

La fédération des infrastructures cloud

Ecole IN2P3 : Maîtriser le cloud

Charles LOOMIS (LAL)

4 juillet 2014

Questions importantes

- Est-ce que la fédération des ressources est nécessaire ?
- Si oui, comment on peut fédérer les ressources ?

L'état actuel

Une environnement avec beaucoup des centres variés.

Plusieurs salles des machines

- Les plus grosses centres peuvent fournir les ressources informatiques plus efficacement et moins cher...
- Pourquoi on a beaucoup des petites salles des machines dans nos laboratoires ?
 - On peut mieux contrôler l'accès au ressources
 - L'accès aux ressources locaux est plus fiable
 - Les ressources propres sont mieux adapter aux besoins
- Avec les technologies actuelles (virtualisation, cloud, réseau) ces raisons sont maintenant discutables !

Consolidation

- Il y a déjà une tendance vers la consolidation des sites ou des services communaux.
- P2IO (LabEx des 6 laboratoires à Orsay) a inauguré une salle des machines mutualisée récemment
- ownCore, un service qui permet le partage et la synchronisation des données pour les agents CNRS, basé sur ownCloud
- PARTAGE, une solution d'outils collaboratifs (Zimbra) déployée sur CloudWatt, créée par RENATER

Un seul cloud commercial ?

Est-ce que toutes nos ressources vont être délocalisés ?

Court terme...

- ▣ Un seul cloud à court terme ? → NON
- ▣ Les instituts veulent capitaliser ces investissements :
 - ▣ Les ressources physiques : salles et machines
 - ▣ Ressources humaines : expertise et aide aux utilisateurs

Moyen terme...

- Un seul cloud à moyen terme ? → NON

- Les problèmes associés avec le gestion des données :
 - Accès aux grosses quantités des données
 - La fiabilité et sauvegarde des données distantes
 - Questions légales : confidentialité et protection

Long terme...

- Un seul cloud à long terme ? → NON
- La « géographie » est un avantage
 - Minimiser la latence : mettre les services près des clients
 - Redondance des services : une plate-forme plus fiable
 - Catastrophe : sauvegarder des données et services

Fédération des ressources

L'utilisation coordonnée les services cloud différents.

Les besoins scientifiques

- Utilisateurs doivent partager leur données et ressources avec les autres dans les collaborations internationales.
- Les analyses scientifiques souvent a besoin d'une capacité de calcul qui dépasse la capacité d'un seul centre.
- Les objectives scientifique requirrent une fiabilité et redondance qui ne peuvent pas satisfaire par un seul centre.

Harmonisation

- Minimiser les différences entre les infrastructures cloud pour rendre la migration d'un centre d'un autre facile.
 - Types et sémantiques des services similaires
 - Des APIs standardisées pour ces services
 - Une méthode de authentification unique
 - Un format unique pour les images des machines
 - Un système de contextualisation uniforme
- L'harmonisation des centres différents demande une bonne volonté des instituts (et sociétés) concernés.

Agrégation

- Gérer les différences entre les infrastructures avec un service intermédiaire
 - Types et sémantiques des services similaires
 - Un interface abstract et des plugins
 - Les « credentials » stocker par le service intermédiaire
 - Formats des images et méthodes de contextualisation différents traiter par le service intermédiaire
 - Potentiellement les services de placement et migration des machines virtuelles et données
- L'agrégation demande plus d'effort aux utilisateurs ou au service intermédiaire.

Harmonisation \leftrightarrow Agrégation

- Les deux techniques sont complémentaires !
- Mais si on utilise les services cloud commerciaux, les techniques d'agrégation sont presque obligatoires.

Harmonisation—Authentication

- Une base d'identification unique (LDAP, AD, ...)
- Réutilisation de l'infrastructure PKI du EGI
 - Un système déjà accepté par plusieurs instituts dans monde
 - Les VOs permettre aussi une gestion des droits distribuée
- Systèmes d'authentification qui utilisent les jetons
 - Shibboleth, eduGAIN, ...
 - Le jeton (SAML) contient l'authentification de la personne ou du services et également les droits

Harmonisation—API Standardisée

- Open Cloud Computing Interface
 - OCCI (<http://occi-wg.org/>)
 - Développé et promu par OGF (<http://www.ogf.org/>)
- Cloud Infrastructure Management Interface
 - CIMI (<http://dmtof.org/standards/cloud>)
 - Développé et promu par DMTF (<http://dmtof.org>)
- OCCI et CIMI ne sont pas bien supportés par les fournisseurs cloud commerciaux ni les distributions cloud « open source ».
- Les APIs d'Amazon EC2/S3 sont toujours les plus implémentées et utilisées → un standard « de facto ».

Harmonisation—Contextualisation

- Les groupes qui développent les standards ont complètement ignoré la contextualisation; le point qui est le plus important !
- Mécanismes existants :
 - Une disque CD-ROM ou VFAT attachée à la machine virtuelle
 - Une serveur web privé avec les informations du contexte
 - Modification directe d'une image pour une machine virtuelles
- CloudInit devient un standard « de facto »
 - Implémenter par le plupart des logiciels cloud
 - Compatible avec Amazon, mais pas trop des autre fournisseurs

Agrégation—Plugins

- Une interface commune pour toutes les infrastructures
 - Peut inclure uniquement les caractéristiques communes entre toutes les infrastructures clouds
 - Les caractéristiques particulières peuvent être visibles aux utilisateurs mais dans un façon non portable
- On utilise (normalement) une architecture “plugin”
 - Traduction entre l'interface commune et l'interface native
 - Doit développer un plugin pour chaque API différentes

Agrégation—Brokering

- Un service « broker » est indispensable avec l'agrégation des ressources cloud
 - Peut optimiser le déploiement initial et son évolution
 - Le « monitoring » des ressources est requis
 - L'utilisateur (ou administrateur) doit fournir une politique explicite
- Un broker permet aussi les infrastructures « hybrid » entre les ressources privées et les ressources publiques

Les cloud fédérés

Les initiatives pour les clouds fédérés

Institut Français de Bioinformatique

- IFB : <http://france-bioinformatique.fr>
 - Une infrastructure cloud fédérée en France
 - Le cœur est hébergé à IDRIS (Orsay)
- Caractéristiques
 - Authentification commune
 - Des APIs différents
 - Appliances et applications communes

European Grid Infrastructure

- EGI FedCloud : <https://www.egi.eu/infrastructure/cloud/>
 - Une infrastructure cloud fédérée en Europe
 - Basé sur l'infrastructure et outils existantes pour la grille
- Caractéristiques
 - Authentification commune (certificats grilles)
 - API commune : OCCl
 - Appliances “commune”

Helix Nebula

- Helix Nebula: <http://www.helix-nebula.eu>
 - Une infrastructure cloud fédérée pour la science
 - Les sociétés cloud européennes fournissent les ressources
 - Valider par CERN, EMBL, ESA, et PIC
- Caractéristiques
 - Authentifications différentes
 - API différentes
 - Appliances différentes (mais similaires)
 - Unique broker pour la fédération (SlipStream)

Questions...

This work is licensed under the Creative Commons Attribution 3.0 Unported License (<http://creativecommons.org/licenses/by/3.0/>).



Copyright © 2014, Charles Loomis