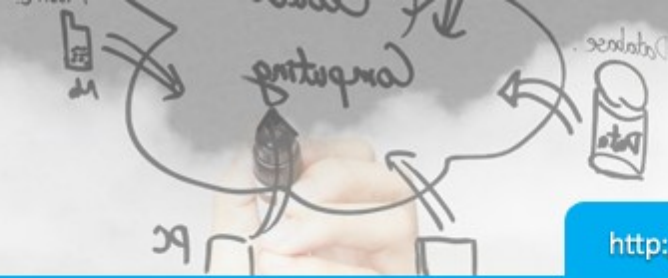




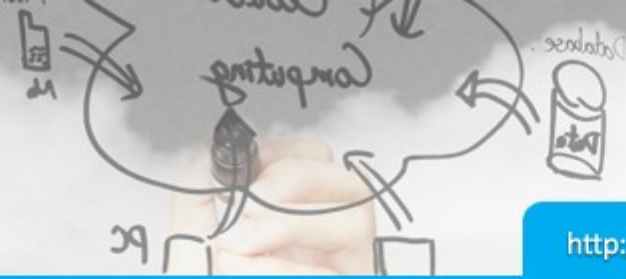
# France Grilles : Infrastructures distribuées pour le calcul scientifique et le stockage de données

**Jérôme PANSANEL** <[jerome.pansanel@iphc.cnrs.fr](mailto:jerome.pansanel@iphc.cnrs.fr)>  
Directeur technique



## Sommaire

- France Grilles : historique et missions
- Infrastructure de grille
- Service DIRAC
- FG-iRODS
- FG-Cloud



# France Grilles : historique et missions

## 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'Énergie Atomique et aux Énergies Alternatives
  - Institut National de la Recherche Agronomique
  - Institut National de la 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, ...)

## Les missions

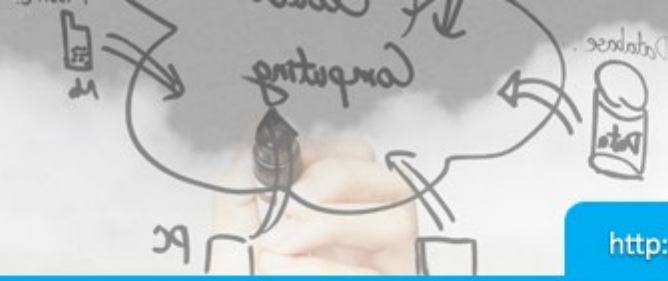
- Établir et opérer une infrastructure nationale de grilles de production, pour le traitement et le stockage de données scientifiques massives.
- Contribuer au fonctionnement de l'infrastructure européenne EGI (European Grid Infrastructure).
- Favoriser rapprochements et échanges entre les grilles de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

=> Missions étendues au domaine du *Cloud Computing*

- Une Infrastructure de Recherche (IR) du Ministère de l'Enseignement Supérieur et de la Recherche (2009) pilotée par le GIS France Grilles
- L'Initiative Nationale de Grille française - « National Grid Initiative » (NGI) représente la France dans EGI (<http://www.egi.eu>)

## France Grilles : Offre de service

- Pour les chercheurs et ingénieurs
  - FG-DIRAC
  - FG-iRODS
  - Infrastructure de production Grille
  - Infrastructure de production Cloud
- Organisation virtuelle
  - FG-DIRAC
- Responsable de portail scientifique
  - Certificat de service (*robot*)
- Enseignement
  - Formations



# Infrastructure de grille

## La grille de calcul

Les *grilles informatiques* sont des infrastructures virtuelles :

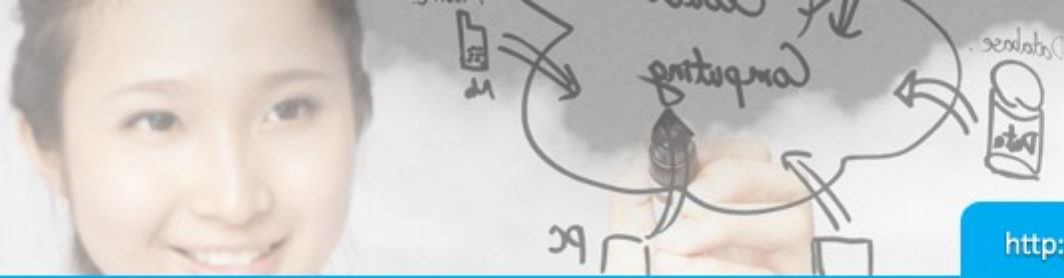
- un ensemble de centres de calcul sur des sites géographiquement distribués
- fonctionnant en réseau
- pour fournir une puissance globale

Permet de disposer à moindre coût :

- de puissance de calcul
- de stockage de données
- d'une infrastructure pour l'analyse de données à haut débit

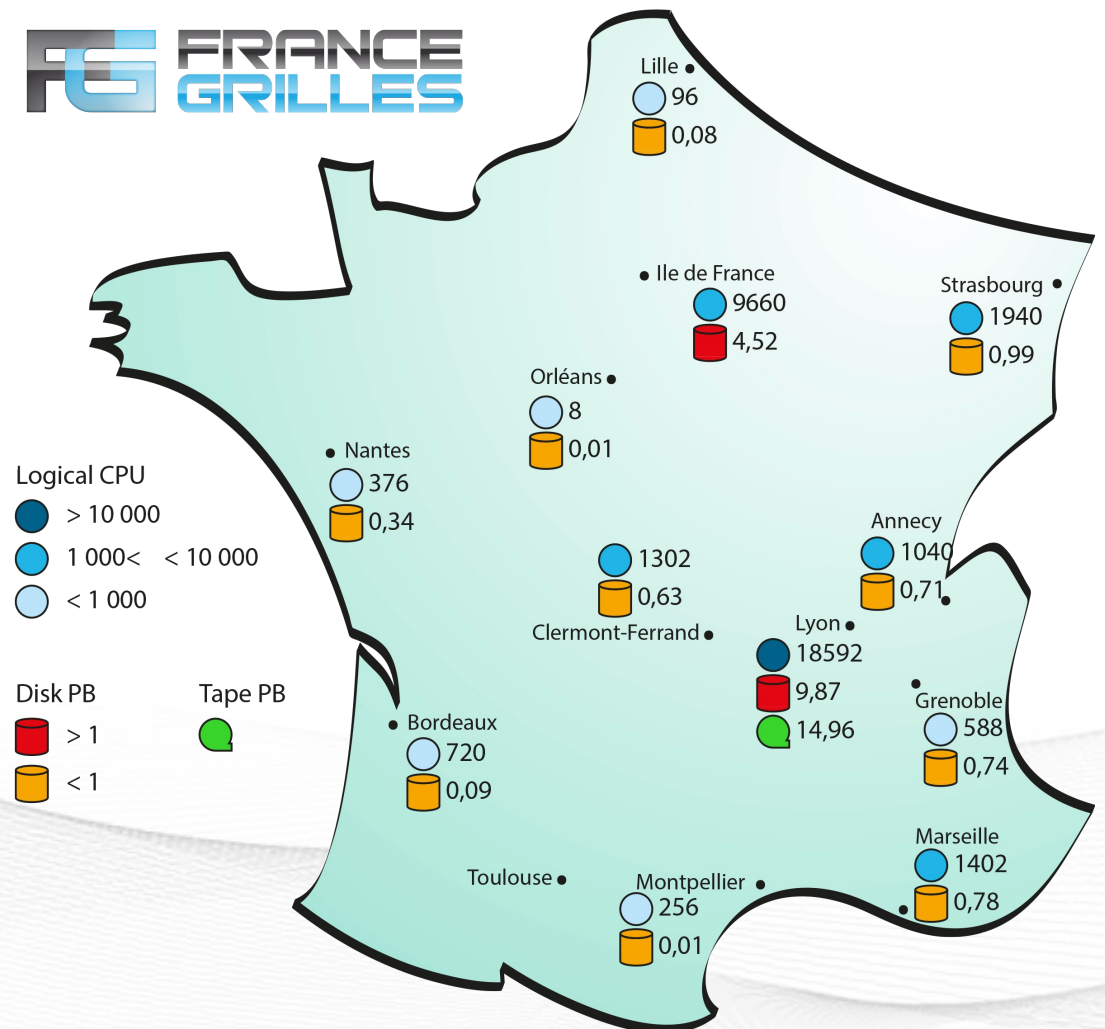
Utilisateurs organisés en VO (Virtual Organisation)

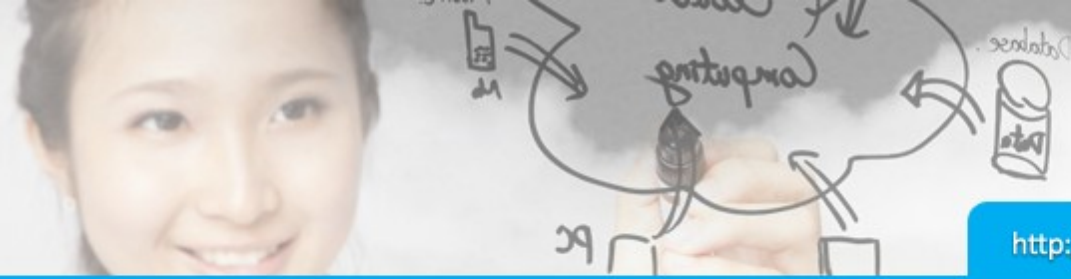




# L'infrastructure française

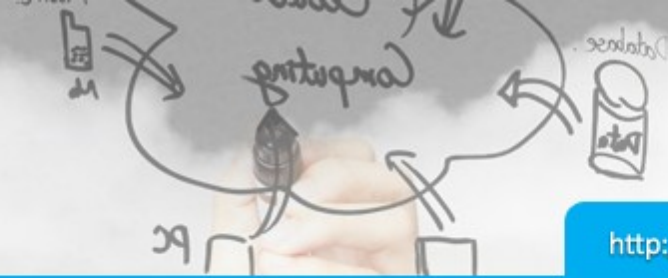
- 16 sites en France
- 39352 CPUs logiques / 396331 HEPSEC06
- 20,52 Po / 14,96 Po stockage sur disque / bande
- ressources mutualisées
- apportées par les communautés
- fiabilité et disponibilité > 99 %





## Utilisation de la grille

- Environ 700 utilisateurs nationaux regroupés en organisations virtuelles (VO)
- Plus de 90 VO supportées (tous les domaines scientifiques sont couverts)
- Accessibles aux partenaires du GIS
- Besoins liés à la recherche scientifique
- Accessibles aux autres organismes ou entreprises dans le cadre de projets communs avec partenaires.
- Plus de 350 publications par an



# Plate-forme DIRAC

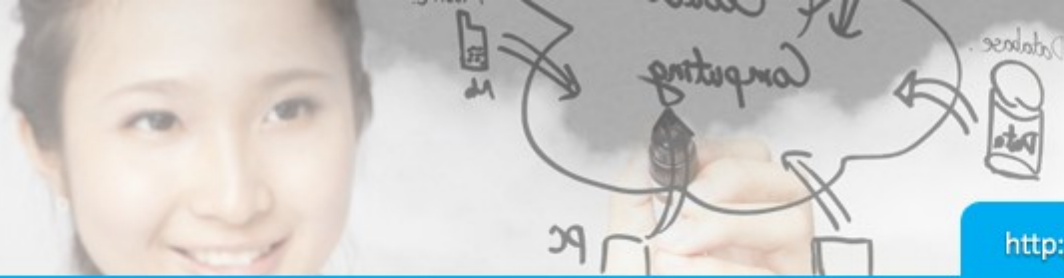
## La plate-forme FG-DIRAC

DIRAC est une solution générale et complète pour une ou plusieurs communautés d'utilisateurs et qui :



- permet de gérer les tâches de calcul et les données générées
- propose d'appliquer des scénarios complexes pour le placement des fichiers
- rend accessible toutes les fonctionnalités à travers un portail Web
- est léger, robuste et extensible
- prend en compte des ressources hétérogènes (grilles de calcul, Clouds, ...)
- est utilisé par plusieurs VO
- permet aux petites expériences de bénéficier facilement des infrastructures de production

L'instance FG-DIRAC :

- hébergement au Centre de Calcul de l'IN2P3 à Lyon
- utilisée par une vingtaine de VO
- administration collective par des administrateurs systèmes issus de 6 laboratoires français




# Interface utilisateur

System ▾ Jobs ▾ Views ▾ Virtual machines ▾ Selected setup: Dirac-Production ▾  

**JobMonitoring** <<  Select All  Select None ✕ Kill ✕ Delete

JobId ▾	Status	MinorStatus	ApplicationStatus	Site	JobName	LastUpdate [UTC]	LastSignOfLife [...]	SubmissionTim...	Owner

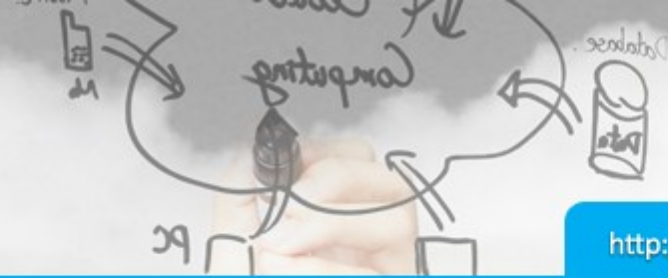
Site: All ▾  
 Status: All ▾  
 Minor status: All ▾  
 Application status: All ▾  
 Owner: Nothing to display ▾  
 JobGroup: All ▾  
 JobType: All ▾

Submit  Reset 

Global Sort + Selected Statistics + Global Statistics +

Page 1 of 1 Refresh Updated: - Items per page: 25 ▾

Jobs > Job monitor DIRAC: v6r12 Anonymous (/O=GRID-FR/C=FR/O=CNRS/OU=IPHC/CN=Jerome Pansanel)

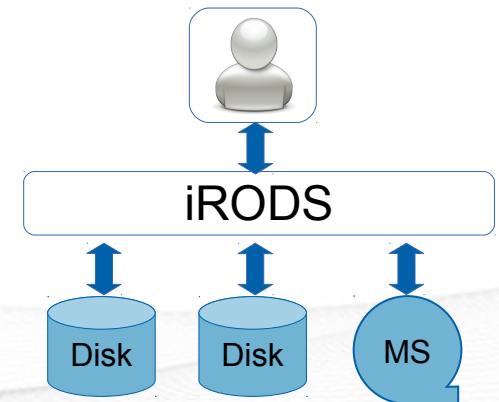


# FG-iRODS

## FG-iRODS : En quelques mots

- Projet démarré en 2006 (basé sur SRB)
- Distribué avec une licence Open Source (BSD)
- Développé par le groupe DICE et plusieurs contributeurs
- Un moteur de règle (*Rule Engine*) applique des actions pré-définies dans des politiques de gestion de données
- Intègre un système de méta-données
- Collections de données pouvant être distribuées sur plusieurs sites et du matériel hétérogène
- Organisation logique des fichiers indépendante de l'organisation physique
- Vérification de la consistance et de l'homogénéité des données

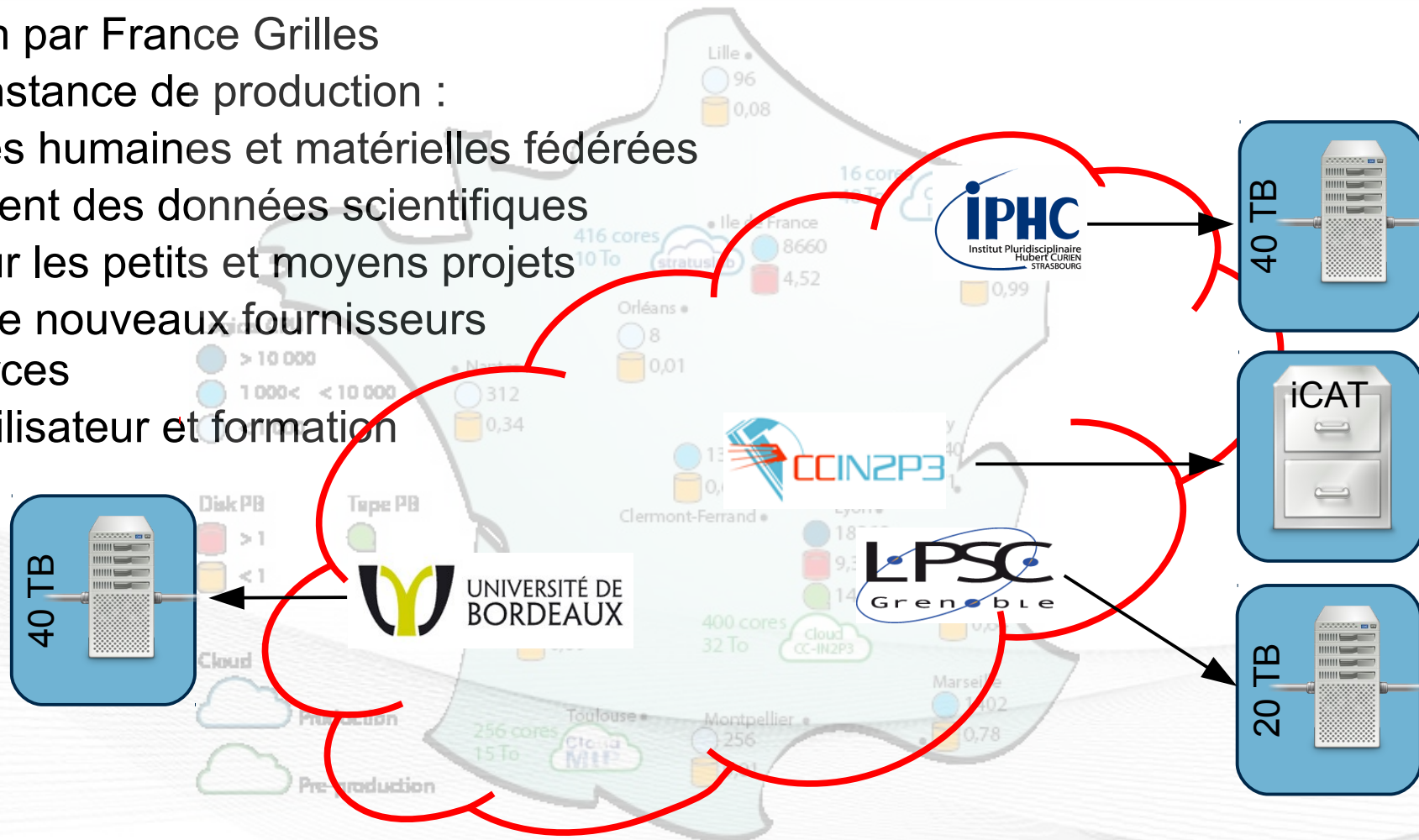
# iRODS



<http://irods.org/>

## FG-iRODS : Aperçu des sites

- Coordination par France Grilles
- Une seule instance de production :
  - Ressources humaines et matérielles fédérées
  - Hébergement des données scientifiques
  - Conçu pour les petits et moyens projets
  - Ouvert à de nouveaux fournisseurs de ressources
  - Support utilisateur et formation

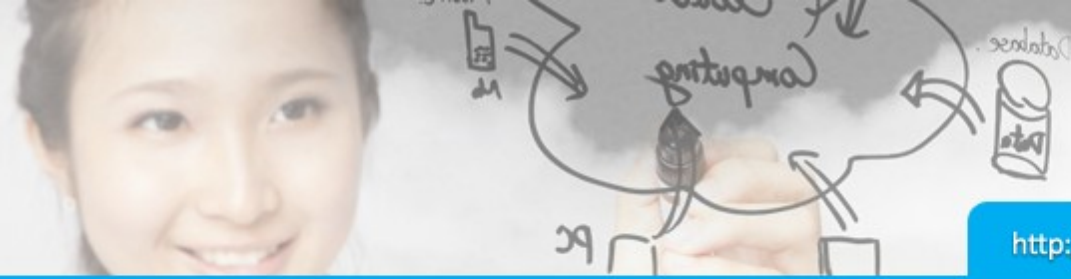


- Contact:  
irods-support-1@france-grilles.fr



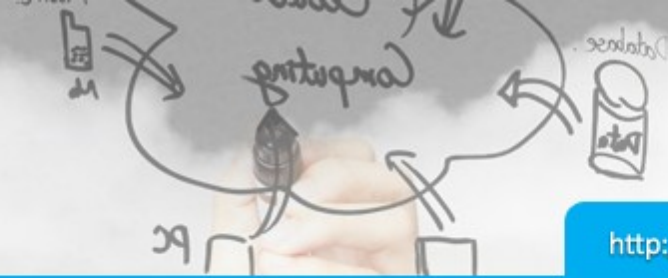
## FG-iRODS : Réalisations

- 100 To de stockage fédéré disponible
- Synergie entre les différents administrateurs
- Stockage disponible et fiable
- Conditions d'utilisation pour les utilisateurs définies
  
- Formation iRODS et sur l'infrastructure France Grilles réalisée en 2014 à Clermont-Ferrand
- Infrastructure utilisée par deux groupes d'utilisateurs (protéomique et données biologiques)
- Interface Web disponible
  
- Client iRODS installé sur tous les sites supportant la VO France Grilles
- Mise à disposition des paquets iRODS avec GSI activé (deb, rpm)
- Mise à disposition d'une VM appliance pré-configurée pour l'utilisation de la grille et iRODS



## FG-iRODS : Perspectives

- Étendre la capacité de stockage avec l'accueil de nouveaux fournisseurs de ressources
- Trouver des ressources de financement pour assurer le fonctionnement pérenne de l'infrastructure
- Déployer le système de surveillance fonctionnelle
- Enrichir le catalogue de service en fonction des premiers retours d'expérience
- Tester le plugin S3
  
- Accueillir de nouveaux utilisateurs
- Partager l'expertise sur la gestion de données et le support utilisateur avec d'autres groupes



# FG-CLOUD

## FG-CLOUD : Le projet

### • Objectifs

- Mettre en place un service de Cloud fédéré IaaS pour la recherche scientifique
- Définir la stratégie et construire l'infrastructure nationale
- Rédiger les conditions d'utilisation, la documentation
- Définir les services aux utilisateurs

### • Obstacles

- Co-existence avec des initiatives existantes (approche *bottom-up*)
- La stratégie France Grilles doit être compatible avec celle des parties tierces (sites, tutelles et partenaires)
- Fonctionnement avec les différents logiciels de déploiement de Cloud
- Les ressources doivent être accessibles simplement
- Démontrer les bénéfices d'un Cloud fédéré

## FG-CLOUD : Pilotage technique

- **Le groupe technique**

- Membres : administrateurs de site, développeurs et utilisateurs experts
- Réunion bimensuelle
- Mission clairement établie
- Contact : `cloud-tech-1@france-grilles.fr`

- **Réalisations**

- Authentification et *Marketplace* communs aux différents sites
- VM-DIRAC et SlipStream sont disponibles pour accéder aux différentes ressources (OpenStack, StratusLab)
- Financement des sites (infrastructure, EGI)
- Règles d'utilisation sont disponibles

- **En cours**

- Développement en cours pour le support de OpenNebula dans SlipStream
- Tests en cours de VM-DIRAC avec OpenNebula et SlipStream
- Développement des tests fonctionnels Nagios

## FG-CLOUD : Aperçu des sites

### • Sites en production

- OpenStack@CC-IN2P3 (\*)
- CloudMip@IRIT
- Stratuslab@LAL (\*)

### • Sites en pre-production

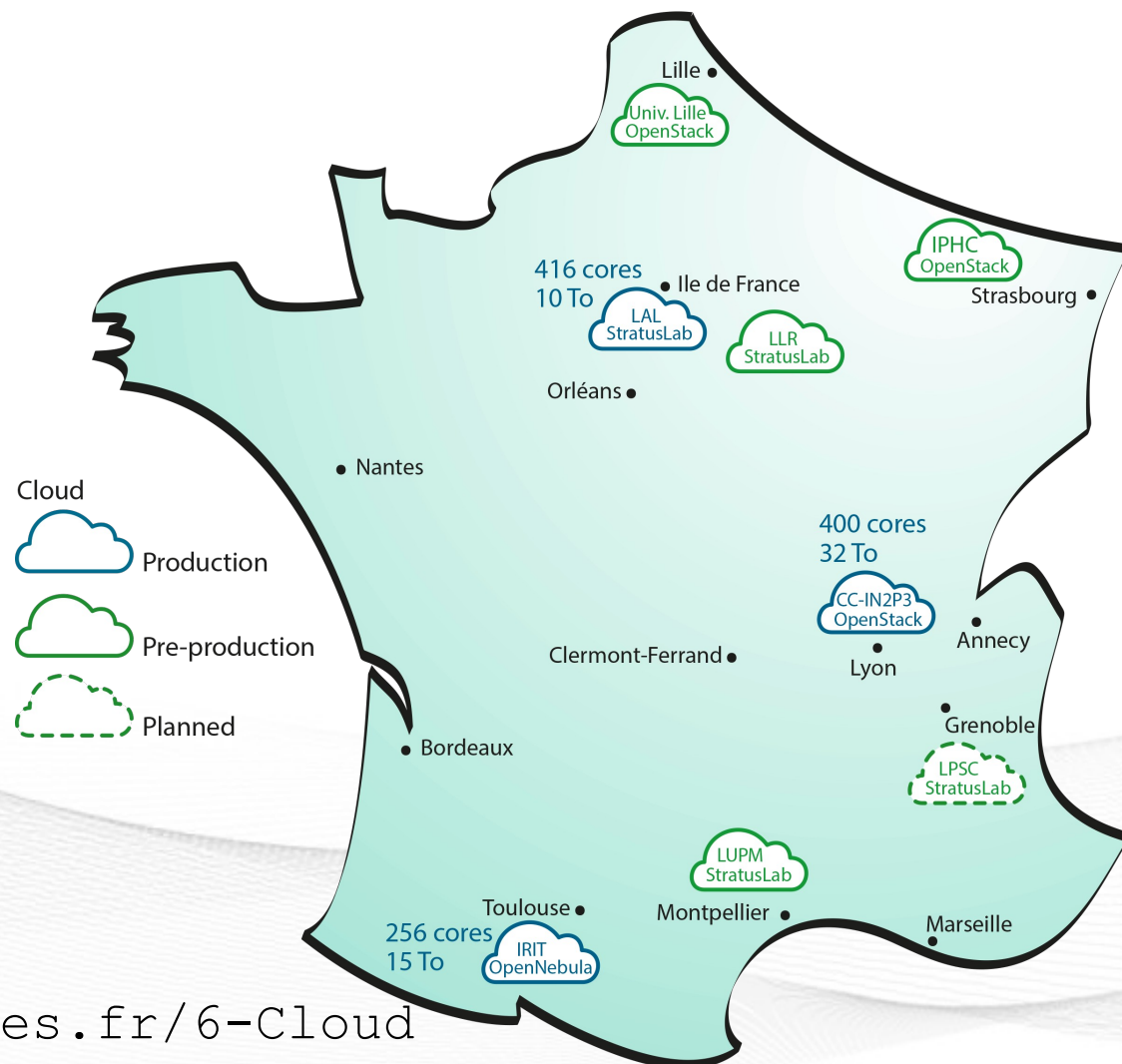
- OpenStack@IPHC (\*)
- StratusLab@LLR (\*)
- StratusLab@LUPM
- OpenStack@UNIV-LILLE

### • En cours d'installation

- StratusLab@LPSC (\*)

### • Plus d'information

- <http://www.france-grilles.fr/6-Cloud>



## FG-CLOUD : Projets en cours & perspectives

### • Projets en cours

- Lancement du Cloud Challenge (utilisation de l'intégralité des ressources par un ou deux projets scientifiques)
- Vérification que le service fonctionne en mode de production (redondance des services, surveillance, catalogue de services, ...)
- Extension des ressources disponibles et de la communauté d'utilisateurs

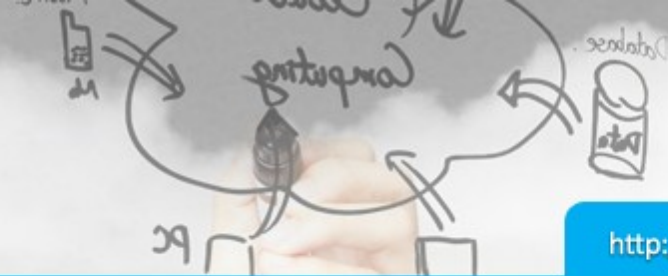
### • Perspectives

- Valider VM-DIRAC avec SlipStream et OpenNebula
- Développer le *backend* rOCCI pour StratusLab (EGI)
- Lancer le Challenge Cloud (novembre 2014)
- Déployer la solution de surveillance et mettre en place les procédures de fiabilité et de disponibilité
- Proposer l'expertise acquise en déploiement de Cloud à d'autres groupes

## FG-CLOUD : Recrutement

- Profil
  - Niveau bac+2 en informatique, avec une première expérience dans le développement d'applications en Python ou dans le domaine de la virtualisation (kvm, libvirt).
  - Familiarisé avec les logiciels libres
  - Bonne maîtrise de la langue anglaise
- Compétences techniques
  - Python
  - GNU/Linux
  - Git
  - Libvirt
  - KVM
  - OpenStack ou StratusLab





**Merci de votre attention !**