

Plateforme d'Hébergement et Traitement de Données "VirtualData"

Rencontre avec le ministère ESRI
5/2/2019

Sommaire

- Les motivations de l'initiative P2IO Virtual Data
- P2IO et ses partenaires
- Datacenter Phase 1 réalisé en 2012
 - Salle Vallée
 - Salle Plateau
- L'extension en cours de la Salle Vallée (CPER)
- Cloud@UPSud
- Stockage distribué (Ceph)
- Gouvernance et business model
- Conclusions

LABEX P2IO

2012

- P2IO = Physique des 2 Infinis et des Origines
 - Physique nucléaire + physique des hautes énergies + astrophysique
 - Y compris leurs applications, par exemple santé
 - 8 laboratoires sur Orsay (CSNSM, IAS, IMNC, IPNO, LAL, LPT) + Saclay (CEA/Irfu) + Palaiseau (Polytechnique/LLR)
 - ~2000 personnes dont 130 informaticiens
- 1 longue tradition de collaboration entre ces laboratoires
 - Les expériences sont souvent des collaborations nationales ou internationales impliquant plusieurs des laboratoires
- 1 volonté de s'affirmer en tant que pôle de physique dans l'université Paris Sud et la future université Paris Saclay
 - Le LABEX vise à structurer la collaboration et les efforts autour des plateformes technologiques

VirtualData : Groupe de travail informatique P2IO

VirtualData : Ambitions

2012

- Produire et exploiter de grandes masses de données
 - LHC, Astrophysique : >10 Peta-Octets/an
- Constituer une infrastructure informatique capable de relever les défis du futur
 - Expériences internationales : forte coordination globale, intégration avec les centres de données majeurs (ex: CCIN2P3)
 - S'appuyer sur l'existant : GRIF (site grille, mésocentre), Cloud, IDOC, ...
 - GRIF : 10K cores, 8 PB, plusieurs dizaines de VOs (communautés)
 - Responsabilité dans le contrôle et le traitement informatique de grands projets internationaux
 - Créer une structure de collaboration entre les "informaticiens" du LABEX P2IO
- **Augmenter la disponibilité des services tout en optimisant le coût et l'impact environnemental (donc la consommation électrique)**
 - Mettre en place des infrastructures mutualisées résilientes et efficaces
 - Permettre une allocation souple des ressources unifiées (ex: cloud)

Positionner P2IO dans les appels à projet

Plateforme Mutualisée : Une solution optimale

2012

- Projet : création de 2 salles informatiques P2IO complémentaires
 - Vallée (Orsay) + Plateau (Ecole Polytechnique)
 - Haute disponibilité : redondance des services critiques

Optimiser les ressources

- **Urgence** dans la Vallée d'Orsay pour les acteurs P2IO en 2011
 - LAL devait engager 150 K€ en 2013 pour réparer sa climatisation
 - IPNO : problèmes récurrents de climatisation
 - IAS : besoin d'extension, au minimum 50 k€ en 2013 et en 2014

Valoriser ces investissements dans une plateforme mutualisée

Datacenter

Salle Vallée : Cible Finale

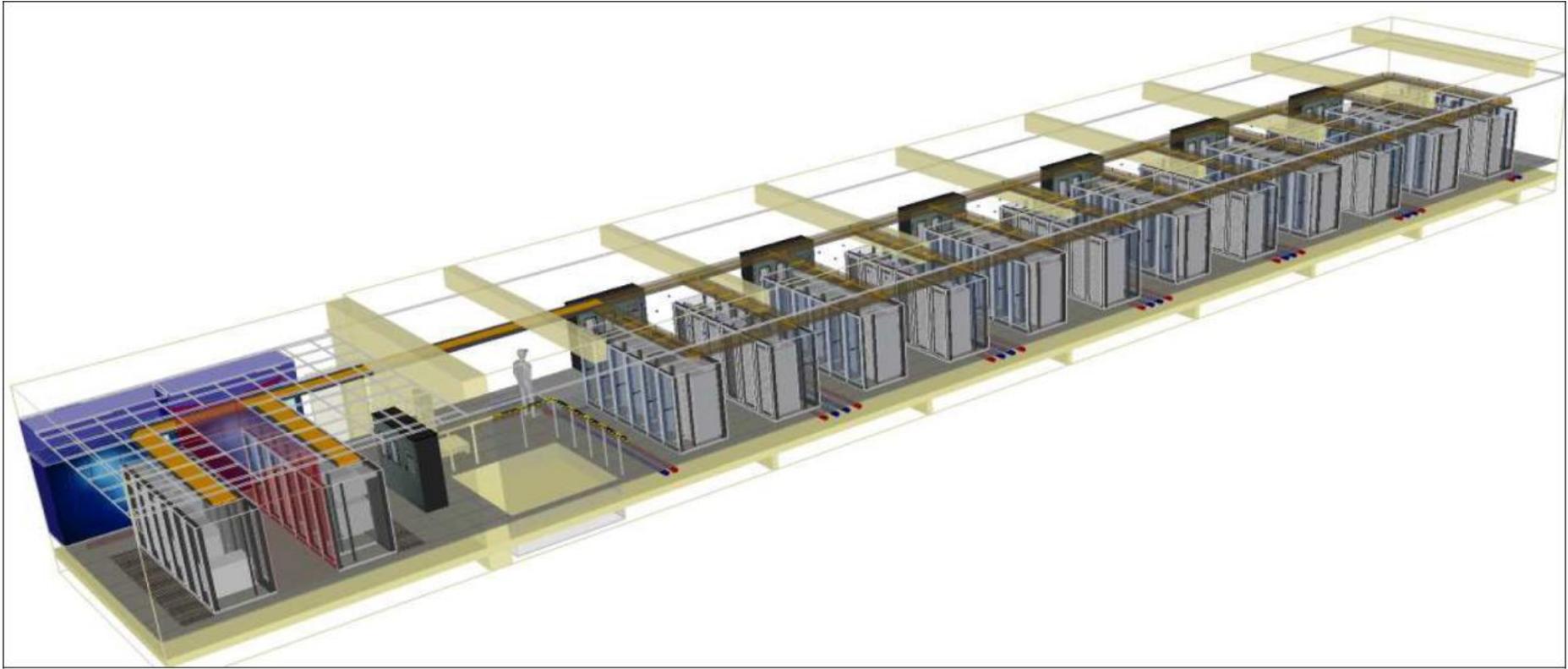
2012

- Répondre aux urgences avec une solution évolutive
 - Pouvoir répondre aux besoins des dix prochaines années
 - Dimensionner l'existant aux besoins réels : une extension modulaire sans impact sur le coût consolidé
- Design centré sur l'efficacité énergétique
 - Principal facteur du coût d'opérations des ressources
 - Pas de surinvestissement dans la redondance : approche globale basée sur 2 salles avec possibilité de redondance
 - Objectif fixé à PUE = 1,3 (depuis la ligne HT !)
- Vallée d'Orsay : réutiliser un bâtiment existant
 - Bâtiment 206 : bâtiment technique sur 2 étages
 - Possibilité de 6 pôles électricité + clim : modularité des extensions

Des investissements mis en perspective

Salle Vallée : Cible Finale

2012



Vue d'Artiste de la Salle 206 (cible finale, juin 2012)

Salle Vallée : au-delà de P2IO

2012

- Laboratoires P2IO Vallée : des acteurs majeurs de UPSud
- Histoire riche de collaboration avec d'autres composantes
 - LRI : visualisation, interprétation de données massives
 - OSU Paris Sud : GEOPS
 - Biologie/bio-informatique : Plateforme eBio UPSud (IGM/I2BC), USR BioEmergences (Gif)
 - Chimie : LPC
- Rencontre régulière avec la DI Paris Sud dès décembre 2011
 - 1 réunion par trimestre autour du projet P2IO pour **l'inscrire dans la cohérence de l'offre calcul scientifique Paris Sud**
 - 1 soutien fort : **s'appuyer sur la dynamique P2IO** pour initier une mutualisation plus globale à UPSud
 - Plusieurs financements ERM soutenus **par toutes les composantes**

Inscrire la réponse à nos besoins dans le contexte global

Design Salle Vallée

2012

- Potentiel de 1,5 MW IT dans 220 m² avec un **PUE ≤ 1,3**
 - 84 racks “denses” tant pour l’électricité que le poids
 - Les racks peuvent potentiellement recevoir tout type de matériel
 - Serveurs : 0,8T/30 kW, stockage : 1,2T/15 kW
- Efficacité énergétique basée sur des racks refroidis à eau
 - Seule solution pour la haute densité thermique
 - Solution la plus dense : pas d’unité de climatisation
 - Modulaire : racks achetés au fur et à mesure des besoins
- Design modulaire basé sur 6 “pôles techniques” fournissant 300 kW IT chacun (électricité et climatisation)
 - Redondance N+1, pas d’onduleur

Design centré sur la haute densité matérielle

Salle Vallée Phase 1

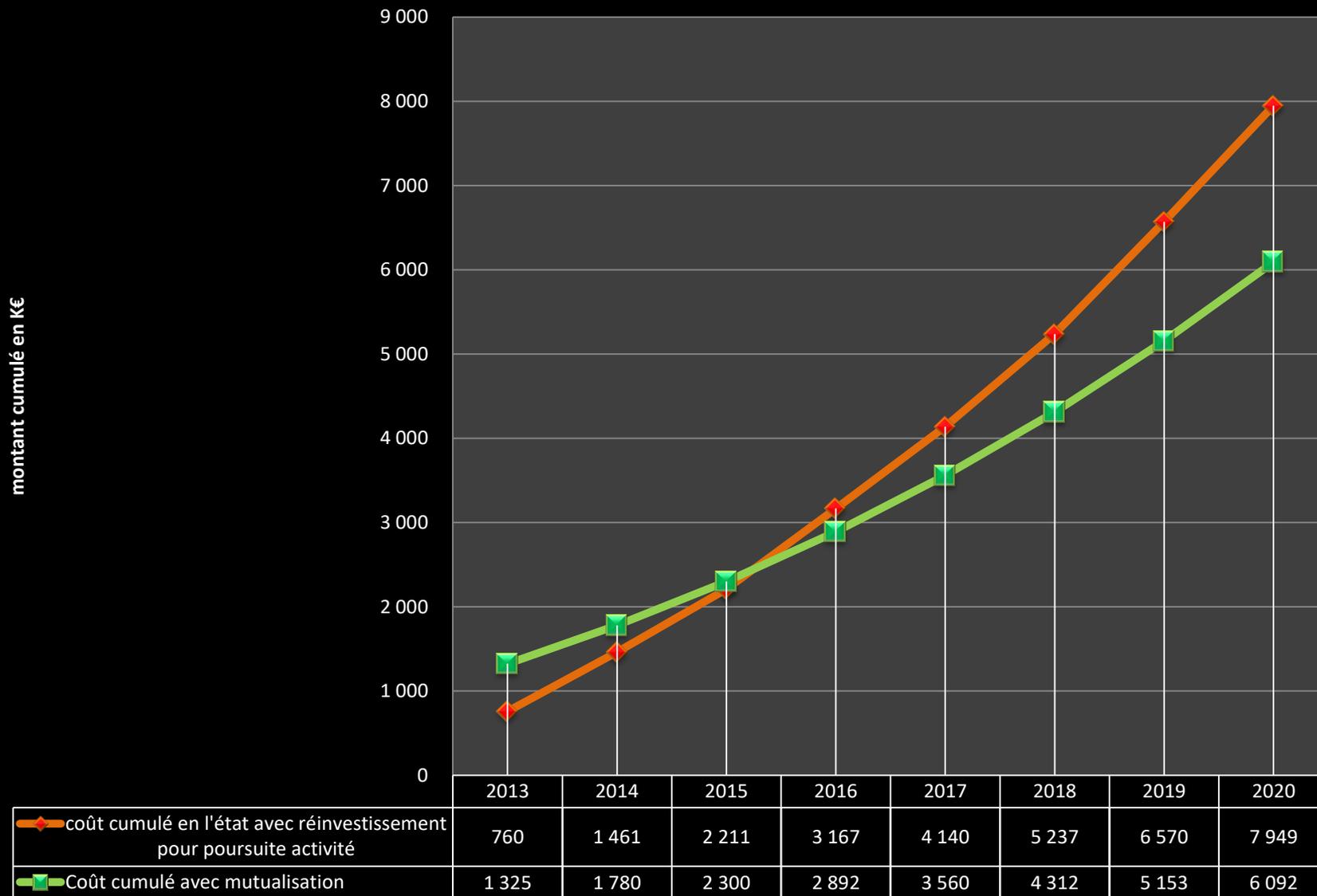
2012

- 1/3 de la cible finale **disponible 1/10/2013**
 - Transformation d'un bâtiment technique existant
 - 100 m² (30 racks), 400 kW IT
 - Au moins 2 groupes de production eau froide (redondance)
 - 1 voie électrique (pas de sécurisation)
- 1 M€ : laboratoires (80%) + soutien du LABEX P2IO (20%)
 - Solution technique permettant d'atteindre les exigences de disponibilité et d'efficacité
 - Design modulaire permettant une extension en fonction des besoins
 - Etudes démarrées début 2012

Coûts engagés très proches de l'estimation initiale
Un design modulaire et extensible

Comparaison des coût cumulés avec ou sans mutualisation (k€) en incluant les réinvestissements obligatoires

2012



Salle Plateau

2012

- Approche différente pour tenir compte du contexte évolutif UPSay
 - Un investissement uniquement P2IO n'aurait eu aucun sens
 - Pas de bâtiment pré-existant utilisable pour une stratégie de moyen-long terme orientée sur l'optimisation du coût de fonctionnement
- Réponse aux besoins de court/moyen terme : collaboration avec l'Ecole Polytechnique en 2012 pour la rénovation de la salle Aile 0
 - P2IO (LLR) a joué un rôle majeur, profitant de l'expérience Salle Vallée
 - P2IO a déployé ~12 racks « haute densité » pour ses propres besoins
 - LLR + redondance avec la salle Vallée
- 2013-2014 : P2IO a été acteur de la discussion avec les autres partenaires de l'IDEX pour une solution de long terme
 - Absence de financement a malheureusement stoppé cette dynamique

Volonté d'inscrire les besoins P2IO dans le contexte UPSay

Extension Salle Vallée...

En cours

- Salle actuellement presque pleine : 1 à 2 racks (sur 30) encore vides
 - 26 racks P2IO, 4 racks UPSud : 250 kW IT
 - **Une demande croissante des laboratoires UPSud** et peut être UPSaclay
 - Renouvellement de matériel ancien a permis de compenser le retard de l'extension (initialement prévue pour été 2018)...
- Objectif : pouvoir héberger 600 kW IT dont 300 kW redondés
 - 3 pôles techniques complets (électricité + climatisation)
 - Suppression des points uniques de défaillance : 1 des pôles sur le 2eme circuit HT de UPSud, 2^{ème} échangeur thermique
- Intègre aussi la fin de l'aménagement 2nd œuvre de la salle
 - Faux plancher, distribution hydraulique...
 - Accès PMR (ascenseur) + diverses utilités
- Extension future : 3 x 300 kW IT (~300 k€ pour 300 kW)
 - Eviter de déployer trop tôt de la capacité électrique/cooling

... Extension Salle Vallée

En cours

- Statut : mise en production à l'automne 2019
 - Même maître d'œuvre que la phase 1 : Critical Building
 - DCE travaux publié début octobre 2018
 - Actuellement : phase finale de la négociation mais offres initiales compatibles avec le budget prévisionnel
 - Début des travaux : mars 2018 (6 mois)
- Financé dans le cadre du CPER UPSud P2IO Vallée (20 M€, Etat + Région IdF)
 - Budget de l'extension : ~2 M€ (MOE & co compris), incluant un peu de jouvence des infrastructures (isolation bâtiment, circuits HT...)
- Extension doit permettre un **accroissement des utilisateurs non P2IO**
 - Etude en cours pour déménager la salle historique de la DSI UPSud
 - Capacité supplémentaire significative (21 racks)
 - Possibilité d'ajouter encore 30 racks

Services Mutualisés

Cloud@UPSud

- Le LAL engagé dans la R&D Cloud depuis 2010
 - Une technologie au service d'une infrastructure de calcul distribué
 - En parallèle avec la ressource de production grille (GRIF)
 - Projet StratusLab (FP7, 2010-2012), CYCLONE (H2020, 2015-2017)
- 2016 : migration du cloud StratusLab -> OpenStack
 - Soutenu par un ERM UPSud pour construire un cloud de production permettant des tests à l'échelle (**toutes les composantes UPSud**)
 - Actuellement : 3500 cœurs, 500 TB de stockage Ceph
- Ouvert à des utilisateurs non UPSud : ISC-PIF, IGN
 - Développer la mutualisation des ressources commencée avec la grille
 - Participation à la fédération de cloud France Grilles

GRIF@Orsay

- GRIF : site de la grille EGI/EOSC réparti sur 4 sites en région IdF
 - Orsay : 5500 cœurs, 2 PB
 - Autres sites : CEA/Irfu (Saclay), CNRS/LLR (Polytechnique), CNRS/LPNHE (UPMC)
 - Ouvert à beaucoup de communautés mais actuellement surtout expérience HEP, astroparticules (HESS, CTA) et ISC-PIF
 - Pour le LHC (WLCG), site Tier2 : pledges de ressources annuels, co-financement LCG-France
- A terme, devrait devenir un service du cloud
 - Forte augmentation (2,5x) du nombre de cœurs du cloud
 - Possibilité d'un partage de ressources plus flexible entre grille et autres services du cloud

Stockage Distribué

- Projet porté par P2IO, susceptible d'une ouverture extérieure à d'autres
 - Financement attribué par le LABEX pour déployer une infrastructure de 1 PB utile sur 3 sites
- Objectif : construire une infrastructure de stockage résiliente et performante à un coût optimisé
 - Basé sur la technologie Ceph : cloud storage (Swift/S3, block device, file system distribué)
 - Basé sur du commodity hardware
 - Résilience à travers de la réplication ou de l'erasure coding + une configuration multisite (Vallée, Plateau, CEA/Irfu)
- Des use cases réels dès la première phase : cloud, LSST, physique nucléaire, IDOC...

JupyterCloud

- JupyterHub (notebook) implémenté dans le cloud
 - Cible : enseignement, développement collaboratif pour l'analyse de données
- Service développé dans le cadre de UPSaclay : UPSud + Polytechnique impliqués dans la définition et la gestion du service
 - Allocation de ressources d'exécution dans le cloud en fonction des besoins
 - Nombreux environnements (kernels) supportés
- Intégré à l'effort national de coordination des services JupyterHub
 - Potentiellement une vocation nationale si des besoins sont exprimés et compatibles avec les besoins locaux
 - En particulier, authentification possible via la fédération RENATER

Outils d'Analyse : Spark

- Spark : technologie de traitement parallélisé de grande masse de données issue du big data (open source)
 - Extension de la technologie Map-Reduce
- Motivé initialement par plusieurs besoins dans UPSud, dont un projet de génomique
- Actuellement, le moteur est devenu une R&D pour l'expérience LSST menée par le LAL depuis 2016
 - Intérêt croissant dans le cadre de LSST
 - Travail spécifique sur les outils d'analyse de « données 3D » : projet [Astrolab Software](#)
- Service Spark multi-tenant avec allocation de ressources à la demande
 - Implémenté dans le cloud, développement en cours

Avant le lancement

0) Conception et tests d'instruments spatiaux

Agences spatiales

1) Instrument Operations

Segments sols

2) Pipelines

Autres fournisseurs de données

Gestion des données

3) Jeux de données, interfaces, outils de traitement, observatoires virtuels

4) Archivage moyen et long terme

Communautés scientifiques

Grand Public

Stratégie : Mutualisation, Standardisation, Qualité
Pilotage : Responsable scientifique et Responsable technique
Comité de pilotage, Coordinateurs scientifiques

Produire

Création des niveaux de données et mise en forme

Production de données à haute valeur ajoutée

Préserver

Disponibilité des données

Confidentialité, sécurité, archivage

Valoriser

Portails

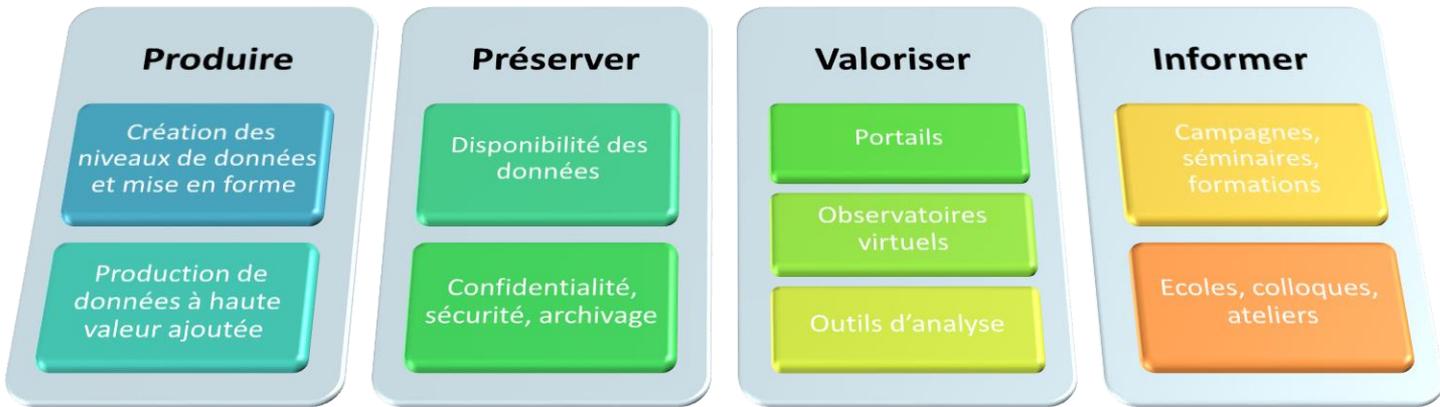
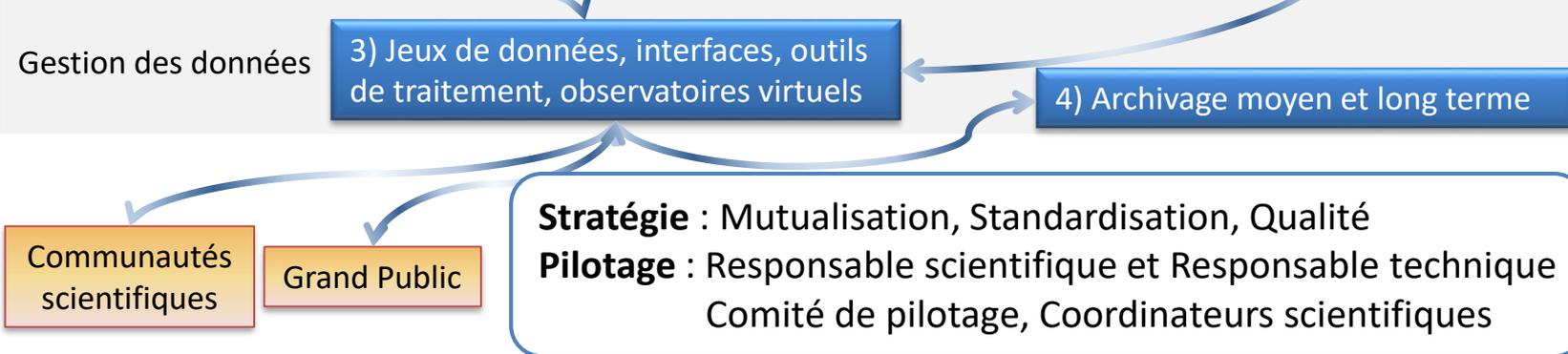
Observatoires virtuels

Outils d'analyse

Informer

Campagnes, séminaires, formations

Ecoles, colloques, ateliers



Gouvernance et Business Model

Mode de Fonctionnement

- Une approche originale basée sur une collaboration volontaire entre des acteurs ancrés dans les laboratoires
 - Volonté explicite de ne pas se transformer en Unité de Service
 - Raison d'être est le lien avec les utilisateurs finaux
 - Chaque ASR/labo gère ses machines : accès à la salle (badge)
- Des groupes de travail techniques spécialisés ouverts à toutes les personnes intéressées
 - Réseau, monitoring, infrastructure
 - Actuellement surtout P2IO + DSI UPSud
- Un tour de garde hebdomadaire assuré à tour de rôle par les différents laboratoires
 - Forte instrumentation des infrastructures
 - Alertes (automate) distribuées à tous les occupants par mail
 - Suivi des éventuels problèmes par la personne de garde (office hours)

Gouvernance Actuelle

- Comité de Pilotage constitué des directeurs et responsables informatiques P2IO
 - 1 réunion/an : décisions budgétaires, validation de l'accueil des extérieurs..
 - Reflète le financement de la construction par P2IO
 - Participation de UPSud (VP Recherche) + coordination en amont
 - Labellisé comme une plateforme IN2P3 et INSU
- Coordination de l'Informatique Scientifique UPSud
 - Créée pour gérer la dynamique UPSud autour de Virtual Data
 - Hébergement et cloud
 - Participation de tous les laboratoires utilisateurs
 - Fait fonction aussi de Comité Utilisateurs dans la phase actuelle
 - Rapport régulier au Conseil de la Recherche et Conseil Scientifique
 - Plateforme formalisée sous la forme d'un mésocentre, DataCenter@UPSud

Gouvernance : Evolution

- Volonté forte d'inscrire la salle vallée et les services mutualisés dans un contexte plus large (UPSud, UPSaclay)
 - Potentiel assez unique pour l'informatique scientifique des laboratoires dans un contexte local très dense
- 5 des 6 laboratoires P2IO vallée sont en train de fusionner (1/1/2020)
 - Besoin de redéfinir le COPIL actuel basé sur 6 laboratoires
- La gouvernance doit s'adapter et permettre la représentation de tous les acteurs financeurs
 - Aucune volonté de P2IO (ou de l'IN2P3) d'avoir le contrôle total de l'infrastructure
 - Besoins P2IO très standards : pas de contraintes particulières sur les autres usages
 - Historique de la collaboration UPSud/P2IO a montré que c'était possible : P2IO par ses besoins et moyens coordonnés est une force d'entraînement dans le contexte global

Business Model

- Hébergement : une tarification basée sur un mixte entre espace occupé et puissance hébergée
 - Favoriser une occupation dense de la salle
 - Intègre une contribution à l'amortissement de la construction
 - S'appuie sur la consommation mesurée (moyenne annuelle)
 - Electricité payée par UPSud (comme pour les autres bâtiments des laboratoires) sauf pour les extérieurs
- Cloud/grille : pas de tarification pay-as-you-go
 - Les utilisateurs qui veulent avoir un niveau de ressources garanties doivent acheter (financer) cette quantité de ressources, y compris les coûts d'hébergement
 - Les autres ont un accès en best-effort
 - Un support horizontal par la mise en réseau des expertises

Réaliser la phase initiale de la Plateforme Mutualisée :

Une évidence

- ✓ Motivée par les besoins immédiats de P2IO
- ✓ Rentable pour tous les acteurs
- ✓ Ouvert à l'accueil de nouveaux demandeurs
- ✓ Extensible par la suite à coût équivalent

Potentiel de la cible finale :

Un atout dans le contexte UPSay

- ✓ Une infrastructure d'hébergement moderne et extensible conçue pour les besoins du calcul scientifique (densité)
- ✓ Un projet porté par les besoins d'une communauté inscrit dès l'origine dans son environnement scientifique global
- ✓ Un modèle extensible à UPSay