

# Equipe Calcul et Données

*A. Tsaregorodtsev*

*C. Carranza*

*E. Knoops*

*A. Lytovchenko*

*C. Meessen*

- ▶ **Présentation du groupe**
- ▶ **Projets**
  - ▶ Développement
  - ▶ Services
- ▶ **Conclusions et perspectives**

# Composition du groupe

---

- ▶ 4 ingénieurs CNRS permanents (temps partiel)
  - ▶ Andrei Tsaregorodtsev, IRHC, responsable scientifique
  - ▶ Edith Knoops, IR I
  - ▶ Carlos Carranza, IR I
  - ▶ Christophe Meessen, IR I
- ▶ 1 ingénieur CDD
  - ▶ A.Lytovchenko, IE

# Responsabilités

---

## ▶ Au niveau international

- ▶ Responsable scientifique du service EGI Workload Manager du projet EOSC-Hub : A.T.
- ▶ Coordinateur de DIRAC Consortium : A.T.

## ▶ Au niveau national

- ▶ Coordinateur du service France-Grilles DIRAC : A.T.
- ▶ Responsable scientifique du Master Project DIRAC@IN2P3 :A.T.

## ▶ Au niveau CPPM

- ▶ Responsable du nœud Cloud : C.Carranza

NB: non-inclus des responsabilités WLCG, T2/CPPM

# Sources de financement

---

- ▶ Horizon 2020, EOSC-Hub

- ▶ ~44 PM
- ▶ A.Lytovchenko à partir du septembre 2018
- ▶ Il reste la possibilité de financer ~16 mois d'ingénieur CDD avant la fin 2020 :
  - ▶ EGI Workload Manager, Eiscat 3D Collaboration
  - ▶ L'offre du poste est en cours de préparation

- ▶ Horizon 2020, ESCAPE

- ▶ 1 FTE à partir de 2019 pour 3 ans, plate-forme DIRAC pour des application Km3NET
  - ▶ L'offre du poste est en cours de préparation

- ▶ CPPM

- ▶ ?

# Activités du groupe

---

- ▶ Développement DIRAC
- ▶ Services DIRAC
- ▶ Cloud CPPM
- ▶ Support aux applications

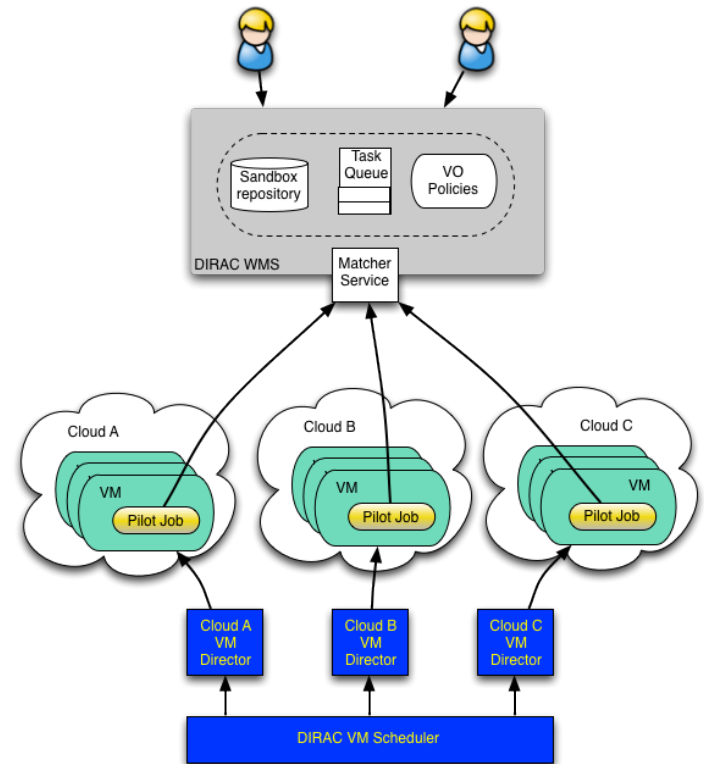
# Développement DIRAC

---

- ▶ Cloud interface
  - ▶ A.T.
- ▶ Cloud metascheduler
  - ▶ A.T.
- ▶ High Level workflows
  - ▶ DIRAC@IN2P3 Project
- ▶ Data Management
  - ▶ A.T., DIRAC@IN2P3 Project
- ▶ DIRAC Client/service advanced framework
  - ▶ C.Meessen
- ▶ SSO/OIDC authentication/autorisation
  - ▶ A.Lytovchenko

# Cloud interface

- ▶ DIRAC intègre de différents clouds avec les plug-ins spécifiques.
- ▶ Les interfaces clouds s'avèrent volatiles. Il faut suivre l'évolution rapide des technologies
  - ▶ apache-libcloud, nova, boto, rOCCI, OCCI deviennent obsolètes
  - ▶ Travail sur communication directe avec
    - ▶ OpenStack
    - ▶ OpenNebula (avec JINR)
    - ▶ EC2(Amazon), Azur(Microsoft), Google, IBM

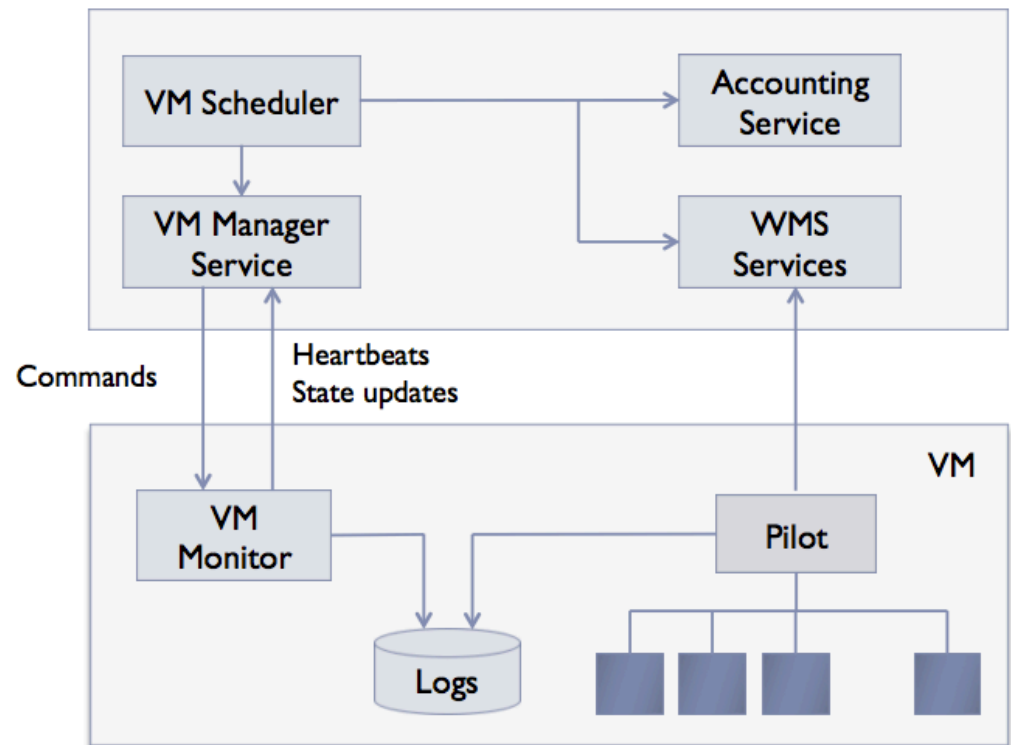




# Cloud Metascheduler

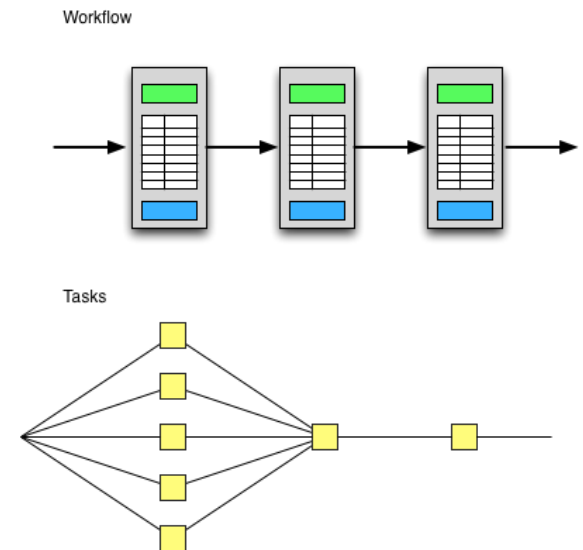
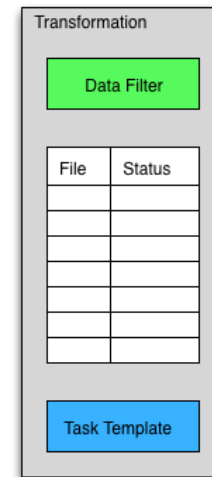
- ▶ Orchestration de travaux de plusieurs communautés sur plusieurs sites cloud afin d'assurer l'utilisation optimale des ressources en respectant les politiques d'utilisation communes

- ▶ En collaboration avec JINR, l'Université de Saint-Petersbourg
- ▶ Fait partie des services DIRAC France-Grilles, EGI Workload Manager



# High level workflows

- ▶ Sujet principale du Master Project DIRAC@IN2P3
  - ▶ Développement commun avec LUPM
- ▶ Généralisation du système des Transformations de DIRAC permettant la définition, l'exécution et le suivi de flux des travaux pilotés par des données à très grand échelle
  - ▶ Utilisateurs potentiels: CTA, Belle II, biomed, etc



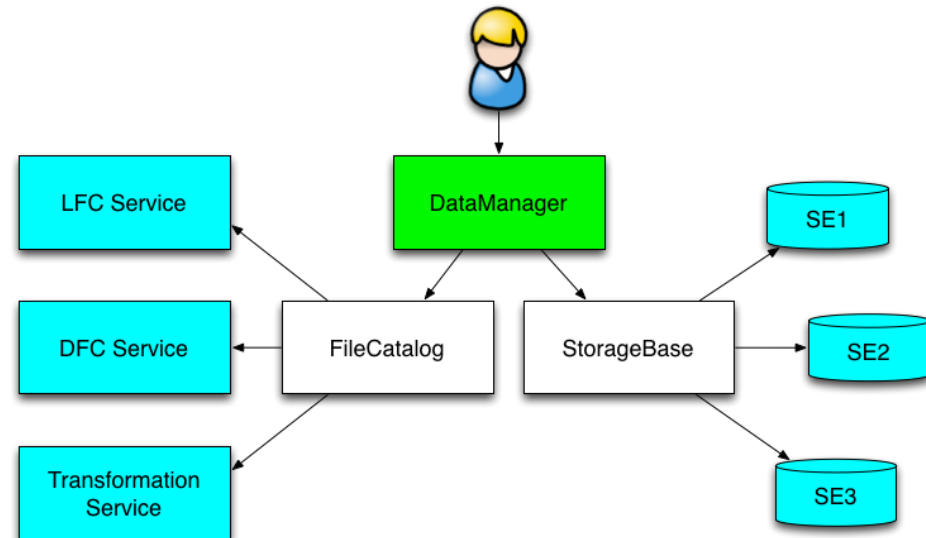
# Data Management

## ▶ Composants de DIRAC Data Management

- ▶ Storage plug-ins
- ▶ Catalogue(s)
- ▶ Interface DataManager à la File System
- ▶ Support pour les opérations en masse

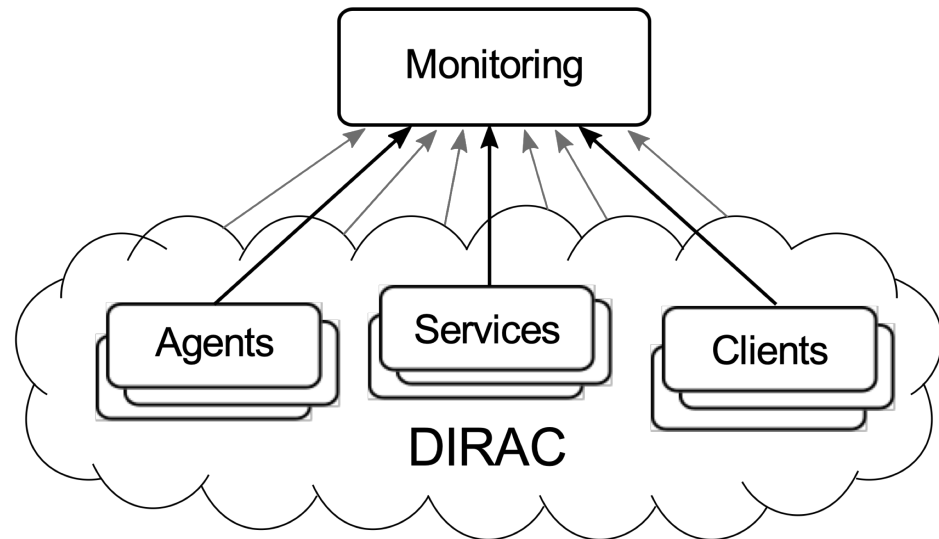
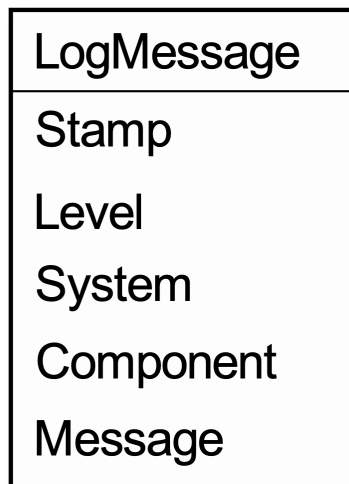
## ▶ Au CPPM

- ▶ Développement du DIRAC File Catalog
  - ▶ Support pour User Metadata
  - ▶ Extensions spécifiques, e.g. pour Eiscat 3D



# Évaluation de Go

- ▶ Approche des limites de mise à l'échelle avec Python
  - ▶ Limite capacité I/O et charge CPU (langage interprété)
- ▶ Évaluation de Go
  - ▶ Compilé et bien adapté pour des applications client-serveur
  - ▶ Comparer un service de logging en Go et Python (perf + code)



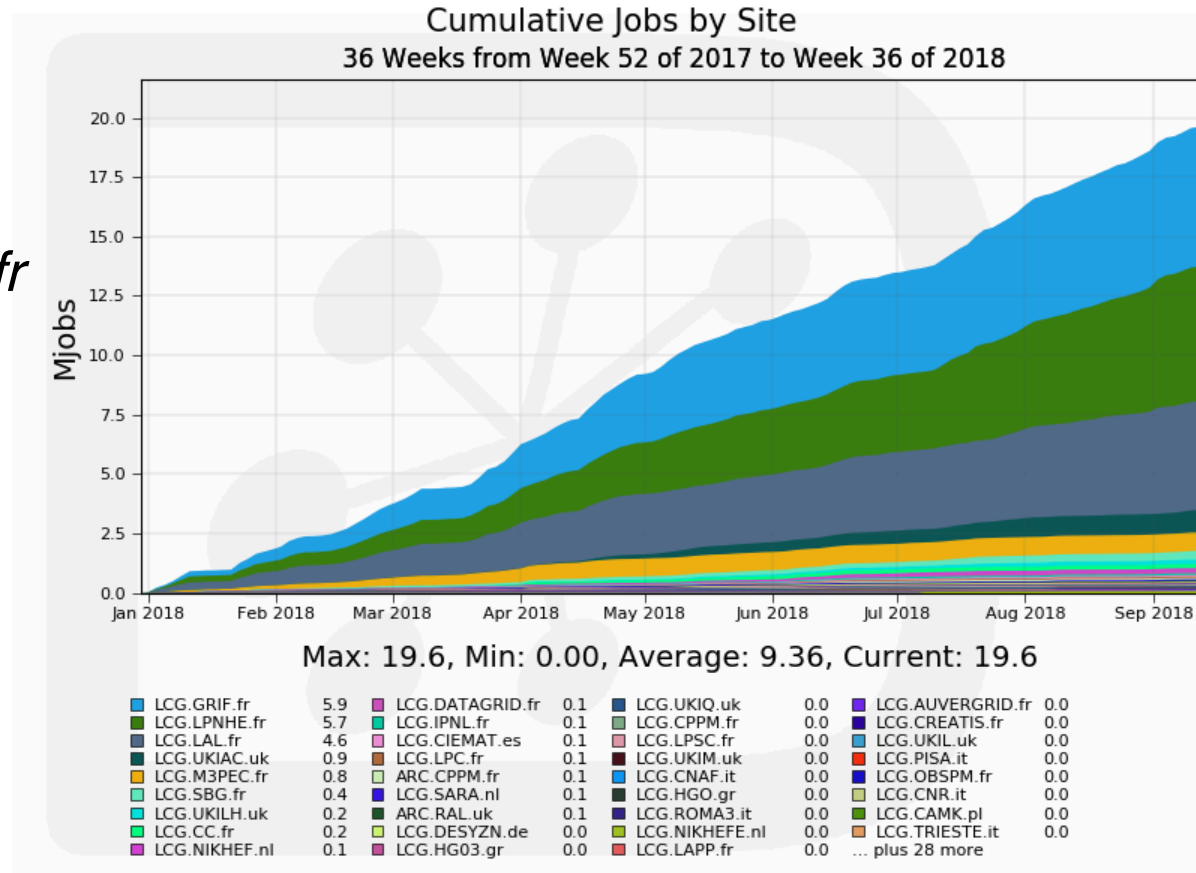
# SSO/OIDC Authentication

- ▶ **Open ID Connect** – standard industrielle pour des fournisseurs d'identité
  - ▶ Remplace progressivement le système d'identification grille basé sur des certificats X509
- ▶ **DIRAC doit supporter pleinement OIDC**
  - ▶ Authentification
  - ▶ Gestion interne des identités : Tous les composants sont touchés
- ▶ **Développement pour EOSC-Hub (EGI Workload Manager)**
  - ▶ Applicable pour d'autres projets: M<sup>3</sup>AMU, FG, etc

The image shows two screenshots of authentication interfaces. The top screenshot is the Janus login page, featuring the CNRS and Janus logos, the text 'dépasser les frontières', and the title 'Janus - Gestion des identités'. It includes a message 'Vous devez vous authentifier pour accéder à...', a 'Username:' field with the value 'atsareg@cppm.in2p3.', a 'Password:' field with masked characters, and a 'LOG' button. The bottom screenshot is the Aix-Marseille University authentication page, titled 'Aix-Marseille Université - Service d'authentification'. It has an 'Identifiant:' field with the value 'tsaregorod.a|', a 'Mot de passe:' field with masked characters, and a prominent blue 'LOGIN' button.

- ▶ Service management
  - ▶ FG-DIRAC
    - ▶ Coordination
    - ▶ En collaboration avec LUPM, CREATIS, CC/IN2P3
    - ▶ Avec la branche M3AMU
  - ▶ DIRAC4EGI
    - ▶ Coordination
    - ▶ En collaboration avec CYFRONET
  - ▶ M<sup>3</sup>AMU
- ▶ Service support
  - ▶ DIRAC@IGI
  - ▶ DIRAC@JINR
  - ▶ DIRAC@IHEP (FCPPL)

- ▶ Hosted by the CC/IN2P3, Lyon
  - ▶ *dirac.france-grilles.fr*
- ▶ Distributed administrator team
  - ▶ 5 participating universities
- ▶ In production since May 2012



Generated on 2018-09-12 19:20:08 UTC

- ▶ About 5 active communities – complex-systems, biomed, vo.france-grilles.fr, ...
- ▶ 15 ▶ > 25M jobs executed in 2018 at 90 sites

# DIRAC4EGI service



## Partners

- ▶ Operated by EGI
- ▶ Hosted by CYFRONET, Krakow
- ▶ DIRAC Project providing software, consultancy
- ▶ Supported via the EOSC-Hub H'2020 grant
- ▶ [dirac.egi.eu](http://dirac.egi.eu)

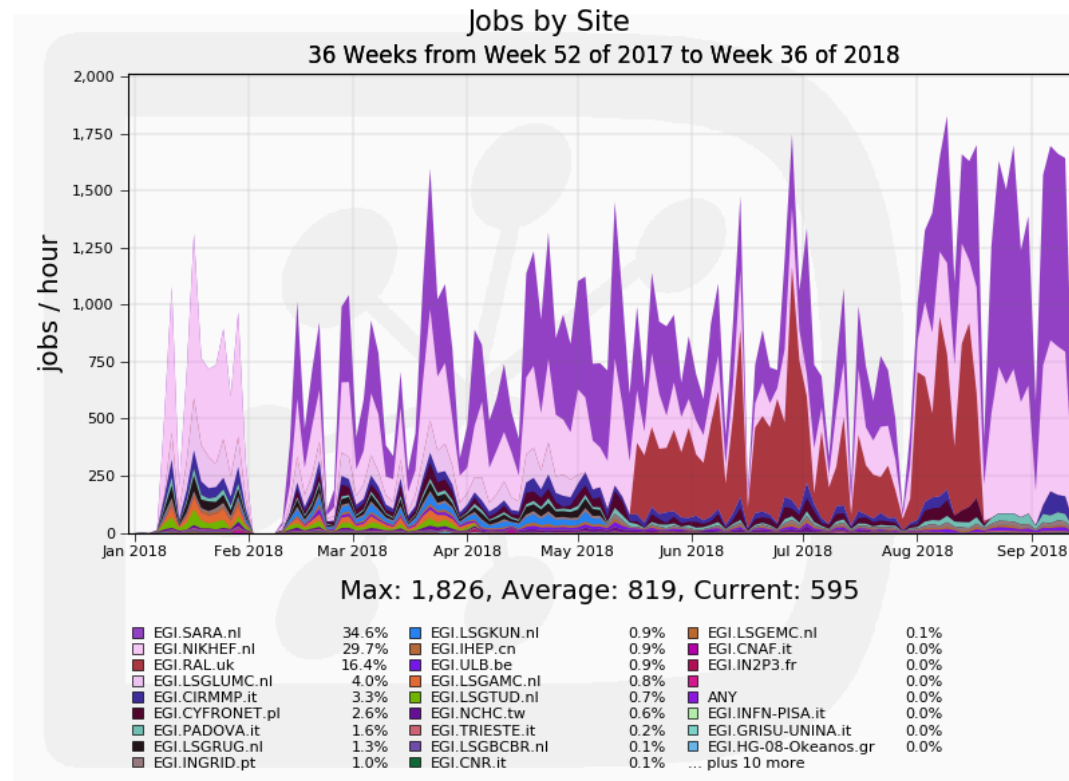
## 20 Virtual Organizations

- ▶ [enmr.eu](http://enmr.eu)
- ▶ [virgo](http://virgo)
- ▶ [eli-beams](http://eli-beams)
- ▶ [eiscat.se](http://eiscat.se)
- ▶ [fedcloud.egi.eu](http://fedcloud.egi.eu)
- ▶ ...

## Usage

- ▶ > 7 million jobs processed in 2018

## DIRAC4EGI activity snapshot



Generated on 2018-09-12 19:23:21 UTC



- ▶ Configuré comme une des communautés servies par France-Grilles DIRAC
  - ▶ Personnalisation AMU si nécessaire avec un portail dédié
- ▶ Ressources connectées
  - ▶ CPPM T2
  - ▶ CPPM Cloud (calcul et stockage)
  - ▶ Mesocentre St-Jérôme
  - ▶ DEC (?)
- ▶ Avant du démarrage
  - ▶ Authentification des utilisateurs par le service OIDC de l'Université
  - ▶ Page d'accueil pour les utilisateurs (guide, enregistrement)
  - ▶ Tutoriel
- ▶ Focus des activités en cours

# Coordination des projets DIRAC

---

- ▶ DIRAC Consortium
  - ▶ Organisation de DIRAC User Workshops
    - ▶ Le prochain et à Londres, 14-17 mai
  - ▶ Boards of representatives
    - ▶ New members, coordinator elections
  - ▶ Maintenance du site [diracgrid.org](http://diracgrid.org)
- ▶ Master Project DIRAC@IN2P3
  - ▶ Coordination des développements
  - ▶ Publications, conférences

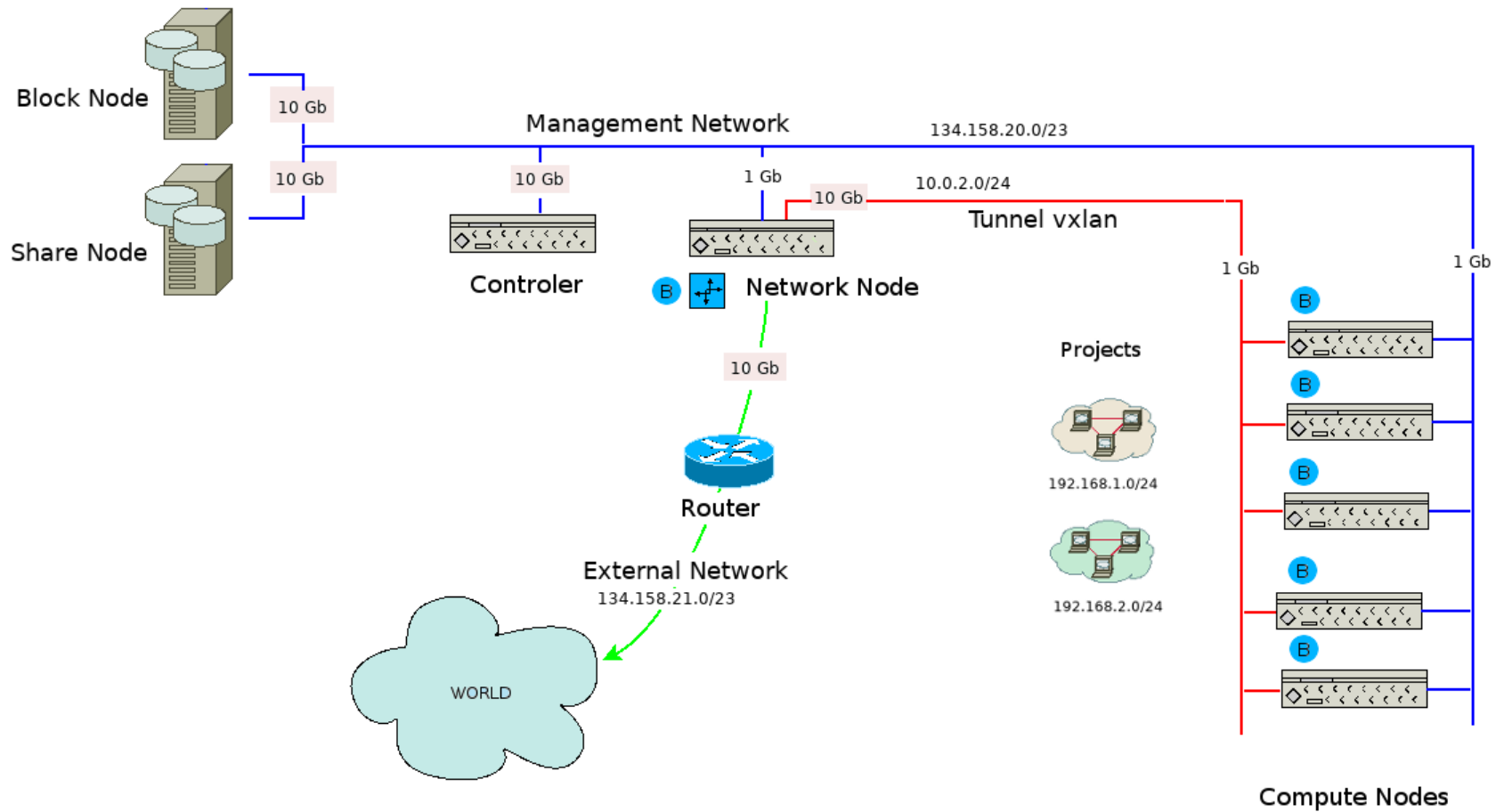
# DIRAC dans le projet FCPPL

---

- ▶ Mise en place du service DIRAC@IHEP
  - ▶ BES III
  - ▶ Juno
  - ▶ CEPC
  - ▶ ...
- ▶ Soutiens des administrateurs du service
- ▶ Coordination des développements
  - ▶ Connexion des clouds Amazon, OCCI
  - ▶ Utilisation des clouds pour les travaux multi-parallèles

- ▶ Le but : services cloud avancés pour des utilisateurs du CPPM, de l'AMU, de France-Grilles
  - ▶ Remplacement des serveurs des groupes CPPM
  - ▶ Les ressources de calcul et de stockage générales
- ▶ Plate-forme fonctionnelle
  - ▶ Préproduction
- ▶ Installation OpenStack
  - ▶ 400 coeurs de calcul
  - ▶ 200 To de stockage
- ▶ Étude de technologie
  - ▶ Composants OpenStack
  - ▶ Stockage CEPH

# Réseau du Cloud CPPM



# Support aux applications

---

- ▶ Dans le cadre de EOSC-Hub
  - ▶ Eiscat 3D
    - ▶ Portail Web dédié, File Catalog dédié
  - ▶ WeNMR, Virgo, ...
    - ▶ Tutoriel, computing school, etc
- ▶ Dans le cadre de ESCAPE
  - ▶ Km3NET
- ▶ “Big Data”
  - ▶ Études de “cloud platform for BigData applications”
  - ▶ Exploration des cas des applications “BigData” au CPPM

- ▶ Groupe jeune et dynamique
- ▶ Plusieurs développements de logiciels de l'intérêt général
- ▶ Plusieurs responsabilités des services en production avec de nombreuses utilisateurs dans le monde
- ▶ Réflexion en cours sur d'autres sujets relatifs aux activités du groupe:
  - ▶ Plate-forme cloud au CPPM
  - ▶ Plate-forme pour les applications BigData