



2^e SéRI 2010-2011

Cloud Computing ?

Date: Jeudi 17 mars 2011, 14h-17h30

Lieu : Amphi-Turing (EISTI - Cergy)



Architectures et performances des solutions *Cloud Open Source*

Conférencier : Hubert KADIMA – Directeur de recherche en informatique EISTI

Le *cloud computing* est un paradigme récent communément considéré comme un regroupement de technologies d'infrastructure visant à déployer les applications sur des conteneurs déportés [Vaquer09]. Les trois objectifs essentiels visés par cette approche sont (1) un passage à l'échelle infini, (2) une accessibilité quasi instantanée et (3) une sécurité et une stabilité des services et des données. Plusieurs solutions des infrastructures *Cloud Open Source* existent actuellement (*OpenNimbus*, *Eucalyptus*, *OpenNebula*). Nous présentons comparativement dans cette conférence les architectures fonctionnelles des principales solutions *Cloud Open Source* actuelles – presque toutes en cours de développement dans divers laboratoires de recherche ou industriels à travers le monde Des études de performances et de robustesse en cours dans notre laboratoire permettront de nous persuader de la maturité industrielle de ces solutions.

[Vaquer09] Vaquero, L. M. ; Roderer-Merino, L. ; Caceres, J. & Lindner, M. (2009), "A break in the clouds : towards a cloud definition", SIGCOMM Comput. Commun. Rev. 39(1), 50-55.

Mise en œuvre du protocole *OpenFlow* dans une grille

Conférencière : Sonia YASSA – Doctorante LARIS/EISTI

Actuellement, la communauté internationale des chercheurs en réseaux est fortement impliquée dans le développement des *réseaux programmables virtualisés* tel que GENI (<http://www.geni.net>). *OpenFlow* est une technologie standardisée basée sur des commutateurs Ethernet qui apporte une réponse satisfaisante à cette problématique. Elle permet aux chercheurs de l'*Université de Stanford* où elle a été initiée d'effectuer des expérimentations sur le réseau sans en perturber le bon fonctionnement.

La dernière spécification du protocole *OpenFlow* en version 1.0 publiée le 31 décembre 2009 est accessible en ligne sur le site www.openflowswitch.org. Cette présentation en est un condensé. Nous présentons les éléments essentiels du protocole *OpenFlow*, son architecture et les diverses configurations pouvant être projetées sur la grille du projet Marguerite qui l'a choisie pour assurer l'interconnexion réseau entre ses composants.

La technologie *OpenFlow* ne serait-elle pas déjà l'une des briques de base de l'Internet du futur ?



MapReduce et exécution des algorithmes de fouille de données dans un Cloud privé

Conférencier : Khaled TANNIR – Doctorant CIFRE LARIS/EISTI



MapReduce est un *framework* de calcul distribué sur de gros volumes de données. C'est un modèle de programmation parallèle dont les *combinators* *Map* et *Reduce* sont inspirés du langage fonctionnel *Lisp*. *MapReduce* a été développé dans Google comme mécanisme pour traiter des gros volumes de données en environnement distribué. *MapReduce* permet de répartir la charge sur un grand nombre de serveurs et gère entièrement, d'une manière transparente, le cluster, la distribution de données, la répartition de la charge et la tolérance aux pannes. *MapReduce* apporte une abstraction quasi-totale de la couche matérielle de l'infrastructure. La librairie *MapReduce* de Google existe dans plusieurs langages de programmation

dont *C++*, *C#*, *Erlang*, *Java*, *Python*, *Ruby*... D'autres librairies implémentent bien ce concept : par exemple *Hadoop* de la fondation Apache et *Elastic* de chez Amazon ...

Ces mécanismes dont nous présentons les principes, dans cette conférence, peuvent être appliqués pour exécuter les algorithmes de fouille de données en mode *PaaS* (*Platform as a Service*) dans un *Cloud privé*.