



Le projet FITS

→ 27/11/2024

Plusieurs problématiques (1)

Explosion du volumes des données des IR/IR*

La feuille de route 2018-2022 des IR/IR* du MESR prévoyait qu'en 2022, le volume de données gérés par les IR/IR* serait multiplié par ~5

Avec une situation est variable d'une IR/IR* à l'autre :

taux d'augmentation très important à partir d'un volume très faible

x500 en partant de ~ 10 To

alors que d'autres partent d'un volume déjà conséquent

x5 en partant de ~ 100 Po

La dernière mise à jour de la FDR indique qu'en 2020 la production des données de l'ensemble des IR est de l'ordre 1 250 Po

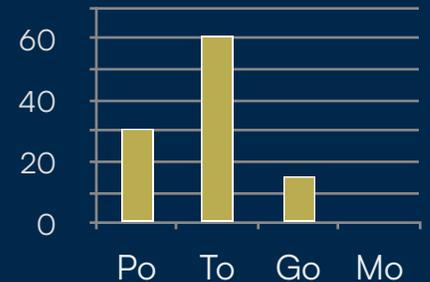
Comment répondre à ce besoin ?

Feuille de route des IR/
IR*

~ 100 infrastructures

> 2/3 des IR/IR* ont
pour porteur principal le
CNRS

Tous les domaines sont
concernés



Nombre d'IR en fonction de la grandeur du
volumes de données produites annuellement

Plusieurs problématiques (2)

Accès aux infrastructures de calcul

Les modèles de traitement de données sont de plus en plus complexes
nécessitent l'accès à des infrastructures de calcul multiples
calcul haut performance (HPC) — calcul haut débit (HTC)
cloud computing / virtualisation
IA

Modernisation - Mutualisation des lieux d'hébergement

La très grande fragmentation des lieux et la variété des conditions d'hébergement
nécessitent la mise en place de politiques de mutualisation et rationalisation afin
d'être en mesure de proposer des datacenters à l'état de l'art.

résilience des installations
réduction de l'empreinte carbone

IDRIS

HPC - IA

Hébergement

CC-IN2P3

HTC

**Grille de
calcul**

Hébergement

Interconnexion

100 Gbs

AMI ESR / Equipex +

Objectifs (extraits)

Mettre en place des équipements ouverts à toutes les communautés scientifiques

Un tarif d'accès à ces moyens doit permettre de couvrir les charges de fonctionnement, de jouvence ou — en partie — de renouvellement.

Actions financées dans le cadre de l'action ESR/EquipEx+ :

des adaptations ou extensions transformant les grands centres de calcul nationaux existants en datacenters opérés, sécurisés, hébergeant les infrastructures et services numériques, y compris le stockage des données des infrastructures partagées: serveurs, stockage, archivage ;

des services dédiés à couvrir les besoins numériques des infrastructures de recherche de la feuille de route nationale et, notamment, des services numériques destinés à prendre en charge le cycle de vie de leurs données;

Projet FITS

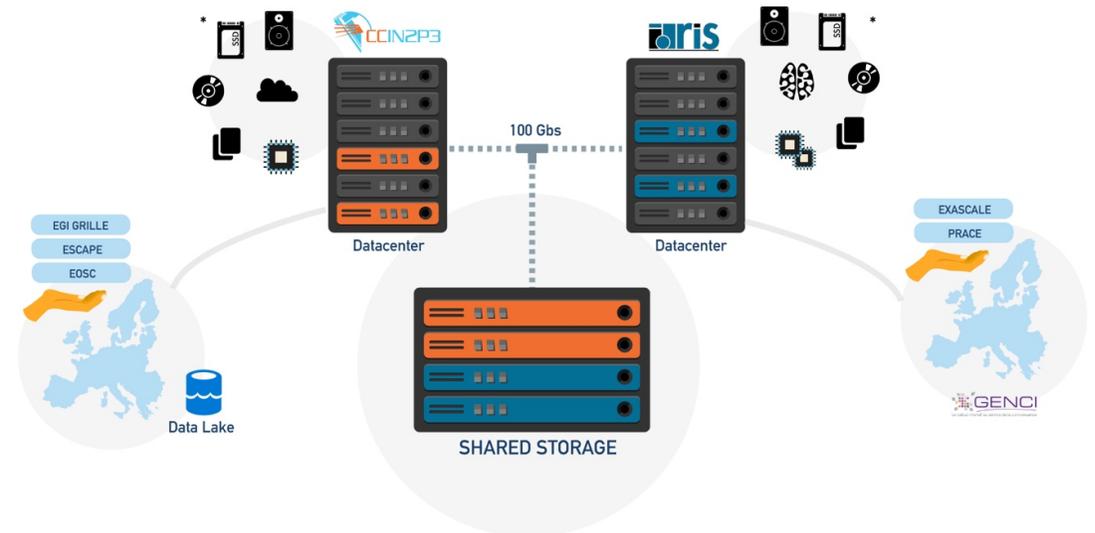
L'ambition du projet FITS est de fédérer les services et savoir-faire de l'IDRIS et du CC-IN2P3, dans le respect de leurs missions spécifiques, à travers la mise en œuvre d'une infrastructure répartie de stockage, de traitement et de mise à disposition, diffusion et valorisation des données, hébergées dans des conditions environnementales à faible empreinte carbone.

Objectifs

Permettre aux IR, disposant d'un modèle de calcul établi et nécessitant l'accès à des technologies variées d'y accéder de manière simple et facile
Leur proposer des moyens d'accès et de distributions à leurs données
Permettre aux IR disposant de leurs infrastructures de les héberger et opérer dans un environnement à l'état de l'art

A travers

- Mise à niveau des capacités d'accueil des deux centres
- Déploiement d'une infrastructure de stockage partagée entre eux
- Déploiement d'un portail d'accès
- Validation par des cas d'utilisation scientifique
- Durable : Modèle économique



FITS

CNRS Federated IT services
for Research Infrastructures



ANR-21-ESRE-0009

IDRIS

CC-IN2P3

partenariat avec GENCI

Budget total

15 520 000 €

75 %

Opérations immobilières

25 %

Hardware

Avancement du projet

Stockage distribué et portail d'accès

Capacités déployées à ce jour

Au CC-IN2P3:

10 To (disque) pour le stockage primaire

22 Po en bandes magnétiques pour stockage

A l'IDRIS :

5 Po stockage chaud : 5 Po

20 Po de stockage froid (bandes magnétiques)

Les briques de bases sont fonctionnelles

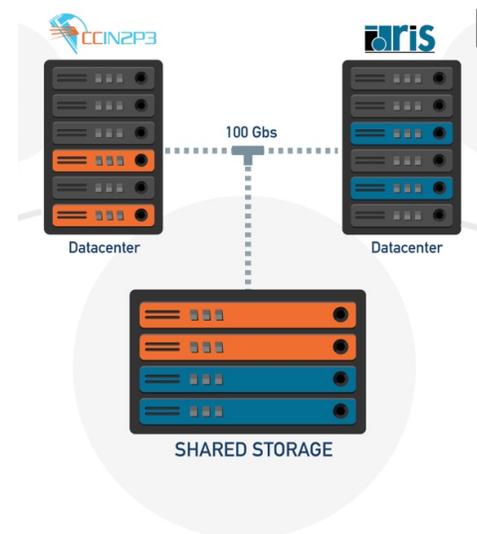
- gestion des identités et des accès
- gestion de la donnée

et accessibles à travers un prototype du portail web

Accès aux ressources de calcul

Prototype prévu pour S1 2025

En collaboration avec GENCI pour l'accès aux ressources HPC/IA



Avancement du projet

Mise à niveau des infrastructures d'accueil à l'IDRIS

Travaux en site occupé sans foncier disponible

Zones concernées

1. Poste de livraison HT — nouvelles cellules HT
2. Postes transformation — nouveaux transformateurs de tension + nouveaux onduleurs et batteries + CTAs
3. Nouvelles pompes et échangeurs de chaleur
4. Nouveau local électrique dédié à la SM
5. Nouveaux tableaux électriques
6. Remplacement des groupes de production d'eau glacée + de nouveaux refroidisseurs adiabatiques

Achèvement prévu : T2 2026

Marchés de travaux (procédures longues, réglementation stricte et contraignante)



Avancement du projet

Mise à niveau des infrastructures d'accueil au CC-IN2P3

Foncier disponible

Réalisation d'une extension destinée à remplacer la salle
« historique »
Doublement des capacités d'hébergement (~150 baies à 15 KVA)
PUE en baisse de 0,5

Planning prévisionnel

Début des travaux : 6 janvier 2025
Durée prévue 12 mois



Opération de conception-réalisation

Prochaines étapes

2025

Achèvement :

- des travaux de mise à niveau des capacités d'accueil des deux centres
- du déploiement d'une infrastructure de stockage partagée et l'intégration des moyens de traitement de données

Le financement du projet permet de se mettre dans la meilleure situation possible pour relever les défis de la gestion du stockage des données des IR/IR*.

Au delà de ce financement initial, la durabilité de l'équipement (jouvence, accroissement de capacité, etc..) et les couts d'opérations devront être supportés par les utilisateurs.

Des tarifications sont à établir :

- différenciant les qualités de services de stockage
- les ressources de calcul
- et l'hébergement

2026

Ouverture progressive