

RAPPORT D'ACTIVITÉS

2021

GENCI

Le calcul intensif au service de la connaissance

GENCI

En charge de **mettre à disposition des moyens performants de calcul et de traitement de données massives,** GENCI a pour mission, aux niveaux national et européen, de **favoriser l'usage du calcul intensif associé à l'Intelligence Artificielle** au bénéfice des communautés de recherche scientifique, académique et industrielle.

Suivez-nous sur www.genci.fr et sur les réseaux sociaux :

www.genci.fr

54,6 PFLOP/s RÉPARTIS DANS 3 CENTRES DE CALCUL

CINES
Supercalculateur **OCCIGEN**
3,5 PFLOP/s - 555 millions d'heures attribuées

IDRIS
Supercalculateur **JEAN ZAY**
28,3 PFLOP/s - 508 millions d'heures attribuées

TGCC
Supercalculateur **JOLIOT-CURIE**
22 PFLOP/s - 1,412 milliard d'heures attribuées

5 ASSOCIÉS

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION
Liberté Égalité Fraternité

cea

cnrs

France Universités

Inria



SOMMAIRE

GENCI : AU CŒUR DES GRANDS DÉFIS

Le mot de la DGRI 04
Edito de Philippe Lavocat 04
2021 vue par les associés de GENCI 06
Le mot de Renaud Vedel 08
Le mot de Neil Abroug 09
On parle de nous 09

NOS ENJEUX MAJEURS EN 2021

VERS L'*EXASCALE* 11
EuroHPC : l'Europe vers l'*Exascale* 12
Exascale, GENCI et le TGCC se préparent 13
La machine précurseur du CINES 14
La France lance son initiative de calcul hybride HPC quantique 15
INTELLIGENCE ARTIFICIELLE ET JEAN ZAY 16
"AI for humanity" 18
Booster Jean Zay 19
L'IA au cœur des enjeux sociétaux 20
Les moyens de GENCI face à la COVID-19 21

LES SERVICES OFFERTS PAR GENCI

Les moyens de calcul disponibles en 2021 24
Les modalités d'accès aux ressources de GENCI 25
Le mot de la Présidente du Comité d'Évaluation 27
Les mots des Présidents des Comités Thématiques 28
Principaux résultats scientifiques 2021 30

GENCI ET SON ÉCOSYSTÈME

Clusster : cloud, formation et industriels 44
MesoNet : la fédération des mésocentres régionaux 45
Competence Center et SiMSEO : au service des industriels 46
Projet FITS 47
PRACE-6IP 48
En route pour un processeur européen Exascale avec les projets EUPEX et EPI SGA2 FR 49

LA SOCIÉTÉ CIVILE GENCI

Des femmes et des hommes au service de la recherche 52
Organisation interne 53
Les membres des Comités et des Commissions 54

GENCI AU CŒUR DES GRANDS DÉFIS

LE MOT DE CLAIRE GIRY

DIRECTRICE GÉNÉRALE DE LA RECHERCHE ET DE L'INNOVATION DU MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION

2021 marque une année importante avec la concrétisation de l'arrivée prochaine de nouveaux moyens de calcul au CINES, et avec une augmentation de puissance GPU à l'IDRIS, financée grâce à la stratégie nationale en IA. GENCI se retrouve au cœur de plusieurs stratégies nationales : intelligence artificielle, technologies quantiques, mais aussi cloud. GENCI est également coordinateur d'un projet Equipex+ rassemblant plusieurs mésocentres de calcul et de traitement de données présents sur l'ensemble du territoire.

Au niveau européen, et en s'appuyant sur cette base solide construite en 2021, GENCI prépare un dossier de candidature pour héberger au TGCC un supercalculateur exaflopique, lors d'un appel qui sera lancé par EuroHPC en 2022.

Plus que jamais, GENCI renforce ainsi son rôle d'opérateur calcul de référence de l'Etat pour la recherche et l'innovation en France.





JEAN ZAY à l'IDRIS

© CNRS - Cyril Fresillon



© CEA - Peinture : C215

JOLIOT-CURIE au TGCC du CEA



© CINES

OCCIGEN au CINES

ÉDITO DE PHILIPPE LAVOCAT

PRÉSIDENT-DIRECTEUR GÉNÉRAL DE GENCI

Malgré une situation sanitaire très contraignante, GENCI a développé en 2021 son activité au service des communautés de recherche scientifiques et industrielles et ce dans ses dimensions régionales, nationales et européennes. En voici quelques faits marquants.

GENCI a continué d'apporter du soutien aux équipes de recherche dans la lutte contre la COVID19 : une trentaine de projets ont tourné avec efficacité en 2021 sur nos calculateurs !

GENCI a poursuivi l'équipement de l'IDRIS et du CINES en précieux moyens de stockage, et signé le marché d'acquisition du calculateur Adastra qui sera pleinement opérationnel au CINES en automne 2022. Poursuivant sa mission de mise en œuvre de la stratégie nationale en intelligence artificielle, via le calculateur Jean Zay opéré par l'IDRIS, déjà un grand succès (près de 600 projets IA dans l'année), GENCI a lancé une opération pour accroitre ses capacités : Jean Zay3 sera opérationnel début 2022. GENCI devrait ainsi disposer sur les 3 centres d'une puissance de calcul de plus de 130 PF à l'automne 2022.

Sélectionné dans le PIA3, le projet MesoNet, piloté par GENCI, préfigurateur d'une infrastructure distribuée des mesocentres régionaux, a démarré à l'automne. La sensibilisation des PME à la simulation numérique s'est poursuivie via SiMSEO dans le cadre du Centre national de Compétences HPC et la relation avec l'industrie s'est développée par des preuves de concept, dans le cadre des programmes quantiques de la Région Ile de France.

L'action européenne croissante de GENCI s'est manifestée dans des projets stratégiques pour la France ; participation au montage du projet EUPEX sélectionné par EuroHPC, démonstrateur de la future architecture de calcul européenne qui sera employée dans la machine *Exascale* qui sera hébergée et opérée par la France au TGCC en 2024. EUPEX s'appuie sur l'initiative EPI, visant à développer un processeur européen, dans laquelle GENCI agit comme interface entre les communautés scientifiques et les concepteurs dans le co-design lancé à l'échelle européenne. Fort logiquement, la grande action de l'année a été la mise en place et le démarrage de l'organisation du Projet *Exascale*-France et la recherche de partenaires nationaux et européens pour créer un consortium en vue de candidater à l'appel à manifestation d'intérêt de EuroHPC qui sera lancé mi-2022. Représentant la France dans le Conseil de PRACE, GENCI a également été moteur dans la mise en place d'une collaboration rapprochée entre EuroHPC et PRACE portant, entre autres, sur les processus de sélection des



« GENCI offre désormais aux équipes de ses Associés, une palette d'outils stratégiques intégrant calcul intensif, intelligence artificielle, et les prémices des technologies quantiques. »

projets pour accéder aux heures de calcul des machines EuroHPC. GENCI est fier d'avoir pu apporter son expertise technique aux Ministères français impliqués dans la gouvernance de EuroHPC.

Enfin, l'année 2021 a vu la concrétisation d'un intense travail d'investissement dans l'instruction de nouveaux projets en réponse aux demandes stratégiques de ses Associés, dans le cadre du Plan National Quantique : GENCI a signé le 1er décembre avec EuroHPC un accord démarrant le projet européen HPCQS, infrastructure hybride HPC et quantique. Cette dernière va servir de brique fondamentale pour démarrer le programme HQI, l'initiative de calcul hybride HPC quantique nationale, pour laquelle GENCI a été sélectionné avec ses Associés dans le cadre des appels à projets du PIA4.

En 2022, la Présidence Française de l'Union Européenne a mis l'accent sur la souveraineté numérique : GENCI offre désormais aux équipes de ses Associés, une palette d'outils stratégiques intégrant calcul intensif, intelligence artificielle, et les prémices des technologies quantiques. Tous ces sujets passionnants développés dans ce numéro, sont dus à l'inlassable investissement de l'équipe associée à celles des centres de calcul. La fructueuse coopération avec nos Associés offre ainsi des perspectives d'avenir très réjouissantes.

2021 VUE PAR LES ASSOCIÉS DE GENCI



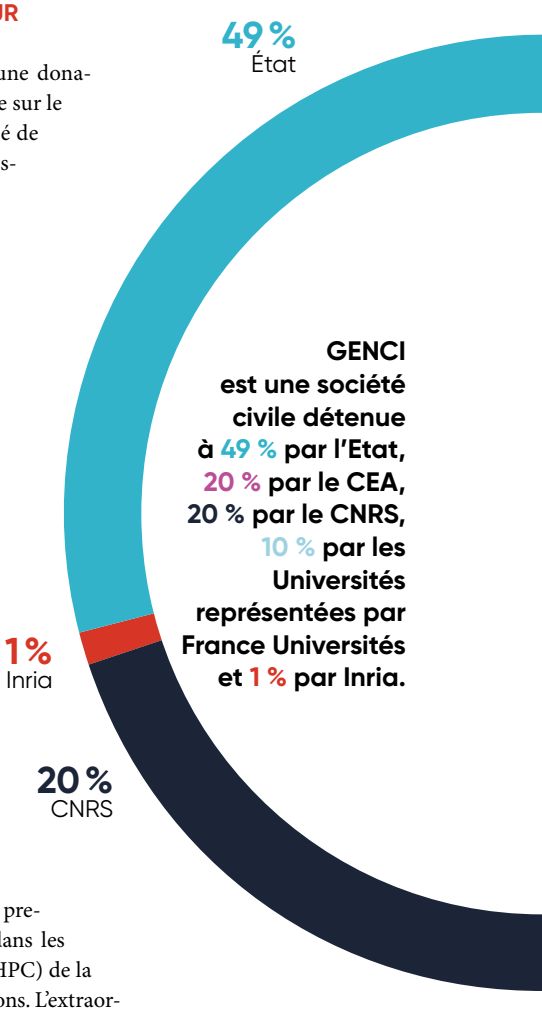
Jean-Frédéric Gerbeau
DIRECTEUR GÉNÉRAL
DÉLÉGUÉ À LA SCIENCE À INRIA

Inria

–– UN ENGOUEMENT POUR LE SUPERCALCULATEUR JEAN ZAY ET L'INTELLIGENCE ARTIFICIELLE

GENCI et Inria avaient été choisis en 2020 pour recevoir une donation en moyens de calcul par AMD pour soutenir la recherche sur le Covid19. L'année 2021 a vu l'installation d'un système, équipé de processeurs AMD EPYC GPU et et d'accélérateurs AMD Instinct, dans l'infrastructure nationale GRID'5000/SILECS sur le site de l'ENS Lyon.

Par ailleurs, les équipes-projets Inria ont confirmé en 2021 leur engouement pour le supercalculateur Jean Zay. Son adoption pour les travaux en IA a conduit à une utilisation d'Inria en très nette augmentation par rapport à ce qui était observé en simulation HPC classique. Le profil des utilisateurs Inria a considérablement évolué en très peu de temps. Les nouvelles procédures mises en place pour accéder aux machines sont particulièrement appréciées.



GENCI est une société civile détenue à 49 % par l'Etat, 20 % par le CEA, 20 % par le CNRS, 10 % par les Universités représentées par France Universités et 1 % par Inria.



Denis Veynante
DIRECTEUR ADJOINT EN CHARGE DES INFRASTRUCTURES NUMÉRIQUES À LA DIRECTION DES DONNÉES OUVERTES DE LA RECHERCHE, CNRS

cnrs

–– UN ENGAGEMENT SANS FAILLE DE NOS ÉQUIPES

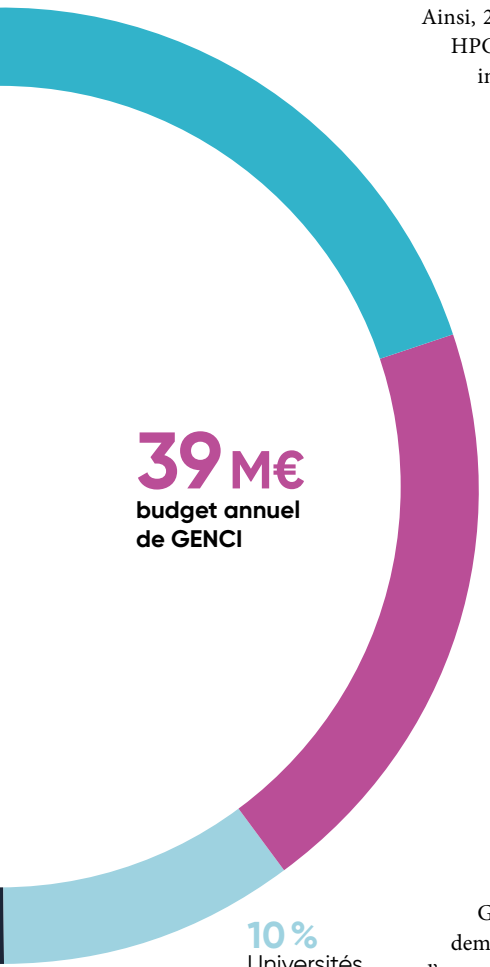
La mise en service et les extensions du calculateur Jean Zay, première architecture convergée en Europe, consacre l'arrivée dans les centres nationaux et le monde du calcul haute performance (HPC) de la communauté de l'intelligence artificielle (IA) et de ses applications. L'extraordinaire appropriation de la machine par cette dernière se traduit aujourd'hui par des projets IA plus nombreux que ceux relevant du HPC. L'engagement sans faille des équipes de l'IDRIS, dont la charge de travail au service des utilisateurs s'est significativement accrue, a été salué par la remise d'un CRISTAL Collectif par le CNRS.

Parallèlement, la dynamique instaurée autour d'IDRIS, du CNRS et de GENCI est confortée par le succès du projet CLUSSTER, coordonné par ATOS et auquel participent notamment Qarnot Computing, OVH et Inria, pour déployer un cloud de recherche dédié à l'IA.

–– DE BELLES PERSPECTIVES POUR 2022

Dès 2020, Genci et le TGCC se mobilisaient pour fournir un accès prioritaire aux ressources de calcul à des équipes de recherche participant à la lutte contre la COVID-19. Ces efforts ont été payants. Des chercheurs du CEA (DRF) ont ainsi pris part à une large collaboration internationale afin de regarder comment le SARS-COV-2 fabrique ses protéines. Une partie des calculs (40 Mh) a été exécutée sur la partition AMD de la machine Joliot-Curie et la modélisation a contribué au développement de nouveaux médicaments antiviraux.

Ainsi, 2021 se distingue comme une très bonne année dans l'écosystème HPC tant au niveau européen que national. Avec GENCI, le CEA s'est impliqué avec succès dans plusieurs appels à projets EuroHPC. Dans le cadre du plan quantique, le TGCC a été choisi pour héberger la première machine de l'initiative de calcul hybride HPC quantique nationale, associant supercalculateurs classiques et des premiers dispositifs accélérateurs quantiques, ouvrant de belles perspectives pour 2022.



–– UNE ACTIVITÉ DYNAMIQUE ET EFFICACE

L'activité de GENCI en 2021 a été particulièrement dynamique et efficace. Elle a notamment impliqué directement France Universités (ex CPU) avec la campagne de remplacement de la machine OCCIGEN du CINES qui sera effectif en 2022 et la coordination du projet MesoNet qui rassemble un grand nombre des mésocentres de nos établissements. Mais la communauté scientifique des membres de France Université s'est également fortement mobilisée dans le cadre de la cellule de veille technologique de GENCI au service du projet d'acquisition et d'hébergement en France d'une machine Exascale pour répondre aux enjeux scientifiques, sociétaux et économiques posés par le calcul intensif. Enfin, grâce à la dynamique collective des centres nationaux animée par GENCI, des évolutions importantes ont pu être réalisées très rapidement dans le mode d'attribution des heures de calcul pour répondre, d'une part, aux nouveaux besoins des utilisateurs dans le domaine de l'intelligence artificielle et, d'autre part, à l'urgence posée par le lancement des projets de recherche dédiés à la COVID19 portés par nos établissements. Les enjeux du numérique sont considérables et touchent en particulier tous les aspects des missions et du fonctionnement des universités.

Dans ce cadre général, depuis sa création, GENCI joue un rôle essentiel dans la construction d'un écosystème national pour répondre aux besoins en calcul et traitement des données de toute la communauté scientifique et a su évoluer en permanence pour se saisir des nouveaux sujets comme le Big Data, l'intelligence artificielle ou le calcul quantique.



Maria Faury
DIRECTRICE INTERNATIONAL ET GRANDES INFRASTRUCTURES DE RECHERCHE CEA/DRF



Guillaume Gellé
PRÉSIDENT DE L'UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE • PRÉSIDENT DU CONSEIL D'ADMINISTRATION DU CINES • PRÉSIDENT DE LA COMMISSION FORMATION DE FRANCE UNIVERSITÉS

LE MOT DE RENAUD VEDEL

COORDONNATEUR NATIONAL POUR L'INTELLIGENCE ARTIFICIELLE



Née d'une proposition du Rapport Villani, la partition dédiée à l'IA du supercalculateur Jean Zay doit être à la fois considérée comme un pilier et comme un succès de la première phase de la stratégie nationale pour l'IA.

Au cours des cinq dernières années, le nombre de publications scientifiques intégrant l'IA a été multiplié par six dans la bibliothèque ArXiv.

Avec son 1 000e projet soutenu au bout d'à peine deux ans, GENCI a démontré, avec l'aide de son équipe-support et le réseau d'ingénieurs de recherche du CNRS et d'Inria, combien elle contribuait de manière déterminante à la diffusion rapide des outils méthodologiques que l'IA contemporaine apporte à la plus grande variété des sciences françaises.

> // GENCI a démontré combien elle contribuait de manière déterminante à la diffusion rapide des outils méthodologiques qu'apporte l'Intelligence Artificielle contemporaine. //

Avec l'accompagnement du projet international exceptionnel en traitement automatisé du langage naturel BigScience, GENCI témoigne de sa force d'attraction au service de la science ouverte.

Avec sa capacité à soutenir de nombreuses entreprises et startups françaises pour des projets donnant lieu à publication, GENCI contribue à la croissance de l'écosystème économique français en IA via le partenariat public-privé.

Et de surcroît, nous sommes déjà assurés que 2022 sera marquée par l'accomplissement de nouveaux défis et l'achèvement de nouveaux projets, accélérés par le dynamisme et la chaleur communicationnelle qui émanent de GENCI.

La forte extension des capacités de calcul de Jean Zay et de nouveaux services mutualisés pour la communauté française de l'IA sont au programme. Nul doute que GENCI répondra présent pour accompagner le nouveau Programme de recherche autour de l'IA embarquée, de l'IA décentralisée, de l'IA de confiance, de l'IA frugale et des fondements mathématiques de l'IA, financé par l'Etat au titre de la « saison 2 » de la Stratégie nationale pour l'IA !

ON PARLE DE NOUS !

CEA, CINECA ET GENCI : MAIN DANS LA MAIN POUR ATTEINDRE L'EXASCALE

Les patrons des trois centres de recherche (deux français et un italien) ont signé un *« engagement commun »* : ce *« protocole d'entente vise à mobiliser ces trois acteurs majeurs du numérique scientifique dans une collaboration sur la route de l'exascale et du numérique post-exascale »*. Il donne un cadre permettant d'*« acter les intérêts communs des parties pour joindre leurs forces dans le cadre des actions/initiatives lancées par l'entreprise commune EuroHPC et de définir la manière de préparer un accord de collaboration à long terme »*, explique le CEA. Pour commencer, cet accord *« vise l'implication des parties dans la réponse à l'appel à candidatures pour héberger un calculateur exascale européen et les futurs appels à candidatures pour le développement de nouveaux processeurs européens ou d'une pile logicielle européenne »*.

nextinpact.com

Supercalculateurs : l'Italie et la France unissent leurs forces vers l'*Exascale* : qu'est-ce que c'est et à quoi cela va-t-il servir ? Dans le sillage du traité du Quirinal signé il y a quelques semaines à Rome par les présidents Sergio Mattarella et Emmanuel Macron, l'Italie et la France se rapprochent de plus en plus. Même dans le domaine de la recherche scientifique.

corriere.it

LE MOT DE NEIL ABROUG

COORDINATEUR NATIONAL POUR LA STRATÉGIE QUANTIQUE

Le plan France 2030 est une opportunité pour répondre aux grands défis actuels sur le plan international, dont font partie les technologies quantiques. Pour positionner la France parmi les premières nations maitrisant ces technologies, le Président de la République a lancé, le 21 janvier 2021, une stratégie nationale pour les technologies quantiques.

Afin de se prémunir d'une forte dépendance à un acteur unique, et répondre aux ambitions françaises en matière de calculateurs NISQ et de solutions logicielles l'État a confié au CEA, à GENCI et à Inria la mise en place au TGCC d'une plateforme hybride HPC-Quantique avec pour objectif :

- Tirer le plus grand bénéfice des accélérateurs quantiques existants en les intégrant dans le contexte plus global du calcul haute performance et en développant des solutions algorithmiques hybrides quantique-classique ;
- Diffuser l'usage du calcul quantique dans les secteurs prioritaires en développant des solutions logicielles et des environnements de développement hardware-agnostiques sur la base de langages de programmations et librairies dédiés aux secteurs différents applicatifs.

En effet, à travers ses missions d'équipement des centres de calcul français et l'application d'une politique nationale unifiée d'attribution d'heures de calcul pour les utilisateurs scientifiques du calcul intensif, GENCI a acquis une expertise reconnue, y compris au niveau européen. Cette reconnaissance transparait à travers la participation à la mise en place de la première infrastructure paneuropéenne de calcul quantique hybride (HPC + quantique) dans le cadre du consortium HPCQS été sélectionné par la JU EuroHPC, et qui sera l'un des piliers de la plateforme française.

Pour ces différentes raisons, l'Etat a décidé de confier à GENCI les volets acquisition et dissémination de la plateforme nationale, le volet R&D étant pour sa part sous pilotage CEA-INRIA.

Evalué positivement par un comité d'experts internationaux, le gouvernement a validé le programme de travail proposé par CEA, GENCI et Inria lors d'une cérémonie organisée à l'institut d'optique, le 4 janvier 2022.

> // L'État a confié au CEA, à GENCI et à Inria la mise en place au TGCC d'une plateforme hybride HPC-Quantique. //



JEAN ZAY, LE SUPERCALCULATEUR QUI AIME (AUSSI) L'IA

Toujours plus gourmande, l'IA a besoin de supercalculateurs. Tout en répondant aux besoins classiques en calcul intensif, l'équipement Jean Zay offre une réponse adaptée aux applications lourdes de l'IA. Il se veut précurseur d'une « convergence » annoncée entre HPC et IA.

dataanalyticspost.com

Pourquoi le supercalculateur Jean Zay connait une augmentation de ses performances en intelligence artificielle.

Le supercalculateur Jean Zay est au cœur de plusieurs projets de recherche de pointe comme Big Science. C'est devenu un outil incontournable et ActuIA y a également consacré un sujet dans le prochain numéro du magazine à paraître dans quelques jours.

actuia.com

ADASTRA : LA FRANCE SE DOTE D'UN NOUVEAU SUPERCALCULATEUR POUR PRÉPARER L'EXASCALE

À horizon 2022, Adastra sera un des supercalculateurs parmi les plus puissants en Europe pour la recherche ouverte académique et industrielle, grâce à son architecture de dernière génération qui complètera les systèmes existants de GENCI disponibles au sein des 2 autres centres nationaux (TGCC du CEA et IDRIS du CNRS).

genci.fr

PLUS D'INFOS

www.genci.fr

NOS ENJEUX MAJEURS EN 2021

VERS L'*EXASCALE* 11
EuroHPC : l'Europe vers l'*Exascale* 12
Exascale, GENCI et le TGCC se préparent 13
La machine précurseur du CINES 14
La France lance son initiative de calcul hybride HPC quantique 15

INTELLIGENCE ARTIFICIELLE ET JEAN ZAY 16
"AI for humanity" 18
Booster Jean Zay 19
L'IA au cœur des enjeux sociétaux 20

Les moyens de GENCI face à la COVID-19 21



VERS L'*EXASCALE*

EUROHPC : L'EUROPE VERS L'*EXASCALE*

Créée en 2018 et localisée au Luxembourg, l'initiative européenne EuroHPC (qui rassemble la Commission Européenne et 34 États Membres) a débuté en 2021 un deuxième mandat avec l'adoption en juillet de la réglementation définissant son mode de fonctionnement et ses objectifs prioritaires pour 2021-2027.

Après la mise en place d'une infrastructure de calcul pan-européenne agrégeant 5 systèmes de classe multi Petascale et 3 systèmes pré *Exascale*, EuroHPC se tourne vers l'étape suivante de son déploiement: l'*Exascale*. Un calculateur de classe *Exascale* est au moins 2 fois plus puissant que le plus puissant des systèmes financés à ce jour par EuroHPC et est comparable à ceux dont le Japon, la Chine et les Etats-Unis se sont dotés.

La feuille de route d'EuroHPC prévoit l'acquisition de 2 calculateurs de classe *Exascale* un en 2023 et un en 2024 qui seront hébergés par des entités sélectionnées en 2022 et 2023.

En ce sens, GENCI prépare avec ses associés et en particulier le CEA qui hébergera ce calculateur au TGCC sa candidature au 2e appel à manifestation d'intérêt *Exascale* attendu en 2022. Pour cela en 2021 :

- des contacts ont été menés avec de potentiels partenaires nationaux (académiques et industriels) et internationaux pour assurer le co-financement à 50% du coût complet sur 5 ans du supercalculateur ;
- l'architecture (calcul, stockage, environnement informatique) qui pourrait composer une machine *Exascale* convergée (HPC, IA et HPDA) a été élaborée, utilisant des ressources scalaires mais surtout des ressources accélérées voire intégrant pour quelques applications ciblées, des accélérateurs quantiques ;
- les besoins des communautés scientifiques utilisatrices ont été recensés et évalués en terme de développement d'applications, de support et de formation.

EuroHPC se tourne vers l'*Exascale* avec l'acquisition prévue de deux calculateurs en 2023 et 2024.



EXASCALE, GENCI ET LE TGCC SE PRÉPARENT



PAR CHRISTINE MÉNACHÉ
RESPONSABLE DU TGCC (CENTRE DU CEA ACCUEILLANT LE SUPERCALCULATEUR JOLIOT-CURIE)

En 2021, tout en poursuivant l'administration et l'exploitation des 4 partitions de calculs de Joliot-Curie et de leur environnement, les équipes du CEA/TGCC ont été mobilisées pour mener les études techniques liées au projet *Exascale France*.



L'équipe du TGCC devant le calculateur Joliot-Curie.

En effet, en fonction des différents scénarios budgétaires envisagés il a fallu produire des études « avant-projet » au niveau des servitudes (électricité, refroidissement, gros œuvre…) pour confirmer la faisabilité technique de l'hébergement du calculateur et de son environnement. En parallèle, des scénarios d'urbanisation des salles machines du TGCC ont été construits pour arriver à poursuivre le plus longtemps possible l'exploitation de Joliot-Curie, au moins partiellement. La définition de l'architecture informatique du calculateur *Exascale* et de son environnement de stockage ont fait l'objet de nombreuses analyses et projections en fonction de la puissance de la machine. Ce travail devra se poursuivre en 2022…

Pour permettre aux chercheurs de se familiariser avec l'architecture ARM (une des architectures possibles pour l'*Exascale*), la partition A64FX de Joliot-Curie a été ouverte à tous les utilisateurs nationaux début octobre 2021. Depuis juin 2021, elle était réservée aux membres de la cellule de veille technologique.

En ce qui concerne la vie du centre de calcul, il faut noter la mise en place fin 2021/début 2022, de nouvelles équipes pour la hotline, l'exploitation et la gestion des serveurs et des réseaux.



La partition A64FX de Joliot-Curie ouverte à tous les utilisateurs nationaux.

+ D'INFOS SUR NOTRE SITE

www.genci.fr

LA MACHINE PRÉCURSEUR **DU CINES**



PAR BORIS DINTRANS
DIRECTEUR DU CINES (CENTRE DE FRANCE UNIVERSITÉS ACCUEILLANT LE SUPERCALCULATEUR OCCIGEN)

L'année 2021 a marqué la fin de la procédure de dialogue commune GENCI/CINES pour le renouvellement du supercalculateur Occigen, mis en production en 2015.



Après 18 mois d'échanges, c'est la solution HPE qui a été retenue. Ce nouveau supercalculateur, dénommé Adastra, sera parmi les plus puissants en Europe pour la recherche ouverte académique et industrielle, avec une performance crête de plus de 70 PFlop/s soit x20 par rapport à Occigen.

Adastra, nom tiré de la locution latine *Per aspera ad astra,* fournira aux chercheurs des capacités de calcul massives et innovantes pour leurs besoins complexes en simulation numérique. Il repose sur un système convergé HPE Cray EX constitué de processeurs Genoa et d'accélérateurs GPU MI250X d'AMD qui offrent notamment le meilleur coût total de possession (TCO), soit un équilibre entre performance soutenue et efficacité énergétique pour un ensemble pertinent d'applications académiques et industrielles de production. De plus, une collaboration innovante sur 2 ans, appelée « contrat de progrès », a été prise en compte et les équipes CINES, HPE et AMD ont démarré dès novembre 2021 les travaux de portage GPU de 5 applications cibles dans l'optique de la mise en production mi-2022 de la partition accélérée d'Adastra. Le système complet GPU+CPU sera définitivement mis en production fin 2022, pour le prochain appel à projets A13 de GENCI, et renforcera par la même la position et les moyens de la recherche française sur la voie passionnante de l'*Exascale* !



La solution HPE Cray retenue pour le renouvellement du supercalculateur OCCIGEN.

© HPE

LA FRANCE LANCE SON INITIATIVE DE CALCUL HYBRIDE HPC QUANTIQUE

Un an après l'annonce par le président de la république du Plan National Quantique, le coordinateur de la stratégie nationale quantique a lancé HQI (France Hybrid HPC Quantum Initiative), un programme de recherche académique et industrielle qui s'appuiera notamment sur des acquisitions de plusieurs technologies quantiques. Ces systèmes seront couplés au supercalculateur Joliot Curie de GENCI, un supercalculateur de classe européenne (Tiers-0), hébergé et opéré par le CEA, servant ainsi les intérêts de la communauté de chercheurs français, européens ainsi que des collaborations internationales pour de la recherche ouverte. Cette hybridation quantique/HPC est une initiative innovante et unique qui bénéficiera de l'expertise reconnue des équipes du CEA/TGCC dans l'exploitation des infrastructures, la sécurité et le support aux scientifiques.

Doté d'un budget de 72,3 M€, ce programme est piloté par le CEA. Le SGPI (Secrétariat général pour l'investissement) a confié à GENCI le volet relatif aux acquisitions de simulateurs et calculateurs quantiques pour un montant de 36,3 M€, qui devrait être abondé au cours des 4 prochaines années par des co-financements européens et industriels. La plateforme, ouverte et évolutive, disposera à terme d'un ensemble de technologies françaises et européennes variées et complémentaires pour répondre au mieux à la majorité des problématiques de recherche et cas d'usage métiers. Le second volet relatif au programme de recherche est lui doté de 36 M€. Il est porté par le CEA et Inria avec le concours du CNRS et de France Universités ainsi que l'acteur cloud souverain OVHcloud, le constructeur informatique Atos et les startups qui concourront, ensemble, à supporter le développement d'une couche logicielle de programmation et d'exécution sur des moyens de calcul hybride incluant aussi le développement de librairies pour des verticales métiers (santé, chimie, finance, etc.) ou transverses (machine learning, optimisation, etc.). GENCI aura en complément la responsabilité d'assurer la mise à disposition et l'usage des moyens de calcul ainsi la dissémination avec l'ensemble des acteurs de l'écosystème quantique (Le Lab Quantique, QuantX, Teratec, Systematic, etc.), collectivités territoriales et industriels métier.



La première brique de HQI est déjà en route : elle repose sur le projet HPCQS cofinancé par la Joint Undertaking EuroHPC et les 6 états membres du consortium (grant agreement No. 101018180) dont la France (via HQI) et l'Allemagne, deux pays dont la candidature à l'*Exascale* est actée. La France devrait ainsi capitaliser sur ce programme et pourrait coupler son futur supercalculateur *Exascale* à des moyens de calcul quantique avec au moins deux objectifs : ouvrir la voie à de nouvelles architectures de calcul pour adresser des enjeux de performance/écoresponsabilité et supporter les utilisateurs dans leur cheminement vers l'ère du post-*Exascale*.

FRANCE HYBRID HPCQUANTUM INITIATIVE

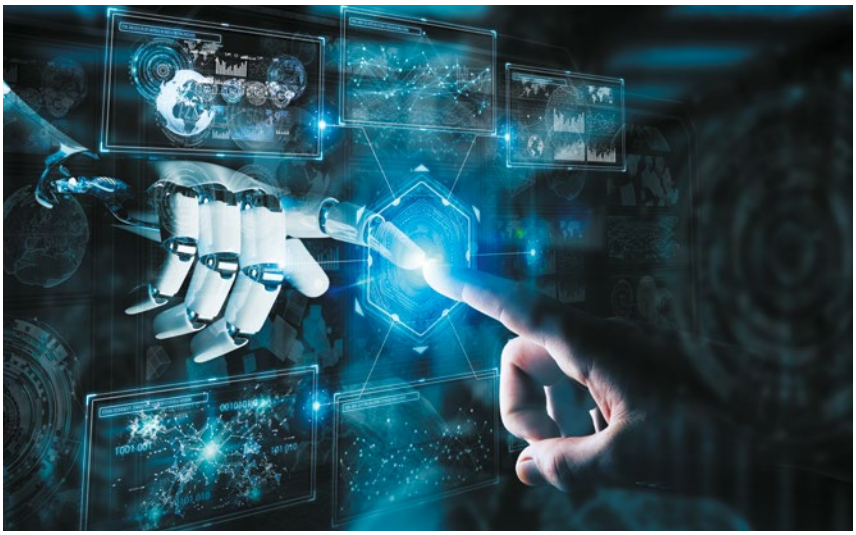


INTELLIGENCE ARTIFICIELLE ET JEAN ZAY



CONTRIBUTION DE GENCI AU PLAN « AI FOR HUMANITY »

En réponse à l'annonce faite en mars 2018 par le Président de la République du plan national « *AIForHumanity* », GENCI a été chargé de mettre en place une des recommandations consistant à mettre à disposition des chercheurs français en Intelligence Artificielle (IA) des moyens de calcul souverains.



En 2019 le supercalculateur convergé nommé Jean Zay - du nom du ministre de l'Éducation Nationale et cofondateur du CNRS avant la Seconde Guerre mondiale - a été mis en service à l'IDRIS, convergé car sa composition matérielle et logicielle, il répond à la fois aux besoins en simulation numérique mais aussi à ceux en intelligence artificielle.

En complément de la machine et afin d'aider les chercheurs à l'utiliser au mieux, les équipes de support applicatif de l'IDRIS se sont renforcées avec une **dizaine d'experts en IA issus du CNRS et d'Inria.**

Deux ans après sa mise en service fin 2019 ce sont plus **800 projets en IA qui ont été soutenus sur Jean Zay,** issus de laboratoires de recherche académique (CNRS, Inria, CEA, Universités, INSERM, INRAE, IFREMER, IRCAM, IRD…), de startups (plus d'une trentaine) ou de grands groupes (Renault, Stellantis, VALEO, Orange…) soit un rythme de presque deux projets par jour ! Au-delà du nombre, il faut noter aussi la **grande variété des domaines d'usage de l'IA sur Jean Zay** avec des projets dans le domaine de la vision, du traitement automatique des langages, de l'aide à la décision, *smart cities*, santé/médecine (notamment autour de la COVID-19), en neurosciences, robotique, sciences sociales, chimie, physique des particules, astrophysique, climatologie…

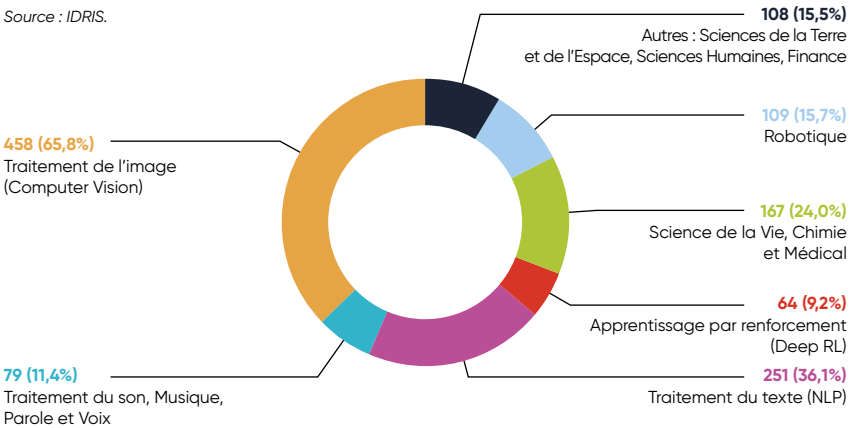
A ce jour, le supercalculateur Jean Zay de GENCI, installé à l'IDRIS dans le cadre de #AIForHumanity, **est un des supercalculateurs convergés les plus puissants en Europe,** prêt à adresser les enjeux du second volet de la stratégie française en IA.

EN SAVOIR + SUR LE PLAN AI FOR HUMANITY

www.aiforhumanity.fr

NOMBRE DE DOSSIERS PAR APPLICATIONS IA
(SUR 696 DOSSIERS)

Source : IDRIS.



BOOSTER **JEAN ZAY**



PAR PIERRE-FRANÇOIS LAVALLÉE
DIRECTEUR DE L'IDRIS (CENTRE DU CNRS ACCUEILLANT LE SUPERCALCULATEUR JEAN ZAY)

Deux ans et demi après sa mise en service, il est temps de faire un premier bilan du supercalculateur Jean Zay.

Première architecture nationale hybride accélérée, Jean Zay a vu sa configuration étendue à plusieurs reprises pour atteindre aujourd'hui une puissance crête de 28 PFLOP/s (2712 accélérateurs GPU V100).

L'accord signé en février 2021 entre le CNRS et l'EPAPS pour récupérer les calories générées par Jean Zay au sein du réseau de chaleur fatale du Campus urbain Paris-Saclay (équivalent du chauffage de 1000 logements neufs) devrait être opérationnel courant 2022, renforçant ainsi le caractère écoresponsable déjà remarquable du calculateur. Première infrastructure nationale dédiée à la communauté de l'Intelligence Artificielle, Jean Zay a rapidement été adopté par les utilisateurs du domaine via les Accès Dynamiques (AD) avec plus de 1 000 comptes ouverts (45% des utilisateurs) et plus de 800 projets (70% des projets). L'accompagnement mis en place dès 2020, avec une équipe de support IA dédiée et aujourd'hui constituée de 12 experts (CNRS, Inria), est indéniablement l'une des clés de ce succès.

Mécaniquement il en résulte une saturation des ressources dévolues à l'IA, encore amplifiée par l'émergence de projets dimensionnants tel la collaboration BigScience dans le domaine du traitement automatique des langues. En parallèle, l'utilisation des ressources par la communauté HPC (principalement les CT4, CT7 et CT9), s'accroit régulièrement au fil des mois. Pour répondre à ce constat, le ministère a accordé une nouvelle extension à Jean Zay de 8 PFLOP/s supplémentaires, composées de 52 nœuds HPE octo-GPU A100 80Go dont la mise à disposition se fera en deux phases, d'abord pour le Grand Challenge BigScience à partir de février, puis pour l'ensemble des utilisateurs à partir de mai.

Le calculateur Jean Zay à l'IDRIS.



© CNRS - Cyril Fresillon

L'IA AU COEUR DES ENJEUX SOCIÉTAUX

BigScience : un modèle de langage de très grande taille, multilingue et ouvert

Le projet BigScience rassemble plus de 900 chercheurs autour d'un but commun : **la création et l'étude d'un modèle de langage de très grande taille et d'un *dataset* multilingue associé.**

Une part importante des progrès actuels découle de l'utilisation de modèles de langue de taille de plus en plus grande et entraînés sur une quantité toujours plus importante de textes.

Cette course au gigantisme est aujourd'hui menée dans un contexte privé par une poignée d'acteurs industriels ce qui pose des problématiques à plusieurs titres : environnementaux, sociétaux, éthiques et d'accessibilité de la recherche.

BigScience montre qu'une autre manière de créer, comprendre, étudier et partager de telles avancées est possible ; une science ouverte, collaborative, holistique et bénéfique pour l'ensemble des parties prenantes.

Les premiers résultats de BigScience ont été publiés en octobre 2021, avec le modèle T0 et le *dataset* P3. **Ils montrent des performances de généralisations supérieures au modèle phare d'OpenAI, GPT-3, avec un modèle de taille 16 fois inférieure.** En parallèle, les travaux sur la création et l'hébergement ouvert du dataset avancent à grande vitesse et des premières publications ont également été réalisées sur la partie des données textuelles relatives aux dialectes d'Afrique du Nord et du Moyen Orient. Le modèle et le dataset finaux seront multilingues, l'objectif étant de couvrir une partie significative des langues les plus parlées dans le monde. Enfin des outils d'analyse et d'évaluation de modèles ont également été déjà publiés en tant que travaux réalisés lors de BigScience.

EN SAVOIR +

bigscience.huggingface.co

OPTIMISER LES OUTILS D'ANALYSE DE DONNÉES D'IMAGERIE MÉDICALE

L'optimisation d'outils d'analyse de données d'imagerie médicale des poumons grâce à des technologies d'intelligence artificielle, est pilotée par Emilie Chouzenoux au sein de l'équipe-projet OPIS (Optimisation, imagerie et santé) du centre de recherche Inria Saclay et de CentraleSupélec-Université Paris-Saclay. Depuis le début de la pandémie de COVID-19, cette équipe a travaillé au développement de plusieurs outils mettant en œuvre des techniques avancées de deep-learning pour l'aide au **pronostic de patients atteints de la COVID-19.** Proposé en juin 2020, dans le cadre de l'accès prioritaire COVID19 mis en place par GENCI, l'outil **AI-Severity** (projet ScanCovIA, partenariat Institut Gustave Roussy, Kremlin Bicêtre, Owkin), **vise à anticiper, dès la première visite, si un patient développera une forme légère ou sévère de l'infection.** Entrainé sur Jean Zay, il a été par la suite mis en open source et est utilisé par les deux hôpitaux parisiens (Gustave Roussy et Kremlin-Bicêtre).

EN SAVOIR +

www.genci.fr



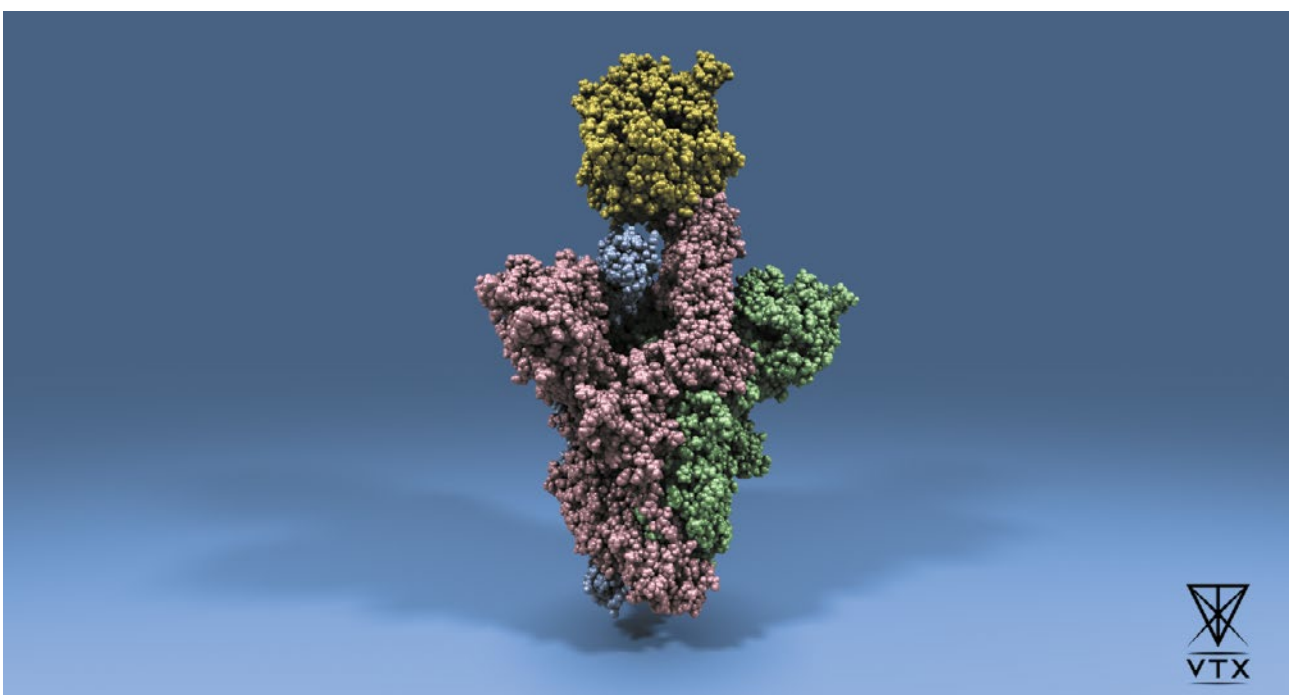
L'équipe partenaire à l'Institut Gustave Roussy

LES MOYENS DE GENCI FACE À LA COVID-19

En 2021, face à l'urgence sanitaire, GENCI et les trois centres nationaux ont continué de soutenir la recherche scientifique dédiée à la lutte contre la COVID-19. Depuis le début de la pandémie, près de 60 projets issus d'équipes de recherche académiques et industrielles dans le domaine de la santé, ont pu bénéficier d'un accès prioritaire à de la puissance numérique de GENCI grâce à des ressources sur les moyens de calcul au niveau national et européen (PRACE), ainsi qu'à de l'accompagnement des équipes de support des centres.

Parmi les différents aspects étudiés, on retrouve : la modélisation de l'entrée du virus dans le corps humain grâce l'étude de la dynamique moléculaire entre la protéine virale Spike du SARS-CoV-2 et la protéine ACE2 présente à la surface des cellules hôtes ; l'étude du variant delta ; la recherche de ligands pouvant inhiber l'interaction entre le virus et la cellule hôte ; le criblage de molécules existantes dans de larges bases de données pour identifier de possibles médicaments ; la modélisation de la transmission du virus dans l'air par le transport de pathogènes contenus dans les gouttelettes de salive émises lors de la respiration, parole, toux… ; l'analyse de l'évolution du taux de reproduction de base du virus R_0, un indicateur permettant de suivre sa circulation au sein de la population ; la simulation des gestions des infections COVID-19 dans les EHPADs et l'exploration des flux Twitter pour l'analyse de l'impact socio-économique de la COVID-19 en France.

Un document sera produit en 2022 pour synthétiser les recherches qui ont été réalisées sur les moyens de GENCI.



Représentation de l'interaction entre la protéine Spike et le récepteur ACE-2 (en jaune) visualisée avec VTX. Données issues de simulations HPC effectuées dans le cadre du projet PRACE COVID-HP.

LES SERVICES OFFERTS PAR GENCI

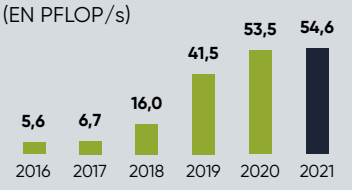
Les moyens de calcul disponibles en 2021 24
Les modalités d'accès aux ressources de GENCI 25
Le mot de la Présidente du Comité d'Évaluation 27
Les mots des Présidents des Comités Thématiques 28
Principaux résultats scientifiques 2021 30

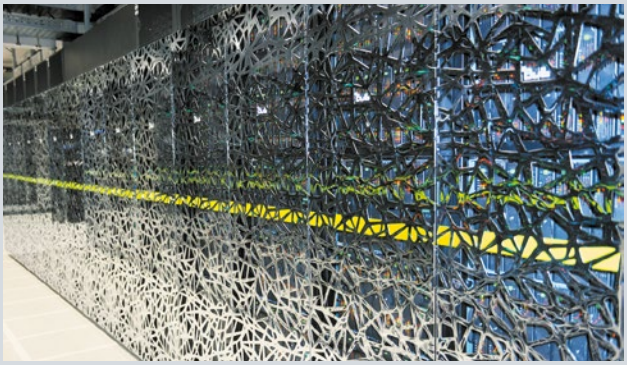




LES MOYENS DE CALCUL DISPONIBLES EN 2021

CAPACITÉS DE CALCUL GENCI
(EN PFLOP/s)





OCCIGEN

SUPERCALCULATEUR ATOS BULLX DE **3,5 PFLOP/s**

Le CINES accueille un calculateur du constructeur Atos d'une puissance crête de 3,5 PFLOP/s. Composé de **3 366 nœuds** pour un total de **85 824 cœurs** avec **2 partitions** en Haswell et en Broadwell.

PERSPECTIVES : A partir de la mi-2022, la machine Occigen sera remplacée par la machine convergée Adastra, machine HPE Cray EX4000 de plus de 70 PFLOP/s avec plus de 120 000 cœurs AMD Genoa et plus de 1 300 GPU AMD MI250X.



JEAN ZAY

SUPERCALCULATEUR HPE SGI 8600 DE **28,3 PFLOP/s**

L'IDRIS accueille un calculateur du constructeur HPE d'une puissance crête de 28 PFLOP/s composé d'une partition scalaire et d'une partition convergée totalisant **86 648 cœurs CPU** et **2 736 GPU V100 ou A100.**

Une partition scalaire de **4,9 PFLOP/s** à base de **1528 nœuds** avec des processeurs Intel Cascade Lake.

Une partition accélérée de **23,4 PFLOP/s** à base de **650 nœuds accélérés** dont 2 712 avec des GPU nVIDIA V100 16/32 Go et 27 avec des GPU A100 40 Go.

PERSPECTIVES : A partir de début 2022, la machine sera complétée par une partition convergée de **8,1 PFLOP/s** de **52 nœuds AMD** Milan octo GPU Nvidia A100 80 Go, soit un total de **3 328 cœurs** et **416 A100.**



JOLIOT-CURIE

SUPERCALCULATEUR ATOS SEQUANA X1000/XH2000 DE **23 PFLOP/s**

Le TGCC accueille un calculateur du constructeur Atos d'une puissance crête de 23 PFLOP/s, composé de **4 partitions** dont 3 partitions scalaires et une partition pour du pré/post traitement à base de GPU NVIDIA.

Une première partition scalaire de **2 292 nœuds AMD** Rome pour 12,2 PFLOP/s et **293 376 cœurs.**

Une deuxième partition scalaire de **1 656 nœuds Intel** à base de Skylake pour 6,9 PFLOP/s et **79 488 cœurs.**

Une troisième partition scalaire de **828 nœuds Intel** à base de Knights Landing pour 2,2 PFLOP/s et **52 992 cœurs.**

Une quatrième partition utilisée pour le pré/post traitement de **32 nœuds NVIDIA** à base de V100 pour 1,2 PFLOP/s et **128 GPU.**

Une cinquième partition prototype de **3 648 cœurs** à base de processeurs Fujitsu ARM A64FX pour 0,3 PFLOP/s.



NEOWISE

PROTOTYPE AMD ISSU DU PLAN MONDIAL DE DONATION D'AMD VISANT À LUTTER CONTRE LA COVID-19

Le site de Grid'5000 de l'ENS Lyon accueille une partition accélérée de **480 cœurs AMD** Rome et de **80 GPU MI50** pour **0,5 PFlop/s.**

LES MODALITÉS D'ACCÈS AUX RESSOURCES DE GENCI

Les moyens de calcul et de stockage nationaux sont mis à disposition gratuitement pour les projets scientifiques en calcul haute performance et en Intelligence Artificielle (IA), émanant des sphères académiques et industrielles. Pour être éligibles, les travaux de recherche ouverte doivent donner lieu à publication à l'issue de la période d'allocation. Toute Demande d'Attribution de Ressources Informatiques (DARI) doit être déposée dans le cadre d'un appel à projets via le portail *www.edari.fr*, commun aux trois centres de calcul nationaux (CINES, IDRIS et TGCC). En 2021, trois types d'accès permettaient d'obtenir des ressources en fonction des besoins des utilisateurs : les accès préparatoires, les accès réguliers et les accès dynamiques. Suite au succès rencontré par les accès dynamiques initialement destinés à la communauté en IA sur Jean Zay à l'IDRIS, nous avons travaillé sur 2021 à élargir ce type d'accès à tous les Comités Thématiques (CT) et à toutes les machines sur les centres nationaux. Depuis le 10 janvier 2022, il est maintenant possible de postuler à tout moment, pour chaque CT, via un formulaire unique à un accès dynamique (demandes inférieures à 50 000 heures GPU/500 000 heures CPU) ou à un accès régulier pour les demandes plus importantes. Il n'est plus possible de déposer un accès préparatoire, les accès dynamiques incluant désormais la fonctionnalité de portage et d'optimisation des codes. Les périodes d'expertises par les CT et les périodes d'allocations des projets pour les accès réguliers restent inchangées.

ACCÈS AU PORTAIL DARI

www.edari.fr

MODES D'ACCÈS

2021	2022
TYPE D'ACCÈS ET BESOINS	
ACCÈS PRÉPARATOIRE (AP) Portage et optimisation de codes	
ACCÈS RÉGULIER (AR) Calcul haute performance et utilisation de l'Intelligence Artificielle (IA)	**ACCÈS RÉGULIER (AR)** Calcul haute performance et IA **Demande > 50 kh GPU/500 kh CPU**
ACCÈS DYNAMIQUE (AD) Développement d'algorithmes et méthodologies en IA et son utilisation	**ACCÈS DYNAMIQUE (AD)** Calcul haute performance et IA Portage et optimisation de codes **Demande ≤ 50 kh GPU/500 kh CPU**
POUR POSTULER	
3 FORMULAIRES À REMPLIR **→ 3 TYPES D'ACCÈS**	**UN SEUL FORMULAIRE À REMPLIR** **→ REDIRECTION VERS AR OU AD**
PORTEURS DE PROJETS ET COMITÉS THÉMATIQUES	
• **AR :** tous Comités Thématiques • **AD :** CT10 dédié à l'AI	**AR et AD :** tous Comités Thématiques
MOYENS NATIONAUX DISPONIBLES	
• **AR :** supercalculateurs de production au CINES, IDRIS et TGCC • **AD :** partition IA à l'IDRIS	**AR et AD :** supercalculateurs de production au CINES, IDRIS et TGCC

SERVICES DISPONIBLES POUR TOUT TYPE D'ACCÈS

• VISUALISATION ET TRAITEMENT DES DONNÉES
• SUPPORT UTILISATEURS ET ACCOMPAGNEMENT DES CENTRES

QUELQUES EXEMPLES DE SIMULATIONS

Ces simulations numériques ont été réalisées dans la cadre de la thèse de Hussein RKEIN. L'objectif est de constituer une base de données très détaillée pour l'étude des propriétés de la turbulence au voisinage d'une paroi. Les paramètres de la simulation ont été choisis pour que l'écoulement soit représentatif d'un écoulement sur une petite partie d'une aile d'avion. Cette base de données doit permettre de mieux comprendre la turbulence et proposer des améliorations sur les modèles qui ne donnent pas entièrement satisfaction pour ce type d'écoulement en présence d'un gradient de pression.

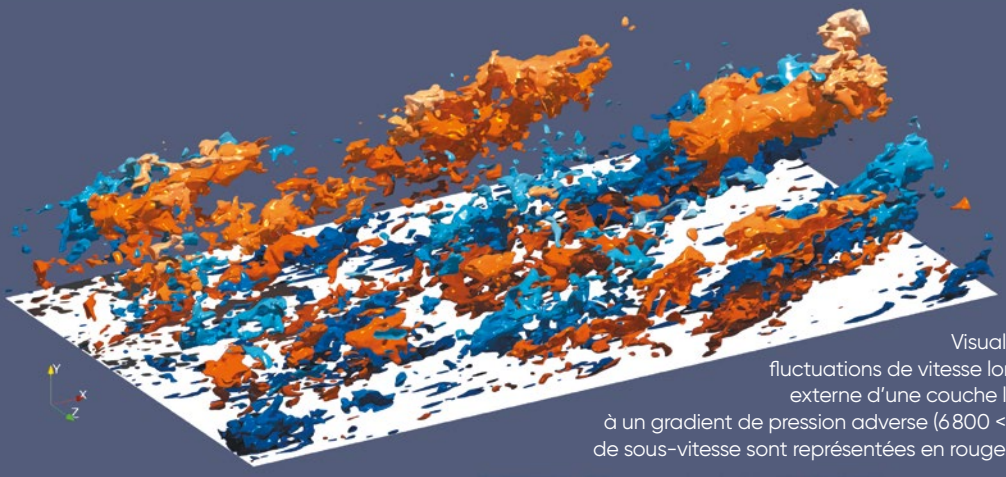


JEAN-PHILIPPE LAVAL
DIRECTEUR DE RECHERCHE CNRS AFFECTÉ AU «LABORATOIRE DE MÉCANIQUE DES FLUIDES DE LILLE - KAMPÉ DE FÉRIET».



HUSSEIN RKEIN
DOCTORANT
À L'UNIVERSITÉ DE LILLE

Visualisation des tourbillons dans une simulation directe de couche limite turbulente soumise à un gradient de pression adverse (6800 < Reθ < 7600).
J.-P. Laval & H. Rkein, LMFL



Visualisation des structures de fluctuations de vitesse longitudinale dans la zone externe d'une couche limite turbulente soumise à un gradient de pression adverse (6 800 < Reθ < 7 600). Les régions de sous-vitesse sont représentées en rouge et de survitesse en bleu.
J.-P. Laval & H. Rkein, LMFL

LE MOT DE LA PRÉSIDENTE DU COMITÉ D'ÉVALUATION

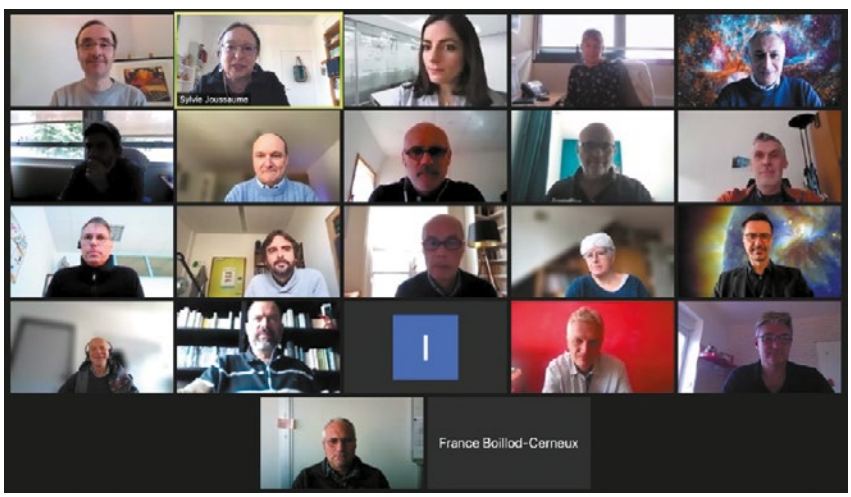
PAR SYLVIE JOUSSAUME,
PRÉSIDENTE DU COMITÉ D'ÉVALUATION DE GENCI

L'année 2021 a été marquée par une pression toujours importante sur les moyens de calcul intensif. Cependant, la disponibilité plus grande des moyens de calcul au niveau national, suite à la diminution des heures dédiées à PRACE, et les possibilités de transfert d'applications entre machines, ont permis d'optimiser ces pressions au cours des allocations DARI A10 et A11. En particulier, cela a permis de compenser l'arrêt temporaire des attributions au CINES pour A11, en attendant la mise en service de la nouvelle machine Adastra au printemps prochain. Cette disponibilité accrue a également permis une session spéciale grands challenges pour 20 projets sur A10.

L'année 2021 a également été consacrée à la préparation du dossier *Exascale* de la France. Beaucoup d'entre vous ont été mobilisés dans le groupe de travail pour présenter au mieux un panorama des principaux codes de calcul de la communauté scientifique française et expliciter leurs besoins futurs, en particulier en soutien pour se préparer aux architectures accélérées, un enjeu pour de nombreuses communautés.

> // 2022 sera une année de changement, avec l'arrivée d'une nouvelle machine au CINES et l'élargissement des accès dynamiques pour les petits projets. //

L'année 2022 sera une année de changement, avec non seulement l'arrivée d'une nouvelle machine mais aussi la mise en place de l'élargissement des accès dynamiques pour les petits projets, auquel les présidents des comités thématiques ont été associés.





Réunion du Comité d'Évaluaton A11, le 7 octobre 2021.

LES MOTS DES **PRÉSIDENTS DES COMITÉS THÉMATIQUES**



CT1
ENVIRONNEMENT
GERHARD KRINNER
Président du CT1

Les projets soumis au CT1 sont d'une très grande diversité en termes de thématiques, de volume, et d'approches. La modélisation du climat est prédominante en ce qui concerne le nombre de projets et les ressources demandées. L'évaluation juste et équilibrée de projets de tailles très différentes - les ressources requises peuvent varier d'un facteur de plus de 1 000 entre les projets – est un point d'attention particulier. L'évolution prévue des modalités d'accès aux ressources GENCI pourrait résoudre certaines difficultés liées à cette diversité.



CT2A
ÉCOULEMENTS NON RÉACTIFS ET MULTIPHASIQUES
GUILLAUME BALARAC
Président du CT2a

La simulation d'écoulements intègre un large spectre d'activités de l'analyse fondamentale aux transferts technologiques. Les 65 projets de cette année témoignent de cette diversité, en approfondissant notre compréhension des écoulements turbulents et des instabilités hydrodynamiques existants dans les applications au cœur des enjeux actuels : énergie, environnement, procédés, transport, santé…



CT2B
ÉCOULEMENTS RÉACTIFS ET MULTIPHASIQUES
PASCALE DOMINGO
Présidente du CT2b

En 2021, 42 projets du CT2b ont reçu une allocation sur les moyens du GENCI, parmi ces projets une petite dizaine étaient des nouveaux projets. La crise du COVID a impacté certains projets de recrutement des doctorants et des post-docs étrangers dont la venue a été rendue difficile.
Le spectre des thèmes de recherche abordés est très large couvrant tous les aspects de la mécanique des fluides réactifs et/ou diphasiques. La thématique de la transition vers une énergie décarbonée est largement présente au travers de projets sur l'hydrogène ou encore le solaire.
La grande majorités des codes utilisés est portable sur toutes les machines disponibles dans l'offre du GENCI, ce qui facilite l'équilibrage des demandes.



CT3
BIOLOGIE ET SANTÉ
LAURENT DESBAT
Président du CT3

Le CT3 (biologie et santé) a évalué 8 dossiers en A10 et 4 demandes complémentaires et 8 en A11 dont 4 nouveaux pour une attribution totale (en Mh) : Jean Zay CPU 13,3 et GPU 0,6, Occigen 2,5, Joliot Curie SKL 10, Rome 26,5. Les projets concernent les thématiques habituelles du domaine : génétique, dynamique moléculaire, dynamique des fluides, biomécanique, imagerie médicale, etc.



CT4
GÉOPHYSIQUE ET ASTROPHYSIQUE
GEOFFROY LESUR
Président du CT4

En 2021, le CT4 a évalué 65 demandes, dans la continuité des années précédentes. Les demandes ont été excellentes et très diverses. Les plus gros dossiers ont porté sur la formation stellaire, les dynamos, et la cosmologie. On note une forte croissance des demandes sur GPU cette année (+ 500 %), résultat de l'arrivée en production de codes adaptés à ces nouvelles architectures dans la communauté.



CT5
PHYSIQUE THÉORIQUE ET PHYSIQUE DES PLASMAS
ÉRIC SERRE
Président du CT5

En 2021, 30 projets de très grande qualité ont bénéficié de 500 millions d'heures de calcul pour couvrir un large spectre de recherches appliquées et fondamentales. Bien que leur nombre diminue, en partie du fait de la mutualisation des codes de calcul, le volume horaire est en forte hausse à environ 25 % de la ressource totale. Ce sont les besoins en QCD sur réseau qui ont conduit à cette très forte croissance.



CT6
INFORMATIQUE ALGORITHMIQUE ET MATHÉMATIQUES
DIDIER AUROUX
Président du CT6

Dans la continuité des années précédentes, le CT6 a expertisé 16 projets en 2021, dont 2 portés par des entreprises. Les thématiques vont des algorithmes numériques aux problèmes de parallélisation (résolution de grands systèmes linéaires, développement de librairies, scalabilité). La majorité des projets sont motivés par des applications provenant d'autres disciplines : physique, biologie, médecine.



CT7
MODÉLISATION MOLÉCULAIRE APPLIQUÉE À LA BIOLOGIE
PATRICK FUCHS
Président du CT7

En 2021, le CT7 a expertisé 30 à 40 dossiers par appel, pour un total d'heures CPU de 60-80 millions et 2-3 millions d'heures GPU. La tendance se confirme avec l'augmentation nette de la demande en GPU grâce à un portage réussi des codes. Toutefois, la demande en CPU est en augmentation également. La plupart des projets concernent dynamique moléculaire de protéines dans leur environnement lipidique ou aqueux. La taille des systèmes étudiés atteint souvent plusieurs millions d'atomes. Ainsi, les moyens du GENCI et la compétence des centres sont décisifs pour mener à bien ces études.



CT8
CHIMIE QUANTIQUE ET MODÉLISATION MOLÉCULAIRE
DAVID LOFFREDA
Président du CT8

Au cours de l'année 2021, la modélisation théorique en chimie quantique a continué de progresser dans la connaissance de la liaison chimique, en gardant ses nombreux domaines d'applications et en œuvrant sur deux fronts : le développement de codes de calcul et leur utilisation. La performance et la robustesse de ceux-ci permettent aujourd'hui de développer des modèles réalistes de systèmes organiques, organométalliques, catalytiques, électrochimiques et biologiques. Il en a résulté cette année une nouvelle augmentation des ressources de calcul demandées, dans un contexte où l'accès permanent à celles-ci a été déterminant. Les développements en cours ouvrent une voix prometteuse vers un portage des codes sur les machines de nouvelle génération. Ceux-ci bénéficient du support applicatif des trois centres du GENCI que le comité remercie pour leur soutien technique.



CT9
PHYSIQUE, CHIMIE ET PROPRIÉTÉS DES MATÉRIAUX
THIERRY DEUTSCH
Président du CT9

Le nombre de dossiers est constant par rapport à l'année dernière de l'ordre de 110 demandes d'allocations avec un taux de nouveaux dossiers de l'ordre de 20 %. La moyenne des demandes d'heures s'établit à un million deux cent mille heures en progression constante. L'apprentissage par réseau de neurone s'établit comme une nouvelle solution dans la simulation des matériaux.

Cette année a vu l'émergence des programmes et équipements prioritaires de recherche (PEPR) notamment dans le domaine des matériaux avec le PEPR Diadème qui met l'accent sur le numérique pour accélérer la découverte de nouveaux matériaux. Ce PEPR devrait proposer de nouveaux services aux utilisateurs et permettre de démocratiser la simulation mais aussi des solutions poussées d'analyse de données vers la communauté des expérimentateurs. L'apparition des accès dynamiques pour les projets inférieurs à 500 000 heures devraient aussi permettre un accès facilité aux moyens de calcul nationaux.
Il reste néanmoins la problématique de la formation à ces nouveaux outils, notamment au langage de script comme python, aux bonnes pratiques de programmation, sans oublier l'aspect gestion des données.



CT10
APPLICATIONS TRANSVERSES ET NOUVELLES APPLICATIONS
BRUNO SCHEURER
Président du CT10

Malgré la pandémie, 2021 se traduit par une croissance du nombre des dossiers (18) et des heures attribuées (100 Mh CPU et 10 Mh GPU).
L'intelligence artificielle (IA) constitue à présent la majorité des dossiers (15), à côté de la minéralogie des planètes, des réacteurs nucléaires et de l'électromagnétisme. Les dossiers d'IA concernent la mise en œuvre de réseaux de neurones dans des domaines très variés. On citera la robotique et les agents autonomes, le traitement de la parole, l'analyse d'images et de vidéos, la réduction de modèles pour la simulation numérique.

2021 • PRINCIPAUX RÉSULTATS SCIENTIFIQUES



Les résultats scientifiques présentés dans les pages suivantes sont issus des travaux de recherche réalisés au cours de l'année 2021 sur les moyens de calcul de GENCI.

Ces projets sont représentatifs des différents domaines scientifiques utilisant les supercalculateurs nationaux mais ils ne constituent toutefois qu'un aperçu des 1 157 projets ayant obtenu des heures de calcul en 2021.

+ DE RÉSULTATS SUR NOTRE SITE

www.genci.fr

31 Physique théorique

- **CT5** Simulations semi-lagrangiennes et multi-échelles de l'équation de Vlasov pour la modélisation des plasmas chauds

32 Sciences de la Terre et de l'Univers

- **CT1** Points de bascule de la calotte Antarctique
- **CT4** Turbulence dans les galaxies primordiales

34 Mécanique des fluides

- **CT2a** Démarche multi-échelle pour les récepteurs solaires : de la physique des couplages température/turbulence au développement de corrélation
- **CT2b** Turbulence induite par un essaim de bulles

36 Biologie, chimie et matériaux

- **CT9** Prédire les propriétés physiques de matériaux
- **CT8** La compréhension des processus physiques et chimiques à l'échelle moléculaire par la modélisation multiméthodes
- **CT7** La simulation éclaire quelques fondements de la défense immunitaire à l'échelle moléculaire

39 Autres applications (informatique, santé...)

- **CT3** Modélisation de complexes anticorps/antigène
- **CT6** Méthodes de décomposition de domaine adaptatives
- **CT10** BigScience : un méga modèle de langage, ouvert et multilingue

CT5 PHYSIQUE THÉORIQUE & PHYSIQUE DES PLASMAS **JEAN ZAY : 5 MILLIONS D'HEURES - JOLIOT-CURIE : 22 MILLIONS D'HEURES**

DANIELE DEL SARTO,
MAÎTRE DE CONFÉRENCES, INSTITUT JEAN LAMOUR, UMR 7198, UNIVERSITÉ DE LORRAINE

ERWAN DERIAZ,
CHARGÉ DE RECHERCHES, INSTITUT JEAN LAMOUR UMR 7198, CNRS

ETIENNE GRAVIER,
PROFESSEUR, INSTITUT JEAN LAMOUR UMR 7198, UNIVERSITÉ DE LORRAINE

MAXIME LESUR,
MAÎTRE DE CONFÉRENCES, INSTITUT JEAN LAMOUR, UMR 7198, UNIVERSITÉ DE LORRAINE

THIERRY RÉVEILLÉ,
MAÎTRE DE CONFÉRENCES, INSTITUT JEAN LAMOUR, UMR 7198, UNIVERSITÉ DE LORRAINE

ALAIN GHIZZO,
PROFESSEUR, INSTITUT JEAN LAMOUR UMR 7198, UNIVERSITÉ DE LORRAINE

PHYSIQUE THÉORIQUE

Simulations semi-lagrangiennes et multi-échelles de l'équation de Vlasov pour la modélisation des plasmas chauds

Le thème central de l'équipe est la modélisation et l'expérimentation numérique en physique des plasmas où les effets cinétiques et microscopiques sont dominants, en particulier dans la génération de champ magnétique due à une anisotropie des fonctions de distribution en vitesse (instabilité Weibel), dans leur dissipation due à la reconnexion magnétique, ou encore au transport des impuretés dans les plasmas de tokamak et dans la fusion à confinement magnétique.

Ce premier exemple concerne la génération d'un champ magnétique à partir d'une instabilité de type Weibel. Une simulation numérique de l'interaction faisceau-plasma en régime relativiste est réalisée sur Jean Zay à l'aide de 2048 cœurs (allocation de 2Mh), en considérant une configuration de deux faisceaux d'électrons se propageant dans des directions opposées. Une telle configuration, rencontrée en astrophysique et dans l'interaction laser-plasma, est instable et conduit à la génération et à l'amplification d'un champ magnétique intense. En utilisant le code semi-lagrangien VLEM (VLasov ElectroMagnetic solver [1, 2]), nous avons montré une corrélation entre la violation de l'entropie et le transfert d'énergie (cinétique vers l'énergie magnétique), induit par les fluctuations microscopiques de la fonction de distribution. Nous observons sur la Fig. 1 une évolution temporelle réversible de la topologie du champ magnétique. Nous avons ainsi pu mettre en évidence le caractère réversible de ce transfert d'énergie [3] en accord avec la théorie de l'information et un nouveau processus de chauffage. Ce type de problème concerne également la reconnexion magnétique où le caractère multi-échelle devient primordial. Un code MHD AMR a été développé pour traiter l'aspect multi-échelle, dans le cadre d'un projet GENCI-DARI *"Nonlinear dynamics of current sheets in magnetic reconnection"* [10-11]. Cette technique AMR pourrait être adaptée pour la résolution du système Vlasov-Maxwell.

Le second exemple concerne le transport des impuretés dans les tokamaks. Le transport des impuretés est complexe et résulte de la turbulence (transport anormal) et des collisions (transport néoclassique). Il peut y avoir une synergie entre les deux canaux, comme récemment mis en évidence [4]. Le code est actuellement développé à l'IRFM-CEA et est nommé GYSELA-5D (GYrokinetic Semi-Lagrangian [6]). D'un point de vue théorique et à partir des résultats des simulations non linéaires, les différents termes impliqués dans le transport (diffusion, vitesses de pincement, contributions néoclassiques et turbulentes, etc.) ont été mis en évidence notamment grâce aux théories néoclassiques et quasi linéaires [7 -8] (voir fig. 2). De plus, une nouvelle source de vorticité qui polarise le système permet de déclencher un cisaillement de vitesse ExB, une barrière de transport et un piédestal (mode à haut confinement, mode H). Pour des taux de cisaillement suffisamment importants, le transport turbulent est supprimé et une barrière de transport est créée. Ces résultats ont été comparés à ceux donnés sans barrière de transport [8], et la barrière de transport s'avère être efficace pour réduire l'amplitude de la turbulence [9].

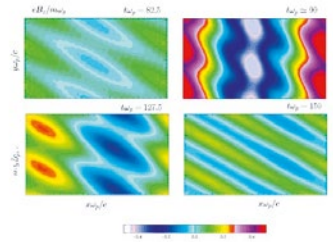


Fig. 1 : Contours de la composante Bz du champ magnétique normalisée à quatre instants différents pour une simulation réalisée à partir du code VLEM) de l'excitation d'une large spectre pour un mode de filamentation oblique. Après saturation de l'instabilité (en haut à gauche), le transfert de l'énergie est inversé.

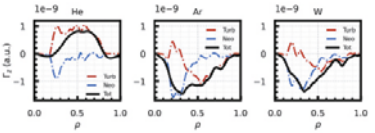


Fig. 2 : Flux de particules pour différentes impuretés (simulations numériques réalisées avec GYSELA). Le flux total d'impuretés (ligne continue noire) est divisé en une partie turbulente (ligne pointillée rouge) et une partie néoclassique (ligne pointillée bleue).

Ces recherches correspondent aux programmes doctoraux de Kyungtak Lim et Guillaume Lo-Cascio, qui associent les équipes de l'Université de Lorraine et de l'IRFM-CEA. Des simulations TERESA [7] ont été réalisées sur Jean Zay-IDRIS pour un total de 5 millions d'heures-cœur. Des simulations GYSELA [8-9] ont été réalisées sur Joliot-Curie-TGCC pour un total de 22 millions d'heures-cœur. Ce travail fait partie d'un projet Eurofusion (TSVV#6).

[1] E. Sonnendrücker, J. R. Roche, P. Bertrand, A. Ghizzo, J. Comput. Physics 149, 201 (1999). [2] M. Sarrat, A. Ghizzo, D. Del Sarto, L. Serrat, Eur. Phys. J. D 71,271 (2017). [3] A. Ghizzo and D. Del Sarto, Plasma Phys. Control. Fusion 63, 055007 (2021). [4] P. Donnel et al., Phys. Plasma Control. Fusion, 61, 044006 (2019). [5] T. Drouot et al., Phys. Plasmas 22, 082302 (2015). [6] V. Grandgirard et al., Comput. Phys. Communications, 207, 35 (2016). [7] K. Lim et al., Plasma Phys. Cont. Fusion 62, 095018 (2020). [8] K. Lim et al., Nucl. Fusion 61, 046037 (2021). [9] G. Lo-Cascio et al., in progress (2022). [10] D. Del Sarto, E. Deriaz, J. Comput. Phys. 351, 511 (2017). [11] H. Betar, D. Del Sarto, M. Ottaviani, A. Ghizzo, Phys. Plasmas, 27, 102106 (2020)

CT1 ENVIRONNEMENT **OCCIGEN : 250 000 HEURES**





BENOÎT URRUTY,
DOCTORANT, UNIVERSITÉ GRENOBLE ALPES
FABIEN GILLET-CHAULET,
CHARGÉ DE RECHERCHE, UNIVERSITÉ GRENOBLE ALPES
OLIVIER GAGLIARDINI,
PROFESSEUR, UNIVERSITÉ GRENOBLE ALPES

Institut des Géosciences de L'Environnement (Univ. Grenoble Alpes, CNRS, IRD, INP)

SCIENCES DE LA TERRE ET DE L'UNIVERS

Points de bascule de la calotte Antarctique

POUR EN SAVOIR PLUS

www.genci.fr

Les calottes polaires sont des composantes importantes du système climatique. Un volume de glace représentant un changement d'élévation du niveau des mers de 58 mètres est stocké sur le continent Antarctique. La glace, qui se forme à partir de l'accumulation des précipitations neigeuses, se comporte comme un fluide visqueux qui s'écoule sous l'effet de son propre poids. La calotte Antarctique a la particularité de reposer sur un socle rocheux qui présente de grandes zones situées sous le niveau des mers et d'être entourée de grandes plateformes de glace qui, du fait de la pression de l'eau, flottent sur les océans. Cette configuration particulière est sujette à une instabilité où une petite perturbation, comme l'augmentation de la fonte sous les plate-formes flottantes, et pourrait conduire à un retrait important et irréversible de la calotte. Les observations récentes ont montré une accélération importante de certains glaciers qui drainent la calotte, notamment dans le secteur de la mer d'Amundsen, suggérant que ce mécanisme ait pu être initié.

Depuis plusieurs années l'IGE co-développe le modèle de calottes polaires Elmer/Ice. C'est surtout au niveau des conditions initiales et aux limites que les difficultés arrivent. Ainsi, par exemple, les conditions à la base sont inaccessibles à la mesure, pourtant la résistance du socle au glissement de la glace peut varier sur plusieurs ordres de grandeur en fonction du type de socle et de la présence d'eau liquide. Il en résulte un écoulement extrêmement hétérogène où la majorité de la glace est drainée par des glaciers émissaires, véritables fleuves de glace au sein de la calotte, pouvant s'écouler à quelques kilomètres par an comme illustré sur la Figure. Un maillage adaptatif et des méthodes d'assimilation de données nous permettent de bien capturer les vitesses d'écoulement observées par satellites. Grâce aux ressources de GENCI, nous avons montré avec nos partenaires du projet européen TiPACCS, que la calotte est actuellement dans une configuration stable face à de petites perturbations.

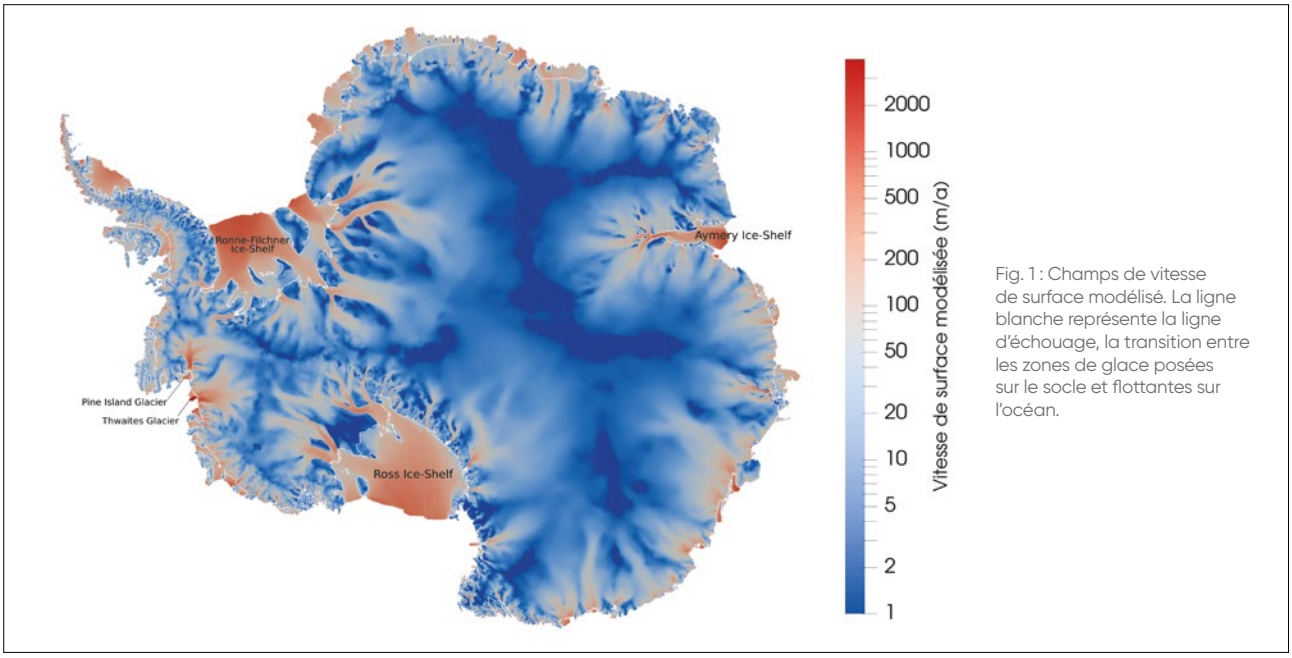


Fig. 1 : Champs de vitesse de surface modélisé. La ligne blanche représente la ligne d'échouage, la transition entre les zones de glace posées sur le socle et flottantes sur l'océan.

CT4 GÉOPHYSIQUE ET ASTROPHYSIQUE **JEAN ZAY : 1,2 MILLION D'HEURES - JOLIOT-CURIE ROME : 2 MILLIONS D'HEURES**



JÉRÉMY FENSCH,
MAÎTRE DE CONFÉRENCES
ECOLE NORMALE SUPÉRIEURE DE LYON

Centre de Recherche Astrophysique de Lyon

SCIENCES DE LA TERRE ET DE L'UNIVERS

Turbulence dans les galaxies primordiales

Les galaxies primordiales sont très différentes des galaxies qui nous entourent. Les disques de ces galaxies ne sont pas composés de spirales, mais de grumeaux qui forment des étoiles de façon soutenue et sont très turbulents. Or, c'est cette turbulence qui engendre les mouvements supersoniques capables de comprimer fortement le gaz et d'enclencher son effondrement gravitationnel et sa conversion en étoiles.

L'origine de cette turbulence est encore débattue : est-elle engendrée par la dynamique de la galaxies ou par les étoiles jeunes, à travers leur rayonnement ionisant ou les explosions de supernovae ? Or, la dynamique locale dépend de la morphologie globale et les étoiles jeunes n'auront pas le même effet si elles sont localisées dans des grumeaux ou réparties dans des bras spiraux. Ces deux processus ont donc toutes les chances de différer entre les galaxies de l'Univers jeune et actuel.

Le but de mon travail est de comprendre ce qui nourrit cette turbulence et quel est son effet sur la formation d'étoiles dans les disques de galaxies au cours de leur histoire cosmique. Pour cela il faut à la fois modéliser l'interaction entre, d'un coté la dynamique globale de la galaxie et la formation stellaire, et de l'autre le gaz du disque. Le défi numérique vient de la gamme dynamique d'échelles spatiales qui rentrent en jeu et du coup mémoire associé.

Pour ce faire, j'ai utilisé les ressources GENCI pour développer une méthode de zoom à l'intérieur d'une galaxie grumeleuse.

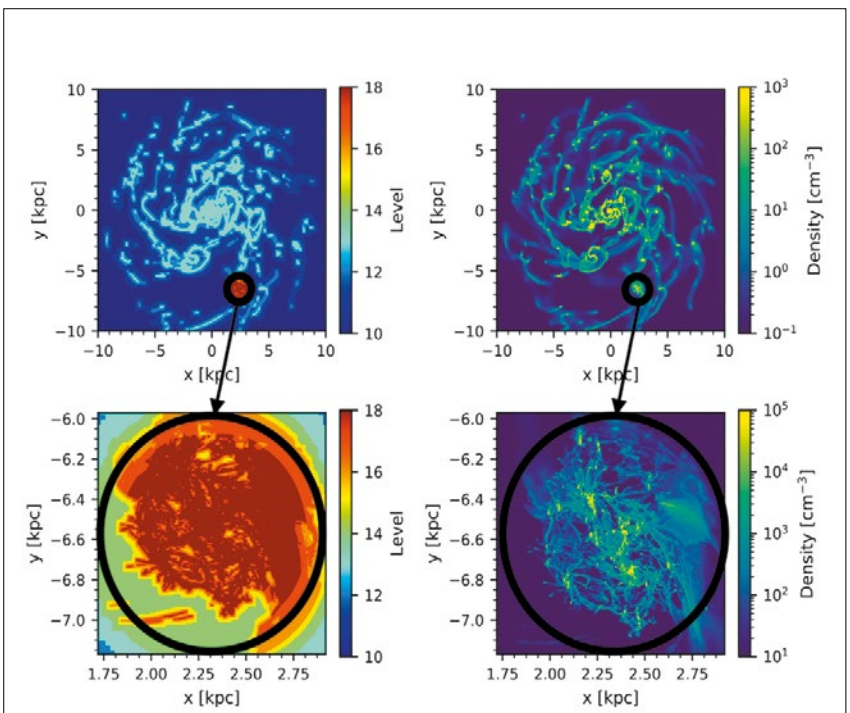


Fig.1 : Niveau de raffinement et carte de densité sur un grumeau dans une des simulations.

Ainsi toute la dynamique de la galaxie est capturée, conjointement à l'effet de la formation d'étoiles, à une résolution maximale de 0,38 parsec (voir Fig. 1).

Afin de différencier l'effet de la dynamique de celui des étoiles, j'ai étudié les conséquences de l'arrêt numérique des effets dus aux étoiles jeunes. D'après ces études préliminaire, l'énergie turbulente ne diminuerait que de 20 %, impliquant une prédominance des effets dynamique sur les effets stellaires.

CT2A ÉCOULEMENTS NON RÉACTIFS ET MULTIPHASIQUES **OCCIGEN : 10 MILLIONS D'HEURES - JOLIOT-CURIE : 8 MILLIONS D'HEURES**





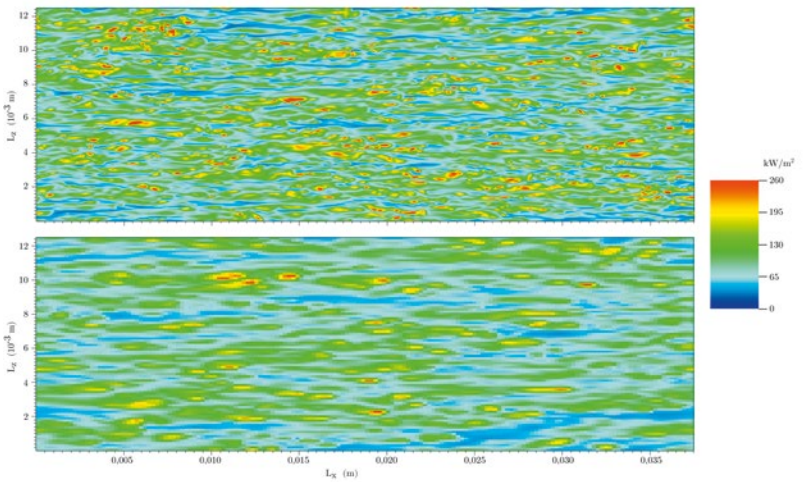
MARTIN DAVID,
ATTACHÉ TEMPORAIRE D'ENSEIGNEMENT ET DE RECHERCHE (ATER)
ADRIEN TOUTANT,
RESPONSABLE DE LA THÉMATIQUE CENTRALES SOLAIRES DE PROCHAINES GÉNÉRATIONS (CSPG)
FRANÇOISE BATAILLE,
DIRECTRICE DU LABORATOIRE PROMES

Laboratoire PROMES-CNRS (UPR 8521), Université de Perpignan via Domitia

MÉCANIQUE DES FLUIDES

Démarche multi-échelle pour les récepteurs solaires : de la physique des couplages température/turbulence au développement de corrélation

Les simulations numériques ont été menées pour la recherche sur les centrales solaires à concentration de nouvelle génération. La résolution numérique des équations couplées de Navier-Stokes et de l'énergie interne dans un canal plan permet l'étude fine des écoulements au sein des récepteurs solaires à gaz pressurisé. Ces écoulements sont très complexes car fortement turbulents et asymétriquement chauffés. Caractérisés par des très hautes températures atteignant 1000 °C, ils sont le siège de transferts de chaleur intenses et de couplages forts entre la thermique et la dynamique. L'approche multi-échelle développée a amélioré les connaissances en comblant des manques de la littérature à chaque niveau et fourni des modèles qui faciliteront l'essor industriel des nouvelles centrales à tour. Les niveaux d'études correspondant à des descriptions détaillées renseignent les niveaux de description plus macroscopique. À l'échelle microscopique, des Simulations Numériques Directes (SND) sont effectuées pour approfondir la compréhension des interactions entre turbulence et température [1] (coût unitaire : 4 millions d'heures de calcul). Au niveau intermédiaire, des Simulations des Grandes Échelles (SGE) sont menées sur des maillages moins fins ce qui réduit drastiquement les coûts de calcul (12000 heures de calcul). Nous cherchons à développer des modèles de sous-maille permettant de reproduire les effets des petites échelles de la turbulence sur les échelles résolues. Les tests sont réalisés en comparant les résultats de chaque SGE avec ceux de SND [2] (voir figure). Grâce à ces évaluations, nous avons identifié un modèle robuste et précis. Ce dernier a été utilisé pour effectuer 70 SGE afin de développer une corrélation permettant d'estimer les transferts thermiques dans des canaux asymétriquement chauffés [3]. Ceci constitue une étude au niveau macroscopique puisque la corrélation pourra être utilisée pour mener des travaux à l'échelle du récepteur solaire [4].



Simulation effectuée par Martin David durant sa thèse, encadrée par Adrien Toutant et Françoise Bataille, lors de l'allocation DARI GENCI 2020-A0082A05099.

Fig. 1 : Champs instantanés de flux conductif à la paroi chaude (1000 °C) pour un nombre de Reynolds de 70000. En haut, les résultats d'une simulation numérique directe sont exposés ; en bas, ceux d'une simulation des grandes échelles de la turbulence.

[1] M. David, A. Toutant, et F. Bataille, « Direct simulations and subgrid modeling of turbulent channel flows asymmetrically heated from both walls », Phys. Fluids, vol. 33, no 8, 2021. [2] M. David, A. Toutant, et F. Bataille, « Investigation of thermal large-eddy simulation approaches in a highly turbulent channel flow submitted to strong asymmetric heating », Phys. Fluids, vol. 33, no 4, p. 045104, 2021. [3] M. David, A. Toutant, et F. Bataille, « Numerical development of heat transfer correlation in asymmetrically heated turbulent channel flow », Int. J. Heat Mass Transf., vol. 164, p. 120599, 2021. [4] M. David, A. Toutant, et F. Bataille, « Impact of asymmetrical heating on the uncertainty propagation of flow parameters on wall heat transfers in solar receivers », Appl. Therm. Eng., vol. 199, p. 117547, 2021.

CT2B ÉCOULEMENTS RÉACTIFS ET MULTIPHASIQUES **OCCIGEN & JEAN ZAY : 3,8 MILLIONS D'HEURES**





FLORIAN LE ROY DE BONNEVILLE,
DOCTORANT À TOULOUSE INPT, INSTITUT DE MÉCANIQUE DES FLUIDES DE TOULOUSE
RÉMI ZAMANSKY,
MAÎTRE DE CONFÉRENCES, INSTITUT DE MÉCANIQUE DES FLUIDES DE TOULOUSE
FRÉDÉRIC RISSO,
DIRECTEUR DE RECHERCHE CNRS, INSTITUT DE MÉCANIQUE DES FLUIDES DE TOULOUSE

Institut de mécanique des fluides de Toulouse, CNRS/Université Paul Sabatier/ Institut National Polytechnique de Toulouse

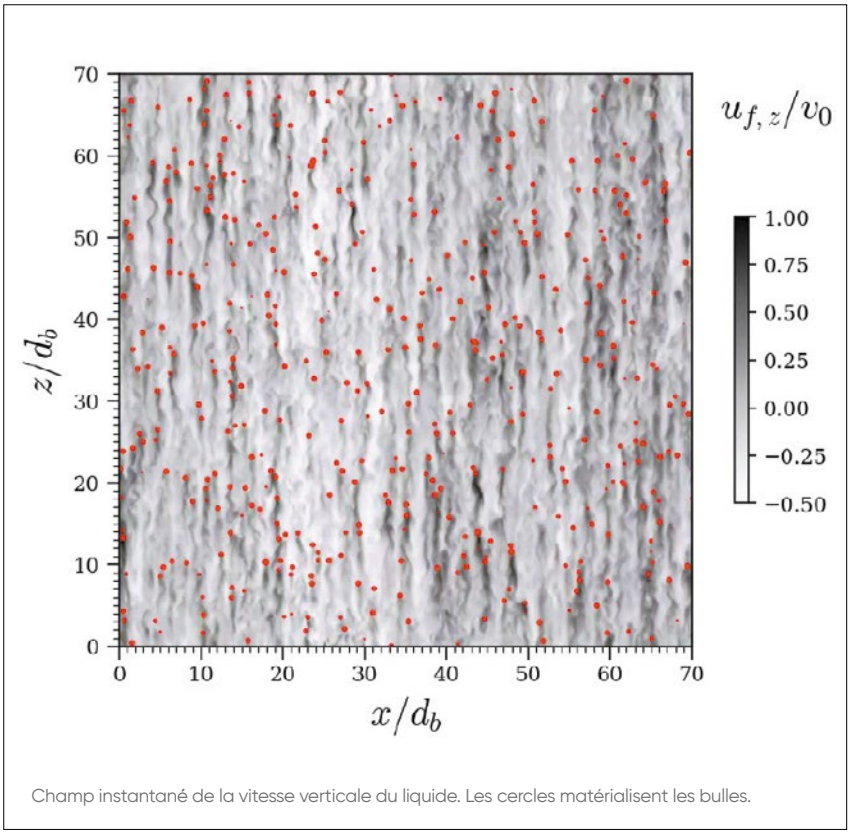
MÉCANIQUE DES FLUIDES

Turbulence induite par un essaim de bulles

On rencontre les écoulements à bulles dans de nombreux phénomènes naturels, procédés industriels, ou des situations accidentelles, comme dans certains scénarios de fusion du cœur d'une centrale nucléaire.

Du fait de leur faible densité, les bulles sont mises en mouvement par la force d'Archimède et l'ascension des bulles entraine le liquide. C'est un système complexe dans lequel les bulles et le liquide sont couplés, conduisant à l'émergence de phénomènes collectifs et des propriétés originales de l'écoulement. Plusieurs paramètres influencent ces mécanismes : le nombre de bulles, leur taille, leur vitesse, la viscosité du liquide, etc. Nous étudions l'ascension d'un essaim de bulles de taille suffisante pour produire des sillages importants, de sorte que l'écoulement est piloté par les interactions entre ces sillages.

Bien que les équations décrivant précisément ce type d'écoulements soient relativement bien connues, leur simulation numérique reste hors de portée, du fait du large spectre d'échelles temporelles et spatiales en jeu. Afin de simuler ces écoulements et de comprendre les mécanismes de transferts turbulents et de mélange y ayant lieu, nous simplifions ces modèles en abandonnant la description précise de la dynamique interfaciale. Cela nous permet de simuler des écoulements ayant un grand nombre de bulles et de mettre l'accent sur les interactions entre sillages. La principale difficulté de ce type de calcul vient de l'auto-interaction, fictive, d'une bulle avec son propre sillage. Nous avons établi une méthode permettant de corriger cet effet et de calculer de façon précise la trajectoire de chaque bulle. Notre allocation sur les machines Occigen et Jean Zay nous a permis d'obtenir des simulations numériques de la turbulence induite par un essaim de bulles en accord avec l'expérience, comme illustrée ci-contre.



Champ instantané de la vitesse verticale du liquide. Les cercles matérialisent les bulles.

CT9 PHYSIQUE, CHIMIE ET PROPRIÉTÉS DES MATÉRIAUX | **JOLIOT-CURIE & OCCIGEN : 6,2 MILLIONS D'HEURES**



NATHALIE VAST, DIRECTRICE DE RECHERCHE AU CEA

De gauche à droite et de haut en bas : **Raja Sen** (Postdoctorant à l'IPP), **Jelena Sjakste** (Chercheuse au CNRS), **Amrita Chakraborti** (Postdoctorante à l'Université de Bayreuth - Allemagne), **Yeonsoo Cho** (Doctorante à l'IPP), **Antoine Jay** (Postdoctorant au LAAS, CNRS, Toulouse), **Olivier Hardouin Duparc** (Chercheur au CNRS)

Équipe de Théorie de la Science des Matériaux du Laboratoire des Solides Irradiés, Institut Polytechnique de Paris (IPP), CEA-DRF-IRAMIS, CNRS UMR 7642

BIOLOGIE, CHIMIE ET MATÉRIAUX

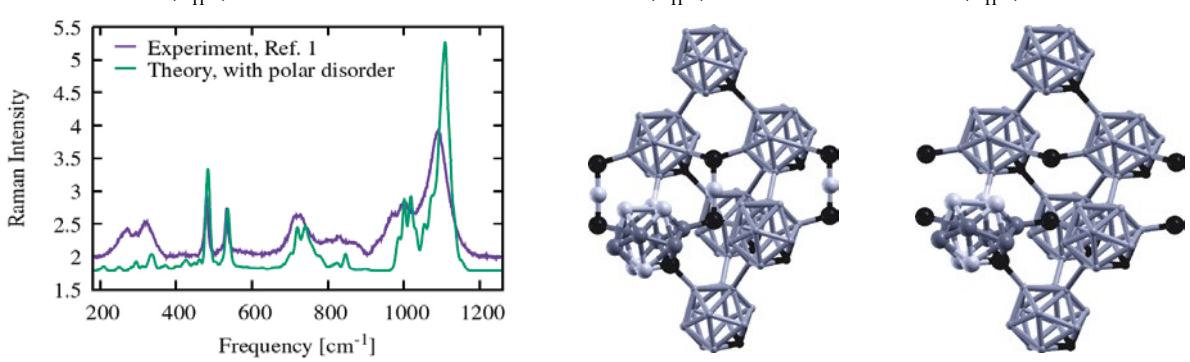
Prédire les propriétés physiques de matériaux

Notre domaine de recherche est la physique théorique et numérique pour la science des matériaux, et consiste en l'étude *ab initio* des propriétés physiques de matériaux, principalement avec des méthodes à l'échelle quantique basées sur la Théorie de la Fonctionnelle de la Densité (DFT). Cette dernière permet de réduire la complexité de l'équation de Schrödinger en utilisant le théorème selon lequel les propriétés de l'état fondamental électronique d'un système d'atomes et leurs électrons sont déterminées de façon unique par la densité électronique de l'état fondamental.

Les calculs DFT sont effectués sans paramètre ajustable et ont permis de prédire avec succès et fiabilité les propriétés physiques de nombreux matériaux, avec une précision maîtrisable. Un exemple est donné *(figure de gauche)* pour le carbure de bore B_4C, l'une des céramiques les plus dures, qui est utilisée pour le blindage d'équipements militaires et la protection des personnes (gilets pare-balle). B_4C est aussi un excellent absorbant de neutrons utile pour contrôler la réaction de chaîne dans de futurs réacteurs à neutrons rapides.

Pour ce matériau, nous avions prédit de façon théorique que le défaut ponctuel lacunaire de moindre énergie consiste en l'éjection de l'atome de bore B de la chaîne C-B-C *(figure du centre)*, donnant des configurations atomiques carbonées C-lacune-C *(figure de droite)* qui amenuisent la tenue mécanique du matériau. Une série d'expériences conduite à l'Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC, Paris), à l'aide de la très récente presse Paris-Edimbourg à enclume rotative pour tomographie (RoToPEC), a permis d'appliquer des déformations en torsion au carbure de bore, le conduisant de façon contrôlée dans le régime plastique.

Les dommages ont été analysés par diffraction des rayons X au synchrotron et par spectroscopie Raman, et interprétés à l'aide de nos calculs en DFT. Les nouveaux pics apparaissant dans les deux méthodes de caractérisation sont la signature de lacunes de bore dans les chaînes, en accord avec les prédictions théoriques. Ce nouveau résultat confirme les pistes de renforcement du B_4C explorées activement avec nos partenaires de l'IMPMC, de l'Institut de Recherche sur les CERamiques (IRCER, Limoges), et de l'Institut de Chimie de la Matière Condensée de Bordeaux (ICMCB).



Exemple de spectre Raman calculé en théorie de la fonctionnelle de la densité (figure de gauche) du carbure de bore avec modélisation du désordre substitutionnel des atomes de carbone dans le site polaire des icosaèdres B11C par une super-maille contenant 405 atomes (courbe bleue), comparé à l'expérience (courbe violette) [Crédit : Physical Review Materials, 6 (2022) 016602]. Structure atomique du carbure de bore dans l'état fondamental (figure du centre) et après déformation plastique (figure de droite) [Crédit : Acta Materialia, 226 (2022) 117553].

CT8 CHIMIE QUANTIQUE ET MODÉLISATION MOLÉCULAIRE **OCCIGEN, JEAN ZAY & JOLIOT-CURIE : 4,3 MILLIONS D'HEURES**



RICHARD OPOKU, DOCTORANT, UMR8523
De gauche à droite et de haut en bas : **Antoine Roose** (Chercheur, UMR8523), **André Severo Pereira Gomes** (Chargé de Recherche CNRS, UMR8523), **Florent Réal** (Maître de Conférences, UMR8523), **Denis Duflot** (Maître de Conférences, UMR8523), **Valérie Vallet** (Directrice de Recherche CNRS, UMR8523), **Céline Toubin** (Professeur des Universités, UMR8523)

UMR8523, Université de Lille

BIOLOGIE, CHIMIE ET MATÉRIAUX

La compréhension des processus physiques et chimiques à l'échelle moléculaire par la modélisation multiméthodes

Les aérosols sont des petites particules solides ou des gouttelettes liquides en suspension. Elles proviennent de sources naturelles (composés organiques volatils émis par la biomasse, sels marins, etc.) ou par l'action humaine (suies issues de la combustion incomplète de carburant).

Ces aérosols jouent un rôle clé en tant que substrats sur lesquels des espèces présentes dans l'atmosphère peuvent s'attacher, réagir en contribuant à leur dispersion autour du globe, y compris les radionucléides tels que l'iode. Ces phénomènes impliquent de multiples processus physico-chimiques qui pour être compris à l'échelle moléculaire requièrent d'utiliser des modèles théoriques combinant différentes approches.

Notre équipe utilise et développe ces modèles et les codes de simulation associés. Nos simulations visent à décrire les propriétés électriques et/ou magnétiques, et/ou thermodynamiques par une approche soluté/solvant en incluant les effets de l'environnement et les effets de température (en moyennant les résultats sur un ensemble de configurations). Nous pouvons notamment simuler: (a) **l'uptake réactif de HO_2 sur des aérosols organiques,** à travers la combinaison de simulations de dynamique moléculaire classique (CMD) et de calculs ab initio QM/MM non-relativistes; (b) les **spectres XPS de l'ion iodure microhydraté,** et l'évolution des spectres d'ionisation de valence d'halogénures avec la taille des gouttelettes, en combinant diverses méthodes développées par l'équipe (champs de force polarisables pour les simulations CMD, et la méthode d'imbrication quantique EOM-IP-CCSD-in-DFT QM/QM relativiste); (c) les spectres XPS de HCl et de l'ion chlorure aux interfaces air/glace couplant des "snapshots" CMD et la méthode EOM-IP-CCSD-in-DFT relativiste; et (d) la solvatation de l'ion iodate (IO_3-), utilisant des simulations Car-Parrinello MD (CPMD) faute de champs de force polarisables classiques.

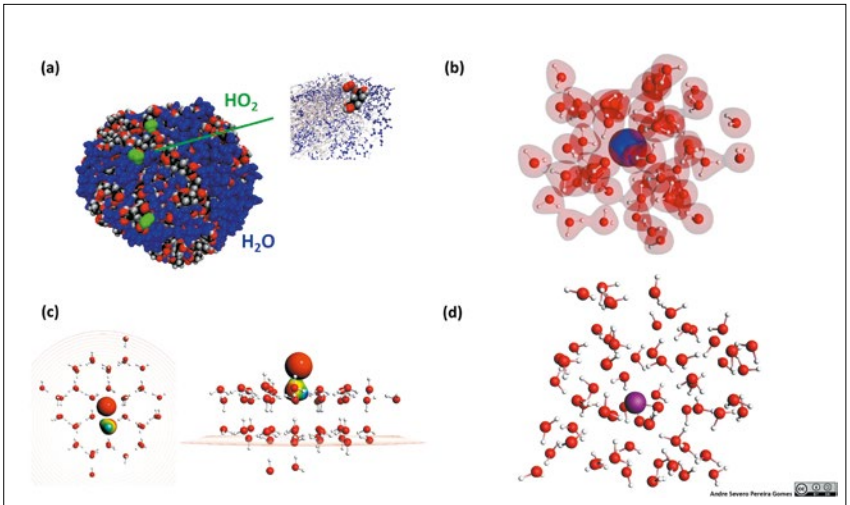


Figure : Illustrations des systèmes modélisés pour (a) distribution de molécules d'eau sur la surface de l'aérosol organique (inset : modèle QM/MM pour le domaine HO_2-phase organique) ; (b) modèle imbriqué pour les halogénures dans des gouttelettes d'eau (rouge : densités électroniques DFT pour l'eau ; bleu : halogénure) ; (c) modèle structurel pour l'interaction chlorure-glace amorphe (d) snapshot de la simulation de l'ion iodate solvaté.

Le très grand volume de simulations requis pour ces exemples d'études a rendu l'utilisation des centres de calcul nationaux indispensable pour atteindre nos objectifs.

CT7 MODÉLISATION MOLÉCULAIRE APPLIQUÉE À LA BIOLOGIE **OCCIGEN : 2 MILLIONS D'HEURES**





XIAOLING WU,
POSTDOCTORANTE, UNIVERSITY COLLEGE LONDON, UK
FABIEN CAILLEZ,
MAITRE DE CONFÉRENCE, UNIVERSITÉ PARIS-SACLAY
AURÉLIEN DE LA LANDE,
DIRECTEUR DE RECHERCHE CNRS, UNIVERSITÉ PARIS-SACLAY

Université Paris-Saclay

BIOLOGIE, CHIMIE ET MATÉRIAUX

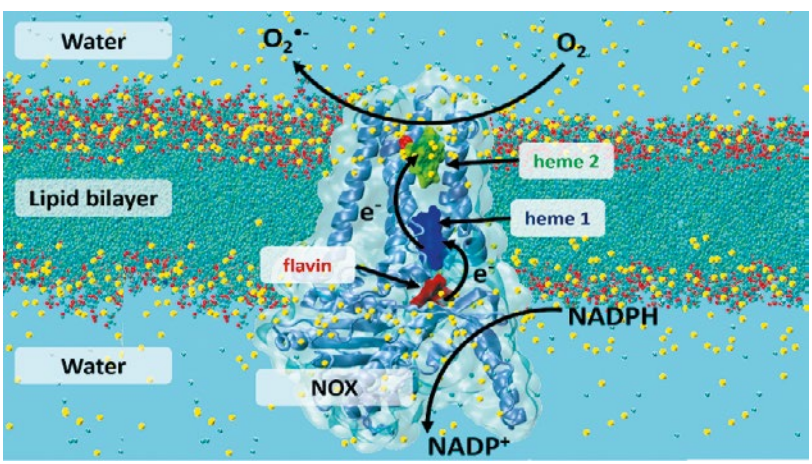
La simulation éclaire quelques fondements de la défense immunitaire à l'échelle moléculaire

Les cellules excellent dans le transport contrôlé des électrons à l'échelle microscopique afin de satisfaire leurs besoins métaboliques. Un exemple est celui des NADPH oxydases de classe 2 (NOX2), qui jouent un rôle primordial dans la défense immunitaire. Les NOX sont ainsi des machines à produire des espèces chimiques toxiques tel que l'anion superoxyde afin de détruire les pathogènes préalablement emprisonnés dans les phagosomes, sortes de sacs formés autour d'eux par une bicouche lipidique. A l'échelle moléculaire, les détails du mécanisme de production des anions superoxydes restent inconnus, bien que les première structures tri-dimensionnelles du cœur catalytique de NOX apparues depuis 2017 permettent d'y voir un peu plus clair. Les électrons transiteraient depuis l'extérieur vers l'intérieur du phagosome via un transfert d'électron transmembranaire.

Les équipes de l'Institut de Chimie Physique (Université Paris Saclay, CNRS) et du Laboratoire de Biochimie Théorique (CNRS) ont joint leurs forces pour tenter d'arracher quelques secrets sur ce transfert d'électron transmembranaire, en se focalisant, pour commencer, sur le transfert inter-hèmes.

Un premier défi à relever a été la construction d'un modèle réaliste du cœur catalytique d'une NOX enchâssé dans une bicouche lipidique. Le modèle généré au LBT inclus plus de 180 000 atomes et se montre stable lors de simulations de dynamique moléculaire classiques sur plusieurs centaines de nanosecondes. Nous avons ensuite déployé les outils dédiés de simulation des transferts d'électrons biologiques développés à l'ICP pour obtenir la première évaluation du bilan thermodynamique de la réaction et identifier des chemins de transferts d'électrons par effet tunnel.

Notre projet a bénéficié du soutien du GENCI et du Labex DYNAMO. Les résultats de notre étude ont été publié dans un numéro spécial de *Frontiers in Chemistry* (2021, 9, 650651), et ouvrent la voie à de nouvelles études passionnantes.



Les NOX sont des protéines membranaires assurant la catalyse de l'oxydation du cofacteur NADPH et de la production d'anions superoxydes ($O_2^{\bullet-}$).

Le transport des électrons NADPH et O_2 est opéré par une succession de sauts entre les cofacteurs flavine, hème 1 et hème 2.

MICHEL MASELLA
CEA, DIRECTION DE LA RECHERCHE FONDAMENTALE,
INSTITUT JOLIOT
LUIGI GENOVESE
CEA, DIRECTION DE LA RECHERCHE FONDAMENTALE,
INSTITUT IRIG

AUTRES APPLICATIONS (INFORMATIQUE, SANTÉ...)

Modélisation de complexes anticorps/antigène

Récompensé par un SANOFI iTech Award 2020

Le développement d'anticorps monoclonaux comme agents thérapeutiques est un domaine en plein essor avec des applications allant du traitement de maladies inflammatoires, des cancers, en passant par la Covid-19. En France, un important vivier de startups travaillent dans ce secteur. De grands groupes pharmaceutiques ont également entrepris un tournant vers les biotechnologies, comme le groupe SANOFI avec la création d'une unité de production d'anticorps monoclonaux sur son centre R&D de Vitry sur Seine.

Si le développement d'anticorps monoclonaux est réalisé en routine, améliorer leur affinité pour un antigène autant que souhaité reste difficile. Les outils numériques disponibles dans ce domaine se concentrent sur l'étude des régions à l'interface de l'assemblage anticorps/antigène (AA). En raison de leur taille, seuls des approches classiques de modélisation moléculaire ou des fonctions de coût empiriques, connues pour leurs approximations grossières limitant leur pouvoir prédictif, sont utilisées pour quantifier les interactions au sein de ces assemblages.

Notre approche, récompensée par SANOFI iTech Awards 2020, ajoute deux nouvelles étapes aux chaînes numériques employées pour optimiser les assemblages AA : (1) étudier leur surface d'énergie potentielle à partir de simulations de type Replica Exchange basées sur une approche de modélisation multi-échelle polarisable avec le code POLARIS(MD) (*http://biodev.cea.fr/polaris/*), et (2) affiner les résultats des simulations à l'aide du code de chimie quantique BigDFT (see *http://bigdft.org/*), qui permet d'évaluer les interactions au sein d'assemblages AA. Ce projet est soutenu par GENCI avec un accès au supercalculateur Joliot-Curie du TGCC et par l'attribution annuelle de 15 millions d'heures de calcul.

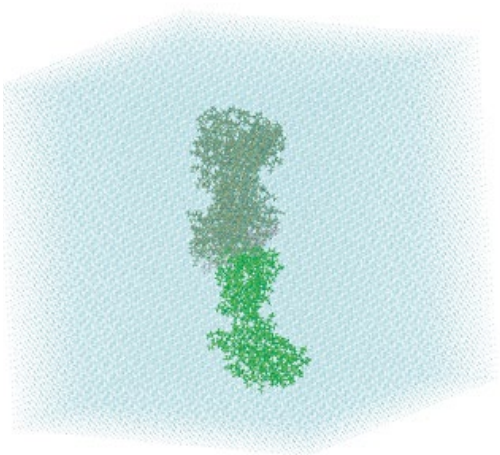


Fig. 1 : un assemblage AA simulé en phase aqueuse par notre approche multi échelle.

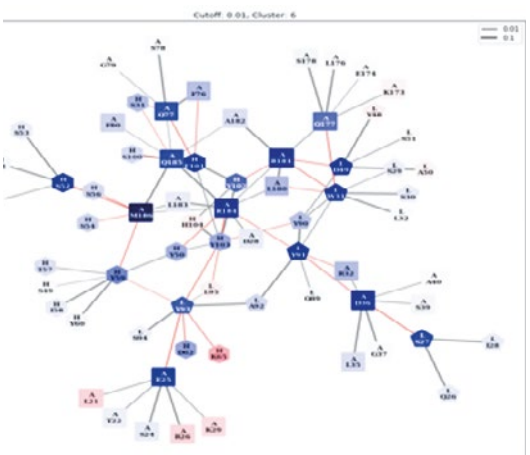


Fig. 2 : résultats du post traitement des données de simulation par une outil de chimie quantique efficace, présentés sous forme de graphe. @CEA



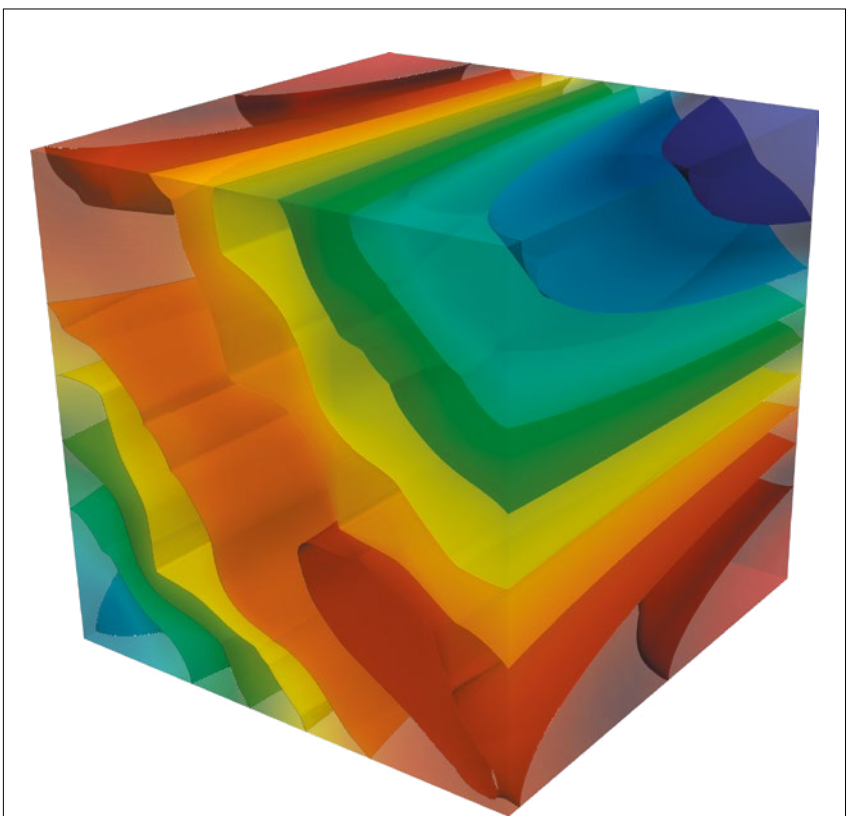
PIERRE JOLIVET,
CHARGÉ DE RECHERCHE,
INSTITUT DE RECHERCHE EN INFORMATIQUE DE TOULOUSE,
CNRS

AUTRES APPLICATIONS (INFORMATIQUE, SANTÉ...)

Méthodes de décomposition de domaine adaptatives

Les méthodes de décomposition de domaine sont basées sur le paradigme « diviser pour régner » afin de résoudre des systèmes linéaires issus de la discrétisation d'équations aux dérivées partielles. Ces méthodes sont dans la même classe d'algorithmes que les méthodes multigrilles. Elles sont quasi-optimales dans le sens où le nombre d'opérations nécessaires à la convergence de la méthode est en O(N), où N est la taille du système creux à résoudre.

Comparé aux coûts légèrement superlinéaires des solveurs de type factorisation LU ou LLT, on comprend bien le pouvoir d'attraction de ces méthodes hybrides. Bien que la théorie des méthodes de décomposition de domaine soient largement étudiées, il existe peu d'implémentations disponibles pour la communauté scientifique. Notre projet est centré autour d'une telle implémentation (*https://github.com/hpddm*), elle-même interfacée avec des librairies de calcul numérique telle PETSc (*https://petsc.org*).



Isovaleurs de la solution d'un problème non-linéaire (équation de Bratu) obtenue avec PETSc en utilisant une méthode de décomposition de domaine adaptative (licence libre).

L'efficacité des méthodes multiniveaux repose sur le développement de grilles « grossières » qui permettent un passage à l'échelle de ces préconditionneurs, indépendamment du nombre d'unités de calcul et de la complexité du modèle physique. Ces grilles peuvent être construites de manière algébrique, ne nécessitant alors que le système linéaire à résoudre. Dans les cas où ce n'est pas possible, des informations sont transmises par le noyau de discrétisation lors de la construction du préconditionneur, e.g., matrice élémentaire éléments finis, pour que la grille grossière s'adapte à la physique. Le cadre théorique des méthodes algébriques et de celles qui ne le sont pas ne permet pas à ce jour de certifier la convergence des méthodes de décomposition de domaine pour un problème arbitraire. Initialement applicable avec garanties uniquement pour certains problèmes spécifiques comme ceux symétriques définis positifs, des avancées récentes nous permettent de traiter de manière adaptative de nouveaux problèmes, comme les moindres carrés.



THOMAS WOLF,
SCIENTIFIC DIRECTOR - CO-FOUNDER,
HUGGING FACE

AUTRES APPLICATIONS (INFORMATIQUE, SANTÉ...)

BigScience : un méga modèle de langage, ouvert et multilingue

L'accélération de l'intelligence artificielle (IA) et du traitement du langage naturel (NLP) aura un impact fondamental sur la société étant donné que ces technologies sont au cœur des outils que nous utilisons au quotidien. En ce qui concerne le NLP, une partie considérable de cette accélération provient actuellement de l'entraînement de modèles de langage de plus en plus grands sur des quantités de plus en plus importantes de textes. Malheureusement, les ressources nécessaires pour créer les modèles les plus performants se trouvent principalement entre les mains des géants de la technologie. La mainmise sur cette technologie pose certains problèmes, et ceci du point de vue de l'avancement de la recherche, de l'environnement, de l'éthique et de la société.

Inspiré par les grandes collaborations de recherche ouvertes comme le LHC, le projet BigScience a été lancé au début de l'année 2021 dans le but de démontrer une autre façon de créer, d'étudier et de partager les grands modèles de langage et les artefacts de recherche au sein des communautés de recherche en IA/NLP.

BigScience est un projet scientifique d'un an qui porte sur les jeux de données multilingues très volumineux et les grands modèles de langage multilingues. Aujourd'hui, il rassemble plus de 900 chercheurs de 60 pays différents et 250 laboratoires universitaires et industriels. Aujourd'hui, 30 groupes de travail se consacrent à l'étude des domaines des données, de la modélisation, de l'ingénierie et de la mise à l'échelle, de l'évaluation ainsi que de l'évaluation des implications éthiques et juridiques et de l'impact social et environnemental liés aux grands modèles linguistiques, ce qui en fait, à ce jour, l'un des plus grands projets scientifiques ouverts et collaboratifs en IA.

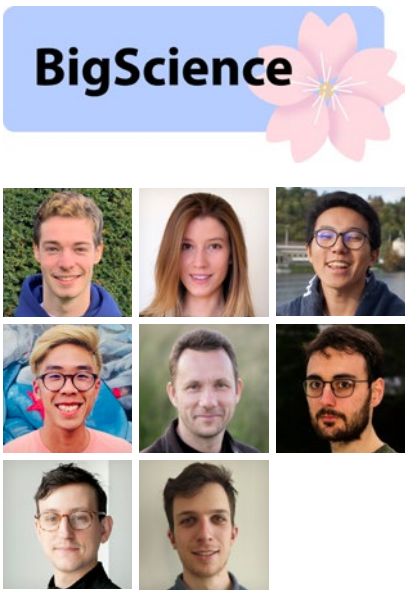
BigScience a pour objectif de créer les artefacts de recherche suivants qui seront partagés avec la communauté des chercheurs :

- Un jeu de données multilingues ;
- Un grand modèle linguistique multilingue (100-200B paramètres) entraîné sur le puissant supercalculateur français Jean Zay ;
- Des outils logiciels, par exemple pour le filtrage des données, le déchargement du modèle, etc. ;
- Documentation des processus, résultats, protocoles, jeux de données et outils développés au cours du projet.

Les principales expérimentations de modélisation ainsi que les efforts complémentaires sont menés sur l'infrastructure du supercalculateur Jean Zay (5 millions d'heures GPU ont été accordées au projet BigScience).

EN SAVOIR PLUS
bigscience.huggingface.co

- Site Web de BigScience
- Introduction, histoire, membres fondateurs, organisation
- Découvrez les artefacts de recherche, les collaborations et les événements organisés en 2021.
- En savoir plus sur les groupes de travail de BigScience



De gauche à droite et de haut en bas :
Hugo Laurençon (Machine Learning Engineer, Hugging Face),
Lucile Saulnier (Machine Learning Engineer, Hugging Face),
Thomas Wang (Machine Learning Engineer, Hugging Face),
Victor Sanh (Research Scientist - Lead, Hugging Face),
Stas Bekman (BigScience team Engineer, Hugging Face),
Pierre Colombo (Post-doctoral fellow, L2S - Centrale Supelec),
Sylvain Viguier (Director of Applications, Graphcore),
Roman Castagné (PhD Student, ALMAnaCH - INRIA Paris)

GENCI ET SON ÉCOSYSTÈME

Clusster : cloud, formation et industriels 44

MesoNet : la fédération des mésocentres régionaux 45

Competence Center et SiMSEO : au service des industriels 46

Projet FITS 47

PRACE-6IP 48

En route pour un processeur européen Exascale avec les projets EUPEX et EPI SGA2 FR 49





CLUSSTER : CLOUD, FORMATION ET INDUSTRIELS

Le projet CLUSSTER, Cloud Unifié Souverain de Services, de TEchnologies et d'infrastructuRes, vise à développer et renforcer la filière française et européenne du Cloud en s'inscrivant dans la stratégie d'accélération CLOUD de la Direction Générale des Entreprises (PIA4).

CLUSSTER fédère ainsi 10 partenaires français parmi lesquels les grands acteurs français publics (GENCI, IDRIS/CNRS, INRIA, CEA-LSCE) et industriels (Atos, Activeeon, OVHcloud, CS GROUP, Qarnot, Hubblo) pour créer un cloud souverain, unifié et évolutif. Ainsi, il répondra dans un premier temps à des besoins en Intelligence Artificielle, puis, il sera amené à couvrir les domaines du HPC et du quantique. Le projet s'efforcera d'offrir aussi un accès vers les moyens et services disponibles en Europe grâce à son intégration prévue avec les services de fédération des plateformes européennes de cloud GAIA-X et EOSC (European Open Science Cloud).

Doté d'un budget de 15,8M€ sur 3 ans, CLUSSTER repose concrètement sur la mise en œuvre d'un portail cloud unifié qui offrira la visibilité complète, à la fois des infrastructures souveraines académiques et industrielles connues, ainsi que des services à valeur ajoutée - Accompagnement, expertise métier, formation (y compris pour les étudiants et enseignants), veille technologique - en IA, pour assurer un continuum unique à ce jour entre recherche ouverte, confidentielle et activités lucratives.

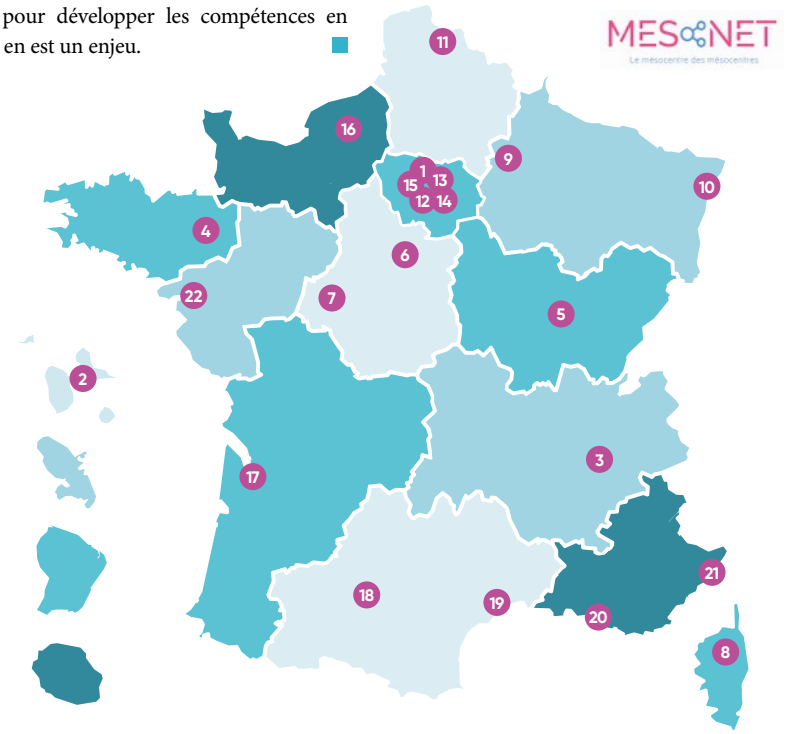


MesoNet : LA FÉDÉRATION DES MÉSOCENTRES RÉGIONAUX

Le projet MesoNet, regroupant 21 mésocentres régionaux et coordonné par GENCI, a officiellement débuté le 1er octobre 2021 pour une durée de six ans et un budget consolidé de 30,4 M€. Ce projet bénéficie d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'avenir portant la référence « ANR-21-ESRE-0051 ».

L'objectif premier du projet est de préfigurer la mise en place d'une Infrastructure de Recherche nationale des mésocentres. MesoNet doit accroitre la visibilité des mésocentres au niveau français et aussi européen, en lien avec l'infrastructure européenne EuroHPC et le Compétence Center français. Cette infrastructure devra offrir aux communautés académiques industrielles et aux étudiants des moyens en calcul intensif HPC mais aussi dans les nouvelles approches comme l'intelligence artificielle et le calcul quantique. Les investissements consentis couvriront tout le spectre, des débutants jusqu'aux besoins nationaux et européens.

Il faut souligner l'importance de la formation et des compétences qui y seront développées, en diversité et en quantité. Ce secteur constitue en effet un gisement d'emplois pour le futur, et la formations des nouvelles générations pour développer les compétences en calcul en est un enjeu.



PARTENAIRES MesoNet

1	GENCI	
2	Université des Antilles	
3	CINAURA (Université Grenoble, Alpes, FLMSN)	
4	GIP numérique Bretagne	
5	Université de Bourgogne Franche-Comté	
6	Université d'Orléans Fédération CaSciModOT	
7	Université de Tours Fédération CaSciModOT	
8	Université de Corse Pasquale Paoli	
9	Université de Reims Champagne-Ardenne	
10	Université de Strasbourg	
11	Université de Lille	
12	Université de Paris Saclay	
13	Centrale Supelec	
14	ENS Paris Saclay	
15	Paris sciences et lettres (dont Observatoire de Paris)	
16	CRIANN (Rouen)	
17	Université de Bordeaux	
18	Université de Toulouse (Calmip)	
19	Université de Montpellier (Meso@LR)	
20	Aix-Marseille Université	
21	Université Côte d'Azur	
22	Centrale Nantes	

COMPETENCE CENTER ET SiMSEO : AU SERVICE DES INDUSTRIELS

Le Projet SiMSEO d'aide aux PME et ETI à la simulation numérique a terminé sa sixième période (1er octobre 2020 au 30 septembre 2021). SiMSEO permet aux PME et ETI de se faire accompagner, y compris financièrement, pour effectuer une preuve de concept autour de la simulation numérique.

Malgré la situation sanitaire complexe, GENCI et les plateformes régionales ont sensibilisé les industriels au travers de diverses initiatives : ateliers, conférences, rendez-vous d'affaires « B2B », démarchage d'agences jouant les intermédiaires (CCI, pôle de compétitivité, acteurs régionaux). Le point important à noter est l'implication de SiMSEO dans le projet européen de *Competences Centers* porté par Teratec, le Cerfacs et GENCI qui a pour conséquence d'imbriquer les actions de SiMSEO dans le *Competence Center* français. Le projet européen de *Competence Center* ayant une durée de fonctionnement de 2 ans, ceci nous a conduit à demander à la DGE une autorisation pour la prolongation de deux ans de SiMSEO, soit jusqu'au 30 septembre 2022.

Au total, depuis le lancement opérationnel du programme « Accompagnement de proximité des PME » de SiMSEO en mai 2016, plus de 2200 PME ont été sensibilisées, 135 PME ont entamé des démarches pour monter un projet et 72 projets ont été lancés.



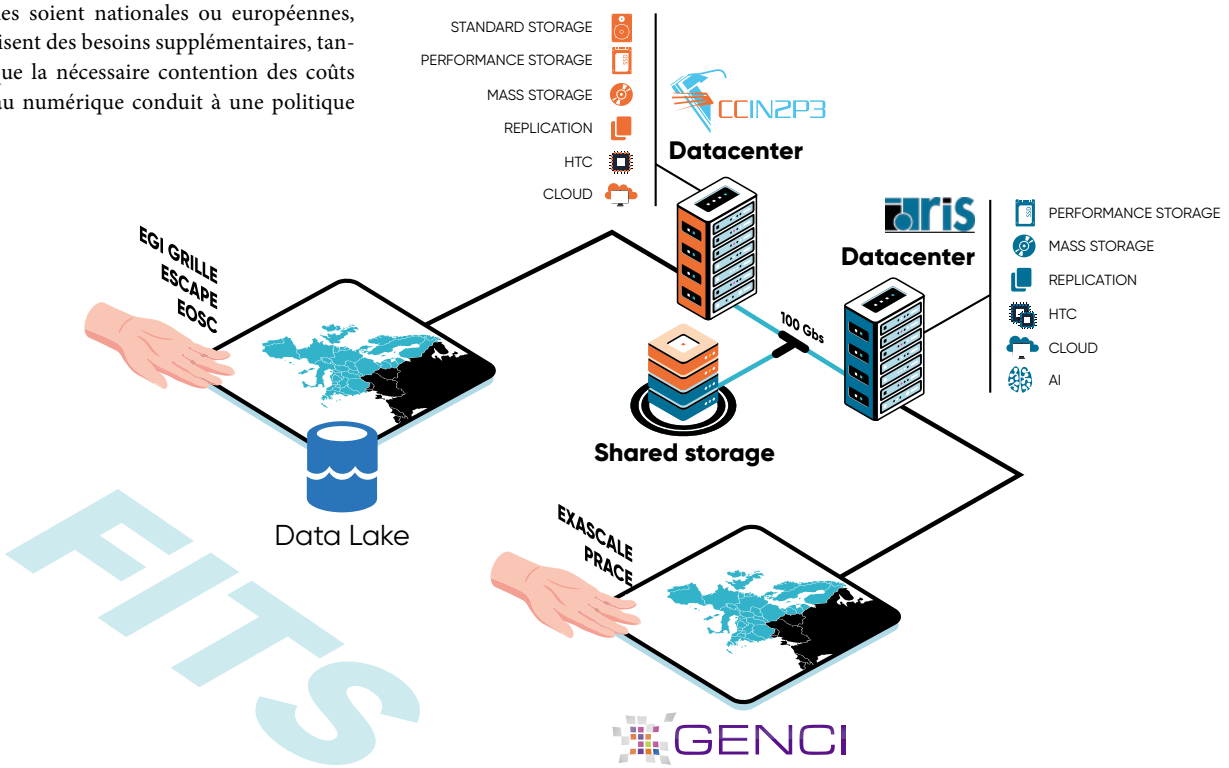
Sur les salons *Les Rendez-Vous Carnot*, les 17 et 18 novembre 2021 et *Big Data & AI Paris* les 28 et 29 septembre 2021.

PROJET FITS : SERVICES NUMÉRIQUES FÉDÉRÉS DU CNRS POUR LES INFRASTRUCTURES DE RECHERCHE

Les infrastructures de recherche (IR) de l'espace de l'enseignement supérieur et de la recherche (ESR) vont devoir faire face à une augmentation exponentielle des données qu'elles produisent. La dichotomie entre données expérimentales et celles issues de la simulation numérique, qui nécessitent chacune des ressources numériques de traitements et d'analyses aux architectures techniques très sensiblement différentes, a vécu. Les politiques menées sur les thématiques de la science ouverte et des données ouvertes, qu'elles soient nationales ou européennes, induisent des besoins supplémentaires, tandis que la nécessaire contention des coûts liés au numérique conduit à une politique de rationalisation et de mutualisation des infrastructures et services.

L'ambition du projet FITS du CNRS est de fédérer grace à ses deux centres de calcul nationaux (IDRIS et le CC-IN2P3) leurs services et savoir-faire, par la mise en œuvre d'une infrastructure répartie de stockage, de traitement et de mise à disposition, diffusion et valorisation des données, hébergées dans des conditions environnementales à faible empreinte carbone.

Cette organisation, en s'appuyant pour partie sur les équipements de GENCI à l'IDRIS, offrira aux IR des moyens de calcul et de traitement de données alliant calcul haute performance et calcul haut débit, supercalculateur disposant d'accélérateurs graphiques et ferme de traitement de données, conteneurisation et « *cloud computing* ». Cette plateforme s'inscrira dans l'offre française comme une composante essentielle du *European Open Science Cloud (EOSC).*



PRACE-6IP

PRACE, « *Partnership for Advanced Computing in Europe* », est l'infrastructure de recherche paneuropéenne de l'ESFRI fournissant depuis 10 ans des services de pointe pour permettre une recherche d'excellence sur des systèmes de calcul hautes performances de classe mondiale.

PRACE aisbl, ses 26 membres ainsi que EUDAT et GÉANT collaborent au projet PRACE-6IP. Cette 6e « *Implementation Phase* » sera la dernière financée par la Commission européenne, à l'avenir elle sera remplacée par un ensemble de projets dédiés à des services spécifiques, financés par EuroHPC. Le projet PRACE-6IP s'appuie sur les réalisations des projets d'implémentation précédents et les poursuit avec de nouvelles activités innovantes et collaboratives, dans le cadre d'un écosystème toujours plus complexe.

Le projet était prévu pour s'achever à la fin de 2021, mais certaines de ses activités, telles que les EuroHPCSummitWeek 2021 et 2022 se prolongeront jusqu'en juin 2022 avant qu'EuroHPC JU ne lance de nouveaux appels à projets.



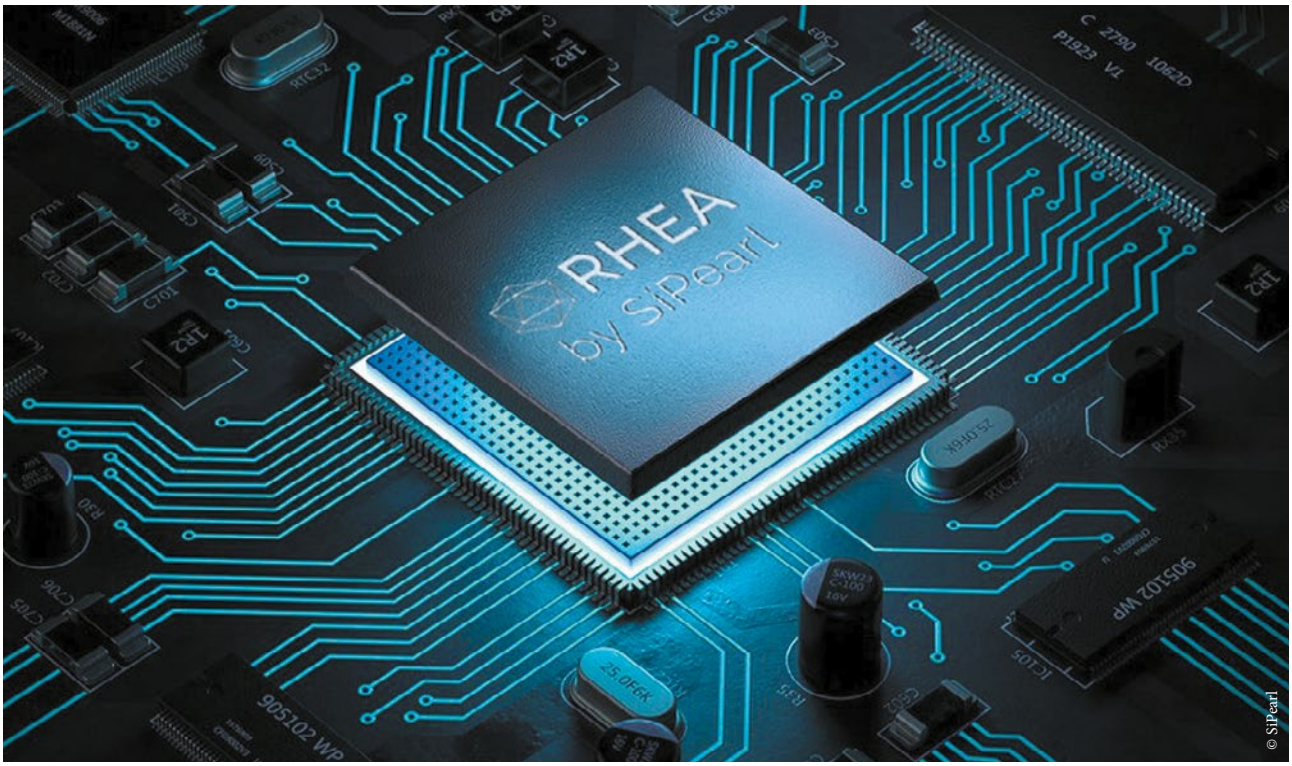
En 2021 le projet PRACE-6IP a également collaboré avec ETP4HPC, les centres d'excellence (CoE) et d'autres organisations européennes et internationales sur les futures architectures, la formation, le support applicatif et les politiques d'accès aux ressources, ainsi qu'avec d'autres e-Infrastructures ou projets tels que EUDAT et FENIX (offrant leurs ressources dans les appels PRACE et collaborant sur l'AAI), et des infrastructures telles que GÉANT, ESRF, SKA et le CERN, tout en favorisent les relations avec le « *European Open Science Cloud* » (EOSC).

Le projet a également poursuivi des activités de veille technologique indépendante, de formation avancée, qui a déjà fourni près de 70 000 jours de formation pour des participants venant de 96 pays, et l'expansion au niveau européen du réseau des centres de formation PRACE, ainsi que la promotion de l'adoption et de l'utilisation du HPC par l'industrie et de nouvelles communautés d'utilisateurs et offrant des services spécifiques aux PME.

TOUT SAVOIR SUR PRACE

www.prace-ri.eu

EN ROUTE POUR UN PROCESSEUR EUROPÉEN EXASCALE AVEC LES PROJETS EUPEX ET EPI SGA2 FR



© SiPearl

Le projet EUPEX, EUropean Pilot for EXascale, s'inscrit dans la vision du développement de technologies souveraines qui pourront équiper les futures architectures *Exascale* européennes. Cofinancé par la Joint Undertaking EuroHPC ainsi que la France (via Bpifrance), l'Allemagne, l'Italie, la Grèce, la République tchèque et la Croatie, il est doté d'un budget de 40,8 M€ sur 4 ans et regroupe 19 partenaires industriels et académiques (dont en France Atos, le CEA, Inria, GENCI et la PME Cybele Tech). EUPEX vise à déployer et valider la première plate-forme européenne pour le HPC intégrant l'ensemble des technologies européennes financées au préalable par l'Europe, du processeur issu de la première phase du projet de recherche européen EPI (*European Processor Initiative,* SGA1) à la couche logicielle et jusqu'aux applications. Le projet EUPEX, qui a démarré en ce début d'année 2022, s'appuiera initialement sur la plateforme innovante A64FX installée sur Joliot Curie au TGCC, en attendant la disponibilité du processeur Rhea de la société SiPearl. Elle permettra aux développeurs de porter leurs applications sur un processeur embarquant des caractéristiques similaires au processeur européen Rhea.

Dans la continuité du projet EPI SGA1, l'Europe poursuit son support à la définition de la deuxième génération du processeur européen et celle d'un accélérateur et engage un financement de 70 M€ dans le projet **EPI SGA2** qui démarrera en février 2022 pour 3 ans.

LA SOCIÉTÉ CIVILE GENCI

Des femmes et des hommes au service de la recherche 52
Organisation interne 53
Les membres des Comités et des Commissions 54



DES FEMMES ET DES HOMMES AU SERVICE DE LA RECHERCHE

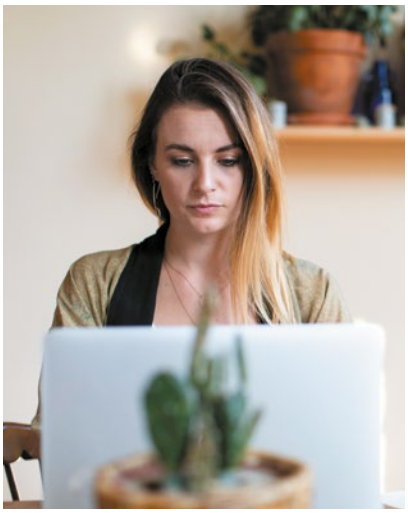
—— **GENCI a poursuivi les efforts engagés en 2020 sur la voie de la dématérialisation des processus internes et leur amélioration continue.**
GENCI a poursuivi la dématérialisation des opérations financières par le traitement et la validation des factures via SAGE automatisation comptable. Ce système permet ainsi de disposer des factures sous format dématérialisé et d'enregistrer l'ensemble des actions (dates de validation et valideur) sur chaque dépense de la société. Le contrôle interne est ainsi renforcé et les risques de fraudes ou d'omission sont considérablement réduits.

A la suite d'une réflexion menée en interne et d'une démultiplication des opérations avec les centres de calcul, une matrice a été créée pour recenser les engagements contractualisés et permettre un suivi opérationnel et financier. L'originalité de cette démarche repose sur le fait que l'outil soit accessible à tous les acteurs de GENCI (juridique, financier, technique et opération) ou du centre de calcul concerné.

Enfin, dans l'optique d'une amélioration continue des processus, GENCI a déployé un nouveau module RH « Pagga » développé par la société Lucca permettant aux salariés de recevoir par voie dématérialisée dans un « coffre-fort » individualisé leurs feuilles de paie en toute sécurité.

—— **Les mesures mises en place en 2020, relatives à la crise sanitaire COVID -19, ont été prolongées jusqu'en septembre 2021.**
Un protocole sanitaire a été mis en place et s'applique au personnel ainsi qu'à tout visiteur extérieur.

GENCI n'a rencontré aucune difficulté à assurer la continuité de service puisque le télétravail avait été mis en œuvre dès 2019. Fort de ce succès sur le plan organisationnel, le télétravail a été étendu sur initiative de l'employeur et après une consultation approfondie avec le personnel, à deux jours par semaine à compter du 1er septembre 2021.



En outre, début 2021, les salariés ont bénéficié d'une enveloppe financière de 450 euros maximum par salarié pour prendre en charge des équipements de bureau nécessaires au télétravail régulier à leur domicile.

—— **Sur le plan des ressources humaines, l'année a été marquée par plusieurs événements.**
En premier lieu, les Associés se sont prononcés à l'unanimité, le 2 juin 2021, en faveur du renouvellement de Philippe Lavocat au poste de PDG pour une durée de 5 ans.

A la suite du départ du représentant du personnel mi-août, il a été procédé à de nouvelles élections du Comité Social et Economique. Ces dernières se sont tenues en octobre. Aucune candidature n'ayant été présentée, un procès-verbal de carence a été rédigé.

Les effectifs se sont établis à vingt personnes pour 2021. GENCI a recruté deux nouveaux salariés en CDI pour les postes de responsable de communication (avril) et de chargé administratif de projets (avril) ainsi qu'un recrutement en intérim pour le poste de chef de projet *Exascale* (février). En parallèle le chef de projet « Competence Center/MesoNet », a quitté ses fonctions mi-août.

—— **Sur le plan des achats publics,** le résultat de la procédure de dialogue compétitif pour le renouvellement des moyens de calcul du CINES a été contesté devant le juge judicaire par un des candidats évincés. Les équipes de GENCI appuyées par celles du CINES ont mené un très important travail d'analyse afin de contester les arguments soulevés par ce candidat dans le cadre d'un référé précontractuel. L'audience s'est tenue le 19 juillet 2021. L'ordonnance du juge rendue le 25 juillet a débouté la demanderesse de l'entièreté de ses demandes.

—— **Enfin, GENCI a finalisé la mise en œuvre des processus découlant de l'étude RGPD** engagée sur la relation de sous-traitance de GENCI et de l'IDRIS par rapport aux utilisateurs des supercalculateurs/responsables de traitement de données personnelles (sensibles ou non). L'étude menée par GENCI ayant permis d'identifier les sous-traitants de 1er et de 2e niveau (respectivement GENCI et les centres), les documents de mise en conformité des relations utilisateurs/IDRIS/GENCI pour l'attribution de ressources de calcul ont été validés par les parties concernées.

Ce travail réalisé avec les équipes du CNRS et de l'IDRIS pour répondre à un besoin urgent va désormais se poursuivre avec les deux autres centres de calcul (TGCC et CINES) en 2022. ■

ORGANISATION INTERNE

COMITÉ DE DIRECTION



Philippe LAVOCAT
Président-directeur général



Edouard BRUNEL
Secrétaire général



Jean-Philippe PROUX
Responsable des Opérations
et de la Sécurité



Stéphane REQUENA
Directeur Technique et Innovation



Marie-Hélène VOUETTE
Responsable de partenariats
Conseiller relations institutionnelles

FONCTIONS SUPPORT ET ADMINISTRATIVES



Maïté CAMPEAS
Assistante de direction



Imène LITIM
Assistante de Direction



Laetitia PHO
Responsable Administratif et Financier



Marieke PODEVIN
Juriste d'affaires senior
Déléguée à la protection des données



Rasa VAICIEKAUSKAITE FALLAHZADEH
Chargée administrative de projets

COMMUNICATION



Nicolas BELOT
Responsable communication
et relations européennes



Annabel TRUONG
Chargée de communication

SERVICES TECHNIQUES ET OPÉRATIONNELS



Christelle PIECHURSKI
Responsable HPC et en charge
de la veille technologique



Elise QUENTEL
Chef de projets nationaux
et européens : MesoNet, SiMSEO



Philippe SEGERS
Chef de projets européens



Corinne BEAL
Chef de projet *Exascale*



Delphine THEODOROU
Chargée de projets
opérations



Pascal VOURY
Chargé de projets opérations



Abraham LUNDA
Chargé de projets
Competence Center

LES MEMBRES DES COMITÉS ET DES COMMISSIONS

AU 31/12/2021

CEA : Commissariat à l'Énergie Atomique et aux Énergies Alternatives. **CNRS :** Centre National de la Recherche Scientifique. **France Universités. Inria :** Institut national de recherche en informatique et en automatique. **MESRI :** Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

Conseil GENCI

Représentant l'État : M. Laurent Crouzet - M. Nicolas Dromel. Représentant le CEA : Mme Maria Faury - M. Hervé Desvaux. Représentant le CNRS : M. Ali Charara - M. Denis Veynante. Représentant les Universités : M. Guillaume Gellé - M. Olivier Simonin. Représentant Inria : M. Jean-Frédéric Gerbeau.

Comité Consultatif Administratif et Financier (CCAF)

Représentant l'État et le MESRI : M. Antoine Perrang, *Président du CCAF.* **Représentant le CNRS : Mme Anne-Françoise Duval. Représentant le CEA : M. Rémy Keller. Représentant Inria : M. Vincent Heyer. Représentant les Universités : M. Michel Dellacasagrande.**

Commission des Marchés (CM)

Représentant l'État : M. Philippe Ajuelos, *Administrateur ministériel des données des algorithmes et des codes sources – Direction du numérique pour l'éducation du Ministère de l'Éducation Nationale et de la Jeunesse, Président de la Commission* - **Mme Lara Montantin,** *Juriste à la direction générale de la recherche et de l'innovation du Ministère de l'Éducation Nationale.* **Représentant les Universités : M. Hervé Celestin,** *Responsable administratif et financier de l'AMUE.* **Représentant le CEA : M. Eric Stehle,** *Directeur des Achats et des Partenaires Stratégiques de Paris-Saclay* - **M. Thibault Pelletier,** *Chef de service du référentiel et de la Commission des marchés.* **Représentant le CNRS : M. Sébastien Turci,** *Directeur délégué aux achats et à l'innovation* - **Mme Marie-Laure Colin,** *Chef du bureau de la réglementation de la direction des affaires juridiques.*

Comité Technique (CT)

Représentant l'État et le MESRI : M. Laurent Crouzet, *Chef du Département Services et Infrastructures Numériques.* **Représentant le CEA : Mme France Boillod-Cerneux,** *Direction de la Recherche Fondamentale* - **M. Emeric Brun,** *Direction de l'Énergie Nucléaire* - **Mme Christine Menaché,** *Responsable du TGCC.* **Représentant le CNRS : M. Michel Daydé,** *Délégué Scientifique* - **M. Pierre-François Lavallée,** *Directeur de l'IDRIS.* **Représentant les Universités : M. Francois Bodin,** *Professeur Université de Rennes 1* - **M. Boris Dintrans,** *Directeur du CINES.* **Représentant Inria : M. Jean Roman,** *Directeur Scientifique Adjoint auprès de la Direction de la recherche en charge du domaine « Mathématiques appliquées, calcul et simulation »* - **M. Lucas Nussbaum,** *Responsable du programme plateformes d'expérimentation à la Direction Générale Déléguée à l'Innovation.*

Comité d'Orientation Stratégique des Investissements (COSI)

Comité en charge de conseiller et d'apporter une assistance à GENCI dans sa stratégie d'investissements en produisant notamment un plan pluriannuel d'investissements à partir des orientations du comité des besoins scientifiques du MESRI et orientant les travaux du CT et du CCAF.

Représentant l'État et le MESRI : M. Laurent Crouzet, *Chef du Département Services et Infrastructures Numériques.* **Représentant le CEA : M. François Robin,** *Adjoint au Directeur DAM-Ile de France* - **M. Christophe Calvin,** *Direction de la Recherche Fondamentale.* **Représentant le CNRS : M. Ali Charara,** *Directeur de l'Institut des Sciences Informatiques et de leurs Interactions* - **M. Denis Veynante,** *Président du comité directeur de la mission Calcul-Données.* **Représentant les Universités : M. Olivier Simonin,** *Président de l'Institut National Polytechnique de Toulouse.* **Représentant Inria : M. Jean Roman,** *Directeur Scientifique Adjoint auprès de la Direction de la recherche en charge du domaine « Mathématiques appliquées, calcul et simulation ».*

Un grand merci aux rédacteurs et contributeurs pour le temps et l'énergie que vous avez consacrés à ce rapport retraçant l'ensemble de l'activité de GENCI pendant l'année 2021.

C'est grâce à tous les acteurs présents dans ce rapport et à bien d'autres encore que GENCI peut atteindre ses objectifs et voir de nouvelles perspectives se dessiner chaque année.



Directeur de la publication : Philippe LAVOCAT - **Coordination :** Annabel TRUONG - **Conception et réalisation :** backdrop
Impression : Welcom - **Photos :** © Cyril FRESILLON/IDRIS/CNRS Photothèque, © Nathanaël SCHAEFFER / ISTERRE / CNRS Photothèque, © CNRS Photothèque, © CEA, © CEA/Peinture C215, © P. Stroppa, © URCA, © AdobeStock, © French Tech Central, DR.

GENCI - 6 bis, rue Auguste Vitu - 75015 Paris - France. Tél. : +33 1 42 50 04 15

© GENCI - mars 2022

GENCI
6 bis rue Auguste Vitu
75015 Paris - FRANCE
www.genci.fr

Suivez GENCI !





Le calcul intensif au service de la connaissance