

# Calcul scientifique dans le Cloud : Des serveurs virtuels aux conteneurs

Jérôme Pansanel

`jerome.pansanel@iphc.cnrs.fr`

Bordeaux – 9 novembre 2016



## **Calcul scientifique**

- Déployer des infrastructures (simples ou complexes)
- Preuve de concepts
- OS ou outils spécifiques
- Intégration avec des outils supportant déjà le Cloud nativement
- 

## **Hébergement et analyse**

- VM pour héberger et analyser des données à la demande
- Visualisation de données

## Virtual Imaging Platform (VIP)

- Portail Web pour la simulation médicale et l'analyse d'images
  - Utilisation de DIRAC et de la VO Biomed
  - Utilisation des ressources de l'IPHC et du CC-IN2P3
- <http://dx.doi.org/10.1109/TMI.2012.2220154>

*Creatis*

## Phénotypage à haut-débit (Phenome)

- Large collection de données phénotypique
- Stockage basé sur iRODS (~ 1PB)
- Analyse de données basée sur un workflow Grille et Cloud

## Biologie des systèmes

- Étude de la transmission d'informations entre animaux
- Analyse de réseaux sociaux
- Basé sur des scripts R et des bibliothèques externes (*RSiena*)
- Nécessite des VMs avec une grande quantité de mémoire (> 200 Go)  
→ <http://dx.doi.org/10.3389/fpsyg.2016.00539>

## Institut Français de Bioinformatique (IFB)

- Noeud français du projet Elixir
- Opère un Cloud communautaire pour la bioinformatique
- IFB-Core, un Cloud PaaS / SaaS hébergé à l'IDRIS adossé à des sites secondaires (satellites)
- Deux de ces satellites sont également membre de FG-Cloud (IPHC et Université de Lille 1)
- Noeud régional strasbourgeois BISTRO



## **Biologie des systèmes**

- Étude de la transmission d'informations entre animaux
- Analyse de réseaux sociaux
- Basé sur des scripts R et des bibliothèques externes (*RSiena*)
- Nécessite des VMs avec une grande quantité de mémoire (> 200 Go)  
→ <http://dx.doi.org/10.3389/fpsyg.2016.00539>

## **Institut Français de Bioinformatique (IFB)**

- Noeud français du projet Elixir
- Opère un Cloud communautaire pour la bioinformatique
- IFB-Core, un Cloud PaaS / SaaS hébergé à l'IDRIS adossé à des sites secondaires (satellites)
- Deux de ces satellites sont également membre de FG-Cloud (IPHC et Université de Lille 1)
- Noeud régional strasbourgeois BISTRO

## SME

- Startup INRIA basée à Rennes
- Expertise avec les workflows Galaxy (génomique)
- Besoin de serveurs de manière occasionnelle pour faire des démonstrations
- Levée de fond

## Analyse génomique

- Serveurs de calcul pour l'analyse de génome humain
- Données anonymisées
- Focus particulier sur la sécurité (tenant dédié, disques chiffrés, ...)
- SLA dédié

## Hackathon

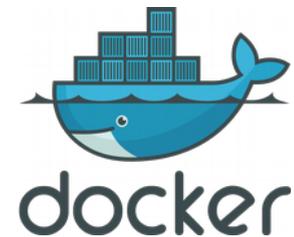
- Hacking Health Camp (Strasbourg, mars 2016)
- Besoin de ressources à la demande le temps d'un week-end
- Développer des preuves de concept
- Docker à la demande pour héberger des micros services



<+> HACKING HEALTH

## Docker

- Automatisation d'applications dans des conteneurs
- Basé sur LXC et cgroups
- Accessible via une API haut niveau
- Orienté micro-services  
→ <https://www.docker.com/>



## Orchestration

- Outils pour gérer de nombreux conteneurs
- Swarm (<https://docs.docker.com/swarm/overview/>)
- Kubernetes (<http://kubernetes.io/>)

## Applications scientifiques

- Assemblage de micro services dans le cadre d'un workflow
- Distribution de logiciels scientifiques
- Environnement léger pour les environnements de calcul à façon

## OpenStack

- Project Magnum
- Orchestration de dockers à travers une API
- Utilisable depuis Kubernetes et SWARM (et mesos)
- Stable depuis Mitaka

## Magnum

- Basé sur Heat (et Ironic pour les performances)
- Documentation Admin :  
<https://wiki.openstack.org/wiki/Magnum>
- Documentation Utilisateur :  
<http://docs.openstack.org/developer/magnum/dev/quickstart.html>  
<http://docs.openstack.org/developer/magnum/userguide.html>
- Déploiement d'ici fin novembre à l'IPHC
- Patch post-mitaka développé au CERN :  
<https://gitlab.cern.ch/cloud-infrastructure/openstack-magnum/tree/cern-mitaka>

# Questions ?