

Projets européens au CC-IN2P3

Ghita Rahal

24/04/2019

- ▶ Incitations politique et stratégique:
 - Stratégies de calcul et financements liés à de grands ensembles supra-CC (France, Europe,...)
 - EOSC, HNSciCloud
 - Mutation rapide de l'environnement du calcul
 - Volonté de lever les barrières informatiques: fédération de centres, utilisation de ressources disséminées, optimisation des ressources et harmonisation des modes de calcul, etc..
 - Volonté soulignée à l'IN2P3

► Incitation pour le CC

- Nous avons des compétences et nous voulons les garder/fructifier
 - Être partenaire de toute (r)évolution technologique
 - Nous pouvons avoir plus de R(H) via ces modes de financements.
- Nous sommes intéressés par les techniques étudiées et mises en place:
 - couches d'interopérabilité et de fédération, transferts de données, virtualisation, containerisation,
 - Base des futurs déploiements sur architecture distribuée (WLCG, astro..)
- Faire partie de la communauté et s'y faire reconnaître pour préserver notre "know how" de grand centre de calcul.
- Nouer des collaborations fructueuses en termes d'échange de savoir

- ▶ Origine variée:
 - Sollicitation de l'IN2P3, des partenaires de projets précédents ou d'autres centres de calcul
 - Décision de participation plus ou moins individuelle
- ▶ Fonctionnement actuel:
 - Projets souvent "isolés" des activités du CC
 - Plusieurs projets sur les même thèmes sans coordination interne.
 - Les projets finis, les compétences disparaissent

► Problèmes relevés:

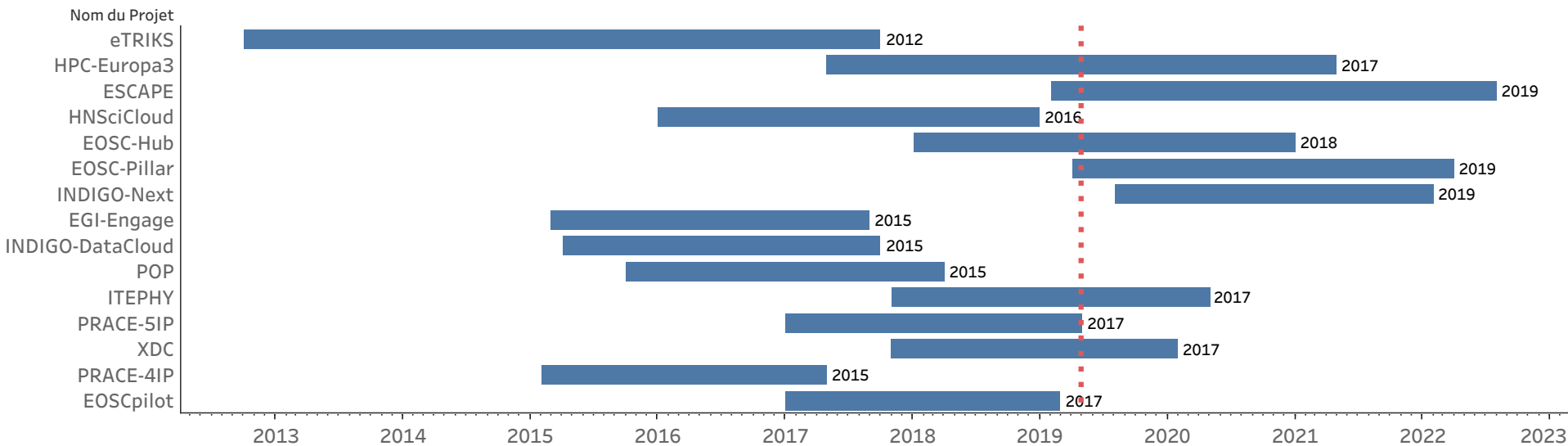
- Participation “passive”: nous attendons qu’on nous sollicite
- Difficulté de trouver des participants lorsqu’il y a de nouveaux projets
 - Temps de prise en charge des projets et formation des nouveaux CDDs non négligeable.
 - Perte de RH dû à cette prise en charge sans contrepartie
- Pas d’appropriation des thématiques au niveau CC:
 - Les CDDs sont cantonnés sur leur projet et travaillent de manière isolée
 - Risque sur les projets si le CDD engagé part ailleurs (aucun backup)
 - Pas d’appropriation des thématiques étudiées, bien que souvent elles peuvent recouper des besoins du CC

▶ Projets Européens en cours au CC

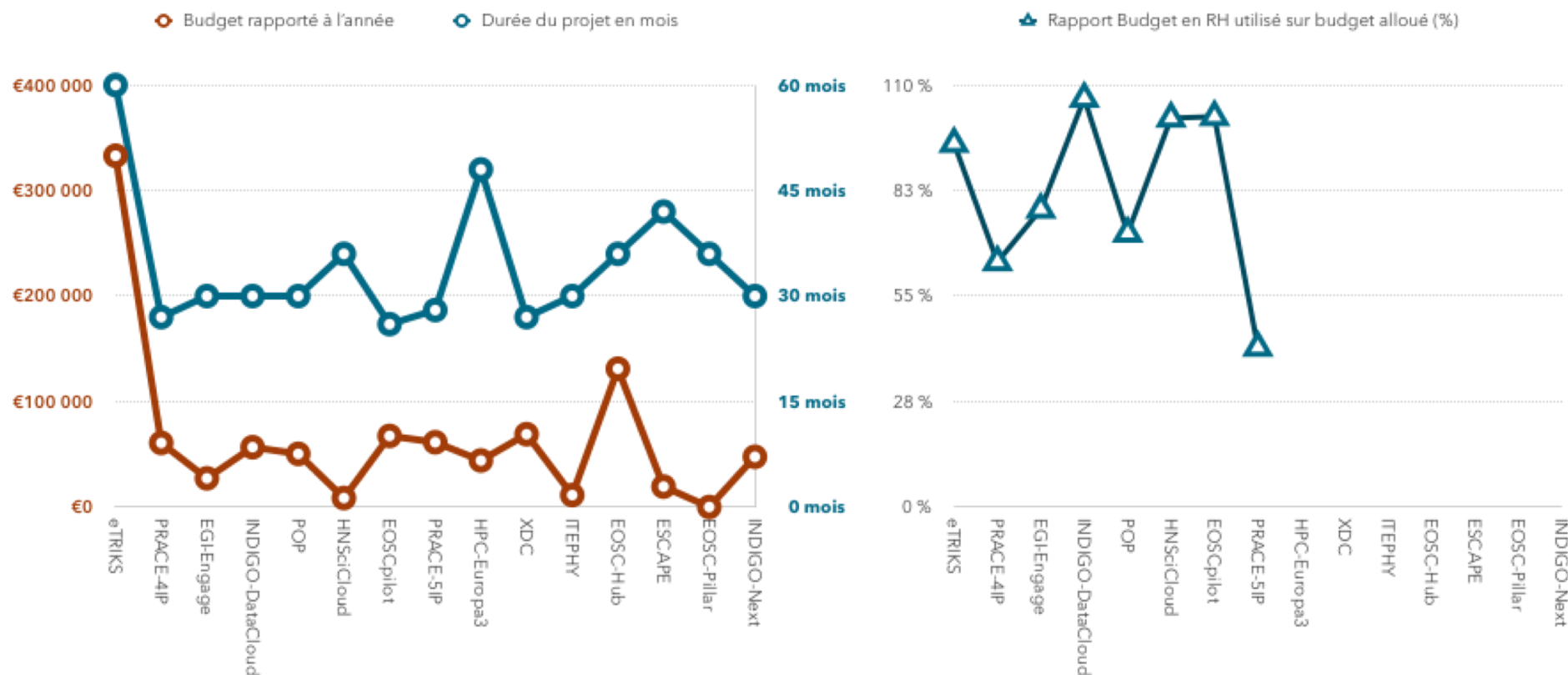
- PRACE5-IP
- HPC-Europa3
- EOSC-Hub
- XDC:
- ITEPHY
- ESCAPE

▶ Projets Européens non commencés ou en cours d'évaluation

- EOSC-Pillar
- Indigo-NEXT



Vue d'ensemble sur les projets



- ▶ Des projets de durée entre 3 et 5 ans avec une moyenne de 34 mois.
- ▶ Des projets avec RH et/ou pas de ressources matérielles
- ▶ Une utilisation moyenne des RH de 89%
 - 2 familles de projets: autour de 70% et autour de 100% → cf. suite.

► Pours:

- Liens des projets avec des activités en cours au CC-IN2P3
 - GPU, orchestrations
 - Apporte un surplus de ressources humaines ressources
- Acquisition d'expertises nouvelles
 - Containers, workflows
- Partage des connaissances
 - Avec les autres membres du CC
 - Avec les utilisateurs et le monde académique
- Participation à des activités avec un réseau d'experts
 - Appropriation des dernières technologies
 - La plupart de ces activités auraient dû être menées de toutes façons.

- ▶ (Non-)Choix du projet auquel participer
 - Pas toujours en adéquation avec les axes stratégiques du CC
- ▶ Flou dans la rédaction des documents de travail (DoW):
 - Mauvaise estimation du travail à livrer et des RH.
- ▶ Processus d'embauches des nouveaux CDDs
 - Grandes difficultés au recrutement (salaire, contrats courts, dernier moment avec un délai important pour le début de l'activité).
 - Difficulté de faire endosser le début du travail par un membre en place: manque de temps, d'intérêt.
 - Utilisation du budget RH:
 - Mauvaise estimation des budgets: manque de connaissance du niveau du CDD engagé.
 - Différentes règles: budget global RH ou nombre de PMs à prendre en compte
- ▶ Dispersion des ressources et manque de coordination globale.
 - Délais dans la mise en place des ressources matérielles pour le projet
 - Interférence avec les plateformes de production
- ▶ Over administratif importants
 - Rigidité dans les règles
 - Qui ne sont pas toujours interprétées de la même manière.
 - Plusieurs « guichets » auxquels s'adresser
 - Empilement des règles (CNRS, EU)

- ▶ **Mieux coordonner → créer une équipe**
 - Identifier les projets d'intérêt pour le CC
 - Centraliser les sollicitations et assurer la mise en place et le suivi
 - Devenir force de proposition pour des thématiques/appels intéressants
- ▶ **Avoir des personnes compétentes dans la connaissance des projets européens**
- ▶ **Mieux participer**
 - Organiser la participation des équipes du CC au projet
 - Identifier les participants et les encadrants selon expertise
 - Identifier les synergies entre activités-projets et activités-CC et les inclure en amont du travail effectué
 - Par exemple, élargir le champ défini par le projet pour répondre à des besoins CC
 - Susciter l'intérêt à participer
- ▶ **Plus et mieux collaborer**
 - Avec l'IN2P3 en amont
 - Avec les autres centres et laboratoires

- ▶ Créée en avril 2017
- ▶ Mission: voir document
- ▶ S'appuie sur les ressources de toutes les équipes CC et leur participation thématique.
- ▶ Former des groupes de travail permettant la «fertilisation croisée » projets-équipes CC.

MISSIONS

- Coordonner la participation du CC-IN2P3 aux différentes réponses à projets européens pour lesquelles il serait soit sollicité par ses partenaires soit force de proposition.
- Assurer la cohérence globale de ces participations et leur cohérence par rapport à la stratégie du centre.
- Former les groupes qui vont répondre et participer à ces projets au sein des autres équipes et selon les compétences.
- Développer au sein des équipes du CC-IN2P3 la capacité pour leurs membres à participer de manière effective aux différentes phases d'un projet (discussions préalables, rédaction, coordination, implication en cas de validation du projet, etc.)
- Valider et coordonner les réponses à projets issues de propositions internes
- Assurer les relations avec les services « Europe » de l'IN2P3 et de la délégation régionale.

- ▶ Un nouveau poste au CC depuis le 8/1/2019:
 - Chargée de l'appui aux projets de recherche: Mathilde Monges
- ▶ Des compétences à mettre en œuvre:
 - Coordination avec toutes les cellules en amont
 - Appui aux responsables des projets
 - Travail et concertation étroite avec le CE&I

- ▶ Gestion plus souple et plus intégrée des ressources humaines
 - Projets et activités menés par plus d'une personne → plus de transfert de connaissances, plus de redondance et de backup
 - Favoriser l'implication de permanents
 - Eviter l'engagement d'un CDD à 100% sur un seul projet.
 - Evite la perte de connaissance en cas de départ.
- ▶ Créer un « testbed »: Plateforme pérenne mais évolutive pour tester des nouveaux produits sans impacter ou mettre à risque la plateforme du CC.
 - Plus d'agilité dans la prise en charge du projet
 - Moins de frilosité à tester des technologies nouvelles disruptives.
 - Bac à sable aussi pour les besoins d'évolution propre de la plateforme du CC

- ▶ Avec les autorités de tutelle
 - Comment connaître plus en amont et plus tôt les appels à projet et participer éventuellement à leur rédaction?
 - Comment devenir initiateur d'une réponse à l'appel à projets?
 - Cellule de montage de la réponse
 - Appui institutionnel
 - Encadrement de la participation avec les partenaires
- ▶ Avec les cellules de projets européen
 - Uniformiser la compréhension des règles de participation
 - Comment faire des circuits de réponse plus courts?
- ▶ Pour les ressources humaines
 - Processus pour engager plus en amont les CDDs sur projets
- ▶ Comment s'organiser au niveau IN2P3?

**Merci à tous ceux qui m'ont fourni leurs
input :
C.L'Orphelin, S. Gadrat, R. Lemrani, P.E.
Macchi, M. Monges, F. Suter**

Les Projets au CC

- ▶ *eXtreme DataCloud*
- ▶ Coordinateur: INFN
 - Durée 27 mois: 1/10/2017-31/1/2020
 - Budget=3.07M€ / CNRS: ??? M€



- ▶ CC-IN2P3
 - 20 PMs
 - Budget 155250 €
- ▶ Objectif global:
 - Gestion des données à grande échelle pour les projets scientifiques futurs : Lifewatch, CTA, ECRIN, LHC

- ▶ WP3: Gestion des releases du software et prototypes (testbeds)
- ▶ Pour le CC-IN2P3: participation à T3.2 :
 - Mise en place d'un prototype (testbed) pour tester les fonctionnalités de Onedata pour le usecase CTA

- ▶ European Science Cluster of Astronomy & Particle physics ESFRI research infrastructure
 - Coordinateur: CNRS/ G.Lamanna
 - Durée 42 mois: 1/2/2019-31/7/2022
 - Budget=15M€ / CNRS: 2.8M€

- ▶ CC-IN2P3
 - 10 PMs
 - Budget 75375€
- ▶ Objectif global:
 - Etablir un cluster commun aux ESFRI dans les domaines de l'astronomie et de la physique des particules sur accélérateurs pour implémenter un programme commun sur l'EOSC.



- ▶ WP2: DIOS – Data Infrastructure for Open Science
 - Concevoir, implémenter, opérer un prototype de “datalake”: infrastructure fédérée avec un accès aux données pour les ESFRI qui font partie du projet ESCAPE.
- ▶ Pour le CC-IN2P3: participation à T2.2 et 2.3:
 - Mise en place d’un prototype (testbed) de stockage de taille modeste mais suffisant pour tester les fonctionnalités.
 - Création d’une architecture fédérée avec tous les composants de base:
 - stockage primaire, cache organisé ou non.
 - Accès, authentification gestion des données.
 - monitoring
 - Service d’orchestration du datalake:
 - Gestion différenciée des répliqués sur plusieurs centres, des QoS, de l’accessibilité.
 - QoS: Optimisation redondance/performance/coûts.
 - Cycle de vie des données: gestion de l’accès, de la protection avec droits d’accès changeant dans le temps et pour le sous groupe d’utilisateurs.
 - RUCIO: outil de DM et orchestration
 - Evaluation et tests avec des données des ESFRI en exerçant leur modèle de données spécifique (exemple LSST(Sic!)).
 - Possibilité d’évaluer d’autres composants venant d’autres projets: ex XDC et OneData.

▶ EOSC-Pillar

- Coordinateur: consortium GARR/Italie
- Durée 36 mois
- Budget=6.9M€ / CNRS: 714k€
- Début du projet non encore fixé



INFRAEOSC-05-2018-2019

▶ CC-IN2P3

- 3 PMs
- Budget ?

▶ Objectif global:

- Aider à la coordination et à l'harmonisation d'initiatives nationales dans le but d'établir un EOSC global qui les intègre.
- Aider à l'adoption et à la mise aux normes des standards de l'EOSC.
- La caractéristique est le focus sur la « longue tail of science » c'est-à-dire à des communautés scientifiques plus petites et plus diverses.

- ▶ Work Packages: WP1 à WP7 (6&7 coordonnés par le CNRS)
- ▶ WP7: – Infrastructure Layer delivering horizontal data storage.
 - Participants français: CINES, INRA, INRIA, IFREMER, INSERM, CNRS(INSU, INS2I, IN2P3)

- ▶ INDIGO-Next Integrating Distributed data Infrastructure for Global ExpLOitation – NEXT:
 - Coordinateur: INFN/Italie
 - Durée 42 mois: 1/2/2019-31/7/2022
 - Budget=6M€
 - Soumis mais non encore évalué.
- ▶ CC-IN2P3
 - 21 PMs
 - Budget 132250€
- ▶ Objectif global:
 - Suite de Indigo data Cloud

INDIGO-Next

- ▶ Work Packages: WP1 à WP6
- ▶ WP2: co-design de services innovants permettant un accès au travers du Hub EOSC aux services répondant aux besoins des communautés scientifiques.
- ▶ WP3: processus DevOps pour gérer la cycle de vie du software: QoS, releases, maintenances et services pilotes.
- ▶ Dans notre cas:
 - Gestion automatique de données distribuées géographiquement: optimisation de l'accès aux données.
 - Exploration des processus déclenchés par l'arrivée des données.
 - Pipelines et workflows complexe de traitement des données (en continuation de HPC-Europa3).

▶ PRACE-5IP

- Coordinateur : Forschungszentrum Jülich
- Durée : 28 mois
- Budget = 16.1M€ / CNRS: 412 k€

▶ CC-IN2P3

- 13 PMs
- Budget : 159k€

▶ Objectif global:

- Préparer les évolutions de l'infrastructure HPC pan-européenne PRACE
- Opération, support, PCP, ...
- Faire le lien avec de grands instruments scientifiques
- Etudier de nouveaux services à déployer en production
 - Deep Learning, containers

- ▶ WP6: Operational Services for the HPC Eco-system
 - Tâche 6.2 : Design and Development of new Services prototypes
 - Service 2 : Link with Large Scale Scientific Instruments (leader)
 - Service : Deep Learning
- ▶ Participants français: CC-IN2P3, CEA, CINES, GENCI, IDRIS, INRIA

- ▶ EOSC-hub: Plateforme ouverte pour l'échange de données de recherche.
 - Coordinateur
 - Durée: 36 mois du 1/1/18 au 31/12/20
- ▶ CC-IN2P3
 - 71.5PM
 - Budget: 393250€



- ▶ T5.2 Marketplace and Order Management Tools - 9 PM
 - CCIN2P3 is in charge of the backend of the Order Management system . The aim is to develop tools to automate the request of ressources done through the marketplace.

- ▶ T5.3 Integrated Business and Operations Support Systems - 10 PM
 - CCIN2P3 is in charge of the continuity and the evolution of this portal developed since 2014

- ▶ T5.4 Monitoring, Accounting, Messaging, Security Tools - 10 PM
 - CCIN2P3 is in charge of the development of the Web User Interface for the ARGO monitoring solution

- ▶ T13.2.9/T13.2.10 Access Provisioning - 41.5 PM
 - CCIN2P3 ensure the provision of the Operations Portal service and of the test instance of ARGO