

Résumé :

Avec l'essor de la technologie infonuagique *Cloud Computing*, les centres de données sont aujourd'hui censés d'occuper une place importante dans le scénario Internet. Malgré cette pertinence, ils sont probablement encore loin de leur véritable zénith en raison de la demande toujours croissante de contenus à stocker et à distribuer dans le Cloud, du besoin de puissance de calcul ou des quantités de plus en plus importantes de données analysées par les grandes entreprises telles que *Google*, *Microsoft* ou *Amazon*. Cependant, avoir un centre de données implique deux problèmes majeurs : ils sont terriblement coûteux à construire, et ils consomment une énorme quantité d'énergie, donc trop chers à entretenir. C'est pourquoi l'augmentation de l'efficacité énergétique (et donc la réduction de l'empreinte de carbone) des centres de données a été l'un des sujets de recherche les plus intéressants au cours des dernières années.

Pour augmenter convenablement l'efficacité énergétique d'un centre de données, nous devons comprendre comment l'énergie est consommée et établir des techniques pour bien gérer le fonctionnement du système informatique. Nous présentons, d'abord, une fonction de corrélation entre le service rendu par un serveur dans un centre de données et la consommation d'énergie de ce serveur en se basant sur le processeur, vu que ce dernier est considéré comme le composant le plus consommateur d'énergie. Ensuite, nous présentons deux différentes stratégies de consolidation des machines virtuelles et des machines physiques dans un centre de données. La première stratégie consiste en un processus de placement statique des machines virtuelles qui prend en compte la dimension énergétique ainsi que la qualité de service rendu aux utilisateurs. La deuxième stratégie de consolidation proposée est la migration en temps réel des machines virtuelles. En effet, afin de respecter les accords au niveau de service, on propose un modèle de migration dynamique permettant de prédire une variété des métriques, et par la suite aider à sélectionner automatiquement l'algorithme de migration le plus convenable pour la situation d'exécution donnée. Enfin, les modèles d'estimation et de gestion des ressources dans un centre de données sont efficaces dans le cas où le système fonctionne correctement; cependant, vue la grande taille et la diversité des composants dans cette structure, une certaine probabilité d'occurrence d'anomalies peut se produire; ainsi, on propose un algorithme d'apprentissage non supervisé pour identifier le comportement anormal des machines virtuelles lors de l'exécution de charges de travail dans un système en fonction des détails de l'utilisation des ressources.

Mots clés :

Cloud Computing, centre de données, efficacité énergétique, consolidation, détection d'anomalies, modèle stochastique, *Machine Learning*, métaheuristique.

ENERGY EFFICIENCY IN THE DATA CENTER: MACHINE-LEARNING AND BIO-INSPIRED ALGORITHMS-BASED MODELING

Abstract :

With the rapid development of Cloud Computing technology, data centers are currently supposed to occupy an important place in the Internet scenario. Despite such relevance, they are probably still far from their true zenith due to the ever-increasing pressure for content to be stored and distributed in the Cloud, the need for computing capacity or the ever-increasing amounts of data being analyzed by large enterprises such as Google, Microsoft or Amazon. Nevertheless, operating a data center involves two major problems: they are terribly expensive to construct, and they consume an enormous amount of energy, therefore too expensive to maintain. As a result, increasing energy efficiency (thereby reducing the carbon footprint) of data centers has been one of the most interesting areas of research over recent years.

To ensure the energy efficiency of a data center, we need to understand how energy is consumed and establish techniques to effectively manage the functioning of the IT system. First, we present a correlation function between the service provided by a server in a data center and the energy consumption of that server based on the CPU, since it is considered to be the most energy-consuming component. Second, we present two different strategies for consolidating virtual machines and physical machines in a data center. The first strategy consists of an energy-based static placement process for virtual machines that takes into account the quality of service provided to users. The second strategy proposed for consolidation is the live migration of virtual machines. To meet service level agreements, a dynamic migration model is proposed to predict a variety of metrics, and then help to automatically select the most suitable migration algorithm for the given execution situation. Finally, resource estimation and management models in a data center are efficient when the system is running properly; however, due to the large size and diversity of components in the data center, a certain probability of anomalies may occur; thus, an unsupervised learning algorithm is proposed to identify the abnormal behaviour of virtual machines when running workloads in a system depending on the details of resource utilization.

Key Words:

Cloud computing, data center, energy efficiency, consolidation, anomaly detection, stochastic model, Machine Learning, metaheuristic.