

# Optimisation multi-critère

Georges Da Costa, Equipe Sepia

December 5, 2016

<http://www.irit.fr/~Georges.Da-Costa/>

dacosta@irit.fr

Stage L3/M1/M2 (adaptable suivant le niveau)

## Impact énergétique des centres de calcul

Sans être au premier plan médiatique, les centres de calcul ont pris une place centrale dans nos vies. Les services tels que ceux fournis par Google, Facebook, Tweeter, Amazon, . . ., reposent sur des grandes infrastructures de calcul et de stockage appelés centres de calcul. Pour fonctionner ils ont besoin de cette grande quantité de ressources qu'il a été nécessaire de regrouper afin de mutualiser les coûts de gestion et de refroidissement.

En 2014 déjà, les centres de calcul classiques pouvaient contenir plusieurs dizaines de milliers de serveurs sur des surfaces de l'ordre de  $10000m^2$ . Cette taille permet d'organiser de manière efficace la gestion des fluides (refroidissement, électricité). Pour un centre classique d'environ 40 000 serveurs utilisés par 500 000 services, la consommation électrique représente environ 10 mégawatts. Cette consommation est telle que le prix d'achat des serveurs est dépassé en quelques années par le prix de leur consommation électrique.

Avec l'explosion de l'utilisation de services sur internet, la consommation des centres de calcul explose. Entre 2000 et 2007, leur consommation annuelle au niveau mondial est passé de 70 à 330 térawatt-heures (TWh), et devrait atteindre 1000 à l'horizon 2020[6]. A l'échelle européenne, on estime que de 2007 à 2020, la consommation passera de 56 à 124 TWh par an[3]. En 2007 cette consommation représentait environ 2% des émissions de gaz carbonique ( $CO_2$ )[8], ce qui plaçait les centres de calcul au 6<sup>ème</sup> rang mondial si il s'agissait d'un pays[5]. Cette tendance à la hausse ne vas pas s'arrêter car les opérateurs de centres de calcul n'étaient que 8.5% à considérer en 2014 que la

capacité de leur centre serait suffisante pour leurs besoins au delà de 2015[4].

L'efficacité de ces centres de calcul est elle-même sujette à caution. En effet, une grande partie des serveurs qui y sont soit peu utilisés soit même devenus inutilisables. L'*Uptime Institute* évalue[9] à 20% les serveurs obsolètes, trop vieux, ou inutilisés. Quand à Google, une étude interne[1] a montré qu'en 2007 une grande majorité de leurs serveurs étaient utilisés à moins de 50% de leur capacité. Comme le fait même d'allumer un serveur consomme beaucoup d'électricité et produit de la chaleur qu'il faut extraire, ces chiffres montrent que sans une gestion très fine, les centres de calcul peuvent se révéler inefficaces.

## Optimisation d'énergie

Le rôle principal de l'intergiciel est donc l'orchestration d'un centre de calcul. C'est ainsi lui qui va décider quels serveurs vont être allumés ou éteints, où les tâches vont être placées. Mais il a aussi un rôle dynamique, il doit se rendre compte des problèmes (panne d'une machine par exemple) et les corriger. Pour chaque décision prise, il y a un large panel de possibilités. C'est de la qualité de ces décisions que va dépendre ensuite la qualité du centre de calcul.

De manière générale ces intergiciels utilisent une boucle autonome pour fonctionner (voir Figure 1). In s'agit d'un cycle où le système est surveillé (*Monitor*) grâce à des capteurs (*Sensor*). Les valeurs obtenues sont analysées (*Analyze*) et on en déduit des actions à effectuer (*Plan*). Les actions sont alors mises en place (*Execute*) grâce à des leviers (*Effector*). Une base de connaissance est manipulée

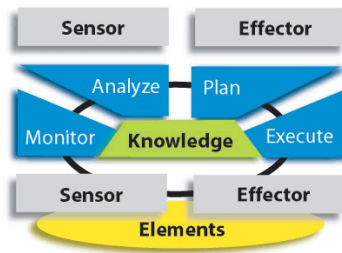


Figure 1: Un gestionnaire de centre de calcul décide des reconfigurations en fonction de l'état du système, de ses connaissances, puis les met en place. (Crédits IBM)

par tous ces éléments comme base de la représentation du système et est donc utilisée pour prendre les décisions.

En utilisant ces données et un modèle mathématique du centre de calcul (la base de connaissance), l'intelligence planifie les reconfigurations nécessaires à la bonne marche du centre de calcul, mais aussi à améliorer son fonctionnement.

## Sujet de Stage

Le but de ce stage va être de proposer de nouvelles heuristiques (algorithmes génétiques, gloutons, recuit simulé,...) ou méthodes de résolutions (linéaire, par contrainte,...) pour le placement et l'ordonnancement de machines virtuelles dans les centres de calculs.

Les optimisations se feront dans un cadre multi-objectif: Performance, énergie, résilience, dynamisme.

## L'équipe SEPIA

L'équipe SEPIA a déjà un grand nombre de publication dans ce domaine, en utilisant des méthodes gloutonnes[2] ou génétiques[7]. Nous nous sommes aussi intéressés lors du projet Européen CoolEmAll à des aspects thermiques[10]. Une thématique forte de l'équipe consiste à tenir compte de la source de l'électricité, par exemple grâce à de la production photovoltaïque et du stockage local (projet national DataZero<sup>1</sup>).

<sup>1</sup><http://www.datazero.org>

## References

- [1] Luiz André Barroso and Urs Hölzle. The case for energy-proportional computing. *IEEE Computer*, 40, 2007.
- [2] Damien Borgetto, Henri Casanova, Georges Da Costa, and Jean-Marc Pierson. Energy-Aware Service Allocation. *Future Generation Computer Systems*, 28(5):769–779, mai 2012. Special section on Energy Efficiency in Large-Scale Distributed Systems.
- [3] Richard Brown et al. Report to congress on server and data center energy efficiency: Public law 109-431. *Lawrence Berkeley National Laboratory*, 2008.
- [4] Emerson. Data center users group: Survey results. Technical report, Emerson Network Power, Emerson Electric Co., 2014.
- [5] Cook Gary. Clicking clean: How companies are creating the green internet. *Greenpeace International*, 2014.
- [6] Make IT Green. Cloud computing and its contribution to climate change. *Greenpeace International*, 2010.
- [7] Tom Guerout, Samir Medjiah, Georges Da Costa, and Thierry Monteil. Quality of Service Modeling for Green Scheduling in Clouds. *Sustainable Computing*, août 2014.
- [8] Cris Pettey. Gartner estimates ict industry accounts for 2 percent of global co2 emissions. <https://www.gartner.com/newsroom/id/503867>, 14:2013, 2007.
- [9] Matt Stansberry. 2014 data center industry survey. Technical report, Uptime Institute, 2014.
- [10] Hongyang Sun, Patricia Stolf, Jean-Marc Pierson, and Georges Da Costa. Energy-efficient and thermal-aware resource management for heterogeneous datacenters. *Sustainable Computing: Informatics and Systems*, pages 1–15, 2014.