

Inria

Refonte des Moyens de Calcul Inria

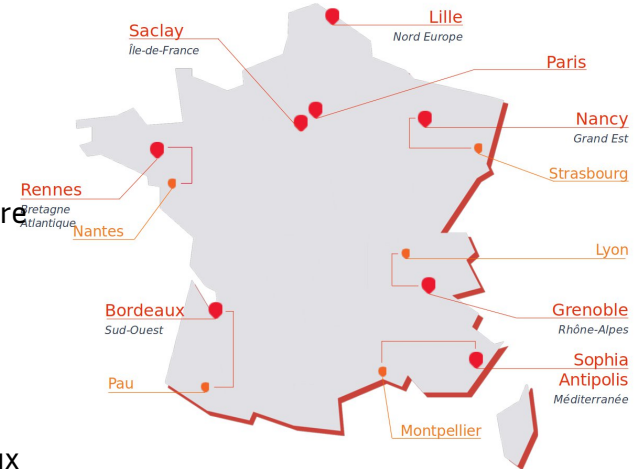
Point d'étape

Lucas Nussbaum

2 octobre 2023

Les Moyens de Calcul de l'Inria en 2019

- Inria : EPST, 9 centres, 4700 personnes (dont 2000 partenaires)
- Des « **clusters de centre** » dans la plupart des centres Inria
 - Mutualisés à l'échelle du centre, parfois ouverts aux extérieurs au centre
 - Mais sans proposer une *vue unifiée* aux utilisateurs
 - Opérés par les mêmes personnes au sein de la DSI unifiée
 - Mais avec des choix techniques historiques différents
- Une **plateforme d'expérimentation (Grid'5000)**
 - En théorie, pour la communauté de recherche en Cloud/IA/HPC/réseaux
 - Déjà utilisée plus largement par des équipes d'autres domaines
- Des clusters d'équipes, non mutualisés
- Des moyens de calcul associés à des plateformes d'expérimentation d'autres domaines (exemple : Cyber, Réalité Virtuelle & Augmentée)
- Recensement en 2019 : 38000 coeurs CPU
 - clusters de centre : 14500 + Grid'5000 : 12000 + équipes : 5000 + plateformes : 6500



Cible à terme : une offre clarifiée de Moyens de Calcul (et de stockage associé) pour les équipes Inria

● Infrastructure nationale de Moyens de Calcul

- > En cours de construction, achevée premier semestre 2026 ?
- > Avantages : accès simple et rapide (quelques minutes) ; adapté aux besoins spécifiques des sciences du numérique ; articulé avec nos plateformes d'expérimentations
- > Inconvénients : pas adapté aux données sensibles

● Plateforme Arcana (données sensibles)

- > Travail en cours → lancement début 2024
- > Avantages : adapté au traitement de données sensibles (y compris de santé) sur CPU+GPU ; adapté aux besoins spécifiques des sciences du numérique
- > Inconvénients : la sécurisation implique forcément des contraintes supplémentaires pour les usagers ; coût élevé

● Supercalculateurs nationaux (GENCI + centres nationaux : Jean Zay, Joliot-Curie/Irene, Aداstra, ...)

- > Déjà disponible, mais modalités d'accès à améliorer pour les équipes Inria (ex : documentation côté Inria)
- > Avantages : beaucoup de ressources disponibles (en particulier GPU) ; gratuit pour Inria
- > Inconvénients : pas adapté aux données sensibles ; processus d'accès assez complexe et long (semaines voire mois) et pas accessible à tous les chercheurs (ZRR) ; pas adapté à tous les besoins spécifiques des sciences du numérique

... à articuler avec l'offre d'Inria sur les données

Cible à terme : une infrastructure nationale de moyens de calcul

- Répondant aux **besoins spécifiques de la recherche en sciences du numérique** (flexibilité, reconfiguration, reproductibilité)
- **Opérée par Inria mais ouverte plus largement** (avec mécanisme de priorisation)
 - Aux chercheurs de sciences du numérique hors Inria
 - Aux chercheurs d'autres domaines avec des besoins spécifiques
 - **Et aussi aux participants aux JCAD, si vous souhaitez suivre notre travail**
- **Grid'5000/SLICES-FR** probablement repositionnée comme une **plateforme d'expérimentation s'appuyant sur l'infrastructure nationale de moyens de calcul**
 - Avec des services spécifiques à SLICES-FR
 - Positionnement similaire pour d'autres plateformes : réalité virtuelle, robotique, génie logiciel ...
- Avenir des **clusters de centre** : intégrés dans l'infrastructure nationale
 - Ou maintenus comme des plateformes d'expérimentation spécifiques, si pertinent (?)

Cible à terme : une infrastructure nationale de moyens de calcul

2019

Clusters de centre
Igrida, NEF, Plafrim, Margaret, ...

Grid'5000

intégration dans

sert de base technologique

sert de base technologique

2024

Arcana
(données sensibles)

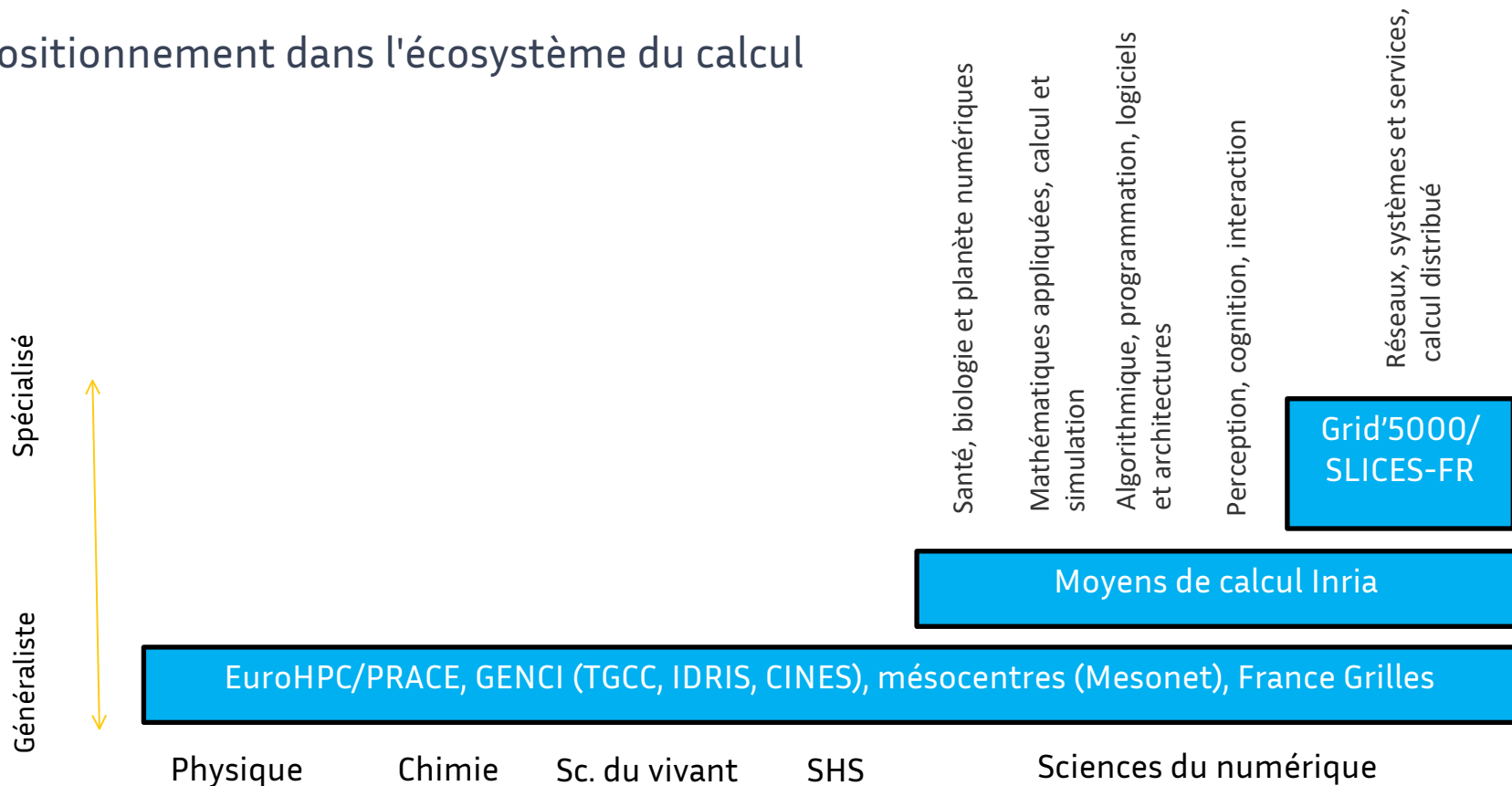
↔
convergence
quand pertinent

Infrastructure nationale
de moyens de calcul Inria

SLICES-FR s'appuiera
probablement sur
l'infra nationale Inria
pour son volet HPC/Cloud

SLICES(-FR)

Positionnement dans l'écosystème du calcul



Pourquoi ne pas s'appuyer sur les mésocentres ?

- Pas forcément de mésocentre proche de chaque centre Inria
- Besoins spécifiques des sciences du numérique (instrumentation, reconfiguration) pour une part importante des utilisateurs
 - Redéploiement bare-metal (avec kadeploy) dans l'offre de service
 - En général pas une priorité pour les mésocentres (et c'est normal)
- Réseau des mésocentres assez informel vs organisme national (DSI unifiée)
 - Obligerait à faire le même travail avec plusieurs mésocentres pour renforcer leur offre de service et supporter les besoins spécifiques → difficile de mutualiser les efforts

De nombreux sujets communs avec les mésocentres et Mesonet

- Hébergement de moyens de calcul dans les DC régionaux labellisés
- Stockage (fichier vs objet ; y compris dans un contexte géo-distribué ; articulation avec archivage/entrepôts)
- Environnement logiciel pour les utilisateurs (modules, Spack, GUIX, NIX, Conda/Mamba, Docker/Podman, Singularity/Apptainer, ...), y compris outils haut niveau (Jupyter, portails)
- Règles de partage de ressources mutualisées
- Sécurité, HDS, authentification, ...
- Impact énergétique et financier
- Outils plus haut niveau pour utiliser les ressources (Jupyter, portails divers)
- Besoins spécifiques (ex : formation)

Augmenter les échanges ?

Avancées depuis 2021 (et la présentation précédente aux JCAD)

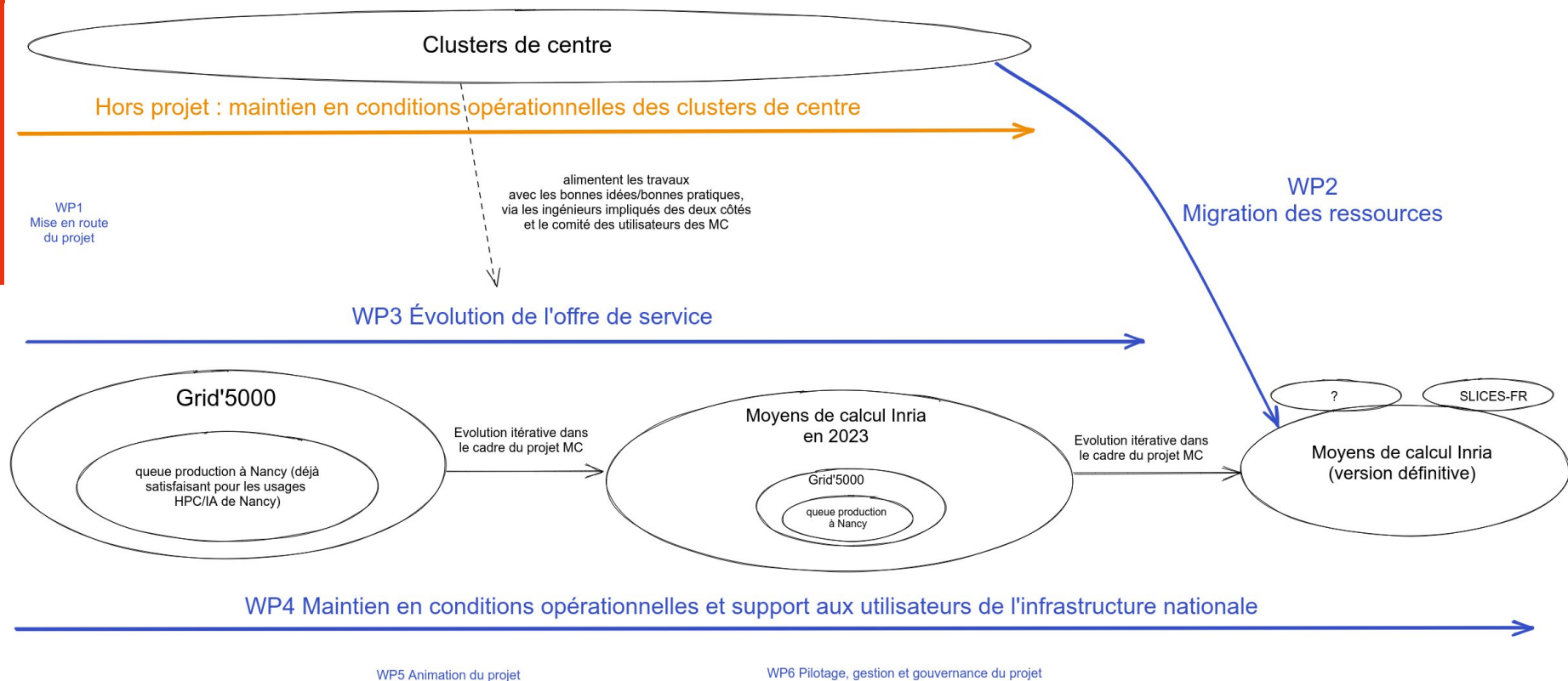
- Constitution de l'équipe d'ingénieurs pour travailler sur le projet
 - Constituée d'ingénieurs de la DSI Inria et de la ligne « développement technologique » Inria (+ 1 CNRS-LORIA)
 - Au 01/10/2023 : 6 ETP IR (permanents) + 1.5 ETP CDD
- Accueil puis montée en compétences des nouveaux contributeurs
- Mise en place de l'organisation du projet
 - Y compris comité des utilisateurs (réuni tous les 1.5 mois)
- Résultats concrets :
 - **Mise en place de l'« offre logicielle » : modules, GUIX, Conda/Mamba, Docker, Singularity/Apptainer, sudo-g5k, images kadeploy**
 - Migration du cluster de centre de Rennes (Igrida) vers l'infrastructure nationale en cours

Merci

Contributeurs actuels au projet :

Yao Chi, Théo Eisenbart, Brice Goglin, Julien Lelaurain, Alexandre Merlin,
Laurent Mirtain, Nicolas Niclausse, Nicolas Perrin, Laurent Pouilloux,
Patrice Ringot, Philippe Virouleau

Projet « Refonte des moyens de calcul Inria »



WP	Tâche
WP1 Mise en route	
WP2 Migration des ressources	
WP3 Évolution de l'offre de service	T3.1 - Offre logicielle
	T3.2 - Partage des ressources
	T3.3 - Gestion ressources alternatif (Slurm)
	T3.4 - Doc utilisateur
	T3.5 - Stockage
	T3.6 - Passage OAR/Slurm
	T3.7 - Autres évolutions / retours utilisateurs
	T3.8 - Gestion comptes utilisateurs
	T3.9 - Notebooks Jupyter
	T3.10 - Impact énergétique & financier
	T3.11 - Convergence outils Inria
WP4 Maintien en conditions opérationnelles et support aux utilisateurs	
WP5 Animation du projet	
WP6 Pilotage, gestion de projet, gouvernance	

Focus : stratégie d'hébergement

- En cohérence avec la **stratégie du MESR de consolidation d'un système fédéré de datacenters**, qui a vocation à concentrer l'hébergement des grandes infrastructures numériques de l'état pour l'ESR dans des **datacenters régionaux labellisés**
- Posture : chercher à **héberger les moyens de calcul dans des datacenters régionaux labellisés** par le MESR, ou à défaut, quand ce n'est pas possible, dans des datacenters d'efficacité énergétique comparable.
 - Possible qu'à terme, la **majorité des ressources soit hébergée dans quelques datacenters régionaux**, en visant par exemple 2 à 4 datacenters régionaux, plutôt que de chercher à utiliser tous les datacenters régionaux à proximité de centres Inria.
- Dans certains cas, il faut toutefois s'autoriser à héberger des moyens de calcul dans les centres, par exemple pour des **raisons techniques de proximité avec une plateforme d'expérimentation** (contraintes de latence réseau ou de débit réseau). Mais ces cas devraient rester limités aux situations où un tel besoin est identifié.

Évolutions de 2021 à l'état cible

Objets technologiques

			Clusters de centre	Infrastructure nationale de moyens de calcul	Grid'5000 → SLICES-FR	
Dimensions	Produit	Fin 2021	6 sites (Bo/Gr/Pa/Re/Sa/So), 600 noeuds	∅ (à l'exception de la queue production de Nancy)	8 sites (Gr/Li/Lu/Ly/Ncy/Ntes/Re/So), 40 clusters, 800 noeuds	
		Cible (2026?)	∅ (Ressources intégrées à l'infra nationale)	Rassemble les ressources de calcul d'Inria, avec une gestion et des mécanismes de réservation homogènes (sauf exceptions). Hébergement en cohérence avec la stratégie nationale de labellisation de datacenters et les besoins de localité des PFE	Ressources « calcul » intégrées à l'infra nationale, probablement avec des règles d'utilisation spécifiques à Grid'5000/SLICES-FR + autres ressources (IoT/Wireless/...) hors périmètre moyens de calcul	
	Offre de service (services offerts aux utilisateurs)	Fin 2021	Des offres de service similaires dans chaque centre, mais avec des implémentations assez différentes	∅ (sur Grid'5000, une offre de service pour les usages « calcul » de qualité limitée)	Une offre de service visant d'abord à répondre aux besoins en expérimentation de la communauté RSD, mais permettant déjà de répondre à des besoins plus génériques	
		Cible (2026?)	Plus d'offre de service spécifique à chaque centre (sauf exception).	Une offre de service pour les usages « calcul » consolidée, définie grâce à l'expérience des clusters de centre	Des services spécifiques aux usages « expérimentation », et un socle commun de services fournis par l'infra nat. de MC	
	Projet	Projet et ressources humaines	Fin 2021	RH: maintenance par DSI ou SED (selon les centres). Pas de projet formalisé pour cadrer le travail	∅	Pas de projet formalisé pour l'instant RH: fournies par des affectations d'ingénieurs permanents CNRS, et des CDD Inria hors ADT
			Cible (2026?)	∅ Plus de besoins de maintenance ou d'évolutions	Projet « Refonte des moyens de calcul Inria ». Ensuite, à définir. Probablement une équipe transverse DSI/Dev-Tech chargée de la maintenance et des évolutions de l'infrastructure	Des moyens humains pour maintenir les services spécifiques, fournis par le PEPR Cloud et peut-être des projets européens liés à SLICES