

Découverte et supervision automatique des serveurs d'un centre de données

Jean-Marc Menaud - FIL 2013

13 Mai 2013

Un centre de données (ou *data center*) est une infrastructure technique où sont regroupés en un même lieu géographique des dizaines, des centaines, voire des milliers d'ordinateurs interconnectés par un réseau haut débit. Un centre de données sert surtout à traiter les informations nécessaires aux activités d'une entreprise, allant d'un simple site web, aux applications SaaS (un concept consistant à proposer un abonnement à un logiciel disponible en ligne plutôt que l'achat d'une licence). Le nombre de centre de données, et le nombre de serveurs composant un centre de données va augmenter dans les prochaines années¹. Cette multiplication est liée à l'accroissement du nombre d'internautes (principalement dans les BRIC : Brésil, Russie, Inde, Chine), des terminaux d'accès², des données disponibles³, et l'apparition de nouveaux usages comme le *Cloud Computing* ou l'Internet des objets⁴.

Repérer, identifier et analyser dynamiquement les serveur présents dans ces salles informatiques devient un point prépondérant pour les administrateurs systèmes. L'objectif de ce projet FIL est de compléter l'application réalisée l'année dernière dans le cadre de cet exercice. Cette application permettait de repérer sur un réseau les machines connectées, puis d'identifier le système d'exploitation présent sur ces machines. Cette application a permis par exemple, de détecter les machines allumées sur le réseau de la MDE. L'objectif de ce projet est qu'après cette phase d'identification (Machine on et système d'exploitation utilisé), l'application devra collecter l'usage courant des machines (principalement le % de cpu utilisé) et repérer les machines allumées en permanence sans activité. Une dernière phase sera de permettre la mise en veille a distance des machines allumées.

1. Chez France Télécom, cette croissance est estimée à 30%, liée à des nouveaux services (Portail Web, Vidéo On Demand)

2. Il y a 1 milliard d'ordinateurs estimés pour 2008 et 2 milliards pour 2013, plus le nouveau marché des smart phones (77 millions vendus sur les trois premiers mois de 2009 soit 14 par seconde)

3. 161 exabytes de données ont été créées en 2006, soit approximativement 3 millions de fois l'information contenue dans tous les livres jamais écrits

4. Il est estimé que d'ici à quelques années le nombre d'appareils communicants pourrait être de plusieurs dizaine de milliards, soit plus que le nombre d'êtres humains

1 Principes

La première phase du projet sera de prendre en main, puis d'améliorer le code d'identification des machines allumées. Le principe est assez simple. Il suffit dans un premier temps de repérer les machines présentes sur le réseau. Pour cela, certains protocoles, comme le protocole "ping", permet de repérer assez facilement les machines allumées. Trouver les machines sur un réseau consiste à "pinguer" une plage d'adresse IP. Attention, il ne s'agit pas simplement de développer un petit client/serveur en java via la classe Socket, mais d'implémenter une partie du protocole ICMP. Une fois les machines repérées, l'application devra envoyer certaines combinaisons de chaînes de caractère sur des ports particuliers, et puis analyser les réponses pour en déterminer les systèmes d'exploitation.

La seconde phase consistera à récupérer les consommations des ressources systèmes (au moins cpu) via l'interface WMI pour Windows, ou vmstat pour linux. Ces consommations seront enregistrées en base. Enfin, et uniquement pour les postes windows, l'application devra permettre l'exécution de la mise en veille à distance de certaines machines.

2 Programme de travail

- Les étudiants devront prendre en main l'application développée l'année dernière.
- A partir de ce code, les étudiants devront rechercher et implanter les protocoles permettant de découvrir le système d'exploitation de la machine identifiée dans la phase précédente.
- Une fois identifié, la consommation des ressources système devra être remontée et sauvegardée.

3 Difficultés

La seule et principale difficulté sur ce sujet sera la montée en compétence sur la partie programmation réseau en Java, et spécifiquement sur ICMP et WMI.

4 Jalons

Un premier jalon dans le projet correspondra à la prise en main de l'application et de son amélioration. Le second jalon sera sur la remontée des consommations des ressources système.