



France Grilles

<http://www.france-grilles.fr>

**INRA - CTIG - Jouy
7 novembre 2013**



Plan

- Présentation de France Grilles
- La stratégie vers les utilisateurs
- Conclusion

Présentation de France Grilles

- Le Groupement d'Intérêt Scientifique France Grilles (GIS) et ses missions
- L'Infrastructure de Recherche
- L'Initiative de Grille Nationale (NGI)
- Infrastructure et utilisateurs
- L'impact scientifique

France Grilles, un Groupement d'Intérêt Scientifique (2010)

- 8 partenaires
 - Ministère de l'Enseignement Supérieur et de la Recherche
 - CNRS, l'[Institut des Grilles et du Cloud](#) (IdGC) est le mandataire du GIS
 - Commissariat à l'Energie Atomique et aux énergies alternatives
 - **Institut National de la Recherche Agronomique**
 - Institut National de Recherche en Informatique et en Automatique
 - Institut National de la Santé et de la Recherche Médicale
 - Conférence des Présidents d'Université
 - Réseau National de Télécommunications pour la Technologie, l'Enseignement et la Recherche
- Des instances spécifiques (conseil de groupement - pour INRA Frédéric Garcia, comité scientifique - pour INRA Alain Franc, comité international...)

Les missions du GIS France Grilles

- Etablir et opérer une infrastructure nationale de grilles et de clouds de production, pour le traitement et le stockage de données scientifiques massives.
- Contribuer au fonctionnement de l'infrastructure européenne EGI (European Grid Infrastructure).
- Favoriser rapprochements et échanges entre les grilles/clouds de production et les grilles/clouds de recherche.

France Grilles, IR et NGI

- Une IR du Ministère de l'Enseignement Supérieur et de la Recherche (2009) pilotée par le GIS France Grilles
 - « Les très grandes infrastructures de recherche sont des instruments de recherche dont le coût et l'usage dépassent les capacités d'un seul organisme de recherche ou les finalités d'une seule discipline. »
- L'Initiative Nationale de Grille française - « National Grid Initiative » (NGI)
 - représente la France dans [EGI \(European Grid Infrastructure\)](#)

Pour mémoire : qu'est-ce qu'une grille ?

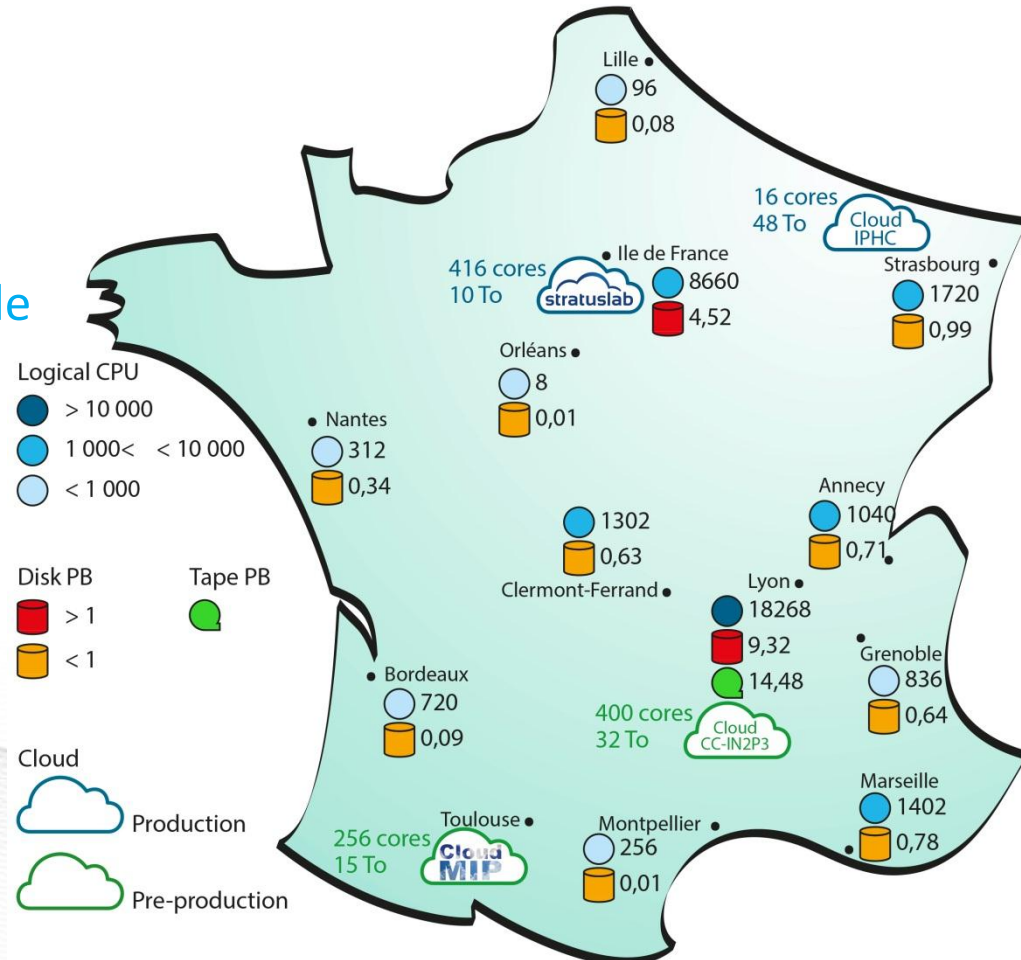
- « **grilles informatiques** » : infrastructures virtuelles
 - ensemble d'ordinateurs géographiquement distribués
 - fonctionnant en réseau
 - pour fournir une puissance globale.
- **disposer à moindre coût de :**
 - puissance de calcul
 - stockage de données
 - pensé pour l'analyse de données haut débit

Utilisateurs organisés en VO (Virtual Organization)

L'infrastructure : les ressources informatiques

- **18 sites grilles** juillet 2013
- **34 620 CPUs logiques** / 342 449 HEPSPC06 (définition)
- **18 PB/14,5 PB** stockage disque/bande

- ressources mutualisées,
- apportées par les communautés,
- accessibles aux partenaires du GIS,
- besoins liés à la recherche,
- accessibles aux organismes ou entreprises dans le cadre de projets communs avec ces partenaires.



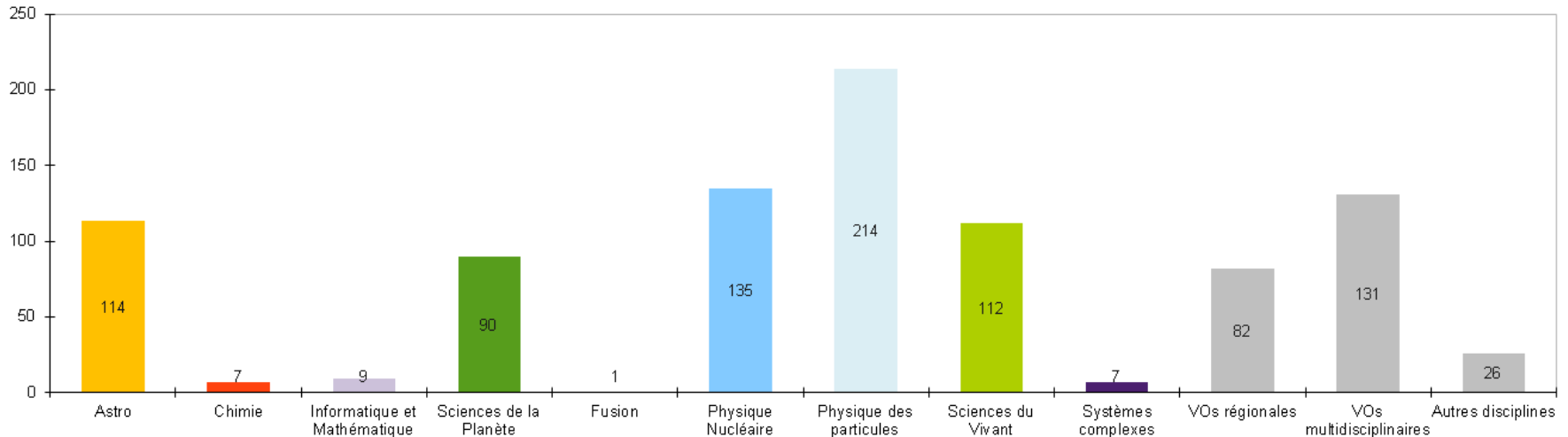
mai 2010 – août 2013 : disponibilité : **97.45%** fiabilité : **98.60%**



Qui utilise les services de France Grilles ?

- environ 750 utilisateurs « nationaux » regroupés en organisations virtuelles (VO)
- 89 Organisations Virtuelles (VO) accueillies sur l'infrastructure dont 27 nationales ou locales

Nombre d'utilisateurs France Grilles par discipline
selon leur inscription aux VOs
31 août 2013



un utilisateur peut être membre de plusieurs VOs



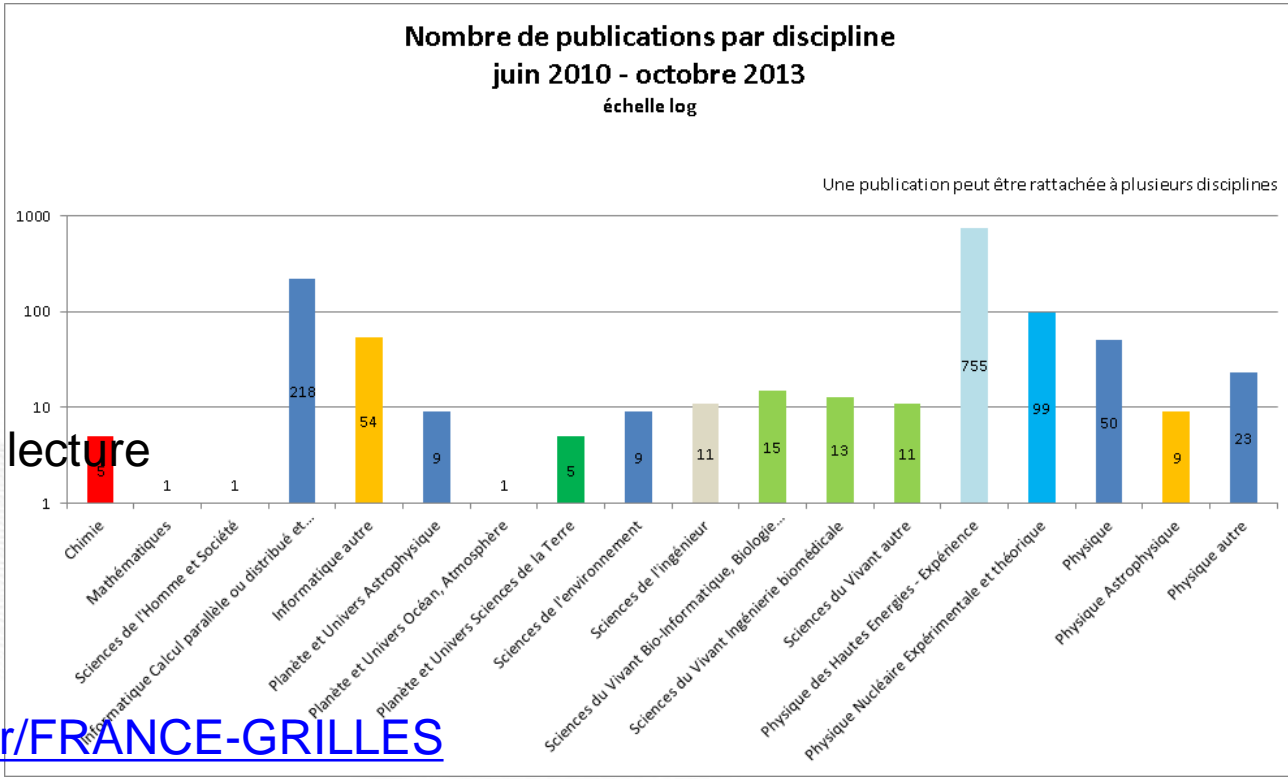
France Grilles : impact scientifique

Indicateurs bibliométriques pour :

- refléter l'activité scientifique et en mesurer l'impact,
- justifier *a priori* ou *a posteriori* l'usage des grilles de production
- établir des statistiques

une collection HAL
pas encore représentative !

1450 référencement
dont 264 articles
dans revues avec comité de lecture



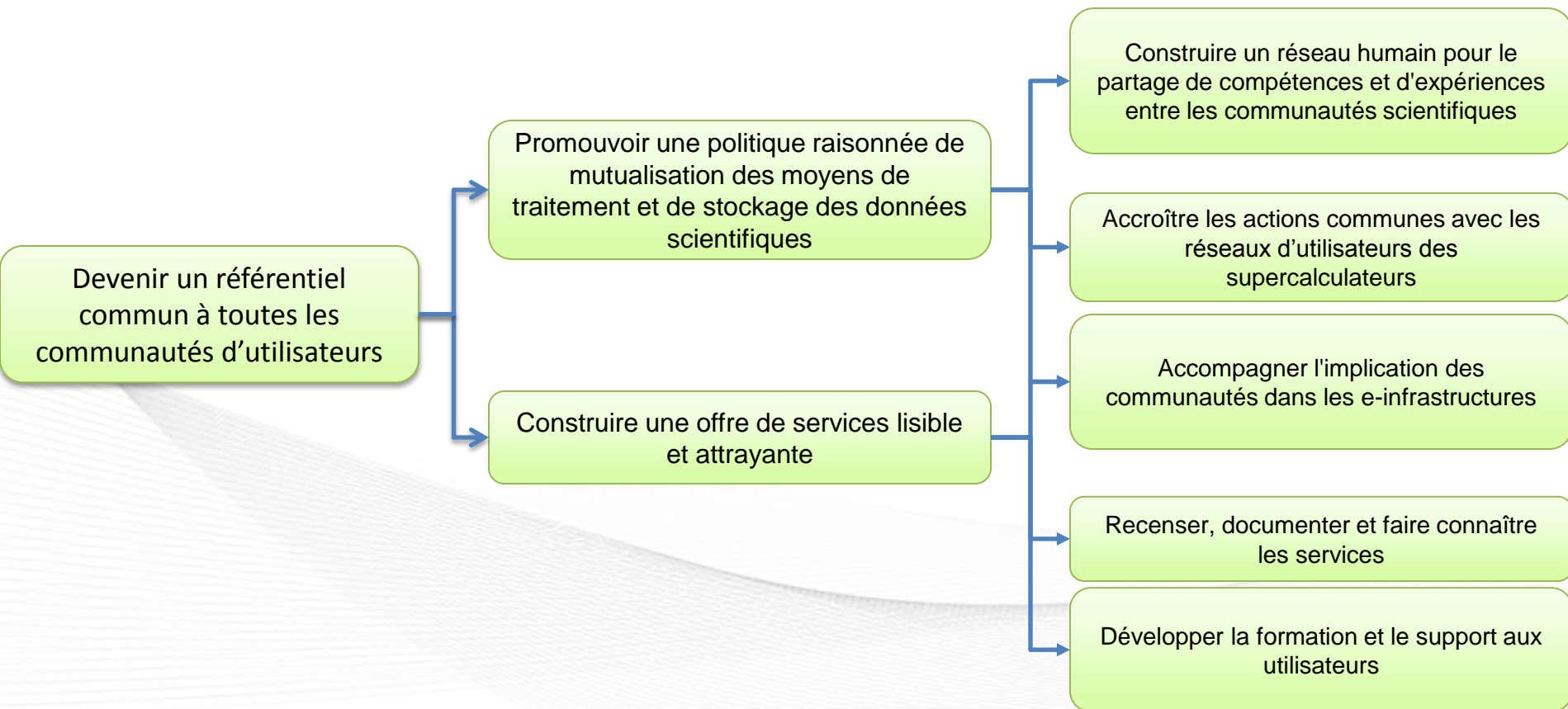
<http://hal.archives-ouvertes.fr/FRANCE-GRILLES>



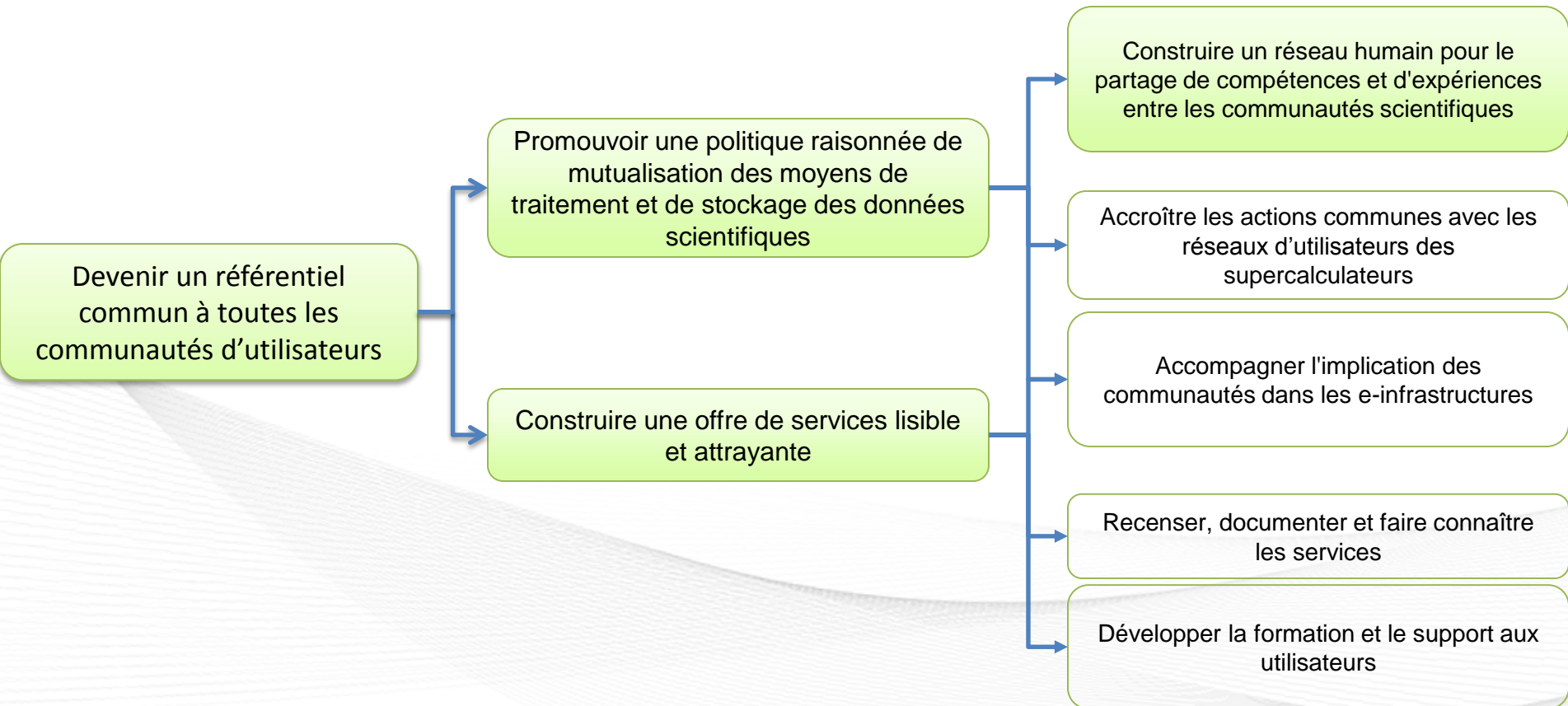
Plan

- Présentation de France Grilles
- La stratégie vers les utilisateurs
- Conclusion

Stratégie vers les utilisateurs



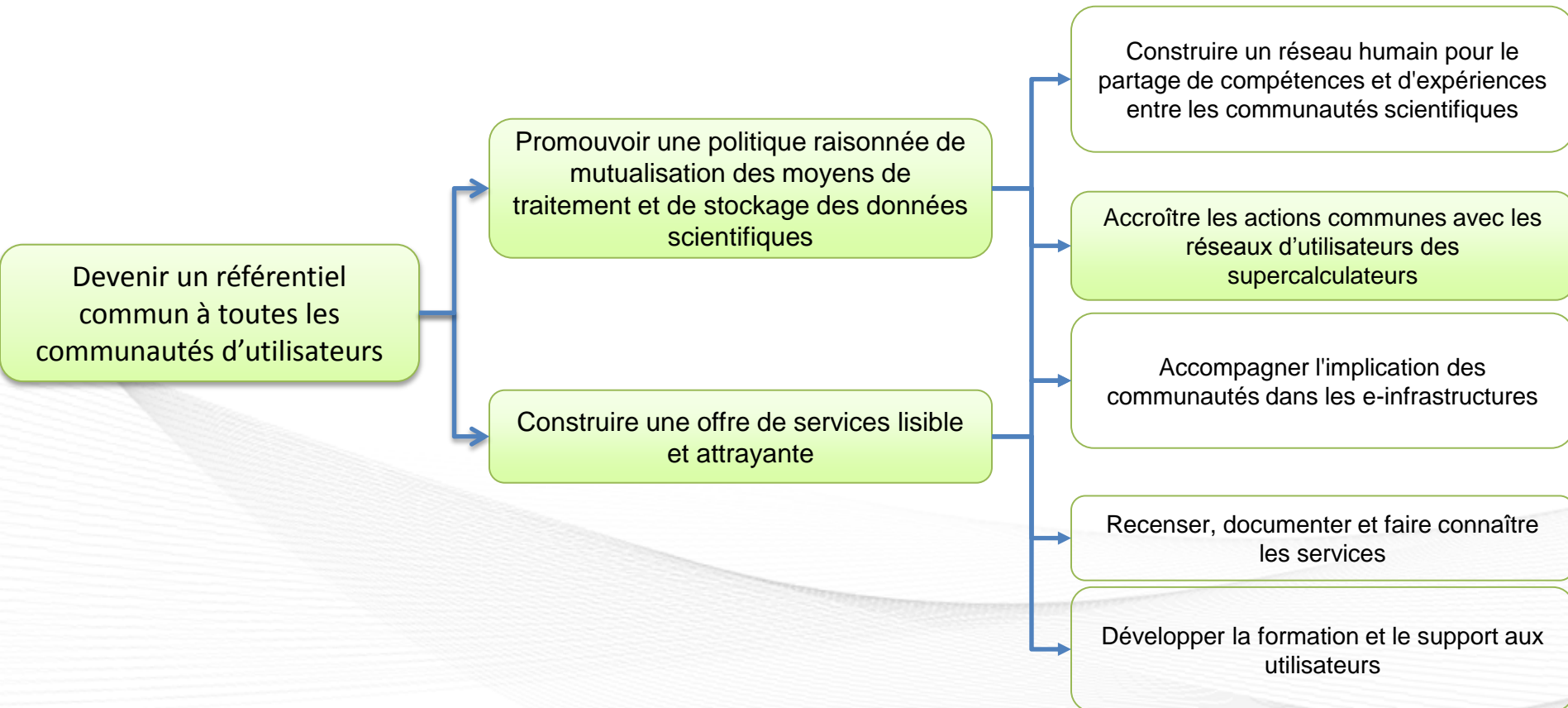
Stratégie vers les utilisateurs



Réseau humain / partage entre communautés scientifiques

- **Implication dans les réseaux**
 - RBDD (rbdd.cnrs.fr), resinfo, Plume (thématique « grilles/cloud » et référencement des logiciels utilisés/produits dans les communautés)
- **Contexte international EGI (www.egi.eu)**
 - groupes de travail thématiques internationaux (fédération de clouds, communautés émergentes sur EGI...)
 - conférences internationales, veille, évolutions techniques...

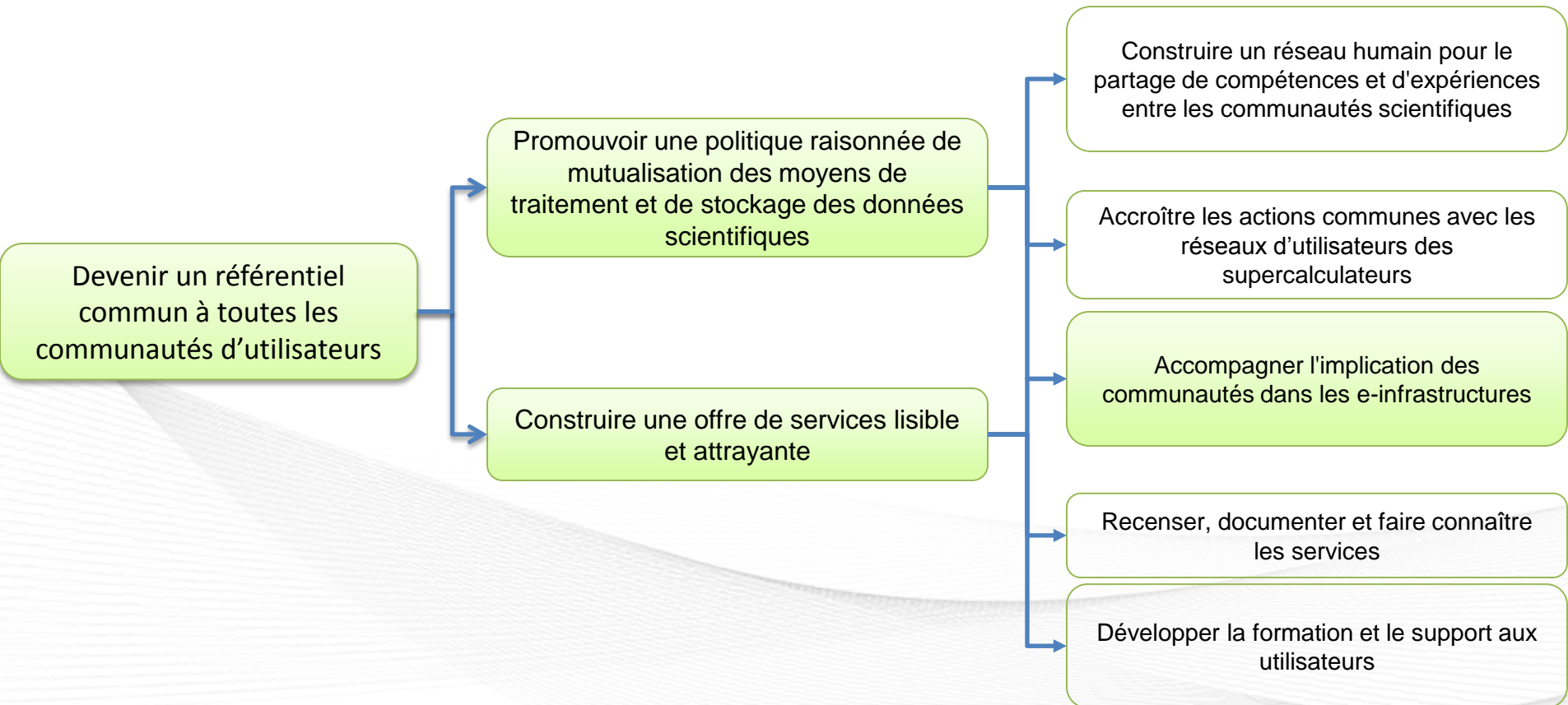
Stratégie vers les utilisateurs



Actions communes

- 13-14 novembre 2013 : <http://succes2013.sciencesconf.org>
 - journées co-organisées avec le [Groupe Calcul](#)
 - programme : 13 présentations orales, 2 posters, autant d'articles
 - webcastée et enregistrées
 - Suivez le webcast !
- 2012 : journées scientifiques mésocentres et France Grilles
 - <http://mesogrilles2012.sciencesconf.org/>
 - 12 présentations orales, 10 posters, autant d'articles en ligne,
 - 147 inscrits + webcast
 - enregistrements en ligne

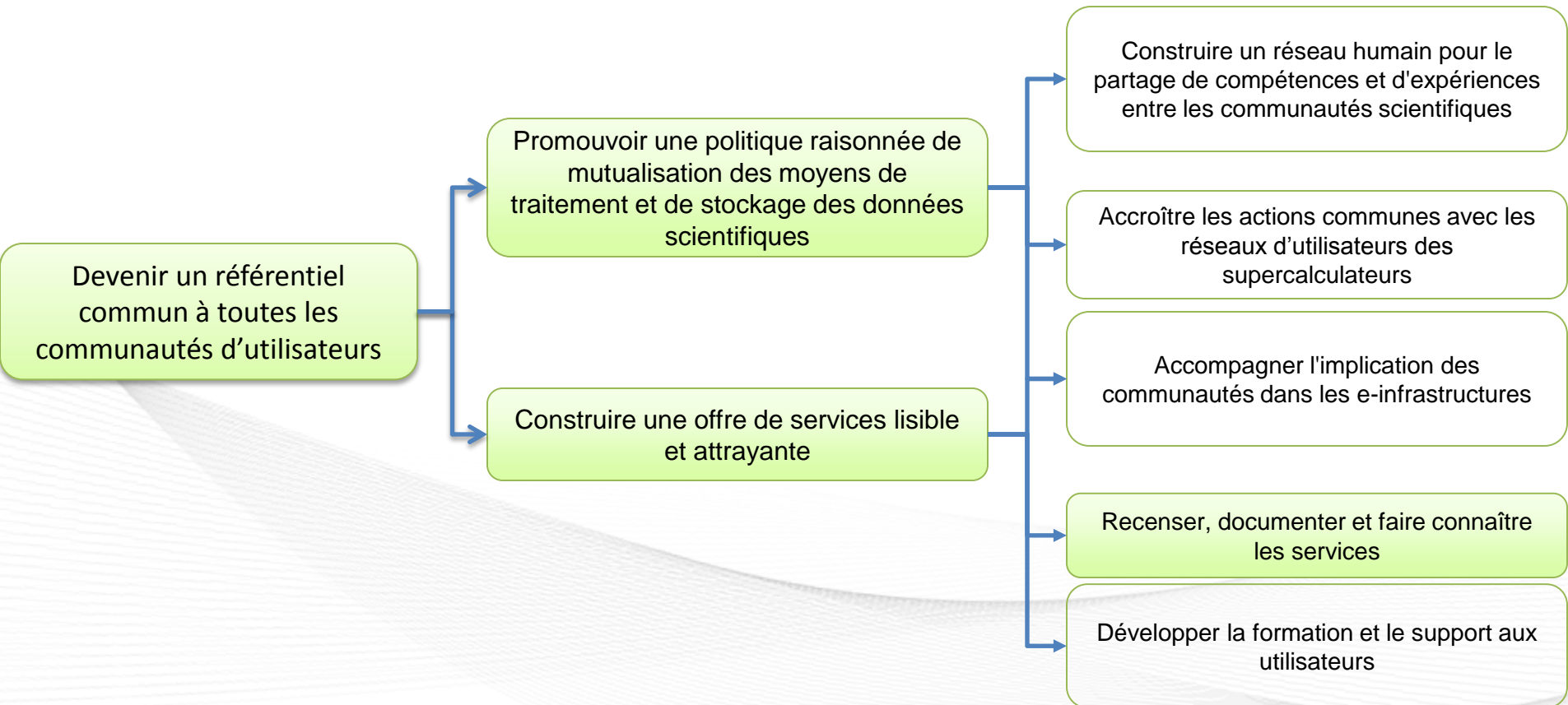
Stratégie vers les utilisateurs



Accompagner l'implication des communautés dans les e-infrastructures

- **Thématique Environnement et Biodiversité :**
 - Yannick Legré, (Yannick.Legre@idgrilles.fr)
 - projet ENVRI <http://envri.eu/>
 - projet Creative-B <http://creative-b.eu/>
- **Budget France Grilles « communautés »**
 - **structuration communautés**
 - **participation à EGI pour profit aux communautés / FG**
 - échanges entre communautés (expertise ...), valorisation logiciels libres utiles à d'autres communautés, financement de publications
- **Recensement des publications dans HAL**

Stratégie vers les utilisateurs



Catalogue de services « utilisateurs »

- Un catalogue de services évolutif
 - services offerts par un/des partenaires France Grilles,
 - services offerts par des communautés sur les infrastructures « France Grilles ».
- Profils adressés :
 - chercheur / ingénieur,
 - responsable ou représentant d'un groupe de chercheurs/ingénieurs,
 - responsable ou administrateur d'un portail scientifique,
 - enseignant (enseignement supérieur).
- N'inclut pas :
 - Les services offerts par France Grilles à d'autres infrastructures ou projets (ex W-LCG),
 - Les services internes (sécurité, monitoring, accounting...).

Vous êtes/représentez un/des chercheur(s) / ingénieur(s) - 1

- Besoin : calcul / traitement de données (éventuellement distribuées / partagées) ; stockage des données
- Service proposé : DIRAC, solution générale et complète pour une ou plusieurs communautés d'utilisateurs
 - fonctionnalités : gestion des tâches de calcul et des données distribuées.
 - léger, robuste et extensible, adaptable aux évolutions rapides de la grille.
 - prend en compte des ressources hétérogènes (grilles de calcul, clouds ou grappes de processeurs indépendants)
 - facilite une bonne expérience utilisateur face aux ressources distribuées
- Communautés utilisatrices : vo.formation.idgrilles.fr, vo.france-grilles.fr, vo.msfg.fr (Montpellier), vo.mcai.fr (Bordeaux)...
- Description du logiciel utilisé : www.projet-plume.org/fiche/DIRAC
- Contact : info@france-grilles.fr

Cas particulier d'un groupe organisé en VO

- Besoin : traitement et/ou stockage de données scientifiques pour un groupe
- Service proposé : FG-DIRAC
 - possibilité si le groupe est organisé en VO de bénéficier du service pour la VO (selon la politique d'accès au service)
- Contact : info@france-grilles.fr

Vous êtes/représentez un/des chercheur(s) / ingénieur(s) - 2

- **Besoin : besoins en terme de stockage ou problématiques de gestion des données.**
- **Service proposé : iRODS - système de virtualisation**
permet :
 - de créer des collections de données et de les gérer simplement ;
 - d'assurer une haute disponibilité des données en les distribuant sur des systèmes distants ou en les répliquant ;
 - d'organiser l'arborescence des fichiers indépendamment de leur organisation physique sur les serveurs de données ;
 - de rechercher des lots de données via un catalogue de métadonnées contenant les informations relatives à l'état des données et leur description ;
 - de contrôler finement l'accès aux données ;
 - de faciliter la gestion de grandes collections de données en appliquant des politiques et des règles de gestion des données définies par l'utilisateur afin d'automatiser les tâches administratives et de renforcer la sécurité et l'intégrité des données ;
 - d'accéder aux données à distance depuis des ordinateurs individuels, des grilles ou des "clouds".

Vous êtes/représentez un/des chercheur(s) / ingénieur(s) - 2

- Description du logiciel utilisé : www.projet-plume.org/fiche/irods
- Service opérationnel basé sur la plate-forme mutualisée FG-iRODS
 - disponible pour les membres de l'organisation virtuelle (VO) vo.france-grilles.fr.
- Contact : info@france-grilles.fr

Vous êtes/représentez un/des chercheur(s) / ingénieur(s) - 3

- **Besoin : cloud IaaS**
- **Stratégie France Grilles pour le cloud :**
 - promouvoir le développement d'une offre cloud académique en France
 - stratégie d'investissement (2011 et 2012)
 - Toulouse, 2011, CloudMIP basé sur OpenNebula
 - CC-IN2P3, 2011, OpenStack, cloud IaaS communautaire, membre de la fédération de clouds EGI
 - soutien à Stratuslab, 2011
 - investissement pour un démonstrateur au CC-IN2P3, 2012
 - dimension européenne à travers la fédération de clouds
- **Services proposés actuellement :**
 - Stratuslab, (description du logiciel www.projet-plume.org/relier/stratuslab)
 - tests possibles sur le cloud CC-IN2P3 pour des utilisateurs experts

Services offerts par des communautés thématiques, VIP 1/2

- Communauté Life-Science Grid Community (LSGC) – VO biomed
- Service proposé : la plate-forme VIP.
 - Accès aux ressources de calcul et de stockage de la grille EGI pour la simulation médicale, en particulier en imagerie.
 - Dans ce portail, les applications sont encapsulées comme des workflows MOTEUR et exécutées sur l'organisation virtuelle biomed en utilisant l'instance DIRAC mise à disposition par France-Grilles.
 - VIP utilise un certificat robot et expose une interface haut-niveau où la grille est totalement cachée.
 - En plus de ce service d'exécution générique, VIP offre aussi des fonctionnalités spécifiques à la simulation d'images médicales : des modèles anatomiques et pathologiques multi-physiques sont disponibles dans un catalogue, et les données simulées sont annotées automatiquement en utilisant les concept d'une ontologie

Services offerts par des communautés thématiques, VIP 2/2

- Pour en savoir plus :
 - En production depuis fin 2010, 476 utilisateurs de 51 pays sont enregistrés dans VIP en octobre 2013 .
 - A virtual imaging platform for multi-modality medical image simulation. Glatard T., Lartizien C., Gibaud B., Ferreira Da Silva R., Forestier G., Cervenansky F., Alessandrini M., Benoit-Cattin H., Bernard O., Camarasu-Pop S. et al
IEEE Transactions on Medical Imaging 32, 1 (2013) 110-8 -
<http://www.hal.inserm.fr/inserm-00762497>
 - Une présentation aux JRES 2013 le 11 décembre : VIP et GateLab : retour d'expérience
 - <https://vip.creatis.insa-lyon.fr/>

Services offerts par des communautés thématiques OpenMOLE

- **Communauté Systèmes complexes**
 - « The goal of the vo.complex-systems.eu is to promote the study of complex systems and complex networks on the Grid infrastructure. The vo.complex-systems.eu Virtual Organization will also serve as the building layer of collaboration among international scientists focusing on the research area of Complexity Science. “
- **Service proposé : OpenMOLE**
- **Description du logiciel utilisé : <https://www.projet-plume.org/fiche/openmole>**

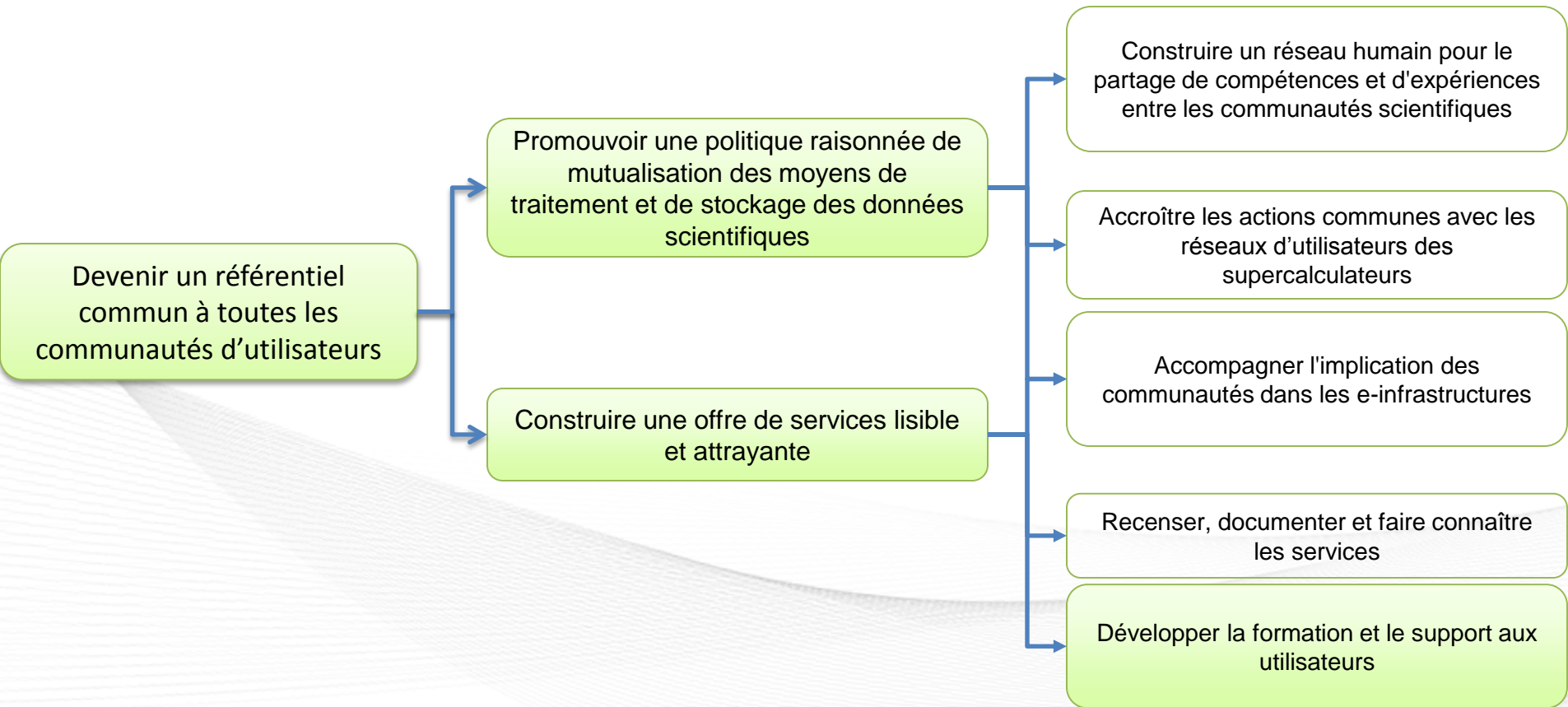
Vous êtes enseignant (enseignement supérieur)

- **Besoin** : vous souhaitez faire réaliser à vos étudiants des travaux pratiques sur une infrastructure distribuée
- **Disciplines** : informatique (étude des infrastructures distribuées) ou discipline scientifique (besoin en calcul scientifique)
- **Service proposé** :
 - Organisation Virtuelle de formation France Grilles et infrastructure
 - certificats pour l'enseignant et les étudiants
- **Contact (au moins 2 mois avant)** : info@france-grilles.fr

Vous êtes administrateur ou responsable d'un portail scientifique

- **Besoin :**
 - vous avez/aurez la charge d'un portail scientifique (site web donnant accès à des applications scientifiques)
 - vous souhaitez faire « tourner » ces applications sur l'infrastructure France Grilles (ou EGI)
 - vous souhaitez faciliter l'accès aux utilisateurs (simplification de l'authentification) en toute sécurité
- **Service proposé**
 - certificat « robot » pour le portail
- **Contact : info@france-grilles.fr**

Stratégie vers les utilisateurs



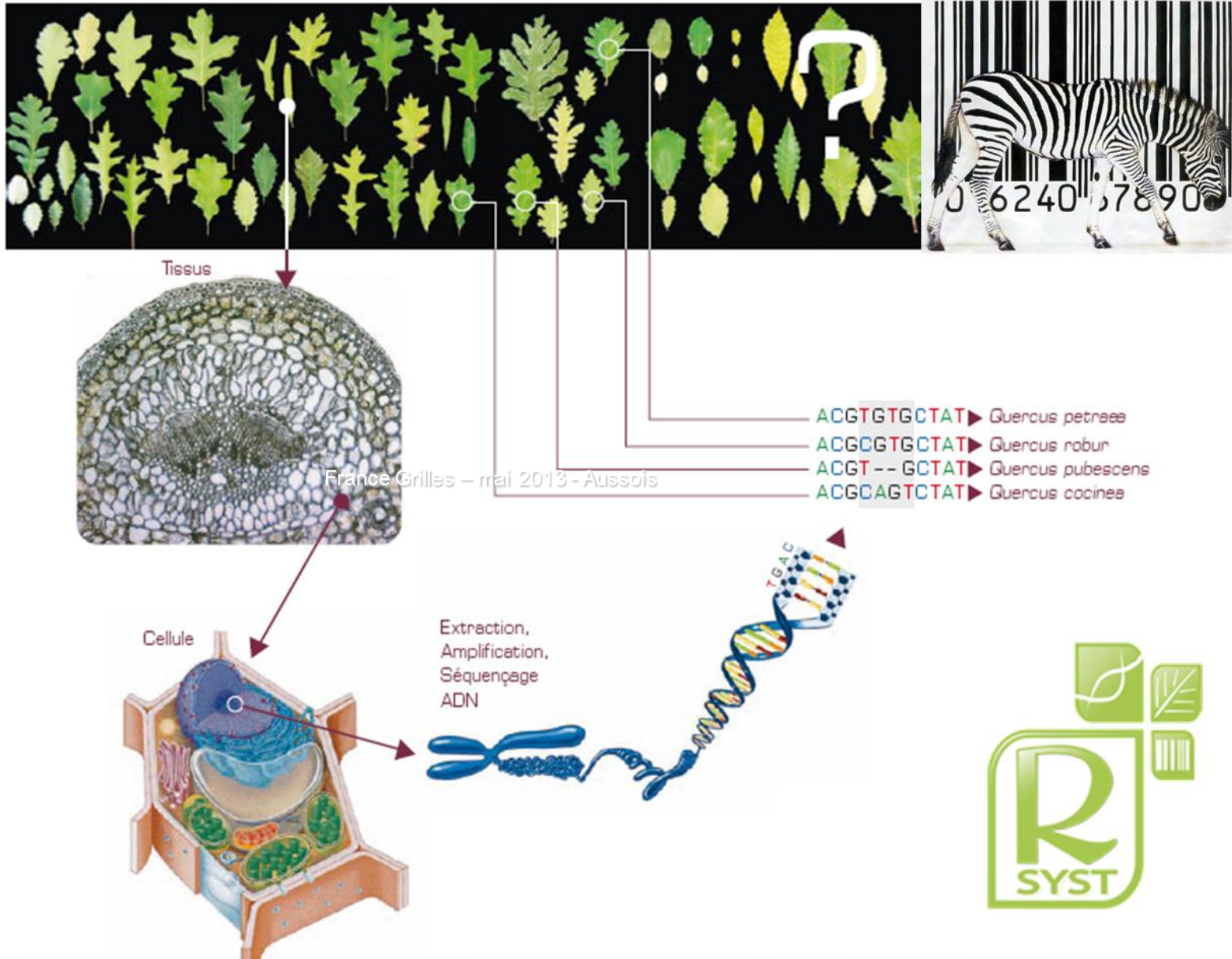
Formation et support

- **Accompagnement des nouveaux utilisateurs selon discipline / besoins :**
 - => communauté scientifique ou régionale ou nationale existante / services associés,
- **Documentation collaborative / forums :**
 - <https://forge.in2p3.fr/projects/france-grilles-documentation>
- **Formations :**
 - Prochaine session : 3 j début février DIRAC-iRODS à Clermont-Ferrand, vous êtes intéressé (e) ? faites-vous connaître !

Plan

- Présentation de France Grilles
- La stratégie vers les utilisateurs
- Conclusion

Un exemple à l'INRA



Barcoding sur la grille

Barcoding : diagnostic taxonomique d'après une séquence (qq centaines bp)

On dispose d'une base de référence, et d'une « query »

On calcule la distance de la query à chaque élément de la base
alignement local

BLAST : heuristique quand la base est volumineuse

exact si taille de la base « raisonnable » : qqs milliers spécimens

On le fait pour tous les reads issus de NGS : qqs centaines de milliers

Problème massivement distribué

Programme C (`disseq`) de calcul,
programme Perl de distribution en paquets
appel par Dirac

Chaumeil, Frigerio, Franc & al,
INRA Bdx, UMR BioGeCo
Journée Grille 2012, et in prep

Et vous ?

- **INRA partenaire du GIS => infrastructures FG accessibles (stockage et traitement des données)**
 - y compris pour des projets communs avec d'autres partenaires (académiques, privés, collaborations internationales...)
- **Pourquoi utiliser les ressources de France Grilles ?**
 - limite des infrastructures locales
 - collaboration avec des équipes distantes et souhait de partager des données.
 - pour faire partie d'une communauté plus large (nationale ou internationale), partager des outils, des ressources...
- **Quelles VOs pour les utilisateurs INRA ?**
 - sciences de la vie (biomed), biodiversité/environnement (en cours de création à l'international), vo.france-grilles.fr, ... selon contexte et besoins

Les conditions d'accès

- **Combien ça coûte ? Votre équipe ou votre laboratoire ne paiera rien !**
 - Infrastructures financées par participation des organismes, subventions régionales, projets, TGIR France Grilles, ...
 - Personnels financés par organismes participants / projets.
 - Si les besoins de votre équipe ou communauté deviennent importants, participation à des ressources humaines à envisager.
- **Que me demandera France Grilles ? Une participation à la vie de la communautés des utilisateurs :**
 - présentation des travaux lors des rencontres scientifiques
 - référencement des publications réalisées grâce à France Grilles dans la collection...
- **Quelles démarches devrai-je faire ? prendre contact avec info@france-grilles.fr**
 - renseigner un questionnaire assez simple pour évaluer vos besoins ou ceux de votre équipe (ressources, formation...),
 - une ou plusieurs réunions (tel ou visio) seront probablement nécessaires avec l'équipe France Grilles pour bien évaluer vos besoins et décider d'un plan de travail avant de pouvoir travailler réellement sur les infrastructures.

www.france-grilles.fr pour en savoir plus...

- La gazette FG : <http://www.france-grilles.fr/-FG-Info-Express->
 - Abonnement à utilisateurs-l@france-grilles.fr pour la recevoir.
- Exemples d'utilisation de l'infrastructure
 - [rencontres scientifiques 2011](#)
 - [journées scientifiques mésocentres et France Grilles 2012](#)
 - <http://succes2013.sciencesconf.org/program>
- Contact France Grilles : info@france-grilles.fr N'hésitez pas !
- Formation : février 2014 et abonnement à formation-l@france-grilles.fr
- Plume : www.projet-plume.org/grille et deux mots-clés spécifiques "Infrastructure FG" et "Utilisateurs FG" outils pour les outils du catalogue FG