



Vers des infrastructures de stockages partagées et mutualisées

AgroDataRing

Christophe Caron *et al.*

Analyse des besoins 2016

Etude: « e-Infra Storage »



- **Un partenariat** Ingenium et un échantillon de laboratoires/plateformes
 - **Pour et avec** les producteurs et utilisateurs !
- **Autour** des besoins convergents **stockage moyen/long terme**
- ... et avec une volumétrie **maîtrisable** à l'échelle d'un Institut





Enjeux stockage moyen/long terme

Diversité



Consolidation !

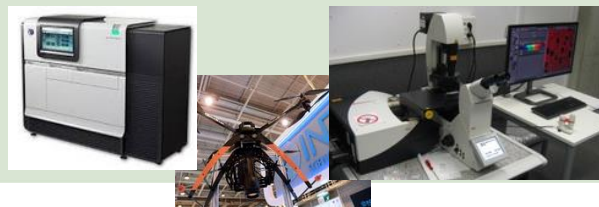
Financement
par projet



Pérennité

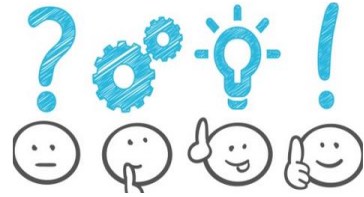
Ruptures
technologiques

•?



Evolutivité

Propositions techno et organisationnelles



- Vers de nouveau **mode d'organisation** grâce au numérique ?
- Nouvelles solutions technologiques **novatrices et pérennes** ?
 - Distribution des données à l'heure de la centralisation ?
 - Systèmes de fichiers parallèles
- **Modèle contributif : partage et mutualisation ?**

AgroDataRing v1 2017

Infrastructure partagée et mutualisée

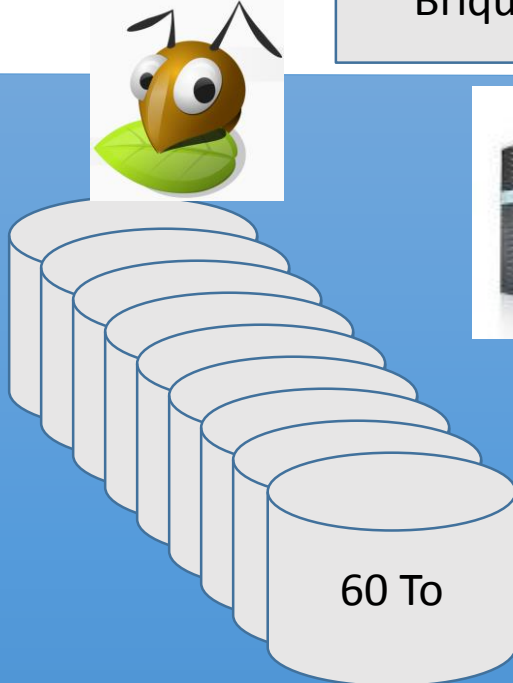
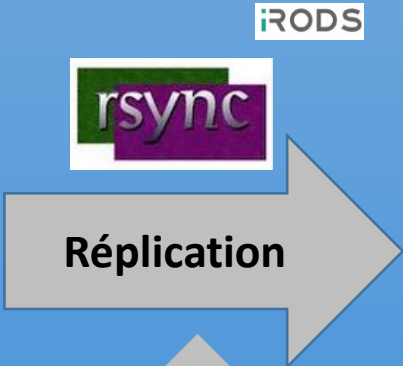
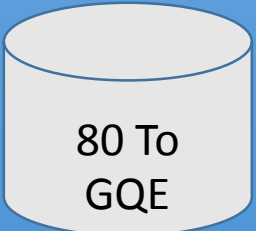
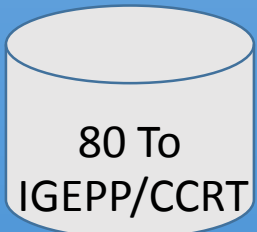
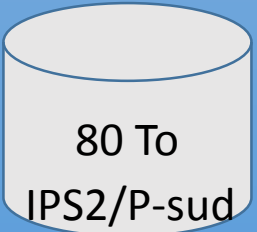
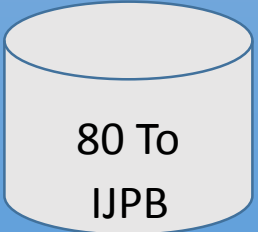
- Par et pour des communautés
 - Des besoins similaires pour des communautés : **biologistes, etc.**
 - Une infrastructure unique et co-administrée par des communautés : **informaticiens**
- **Architecture distribuée** : agrégat de briques de 80 To utiles (1..n briques par contributeur)
- **Mutualisation** ressources humaines: partage des expertises et compétences
- Ouverte : prise en compte les partenariats de recherche et **synergies avec les structures régionales/nationales**
 - Financement régionaux, MesoCentre, etc.

Architecture modulaire/évolutive

Briques de niveau 1



Brique de niveau 2



Échanges : big file sender



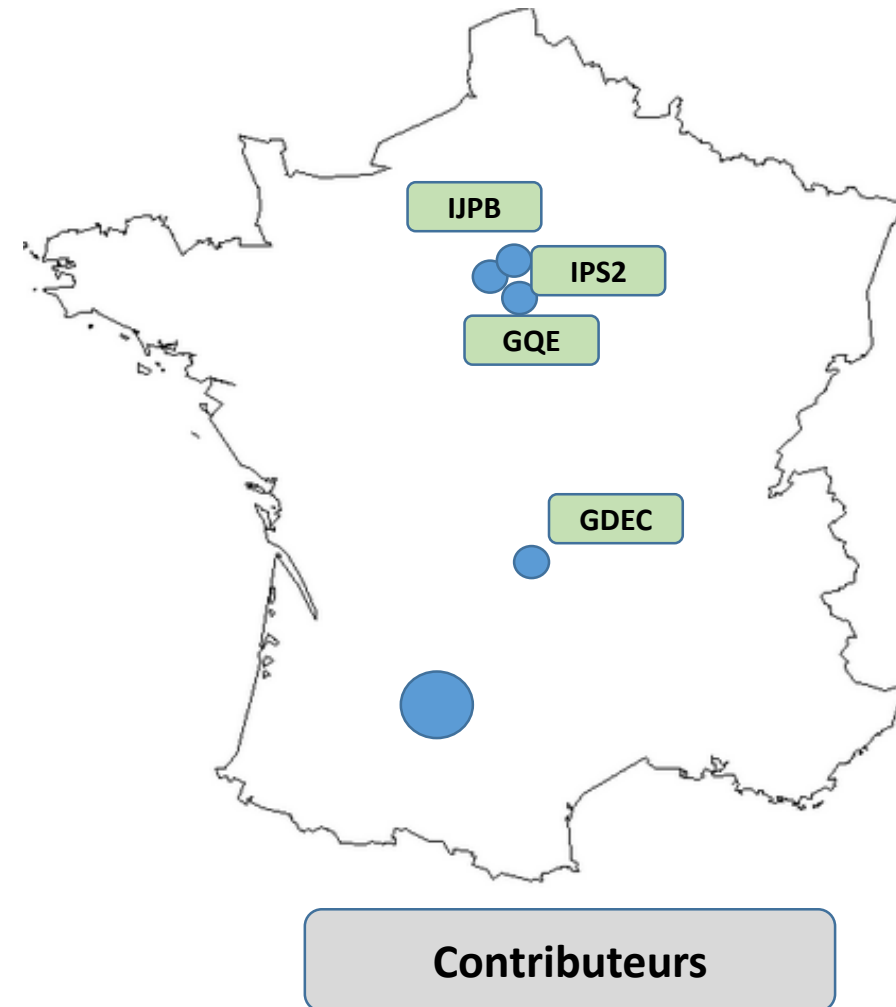
Réplication + rapprocher les données des plateformes bioinfo Genotoul et URGI

Hébergement « local »

DataCenter INRA Toulouse

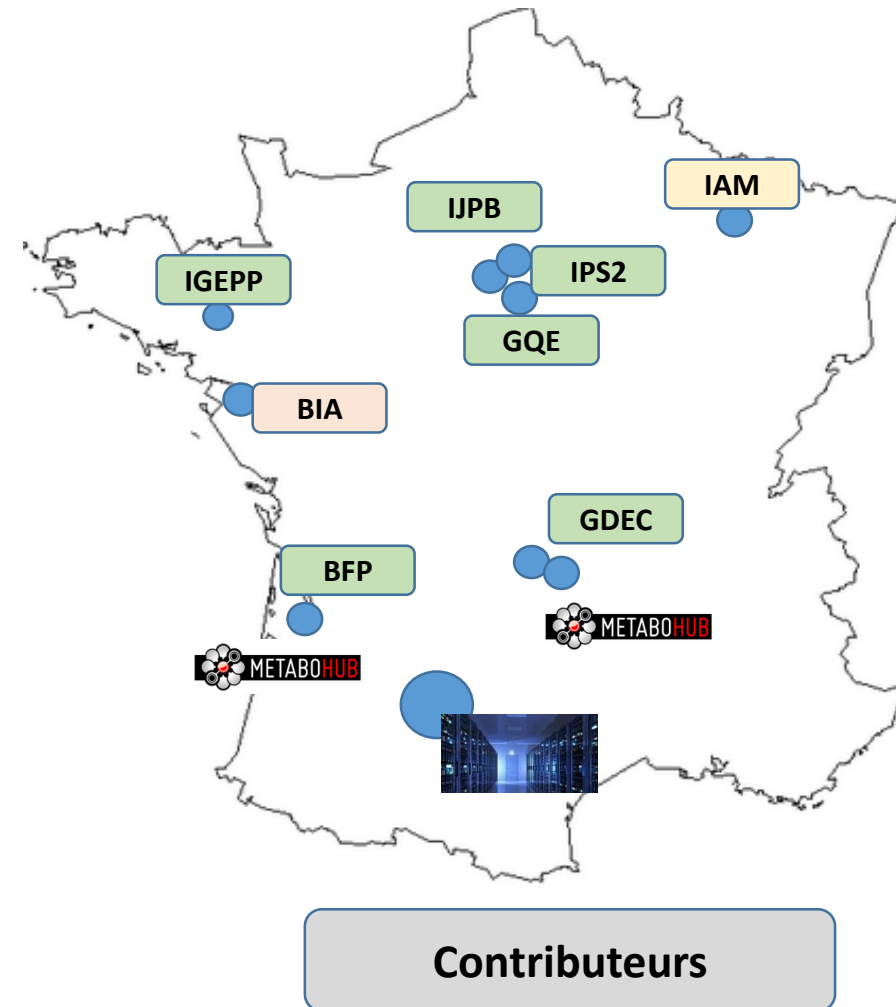
Instanciación : 2017

- Février - Mars
- Premiers financements
- Département BAP et Unité Ingenum



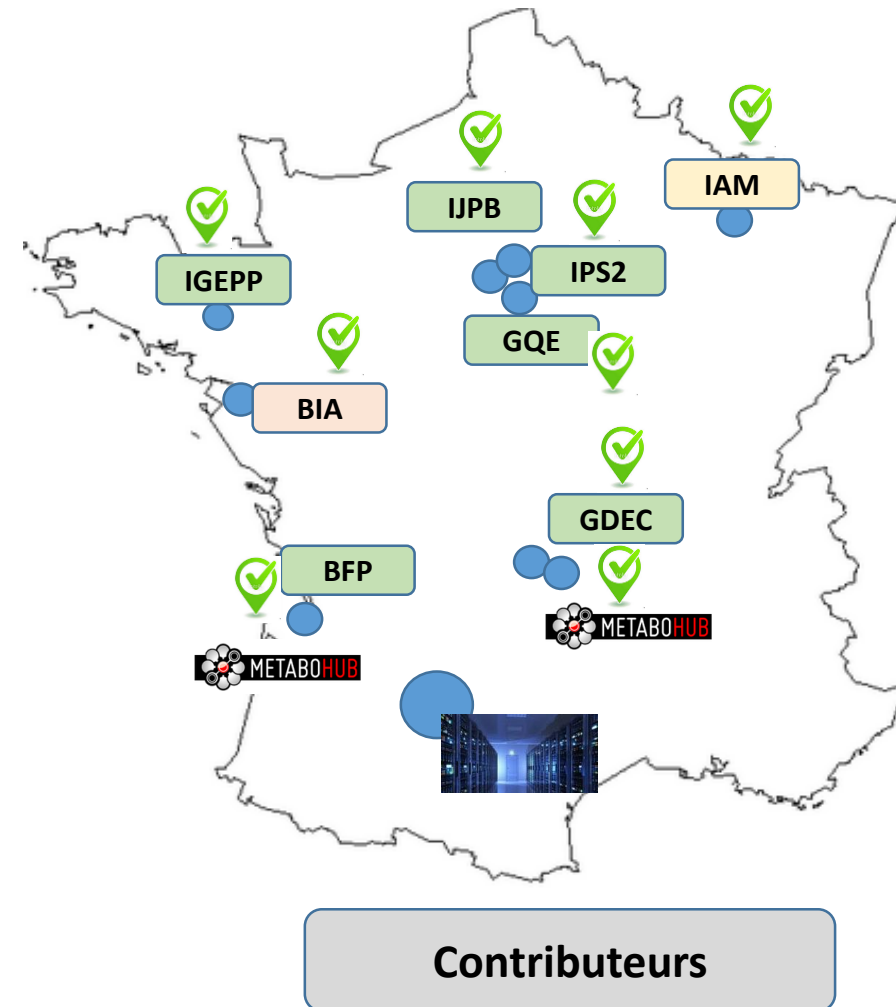
Instanciación : 2017

- Avril - Juillet
- Nouveaux contributeurs
 - BAP : IGEPP
 - Plateforme BIBS / BIA Nantes
 - Unité IAM / Nancy
 - Infrastructure Nationale MetaboHUB (2 sites)



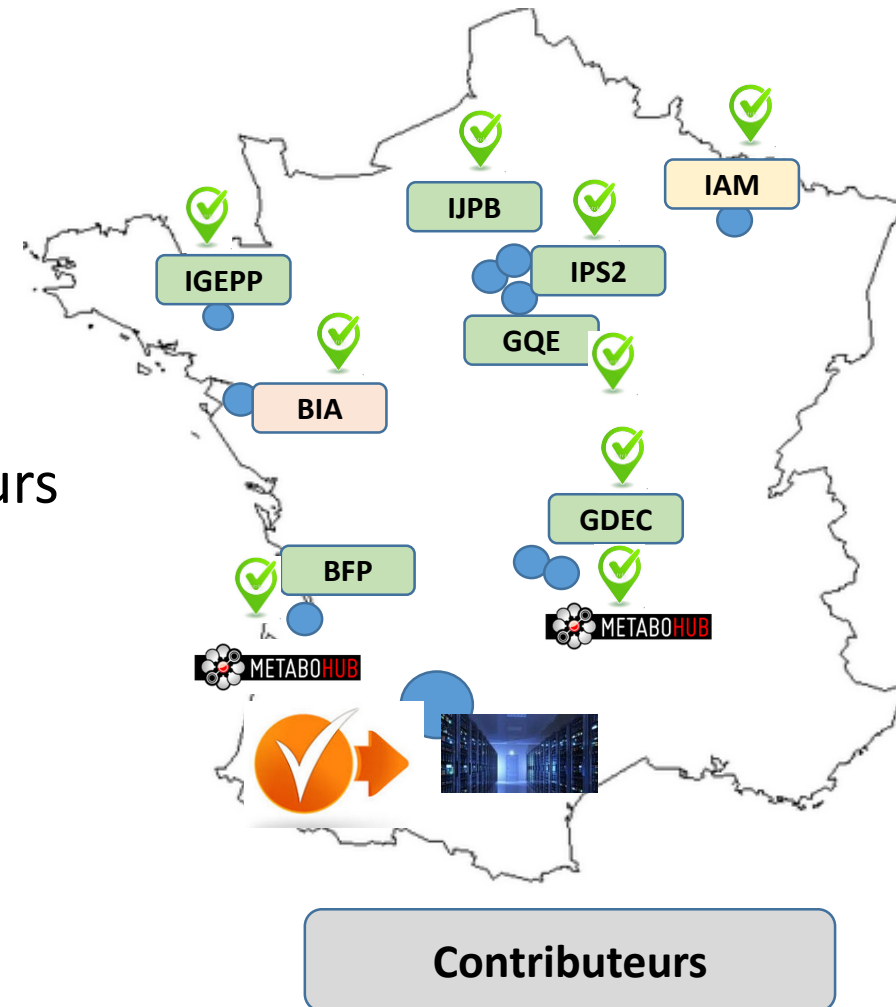
Instanciación : 2017

- Décembre
- Sites de niveau 1 opérationnels
 - 80 To utiles





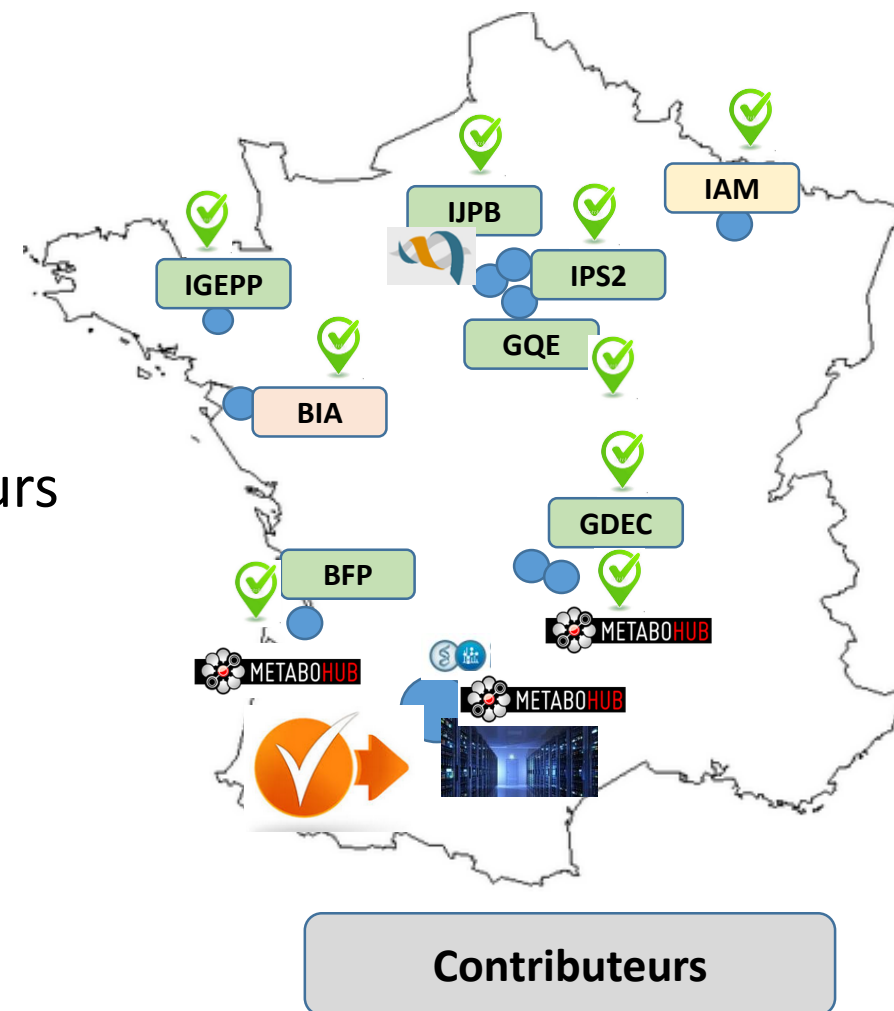
Instanciación : 2018

- Site de niveau 2 en cours de déploiement
 - Meta-brique : agrégat sous GlusterFS de n serveurs de stockage
- **Réplication !**



Instanciación : 2018

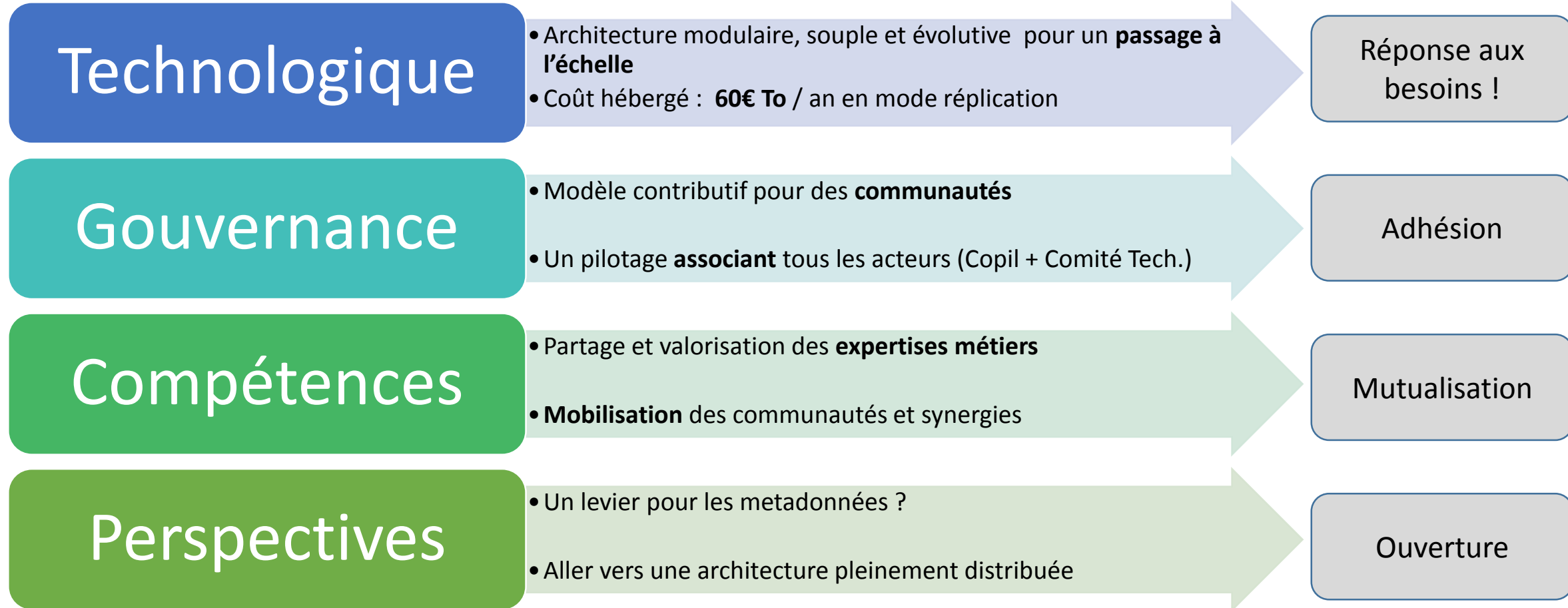
- Sites de niveau 2 en cours de déploiement
 - Meta-brique : agrégat sous GlusterFS de n serveurs de stockage
- Nouveaux contributeurs
 - Jouy-en-Josas: 
 -  / LISBP : 3^e site MetaboHUB



Gouvernance

- **Stratégique** : les directions des laboratoires concernés décident des moyens affectés : unité de stockage et/ou RH pour l'année n+1. Leur rôle est l'identification des besoins, l'orientation des choix d'organisation et d'évolution. Ce comité est en lien avec le coordinateur Ingenum de ce projet fédérateur.
- **Comité de pilotage** assure les choix techniques et a un rôle décisionnaire. Il proposera les règles au niveau stratégique : accès, adhésion, gestion des RH, usage des espaces, etc. Ce comité de pilotage se réunit tous les mois durant la mise en place de l'infrastructure, puis une fois par trimestre pour la suite du projet.
- Le **comité technique** assure le côté opérationnel. Il est représenté par un collectif d'ingénieurs système qui administre et fait évoluer l'infrastructure en synergie avec les autres axes.

Bilan





- Cela ne marchera jamais !
- Modèle réparti/distribué :
 - Techniques
 - mais surtout organisationnelles
- Mutualisation = Restriction !?
 - Ne pas y voir que les aspects « économiques »



- Dynamique de groupe (~20) et synergies
- Montée en compétence
- Mutualisation = travailler ensemble et partager

Merci

Auteurs : [Pierre Adenot¹](#), [Stéphane Bansard²](#), [David Benaben³](#), [Véronique Brunaud⁴](#), [Christophe Caron⁵](#), [Alexandre Dehne-Garcia⁶](#), [Christophe Dupérier⁷](#), [Adrien Falce⁸](#), [Olivier Filangi⁹](#), [Franck Giacomoni⁷](#), [Fabienne Granier¹⁰](#), [P. Grevet⁴](#), [Nicolas Guilhot¹¹](#), [Annie Hofstetter¹²](#), [Johann Joets⁸](#), [Thierry Hotelier¹³](#), [Olivier Langella⁸](#), [Ludovic Legrand¹⁴](#), [Mikael Loaec¹⁵](#), [Virginie Lollier²](#), [Patrick Moreau¹⁶](#), [Emmanuelle Morin¹⁷](#), [Hadi Quesneville¹⁵](#), [Tovo Rabemanantsoa¹⁹](#), [Gérald Salin¹⁹](#), [Dominique Tessier²](#)

¹UMR Biologie du Développement et Reproduction Jouy-en-Josas, ²UR1268 BIA Biopolymères, Interactions Assemblages Nantes, ³UMR1332 BFP Biologie du Fruit et Pathologie Bordeaux, ⁴UMR1403 IPS2 Institut des Sciences des Plantes de Paris Saclay Versailles-Grignon, ⁵UAR1413 IngeNum Ingénierie Numérique en Recherche Toulouse, ⁶UMR1062 CBGP Centre de Biologie pour la Gestion des Populations Montpellier, ⁷UMR1019 UNH Unité de Nutrition Humaine Clermont-Ferrand, ⁸UMR0320 GQE Génétique Quantitative et Evolution - Le Moulon, ⁹UMR1349 IGEPP Institut de Génétique Environnement et Protection des Plantes Rennes, ¹⁰UMR1318 IJPB Institut Jean-Pierre BOURGIN Versailles-Grignon, ¹¹UMR1095 GDEC Génétique Diversité et Ecophysiologie des Céréales Clermont-Ferrand, ¹²UMR1135 LAMETA Laboratoire Montpelliérain d'Economie Théorique et Appliquée Montpellier, ¹³Montpellier SupAgro, ¹⁴UMR0441 LIPM Laboratoire des Interactions Plantes Micro organismes Toulouse, ¹⁵UR1164 URGI Unité de Recherche Génomique-Info Versailles-Grignon, ¹⁶UMR1248 AGIR AGroécologie, Innovations, teRritoires Toulouse, ¹⁷UMR1136 IaM Interactions Arbres/Micro-organismes Nancy, ¹⁸UMR1391 ISPA Interaction Sol Plante Atmosphère Bordeaux, ¹⁹US1426 GeT-PlaGe Génome et Transcriptome - Plateforme Génomique Toulouse.

Remerciements : Mathieu Andro, Philippe Cheneau, Eddie Iannuccelli, Lionel Million, Sylvie Nugier, Stéphane Paris, Rémi Person, Lina Sbeih, Florian Trincal