

NOTE SUR LE RENOUVELLEMENT DU GIS FRANCE GRILLES

V. Breton

Ce document a pour but de résumer les principaux accomplissements du Groupement d'Intérêt Scientifique France Grilles depuis sa création en Juin 2010 et de présenter des éléments stratégiques dans la perspective de son renouvellement.

1 INTRODUCTION

France Grilles est un Groupement d'Intérêt Scientifique créé en Juin 2010 par 8 partenaires : le Ministère de l'Enseignement Supérieur et de la Recherche, le CEA, le CNRS, la Conférence des Présidents d'Université, l'INRA, l'INRIA, l'INSERM et RENATER. Les missions principales de France Grilles sont les suivantes :

- établir et opérer une infrastructure nationale de grille de production, pour le traitement et le stockage de données scientifiques massives
- contribuer avec les autres états membre impliqués au fonctionnement de l'infrastructure européenne EGI et définir les modalités de la participation française à EGI.eu
- favoriser les rapprochements et les échanges entre les équipes travaillant sur les grilles de production et les grilles de recherche

Ces missions s'étendent depuis 2012 au domaine du Cloud Computing.

2 ACCOMPLISSEMENT DES OBJECTIFS

Trois ans après la création du GIS France Grilles, la France dispose aujourd'hui d'une infrastructure de grille de production (figure 1) fiable et performante offrant plus de 35000 processeurs et 18 petaoctets de disque à plus de 750 utilisateurs issus de multiples communautés scientifiques (figure 2). Plus de 900 publications (figure 3) issues des travaux accomplis sur cette infrastructure sont recensées dans la base de données HAL du CNRS pour la période 2010-2013. Au niveau international, la France a fourni depuis 3 ans environ 13% des ressources de la grille EGI. La France siège au conseil d'EGI et est représentée depuis Décembre 2012 au conseil d'administration de la fondation EGI.eu.

2.1 UNE INFRASTRUCTURE TRES PERFORMANTE SUR L'ENSEMBLE DU TERRITOIRE

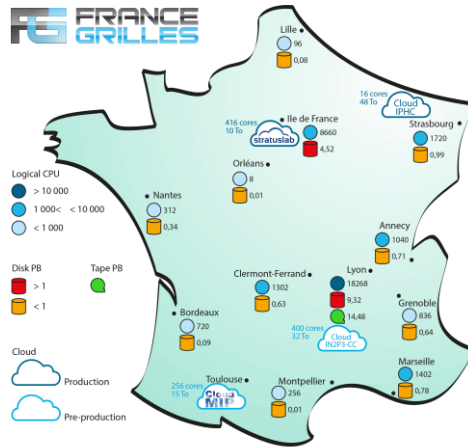


Figure 1 : carte des ressources de la grille de production et de la fédération de clouds académiques

Sous la coordination technique du CC-IN2P3, l'infrastructure de grille de production proposée par France Grilles (figure 1) met à la disposition de ses utilisateurs environ 35000 processeurs et 18 petaoctets de stockage dans 18 sites sur le territoire de la France métropolitaine. Ces ressources sont fournies à 85% par des laboratoires de l'IN2P3, 10% par le CEA et 5% par les autres partenaires (CNRS hors IN2P3, universités, BRGM).

Les indicateurs de fiabilité et de disponibilité (tableau 1) sont excellents pour l'année 2013, dans la continuité des années précédentes.

Métrique	MOYENNE
Disponibilité tous services confondus	0,97439
Fiabilité tous services confondus	0,98634
Disponibilité services Critiques (TopBDII)	0,99996
Fiabilité services Critiques (TopBDII)	0,99996

Tableau 1. Indicateurs de performances des opérations

Depuis Janvier 2013, France Grilles a créé une fédération de clouds académiques avec 2 sites en production (Grille Régionale Ile-de-France, Strasbourg) et deux sites en pré production (Toulouse, Lyon). Une partie des ressources du site de Toulouse est intégrée à Grid5000¹.

¹ Grid5000 est une plate-forme expérimentale pour la recherche sur les systèmes distribués composée de 1600 machines dans 26 clusters et 10 sites en France (<http://grid5000.fr>)

Le traitement des données scientifiques à haut débit est aujourd’hui indissociable de la gestion de ces données. France Grilles lance une offre distribuée de stockage et de manipulation de grands volumes de données scientifiques basée sur la technologie iRODS².

2.2 UNE VRAIE PLURIDISCIPLINARITE DANS L’UTILISATION DES SERVICES ET DES RESSOURCES

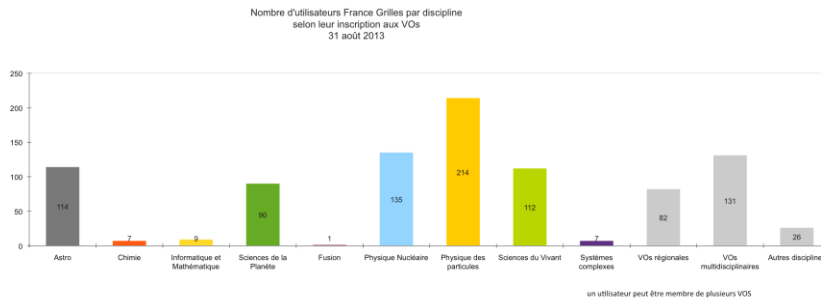


Figure 2 : distribution des utilisateurs par organisations virtuelles

Les utilisateurs des ressources de la grille de production nationale titulaires de certificats sont plus de 750 si l’on ne tient compte que de ceux inscrits dans des organisations virtuelles disciplinaires (figure 2). De plus en plus d’utilisateurs accèdent aux ressources de la grille à travers des plates-formes qui visent à devenir de véritables laboratoires virtuels, sans être titulaires de certificat.

2.3 UNE PRODUCTION SCIENTIFIQUE SIGNIFICATIVE MAIS SOUS-ESTIMÉE

Sur les trois dernières années, plus de 900 publications ont été recensées dans la base de données HAL du CNRS (figure 3). Le prix Nobel 2013 attribué aux physiciens François Englert et Peter Higgs récompense aussi les personnels du CEA et du CNRS qui ont participé à la découverte du boson de Higgs dans les données du collisionneur LHC au CERN analysées sur la grille LCG (LHC Computing Grid).

² iRODS est un système de virtualisation du stockage qui gère aujourd’hui plusieurs petaoctets de données dans de nombreux projets de la biologie aux archives nationales françaises en passant par la physique des particules et l’astronomie

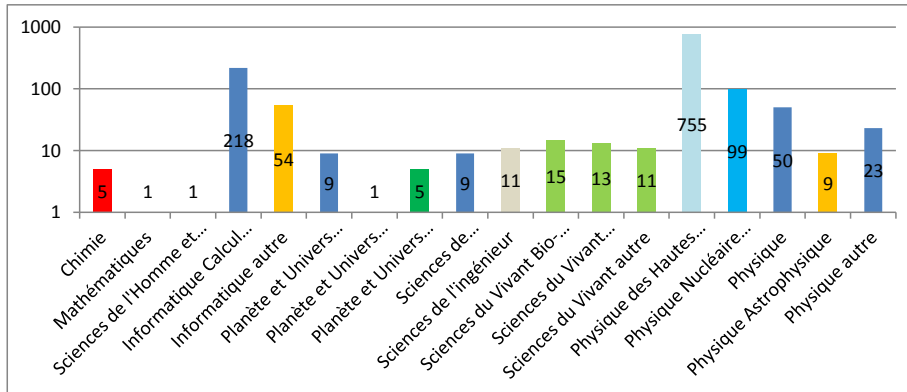


Figure 3 : nombre de publications recensées dans la base de données HAL avec le label France Grilles

Cette production scientifique significative ne doit pas masquer la difficulté à collecter les publications dans les domaines scientifiques où l'enregistrement dans HAL n'est pas automatique ou très fortement recommandé (physique des hautes énergies et informatique).

2.4 UNE FORTE IMPLICATION DANS EGI

La France siège au conseil d'EGI depuis sa création avec le même droit de vote que l'Allemagne et le Royaume-Uni. Elle est représentée au conseil d'administration de la fondation EGI.eu depuis Décembre 2012 par Vincent Breton. Elle opère un des principaux services centraux d'EGI, le portail des opérations, développé et maintenu par le CC-IN2P3. Les ingénieurs et chercheurs français ont aussi été très actifs dans les groupes de travail (Virtual Teams) mis en place par la fondation EGI.eu pour améliorer la réponse aux besoins des utilisateurs.

Dans le cadre du projet européen EGI-Inspire, un appel à projets a été lancé à l'été 2013 pour des services innovants à déployer en 2014 sur EGI. Plus d'un tiers des projets sélectionnés sont portés par ou impliquent des équipes françaises, ce qui les positionne de façon très favorable dans la perspective d'Horizon 2020.

2.5 LE DEVELOPPEMENT DE SYNERGIES ENTRE UTILISATEURS DES RESSOURCES ET CHERCHEURS EN INFORMATIQUE DISTRIBUE

La recherche en informatique est au coeur de la stratégie France Grilles. Le GIS soutient et promeut le transfert de technologies issues des équipes de chercheurs en informatique vers les utilisateurs. Un excellent exemple est la plate-forme VIP (<http://www.creatis.insa-lyon.fr/vip>), développée par les laboratoires CREATIS et I3S pour servir la communauté de recherche en imagerie médicale et en neurosciences, qui compte 480 utilisateurs dont 175 français et qui représente un quart de l'utilisation des ressources de la plus grosse organisation virtuelle internationale en sciences du vivant sur EGI (figure 4).

Code de champ modifié

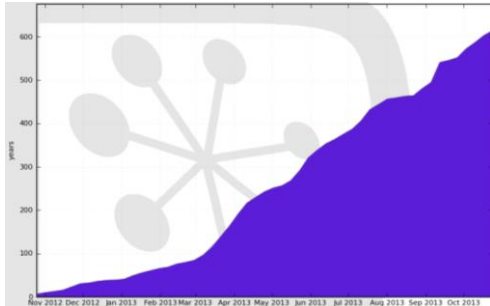


Figure 4 : évolution de l'utilisation mensuelle des ressources de la grille à travers la plate-forme VIP (crédit : T. Glatard)

Un autre exemple est le projet de plate-forme d'expérimentation ESPERANCE dans lequel tous les acteurs académiques français du calcul intensif et distribué se sont impliqués. Très fortement soutenu par le Ministère de l'Enseignement Supérieur et de la Recherche, le projet a été construit en totale synergie entre France Grilles et tous les autres acteurs académiques nationaux du calcul intensif.

2.6 AUTRES ACTIONS

2.6.1 COLLABORATION AVEC LES MESOCENTRES

France Grilles a mené une politique d'ouverture vers les mésocentres : les journées SUCCEs organisées pour la deuxième année conjointement par France Grilles, le groupe Calcul et la coordination des mésocentres ont suscité beaucoup d'intérêt dans la communauté des utilisateurs de ces ressources.

2.6.2 ACTIONS SPECIFIQUES VERS LES PARTENAIRES DU GIS

France Grilles développe une action vers les partenaires du GIS moins impliqués au départ dans les grilles que le CEA et le CNRS. C'est notamment le cas avec l'INRA avec un projet de laboratoire virtuel porté par des chercheurs de l'INRA et soutenu par France Grilles. Suite aux journées SUCCEs, une étude des besoins de l'INSERM est aussi programmée pour préparer Horizon 2020.

3 RENOUVELLEMENT DU GIS : CONTEXTE INTERNATIONAL

Le renouvellement du GIS France Grilles s'inscrit dans un contexte international complexe.

3.1 EGI

En lien avec le démarrage d'Horizon 2020, plusieurs pays européens ont redéfini ou sont en cours de redéfinition de leurs politiques nationales en matière de grilles et de clouds. EGI s'en trouve fragilisé. Suite au retrait de l'Allemagne de la fondation EGI.eu, les autres pays partenaires d'EGI ont voté le 24 Octobre 2013 un budget de 1,1 Millions d'euros pour 2014 pour la fondation EGI.eu. Parmi les 37 pays membres d'EGI, deux autres grands pays européens (Italie, Royaume-Uni) n'ont pas voté le budget 2014 au dernier conseil, pour des raisons différentes de l'Allemagne et qui sont explicitées ci-dessous.

3.1.1 SITUATION EN ALLEMAGNE

Le 27 Septembre, le président de la Gauss Allianz, le Prof. Dr. Wolfgang E. Nagel, a adressé un courrier au chairman du conseil d'EGI pour l'informer du retrait de l'Allemagne de la fondation EGI.eu, structure légale qui anime et coordonne les activités de l'initiative de grille européenne EGI. La raison invoquée est l'impossibilité pour l'Initiative de Grille Allemande, représentée par la Gauss Allianz, de payer en 2014 la contribution allemande au niveau de 2013, soit 173 KEuros.

Dans son courrier, il précise que l'Allemagne est toujours très intéressée par EGI et soutient ses missions. Il a aussi précisé que les partenaires allemands impliqués dans le projet EGI-Inspire tenaient à remplir leurs obligations contractuelles.

Des négociations sont en cours entre les représentants allemands au conseil d'EGI et le bureau exécutif de la fondation EGI.eu pour explorer les modalités d'une continuité de la représentation allemande en 2014 au conseil d'EGI.

Code de champ modifié

3.1.2 SITUATION EN ITALIE ET AU ROYAUME-UNI

Le représentant de l'Italie au conseil d'EGI a indiqué que l'Italie n'était pas en mesure de payer sa contribution à EGI en 2014 au niveau de 2013. La gestion de l'initiative de Grille Italienne qui incombait depuis sa création à l'INFN (Institut National de Physique Nucléaire) va être transférée à GARR, l'opérateur des réseaux de recherche et éducation italien (NREN).

Côté anglais, le JISC (Joint Information Systems Committee) a payé pendant trois ans la contribution anglaise à EGI. Du fait de sa restructuration et d'un changement du modèle de souscription, des discussions sont en cours pour que les Research Councils reprennent la charge de la cotisation anglaise. Elles doivent déboucher d'ici la fin de l'année.

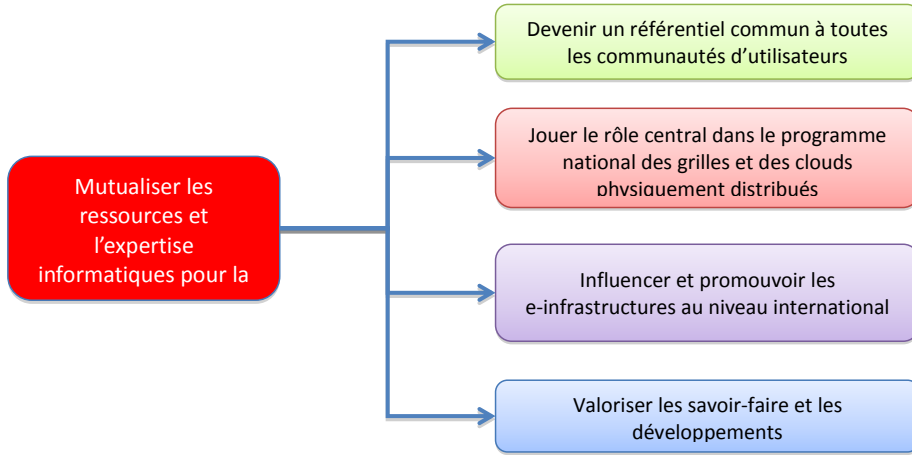
3.2 HORIZON 2020

Le programme de travail 2014-2015 sur les infrastructures de recherche européennes offre de nombreuses opportunités aux équipes françaises impliquées dans France Grilles de répondre dès les premiers appels à projets prévus en Décembre 2013 au démarrage d'Horizon 2020. Si la ligne de financement dédiée à EGI a été réduite à 8 Millions d'euros, la Commission Européenne a encouragé le développement de collaborations entre EGI, EUDAT et HelixNebula (porté par le CERN) pour des réponses concertées aux autres appels. France Grilles participe activement aux réunions de concertations au niveau européen et s'attache à stimuler l'implication des acteurs français dans le programme de travail de l'unité eInfrastructure (c1) de la DG Connect, soit en soumettant des projets, soit en devenant experts.

4 PROPOSITION DE STRATEGIE

4.1 INTRODUCTION

La vision de France Grilles est de construire et opérer une infrastructure informatique distribuée ouverte à toutes les sciences et aux pays en développement qui constitue un espace ouvert de collaboration au sein et entre les disciplines.



La stratégie proposée pour mettre en œuvre cette vision se décline selon quatre axes (figure 5) :

- Devenir un référentiel commun à toutes les communautés d'utilisateurs (figure 6)
- Jouer le rôle central dans le programme national des grilles et clouds physiquement distribués utilisateurs (figure 7)
- Influencer et promouvoir les e-infrastructures au niveau international (figure 7)
- Valoriser les savoir-faire et développement (figure 8).

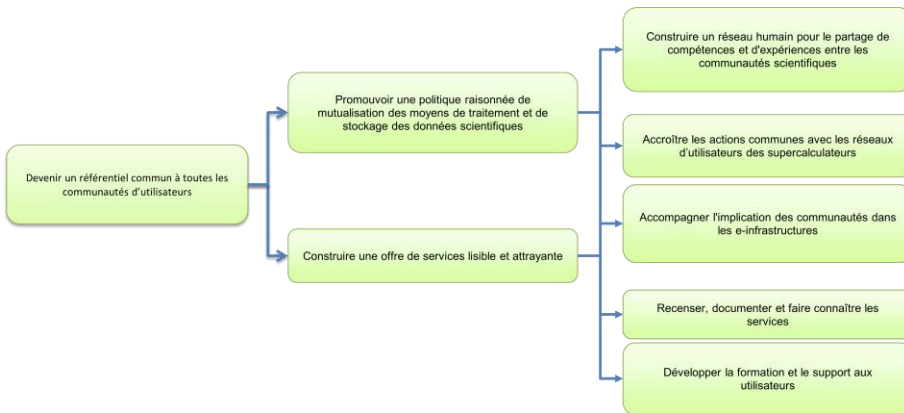


Figure 6 : stratégie vers les utilisateurs

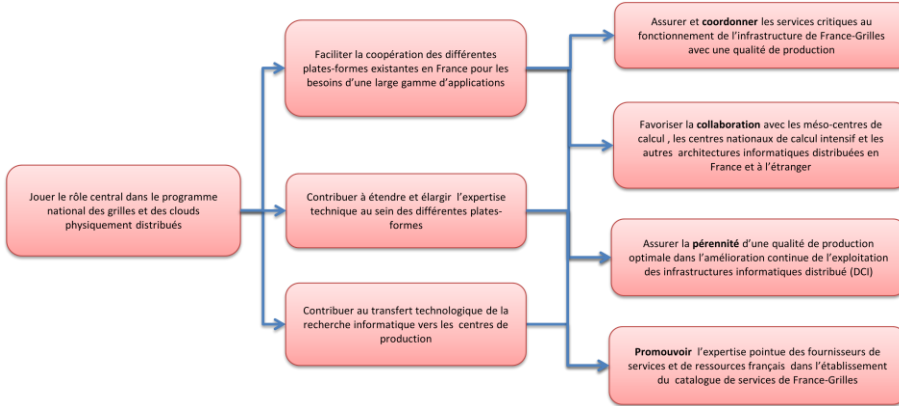


Figure 7 : stratégie vers les fournisseurs de service

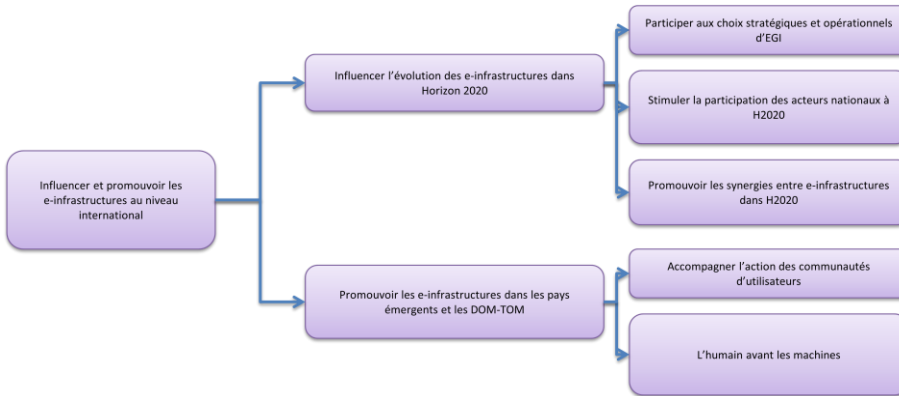


Figure 8 : stratégie internationale

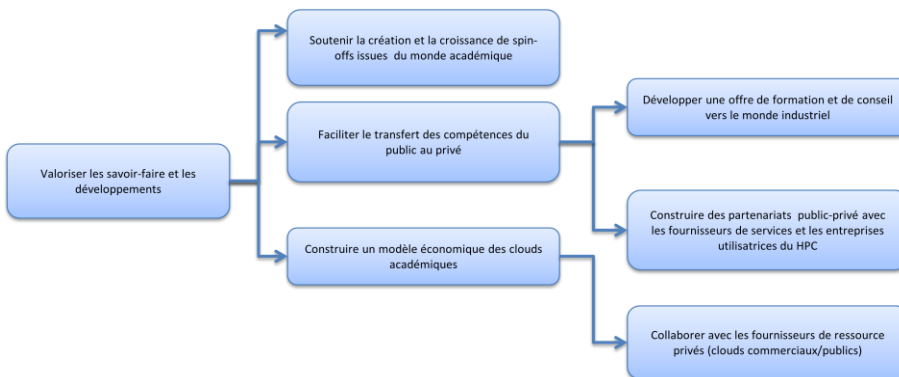


Figure 9 : stratégie pour valoriser les savoir-faire et les développements

4.2 STRATEGIE EN MATIERE DE CLOUD ET DE GESTION DE DONNEES SCIENTIFIQUES

Dans le domaine du cloud, France Grilles propose un cadre pour l'échange de compétences, d'outils et de savoir-faire entre la communauté de recherche en informatique et les communautés d'utilisateurs des ressources au niveau national. La création de la fédération de clouds académiques en 2013 a permis la mise en place d'un groupe de travail multidisciplinaire qui rassemble des ingénieurs impliqués dans le déploiement d'infrastructures clouds en France. La fédération n'a pas vocation à se substituer aux communautés pour faire des choix techniques mais plutôt à proposer un espace de dialogue entre les communautés pour mutualiser les compétences et identifier des synergies.

France Grilles a investi en 2011 et 2012 dans l'achat d'équipement pour le démarrage de clouds académiques à Toulouse et au CC-IN2P3 : leur mise en production est prévue en 2014. Deux autres sites, GRIF depuis 2012 et Strasbourg depuis Octobre 2013, sont dès à présent ouverts à recevoir de nouveaux utilisateurs.

Bien que le Big Data ne soit pas explicitement dans les missions du GIS France Grilles, offrir aux utilisateurs des services pour le traitement à haut débit des données est indissociable d'un service pour le stockage de ces données. C'est ainsi qu'une infrastructure distribuée de stockage de données scientifiques basée sur le service iRODS est proposée depuis l'automne 2013. Un programme de formation à iRODS est prévu en 2014.

4.3 LIENS AVEC RENATER ET GENCI

Les missions de France Grilles sont complémentaires de celles de RENATER et de GENCI:

- RENATER est l'un des huit partenaires du GIS France Grilles. RENATER assure notamment au sein de France Grilles la gestion de l'autorité de certification nationale qui délivre les certificats aux utilisateurs de la grille sur le territoire français. RENATER contribue aussi à la définition et à la mise en œuvre de la politique de sécurité de la grille et des clouds académiques
- Des liens se développent entre GENCI et France Grilles, notamment à travers les rencontres communes entre utilisateurs des grilles/clouds et des mésocentres.

5 CONCLUSION

La convention qui lie les partenaires du GIS France Grilles a été signée en Juin 2010 pour une durée de 4 ans. Le renouvellement du GIS France Grilles constitue une opportunité de confirmer le rôle du GIS dans le paysage national du calcul intensif pour l'animation et la coordination des activités sur la grille et le cloud mais aussi d'étendre ses missions à la gestion des données scientifiques. A l'ère du Big Data, la gestion des grands volumes de données scientifiques est en effet indissociable de leur traitement.

Le renouvellement du GIS s'inscrit aussi dans le contexte du démarrage d'Horizon 2020 où les laboratoires français doivent accroître leur visibilité, notamment dans le secteur des infrastructures de recherche et des e-infrastructures en particulier. Aux côtés de GENCI, le GIS France Grilles constitue un outil stratégique pour positionner la France dans les e-infrastructures de l'Espace Européen de la Recherche.