FG-Cloud : un Cloud fédéré pluridisciplinaire pour le déploiement et l'orchestration de conteneurs

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

ANF UST4HPC - 21 janvier 2021

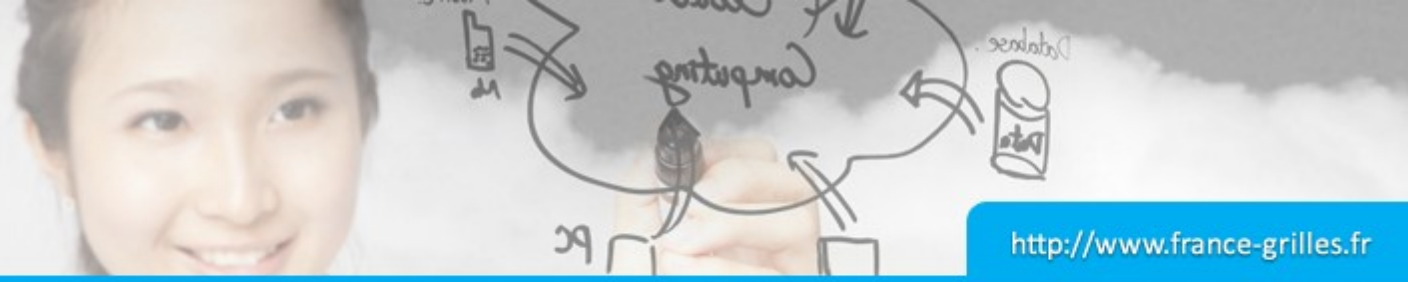
À propos

Au sommaire

- Cloud OpenStack de production
- La fédération FG-Cloud
- Utilisation d'OpenStack
- Déployer Kubernetes sur OpenStack
- Questions / discussions

Qui suis-je ?

- Jérôme Pansanel – ingénieur de recherche à l'IPHC
- Responsable de la plateforme SCIGNE
  `https://grand-est.fr`
- Directeur technique de France Grilles
  `http://www.france-grilles.fr`

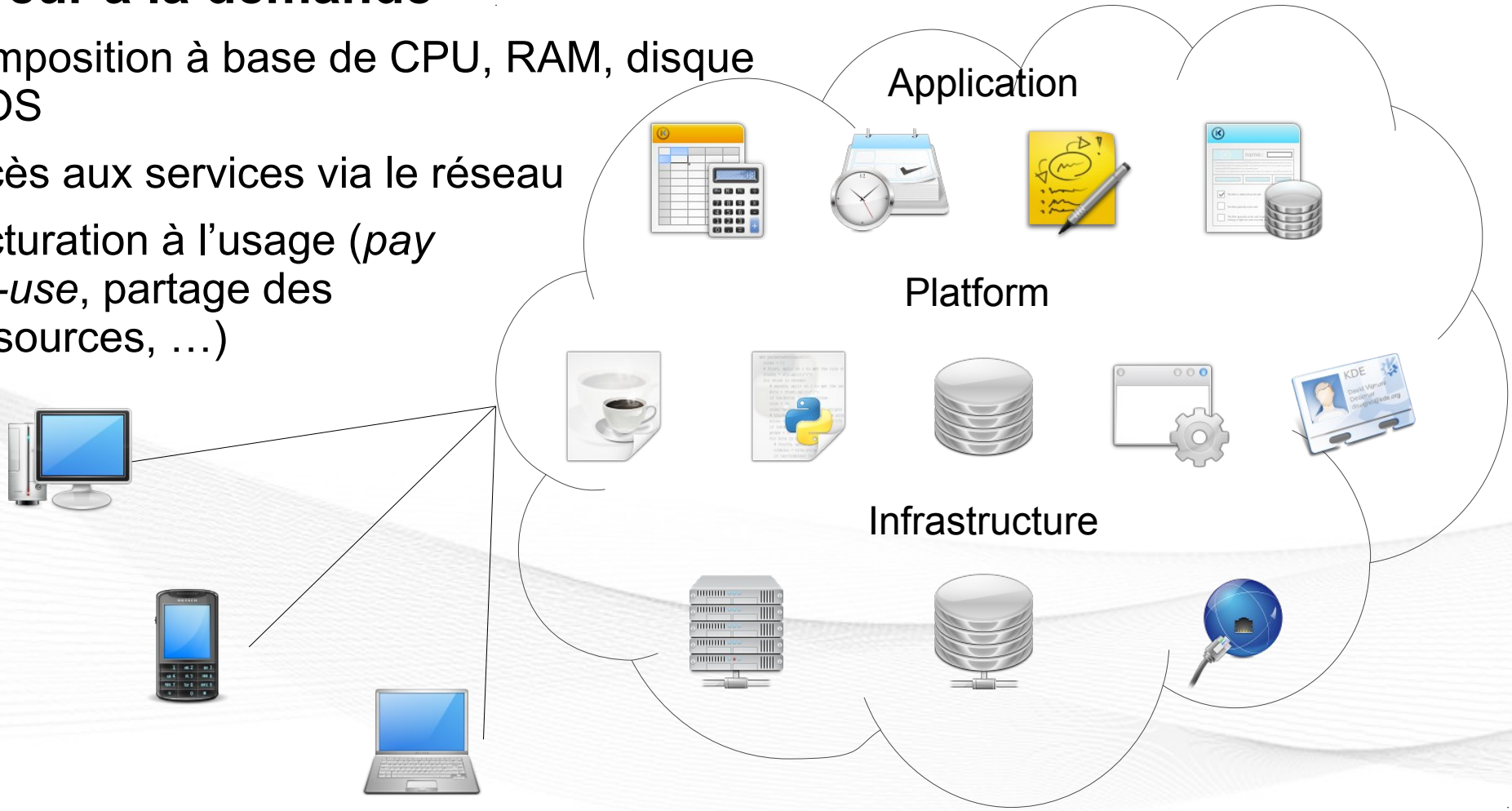


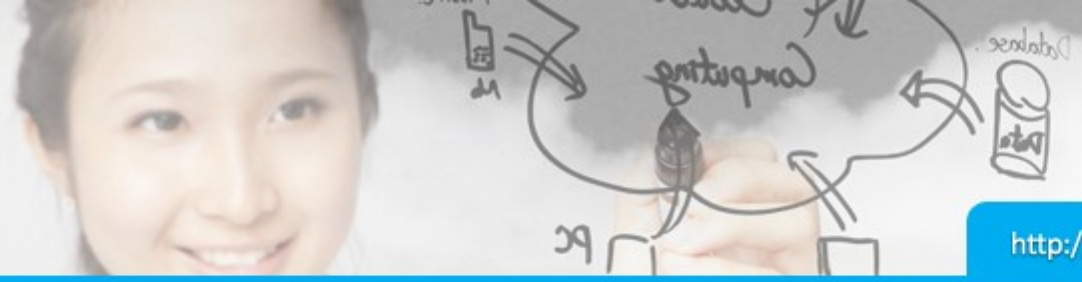
Cloud OpenStack de production

Le Cloud IaaS

Serveur à la demande

- Composition à base de CPU, RAM, disque et OS
- Accès aux services via le réseau
- Facturation à l'usage (*pay per-use*, partage des ressources, …)

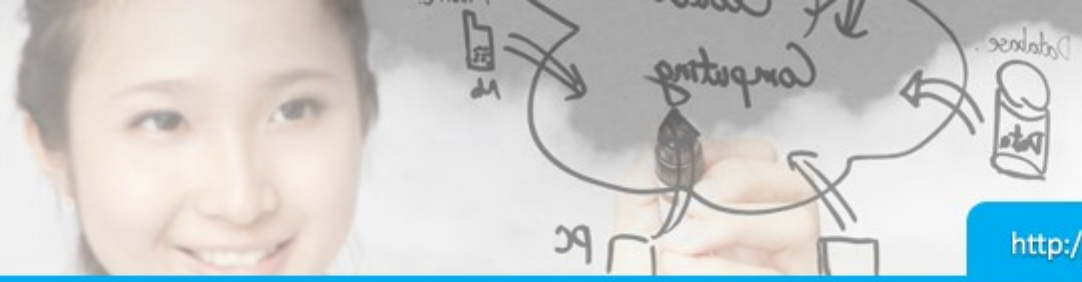




Déployer Docker et Kubernetes sur OpenStack

Des avantages intéressants …

- Utiliser un OS ou des outils spécifiques
- Déployer des infrastructures de test (rapidement et simplement)
- Intégration avec des outils supportant le Cloud nativement (par ex. docker-compose)
- Anticiper les besoins (facilité d'ajout de services, …)
- Mettre en place des plateformes d'analyse
- Savoir facilement intégrer les réponses aux besoins de calcul hors norme
- Gestion des logiciels propriétaires par équipe
- Pouvoir déborder sur les centres partenaires / cloud commerciaux



Choix de l'infrastructure

Accès

- Facilité d'utilisation et fiabilité
- Coût, contraintes / facilités d'accès (partenariats, type d'accès, etc)
- Compatibilité des APIs avec l'outil de gestion de conteneurs
- Ressources suffisantes pour votre projet, interopérabilité

Gestion des images

- Disponibilité d'images spécifiques (Fedora CoreOS ou RHEL CoreOS)
- Possibilité de charger ses propres images

Gestion du réseau

- Définition des groupes de sécurité / pare-feu
- Création et type de réseaux

Ressources dans le Cloud

Différents types de ressource dans le Cloud

- Machine virtuelle (VM)
- Bare metal
- Conteneurs

VM vs bare metal

- Performance (coût CPU et mémoire, latence réseau)
- Roll-back / snapshot
- Sécurité (machine mono projet vs flash bios, re-configuration des hyperviseurs, sécurité réseau)

Infrastructures de production

Différents types d'infrastructure

- Cloud commerciaux
  - Amazon, Google, Azur, … (US)
  - OVH, Outscale, Open Telekom Cloud, ~~Cloudwatt~~ († 2020), ~~Numergy~~ († 2016)
- Cloud académiques
  - Cloud du laboratoire
  - Plateforme régionale (mésocentre)
  - Fédérations de Cloud (FG-Cloud, IFB, EGI FedCloud, ...)

Cloud commerciaux

Clarifying Lawful Overseas Use of Data Act

- Divulgation des informations personnelles dans le cadre d'enquêtes
- Les données n'ont pas besoin d'être stockées sur le territoire américain
- Pas de validation des demandes par un juge
- Normalement encadré par un protocole cadre d'échange des données

Cloud commerciaux

OVH

- Cloud basé sur OpenStack
- Possibilité de charger ses propres images, K8S à la demande
- Instance spécifique CPU et/ou GPU
- Tarif : 8 vCPUs, 45 Go de RAM, NVIDIA v100 → 800 € / mois
- https://www.ovhcloud.com/fr/public-cloud/orchestration/

Open Telekom Cloud (T-System)

- Cloud public de T-Systems (filiale Deustche Telekom), basé sur OpenStack
- Possibilité de charger ses propres images, K8S à la demande
- Instance spécifique CPU et/ou GPU, réseau infiniband
- Tarif : 8 vCPUs, 64 Go de RAM, NVIDIA V100 → 2000 € / mois
- https://open-telekom-cloud.com/en

https://www.data-infrastructure.eu/

Cloud académiques

Plateforme Cloud @ Mésocentre AMU

- Projet porté par l'initiative Mésocentre Aix Marseille Université
- OpenStack / CEPH
- ~ 1536 cœurs / 400 To de stockage, GPUs
- Également accessible à travers la fédération France Grilles
- `https://mesocentre.univ-amu.fr/ressources-cloud/`

Cloud académiques

Plateforme OpenStack @ GRICAD

- Hébergée et gérée par GRICAD, infrastructure de calculs intensifs et de données à Grenoble
- OpenStack / CEPH
- ~ 1152 cœurs / 519 To de stockage, GPUs
- Déploiement basé sur Kolla
- Également accessible à travers la fédération France Grilles
- `https://gricad-doc.univ-grenoble-alpes.fr/nova/`

Cloud académiques

Plateforme OSCAR (mésocentre UCA)

- Hébergée par le mésocentre de l'Université de Clermont-Auvergne
- Projet en commun avec le LPC Clermont
- OpenStack / CEPH
- ~ 2000 cœurs / 300 To de stockage
- Également accessible à travers la fédération France Grilles et IFB
- `https://mesocentre.uca.fr/ressources`

Cloud académiques

Plateforme SCIGNE

- Hébergée et gérée par l'IPHC (Strasbourg)
- Appel à projet en commun avec le mésocentre de l'Université de Strasbourg
- OpenStack / CEPH
- ~ 1064 cœurs et 432 To de stockage, GPUs
- Kubernetes-as-a-service, calcul scientifique, plateforme d'analyse, etc
- Également accessible à travers les fédérations France Grilles, IFB et EGI
- https://grand-est.fr/

Cloud académiques

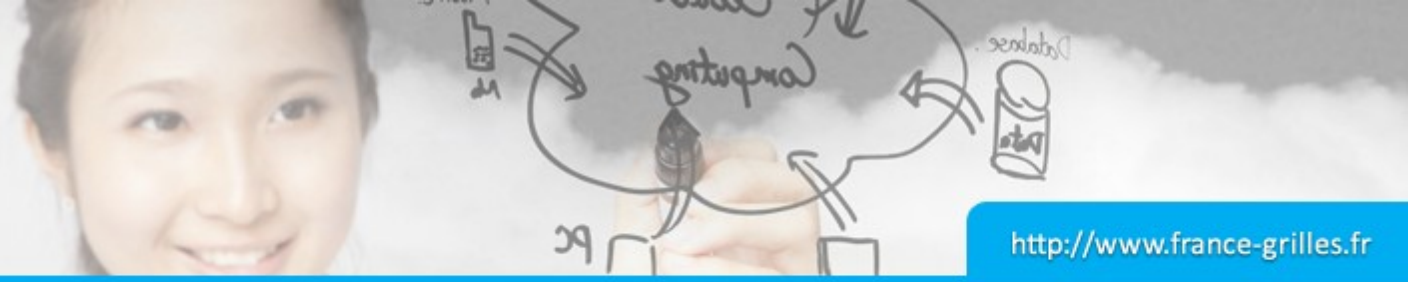
Plateforme OpenStack du mésocentre de Lille

- Hébergée par le mésocentre de l'Université de Lille
- OpenStack / CEPH
- ~ 1436 cœurs et 215 To de stockage
- Également accessible à travers les fédérations France Grilles, IFB et EGI
- `http://hpc.univ-lille.fr/cloud-openstack`

Cloud académiques

Plateforme Cloud @ Virtual Data

- Hébergée à P2IO (Orsay) et gérée par IJClab
- OpenStack / CEPH
- ~ 8000 cœurs / 400 To de stockage
- Hébergement d'infrastructure de calcul scientifique : Spark (300 cœurs / 40 To), JupyterHub (80 coeurs), DW4NP (Data Workflow 4 Nuclear Physics - projet CSNSM-IPNO), Kubernetes-as-a-service
- Également accessible à travers la fédération France Grilles
- `https://openstack.lal.in2p3.fr/`



La fédération FG-Cloud

France Grilles

Le GIS France Grilles

- Créé en 2010 par 8 partenaires
- Opérer et fédérer des moyens de calcul et de stockage géographiquement distribué pour la recherche scientifique
- Une communauté d'experts
- Un ensemble de services (FG-Cloud, FG-DIRAC, FG-iRODS, formation, etc)
- Évolution naturelle : des grilles de calcul au « *Cloud computing* »
- `http://www.france-grilles.fr`

FG-Cloud

Les objectifs

- Établir une infrastructure nationale de *Cloud* de production, pour le stockage et le traitement de données massives pour la recherche scientifique
- Promouvoir son usage dans toutes les communautés scientifiques
- Disposer d'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

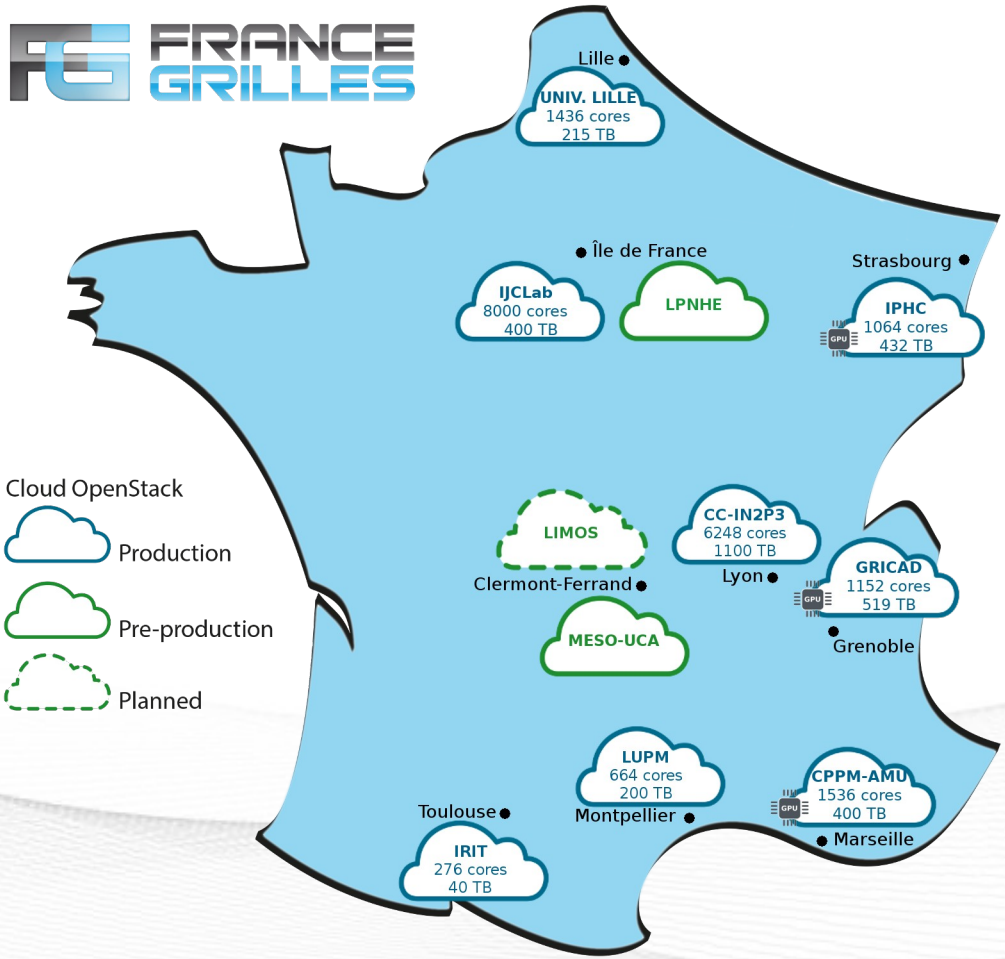
FG-Cloud

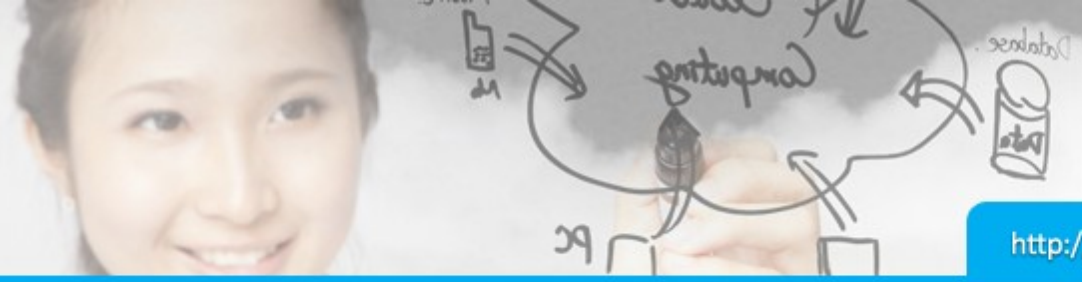
La stratégie

- Organiser la fédération en se basant sur la participation volontaire des sites, en adéquation avec leur stratégie locale
- Ouvrir le service à l'ensemble des communautés scientifiques et organiser des formations
- Créer et animer un réseau d'experts
- Faire le lien avec les infrastructure et projets internationaux
- Pas d'*a priori* sur les technologies Cloud
- Utiliser des APIs interopérables
- Assurer une veille technologie et scientifique

L'infrastructrure FG-Cloud

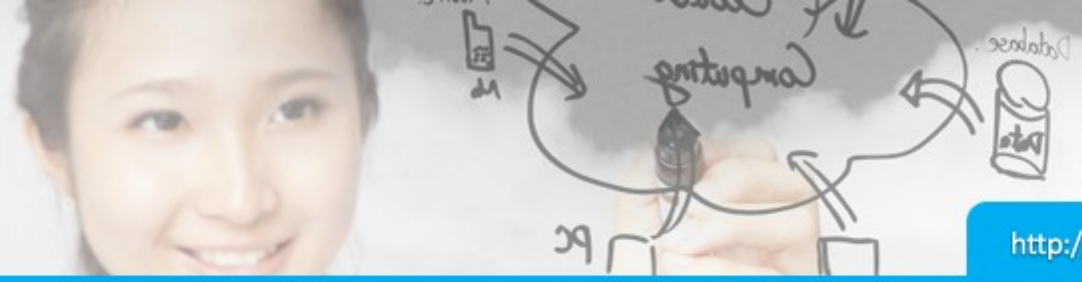
- Infrastructure fédérative de Cloud
- Pilotage par le groupe FG-Cloud
- En accord avec les stratégies locales
- Accessible via le catalogue de services France Grilles
- Géré avec OpenStack (API & CLI, Web)
- 6 sites en production, 7500 cœurs, 1500 To de stockage
- Surveillance fonctionnelle, distribution des images, authentification centralisée (outils non invasifs)





Organisation

- Groupe technique : 2 réunions / mois
  - Tour de table des sites
  - Suivi des utilisateurs
  - Développements en cours
  - Liste de discussion
  - Canal RocketChat
- Mutualisation : expertise partagée, développements
- Compétences destinées à :
  - Utilisateurs des Clouds
  - Développeurs de solutions Cloud
  - Administrateurs de site voulant s'y mettre



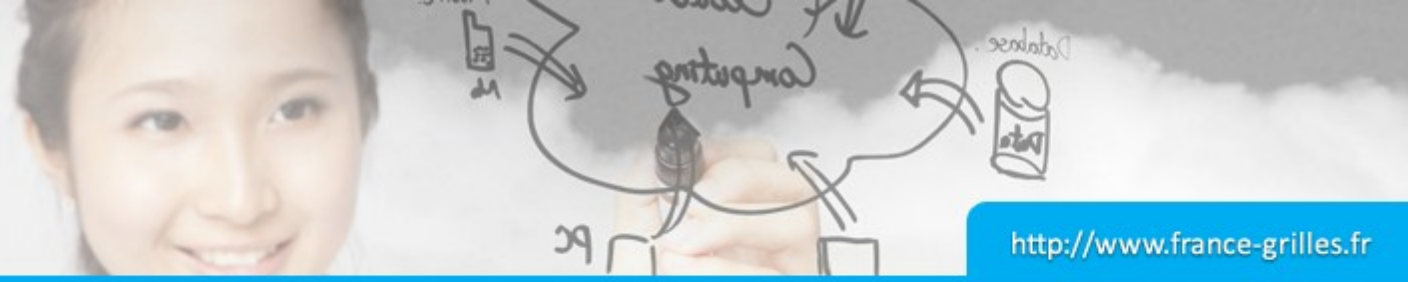
Accueil de nouveaux sites

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

Accès pour les utilisateurs

- Qui : tous les partenaires du GIS et tiers qui ont des projets en commun
- Prendre contact avec France Grilles → info@france-grilles.fr
- Répondre à un questionnaire
- Accès via tableau de bord, API ou CLI
- Accompagnement
- Pas de contre-partie financière (*)
- Participation à la communauté
  - Présenter vos travaux avec remerciements
  - Référencer les publications dans HAL

(*) Sauf en cas de besoins importants car FG ne finance pas l'infrastructure



Utilisation d'OpenStack

OpenStack

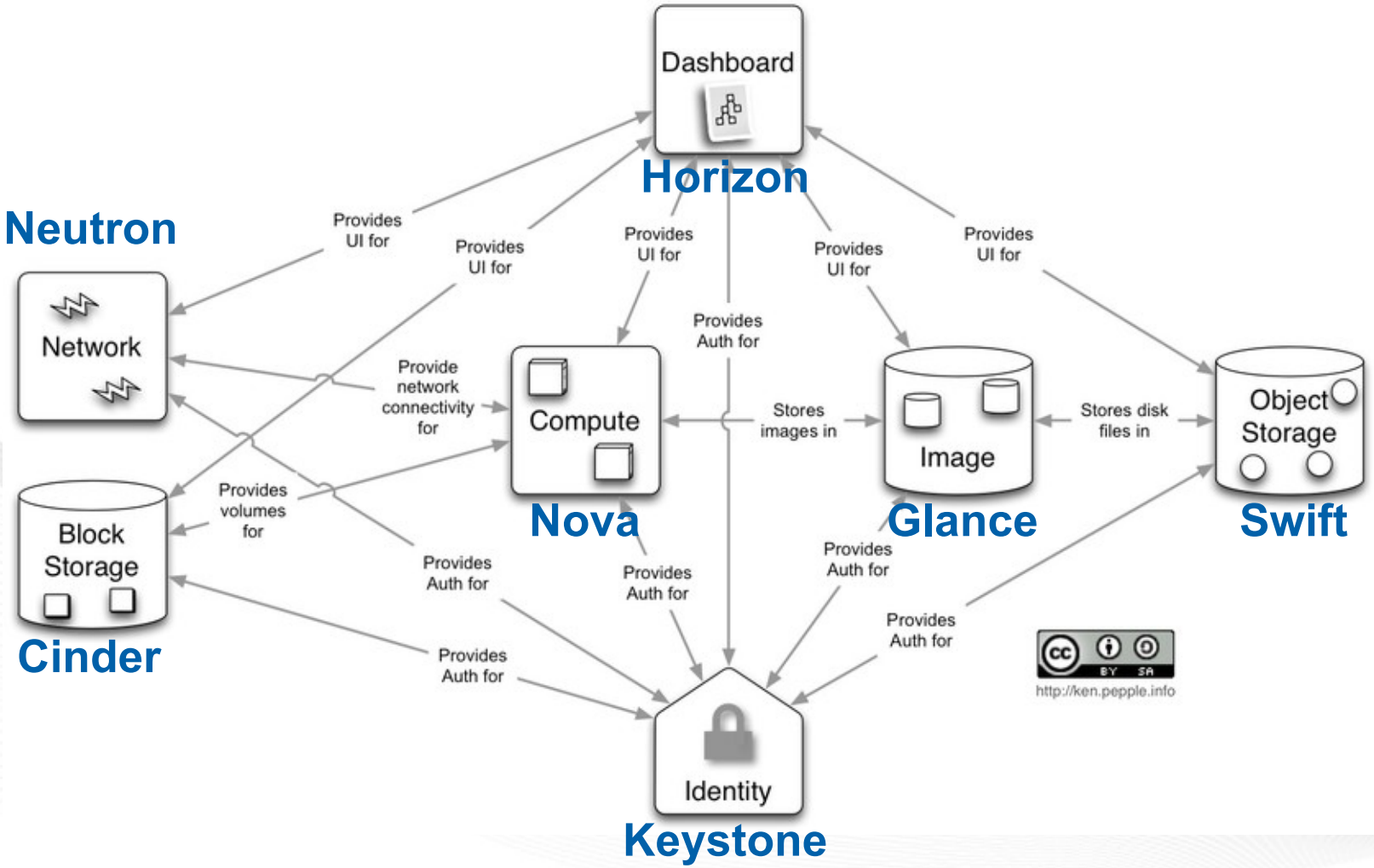
En quelques mots

- *Middleware* Cloud ouvert / libre (licence Apache 2.0)
- Rackspace (stockage) + NASA (calcul)
- Développement Python, très actif
- RedHat, IBM, Dell, Intel, Cisco, Juniper, NetApp, HP, VMWare, …
- Disponible dans de nombreuses distributions

Modules

- Keystone : Identité
- Glance : Images VM
- Cinder : Stockage bloc
- Neutron : Réseau (SDN)
- Nova : Calcul
- Horizon : UI web
- Swift : Stockage objet
- Heat : Orchestration
- Magnum : Conteneurs
- Ironic : Bare-metal
- Octavia : LBaaS
- Barbican : Gestion de clés
- Manila : Systèmes de fichiers partagés

Architecture OpenStack



Architecture OpenStack *IRL*

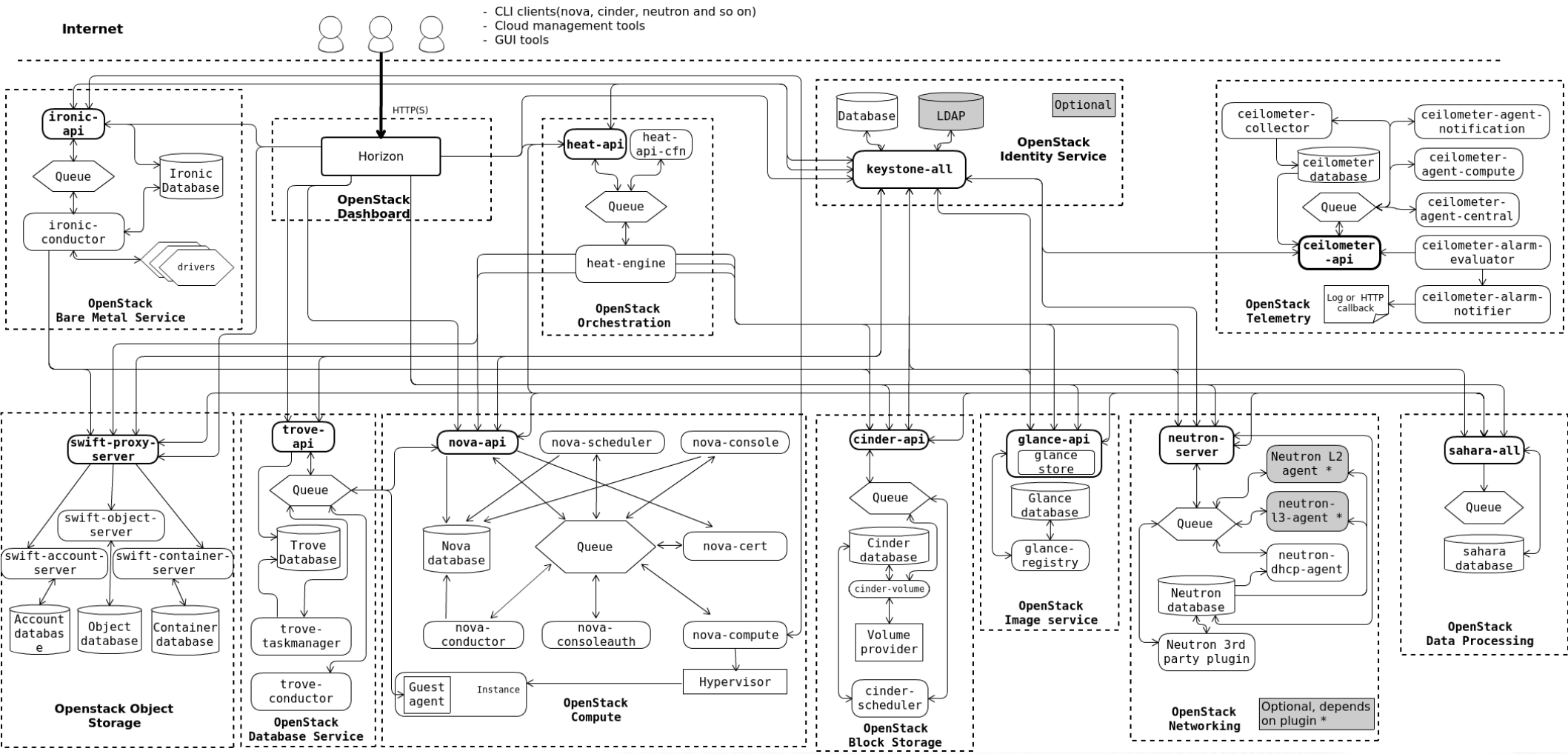


Tableau de bord Horizon

Type d'authentifcation

Domaine OpenStack

Nom d'utilisateur

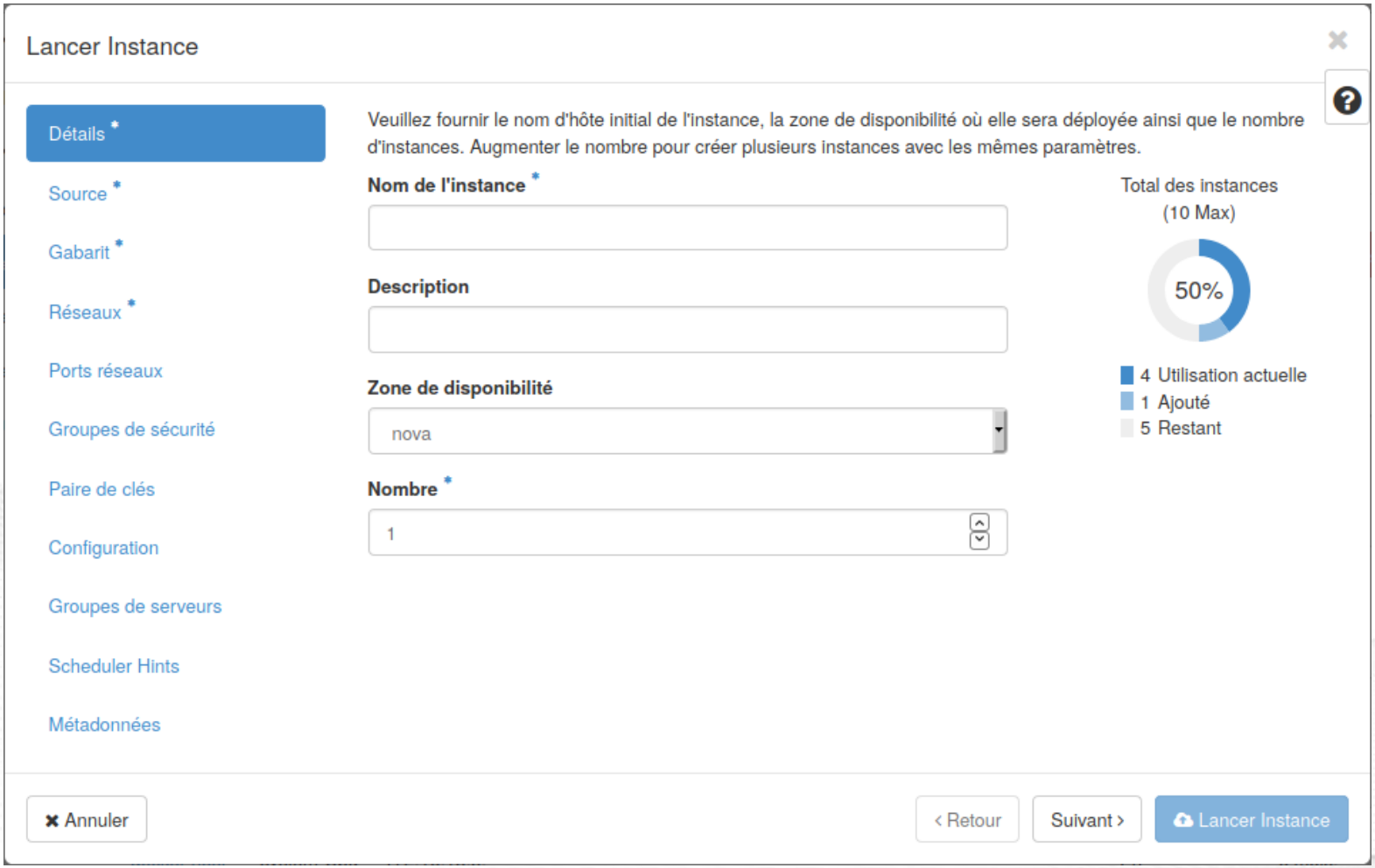
Mot de passe

Et en ligne de commande

Définir les variables d'environnement

```
export LANG=en_US.utf-8
export LC_ALL=en_US.utf-8
export OS_USERNAME=username
export OS_PASSWORD=password
export OS_PROJECT_NAME=FG_Cloud
export OS_USER_DOMAIN_NAME=default
export OS_PROJECT_DOMAIN_NAME=default
export OS_AUTH_URL=https://sbgcloud.in2p3.fr:5000/v3
export OS_IDENTITY_API_VERSION=3
```

Lancement d'une machine virtuelle



Et en ligne de commande

Lancer une VM

```
$ openstack server create --key-name cloudkey \
  --image 92876744-962e-4831-851e-b57e68bc8cdc \
  --flavor m1.small \
  --nic net-id=16ddcf0e-05c7-4023-8fc6-9cb49fd93aa4 \
  nom_de_ma_vm
```

Les paramètres importants

- Nom de clé SSH
- L'image de base
- Le gabarit
- L'identifiant du réseau
- Et le nom de la VM ...

Et en ligne de commande

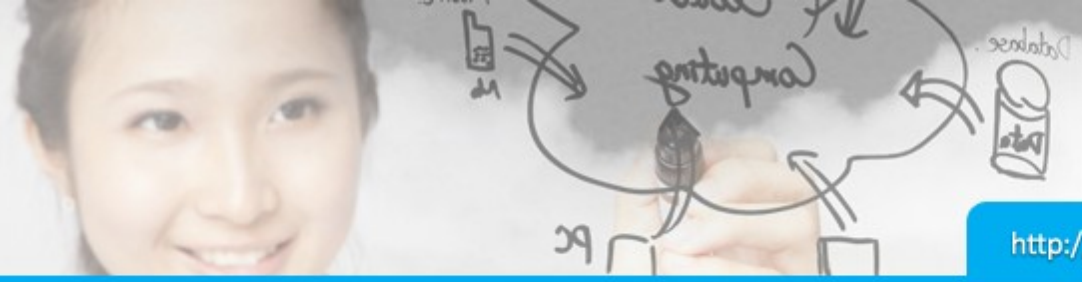
La VM

Gestion du réseau

```
$ openstack server show -c addresses -c flavor -c id -c image \
  -c key_name -c name -c security_groups -c volumes_attached \
  cbdf6d91-1317-4209-9985-8eb126f2b58a
+-------------------+--------------------------------------+
| Field             | Value                                |
+-------------------+--------------------------------------+
| addresses         | iphc-net=172.16.160.15               |
| flavor            | m1.small                             |
| id                | cbdf6d91-1317-4209-9985-8eb126f2b58a |
| image             | Image for CentOS Server 8 [CentOS/8/KVM] |
| key_name          | cloudkey                             |
| name              | nom_de_ma_vm                         |
| security_groups   | name='default'                       |
| volumes_attached  |                                      |
+-------------------+--------------------------------------+
```

Gestion du stockage permanent

Gestion de la sécurité



Déployer Kubernetes sur OpenStack

Cloud Provider OpenStack

Déploiement avec kubeadm

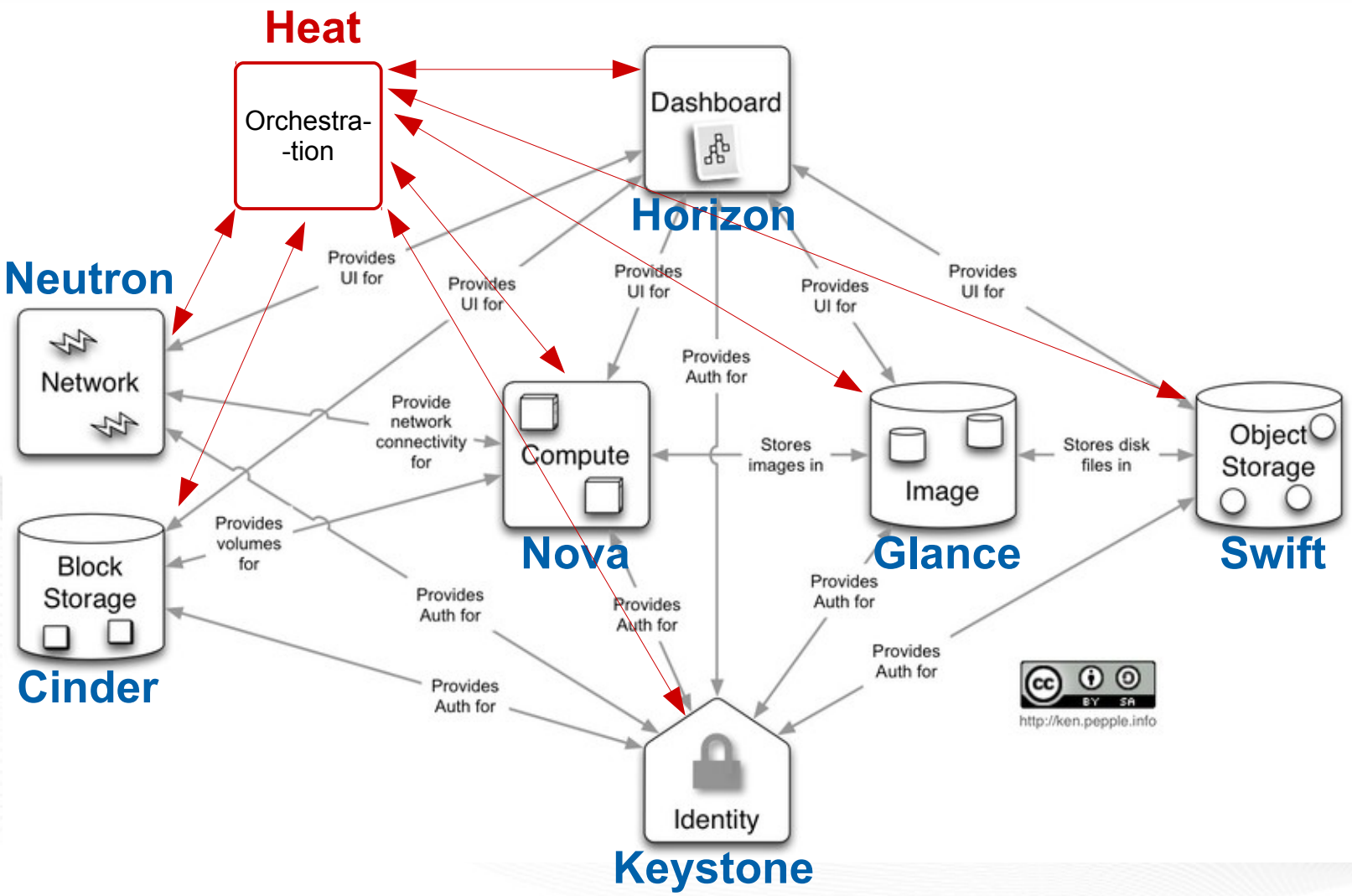
- Intégration d'OpenStack dans Kubernetes
- Plugin pour **kubeadm**
- Nécessite le module OpenStack Octavia pour la gestion de l'équilibrage de charge
- Nécessite l'ouverture de ports
- https://github.com/kubernetes/cloud-provider-openstack/blob/master/docs/openstack-cloud-controller-manager/using-openstack-cloud-controller-manager.md
- https://github.com/kubernetes/cloud-provider-openstack
- https://kubernetes.io/blog/2020/02/07/deploying-external-openstack-cloud-provider-with-kubeadm/

Orchestration avec Heat

Orchestration de machines virtuelles

- Inspiré de CloudFormation (AWS)
- Description de l'infrastructure dans un fichier *template* structuré (YAML)
- Format HOT (Heat Orchestration Template)
- Ce fichier peut être instancié en stack (ensemble de ressources)
- Heat s'occupe de l'allocation des ressources en faisant les demandes auprès des différents services OpenStack
- Possibilité de passer des variables au lancement d'un stack
- Possibilité de récupérer des valeurs (IPs, …) à la fin du lancement
- Possibilité d'*auto-scaling* (interaction avec les outils de métrologie)
- `https://docs.openstack.org/heat/latest/template_guide/`
- `https://gricad-gitlab.univ-grenoble-alpes.fr/kubernetes-alpes/nova-k8s/-/tree/master/`

Orchestration avec Heat



OpenStack Magnum

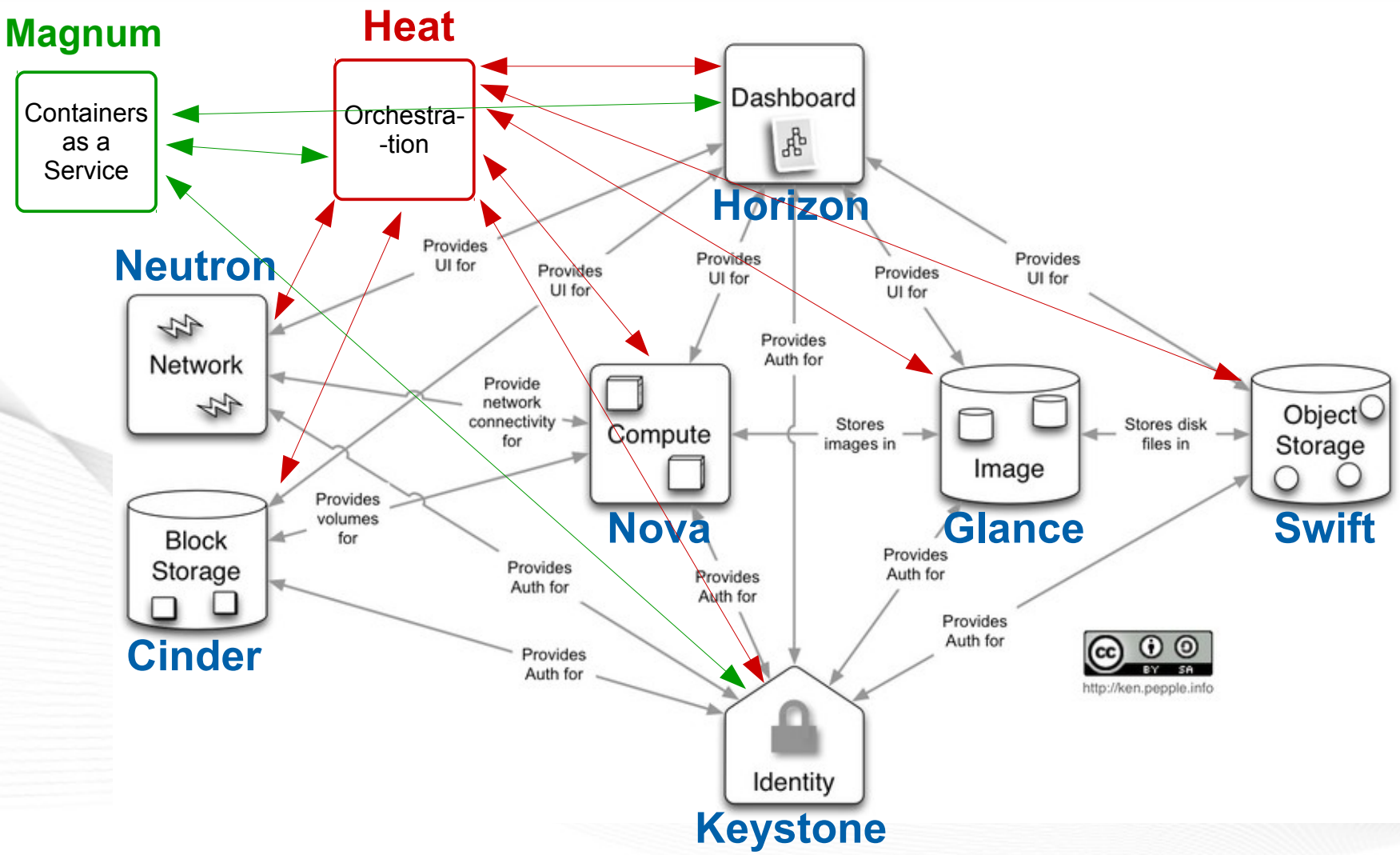
Des avantages

- Déploiement simple de cluster Kubernetes
- Une certaine élasticité
- Suite logiciel maintenu
- Aspect sécurité bien traité (chiffrement et authentification)
- Accessible via le *dashboard*

Et des inconvénients

- Stress important sur le service de messagerie RabbitMQ
- Nombre de couches importantes ne facilitant pas l'analyse des problèmes
- Toutes les versions de Kubernetes ne sont pas supportées

OpenStack Magnum



Rancher

Rancher 2.x

- Création de clusters Kubernetes
- Gestion de l'authentification et des droits d'accès
- Déploiement d'applications
- Supervision des clusters et des applications
- Collecte des logs
- Installation simple (paramétrage un peu plus ardu) :

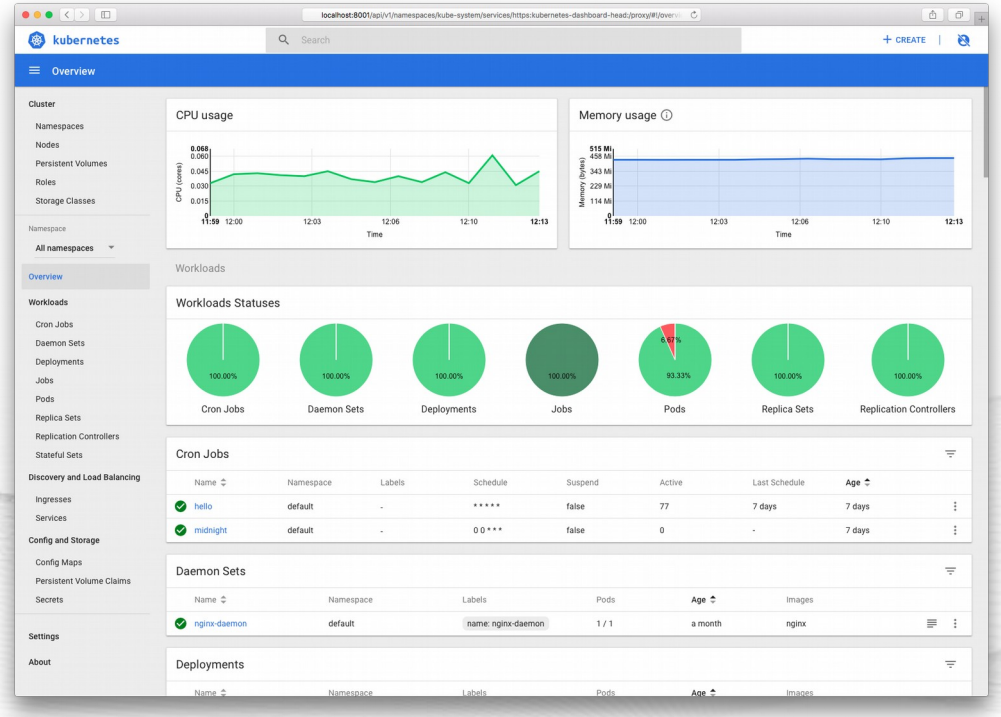
```
$ sudo apt install docker.io
$ docker run --privileged -d --restart=unless-stopped -p 80:80 \
  -p 443:443 rancher/rancher
```

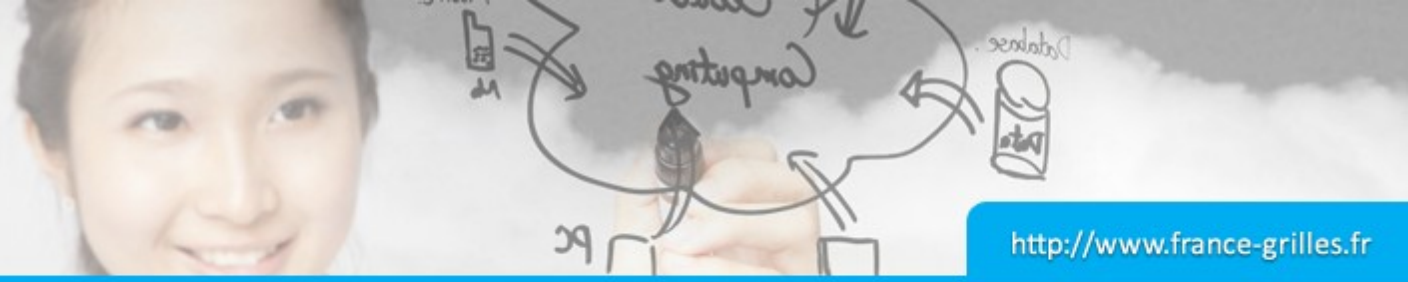
- Déploiement sur différents fournisseurs de Cloud, dont OpenStack : https://rancher.com/docs/rke/latest/en/config-options/cloud-providers/openstack/

Sécurité

Protection de son déploiement Kubernetes

- Pour la partie conteneur, voir la présentation de Martin : https://indico.mathrice.fr/event/225/session/1/contribution/5
- Vérifier les *security group*
- Mettre en place un système d'authentification sur tableau de bord Kubernetes
- Vérifier les images Docker disponibles
- Maintenir les systèmes à jour





Question / Discussion