions, dans l'optique d'une création du mésocentre au 1<sup>er</sup> juillet 2022.

➡ Validation au CoDiRE de juin 2022

- responsables de Fusion/Ruche, Virtual Data et Lab-IA : P.-A. Boucard et R. Vicquelin, M. Jouvin et F. Yvon,
- directeur de la MdS : M. Baboulin,
- directeur de la DSI de l'UPSaclay ou son représentant,
- représentant de la VP Recherche de l'UPSaclay : M. Guidal,
- VP adjoint Recherche en charge de la Science Ouverte à l'UPSaclay : E. Augé,
- chargé de mission « mésocentre UPSaclay » : G. Bernadat.

Gif, le 14 février 2022

Chers collègues

Le projet de mésocentre mutualisé de l'Université Paris-Saclay (UPSaclay) consiste à construire un outil d'appui à la recherche dans les domaines de la simulation numérique, de l'hébergement de données, du traitement de données et de l'IA au bénéfice des communautés de l'Université Paris-Saclay. Le mésocentre de l'Université Paris-Saclay permettra de mutualiser des moyens de calcul et de stockage ainsi que des moyens humains autour d'un objet commun au service de la communauté scientifique, de la recherche et de la formation.

Cette structure légère a vocation à être créée au 1<sup>er</sup> juillet 2022, en fédérant les structures existantes de l'Université Paris-Saclay : Virtual Data (pour sa partie cloud et stockage), Fusion/Ruche et Lab-IA. Ce mésocentre sera caractérisé par :

- Un périmètre d'infrastructures concernées,
- Un fonctionnement coordonné de ces machines, au bénéfice de la communauté de l'UPSaclay (sens large), avec des règles de fonctionnement opérationnelles et budgétaires transparentes,
- Une gouvernance collective en charge de définir, en consultation avec les tutelles, les orientations scientifiques et stratégiques du mésocentre, l'allocation des moyens et les investissements à effectuer, et d'assurer le fonctionnement au quotidien du mésocentre,
- Un modèle économique,
- Un nom.

Il vous est demandé de faire partie du Comité de Pilotage transitoire du mésocentre, avec pour mission de préparer la création du mésocentre pour le 1<sup>er</sup> juillet 2022. En ayant régulièrement tenu au courant le CoDiReV de l'avancement de vos travaux, vous ferez ainsi des propositions au CoDiRE de l'Université avant le 30 avril 2022 sur les cinq points ci-dessus, établissant le mode et les règles de fonctionnement du mésocentre. Dans votre mission vous n'omettrez pas d'accorder une attention aux dimensions de science ouverte et de protection des données personnelles, à l'articulation du mésocentre avec les autres outils existant dans l'écosystème de l'UPSaclay, et au positionnement du mésocentre par rapport aux dispositifs régionaux et nationaux.

En vous remerciant pour votre investissement, bien cordialement,

S. Retailleau  
Présidente de l'Université Paris-Saclay



## Périmètre

➡ Tous les personnels scientifiques et techniques des unités rattachées à l'Université Paris-Saclay pourront bénéficier des ressources du mésocentre. **Dans un premier temps**, pour ne pas saturer les ressources et tester les capacités et le mode de fonctionnement du mésocentre, les projets soumis peuvent être limités par le Comité Exécutif.

**(les coûts d'accès/contributions au fonctionnement peuvent cependant être différents selon les tutelles)**

➡ Le mésocentre a pour but de couvrir les utilisations ne relevant pas d'un dépôt de projet dans un centre de calcul national (mise au point d'un code de calcul ou de jeux de paramètres adaptés, utilisation pédagogique de certaines ressources, etc.)

➡ La raison d'être du mésocentre est la mise en oeuvre de ressources mutualisées de calcul et de stockage. Il existe aussi des possibilités d'hébergement de matériel dans une salle informatique moderne, résiliente et à la consommation énergétique optimisée pour un coût de fonctionnement le plus réduit possible



# Gouvernance

**Chacune des trois plateformes garde sa gouvernance. La gouvernance du mésocentre est destinée à coordonner les actions des plateformes, tout en respectant leur autonomie. Le mésocentre est du type « fédération ».**

## ➡ Le Comité Exécutif (CE) est composé :

- des responsables de Ruche, VirtualData et Lab-IA, actuellement : P.-A. Boucard et R. Vicquelin, M. Jouvin et F. Yvon,
- du directeur de la Maison de la Simulation, actuellement : P. Tremblin,
- du représentant de la DSI de l'UPSaclay, actuellement : M. Leoni,
- du représentant de la VP Recherche de l'UPSaclay, actuellement : M. Guidal,
- du VP adjoint Recherche en charge de la Science Ouverte à l'UPSaclay, actuellement : E. Augé,
- du chargé de mission « mésocentre UPSaclay »: G. Bernadat.

**+ Comité des Tutelles**  
**(1<sup>ère</sup> réunion 28/11/22)**

## ➡ Les missions (quelques-unes) du CE sont :

- veiller au **bon fonctionnement** d'ensemble du mésocentre tout en respectant l'autonomie des plateformes,
- **analyser les demandes** d'utilisateurs (ou consortiums d'utilisateurs) pour les **orienter** vers la ou les plateforme(s) pertinente(s),
- si nécessaire, aider les plateformes à **arbitrer** les demandes d'utilisateurs,
- définir, attribuer et suivre les missions et tâches des **personnels « mutualisés »** (Springboard, MesoNet,...),
- dresser un **bilan budgétaire** et fonctionnel annuel de chacune des trois plateformes,
- **proposer les investissements** à effectuer (par exemple, en utilisant la dotation CPER),
- **partager l'information** sur le fonctionnement des plateformes (modalités financières, contrats, prestations, RH...)
- présenter au CoDiReV une fois par an le **bilan** des actions et les perspectives.

## Soumettre un projet

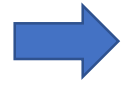
➡ La procédure standard pour accéder aux services du mésocentre consiste à se rendre sur le site web (« <https://mesocentre.universite-paris-saclay.fr> ») où figure une présentation des principaux services du mésocentre et un **formulaire électronique** commun de demande d'accès dans lequel l'utilisateur décrit la **nature du projet**, indique un **correspondant**, les **ressources demandées**, le **nombre d'heures de calcul nécessaires**, etc.

➡ Les éléments clés du projet sont **examinés** rapidement après la réception du formulaire par le Comité Exécutif, qui pourra éventuellement proposer une rencontre avec un/des spécialistes du domaine scientifique visé dans le but d'aider à optimiser l'utilisation des ressources fournies et émettre un avis sur les demandes d'accès direct aux plateformes en fonction du niveau de ressources demandé. A la fin de ce processus (léger), un certain **quota de coeurs de calcul et de stockage est alors alloué** sur une/des plateforme/s.

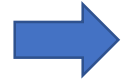
➡ Un projet doit bénéficier d'un minimum de **soutien technique**, selon la nature et la complexité du projet, **hors de l'équipe opérationnelle du mésocentre**. Tout nouveau projet doit venir avec ses propres forces si de nouvelles technicités doivent être développées. Le soutien fourni par l'équipe opérationnelle du mésocentre est significatif, mais **limité**. En particulier, le contact d'un responsable scientifique et d'un responsable technique permanents doivent obligatoirement être fournis avant la création effective d'un projet au sein du mésocentre.

(pour les utilisateurs déjà usagers d'une des trois plateformes, familiers avec leur fonctionnement et ayant déjà identifié leurs propres besoins spécifiques, il est possible d'éviter le traitement en « central » décrit ci-dessus et de s'adresser directement à l'une des trois plateformes constitutives du mésocentre)

## Moyens financiers supplémentaires



Contribution d'UPSaclay au mésocentre **Fusion/Ruche (50k€)**



**CPER 2021-2027: 4 M€** ont été obtenus (*sur 8,35 M€ demandés*) à répartir en 4 volets :

- Volet 1 : **renforcement des plateformes Fusion et LabIA**, dans la perspective de leur convergence en une machine unique ou une intégration plus forte des services offerts,
- Volet 2 : **extension de la plateforme Cloud à VirtualData**,
- Volet 3 : **renforcement de la plateforme globale de stockage des données**,
- Volet 4 : **extension des infrastructures d'hébergement à VirtualData et à l'IDRIS.**

(obtention, a priori, de 500k€ dès fin 2022 pour du stockage)



**MesoNet: 100 k€** pour du stockage



**SESAME COVADO-SHS: ~100 k€** pour du stockage

## Moyens humains supplémentaires

(pour rappel: 1 ingénieur dédié au support de l'informatique scientifique à la DSI d'UPSaclay (ex UPSud))

Pour démarrer le projet et faire fonctionner l'ensemble, il faut, au moins dans un premier temps, **des moyens, en particulier humains, supplémentaires**, le personnel de chacune des plateformes étant dédié à l'opération de chacune des plateformes et à leurs actions propres.

➡ Dans le cadre du projet d'Equipex+ « MesoNet »: un poste CDD d'ingénieur de 4 ans (qu'UPSaclay s'est engagé à pérenniser à terme). Une dizaine d'entretiens ont été conduits depuis la rentrée. Une offre faite à un candidat, en attente de réponse.

➡ Dans le cadre de l'appel à projets PIA4 « ExcellenceS/Springboard »: un poste CDD d'ingénieur de 3 ans en informatique, pour le volet « soutien aux infrastructures communes de recherche pour l'Université Paris-Saclay » (avec un volet sur le partenariat industriel et le transfert technologique). Dans ce même projet « ExcellenceS », 5 recrutements en CDD ont été demandés pour contribuer au virage de l'université vers la science ouverte, en ciblant plus particulièrement l'ouverture des données de la recherche. Ces personnes travailleront en lien étroit avec le mésocentre.

➡ Un poste d'ingénieur de recherche affecté à l'IDRIS, en partie destiné à soutenir les infrastructures Ruche et LabIA, a été ouvert ces deux dernières années par le CNRS. Le poste est cependant toujours non-pourvu.

## Mise en place d'outils/actions spécifiques au mésocentre

- ➔ Un site web (« <https://mesocentre.universite-paris-saclay.fr> ») a été créé, avec une présentation des principaux services du mésocentre et un **formulaire électronique** commun de demande d'accès (« <https://mesocentre.universite-paris-saclay.fr/formulaire.html> »), à destination de nouveaux utilisateurs d'UPSaclay. Un **système de demande d'intervention commun** par « tickets » est également envisagé et actuellement en phase de discussion.
- ➔ Une adresse mail de contact: « mesocentre-copil@universite-paris-saclay.fr »
- ➔ Une charte utilisateurs
- ➔ Un comité des tutelles (un représentant, expert du domaine HPC, de chacun des 14 établissements/organismes qui composent UPSaclay et souhaitant participer à la conduite du mésocentre; réunion une fois par an)
- ➔ Une journée de lancement/présentation du mésocentre (6 décembre après-midi) ouverte à toute la communauté UPSaclay (annonce via les GS)
- ➔ Des développements dans le domaine de **l'open data/science ouverte** doivent aussi être menés. Deux projets, qui exploitent l'approche LinkedOpenData, existent déjà au sein des facultés de Droit (« <https://gregorius.dsi.universite-paris-saclay.fr> ») et de Pharmacie (« <https://daap.dsi.universite-paris-saclay.fr/> ») de l'UPSaclay et les moyens obtenus dans le projet « ExcellenceS » doivent permettre d'accompagner leur évolution.

## Modèle économique



Les trois plateformes VirtualData (cloud) (CNRS/IN2P3, UPSaclay), Ruche (ENS et CS) et LabIA (CNRS/INS2I et UPSaclay) sont financées et opérées par les établissements indiqués entre parenthèses et leurs équipes techniques respectives. **Les trois plateformes ont actuellement un modèle économique autonome et équilibré. Les utilisateurs doivent se conformer aux modèles économiques des trois plateformes qu'ils comptent utiliser.**

**RUCHE:** unités avec tutelles CS, ENS ou UPSaclay accès gratuit (convention frais de fonctionnement: 150k€/3).

Pour les unités n'ayant pas de tutelle ENS, CS, UPSaclay, l'accès à Ruche est ouvert avec le coût indiqué dans la deuxième ligne du tableau ci-dessous.

Coût horaire par cœur (HT)	CPU	GPU
Coût horaire par cœur utilisation Mésocentre en interne [CS + ENS + Université], faisant partie prenante à la convention d'utilisation du Mésocentre	0,008 €	0,771 €
Coût horaire par cœur utilisation Mésocentre pour académique externe non-partie prenante à la convention d'utilisation du Mésocentre	0,011 €	1,156 €
Coût horaire par cœur utilisation Mésocentre hors académique	0,019 €	1,927 €

**Virtual Data (Cloud):** Les utilisateurs doivent acquérir et maintenir les ressources de calcul et de stockage et en payer les frais d'hébergement. Ces coûts d'hébergement sont facturés annuellement et sont basés sur l'espace occupé et la puissance électrique consommée. L'électricité est facturée pour les unités dont l'hébergeur n'est pas UPSaclay.

(à noter que le CPER permet d'acheter un nombre significatif de CPUs et de disques en plus de ce qu'achètent les utilisateurs)



Les projets **bénéficiant d'un financement sur contrat** (ANR, Européen, industriel...) sont encouragés à intégrer les coûts d'utilisation des ressources dans leurs contrats. En particulier, le message de bien intégrer en amont par les chercheurs, dans le montage des projets lors des demandes de financement auprès de l'Europe, ANR, etc..., les coûts associés au mésocentre, **devra être bien relayé** aux différents services des établissements (SPV, DIREV, etc...).