France Grilles : Un GIS en pleine évolution

Jérôme Pansanel <jerome.pansanel@iphc.cnrs.fr>

Journées Mathrice 2016

Sommaire

- France Grilles
- FG-DIRAC
- FG-iRODS
- FG-Cloud
- Conclusion

France Grilles

France Grilles

GIS – 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'Énergie 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
- Comité scientifique
- Comité technique
- Comité international

France Grilles, IR et NGI

Infrastructure de Recherche (IR)

Une infrastructure de recherche 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. »

National Grid Initiative (NGI)

L'Initiative Nationale de Grille française représente la France dans EGI

France Grilles

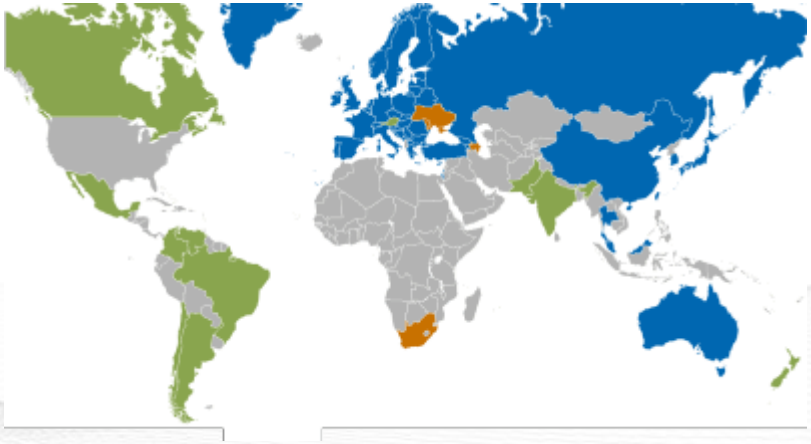
Objectifs

- Établir et opérer une infrastructure nationale de grille de production, pour le traitement et le stockage de données scientifiques massives
- Contribuer au fonctionnement de la fondation européenne EGI (EGI.eu)
- Favoriser rapprochements et échanges entre les grilles de production et les grilles de recherche
- Missions étendues au domaine du *Cloud Computing*
- Offre de services complète autour du traitement de données, du calcul et du stockage sur les infrastructures distribuées pour les utilisateurs, les organisations virtuelles (VO), les responsables de portail scientifique et les formations

France Grilles

EGI – `http://www.egi.eu`

- Fondation créé en 2010 basée à Amsterdam
- Opère des infrastructures distribuées à travers l'Europe et au-delà
- Développer et maintenir le *middleware*
- Soutien technique et formation
- Partenariat fort avec les autres projets européens (EUDAT, GEANT, PRACE, Elixir, ...)
- Surveillance fonctionnelle et *accounting* des sites certifiés
- Équipe de sécurité (CSIRT)
- Deux conférences annuelles



→ `https://indico.egi.eu/indico/event/2875/`

France Grilles

Mise en œuvre d'infrastructures distribuées

- Infrastructures virtuelles (grille et Cloud)
- Ensemble d’ordinateurs géographiquement distribués
- Travail en équipe distribué
- Fonctionnant en réseau
- Pour fournir une puissance globale
- Support utilisateur

Disposer à moindre coût

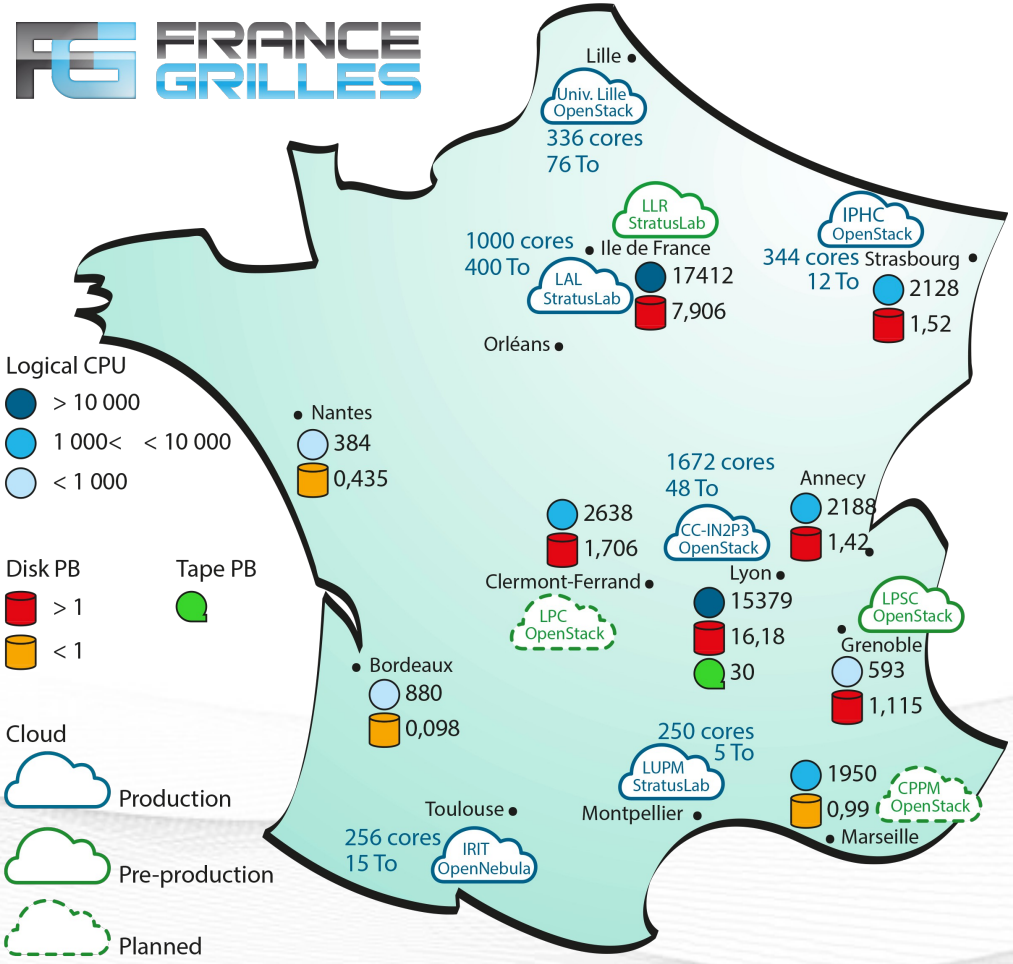
- Puissance de calcul
- Stockage de données
- Service pour l’analyse de données à haut débit (HTC)

France Grilles

Quelques chiffres

- **17** sites
- **31 900** CPUs logiques / **280 300** HEP-SPEC06
- **45,5 PB** / **30,0 PB** stockage disque / bande
- Disponibilité > 99 %

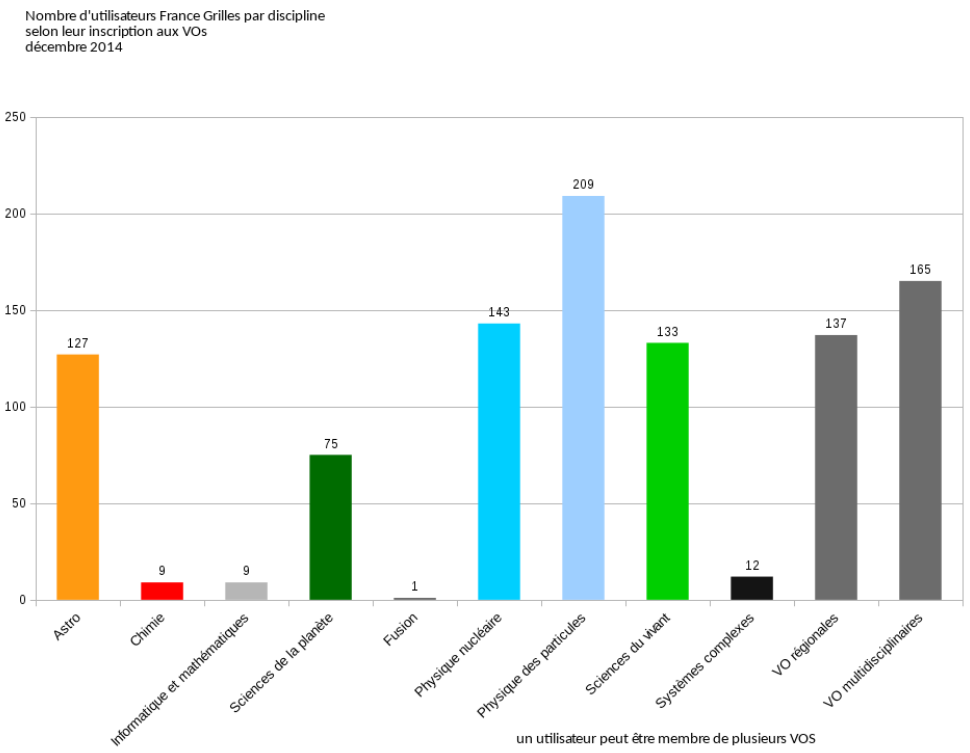
- Ressources mutualisées
- Apportées par les communautés
- Accessibles aux partenaires du GIS
- Besoins liés à la recherche
- Accessibles aux organismes ou entreprises dans le cadre de projets communs avec ces partenaires



France Grilles

Plus de 1000 utilisateurs

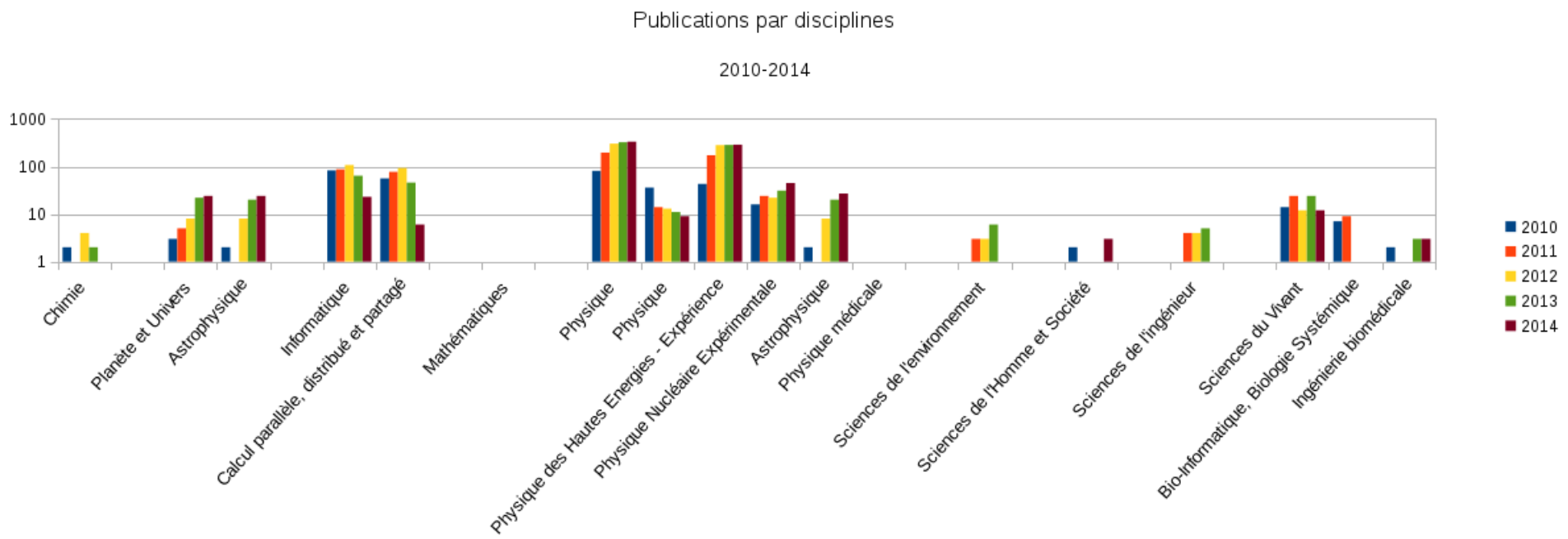
- Pas de sélection à l'entrée
- Issus de nombreuses disciplines
- L'organisation virtuelle France Grilles n°1 de la pluridisciplinarité, top 10 de toutes les organisations virtuelles d'EGI
- Portail VIP (CREATIS) pour l'imagerie médicale (+ 800 utilisateurs, +25%/an))
- Conseil/expertise vers les partenaires du GIS



France Grilles

Une forte production scientifique

- Plus de 2000 publications en 5 ans (HAL)
- Majoritairement en physique des particules
- Largement sous-estimée
- https://hal.archives-ouvertes.fr/FRANCE-GRILLES

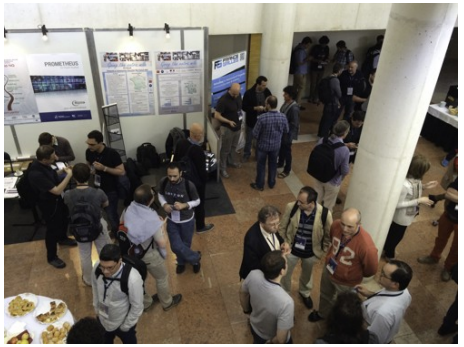


Relation avec les autres infrastructures

Intégré dans EGI

Collaboration technique

- RENATER
- CC-IN2P3
- LCG-France
- Mésocentres
- IDRIS



Travail d'animation avec le groupe calcul et GRID 5000

ASR Architecture Systèmes Réseaux — FRANCE GRILLES — Groupe CALCUL — Grid'5000

Journées SUCCES

Rencontres

Scientifiques des Utilisateurs de Calcul intensif, de Cloud Et de Stockage

5 et 6 novembre 2015, IPGP, Paris 5

France Grilles : offre de services

Pour les chercheurs et ingénieurs

- FG-DIRAC
- FG-iRODS
- FG-Cloud

Organisation virtuelle

- FG-DIRAC

Responsable de portail scientifique

- Certificat de service (*robot*)

Enseignement

- Formations

FG-DIRAC

FG-DIRAC

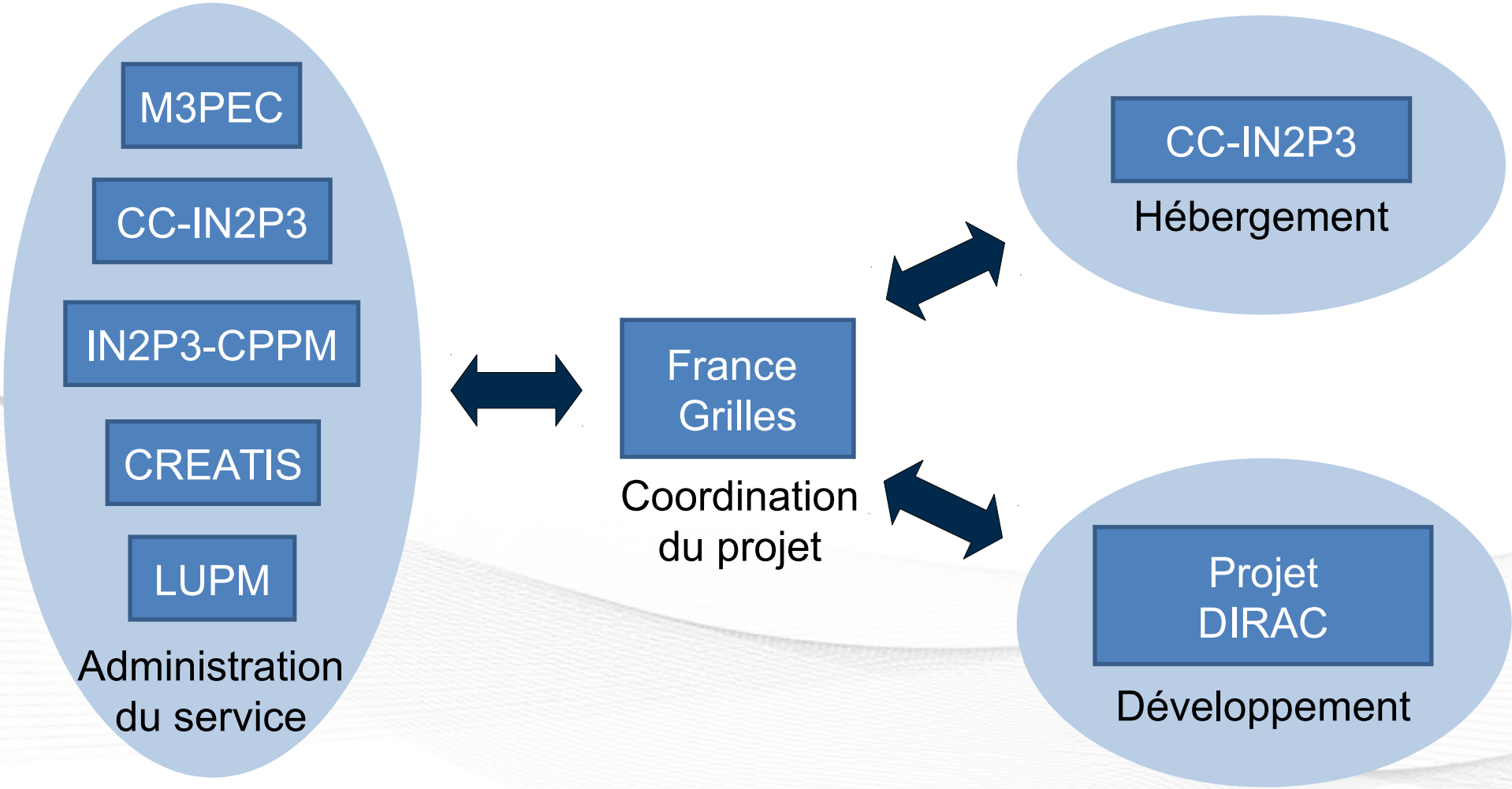
DIRAC en quelques mots

- Gestion des tâches de calcul et les données générées
- Utilisation des scénarios complexes pour le placement des fichiers
- Fonctionnalités accessibles à travers un portail Web
- Léger, robuste et extensible
- Utilisé par plusieurs VOs
- Intègre des ressources hétérogènes (grilles de calcul, Clouds, …)
- Faciliter aux petites expériences de bénéficier des infrastructures de production

L'instance FG-DIRAC

- Destiné aux VOs
- Hébergement au Centre de Calcul de l'IN2P3 à Lyon
- Utilisée par une vingtaine de VOs
- Administration collective par des administrateurs systèmes issus de 6 laboratoires français
- Expertise France Grilles
- `https://dirac.france-grilles.fr/DIRAC/`

FG-DIRAC


M3PEC
CC-IN2P3
IN2P3-CPPM
CREATIS
LUPM
Administration du service
France Grilles
Coordination du projet
CC-IN2P3
Hébergement
Projet DIRAC
Développement

FG-DIRAC

Réalisations

- Accueil de plus d'une dizaine de VOs
- Accord d'hébergement avec le CC-IN2P3
- Formations
- Communauté d'utilisateurs et d'administrateurs structurée
- Simplification des outils CLI (COMDIRAC)

Perspectives

- Accueil de nouvelles communautés d'utilisateurs
- Automatisation d'une partie des shifts avec Nagios
- Transfert technologique vers les partenaires européens

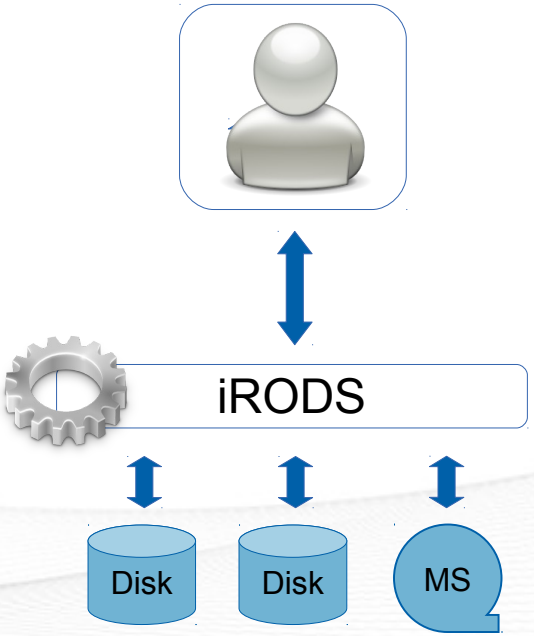
Workshop Utilisateurs DIRAC du 23 au 25 mai à Montpellier :
→ `https://indico.cern.ch/event/477578/`

FG-iRODS

FG-iRODS

iRODS

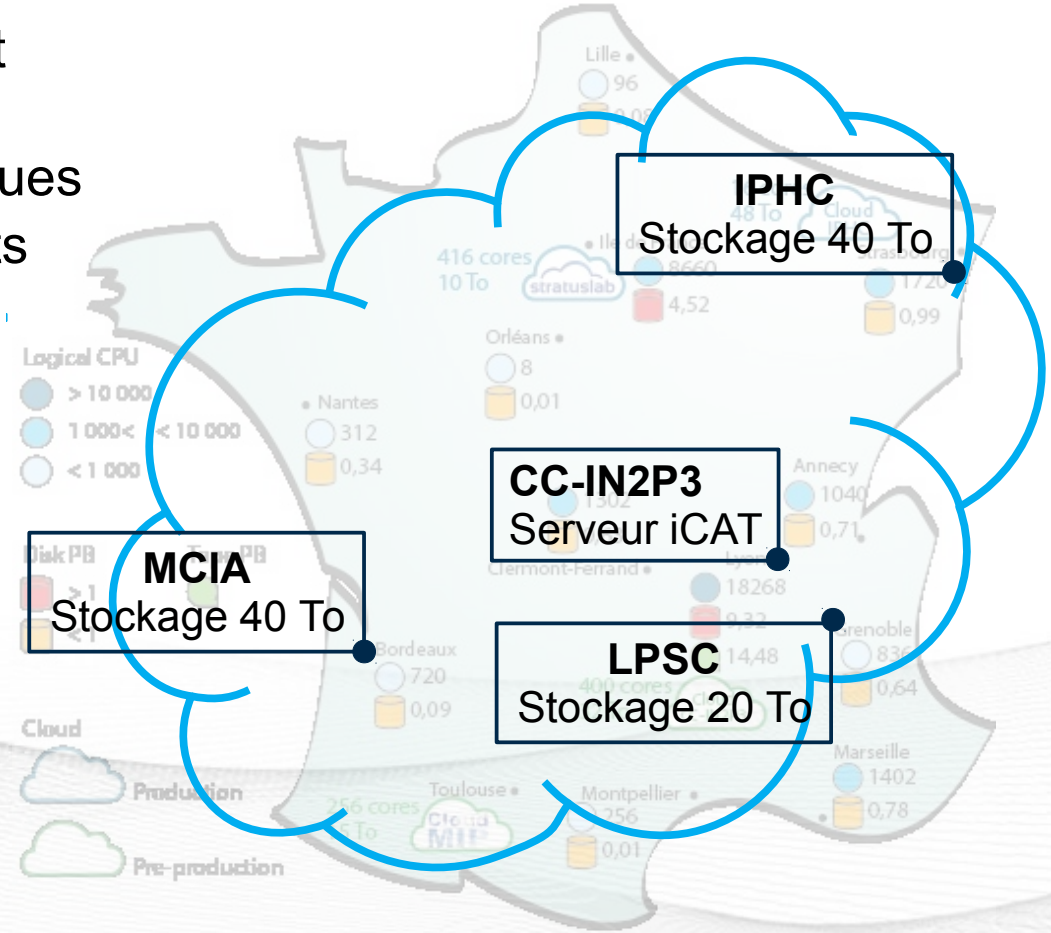
- Projet lancé en 2006 distribué sous licence libre (BSD)
- Développé et maintenu par le groupe DICE et de nombreux contributeurs, plusieurs milliers d'utilisateurs
- Moteur de règles interne exécutant des actions pré-définies
- Intègre un système de métadonnées descriptives
- Possibilité d'avoir les données sur des sites géographiquement répartis et sur du matériel hétérogène
- Organisation logique indépendante de l'implémentation matérielle
- Consistance et homogénéité des formats peuvent être contraintes
- Documentation complète
- `http://irods.org`



FG-iRODS

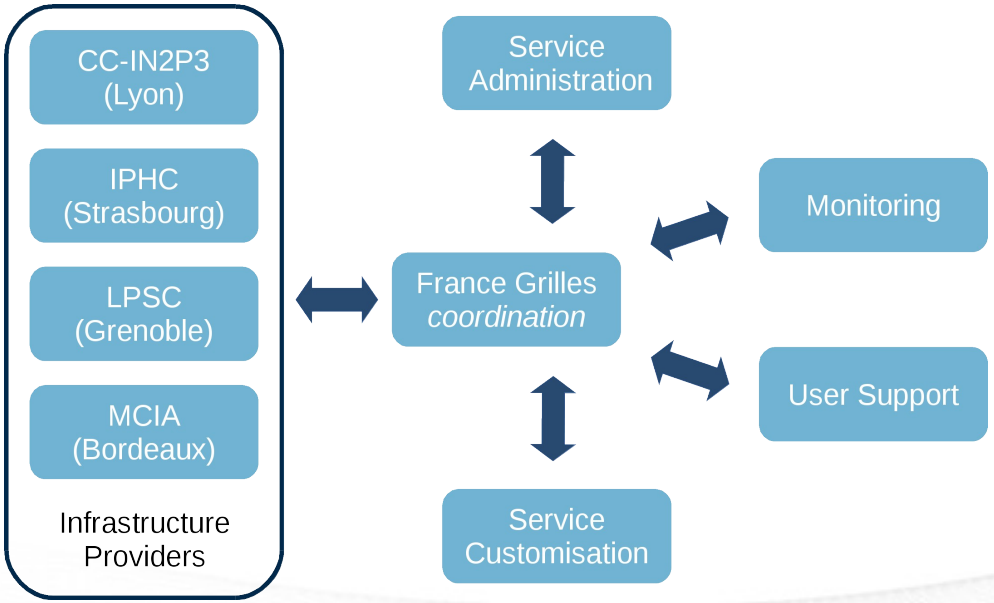
Une instance de production :

- Fédération de ressources humaines et matérielles
- Ouverte à tous les domaines scientifiques
- Taillée pour les petits et moyens projets
- Service de niveau production
- Support utilisateur et formation
- Investissement financier minimal



FG-iRODS

- Projet démarré en 2013 et piloté par l'équipe FG-iRODS de France Grilles
- Authentification par identifiants ou certificats
- Administré collectivement par les différents partenaires
- Moteur de règle et catalogue centralisé permettant de contraindre la cohérence et la gestion des données
- Ressources de stockage distribuées géographiquement
- Support utilisateur
- *Monitoring* et *accounting* du service



FG-iRODS : monitoring

Status Grid For All Host Groups

FG-iRODS iCAT (fg-irods-icat)

Host	Services	Actions
ccirods.in2p3.fr	org.irods.irods3.Icat-Connection	

FG-iRODS Resources (fg-irods-resources)

Host	Services				Actions
fgirods0.mcia.univ-bordeaux.fr	org.irods.irods3.Resource-All	org.irods.irods3.Resource-Iget	org.irods.irods3.Resource-Iput	org.irods.irods3.Resource-Irm	
lpsc-datagrid2.in2p3.fr	org.irods.irods3.Resource-All	org.irods.irods3.Resource-Iget	org.irods.irods3.Resource-Iput	org.irods.irods3.Resource-Irm	
sbgse25.in2p3.fr	org.irods.irods3.Resource-All	org.irods.irods3.Resource-Iget	org.irods.irods3.Resource-Iput	org.irods.irods3.Resource-Irm	

Service Status Details For All Hosts

Limit Results: 100

Host	Service		Status	Last Check	Durat
ccirods.in2p3.fr	org.irods.irods3.Icat-Connection		OK	04-29-2015 12:05:46	2d 23h
fgirods0.mcia.univ-bordeaux.fr	org.irods.irods3.Resource-All		OK	04-29-2015 11:29:49	2d 23h
	org.irods.irods3.Resource-Iget	?	OK	04-29-2015 11:29:50	2d 23h
	org.irods.irods3.Resource-Iput	?	OK	04-29-2015 11:29:49	2d 23h
	org.irods.irods3.Resource-Irm	?	OK	04-29-2015 11:29:50	2d 23h
lpsc-datagrid2.in2p3.fr	org.irods.irods3.Resource-All		OK	04-29-2015 11:30:01	2d 23h
	org.irods.irods3.Resource-Iget	?	OK	04-29-2015 11:30:02	2d 23h
	org.irods.irods3.Resource-Iput	?	OK	04-29-2015 11:30:01	2d 23h
	org.irods.irods3.Resource-Irm	?	OK	04-29-2015 11:30:03	2d 23h
sbgse25.in2p3.fr	org.irods.irods3.Resource-All		OK	04-29-2015 11:29:39	2d 23h
	org.irods.irods3.Resource-Iget	?	OK	04-29-2015 11:29:40	2d 23h
	org.irods.irods3.Resource-Iput	?	OK	04-29-2015 11:29:39	2d 23h
	org.irods.irods3.Resource-Irm	?	OK	04-29-2015 11:29:40	2d 23h

FG-iRODS : exemples d'utilisation

Réplication des données hors site

- 2 To actuellement utilisés
- Accessible à un nombre limité d'utilisateurs

Stockage de données massives

- Hébergement de 1 Po d'ici 5 ans
- Réplication partielle entre les sites
- Transfert de compétences concernant l'administration iRODS
- Gestion fine des autorisations et des accès
- Scénario flexible pour l'hébergement

FG-iRODS

Réalisation

- Ressources fédérées de 100 To
- Synergie entre les administrateurs
- Stockage disponible et sécurisé
- Règle d'utilisations publiées
- Formations utilisateur (2014, 2015 et 2016)
- Formation administrateur (2016)
- Client iRODS déployé sur tous les sites de grille supportant la VO France Grilles
- Distribution des clients iRODS avec le support GSI (deb, rpm)
- Interface Web
- Machine virtuelle permettant d'accéder à iRODS (et également DIRAC, Grille et Cloud)

FG-iRODS

Perspectives

- Étendre la capacité disponible avec des ressources internes et de nouveaux partenaires
- Augmenter l'utilisation des ressources et en particulier les services disponibles
- Tester le plugin S3
- Migration vers la version 4
- Assurer la pérennité de l'infrastructure
- Partager l'expertise acquise avec d'autres groupes d'administrateurs / utilisateurs, que ce soit à propos de la partie technique d'iRODS ou dans la gestion des données scientifiques.

→ `http://www.france-grilles.fr/Pour-les-chercheurs-ou-ingenieurs#iRODS`

FG-Cloud

FG-Cloud

Stratégie

- Construction et opération d'un *Cloud* pour l'analyse des données scientifiques
- Basé sur le volontariat des sites participants
- Accès aux ressources ouvert : toutes disciplines
- Intégré dans une infrastructure internationale
- Pas d'a-priori sur les technologies Cloud, suffisamment
- souple pour être compatible avec les stratégies locales
- Représente la France dans EGI (NGI)

Objectifs

- Établir une infrastructure nationale de *Cloud* de production
- Promouvoir son usage dans toutes les communautés scientifiques
- Créer une communauté d'experts *Cloud*
- Valoriser cette expertise : publications, interventions
- Mutualiser les expériences, les compétences
- Diffuser la connaissance : documentation, formation
- Facilitateur pour les relations entre les communautés française et internationales

FG-Cloud : Organisation

Groupe technique

- Deux réunions / mois
- Évolutions des sites
- Suivi des utilisateurs
- Développements en cours

Communication

- Liste de diffusion
- Canal IRC

Mutualisation

- Expertise partagée
- Développements
- Utilisateurs des Clouds, administrateurs de site et développeurs de solution
- Intégration avec le Cloud européen EGI (technique, utilisateurs)

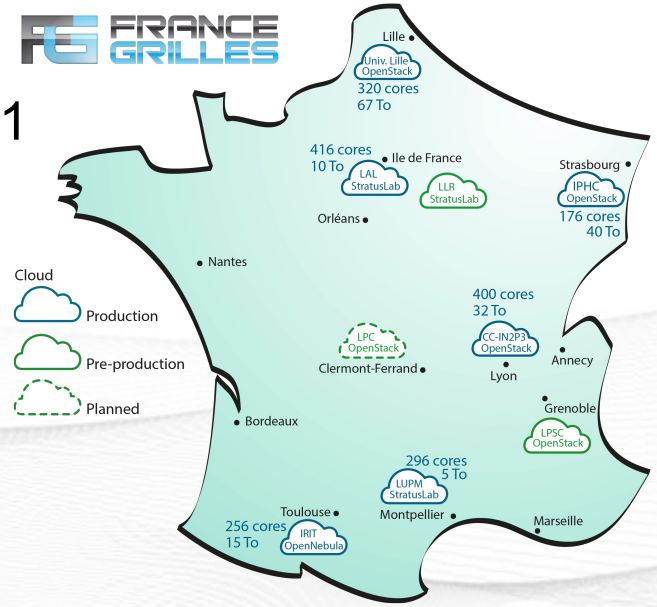
FG-Cloud

L'infrastructure

- Type IaaS : instanciation de machines virtuelles
- Ressources distribuées : calcul et stockage
- Extensible : Intégration progressive de nouveaux sites
- Interface homogène d'accès à des ressources hétérogènes
- Intégration facilitée : pas de contraintes fortes imposées aux sites

Fournisseurs de ressources

- Production : CC-IN2P3, LAL, IPHC, LUPM, Univ. Lille1
- Pré-production : LPSC, IRIT
- En cours d'installation : LPC, CPPM, LLR



Quelques chiffres

- 3882 cœurs de CPU
- 14 TB de RAM
- 560 TB de stockage disque

FG-Cloud

Les technologies

- OpenStack – *Cloud middleware*
- StratusLab – *Cloud middleware*
- SlipStream – Déploiement d'applications (PaaS)
- DIRAC – Utilisation des ressources Cloud
- Qemu / KVM / Libvirt – Couches de virtualisation
- OpenVSwitch – Virtualisation de réseaux (SDN)
- Linux (CentOS 7)

FG-Cloud

Accueil des nouveaux sites

- Aide à la mise en place logicielle
- Conseils sur l'architecture matérielle et les performances
- Réunions de suivi
- Transfert d'expertise aux nouveaux venus
- Intégration progressive en trois phases :
  - Test
  - Pré-production
  - Production (maintenance, surveillance, évolution)
  - Intégration au Cloud européen EGI

FG-Cloud

Cloud Challenge

- Test grandeur nature d'utilisation massive (courant 2015)
- Utilisateur expert sélectionné après candidatures
- Simulation MC en chimie quantique
- Équipes des Univ. Paul Sabatier & IRIT à Toulouse
- Résultats présentés aux JDEV(1) et SUCCES(2)
- Validation de la pertinence du *Cloud*
- Estimation du niveau de performance / *bare-metal*
- Certains calculs sont dépendants de fonctionnalités CPU précises
- La disponibilité peut aider à compenser

IRIT CNRS - INPT - UPS - UT1 - UT2J
UNIVERSITÉ TOULOUSE III PAUL SABATIER
JDEV2015

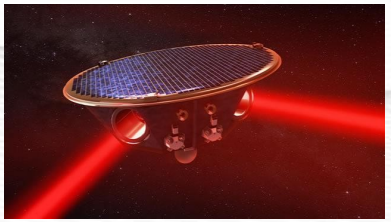
(1) http://devlog.cnrs.fr/jdev2015 (2) http://succes2015.sciencesconf.org

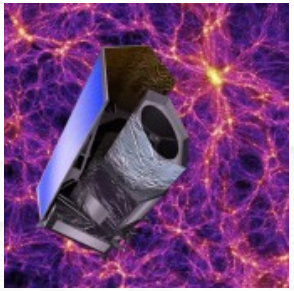
FG-Cloud

Autres exemples d'utilisation

- Astrophysique, données des missions spatiales pour les futurs projets Euclid et eLisa
- Cluster Torque / Maui
- Cluster Hadoop
- Outils pour les formations (SGBD, OpenStack, DIRAC, iRODS, ...)
- Analyse de réseaux sociaux
- Reconstruction et analyse de séquences génomiques
- Hébergement de portails scientifiques
- Calcul Monte-Carlo en physique des particules
- ...

UMR 5175
CENTRE D'ECOLOGIE FONCTIONNELLE & EVOLUTIVE





FG-Cloud

Conclusion

- La phase de production est atteinte
- Le support utilisateur est opérationnel
- Marketplace centralisé au LAL
- Tremplin pour l'utilisation du Cloud EGI
- Possibilité d'héberger / mutualiser

Perspectives

- Extension de la fédération
- Intégration progressive dans EGI
- Accueil de nouveaux utilisateurs
- Amélioration continue des procédures de sécurité
- Développement de nouveaux outils (supervision, *accounting*, gestion des VMs)
- Tests d'intégration docker (magnum)

Formation Utilisateur FG-Cloud du 27 au 29 avril à Lyon :
→ `https://indico.in2p3.fr/event/12720/`

Conclusion

Accueil des nouveaux utilisateurs

- Premier contact : info@france-grilles.fr
- Expression / étude des besoins
- Aide à la prise en main du service
- Tutoriels planifiés : http://www.france-grilles.fr/-Formation-
- Documentation collaborative pour les utilisateurs (wiki)
- Aide personnalisée pour adapter le service aux besoins

Merci de votre attention !